Multi-criteria optimization algorithms for high dose rate brachytherapy

Thèse

Songye Cui

Doctorat en physique
Philosophiæ doctor (Ph. D.)

Québec, Canada

Multi-criteria optimization algorithms for high dose rate brachytherapy

Thèse

Songye Cui

Sous la direction de:

Luc Beaulieu, directeur de recherche
Philippe Després, codirecteur de recherche

Résumé

L’objectif général de cette thèse est d’utiliser les connaissances en physique de la radiation, en programmation informatique et en équipement informatique à la haute pointe de la technologie pour améliorer les traitements du cancer. En particulier, l’élaboration d’un plan de traitement en radiothérapie peut être complexe et dépendant de l’utilisateur. Cette thèse a pour objectif de simplifier la planification de traitement actuelle en curiethérapie de la prostate à haut débit de dose (HDR).

Ce projet a débuté à partir d’un algorithme de planification inverse largement utilisé, la planification de traitement inverse par recuit simulé (IPSA). Pour aboutir à un algorithme de planification inverse ultra-rapide et automatisé, trois algorithmes d’optimisation multicritères (MCO) ont été mis en œuvre. Suite à la génération d’une banque de plans de traitement ayant divers compromis avec les algorithmes MCO, un plan de qualité a été automatiquement sélectionné.

Dans la première étude, un algorithme MCO a été introduit pour explorer les frontières de Pareto en curiethérapie HDR. L’algorithme s’inspire de la fonctionnalité MCO intégrée au système Raystation (RaySearch Laboratories, Stockholm, Suède). Pour chaque cas, 300 plans de traitement ont été générés en série pour obtenir une approximation uniforme de la frontière de Pareto. Chaque plan optimal de Pareto a été calculé avec IPSA et chaque nouveau plan a été ajouté à la portion de la frontière de Pareto où la distance entre sa limite supérieure et sa limite inférieure était la plus grande.

Dans une étude complémentaire, ou dans la seconde étude, un algorithme MCO basé sur la connaissance (kMCO) a été mis en œuvre pour réduire le temps de calcul de l’algorithme MCO. Pour ce faire, deux stratégies ont été mises en œuvre : une prédiction de l’espace des solutions cliniquement acceptables à partir de modèles de régression et d’un calcul parallèle des plans de traitement avec deux processeurs à six cœurs. En conséquence, une banque de plans de traitement de petite taille (14) a été générée et un plan a été sélectionné en tant que plan kMCO. L’efficacité de la planification et de la performance dosimétrique ont été comparées entre les plans approuvés par le médecin et les plans kMCO pour 236 cas.

La troisième et dernière étude de cette thèse a été réalisée en coopération avec Cédric Bélanger.

Un algorithme MCO (gMCO) basé sur l'utilisation d'un environnement de développement compatible avec les cartes graphiques a été mis en œuvre pour accélérer davantage le calcul. De plus, un algorithme d'optimisation quasi-Newton a été implémenté pour remplacer le recuit simulé dans la première et la deuxième étude. De cette manière, un millier de plans de traitement avec divers compromis et équivalents à ceux générés par IPSA ont été calculés en parallèle. Parmi la banque de plans de traitement généré par l'agorithme gMCO, un plan a été sélectionné (plan gMCO). Le temps de planification et les résultats dosimétriques ont été comparés entre les plans approuvés par le médecin et les plans gMCO pour 457 cas.

Une comparaison à grande échelle avec les plans approuvés par les radio-oncologues montre que notre dernier algorithme MCO (gMCO) peut améliorer l'efficacité de la planification du traitement (de quelques minutes à 9.4 s) ainsi que la qualité dosimétrique des plans de traitements (des plans passant de 92.6 % à 99.8 % selon les critères dosimétriques du groupe de traitement oncologique par radiation (RTOG)).

Avec trois algorithmes MCO mis en œuvre, cette thèse représente un effort soutenu pour développer un algorithme de planification inverse ultra-rapide, automatique et robuste en curiethérapie HDR.

Abstract

The overall purpose of this thesis is to use the knowledge of radiation physics, computer programming and computing hardware to improve cancer treatments. In particular, designing a treatment plan in radiation therapy can be complex and user-dependent, and this thesis aims to simplify current treatment planning in high dose rate (HDR) prostate brachytherapy.

This project was started from a widely used inverse planning algorithm, Inverse Planning Simulated Annealing (IPSA). In order to eventually lead to an ultra-fast and automatic inverse planning algorithm, three multi-criteria optimization (MCO) algorithms were implemented. With MCO algorithms, a desirable plan was selected after computing a set of treatment plans with various trade-offs.

In the first study, an MCO algorithm was introduced to explore the Pareto surfaces in HDR brachytherapy. The algorithm was inspired by the MCO feature integrated in the Raystation system (RaySearch Laboratories, Stockholm, Sweden). For each case, 300 treatment plans were serially generated to obtain a uniform approximation of the Pareto surface. Each Pareto optimal plan was computed with IPSA, and each new plan was added to the Pareto surface portion where the distance between its upper boundary and its lower boundary was the largest.

In a companion study, or the second study, a knowledge-based MCO (kMCO) algorithm was implemented to shorten the computation time of the MCO algorithm. To achieve this, two strategies were implemented: a prediction of clinical relevant solution space with previous knowledge, and a parallel computation of treatment plans with two six-core CPUs. As a result, a small size (14) plan dataset was created, and one plan was selected as the kMCO plan. The planning efficiency and the dosimetric performance were compared between the physician-approved plans and the kMCO plans for 236 cases.

The third and final study of this thesis was conducted in cooperation with Cédric Bélanger. A graphics processing units (GPU) based MCO (gMCO) algorithm was implemented to further speed up the computation. Furthermore, a quasi-Newton optimization engine was implemented to replace simulated annealing in the first and the second study. In this way, one thousand IPSA equivalent treatment plans with various trade-offs were computed in parallel. One plan was selected as the gMCO plan from the calculated plan dataset. The planning time and the dosimetric results were compared between the physician-approved plans and the gMCO plans for 457 cases.

A large-scale comparison against the physician-approved plans shows that our latest MCO algorithm (gMCO) can result in an improved treatment planning efficiency (from minutes to 9.4 s) as well as an improved treatment plan dosimetric quality (Radiation Therapy Oncology Group (RTOG) acceptance rate from 92.6% to 99.8%).

With three implemented MCO algorithms, this thesis represents a sustained effort to develop an ultra-fast, automatic and robust inverse planning algorithm in HDR brachytherapy.

Table des matières

Liste des tableaux

Liste des figures

Liste des abréviations

1D	1 Dimensions
2D	2 Dimensions
3D	3 Dimensions
3DUS	3D Ultrasound System
4D	4 Dimensions
AAPM	American Association of Physicist in Medicine
ABS	American Brachytherapy Society
ANN	Artificial Neural Network
AP	Auto Planning
API	Application Programming Interface
BFGS	Broyden–Fletcher–Goldfarb–Shanno (optimization method)
CHU	Centre Hospitalier Universitaire
CPU	Central Processing Unit
CLRP	Carleton Laboratory for Radiotherapy Physics
CRSNG	Conseil de Recherches en Sciences Naturelles et en Génie du Canada
cSA	CPU-based Simulated Annealing
CT	Computed Tomography
CTV	Clinical Target Volume Tomography
CVT	Centroidal Voronoi Tessellations
CUDA	Compute Unified Device Architecture
DTGR	Dwell Time Gradient Restriction
DVH	Dose Volume Histogram
DVHO	Dose Volume Histogram Objective
DSP	Digital Signal Processor
EBRT	External Beam Radiation Therapy
EM	Electromagnetic
ENIAC	Electronic Numerical Integrator And Computer
EUD	Equivalent Uniform Dose

FPGA	Field-Programmable Gate Array
GPU	Graphics Processing Unit
gL-BFGS	GPU-based Limited memory Broyden–Fletcher–Goldfarb–Shanno
gMCO	GPU-based Multi-Criteria Optimization
GPGPU	General Purpose GPU
gSA	GPU-based Simulated Annealing
GTV	Gross Target Volume
Gy	Gray
GYN	Gynecologic
HEBD	High Energy Brachytherapy Source Dosimetry
HIPO	Hybrid Inverse Planning Optimization
HDR	High Dose Rate
HU	Hounsfield Unit
IBM	International Business Machines corporation
iCycle	IMRT Cycle
ICRU	International Commission on Radiation Units and Measurements
IMPT	Intensity-Modulated Proton therapy
IMRT	Intensity-Modulated Radiation Therapy
IROC	Imaging and Radiation Oncology Core
IPSA	Inverse Planning Simulated Annealing
KBP	Knowledge-Based Planning
kMCO	knowledge-based Multi-Criteria Optimization
L-BFGS	Limited memory Broyden–Fletcher–Goldfarb–Shanno
LDR	Low Dose Rate
LINAC	Linear Accelerator
MCO	Multi-Criteria Optimization
MPRTN	Medical Physics Research Training Network
MRI	Magnetic Resonance Imaging
nvcc	NVIDIA CUDA Compiler
OAR	Organ At Risk
OpenCL	Open Computing Language
OpenMP	Open Multi-Processing
OR	Operating Room
PDR	Pulse Dose Rate
PTV	Planning Target Volume
RAM	Random-Access Memory

ROI	Region Of Interest
RTOG	Radiation Therapy Oncology Group
SA	Simulated Annealing
SBRT	Stereotactic Body Radiation Therapy
SRS	Stereotactic Radiosurgery
TG-186	Task Group 186
TG-43	Task Group 43
TPS	Treatment Planning System
TRUS	Transrectal Ultrasound
US	Ultrasound
VMAT	Volumetric Modulated Arc Therapy

To my family

Remerciements

I once thought it would be too sensational to acknowledge a bunch of people in this section. After 5 years' study abroad, I have learned to appreciate every supports from mentors, co-workers, friends and families.

I feel fortunate and honored that I chose medical physics as a profession. As such, I will begin my acknowledgements by expressing my gratitude to the people who guided me into this field. I first learned about medical physics in junior year at university, when Prof. Zhensheng Deng introduced its role in cancer treatment during a medical imaging course. I also want to mention my supervisor and co-supervisor during my master's study, Prof. Wenyi Zhang and Prof. Ling Jiao. They not only taught me how to conduct a research project, but also gave me so much supports regardless of the geographic distance.

Being trained in a CAMPEP accredited medical physics program was once my major career goal and I feel lucky that I have a chance to learn from the experts in this field. I am grateful that my supervisor Prof. Luc Beaulieu gave me a chance to start the adventure. As time goes by, I started to understand how mentorship works in an open society. I once wondered why western countries are so good at science and I guess some of the reasons must include more respects and open discussions, less condescension and hierarchy.

At the beginning of my study, I was overwhelmed with the courses and the research. The sub-group meeting with my co-supervisor Prof. Philippe Després, made me feel integrated and I received tremendous helps from Jean-François Montégiani, Daniel Maneval, Dmitri Matenine, Julia Mascolo-Fortin and Marie-Annie Saucier.

I certainly appreciate numerous discussions with Yunzhi Ma, Daline Tho and Cédric Bélanger. I would like to thank Ghyslain Leclerc and Carly Whittaker for the grammatical corrections. I want to express my gratitude to all the teachers who were engaged in the medical physics courses and to all the medical physicists I had chances to learn from.

Furthermore, I would like to thank the Chinese Scholarship Council for the major financial sponsorship during my study and thank Calcul Québec and Compute Canada for organizing the excellent workshops on high performance computation.

I enjoyed the moments that I spent with my friends in Quebec, Xiaotong Guo, Xiaoye Sang, Shiqi Liu and many others. I would like to give a special thanks to the Mas' family, with whom I can feel like home. I would like to thank two close friends who have been giving me massive supports especially during some of the toughest moments: Yongrui Gao, who I often exchanges the opinions about TV shows, movies as well as different cultures and Dabin Tang, who I often discusses about football, techniques, politics and general news.

Last, but certainly not the least, my strongest supports are certainly from my family: especially my father Dewei Cui, my mother Fengmin Li, my elder sister Dan Cui, my brother-in-law Jundong Shi and my nephews Fuxuan and Fuen Shi. Without their supports, it would be impossible for me to walk until the end of my study.

Avant-propos

This thesis includes three articles that have been published in scientific journals. The information on each of the authors as well as their respective contributions is given below. Songye Cui is the first author (or co-first author) of each article.

Chapter 2: A multi-criteria optimization approach for HDR prostate brachytherapy: I. Pareto surface approximation

Songye Cui[1,2], Philippe Després[1,2], Luc Beaulieu[1,2]

Publié: Physics in Medicine and Biology **63** (20) 205004, October 2018

Contributions: Songye Cui completed the design of the Pareto surface approximation algorithm based on the existing planning algorithm (IPSA), performed the C/C++ programming of the Pareto surface approximation algorithm, generated the input data for the algorithm, generated the raw results, wrote the article and corrected the article based on the comments of the co-authors, the referees appointed by the PMB editors, and the PMB editors. Luc Beaulieu participated in the development of the approach through numerous discussions, provided the opportunities to participate in workshops and conferences that lead to the implemented approach, participated in interpretation of results and the revision of the article. Philippe Després participated in the development of the approach through numerous discussions, offered advice about programming and bibliography management, participated in interpretation of results and the revision of the article.

Chapter 3: A multi-criteria optimization approach for HDR prostate brachytherapy: II. benchmark against clinical plans

Songye Cui[1,2], Philippe Després[1,2], Luc Beaulieu[1,2]

Publié: Physics in Medicine and Biology **63** (20) 205005, October 2018

Contributions: Songye Cui completed the design of multi-criteria optimization (MCO) algorithm, performed the C/C++ programming as well as the openMP (open Multi-Processing) implementation of the MCO algorithm, generated the input data for the algorithm, generated the raw results, wrote the article and corrected the article based on the comments of the co-authors, the referees appointed by the PMB editors, and the PMB editors. Luc Beaulieu participated in the development of the approach through numerous discussions, provided the opportunities to participate in workshops and conferences that lead to the implemented approach, participated in interpretation of results and the revision of the article. Philippe Després participated in the development of the approach through numerous discussions, provided the opportunities to participate in a workshop that leads to the openMP implementation, participated in interpretation of results and the revision of the article.

Chapter 4: A GPU-based multi-criteria optimization algorithm for HDR brachytherapy

Cédric Bélanger[1,2*], Songye Cui[1,2*], Yunzhi Ma[2], Philippe Després[1,2], J. Adam M. Cunha[3], Luc Beaulieu[1,2]

Publié: Physics in Medicine and Biology **64** (10) 105005, May 2019

Contributions: Songye Cui completed the design of multi-criteria optimization (MCO) algorithm, implemented the gradient method (including the design of mathematical formulations and a compatible class solution), generated the input data for the algorithm, generated the raw results, wrote the article and corrected the article based on the comments of the co-authors. Cédric Bélanger participated in the development of the gradient method, implemented the MCO algorithm on GPU, generated the raw results, wrote the article and corrected the article based on the comments of the co-authors. Yunzhi Ma completed the design of the gSA (GPU-based simulated annealing) prototype. Luc Beaulieu participated in the development of the gradient method as well as the GPU implementation, participated in interpretation of results and the revision of the article. Philippe Després participated in the development of the gradient method as well as the GPU implementation, participated in interpretation of results and the revision of the article.

Affiliations of authors:

1. Department of Physics, Engineering Physics and Optics and Cancer Research Center, Université Laval, Quebec City, QC, G1V 0A6, Canada
2. Department of Radiation Oncology and Research Center of CHU de Québec - Université Laval, Quebec City, QC, G1R 2J6, Canada
3. Radiation Oncology, University of California, San Francisco, CA, 94115, USA

∗ These authors contributed equally to this work.

Introduction

Background

According to the analysis from the Global Burden of Disease Study [1], there were 52.8 million deaths in 2010 in the world, the top three causes of which were attributed to: (1) Communicable, maternal, neonatal and nutritional causes (13.1 million, or 24.9%), (2) Ischaemic heart disease and stroke (12.9 million, or 24.4%), and (3) Cancers (8 million, or 15.2%) [1].

Even though the age-standardized death rates of cancers in 2010 has dropped by 13.8% compared to 1990 [1], it is important to improve current treatments and to accelerate the development of new therapeutic approaches.

While immunotherapy, hormone therapy, or targeted therapy may be effective for some cancer patients, radiotherapy (also called radiation therapy), together with surgery and chemotherapy remain three common cancer treatment approaches nowadays [2].

Studies [3–5] estimated that about 52.3% of non-skin cancer patients receive radiation therapy (alone or in combination with other approaches) during the course of their illness. Radiation can be administered either in the form of photons (x-rays and gamma rays) or particles (protons, heavy ions, neutrons, and electrons) [3]. Photons are currently the main type of therapeutic radiation beam, and they can be used to kill the tumor cells by damaging the DNA in chromosomes [3]. Specifically, photon interactions involve transferring part of their energy to free electrons (Compton scattering or photoelectric effect), and these electrons can induce DNA damage either directly or indirectly (by producing hydroxyl radicals through interacting with water molecules) (Figure 0.1).

Photons can be delivered either externally from a medical linear accelerator (LINAC) in external beam radiation therapy (EBRT) or internally from an inserted small radioactive source in brachytherapy [5]. Treatment-delivery machines in radiation therapy are shown in Figure 0.2. This project focuses on the second type of radiation therapy: brachytherapy.

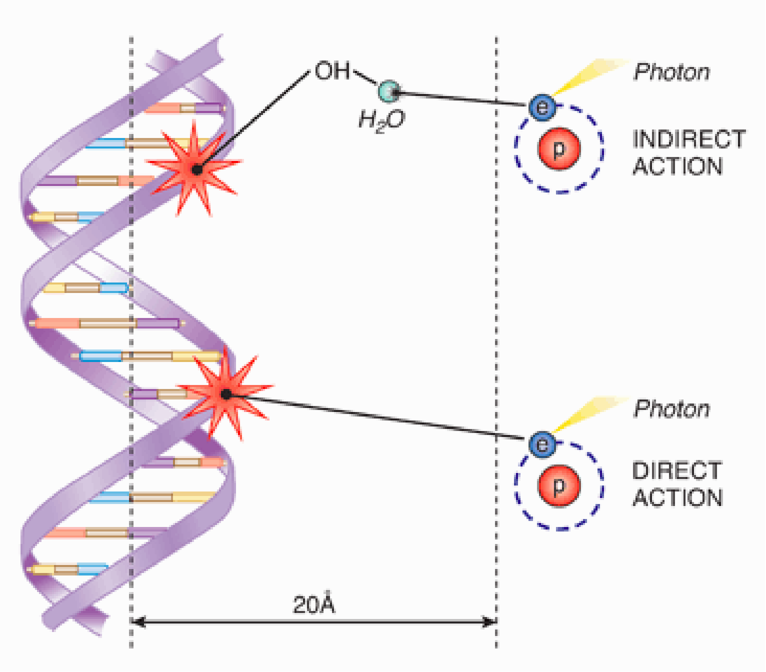


FIGURE 0.1 – The direct and indirect effects of photon beams on DNA, image from [3].

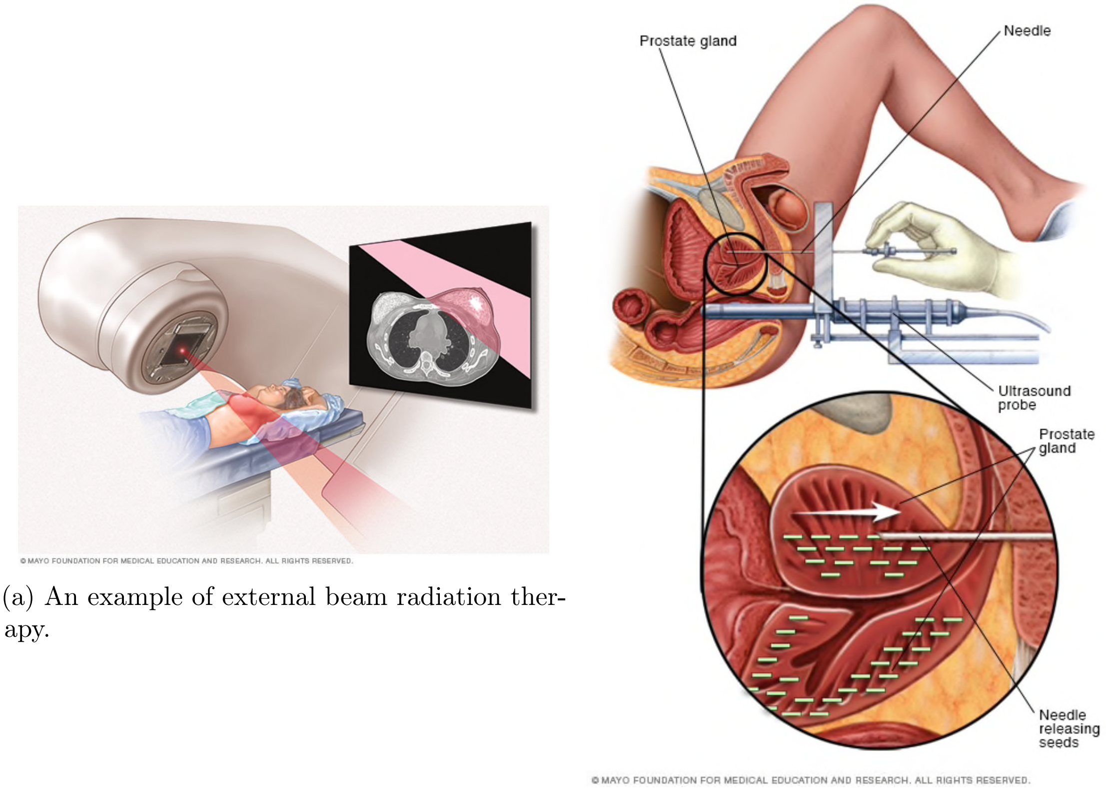
Brachytherapy

Definition

Shortly after Henri Becquerel discovered radioactivity in 1896, the idea of shrinking tumors with a radioactive source was suggested by Pierre Curie in 1901 [6]. This treatment method, later known as internal radiation therapy or brachytherapy (curiethérapie in French), involves directly placing a sealed container of radioactive substance into a tumor. These characteristics can be derived from the word brachy, which originates from the Greek word $\beta\rho\alpha\chi\upsilon\varsigma$ (brachys, meaning "short-distance").

Based on the dose rate of radioactive sources, brachytherapy can be classified into low dose rate (LDR) brachytherapy and high dose rate (HDR) brachytherapy. LDR brachytherapy delivers durable and continuous low doses of radiation. The dose rate of radioactive seeds in LDR brachytherapy is less than $2\,\mathrm{Gy\,h^{-1}}$, and common LDR seeds include ^{103}Pd, ^{125}I and ^{131}Cs [7,8]. The seeds (Figure 0.3a) are usually implanted through inserted needles (Figure 0.3d), and their duration inside the patient's body can either be hours, days (1 to 7 days before being taken out) or they can be permanent [9].

On the other hand, HDR brachytherapy usually delivers a single temporary high dose of radiation (although some treatments may be conducted in multiple fractions). The dose rate of radioactive sources in HDR brachytherapy is larger than $12\,\mathrm{Gy\,h^{-1}}$, and classic HDR



(a) An example of external beam radiation therapy.

(b) An example of permanent prostate brachytherapy.

FIGURE 0.2 – Treatment facilities in radiation therapy. Images downloaded from https://www.mayoclinic.org/tests-procedures in December 2018.

sources include ^{60}Co and ^{192}Ir [7,8]. The sources are attached with wires (Figure 0.3b), and travel through plastic catheters and applicators using a remote afterloader (Figure 0.3c). The duration inside a patient's body is usually a few minutes [9].

Both LDR brachytherapy and HDR brachytherapy are currently widely used in clinics. The anatomic sites treated with LDR brachytherapy usually include prostate, eye, lung and brain cancers. HDR brachytherapy, on the other hand, is commonly used to treat prostate, cervical, breast, and head and neck cancers.

Pros and cons of brachytherapy

In brachytherapy, radioactive sources are placed directly on, in, or within the treatment target volume, and tumor cells are killed from a short distance. Due to this characteristic, there are both advantages and disadvantages of brachytherapy compared with EBRT.

Based on the fact that radioactive sources are placed close to the target volume and that

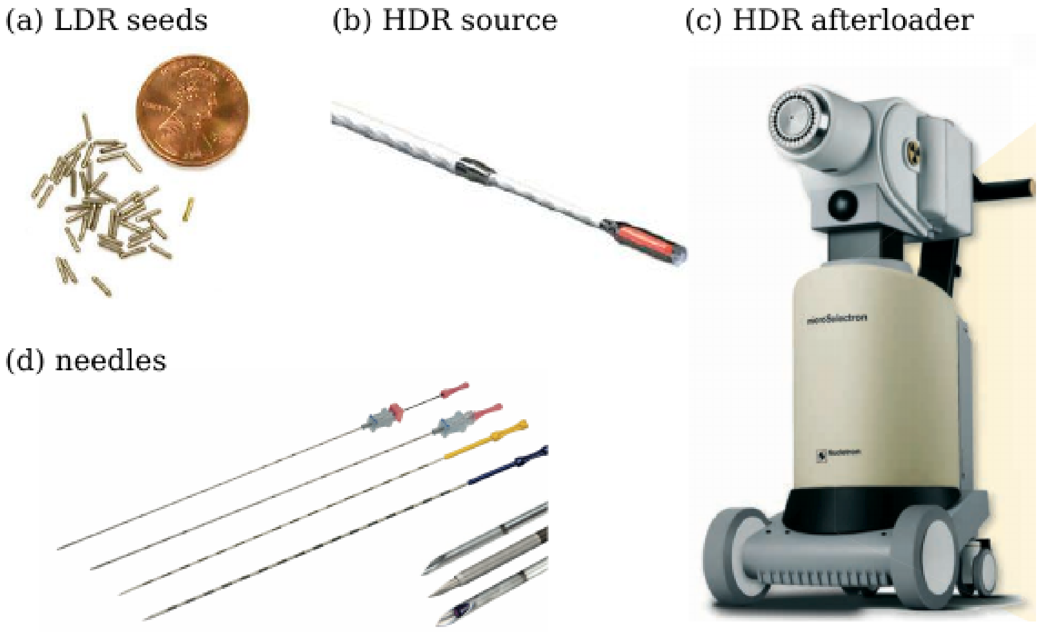
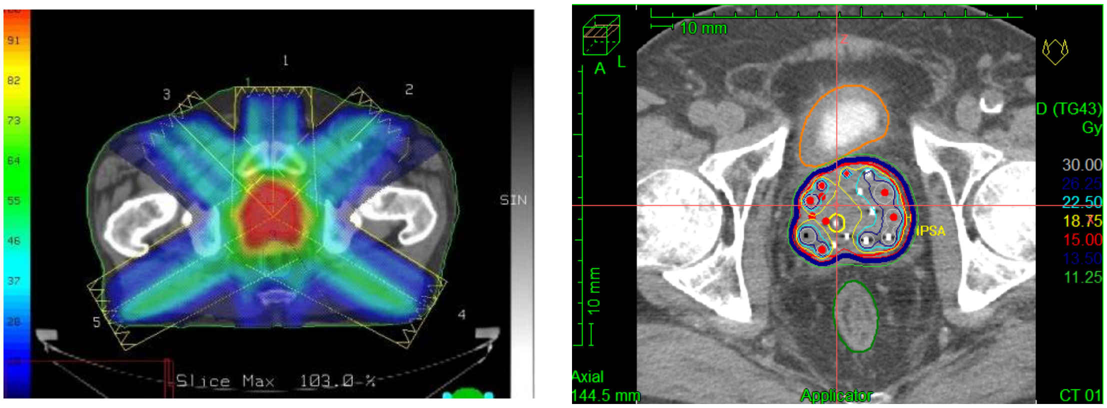


FIGURE 0.3 – An overview of the Devices used in brachytherapy: (a) LDR seeds for implantation, (b) A HDR radioactive source attached with wires, (c) A MicroSelectron remote afterloader (Elekta Brachy, Veenendaal, The Netherlands), and (d) Needles for LDR brachytherapy, images from [10].



(a) Dose distribution of EBRT

(b) Dose distribution of brachytherapy

FIGURE 0.4 – A comparison of dose distribution between EBRT prostate (image downloaded from https://www.varian.com in December 2018) and HDR prostate brachytherapy (a screenshot of Oncentra Brachy).

the dose distribution simply follows the inverse square law, brachytherapy treatments can result in (1) A sufficiently high dose in the target, and (2) A rapid dose falloff to surrounding healthy tissue [9]. In other words, the resulting dose distribution is often considered an ideal outcome because it can be both high and conformal within the target volume and low in healthy organs (Figure 0.4). The dosimetric advantages of brachytherapy over proton therapy (with significant skin entry dose) and stereotactic radiosurgery (SRS) (with entry dose and exit dose) are detailed in the following reference [9]. In addition, brachytherapy has a clinical advantage in the accuracy and reliability of its dose delivery because the relative position between the target and the inserted applicator/catheters is fixed, and the setup error caused by internal organ movement can be minimized [9]. Furthermore, the overall treatment duration of HDR brachytherapy is generally shorter than external beam treatment, therefore the errors caused by the target volume change can be mitigated [9].

However, directly delivering radioactive sources to the target volume can also lead to several clinical disadvantages. For example, one main disadvantage is the invasive procedure of inserting catheters or sources. Even though some form of anesthesia may reduce the pain during surgery, the brachytherapy treatment procedure is relatively complicated, and usually involves a multidisciplinary cooperation. Furthermore, the clinical application of brachytherapy is relatively limited compared with EBRT because the treatment sites need to be suitable for source placements. In particular, the tumors treated in brachytherapy is usually limited to surface tumors (skin), body cavities (gynecological, esophagus and bronchus), and a few interstitial sites (prostate, breast, rectal, gynecological and head-and-neck region) [9]. Finally, specific skills and training are required to conduct brachytherapy, which could also limit its application [9].

Treatment procedure in HDR prostate brachytherapy

HDR prostate brachytherapy can either be delivered as a boost to EBRT or delivered alone (or called monotherapy) [8]. Prior to a brachytherapy procedure, it is necessary for the patient to make an initial consultation with a radiation oncologist. Clinical examinations including an imaging scan may be given to identify the characteristics of the tumour in a digital visualized form. The options of imaging scans usually include ultrasound (US), computerized tomography (CT) and magnetic resonance imaging (MRI) (detailed in Section).

Once a decision to proceed with brachytherapy is made, it is important to coordinate efforts from different specialists including urologists, radiation oncologists, anaesthetists, medical physicists, nurses, therapists and dosimetrists.

Treatment setups in the operating room (OR) begin with a general or regional (spinal, epidural) anesthesia to reduce the pain, and patients are placed on a treatment couch in the lithotomy position (Figure 0.5). Radiation oncologists then insert plastic catheters/needles (Figure 0.5) into the prostate through the perineum area, guided by a transrectal ultrasound (or TRUS) (Figure 0.5). Through a cystoscopy, physicians are able to make sure the position of catheters will not damage bladder and urethra, while ensuring a good coverage of the base of the prostate.

After the treatment setup, it is necessary to begin a treatment planning procedure (described in the next section).

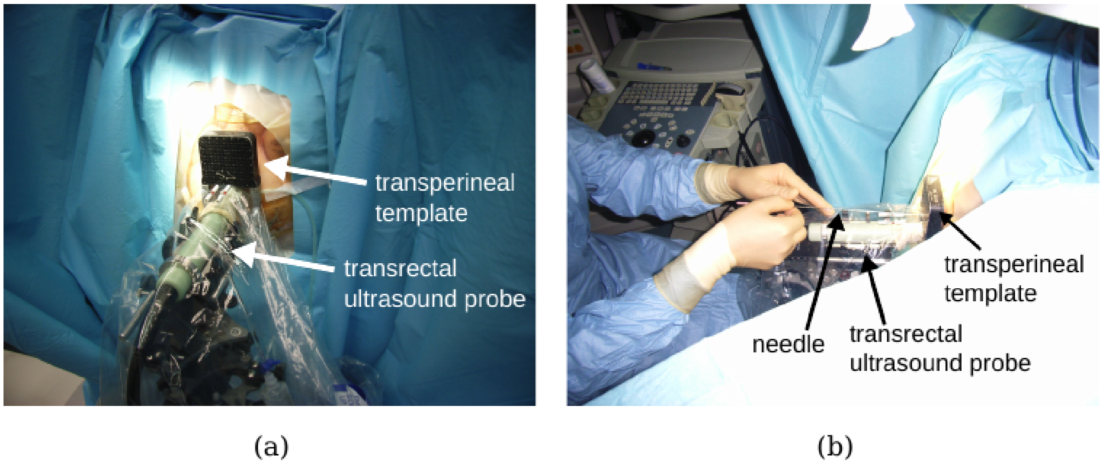
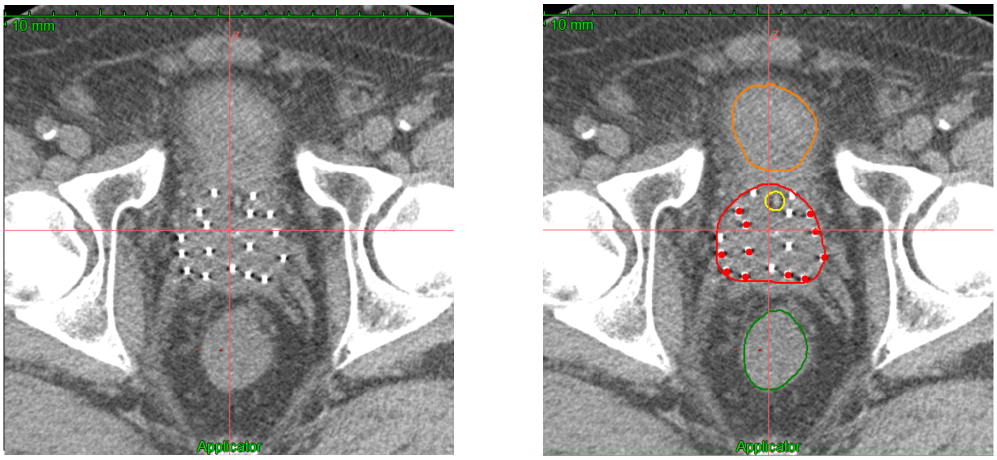


Figure 0.5 – Treatment setups for prostate brachytherapy in the OR, images from [10].

Treatment planning process in HDR brachytherapy

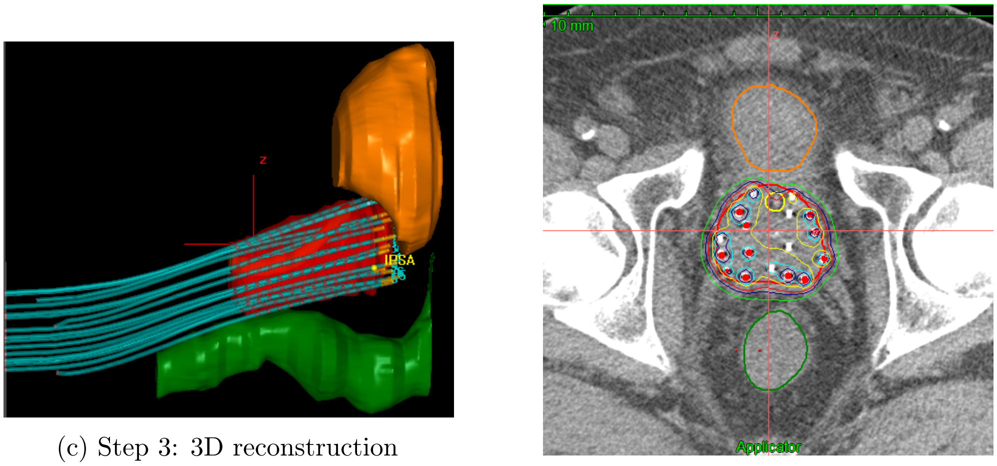
The purpose of radiation therapy is to deliver a therapeutic dose to the target volume without over-irradiating the surrounding healthy tissue. Oncologists are responsible for the prescription of the treatment. In order to meet the clinical prescription, it is necessary to develop an optimal (ideal) treatment plan that is deliverable using the equipment on hand (Figure 0.2). While an EBRT plan may consist of a beam angle arrangement and a fluence map distribution, a brachytherapy plan usually involves in determining dwell positions (source positions in catheters) and dwell times (source stopping duration).

The process of designing such a plan is known as treatment planning, and a variety of treatment planning algorithms [11] installed in workstations are available to facilitate the process. In Figure 0.6, the process of treatment planning for HDR prostate brachytherapy is described in the scenario of Oncentra Brachy IPSA (Inverse Planning Simulated Annealing, detailed in Section 1.2.1.1) (Elekta Brachy, Veenendaal, The Netherlands). Each step will be described in the following sections.



(a) Step 1: imaging

(b) Step 2: contouring



(c) Step 3: 3D reconstruction

(d) Step 4: planning and evaluation

FIGURE 0.6 – Workflow of treatment planning in HDR prostate brachytherapy: (1) CT image import, (2) Contouring of organs, (3) 3D reconstruction of organs and catheters, and (4) Treatment planning and plan evaluation.

Medical imaging modalities used in brachytherapy

The emergence of 3D imaging modalities is often marked as an evolution in modern radiation oncology because of its non-invasive approach to visualize internal organs and identify abnormalities inside patient's body. The electromagnetic (EM) spectrum (Figure 0.7) of medical imaging techniques includes x-rays in mammography and CT, radiofrequency in MRI, and gamma rays in nuclear medicine [12].

In radiography, x-rays are emitted from a fixed vacuum X-ray tube, where the high-speed electrons (released by a hot cathode, accelerated with a high voltage) collide with a metal target (usually tungsten) [12]. As the attenuation properties of tissues are different, the entering homogeneous x-rays become heterogeneous when emerging from the patient [12]. An image is created when the heterogeneous x-ray distribution is received by a detector opposite the x-ray

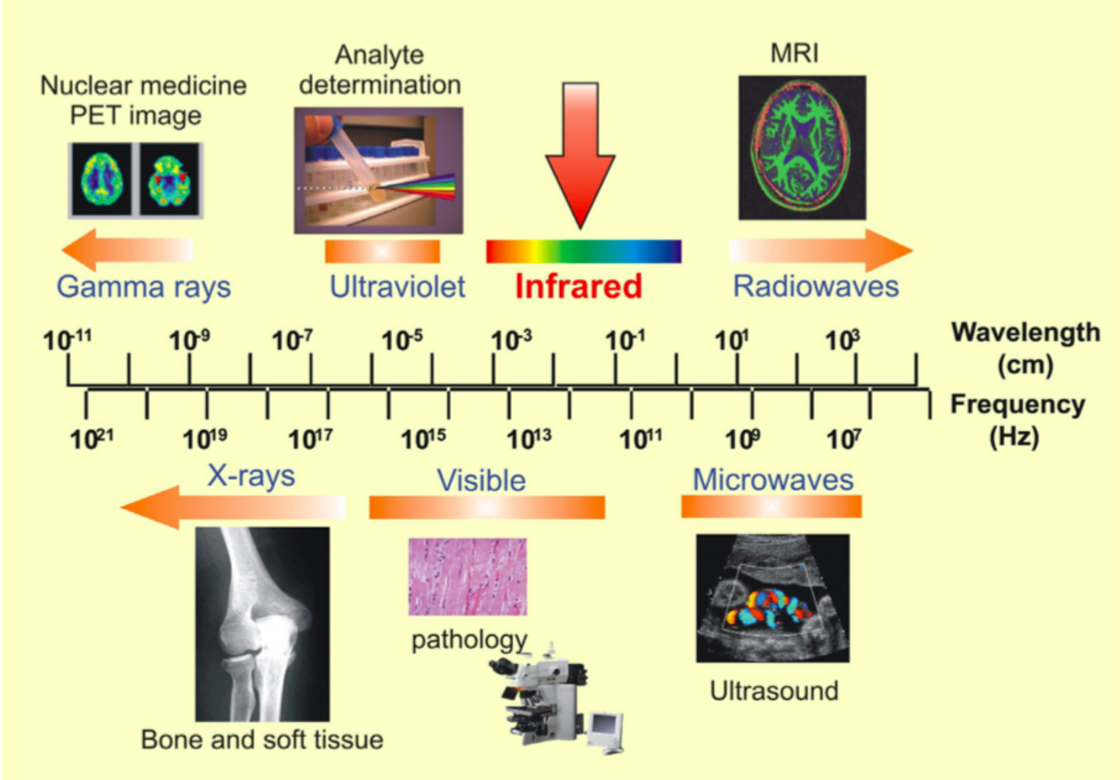


FIGURE 0.7 – The electromagnetic spectrum in medical imaging modalities, image from [13].

tube [12].

In computed tomography (CT), images are synthesized from the transmission projection data of a large number of angles by a computer [12]. The transmission projection data are produced by rotating the x-ray tube and the detector array around the patient [12]. Compared with MRI, a CT scan has several advantages as they are: (1) Less expensive, (2) Require a shorter examination time, and (3) Are able to better distinguish between bone structure, soft tissue and air [12]. In Figure 0.6a, a set of CT images from a prostate cancer patient were obtained.

In clinical magnetic resonance imaging (MRI), patients (whom contain abundant water inside the body) are placed in a strong magnetic field (1.5 to 3.0 T). Based on the fact that some hydrogen nuclei within water molecules can absorb and emit radio signals at a given frequency in a magnetic field, MRI can produce a set of tomographic images from the emitted radio signals, which are then detected by the antennas around the patient [12]. Compared with CT, MRI has three advantages: (1) Soft tissue and internal organs can usually be better distinguished, (2) The difference between normal and abnormal tissue is often clearer, and (3) No ionizing radiation is required (hence usually a good choice for pregnant women or children).

Ultrasound (US) is sound waves of which the frequency is too high (above 20 kHz) to be heard by humans [12]. The frequency of medical ultrasound usually ranges from 2 MHz to

10 MHz [12]. With this imaging technique, a probe (ultrasonic transducer) is used to produce a sound wave from electrical signals. Once the sound wave is delivered, it travels towards the tissues. Reflection of ultrasound (returns to the probe as an echo) occurs at tissue boundaries because of the differences in the acoustic impedances of the two tissues [12]. The distance from the probe to the tissue can be calculated based on the speed of sound and the travel duration [12]. After adjusting the angle of the probe several times, a set of tomographic images can be synthesized [12]. Compared with other imaging modalities, there are a few advantages in using US imaging techniques: (1) The portability of the imaging device, (2) The ability to perform real time examinations, (3) A lower treatment cost and (4) No ionizing radiation is required.

In brachytherapy, CT and ultrasound are currently the two major imaging modalities because of their widespread availability in the majority of radiation oncology departments. However, MRI guided brachytherapy is promising, as it can offer an excellent distinction between soft tissues, especially in prostate and gynecologic cancers [14–16].

Contouring

With digital visualizations from imaging modalities it is possible to delineate the regions of interest (ROIs). In radiation therapy, ROIs include both the target volume for treatment and its surrounding normal critical structures or organs at risk (OARs).

For the target volume, there are three definitions according to the International Commission on Radiation Units and Measurements (ICRU) Report No. 50 [17]. The Gross Target Volume (GTV) is defined by the physician as all visible disease, as determined after reviewing medical images and clinical information [17, 18]. The Clinical Target Volume (CTV) includes areas at risk for microscopic disease extension [17, 18]. Usually, the scan must include all of the CTV with at least 9 mm superior and inferior margin [18]. The Planning Target Volume (PTV) aims to compensate for the variability of treatment setups and internal organ motion [17, 18]. In brachytherapy, the PTV is either the same as the CTV or a 2-3 mm margin is added anteriorly and laterally [18] because there are no significant opportunities for setup errors [9].

In addition to treating the disease, it is also important to keep the OARs at a low-level toxicity. Due to the different deliveries of the radiation between EBRT and brachytherapy (see Figures 0.3 and 0.4), the contouring of OARs in prostate cases slightly vary (bladder, rectum and femoral heads in EBRT [19, 20], bladder, rectum and urethra in HDR brachytherapy [18]). In Figure 0.6b, the target volume and OARs were contoured by an experienced radiation oncologist.

3D reconstruction of organs and catheters

The determination of dwell positions (source positions in catheters) is part of the plan optimization in brachytherapy. Therefore, it is necessary to first reconstruct the catheter positions (Figure 0.6c) from the CT images before the plan optimization. The reconstruction can either base on the use of the radio-opaque markers with metal wires [21], or the air filled inside the catheters on the CT images, as the Hounsfield Unit (HU) value of air is significantly different to the surrounding soft tissue. Recent studies show that it is possible to use electromagnetic (EM) tracking systems for catheter reconstructions [22, 23].

In order to keep the normal tissue at a minimum level of irradiation, there is no need to stop the radioactive sources outside the target [18]. In other words, source dwell positions are only activated (or turned on) if these positions are located inside the target volume.

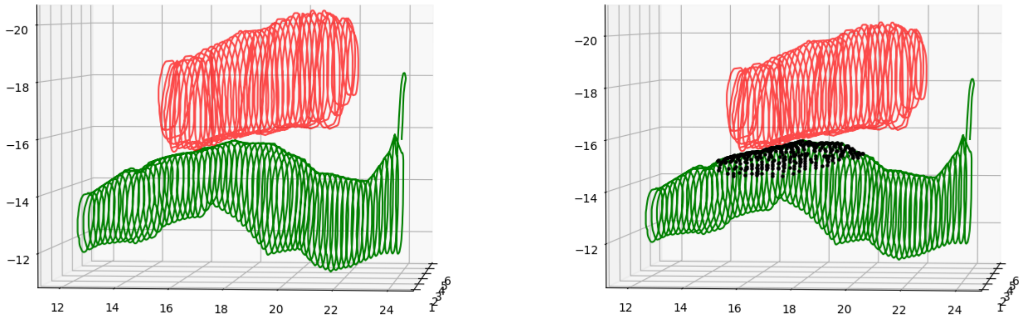
Treatment planning and plan evaluation

Dose calculation points creation

From an image dataset, 3D Cartesian coordinates of the contoured organs and the reconstructed catheters are extracted. In order to evaluate if a dose distribution is adequate, it is necessary to create a set of dose calculation points and make an evaluation based on the dose values at these points. The number of dose calculation points, which are created from contours such as those shown in Figure 0.8 should be minimized in order to accelerate the calculations [24]. Three types of dose calculation points were therefore introduced in IPSA [24]: surface, volume, and dose-volume histogram (DVH) dose calculations (Figures 0.8b-d). Surface dose calculation points (uniformly generated on the contours) and volume dose calculation points (created inside the organ volumes) are required only where the dose needs to be controlled [24]. For example, the surface and the volume dose calculation points of the rectum/bladder are generated only within a limited area around the catheters to reduce their numbers. As a result, the numbers of surface and volume dose calculation points are much less than the number of DVH dose calculation points that are uniformly generated throughout the whole volume (e.g. in Figure 0.8, surface: 392, volume: 566, DVH: 8201). This strategy can result in a high computational efficiency in IPSA because only a small number of surface and volume dose calculation points are repetitively used during the plan optimization process.

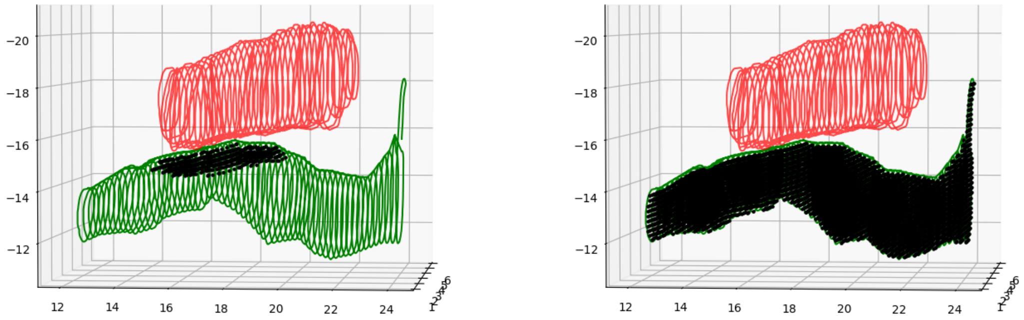
Source definition and dose calculation

A wide range of brachytherapy sources are available [25], and the type of source needs to be specified before the treatment. Once the source type, active dwell positions and dose calculation



(a) Contours

(b) Surface dose calculation points



(c) Volume dose calculation points

(d) DVH dose calculation points

FIGURE 0.8 – An example to illustrate three types of dose calculation points in IPSA: (a) Only contours (target in red, rectum in green), (b) Surface dose calculation points (number of points: 392), (c) Volume dose calculation points (number of points: 566), and (d) DVH dose calculation points (number of points: 8201).

points are determined, it is possible to calculate the dose kernel for each dwell position. A dose kernel is sometimes referred as a dose rate matrix (the dose rate contribution of a dwell position to the surrounding dose calculation points).

In HDR prostate brachytherapy, the dose calculation formalism was proposed in the American Association of Physicists in Medicine Task Group No. 43 (AAPM TG-43) report [26] and its updates [27, 28]. The TG-43 formalism is based on the dose rate measured at a reference point in a spherical water phantom from a single source geometry [9]. The dose rates at other points are corrected for relative attenuation in water, through the radial dose function and anisotropy function [9]. The 1D and 2D equations used to calculate dose rate in the TG-43 report update [27] are described in Equation (1-2) respectively:

$$\dot{D}(r) = S_k \cdot \Lambda \cdot \frac{G_L(r,\theta_0)}{G_L(r_0,\theta_0)} \cdot g_L(r) \cdot \phi_{an}(r) \tag{1}$$

$$\dot{D}(r,\theta) = S_k \cdot \Lambda \cdot \frac{G_L(r,\theta)}{G_L(r_0,\theta_0)} \cdot g_L(r) \cdot F(r,\theta) \tag{2}$$

with dosimetry parameters defined below:

- $\dot{D}(r)$ is the 1D calculated dose rate at radial distance r.
- $\dot{D}(r,\theta)$ is the 2D calculated dose rate at radial distance r, and polar angle θ.
- S_k is the air-kerma strength of the brachytherapy source, with a unit of $\mathrm{U{=}cGy\,cm^2\,h^{-1}}$.
- Λ is the dose rate constant, with a units of $\mathrm{cGy\,h^{-1}U^{-1}}$.
- $G_L(r,\theta)$ is the geometry function, defined as $\frac{\beta}{L \cdot r \cdot sin\theta}$, if $\theta \neq 0°$, and $(r^2 - L^2/4)^{-1}$, if $\theta = 0°$.
- L is the active length of a brachytherapy source.
- r_0 is the radial distance of a reference position, with a value of 1 cm.
- θ_0 is the polar angle of a reference position, with a value of 0°.
- $g_L(r)$ is the line-source radial dose function, defined as $\frac{\dot{D}(r,\theta_0)}{\dot{D}(r_0,\theta)} \cdot \frac{G_L(r_0,\theta_0)}{G_L(r,\theta_0)}$.
- $\phi_{an}(r)$ is the 1D anisotropy function.
- $F(r,\theta)$ is the 2D anisotropy function.

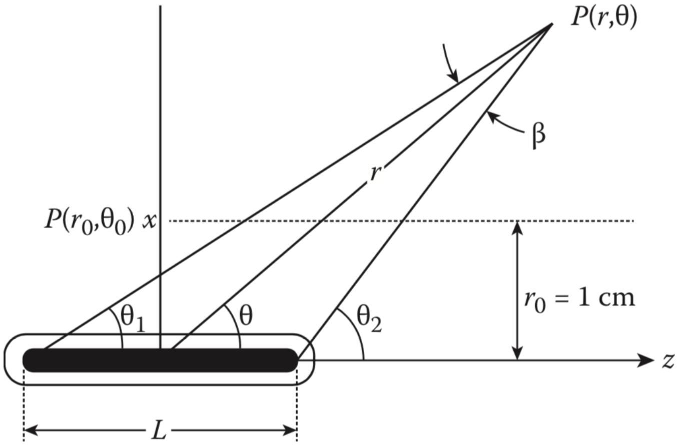


FIGURE 0.9 – Polar coordinate system used in 2D TG-43 dose calculation formalism. The dose value at any point $P(r,\theta)$ is a function of its radial distance r and polar angle θ. A reference position is defined at $P(r_0,\theta_0)$, where $r_0 = 1$ cm and $\theta_0 = 90°$. Image from [9].

More details about these dosimetry parameters can be found in the TG-43 report [26] and its updates [27, 28]. Recommended dosimetry parameters can be found in the MD Anderson Imaging and Radiation Oncology Core (IROC) database [25] and the Carleton Laboratory for Radiotherapy Physics (CLRP) database [29]. These databases are based on the report from the High Energy Brachytherapy Source Dosimetry (HEBD) Working Group [30]. An uncertainty analysis of these parameters can be found in the AAPM Task Group No. 138 report [31].

Briefly, the TG-43 dose calculation formalism aims to estimate the dose distribution of a source that is placed in an environment similar to a water phantom [32]. The dose distribution in brachytherapy treatment planning can be calculated by superposing the precalculated TG-43 dose distributions for a single source from all dwell positions [32]. The TG-43 formalism and the superposition strategy are fast and therefore practical, but the effect of heterogeneity (tissue or applicator) and interseed attenuation are not considered [32].

To address the limitations of the TG-43 formalism, various model-based dose calculation algorithms (MBDCAs) such as collapsed-cone convolution [33], superposition- convolution [34], Monte Carlo methods [35, 36], and grid-based Boltzmann solver [37, 38] have been proposed. Among these algorithms, Monte Carlo methods are often considered as the gold standard for 3D dose calculations [36]. Further details as well as the guidance to ensure practice uniformity when implementing MBDCAs can be found in the AAPM Task Group No. 186 report [32]. In this work, such algorithms are not used for optimization purposes, only TG-43. However, a demonstration was conducted where MDBCAs could be used to recalculate 3D dose kernels for all possible dwell-positions in the geometry and in place of TG-43 in the optimization process [39].

Plan optimization

For each active dwell position (inside the target), the dose distribution is a multiplication of the dose kernel (calculated with Equation (1)/(2)) with the corresponding dwell time value. As a result, the final dose distribution is a superposition of dose distributions of all dwell positions (Equation (3)),

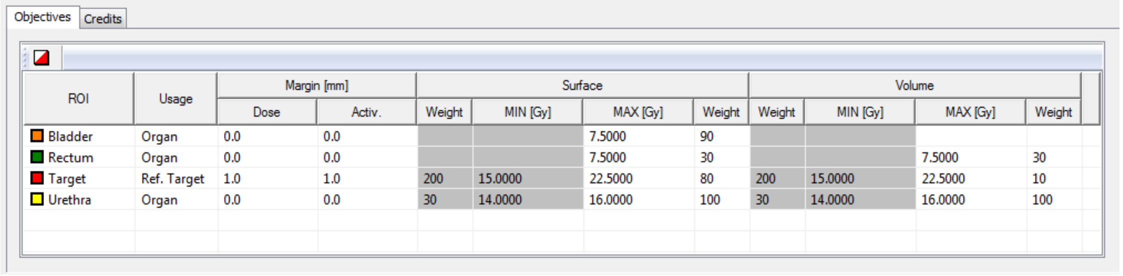
$$D_{ij} = \sum_{l=1}^{N_{\text{act}}} \dot{D}_{ijl} t_l \tag{3}$$

where D_{ij} is the dose at the i^{th} dose calculation point of the j^{th} organ, N_{act} is the number of active dwell positions, $\dot{D}_{ijl}$ is the dose rate contribution of the l^{th} dwell position to the i^{th} dose calculation point in the j^{th} organ, and t_l is the dwell time of the l^{th} dwell position.

As mentioned previously, clinical prescriptions in radiation therapy often consist of a goal for the target volume, and several goals for OARs. The purpose of the plan optimization is to find a final dose distribution that meets these goals, through determining the dwell times of all dwell positions. After translating these goals into mathematical objective functions and accumulating them into a single function (known as the weighted sum method, in Equation (4)), an optimization algorithm can be used to obtain a solution to the summed function,

$$OF = \sum_{j=1}^{N_{\mathrm{O}}} w_j \cdot \frac{1}{N_{pnt,j}} \sum_{i=1}^{N_{\mathrm{pnt,j}}} f_j(D_{i,j}, D_{limit,j}) \tag{4}$$

where OF stands for the summed objective function and N_O for the number of organs. For the j^{th} organ, w_j is the weight (or called importance factor, to steer the trade-offs among organs), $N_{pnt,j}$ is the number of dose calculation points, f_j is the objective function, $D_{i,j}$ is the dose value at the i^{th} dose calculation point, and $D_{limit,j}$ is the dose limit in the prescription. These defined objective functions can be initialized with a planning template, called the IPSA class solution (Figure 0.10). More details about the plan optimization in HDR brachytherapy can be found in Chapter 1.



ROI	Usage	Margin [mm]		Surface				Volume			
		Dose	Activ.	Weight	MIN [Gy]	MAX [Gy]	Weight	Weight	MIN [Gy]	MAX [Gy]	Weight
🟧 Bladder	Organ	0.0	0.0			7.5000	90				
🟩 Rectum	Organ	0.0	0.0			7.5000	30			7.5000	30
🟥 Target	Ref. Target	1.0	1.0	200	15.0000	22.5000	80	200	15.0000	22.5000	10
🟨 Urethra	Organ	0.0	0.0	30	14.0000	16.0000	100	30	14.0000	16.0000	100

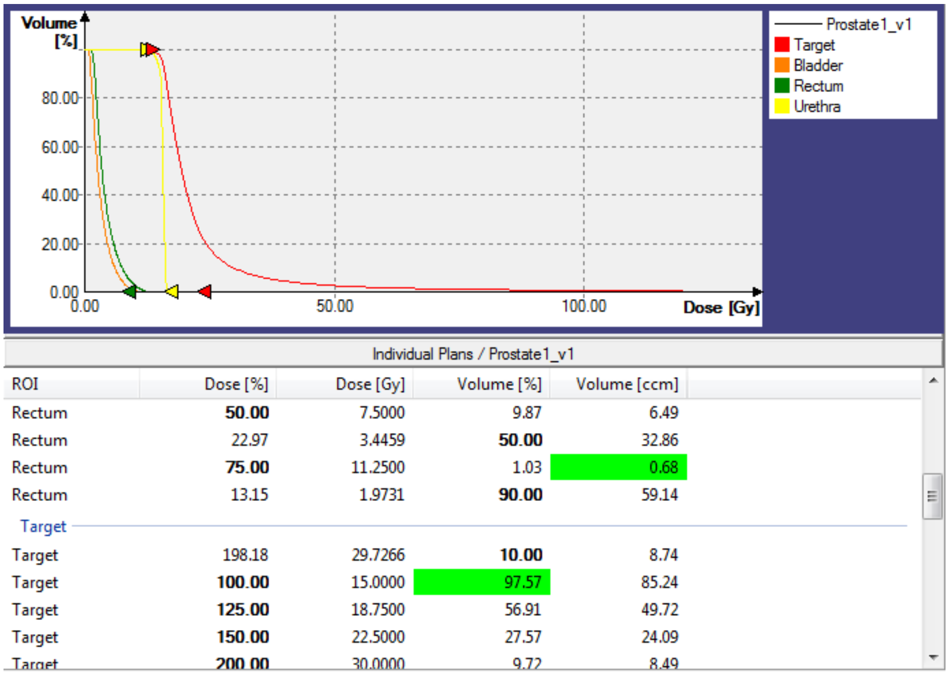
FIGURE 0.10 – The IPSA class solution (the weights and the dose limits) used for HDR prostate brachytherapy in our center (a screenshot of Oncentra Brachy).

Plan evaluation

It is important to evaluate if the dose distribution of the obtained treatment plan is acceptable. The most commonly tool used to evaluate dose distributions for treatment planning is the dose-volume histogram (DVH), a histogram that shows the relationship between radiation dose and tissue volume. There are two types of DVH: the differential DVH and the cumulative DVH. While the differential DVH is defined as the volume of the organ receiving a dose within a specified dose interval, the cumulative DVH represents the volume receiving a minimum specified dose value [40]. The latter is the preferred format in clinical practice because it is easier to evaluate the total size of tissue volume in hot spots (overdose, dose above the maximum dose value) or cold spots (underdose, dose below the minimum dose value) [40].

To ease the evaluation procedure, a variety of dose-volume metrics can be extracted from the DVH results. These metrics are sometimes referred as dosimetric parameters, and can be categorized into $V_x\%$ (the volume of tissue receiving x% of the prescription dose) and $D_x\%$ (the dose covering at least x% of the tissue volume). The Radiation Therapy Oncology Group (RTOG) recommended parameters for HDR prostate brachytherapy can be found in the following protocols [18, 41]. An example of a plan evaluation based on DVH results and dosimetric parameter results is shown in Figure 0.11. There is one major disadvantage in using

DVHs which is the lack of anatomical location of hot or cold spots. To solve this issue, graphic dose distributions such as isodose distributions (curves that indicate an equal radiation dose, see Figure 0.6d) are often used [40]. In addition, dose homogeneity (the degree of homogeneity of a dose distribution over the volume) and dose conformity (the degree of the entire volume receiving the optimal dose) are sometimes used to evaluate treatment plans.



Individual Plans / Prostate1_v1

ROI	Dose [%]	Dose [Gy]	Volume [%]	Volume [ccm]
Rectum	**50.00**	7.5000	9.87	6.49
Rectum	22.97	3.4459	**50.00**	32.86
Rectum	**75.00**	11.2500	1.03	0.68
Rectum	13.15	1.9731	**90.00**	59.14
Target				
Target	198.18	29.7266	**10.00**	8.74
Target	**100.00**	15.0000	97.57	85.24
Target	**125.00**	18.7500	56.91	49.72
Target	**150.00**	22.5000	27.57	24.09
Target	**200.00**	30.0000	9.72	8.49

Figure 0.11 – DVH results and dosimetric parameter results (the recommended ones for evaluations are highlighted) of an optimal treatment plan for prostate cancer in Oncentra Brachy (top: cumulative DVHs, bottom: dosimetric parameters).

High performance computation

A vision 20/20 paper [42] reviewed current applications of automation and advanced computing in radiotherapy. The authors envisioned that four main areas including cloud computing, aggregate data analysis, parallel computation and automation might change the practice in radiation oncology. This thesis involves an application of parallel computation to speed up the computational tasks in treatment planning, and an application of automated planning to mitigate the repetitive tasks and to improve productivity.

The history in computation and main parallel computation techniques are briefly reviewed in this section, and a detailed literature review on automated planning will be discussed in Chapter 1.

A brief history of computation

The idea of a digital computer can be dated back to 1931, when C. E. Wynn-Williams suggested that digital electronics can be used for computation [43].

In 1946, the announcement of ENIAC (Electronic Numerical Integrator And Computer) was known as the first electronic general-purpose computer available to the public. With ENIAC, computing problems can be solved through programming for the first time.

In 1971, Intel (Santa Clara, CA) introduced the world's first commercial microprocessor chip (Intel 4004), which led to the development of microcomputers. Due to technical advances, the dimension of transistors has been reduced. As a result, more and more transistors (Figure 0.12) can be fit on the chip of the same size, and the computing efficiency of microprocessors experienced significant improvements.

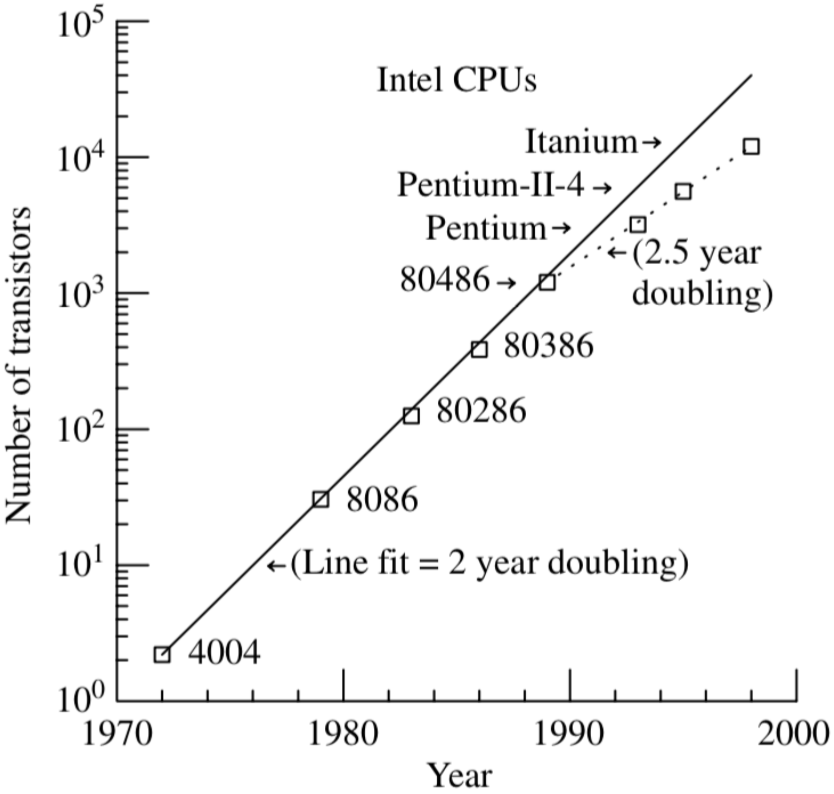


FIGURE 0.12 – Moore's Law: the number of transistors per chip doubles about every two years, image from [44].

As chip-making technologies are approaching their physical limits (i.e. the number of transistors per chip), keeping up with Moore's law is a challenge, however, the industry managed to improve the computational performance by adding extra microprocessor cores on a single chip. For example, IBM (Armonk, New York) introduced the first commercial multi-core processor

(POWER4) in 2001. The performance of a single chip can be increased significantly (up to a factor of the number of physical cores), as each core can simultaneously execute instructions.

Central processing unit (CPU) parallel computation

In order to fully utilize the capabilities of the multi-core architecture, two types of application programming interfaces (APIs) are usually used: Open Multi-Processing (OpenMP) and Message Passing Interface (MPI).

OpenMP is not a new programming language, and it instead, consists of a set of compiler directives that describe how the work will be shared by different cores [45,46]. As a result, OpenMP greatly simplifies writing multi-threaded programs in Fortran, C and C++ for programmers.

Here is an example to explain the OpenMP core syntax (written in C, executed with a four-core CPU):

```
#include <omp.h>

int main(void)
{
    int var1, var2;
    execute_serial_code();

    #pragma omp parallel num_threads(4) shared(var2)
    {
        execute_parallel_code();
    }

    resume_serial_code();
}

$ gcc -fopenmp main.c -o main
```

In this example, the main code is executed serially, like any C or C++ program. In the beginning, it is mandatory to have an include file that defines the functions used by OpenMP. OpenMP compiler directives always start with *#pragma omp*, followed by a specific keyword that identifies the directive [45]. For example, *#pragma omp parallel* instructs the compiler to run the enclosed code in parallel, and the following keywords specify the number of threads and the variable to be shared by all threads. In the end, it is a command to compile this OpenMP program.

OpenMP is used for parallelism on the multi-core level within a node (a single computation device). On the other hand, MPI is used for parallelism within a distributed computing system, which usually consists of thousands of computation devices often known as massively parallel computing, distributed computing, or grid computing [46].

Graphics processing unit (GPU) parallel computation

While a CPU is a general-purpose processor, a GPU has been a special-purpose processor optimized for image or video data processing. This is because GPUs can have more arithmetic logic units (ALUs) compared to CPUs (Figure 0.13), and these ALUs are responsible for all the required arithmetic and logic operations. Therefore, GPUs are more efficient in processing massive graphic data in parallel.

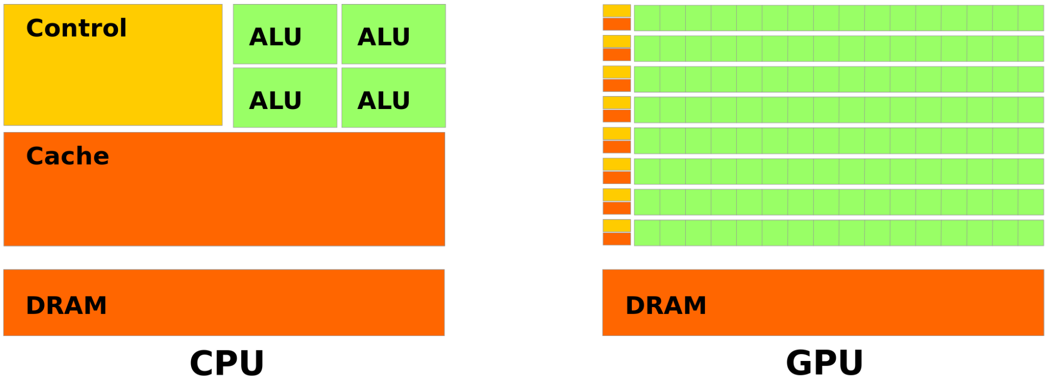


FIGURE 0.13 – CPU architecture versus GPU architecture. Image downloaded from https://nvidia.com in April 2019.

NVIDIA (Santa Clara, CA) is a major player in the GPU field. In 1999, the company launched the world's first GPU (GeForce 256), and the term of GPU became popular afterwards [47]. The company launched the first programmable GPU (GeForce 3) in 2001, and released its GPU programming model (CUDA, Compute Unified Device Architecture) in 2007 [48]. CUDA allows researchers and developers to design programs without understanding the computing architecture of GPU. They can write a kernel function in C or C++, with additional key words to express parallelism, and execute it simultaneously by thousands of threads [46]. It is worthwhile to note that CUDA is specific to NVIDIA GPUs [48].

Since the launch of CUDA, there has been a tendency for programmers to fully utilize the computing capabilities of GPUs in general-purpose computing (including non-graphic tasks) [48]. For example, the number of GPU-related research articles published in the medical physics community has been monotonically increasing since 2007 [49]. The authors indicated that GPUs can offer a high computing power suitable for both imaging and therapy related problems [49].

In addition to CUDA, OpenCL (Open Computing Language) is another GPU programming model. One main advantage of OpenCL is its ability to cross various platforms including CPUs, GPUs, digital signal processors (DSPs), and field-programmable gate arrays (FPGAs) [48, 50].

Description of the research project

Problem

A desirable output of radiation therapy is a dose distribution that meets the criteria recommended by the Radiation Therapy Oncology Group (RTOG) or other organizations. To achieve this, it is necessary to determine an optimal planning parameters to configure treatment modalities. This process is essentially a computerized optimization which is known as inverse planning.

Current inverse planning algorithms usually require dosimetrists or medical physicists to determine the patient-specific trade-offs between the target and healthy organs under the supervision of oncologists. The process usually starts with a population based planning template (Figure 0.10), and involves an inevitable fine-tuning of the corresponding weights in the template.

This process could result in the following issues:

1. There are large dosimetric variations in treatment planning among medical institutions [51].
2. It is likely that the dosimetric results of treatment plans are superior in experienced centers than in less-experienced centers [52].
3. The total average time of planning, in particular in an operating room setting, can be become prohibitively long if too many iterations are performed [53].

Consequently, these issues can limit further advancements, as the dosimetric variations would complicate multi-institutional clinical trials [51], and furthermore the low planning efficiency can increase the overall cost of the operation [53].

Purpose and objective of the thesis

The goal of this thesis is to design and validate an operator-free inverse planning framework for HDR brachytherapy. The framework aims to improve the treatment planning efficiency and the treatment plan quality comparing to the available inverse planning algorithms.

After a literature review on traditional and operator-free inverse planning algorithms in Chapter 1, we can make two important hypotheses. First, patient-specific treatment plans can always be found after exploring the solution space with a multi-criteria optimization (MCO) algorithm. Second, an appropriate optimization algorithm and high performance computing techniques can fundamentally boost the computational performance of the implemented MCO algorithm.

In order to respond to the hypotheses, the project was divided into four objectives:

1. Develop and integrate an MCO algorithm on top of a traditional dose optimization engine to automatically generate a set of treatment plans with various trade-offs.
2. Upgrade the implemented MCO algorithm to a knowledge-based MCO (kMCO) algorithm, after incorporating prior knowledge and parallel computation techniques. Finally, the kMCO plans were compared with the IPSA physician-approved plans.
3. Develop and validate a quasi-Newton method to improve the computational performance without jeopardizing the quality of solutions.
4. Propose a GPU-based MCO (gMCO) algorithm on top of the quasi-Newton method to compute hundreds of plans in parallel, and compare the gMCO plans with the IPSA physician-approved plans.

Chapitre 1

Methodology

Optimization is the action of finding the best solution for a given problem (Figure 1.1). The process of identifying objectives, optimization variables, and constraints of a problem is known as modeling [54]. Constructing an appropriate model is the first, and sometimes most important step in the optimization process because (1) If the model is too simplistic, it will not give useful insights into the practical problem, and (2) If the model is too complex, it may be too difficult to solve the problem [54].

Once the problem has been formulated into a model, an optimization algorithm can be used to find its solution, usually with the help of a computer [54]. Choosing an appropriate algorithm is important because it can determine whether the problem can be solved at all, or whether it can be solved rapidly or slowly [54]. Usually, mathematical objective functions are used for checking if the current solution is indeed the optimal solution of the problem [54].

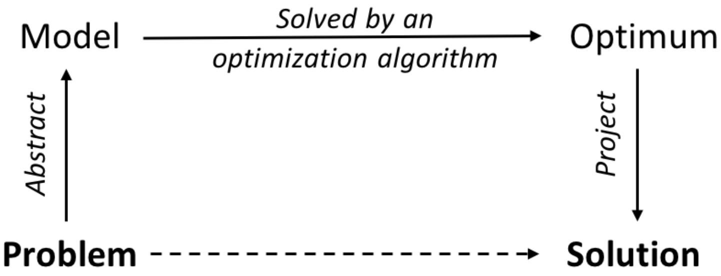


Figure 1.1 – The process of building a mathematical model of a problem and solving the model with an optimization algorithm. Image downloaded from https://www.eudoxus.com in April 2019.

In the context of high dose rate (HDR) brachytherapy, treatment planning is often considered as an optimization problem where the solution consists of a set of dwell times (source stopping duration) and dwell positions (source positions in catheters). The first part of this chapter consists of a literature review on mathematical objective functions and optimization algorithms used in HDR brachytherapy over the past two decades. The second part of this chapter includes a literature review on operator-free inverse planning algorithms that were proposed to facilitate the manual planning procedure in external beam radiation therapy (EBRT) and brachytherapy. While the first part of the literature review laid a foundation for finding a fast optimization algorithm, the second part gave an inspiration of how to mitigate the trial-and-error efforts in traditional planning, which is the main purpose of this study.

1.1 A review of mathematical objective functions in HDR brachytherapy planning

1.1.1 Dose-volume based objective functions

The ultimate goal in treatment planning is to find an optimal dose distribution inside the treatment target volume and OARs. In order to achieve this goal, an intuitive approach for both physicists and physicians is to formulate a set of objective functions to illustrate the relationship between dose and volume. These function are dose-volume based, and are widely used in HDR brachytherapy [15, 24, 55–61].

1.1.1.1 Underdose and overdose

An optimal dose distribution can be specified as: a dose which is not only high enough to kill the cancerous cells inside the target volume, but also low enough to keep the surrounding organs at a low-level toxicity.

A collection of dose-volume terminologies, including the maximum dose, the average dose, the underdose and overdose limitations can be used to find an optimal dose distribution. Among them, underdose and overdose, referring to the dose below and above a threshold (or dose limit) (Figure 1.2b-d), respectively, are often used in brachytherapy.

For the target, applying underdose and overdose limitations can steer the dose distribution towards the prescription dose because any dose values outside the scope will be penalized (Figure 1.2b-c). On the other hand, an underdose limitation is counterproductive for healthy OARs, for which lower dose values are acceptable. Therefore, applying an overdose limitation to OARs is necessary, and will restrict a high dose distribution inside (Figure 1.2d).

It is worthwhile to mention that the urethra, being located within the treatment volume, is a special OAR in HDR prostate brachytherapy because it is necessary to protect this organ from being over-radiated by the implanted radioactive sources. For example, if the urethra is only subject to an overdose limitation, it may have cold spots (low dose regions) and the adjacent prostate may not be covered with adequate doses [62, 63]. Therefore, applying both underdose and overdose limitations to this special OAR is very common [15, 58, 60]. For the target volume, the underdose and the overdose limitations are typically 100 % and 150 % of the prescription dose, correspondingly [15, 58, 60]. For the urethra volume, the underdose limitation is 100 % of the prescription dose, and the overdose limitation is either 113 % or 120 % of the prescription dose [15, 58, 60].

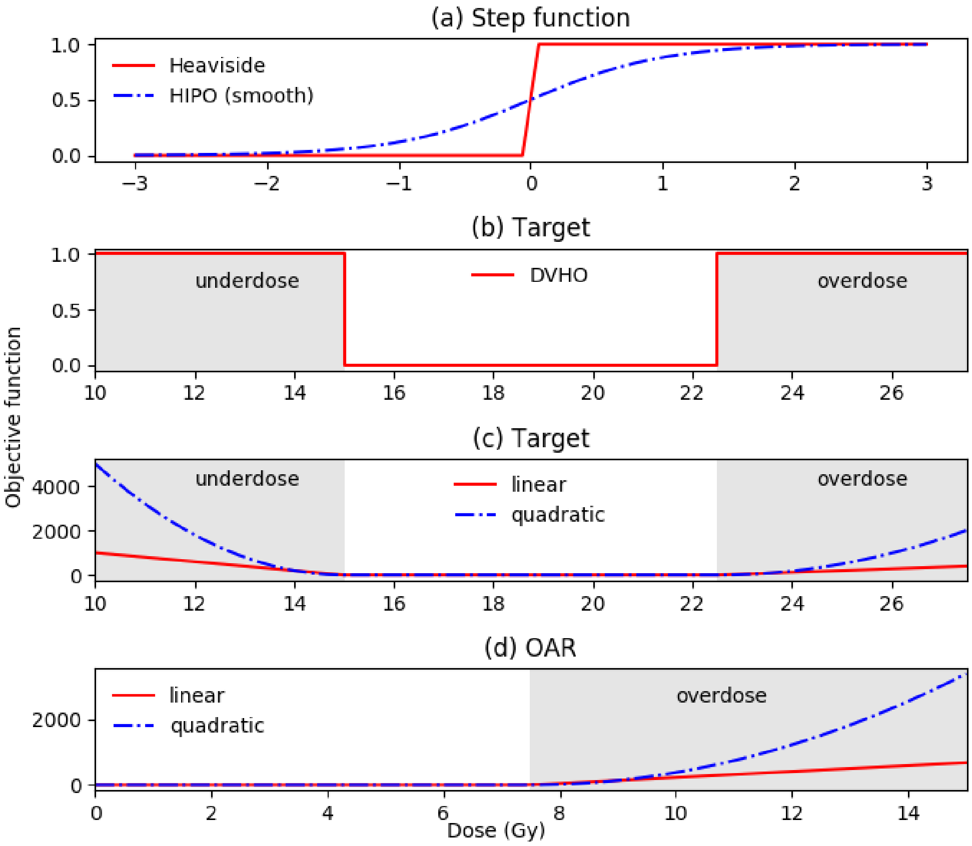


FIGURE 1.2 – Underdose and overdose objective functions of: (a) Step functions (solid line: Heaviside, dashed line: A smooth approximation (used in HIPO) defined in Section 1.2.2.1), (b) Target (DVH objective, or DVHO), (c) Target (linear and quadratic) and (d) OAR (bladder). The values of dose limits and weights in (c) and (d) are identical to values used in references [64–66].

Based on how to penalize the difference between the ideal and actual dose distribution, objective functions can be divided into three categories: dose volume histogram objectives (DVHO), linear objectives and quadratic objectives [67, 68], depending on the α value in Equation (1.1-1.2),

$$f_U(x) = \sum_{i=1}^{N} H(D_U - d_i(x))(D_U - d_i(x))^{\alpha} \tag{1.1}$$

$$f_O(x) = \sum_{i=1}^{N} H(d_i(x) - D_O)(d_i(x) - D_O)^{\alpha} \tag{1.2}$$

where f_U and f_O are the objective functions to be minimized, with underdose and overdose limitations correspondingly. N is the number of dose calculation points (can be created based on the strategy in Section). x is a treatment plan that contains dwell positions and dwell times. H is the Heaviside step function (solid line in Figure 1.2a). Underdose limit D_U and overdose limit D_O are assigned by an oncologist before the optimization. $d_i(x)$ is the dose value for the i^{th} dose calculation point.

If $\alpha = 0$, Equation (1.1-1.2) are in the form of DVH objectives [59–61, 67], as the objective value f_O is the cumulative DVH value at the dose D_O (Figure 1.2b). One advantage of this category is that the objective value, which equals the number of voxels in which the dose value is higher or lower than the dose limit value, is more intuitive to understand than the linear or quadratic objectives.

If $\alpha = 1$, Equation (1.1-1.2) are in the form of linear objectives [24, 55], as the difference between the ideal and actual dose distribution is penalized linearly (solid lines in Figure 1.2c-d). This category of objective functions is sometimes coupled to a stochastic optimizer such as simulated annealing in IPSA (Inverse Planning Simulated Annealing) (Elekta Brachy, Veenendaal, The Netherlands).

If $\alpha = 2$, Equation (1.1-1.2) are in the form of quadratic objectives [56, 57], as the difference between the ideal and actual dose distribution is penalized quadratically (dashed lines in Figure 1.2c-d). This category of objective functions is usually coupled to a deterministic optimizer [56, 57].

1.1.1.2 Dose homogeneity index and dose conformity index

In 1988, Wu *et al.* [69] proposed a dose homogeneity index (HI) to evaluate the dose homogeneity of treatment plans in HDR breast brachytherapy. As shown in Equation (1.3), HI is equivalent to the fraction of the total treatment volume which receives a dose between 100 % and 150 % of the prescribed dose.

$$HI = \frac{V_{100} - V_{150}}{V_{100}} \tag{1.3}$$

Where V_{100} and V_{150} are the fraction of the total treatment volume which receives at least 100 % and 150 % of the prescription dose, respectively.

Ten years later, Baltas *et al.* [70] proposed a conformal index (COIN) to replace the DVH when evaluating the plan quality in HDR brachytherapy. The definition of COIN is expressed by Equation (1.4),

$$\begin{aligned} COIN &= c_1 \times c_2 \\ &= \frac{PTV_{ref}}{PTV} \times \frac{PTV_{ref}}{V_{ref}} \end{aligned} \tag{1.4}$$

where PTV is the planning target volume (PTV), V_{ref} is the volume that is enclosed by the reference isodose curves and PTV_{ref} the overlap volume between PTV and V_{ref} (Figure 1.3). As such, the coefficient c_1 is the fraction for the PTV that is enclosed by the reference isodose curves, and the coefficient c_2 is the fraction of volume V_{ref} that is enclosed by the PTV.

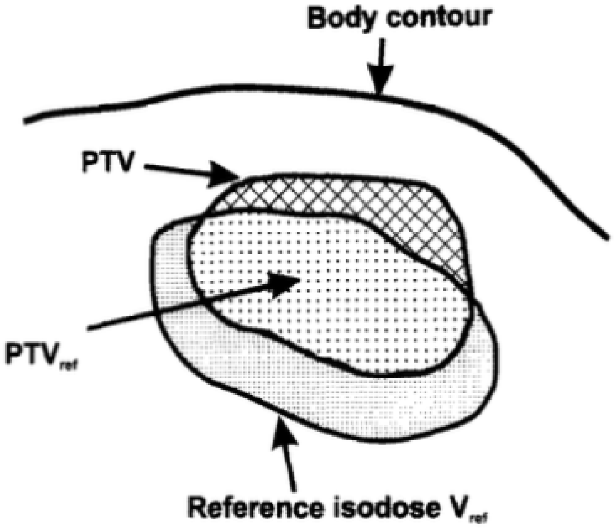


FIGURE 1.3 – Schematic diagrams of volumes in COIN computation, image from [70].

In summary, a high HI value means a uniform dose distribution within the treatment volume, and a high COIN value means a conformal dose distribution within the treatment volume. These indices can be used to select a final solution after generating a dataset of optimal solutions [56, 67].

1.1.2 Dwell time in objective functions

Dwell time is less common when formulating objective functions [71–73] compared to the widely studied dose-volume based ones [15, 24, 55–61].

Renner *et al.* [71] formulated an objective to minimize the total dwell time under the constraint that all the points inside the target receive at least the target dose. The purpose is to minimize the total reference air kerma (TRAK) or the accumulated dose inside the patient's body while meeting the treatment requirements, as these quantities are directly proportional the total dwell time.

Baltas *et al.* [72] proposed a dwell time gradient restriction (DTGR) in HDR prostate brachytherapy which was integrated into HIPO (Elekta Brachy, Veenendaal, The Netherlands) [74]. The purpose is to achieve a smooth source movement or a smooth dwell time distribution within the catheters. This DTGR parameter was implemented by considering the variation of the source dwell times as an additional objective function to the dose-volume ones. The DTGR parameter has a range from 0 (no dwell time restriction) to 1 (maximum dwell time restriction).

Cunha *et al.* [73] applied a similar strategy to obtain a smooth distribution of dwell times with IPSA. A plan was computed with IPSA, and the mean dwell time in each catheter was calculated from the optimized plan. A new plan of dwell times that were tightly clustered around the mean value in each catheter were obtained.

1.2 A review of optimization algorithms

In order to obtain a solution to the planning problem, it is necessary to solve it with an optimization algorithm. Optimization algorithms can either be stochastic or deterministic. Stochastic algorithms include randomness in the search-process. Such randomness not only mitigates the effect from modeling errors, but also enables the algorithms to approach a global solution. A global solution is the point with the lowest objective function of all feasible points [54].

On the other hand, in deterministic algorithms, the search-process is completely known (without randomness). These algorithms are usually superior in computing efficiency, but might be more sensitive to modeling errors and can be trapped in a local solution. A local solution is a point at which the objective function is smaller than all other feasible nearby points [54].

Convexity is an important concept in optimization. If the objective function and the feasible set of the optimization variables are both convex, then the optimization problem is convex, and local solutions are also global solutions [54, 75].

The objective function is convex if the line segment (or the chord) between any two points of the function lies above the function (illustrated in Figure 1.4) [75]. Mathematically, a convex function needs to be satisfied the following expression (Equation 1.5).

$$f(\alpha x + (1-\alpha)y) \leq f(\alpha x) + (1-\alpha)f(y) \tag{1.5}$$

Where f is the objective function. x, y are any two input optimization variables, and $0 \leq \alpha \leq 1$.

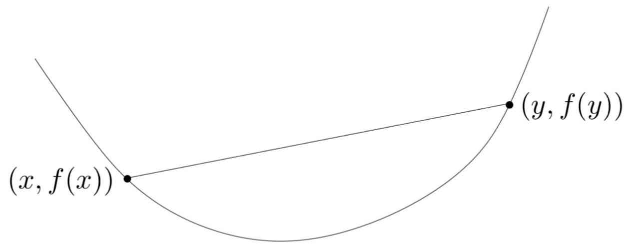


FIGURE 1.4 – Graph of a convex function. The chord (or line segment) between any two points on the graph lies above the graph, image from [75].

The feasible set of points that satisfying the constraints (if they exist) is convex if the line segment between any two points from this set lies in this set (satisfying Equation (1.6)) [75].

$$\alpha x + (1 - \alpha)y \in S \tag{1.6}$$

Where x, y are any two points that belong to the feasible set S, and $0 \leq \alpha \leq 1$.

1.2.1 Stochastic optimization algorithms

Simulated annealing and evolutionary algorithms are two widely studied stochastic optimization algorithms in HDR brachytherapy [9].

1.2.1.1 Simulated Annealing

In 1983, Kirkpatrick first proposed simulated annealing (SA) as a stochastic optimization algorithm [76]. As seen from its name, SA was inspired from annealing in metallurgy, a technique that involves heating and controlled cooling of a material to increase the size of its crystals and reduce their defects [77].

Sloboda [78] first adapted SA in brachytherapy. The author used SA to optimize the dose distribution in LDR vaginal brachytherapy. Pouliot *et al.* [79] implemented SA to rapidly produce plans for permanent prostate implant treatments. Later, Lessard and Pouliot [55] developed the Inverse Planning Simulated Annealing (IPSA) algorithm to implement an anatomy-based inverse planning algorithm in HDR brachytherapy. IPSA is a commercialized stochastic optimization algorithm (Elekta Brachy, Veenendaal, The Netherlands) for HDR brachytherapy. As SA can approximate an acceptable global optimal solution [55], IPSA has gained popularity both in clinic and in research studies [15, 24, 56, 58, 60, 61, 80, 81].

In IPSA, linear dose-volume based objective functions are used for the target volume and OARs because the translation from the physician's idea of the ideal treatment to a mathematical form is straightforward [24]. For each organ, the dose limits and the corresponding slope values are assigned before starting the optimization. By minimizing the value of the objective function, a set of optimized dwell times can be found with SA.

Major steps in IPSA [55] are:

1. The initial solution of dwell times are set with arbitrary values (i.e. 1 s for all dwell positions within the target).
2. At the k^{th} iteration, the objective function value f_k can be calculated based on Equation (1.1-1.2) ($\alpha = 1$).
3. At the $k+1^{th}$ iteration, one random dwell position is selected, and the dwell time at this position is modified (increased or decreased, which is randomly decided) with a step size (0.1 s). The objective function value f_{k+1} is recalculated, and the change in the objective function during the transition is $\Delta f = f_{k+1} - f_k$.
4. While a solution that results in a smaller objective function value ($\Delta f < 0$) is always accepted, a solution that results in a larger objective function value ($\Delta f > 0$) will only be accepted with a certain level of probability, given by Equation (1.7),

   $$P(\Delta f) = exp[-\Delta f / T(k)] \tag{1.7}$$

   where $P(\Delta f)$ is the probability to accept a positive transition, Δf is the change in the objective function during a transition from the k^{th} iteration to the $k+1^{th}$ iteration, and $T(k)$ is the pseudo-temperature parameter at the k^{th} iteration, with an annealing schedule expressed in Equation (1.8),

   $$T(k) = T_0 / k^{\alpha} \tag{1.8}$$

   where $T(k)$ is the current temperature, T_0 is the initial temperature, which is set proportional to the mean Δf values over 500 random transitions f_k to f_{k+1} [55] (as the first transition is chosen at random, so this transition is repeated until a good statistical sample of the initial optimization problem condition is accumulated [24]). α is the annealing speed parameter, which is often empirically set to 0.6 to optimize the results while preserving a reasonable time [55].
5. The search process (steps 2-4) is repeated until termination (a large number of iterations, for example 100 000 in reference [82]).

1.2.1.2 Evolutionary algorithms

Evolutionary algorithms (EAs) represent a category of population-based optimization algorithms. As seen from the name, EA was inspired by biological evolution including the processes of selection, crossover and mutation.

In HDR brachytherapy, Lahanas *et al.* [83] formulated a multi-objective problem using dose-volume based objective functions. Each chromosome represents several dwell weights, or the fraction of a specific dwell time in the total dwell time. A set of optimal solutions were obtained with an EA. Finally, the decision maker chose one solution depending on the COIN and DVH results. Milickovic *et al.* [84] also formulated a multi-objective problem using dose-volume based objective functions. The problem was solved with two variants of EAs and one deterministic gradient based optimization algorithm. The authors claimed that the computational performance can be improved by initializing the population of EAs with a deterministic algorithm. Other studies about EAs can be found in the following references [67, 85–87].

1.2.2 Deterministic optimization algorithms

Due to the high computing efficiency, deterministic optimization algorithms are widely studied [56, 72, 88–91] and used in treatment planning systems (e.g. the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) (Elekta Brachy, Veenendaal, The Netherlands) and the Nelder-Mead simplex (NMS) (Varian, Palo Alto, CA)). A deterministic optimization algorithm is either gradient-based or gradient-free, depending on if the gradient of the objective function is used. In the following sections, the applications in each category are reviewed in the context of HDR brachytherapy.

1.2.2.1 Gradient-based algorithms

Gradient-based optimization algorithms are sometimes coupled with dose-volume based quadratic objective functions. For example, Milickovic *et al.* [56] formulated dose-volume based quadratic objective functions and compared the solutions obtained with three optimization algorithms: two gradient-based algorithms (L-BFGS and the Fletcher-Reeves-Polak-Ribiere (FRPR) algorithm) and one gradient-free algorithm (the modified Powell method). The authors concluded that the solutions obtained with deterministic optimization algorithms are global optimal, due to the constructed convex objective functions.

The Hybrid Inverse Planning Optimization method (HIPO) is a well-known example of a gradient-based optimization algorithm in brachytherapy. HIPO was introduced by Karabis *et al.* [88] and has been widely studied [72, 89–91] after being integrated into Oncentra (Elekta

Brachy, Veenendaal, The Netherlands). HIPO starts from an initial set of catheters and uses a simulated annealing algorithm to select and adjust catheter positions [88]. For each combination of catheters, HIPO uses L-BFGS for dose optimization. As in IPSA, a linear dose-volume objective function was used in HIPO, based on Equation (1.1-1.2) (when $\alpha = 1$) [74]. In order to avoid the discontinuity of the Heaviside step function H in Equation (1.1-1.2), HIPO employs the logistic smooth approximation to the step function (dashed line in Figure 1.2a): $H(y) = 0.5(1 + \tanh(\beta y))$, where β is a parameter corresponding to the sharpness of transition [74, 89].

Major steps of L-BFGS [54, 57, 92] are reviewed:

1. Starting with x_0, an initial solution, and H_0, an initial approximation of the inverse Hessian of f.
2. At the k^{th} iteration, determine a descent direction $p_k = -H_k \nabla f(x_k)$.
3. Line search: choose a step size $\alpha_k = arg \min_{\alpha>0} f(x_k + \alpha_k p_k)$.
4. Update $x_{k+1} = x_k + \alpha_k p_k$.
5. Update g_{k+1}, the gradient of f at x_{k+1}.
6. Update H_{k+1}, the approximation of the inverse Hessian of f at the $k + 1^{th}$ iteration.
7. $k = k + 1$, steps 2-6 are repeated until termination.

There are two important features of HIPO [72, 74, 89, 90]. First, operators can make a local adjustment of dose distribution by locking certain catheters (no change of dwell times at dwell positions for these catheters). The second feature is the aforementioned dwell time gradient restriction (DTGR) to achieve a smooth dwell time distribution within a catheter.

Poulin *et al.* [60] compared HIPO with IPSA for eight HDR prostate cases. The authors concluded that HIPO and IPSA gave similar dosimetric results under fixed catheter configurations.

1.2.2.2 Gradient-free algorithms

In addition to the gradient-based algorithms described above, other deterministic optimization algorithms are categorized into gradient-free algorithms in this section. As this category is not common in clinical brachytherapy, only linear programming (LP) algorithms will be discussed.

A LP problem is usually defined by a linear objective function and constraints (sometimes using clinical criteria) [82]. The global solution for a LP problem can be solved exactly and deterministically [82]. Alterovitz *et al.* [82] formulated the dwell time optimization problem as a LP problem and conducted a comparison against IPSA for twenty prostate cases. The results showed that the planning time with the LP method was less than 15 s, and that no significant difference in dosimetric indices ($P<0.01$) existed between the LP method and IPSA.

1.3 Operator-free inverse planning algorithms in EBRT

A treatment plan can be obtained through inverse planning that consists of an appropriate mathematical objective function and an optimization algorithm. As mentioned in Section , however, there is a limitation of inverse planning: trial-and-error adjustments of planning parameters to approach a clinically acceptable treatment plan.

In order to improve the planning efficiency, it is necessary to streamline the planning process and develop an inverse planning algorithm that can automatically generate a deliverable plan. Planning is a common problem shared in HDR brachytherapy and EBRT. Due to the limited references of this topic in brachytherapy, four major approaches to solving the planning problem in EBRT were firstly reviewed. These approaches were categorized as operator-free inverse planning algorithms because they do not require manual weight adjustment (although they might still require minor user intervention by picking a suitable plan).

1.3.1 Knowledge-based planning algorithm

In general, knowledge-based planning (KBP) algorithms use previous knowledge to gain a high computational efficiency and/or a high quality treatment plan. To be specific, KBP algorithms aim to create a valid plan by using an established relationship between patient-specific anatomical features and high quality physician-approved treatment plans. KBP algorithms have been commercialized in RapidPlan (Varian Medical Systems, Palo Alto, CA), and have been studied by many groups (a recent review [93]). In this section, we will review the KBP studies from the group at University of California San Diego (UCSD) [94] to illustrate the evolution of this approach.

In 2011, Moore *et al.* [95] proposed an algorithm to predict achievable OAR mean doses based on the overlap volume of the OAR with the PTV. This algorithm has been tested with intensity modulated radiotherapy (IMRT) prostate cases and IMRT head and neck cases. The authors concluded that this algorithm can either be an automated quality control tool or a plan generation tool in IMRT planning.

In 2012, Appenzoller *et al.* [96] proposed a quality control tool for IMRT planning to predict achievable OAR DVHs. The prediction was based on the correlation of the expected dose with the Hausdorff distance (the minimum distance between a point outside the PTV and the PTV surface). The authors trained the prediction model with 20 prostate plans and 24 head-and-neck plans. The KBP plans were evaluated by the physician who approved the original plan. The authors concluded that their mathematical framework can successfully predict achievable OAR DVHs based on individual patient anatomy.

In 2015, Shiraishi *et al.* [97] proposed an algorithm to predict DVH-based quality metrics (QMs) based on anatomical features in intracranial stereotactic radiosurgery (SRS). QMs include brain V_{10Gy} and the conformity index (COIN). The training cohort consists of 223 SRS plans, and the validation cohort consists of 20 plans. The authors concluded that the algorithm can predict SRS QMs precisely.

In 2016, Shiraishi *et al.* [98] established an artificial neural network (ANN) to predict 3D dose distributions based on anatomical and plan parameters. The network was trained with 12 prostate plans and 23 SRS plans, and the results were evaluated with 11 prostate plans and 20 SRS plans. The authors concluded that their knowledge-based 3D dose predictions for radiotherapy plans were highly accurate.

1.3.2 Multi-criteria optimization algorithm

In a single-objective optimization, the optimal solution is usually clearly defined [99]. However, in a multi-objective problem (conflicting objectives) there are a set of trade-off solutions because a single solution is hardly the best for all the objectives simultaneously [99]. A multi-objective optimization problem can be expressed in Equation (1.9),

$$\text{Minimize } F(x) = [f_1(x), f_2(x), ..., f_m(x)] \tag{1.9}$$

where F represents the vector of objective functions, m is the number of objective functions and x is the input optimization variable.

As shown in Figure 1.5, a solution is Pareto optimal if one objective can not be improved without decreasing at least one of the other objectives, and a collection of all Pareto optimal solutions is called the Pareto (optimal) surface or the Pareto frontier [99, 100].

As previously mentioned, treatment planning in radiation therapy is usually a multi-objective problem, and the objective of treating the target and the objectives of protecting the OARs are often conflicting. The multi-objective problem can be solved with a multi-criteria optimization (MCO) algorithm by computing the Pareto surface for each patient and selecting a plan that meets the clinical goals [102].

To the best of our knowledge, the concept of MCO in radiation therapy can be dated back to 1997, when Yu *et al.* [103] combined a decision-theoretic steering scheme and iterative optimization to explore the Pareto surfaces in the context of stereotactic radiosurgery and prostate implantation. Hamacher *et al.* [104] described the first use of MCO for intensity-modulated radiation therapy. The author started the optimization process with an initial solution, and the algorithm explored the neighborhood solutions until the Pareto surface is

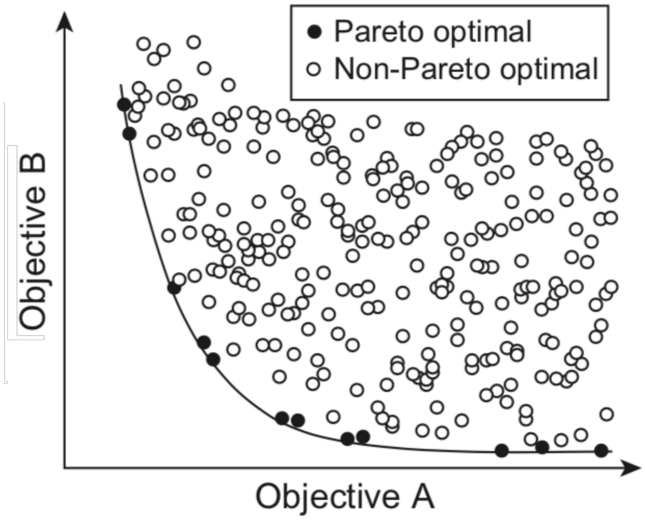


Figure 1.5 – Illustration of Pareto optimal solutions (black dots) and the Pareto surface (black curve) for a bi-objective optimization problem, image adapted from [101].

well reconstructed. With the aid of an appropriate online tool, the physician can select a plan that meets his/her wishes and preferences from the plan dataset in a few minutes.

One year later, Küfer *et al.* [105] specified more details about the method to populate the Pareto surface in the previous study [104]. There are two phases in this method: (1) A calculation of all extreme compromises by only optimizing one objective, and (2) An iterative interpolation based on triangulation. In the end, an Equivalent Uniform Dose (EUD) navigation tool is used to select the final plan. The computing time for constructing the plan dataset ranges 2 to 5 h on a 1.7 GHz Pentium IV platform with 2 GB RAM.

In 2006, Craft *et al.* [102] presented an algorithm to approximate Pareto surfaces. The algorithm automatically generates a Pareto optimal solution based on the geometric lower and upper bounds on the Pareto surfaces. The author illustrated the mechanism of the algorithm in details with one IMRT prostate case and one skull base case.

Applications of the MCO algorithm proposed by Craft *et al.* [102] has been extended to other clinical sites, to volumetric modulated arc therapy (VMAT) and to proton therapy. For example, Chen *et al.* [106] adapted an MCO algorithm to solve the intensity modulated proton therapy (IMPT) planning problem for three clinical cases (a pancreas case, an esophagus case, and a case where the tumor is located along the rib cage). Craft *et al.* [107] used the MCO algorithm to solve the fluence map optimization and angle configurations simultaneously. The algorithm was tested with one IMRT pancreas case under 100 different five-beam configurations [107]. Craft *et al.* [108] adapted the MCO algorithm to solve the VMAT planning problem (an angular grid of 180 equi-spaced beams). The algorithm was tested with one prostate, one pancreas, and one brain case. The authors concluded that a high quality plan with a single arc can be created within 5 min [108].

After RaySearch Laboratories (Stockholm, Sweden) integrated the aforementioned MCO algorithm [102] into its Raystation treatment planning system, several studies [109–111] reported their comparisons between the MCO algorithm and a standard treatment planning system. For example, Wala *et al.* [109] compared the Raystation MCO algorithm against a standard planning algorithm (Corvus, Nomos, PA) for 9 IMRT prostate cases. The results showed that the MCO algorithm is superior to the standard planning algorithm for all 9 cases. The planning times per case are 1 to 8 h (standard) and 35 min (MCO). The authors concluded that the MCO algorithm is a valuable tool for IMRT prostate cases.

Wo *et al.* [110] compared the Raystation MCO algorithm against a standard planning algorithm (XiO, CMS, St. Louis) for 10 IMRT anal cases. The results showed that there was an increase in target dose heterogeneity in MCO plans. With MCO, the OAR mean dose can be reduced (6/10 in the large bowel, and 9/10 in femoral heads). The navigation time was less than 10 min for each case. The authors concluded that the MCO implementation is feasible in treatment planning for anal cancer patients.

Craft *et al.* [111] compared the Raystation MCO algorithm against a standard planning algorithm (XiO, CMS, St. Louis) for 10 IMRT cases (5 with glioblastoma and 5 with locally advanced pancreatic cancers). Results indicated that the MCO plan was preferred in all cases. The mean planning times were 156 min (standard) and 12.4 min (MCO) for glioblastomas, 114 min (standard) and 11.6 min (MCO) for pancreatic cancers. The authors concluded that the MCO algorithm is a new and efficient clinical tool that is able to aid the decision making process for physicians in treatment planning.

1.3.3 Auto-planning algorithm

In order to improve the planning efficiency, various class solutions are usually developed for simple cases (e.g., localized prostate, whole breast) [112], however, treatment planning can still be time-consuming for complex cases which have multiple target volumes (e.g. head and neck, pelvic nodes). Sometimes, treatment planning can include extra optimization structures, beam geometry adjustments, and changes in the objectives to reduce cold and hot spots in the dose distribution [112].

The Auto-Planning (AP) module, an option in Pinnacle (version 9.10, Philips Medical Systems, Best, The Netherlands), was developed to mimic these planning tricks mentioned above. The core part of the AP algorithm can be derived from the regional optimization introduced by Cotrutz and Xing [113]. It begins with defining and automatic segmentation of hot and cold spots from an optimized treatment plan, introducing new objectives and relative weights of these spots into the optimization process, re-optimizing the plan, and repeating these steps until a termination criteria [112, 114]. Therefore, in the AP algorithm, there are multiple loops that automatically adjust the planning configurations, which includes not only changing the objectives but also adding tuning structures, beams, or arcs.

Since the AP module has been integrated into the Pinnacle system [112], several comparison studies were conducted to compare the AP algorithm and a standard planning algorithm. For example, Hazell *et al.* [115] re-optimized 26 IMRT head and neck (oropharynx) plans with the AP algorithm. The results showed that all 26 AP plans were clinically acceptable. While the target coverage was similar between AP plans and manual plans, AP plans could better protect the healthy tissues. Gintz *et al.* [114] re-optimized 10 challenging VMAT head and neck plans with the AP algorithm. The results showed that AP plans gave a better OAR sparing (lower doses), but manual plans gave a higher dose homogeneity. The authors concluded that the AP algorithm was a promising clinical tool, but it could benefit from a better process for shifting the balance between the target dose coverage/homogeneity and OAR sparing.

Hansen *et al.* [116] re-optimized 30 VMAT head and neck plans with the AP algorithm. The results showed that 29/30 of AP plans were preferred, as the AP plans gave a higher dose homogeneity for the target and gave lower OAR doses. The AP planning time was half of the manual planning time (64 min). The authors concluded that the target coverage were similar between two methods, but AP plans better spared all OARs. One year later, Hansen *et al.* [117] validated the AP algorithm with 32 VMAT esophageal plans. The results showed that 31/32 of AP plans were preferred over the standard plans. While the two planning methods gave similar target coverage, AP plans resulted in a higher mean spinal cord dose and a lower mean lung dose. The planning time was faster with auto planning (117 min). The authors concluded that auto plans were usually preferred, and gave a better protection for lung.

Vanderstraeten *et al.* [118] re-optimized 56 stereotactic body radiation therapy (SBRT) lung plans with the AP algorithm. The results showed that AP plans gave a significant reduction in OAR dose and a similar target coverage. The authors concluded that the AP plans were not always preferred over manual plans.

1.3.4 iCycle algorithm

iCycle (IMRT Cycle) was mainly developed and studied at the Erasmus Medical Center (Rotterdam, The Netherlands). In traditional inverse planning algorithms, the planning problem was solved through a weighted sum method where the objective function of each organ was first set up and aggregated into a single objective function which was then solved with an optimization algorithm. iCycle proposed to solve the problem with goal programming (an extension of linear programming to handle multiple conflicting objectives) by optimizing one objective (by ordered prioritization) at a time while keeping the other objectives as constraints [119].

Breedveld *et al.* [119] developed a prototype of iCycle that optimizes fluence profiles under a fixed beam arrangement. Breedveld *et al.* [120] improved iCycle by integrating a wish-list of beam orientations into the fluence profile optimization algorithm. In this way, iCycle is able to generate coplanar and non-coplanar IMRT plans.

Several studies reported the comparisons between iCycle and a standard planning algorithm. For example, Voet *et al.* [121] compared iCycle with a standard planning algorithm (Monaco, Elekta AB, Stockholm, Sweden) for 20 IMRT head and neck cases. Plans created with these two methods were blindly selected by a physician. Results showed that 32/33 selected plans were iCycle plans, for which, the OARs were in general better protected, and target coverage was improved in 15 out of 20 cases. The authors concluded that iCycle was able to improve plan quality and reduce workload.

Sharfo *et al.* [122] validated iCycle with a standard planning algorithm (Monaco, Elekta AB, Stockholm, Sweden) for 10 cervical cases. For each case, three IMRT plans (9, 12, and 20 beams) and two VMAT plans (single and dual arc) were created with iCycle. Results showed that all iCycle plans were clinically acceptable with a high and similar PTV coverage, and that OAR sparing increased when going from 9 to 12 to 20 IMRT beams, and from single to dual arc VMAT. The authors concluded that the plan quality of a 20 beam IMRT was superior to VMAT (single and dual arc) for all tested cervical cases.

1.4 Operator-free inverse planning algorithms in brachytherapy

In brachytherapy, several studies have been published to simplify the treatment procedure. For example, Lahanas *et al.* [67] proposed a multi-objective optimization algorithm called nondominated sorting genetic algorithm II (NSGA-II). The authors calculated 1000-2000 solutions in 2-5 min to represent the trade-off surfaces. A decision making tool based on the COIN results can be used to select the final treatment plan from the computed plans.

Nicolae *et al.* [123] applied a machine learning algorithm to generate a high quality treatment plan in LDR prostate brachytherapy. The authors established a framework to match 100 high quality LDR treatment plans based on the similarity of geometric features. The authors concluded that this algorithm was expected to improve LDR treatment plan uniformity while reducing planning time.

Shen *et al.* [124] used deep reinforcement learning to mimic manual planning in HDR cervical brachytherapy. This approach is named as weight-tuning policy network (WTPN). The plan quality was evaluated based on a reward function (related to clinical objectives) for 5 cases. The results showed that the plan quality was improved by 10.7 % with WTPN compared to the plan generated by human planners. For each patient case, it took 25 steps of weight adjustment, and the total time was 4-5 min. The author concluded that it was feasible to develop an intelligent treatment planning approach via deep reinforcement learning.

1.5 Conclusion

After reviewing the mathematical objective functions, optimization algorithms, and the operator-free inverse planning algorithms that aim to mitigate current workload, it is necessary to justify the approaches used in this study.

From all the reviewed operator-free inverse planning algorithms, the multi-criteria optimization (MCO) algorithm (Section 1.3.2) was chosen throughout this thesis. Traditionally, clinical plans are obtained by experienced physicists, through searching the solution space around a population-based planning template. We hypothesize that MCO algorithms can mimic manual planning, by (1) Automatically generating a high quality plan dataset around a population-based planning template, and (2) Selecting the RTOG acceptable plans from the plan dataset. As such, this thesis is working toward implementing a robust and ultra-fast MCO algorithm.

The mathematical objective functions used in the subsequent Chapters (2-4), are dose-volume based (Section 1.1.1.1), similar to IPSA. These functions are preferred because they are relevant to the dosimetric parameters specified in the Radiation Therapy Oncology Group (RTOG) 0924 protocol [18] and are intuitive to physicians and physicists.

The optimization algorithm used in Chapters 2 and 3 is simulated annealing (SA) (Section 1.2.1.1). There are two reasons for choosing SA: (1) The algorithm is known to the brachytherapy community and is already part of a commercial inverse planning system, (2) We believe that applying the same optimization algorithm (SA) can smooth the transition from a traditional inverse planning algorithm (IPSA) to an operator-free inverse planning algorithm (MCO).

SA was found to be a major obstacle in increasing the planning efficiency, and as such, a deterministic optimization algorithm (L-BFGS in Section 1.2.2.1) was implemented to achieve a high computing efficiency in Chapter 4. Like SA, the popularity of L-BFGS in clinical brachytherapy is the first reason to consider this algorithm. Admittedly, the HDR brachytherapy optimization problem is relatively small-scale for one plan (a few hundreds of dwell positions), the computational burden proposed by our approach is large scale for thousands of plans. For this reason, a limited-memory quasi-Newton optimizer such as L-BFGS is well suited to reduce the computational burden associated with the Hessian matrix evaluations.

Chapitre 2

A multi-criteria optimization approach for HDR prostate brachytherapy : I. Pareto surface approximation

Songye Cui[1,2], Philippe Després[1,2] et Luc Beaulieu[1,2]

1. Department of Physics, Engineering Physics and Optics and Cancer Research Center, Université Laval, Quebec City, QC, G1V 0A6, Canada
2. Department of Radiation Oncology and Research Center of CHU de Québec - Université Laval, Quebec City, QC, G1R 2J6, Canada

2.1 Résumé

La planification de la curiethérapie à haut débit de dose (High Dose Rate, HDR) implique généralement un processus itératif d'affinement des objectifs de planification jusqu'à ce qu'un plan cliniquement acceptable soit élaboré. Le but de cette étude en deux parties est d'améliorer les pratiques de planification actuelles en concevant un nouvel algorithme de planification inverse basé sur l'optimisation multicritères (Multi-Criteria Optimization, MCO). Dans la première partie, les surfaces complètes de Pareto ont été approximées et étudiées pour les cas de prostate.

Un algorithme d'approximation de la surface de Pareto a été implémenté dans le cadre du recuit simulé de planification inverse (Inverse Planning Simulated Annealing, IPSA). Les surfaces de Pareto de 140 cas de prostate ont été approximées avec l'algorithme MCO proposé. Pour chaque cas, la surface de Pareto était représentée en générant automatiquement 300 plans optimaux de Pareto, et la région cliniquement acceptable était identifiée. Ainsi, les plans optimaux de 42 000 Pareto ont été créés pour caractériser les surfaces de Pareto pour tous les cas. En outre, la relation entre la région cliniquement acceptable et 4 plans d'ancrage a été étudiée. En conséquence, un ensemble de modèles de régression polynomiale a été extrait afin de prédire rapidement la région cliniquement acceptable sur la surface de Pareto sur la base de plans d'ancrage.

Les surfaces de Pareto pour les cas de prostate de curiethérapie HDR ont été bien caractérisées dans cette étude. Les modèles de régression proposés peuvent aider à définir l'espace de phase de solution le plus pertinent.

2.2 Abstract

High dose rate (HDR) brachytherapy planning usually involves an iterative process of refining planning objectives until a clinically acceptable plan is produced. The purpose of this two-part study is to improve current planning practice by designing a novel inverse planning algorithm based on multi-criteria optimization (MCO). In the first part, complete Pareto surfaces were approximated and studied for prostate cases.

A Pareto surface approximation algorithm was implemented within the framework of Inverse Planning Simulated Annealing (IPSA). The Pareto surfaces of 140 prostate cases were approximated with the proposed MCO algorithm. For each case, the Pareto surface was represented by automatically generating 300 Pareto optimal plans, and the clinically acceptable region was identified. Thus, 42 000 Pareto optimal plans were created to characterize Pareto surfaces for all the cases. In addition, the relationship between the clinically acceptable region and 4 anchor plans was studied. As a result, a set of polynomial regression models was extracted to rapidly predict the clinically acceptable region on the Pareto surface based on anchor plans.

Pareto surfaces for HDR brachytherapy prostate cases were well characterized in this study. The proposed regression models may help define the most relevant solution phase space.

keywords: brachytherapy, treatment planning, prostate cancer, optimization, Pareto surface

2.3 Introduction

High dose rate (HDR) brachytherapy is an effective radiotherapy treatment modality. It consists in temporarily implanting a small radioactive source directly into or near the tumor [8]. Several studies have demonstrated an excellent therapeutic outcome for HDR brachytherapy in combination with external beam radiation therapy (EBRT), in particular for intermediate and high risk localized prostate cancer [125–127] and for gynecologic cancer [128].

Treatment planning in radiation therapy directly affects patient treatment quality [129]. The optimized quantities in HDR brachytherapy is a set of dwell times and dwell positions necessary to achieve good therapeutic results, i.e. high enough dose inside the target to kill cancer cells and low enough dose in surrounding organs at risk (OARs) to keep organ function [9]. Most modern treatment techniques involve some forms of inverse planning with often conflicting objectives related to both tumor coverage and OAR sparing. Currently available inverse planning approaches usually solve the multi-objective problem using a weighted sum method [130]. With this method, all individual objective functions are aggregated into a single objective function. Since the appropriate weights are patient specific [9], the treatment planner has to make several attempts to seek the optimal set of weights for the specific case. This is usually done by tuning an empirically found class solution containing the recommended objectives (target doses and constraints) and their weights [55]. This is time consuming and the result is dependent on the treatment planner's experience [60, 129, 131].

A treatment plan is Pareto optimal if no plan resulting in an improvement of all individual objective functions exists, and the set of Pareto optimal solutions is called the Pareto surface. In order to better understand the trade-off in the multi-objective problem, it is necessary to explore the Pareto surface. It can be solved by enumerating a set of weights, but there is no guarantee of an evenly distributed representation of the Pareto surface. Several studies have proposed effective Pareto surface approximation algorithms, such as the inner approximation [132], the outer approximation [133] and the sandwich approximation [102, 134, 135]. In the inner approximation, a convex hull of current Pareto optimal solutions is first computed, and a Pareto surface is approximated by iteratively using inside lower hulls as bounds of the Pareto surface. In the outer approximation, a Pareto surface is approximated by iteratively finding supporting hyperplanes of the Pareto surface from outside [136]. Sandwich algorithm is a combination of these two algorithms: it iteratively approaches a Pareto surface by minimizing the distance between lower hulls found with inner approximation and supporting hyperplanes found with outer approximation [136]. Since large distance reflects the parts of the Pareto surface that need more accurate approximation, the property of directly minimizing the distance is often considered as a major advantage of Sandwich algorithm [135, 136].

While effective Pareto approximation algorithms have been applied in EBRT [102, 105, 137, 138], to the best of our knowledge, only enumerative algorithms including both randomly and uniformly distributed weights have been applied in brachytherapy [56, 57].

In this study, the concept of multi-criteria optimization (MCO) [102] is introduced within the framework of the well-known Inverse Planning Simulated Annealing (IPSA) algorithm [55]. First, an effective Pareto surface approximation algorithm based on a sandwich algorithm is illustrated with an example. The algorithm is used to characterize Pareto surfaces and to identify clinically acceptable solution space for 140 prostate patients. In the end, a set of regression models is proposed to rapidly predict clinically acceptable region on Pareto surfaces for new cases.

2.4 Materials and Methods

This section is divided into three parts: Pre-, Peri- and Post-Pareto surface approximations. In Pre-Pareto surface approximation, the design of experiments and the formulation of objective functions are described. In Peri-Pareto surface approximation, the implementation of a sandwich scheme Pareto surface approximation is detailed. In Post-Pareto surface approximation, the distribution of clinically acceptable solutions on the Pareto surface is analysed.

2.4.1 Pre-Pareto surface approximation

2.4.1.1 Design of experiments

One hundred and forty prostate cases that have received HDR brachytherapy as a boost to EBRT were randomly selected from a research anonymized database at our institution. These cases are a subset of data previously published [139, 140]. Sixteen to eighteen catheters were inserted into the prostate using a template guide (Elekta Brachy, Veenendaal, The Netherlands) under transrectal ultrasound guidance. The target (the whole prostate) and OAR structures (the bladder, the rectum and the urethra) were contoured on a CT dataset with a slice thickness of 2 mm obtained with a Brilliance Big Bore CT (Philips, Amsterdam, The Netherlands). The mean prostate volume was $60.5\,\mathrm{cm}^3$ (SD: $18.3\,\mathrm{cm}^3$). Each patient received a 15 Gy one fraction HDR brachytherapy treatment delivered using Ir-192 from a Flexitron afterloader (Elekta Brachy, Veenendaal, The Netherlands).

The computations of the proposed algorithm were performed on a 4-core CPU (Intel Core i7-930, 2.8 GHz, 8M Cache).

2.4.1.2 Traditional planning with IPSA

In our current clinical planning procedure with IPSA, a single joint objective function (Equation (2.1)) is minimized by a fast simulated annealing optimizer [55].

$$f(x) = \sum_{i=1}^{N} \frac{1}{M_i} \sum_{j=1}^{M_i} f_i(x_j) \tag{2.1}$$

where N is the number of organs, M_i is the number of dose calculation points of the ith organ, f_i is an individual objective function (a piecewise linear function, Equation (2.2)) of the ith organ, and x_j is the dose value at the jth dose calculation point. The individual objective functions are defined as follows:

$$f_i(x_j) = \begin{cases} w_{min}(D_{min} - x_j) & x_j < D_{min} \\ 0 & D_{min} \leq x_j \leq D_{max} \\ w_{max}(x_j - D_{max}) & x_j > D_{max} \end{cases} \tag{2.2}$$

where D_{min} is the minimum dose (or underdosage) limit of the ith organ and w_{min} is the weight of minimum dose limit of ith organ. Similarly, D_{max} is the maximum dose (or overdosage) limit of the ith organ and w_{max} is the weight of maximum dose limit of the ith organ. The population-based class solution used as an initial objective function in our center is listed in Table 2.1. Dose calculation points used in the objective function are uniformly generated either on the contours (surface dose calculation points) or inside the volume (volume dose calculation points) [55].

Table 2.1 – The IPSA class solution for 15 Gy prostate boost HDR treatment (Surface: surface dose calculation points. Volume: volume dose calculation points).

Organ	Surface				Volume			
	w_{min}	D_{min}(Gy)	D_{max}(Gy)	w_{max}	w_{min}	D_{min}(Gy)	D_{max}(Gy)	w_{max}
Target	200	15	22.5	80	200	15	22.5	10
Urethra	30	14	16	100	30	14	16	100
Bladder	0	0	7.5	90	-	-	-	-
Rectum	0	0	7.5	30	0	0	7.5	30

2.4.1.3 The new joint objective function

Hidden weight was introduced into the original IPSA objective function (Equations (2.1) and (2.2)). Hidden weight is an organ specific scaling of the IPSA weights (in Table 2.1), conveying the user's preferences between treating the prostate and sparing the OARs. These weights are hidden, in contrast to the IPSA weights that were exposed to the user for modifications. The new joint objective function based on hidden weights is given by:

$$f = w_{TU}(f_T + f_U) + w_B f_B + w_R f_R, \tag{2.3}$$

$$w_{TU} + w_B + w_R = 1 \tag{2.4}$$

Equation (2.3) represents how the total objective function (f) is calculated. In it, volumes are denoted by their initials (T for target, U for urethra, B for bladder and R for rectum). Hidden weights w_{TU}, w_B, w_R are always non-negative, and their summation is 1. Individual objective functions f_T, f_U, f_B, f_R are the same objective functions as those of the original IPSA algorithm (Equations (2.1), (2.2) with data from Table 2.1). Based on the dose calculation points of each organ, four individual objective functions are calculated, and are aggregated into a single function after multiplication with hidden weights.

Two assumptions were made to formulate the new joint objective function (Equation (2.3)). First, it assumes that the current class solution used in the clinic is a good starting point, and clinically acceptable trade-offs will be found around that class solution by fine-tuning the hidden weights values. Furthermore, the clinical class solution used for prostate and urethra always results in a well-protected urethra. Over the 140 training cases, the mean V_{125} and D_{10} results are 0.08% and 105.7%, respectively (both well below the RTOG recommended values described in Section 2.4.2.2). As a result, there is no motivation to leave the urethra weight as a free parameter. Instead, it assumes that it is sufficient to let it vary from its class solution value by the same magnitude as the target, with the added benefit of a reduction of the search space (4 to 3 dimensions).

2.4.1.4 RTOG valid solutions

As recommended by the Radiation Therapy Oncology Group (RTOG) 0924 protocol [18], RTOG valid solutions or clinically acceptable plans should meet the following requirements:

- Target V_{100} of at least 90%,
  - In practice, a value above 90% (95%) is desirable because it is usually attainable in the clinic without sacrificing the OAR protection.
- Bladder and rectum $V_{75} < 1\,\mathrm{cm}^3$.
- Urethra $D_{10} < 118\%$.

In the above items, V_x stands for the volume receiving x% of the prescription dose and D_x stands for the dose covering x% of the volume.

2.4.1.5 Determination of the number of iterations to reach the Pareto surface

Simulated annealing is a stochastic optimization approach, and is able to escape local minima solutions. The selection of the number of iterations is a key element to reach the Pareto surface. Here, the effect of the number of iterations on Pareto surface approximation was evaluated.

The Pareto surface of a bi-objective function (objective "subfunctions" for the target and the bladder) was visualized and compared to illustrate the effect of the number of IPSA iterations (200 000, 100 000, 50 000, 40 000, and 30 000 iterations per plan). The formulation of the global objective function was based on Equation (2.3). In addition, the convexity of Pareto surfaces of all 140 patients were verified.

2.4.2 Peri-Pareto surface approximation

In order to effectively approximate Pareto surfaces, a sandwich algorithm [102] has been implemented in a research version of IPSA [55]. This IPSA is similar to the one implemented in Oncentra Brachy, and allows us to use the algorithm as we see fit and integrate it within other pieces of code. A treatment plan is mathematically an optimal solution (although SA does not necessarily converge to a global optimum), or geometrically a vertex. A convex hull algorithm (qhull) [141] was used to find convex hulls of current vertices. Both upper hulls and lower hulls are found, and they are geometrically referred as facets.

The mechanism of the Pareto surface approximation can be divided into four steps: (1) anchor plans computation, (2) current plan convex hull computation, (3) lower bounds and intersection points computation, and (4) next plan computation.

Each step begins with a general description of how the mechanism works for clinical cases (three objectives), and ends with an illustration of a bi-objective problem (visualization-friendly, but not a clinical case) in Figure 2.1.

2.4.2.1 Computation of anchor plans

Usually, an anchor plan is defined as a Pareto optimal solution obtained by minimizing one single objective [135]. In this study, the definition of anchor plan was extended slightly adding the class solution as an anchor plan (can be seen as equally minimizing all organs) and an anchor was not defined for a plan minimizing the urethral dose, again since it is inside the target. The definition of anchor plans was defined as follows:

Anchor plan CS:

class solution or equally minimizing all organs (a planning template).
$(w_{TU}, w_B, w_R) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$

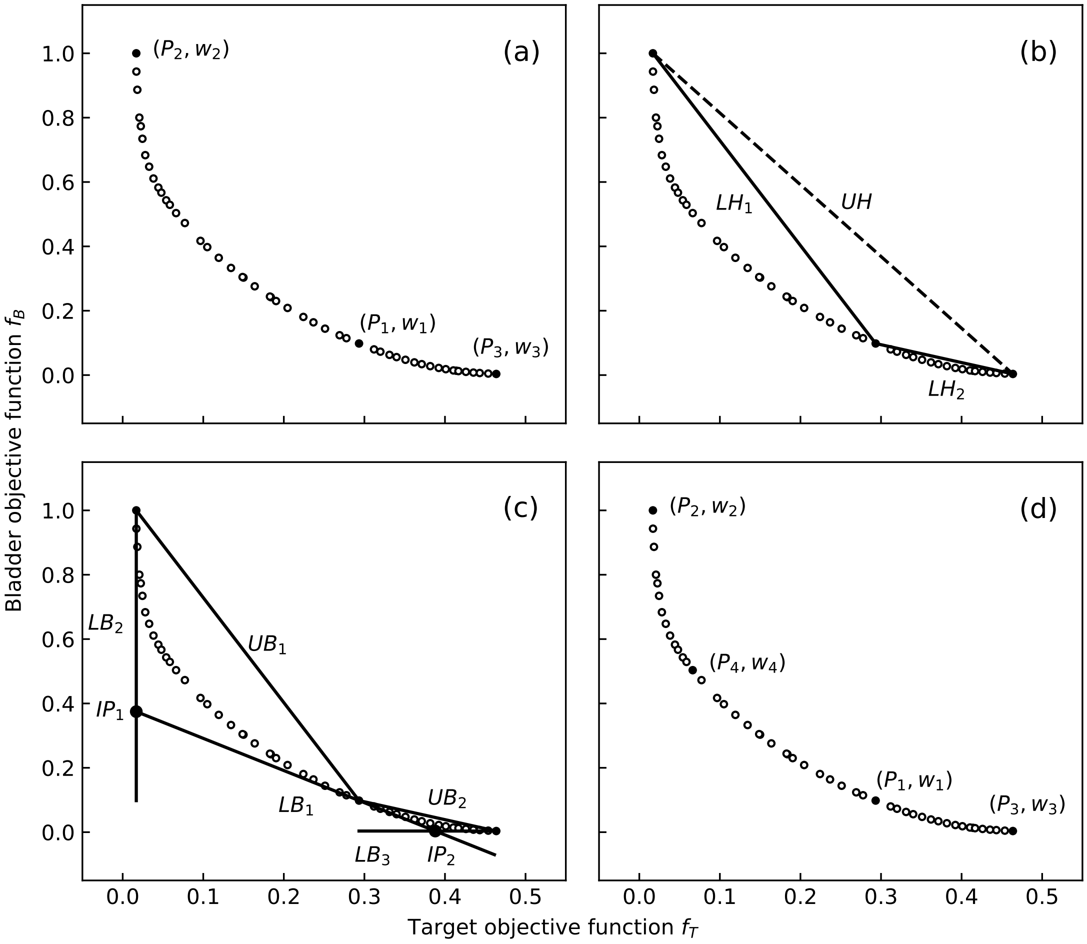


FIGURE 2.1 – The mechanism of Pareto surface approximation (expected in open dots). (a) Computation of anchor plans (P_1, P_2, P_3). (b) Convex hull (including lower hulls LH_1, LH_2 in solid line, and rejected upper hull UH_1 in dashed line) of current vertices. (c) Generation of lower bounds (LB_1, LB_2, LB_3 in solid lines) of Pareto surface, and intersection points (IP_1, IP_2 in closed dots) of lower bounds. Note: LH_1, LH_2 in b were renamed as upper bounds UB_1, UB_2. (d) Computation of next plan (P_4) after optimization with $w_4 = 0.5(w_1 + w_2)$.

Anchor plan minT:

only minimizing the target volume.
$(w_{TU}, w_B, w_R) = (1, 0, 0)$

Anchor plan minB:

only minimizing the bladder volume.
$(w_{TU}, w_B, w_R) = (0.1, 0.9, 0)$

Anchor plan minR:

only minimizing the rectum volume.
$(w_{TU}, w_B, w_R) = (0.1, 0, 0.9)$

Note that in the third and fourth scenarios, the value of w_{TU} was not set to zero as the optimization would be trapped at an invalid solution (zero dwell times for all dwell positions) and thus a small value of 0.1 was selected to represent an optimization that almost completely favors the bladder or the rectum.

For a bi-objective problem in Figure 2.1a, three anchor plans (CS: P_1, minT: P_2, minB: P_3) were first computed after optimization with hidden weights $w_1 = (w_T, w_B) = (0.5, 0.5)$, $w_2 = (w_T, w_B) = (1, 0)$, $w_3 = (w_T, w_B) = (0.1, 0.9)$. The Pareto surface was approximated with 50 vertices (open dots in Figure 2.1).

2.4.2.2 Convex hull computation of current plans

A C++ version of the `qhull` algorithm [141] was integrated into IPSA to perform a convex hull calculation. The algorithm reads an input file that contains individual objective function values of current plans, and writes an output file that contains the convex hull information. Both upper hulls and lower hulls obtained were expressed with a normal vector and an offset. Upper hulls were identified if the output hulls have a negative offset, and were rejected in following steps. A hull facet is identified by 3 vertices for a clinical case, or 2 vertices for the bi-objective problem in Figure 2.1 (i.e. The hull LH_1 is identified by P_1 and P_2).

For the bi-objective problem in Figure 2.1b, both upper hulls (dashed line, noted as UH) and lower hulls (solid line, noted as LH_1, LH_2) were created by the `qhull` algorithm. Only lower hulls LH_1, LH_2 will be used in following steps.

2.4.2.3 Computation of lower bounds and intersection points

Each lower hull facet found in Section 2.4.2.2 was used as an upper bound of the Pareto surface, and was paired with the 3 vertices identifying this facet. For each vertex, a supporting hyperplane was specified, and was used as a lower bound of the Pareto surface. The normal vector of the supporting hyperplane is the hidden weight of this vertex. Three lower bounds linked with an upper bound intersected at one point. The distance between the intersection point and the upper bound was calculated, and the upper bound facet with maximum distance was chosen as the next-to-run facet.

For the bi-objective problem in Figure 2.1c, the lower hulls LH_1 and LH_2 in Figure 2.1b were re-named as $UB_1$and UB_2 respectively to show that they served as upper bounds of the Pareto surface. For the upper bound UB_1, vertices P_1 and P_2 were on UB_1, and therefore were associated with it. For vertex P_1, lower bound LB_1 was found, and for vertex P_2, lower bound LB_2 was found. Lower bounds LB_1 and LB_2 intersected at a point IP_1, and the distance between this intersection point IP_1 and the upper bound UB_1 was calculated. Same method was implemented for UB_2. The results are displayed in Figure 2.1c and show that UB_1 was clearly the next-to-run facet because the distance between IP_1 and UB_1 is larger than the distance between IP_2 and UB_2.

2.4.2.4 Computation of the next plan

The average hidden weight of the three associated vertices on the next-to-run facet was used to create the next plan. Steps described in Sections 2.4.2.2, 2.4.2.3, and this section were repeated until a fixed number (an arbitrary value of 300 was used) of plans had been generated.

For the bi-objective problem in Figure 2.1d, P_1 and P_2 are the vertices associated with the next-to-run facet UB_1. The hidden weight of the next plan is $w_4 = 0.5(w_1 + w_2)$, and P_4 is the next plan that is obtained. Steps described in Sections 2.4.2.2, 2.4.2.3, and this section were repeated until generating a total of 50 plans (open dots in Figure 2.1) to represent the Pareto surface.

2.4.3 Post-Pareto surface approximation

2.4.3.1 Relationships between hidden weights and dosimetric parameters

For each case, the Pareto surface was represented with 300 Pareto optimal plans generated using the method described in Section 2.4.2. For each Pareto optimal plan, dose-volume histograms (DVH) and dosimetric parameters were computed for both target and OARs. For the target, the relationship between the hidden weight value w_{TU} and the dosimetric parameter value V_{100} was studied, and the requirement of w_{TU} to find RTOG valid solutions (only in terms of the criteria of the target) was identified. For OARs, the relationship between the hidden weight value (both absolute ones w_B, w_R and the relative ones $w_B/w_{TU}, w_R/w_{TU}$) and the dosimetric parameter value V_{75} was studied, and the requirement of the hidden weight to find RTOG valid solutions (only in terms of the criteria of the related OAR) was identified.

2.4.3.2 Regression models to obtain RTOG valid solutions

Even though it is possible to identify RTOG valid solutions after characterizing the Pareto surface with numerous plans, the costly computation (1.1 h per patient) makes the implementation unrealistic in clinical practice.

The concept of dosimetric indicators were therefore introduced to predict the clinically acceptable solution space based on a few anchor plans. Dosimetric indicators are organ-specific, and indicate the difficulty to reach the RTOG recommended dosimetric criteria from the planning template (class solution). The dosimetric indicator value are determined based on the hidden weights and the dosimetric results of anchor plans (Equations (2.5a) - (2.5c)).

$$di_{TU} = \frac{V_{100,RTOG} - V_{100,CS}}{(V_{100,minT} - V_{100,CS})/(w_{TU,minT} - w_{TU,CS})} \tag{2.5a}$$

$$di_{B} = \frac{D_{1ccB,RTOG} - D_{1ccB,CS}}{(D_{1ccB,minR} - D_{1ccB,CS})/(w_{B,minR} - w_{B,CS})} \tag{2.5b}$$

$$di_{R} = \frac{D_{1ccR,RTOG} - D_{1ccR,CS}}{(D_{1ccR,minB} - D_{1ccR,CS})/(w_{R,minB} - w_{R,CS})} \tag{2.5c}$$

In the definition of target-urethra dosimetric indicators di_{TU} (Equation (2.5a)), for instance, variables $V_{100,RTOG}, V_{100,CS}, V_{100,minT}$ are the minimal target V_{100} values required by RTOG, resulted in anchor plan CS, and resulted in anchor plan minT, respectively. A value of 95% (a common criteria in clinic) was assigned to $V_{100,RTOG}$. Variables $w_{TU,minT}, w_{TU,CS}$ (defined in Section 2.4.2.1) are the target-urethra hidden weights in anchor plan minT and in anchor plan CS, respectively.

The numerator of di_{TU} reflects the difference between $V_{100,RTOG}$ (expectation) and $V_{100,CS}$ (starting point), and the denominator of di_{TU} reflects the change of V_{100} per unit of hidden weight w_{TU}(or $\triangle V_{100}/\triangle w_{TU}$). As a result, the value of di_{TU} indicates the level of difficulty to obtain a 95% of target V_{100} value from the class solution. The same strategy was applied to the OARs in Equation (2.5b-c).

With the method described in Section 2.4.3.1, Pareto surfaces were approximated and the requirement of the hidden weight to find RTOG valid solutions were identified for 140 prostate cases. In addition, the dosimetric indicator values were also determined for those cases.

The relationship between the dosimetric indicator values and the hidden weight values was observed. Simple regression models (Microsoft Excel Version 15.27) including linear and polynomial models were applied to illustrate this relationship, and one appropriate model with the highest R^2 result was chosen.

2.5 Results

2.5.1 Pre-Pareto surface approximation

2.5.1.1 Determination of the number of iterations to reach the Pareto surface

In Figure 2.2, Pareto surfaces of a bi-objective problem were approximated under 5 different number of iterations (200 000, 100 000, 50 000, 40 000, 30 000) were compared. From the result, the Pareto surface can already be well represented with 100 000 iterations because there was not significant improvement with further iterations (200 000 iterations). For an average size prostate clinical case, the optimization time is 2.5 s per 10 000 iterations. Therefore, in order to be clinically pragmatic, 50 000 iterations per solution was chosen to approximate the Pareto surface in this study.

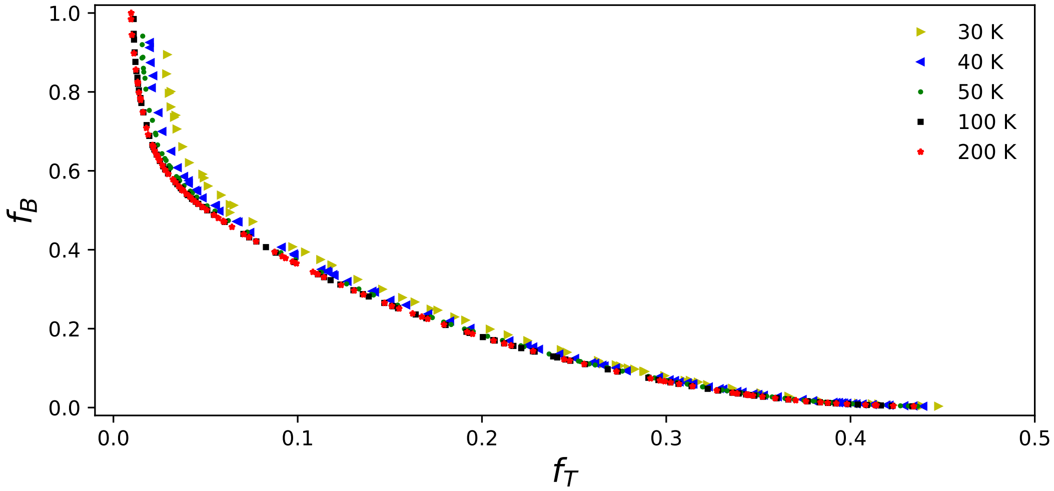


FIGURE 2.2 – Comparison of Pareto surface approximations under 5 scenarios (200 000, 100 000, 50 000, 40 000, 30 000 iterations).

Pareto surfaces of 140 patients were approximated under 50 000 iterations of IPSA. On average, about 98 % solutions were non-dominated (no solution that improves all individual objective function values exists). Therefore it seems safe to assume that our method is capable of sufficiently approximating the Pareto surface.

2.5.2 Peri-Pareto surface approximation

2.5.2.1 Distributions of objective functions and hidden weights

The Pareto surface of one clinical case was characterized with 300 Pareto optimal plans. The distribution of hidden weights leading to such a Pareto surface characterization is shown in Figure 2.3a, and the distribution of resulting dosimetric parameter results is shown in Figure 2.3b.

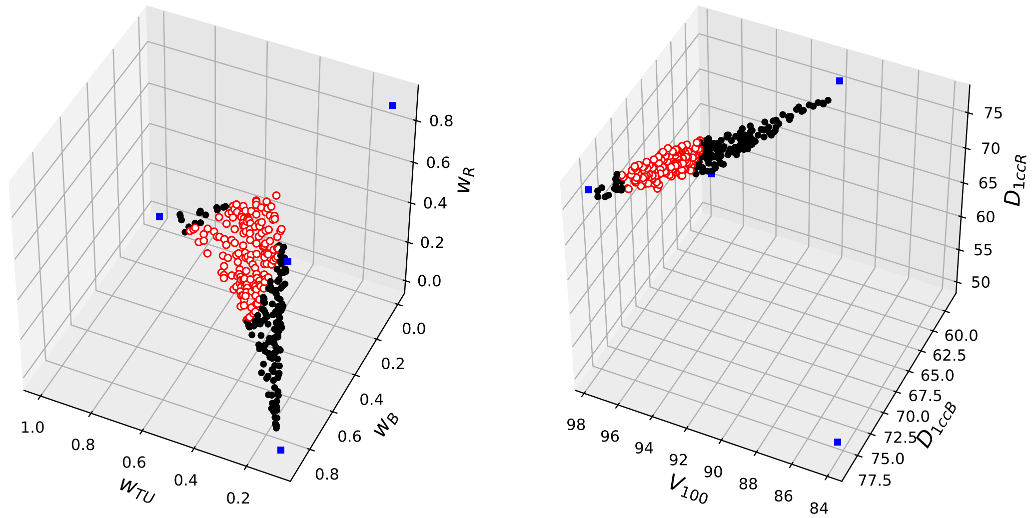


FIGURE 2.3 – Pareto surface approximation of one case (anchor plan in squares, regular optimal plans in solid dots, clinical acceptable plans in open dots) (a) Distribution of hidden weights (b) Distribution of dosimetric parameters.

2.5.3 Post-Pareto surface approximation

2.5.3.1 Relationships between hidden weights and dosimetric parameters

The relationships between the hidden weights and the dosimetric parameters for the same case in Section 2.5.2.1 are shown in Figure 2.4. From the result in Figure 2.4a, there is a positive correlation between the absolute hidden weight value and the dosimetric parameter value for the target. However, as shown in Figure 2.4b-c, there are negative correlations between the relative hidden weight value and the dosimetric parameter value for the bladder or the rectum (closed dots in Figure 2.4). Similar behaviors were not found between the absolute hidden weight value and the dosimetric parameter value for the bladder or the rectum (open squares in Figure 2.4).

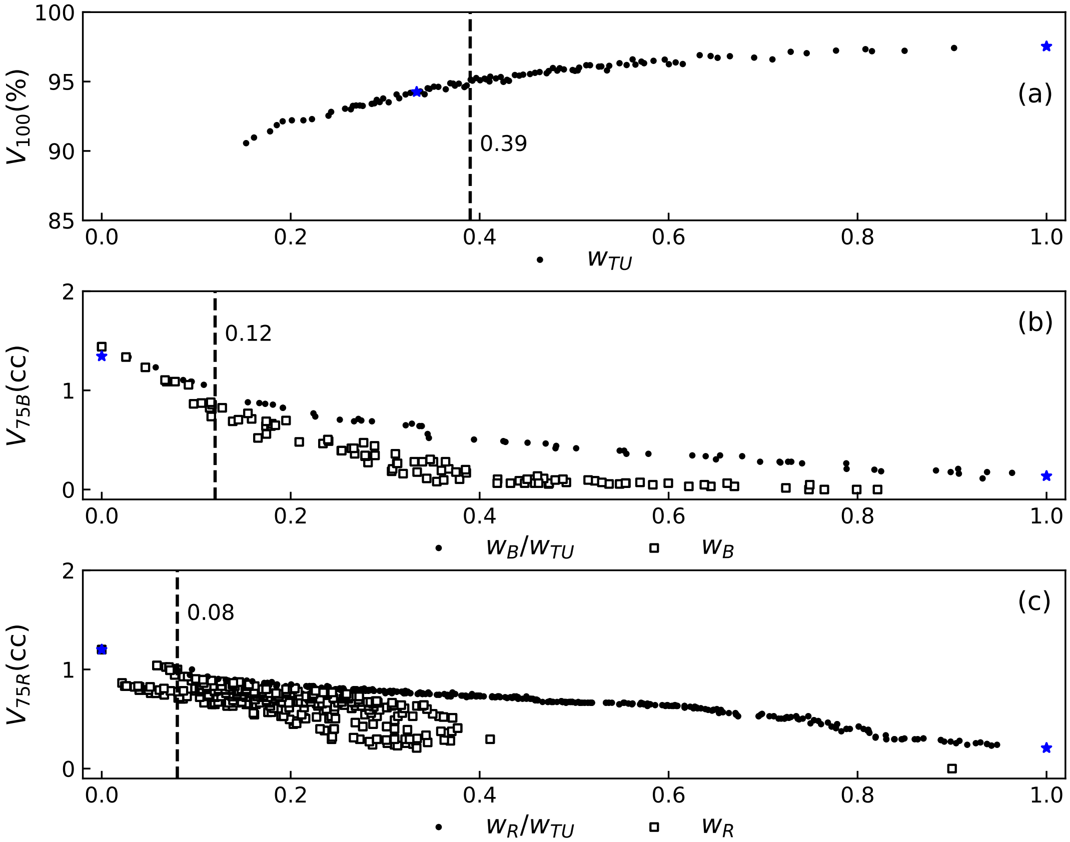


FIGURE 2.4 – Relationships between hidden weights and dosimetric parameters (the requirements of hidden weights are in dashed lines, the dosimetric results of anchor plans are in closed stars). (a) The relationship between w_{TU} and V_{100} for the target. (b) The relationship between w_B/w_{TU} (closed dots), w_B (open squares) and V_{75} for the bladder. (c) The relationship between w_R/w_{TU} (closed dots), w_R (open squares) and V_{75} for the rectum.

From the dashed lines in Figure 2.4, the requirements to obtain RTOG valid solutions can be identified as: $w_{TU} \geq 0.39$, $w_B/w_{TU} \geq 0.12$, $w_R/w_{TU} \geq 0.08$. Since the criteria of the urethra can be met for all the cases studied, its result was ignored in Figure 2.4.

2.5.3.2 Regression models to obtain RTOG valid solutions

A set of polynomial models to predict the requirement of hidden weights to obtain RTOG valid solutions for new cases is illustrated in Figure 2.5. Dosimetric indicators were calculated with Equation (2.5) simply based on anchor plans (stars in Figure 2.4).

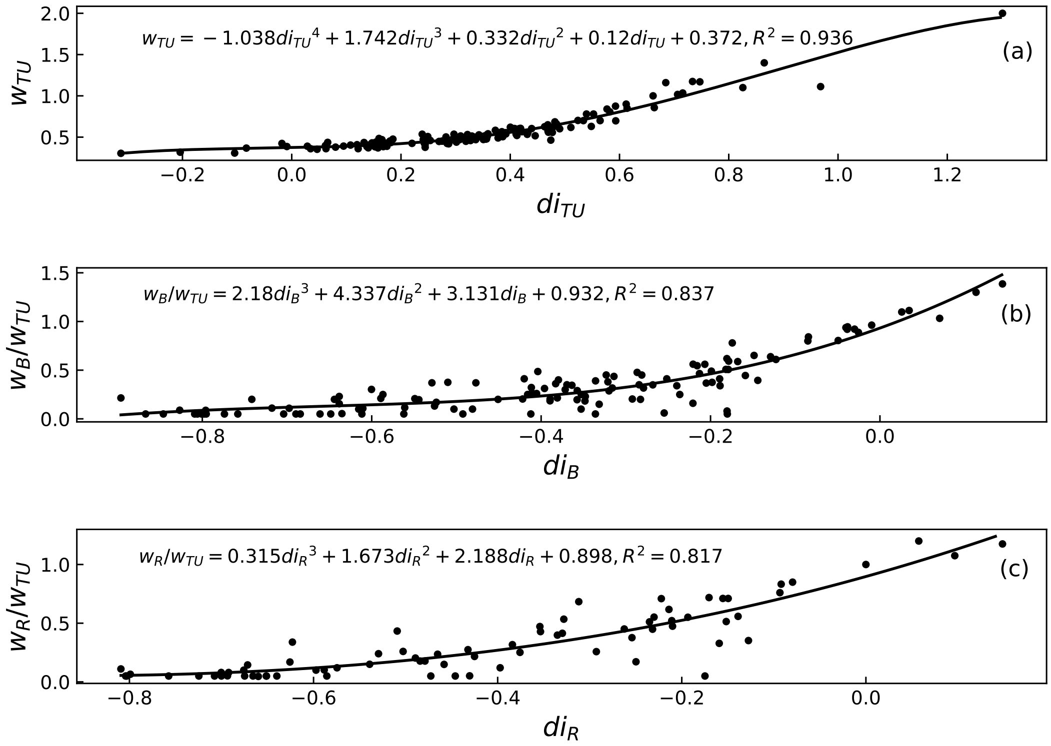


FIGURE 2.5 – Regression models ((a) prostate, (b) bladder, (c) rectum) to have RTOG valid solutions based on dosimetric indicators (In each subfigure, one dot refers to the dosimetric indicator value and the hidden weight value of one patient, and the line is the fit based on all dots).

2.5.3.3 Performance

Pareto surfaces for 140 prostate cases were characterized. For each case, 300 Pareto optimal plans were created. The average time to obtain one plan is 13.3 s (50 000 iterations optimization + DVH computations).

Therefore, in order to characterize Pareto surface for 140 cases, a total number of 42 000 plans were generated, for a total time of approximately 155 h (140 × 300 × 13.3 s).

2.6 Discussion

The first part of the study serves as a preprocessing stage of developing an operator-free inverse planning algorithm for HDR brachytherapy.

In EBRT, DVH based objective functions usually lead to a non-convex solution space [142]. In our case, convex Pareto surfaces were observed. A sandwich algorithm was implemented within the framework of IPSA to iteratively populate the Pareto surface. This algorithm approaches both upper bounds and lower bounds of the surface, and has the capacity of steering the new points towards the part of the Pareto surface with a high approximation error [135, 136].

The objective function of commercial brachytherapy inverse planning algorithms, including IPSA, is usually based on limited dose calculation points, rather than all possible dose calculation points (i.e. DVH or dosimetric parameters). For example, only the bladder surface dose calculation points near the target (or dwell positions) are used in optimization in IPSA. However, such a compromise of efficiency [143] does not necessarily lead to a disadvantage, since a direct/indirect relationship between hidden weights and dosimetric parameters can already be achieved with IPSA (Figure 2.4).

The characterization of Pareto surfaces is a compute-intensive process (1.1 h per patient in this study), and a Pareto surface approximation algorithm may thus not be compatible to a real-time inverse planning system. However, the resulting regression models (Figure 2.5) enable an extremely fast navigation for any new prostate HDR brachytherapy cases.

In a companion paper, the regression models obtained in this study were implemented within IPSA to automatically generate a plan. This plan was benchmarked against a physician-approved plan in terms of dosimetric parameters and performance.

2.7 Conclusion

A sandwich algorithm was implemented to automatically populate Pareto surfaces in the HDR brachytherapy planning process. The algorithm was tested with 140 prostate cases. The algorithm could help treatment planners and researchers better understand the trade-off on Pareto surfaces, and the results may help in obtaining a clinically acceptable plan without manual tuning of the class solution on a case-by-case basis.

2.8 Acknowledgements

This work was supported in part by the National Sciences and Engineering Research Council of Canada (NSERC) via the NSERC-Elekta Industrial Research Chair (Grant number: 484144-15). The authors acknowledge a scholarship from the Chinese Scholarship Council, and acknowledge partial support by the CREATE Medical Physics Research Training Network grant of the Natural Sciences and Engineering Research Council (Grant number: 432290). The authors acknowledge the supports from their colleagues, especially Frédéric Lacroix, Éric Poulin, Nicolas Varfalvy, Marie-Claude Lavallée for answering the questions and providing the clinical class solution, Yunzhi Ma for the brainstorming discussions, Paul Edimo for the data collection, and Ghyslain Leclerc for the careful grammatical review.

Chapitre 3

A multi-criteria optimization approach for HDR prostate brachytherapy : II. benchmark against clinical plans

Songye Cui[1,2], Philippe Després[1,2] et Luc Beaulieu[1,2]

1. Department of Physics, Engineering Physics and Optics and Cancer Research Center, Université Laval, Quebec City, QC, G1V 0A6, Canada
2. Department of Radiation Oncology and Research Center of CHU de Québec - Université Laval, Quebec City, QC, G1R 2J6, Canada

3.1 Résumé

L'approche itérative actuelle de la planification inverse de la planification du traitement à haut débit de dose peut prendre beaucoup de temps. Cette étude en deux parties vise à rationaliser le processus de planification tout en maintenant la qualité du plan. Dans cette seconde partie, un algorithme de planification d'optimisation multi-critères (Multi-Criteria Optimization, MCO) est proposé et comparé à un algorithme de planification standard.

Avec un ensemble de modèles de régression préalablement établis, un espace de solution valide spécifique au patient sur la surface de Pareto a été prédit en fonction des résultats du plan d'ancrage. Des plans alternatifs générés parallèlement au front partiel de Pareto ont été présentés au planificateur, et un plan a été choisi comme plan de MCO. Les résultats des paramètres dosimétriques ainsi que le temps de planification ont été comparés entre les plans de MCO et les plans standard approuvés par le médecin pour 236 cas de prostate.

Les résultats montrent que la planification de l'MCO protège mieux l'urètre que la planification standard (valeur urétrale moyenne inférieure D_{10} de 2.25%). La qualité globale du plan MCO est également supérieure à la qualité du plan standard, car la planification MCO est capable d'augmenter la fréquence des plans cliniquement acceptables répondant à tous les critères RTOG simultanément sans intervention humaine (de 83.05% à 97.46%). Enfin, le temps de planification moyen d'une MCO est 41 s sans aucune intervention des planificateurs de traitement.

L'algorithme de planification MCO présenté constitue un moyen robuste et automatisé d'améliorer la qualité du traitement en curiethérapie.

3.2 Abstract

The current iterative approach to inverse planning of high dose rate (HDR) treatment planning can be time consuming. The purpose of this two-part study is to streamline the planning process while maintaining plan quality. In this second part, a multi-criteria optimization (MCO) planning algorithm is proposed and benchmarked against a standard planning algorithm.

With a set of previously established regression models, a patient-specific valid solution space on the Pareto surface was predicted based on the anchor plans results. Alternative plans generated alongside the partial Pareto front were presented to the planner, and one plan was selected as the MCO plan. The dosimetric parameters results as well as the planning time were compared between the MCO plans and the physician-approved standard plans for 236 prostate cases.

Results show that the urethra is better spared with MCO planning than with standard planning (a lower mean urethral D_{10} value of 2.25%). The overall MCO plan quality also outperforms the standard plan quality, since MCO planning is able to increase the frequency of clinically acceptable plans meeting all of RTOG criteria simultaneously without any human intervention (from 83.05% to 97.46%). Finally, the average MCO planning time is 41 s without any interventions of treatment planners.

The presented MCO planning algorithm constitutes a robust and automated way to improve treatment quality in brachytherapy.

keywords: brachytherapy, prostate cancer, patient-specific, treatment planning, optimization

3.3 Introduction

Radiotherapy, along with surgery and chemotherapy are three major methods of cancer treatment [144]. Radiotherapy can be categorized into external-beam radiotherapy (EBRT) and internal radiotherapy (brachytherapy). With EBRT, tumor cells are killed by radiation beams generated from a linear accelerator, and with brachytherapy, small radioactive sources are directly implanted into or near a tumor [8].

High dose rate (HDR) brachytherapy consists of a temporary implant of a radioactive source with $\geq 12\,\mathrm{Gy\,h^{-1}}$ dose rate, and it can either serve as a monotherapy or a boost to EBRT [145]. During treatment planning, treatment planners need to determine a set of dwell times and dwell positions that results in a dose distribution, which is high enough to kill cancerous cells but low enough to spare organs at risk (OARs) [9].

Commercially available inverse planning algorithms used in HDR brachytherapy such as Inverse Planning Simulated Annealing (IPSA) [55] and Hybrid Inverse Planning Optimization (HIPO) [57] are able to help treatment planners optimize the dose distribution. However, the underlying objective function of these algorithms is population based and thus not tailored specifically to each patient.

Several patient-specific inverse planning algorithms such as multi-criteria optimization [102,146], knowledge-based planning [147–149] and artificial intelligence based approaches [123,150,151] have been proposed to improve current planning practice in EBRT and low dose rate (LDR) brachytherapy. MCO first populates Pareto surface by constructing a plan dataset and lets treatment planners navigate among these plans [152]. Knowledge-based planning matches one or several physician-approved plans based on the similarity of anatomic geometric features. Artificial intelligence approaches try to solve the problem using a machine learning algorithm. To the best of our knowledge, the only two HDR brachytherapy studies about patient-specific inverse planning algorithms [153,154] were specific to gynecological (GYN) cancers.

MCO developed for external beam therapy [102,146] may not be compatible to the planning in HDR prostate brachytherapy that usually occurs in the operation room because interactively populating the Pareto surface is time consuming.

In this study, a patient-specific inverse planning algorithm called multi-criteria optimization is proposed for HDR brachytherapy. The algorithm was benchmarked in terms of the dosimetric parameters and the planning time against physician-approved plans for 236 prostates cases.

3.4 Materials and Methods

3.4.1 Patient selection

Two hundred and thirty-six prostate cases who received HDR brachytherapy as a boost to EBRT were randomly selected. These cases are different from the cases used in the first part study, but they are both from a larger patient database [139, 140]. For each case, 16-18 catheters were implanted in the prostate under the guidance of a transrectal ultrasound while the patient was under anesthesia, and one fraction of 15 Gy was delivered using Ir-192 from a Flexitron afterloader (Elekta Brachy, Veenendaal, The Netherlands). The structures of the target, the bladder, the rectum and the urethra were contoured on a CT dataset by a physician. The mean prostate volume was 57.4 cm^3 (SD: 20.6 cm^3).

3.4.2 RTOG valid solutions

Plan evaluations may vary across centers as described in the American Brachytherapy Society (ABS) Consensus Guidelines [41]. The Radiation Therapy Oncology Group (RTOG) 0924 protocol [18] recommended that valid solutions or clinically acceptable plans for CT based HDR prostate brachytherapy boost should meet the following requirements:

- Target V_{100} of at least 90%,
    - A value above 90% (95%) is desirable because it is usually attainable in the clinic without sacrificing the OAR protection.
- Bladder and rectum $V_{75} < 1\,\mathrm{cm}^3$.
- Urethral $D_{10} <$ 118% of the prescription dose.

In the above items, V_x stands for the volume receiving x% of the prescription dose and D_x stands for the dose covering x% of the volume.

3.4.3 Standard planning

Standard planning involved creating a treatment plan with the traditional inverse planning algorithm IPSA in Oncentra Prostate (Elekta Brachy, Veenendaal, The Netherlands). The objective functions of the inverse planning algorithm include several linear piecewise functions which are formulated based on Equation (3.1). The minimum dose limit D_{min} and the maximum dose limit D_{max} are the underdosage limit and the overdosage limit, respectively. The weight of minimum dose limit w_{min} and the weight of maximum dose limit w_{max} are also referred as the slope of the piecewise function. The objective function value depends on the dose values

x at dose calculation points. There are two types of dose calculation points: surface dose calculation points (on the contours) and volume dose calculation points (inside the volume).

$$f(x) = \begin{cases} w_{min}(D_{min} - x) & x < D_{min} \\ 0 & D_{min} \leq x \leq D_{max} \\ w_{max}(x - D_{max}) & x > D_{max} \end{cases} \tag{3.1}$$

In this approach, treatment planners first set up the objective function with a planning template called class solution (Table 3.1), which contains recommended dose limits and weights values [55]. After the initial optimization, a plan was created and evaluated based on dosimetric parameters and the isodose curves. Usually, the class solution would be modified and the aforementioned steps would be repeated until a RTOG valid solution is obtained. This trial-and-error procedure of the standard planning is illustrated in Figure 3.1.

Table 3.1 – The IPSA class solution for 15 Gy prostate boost HDR treatment (Surface: surface dose calculation points. Volume: volume dose calculation points).

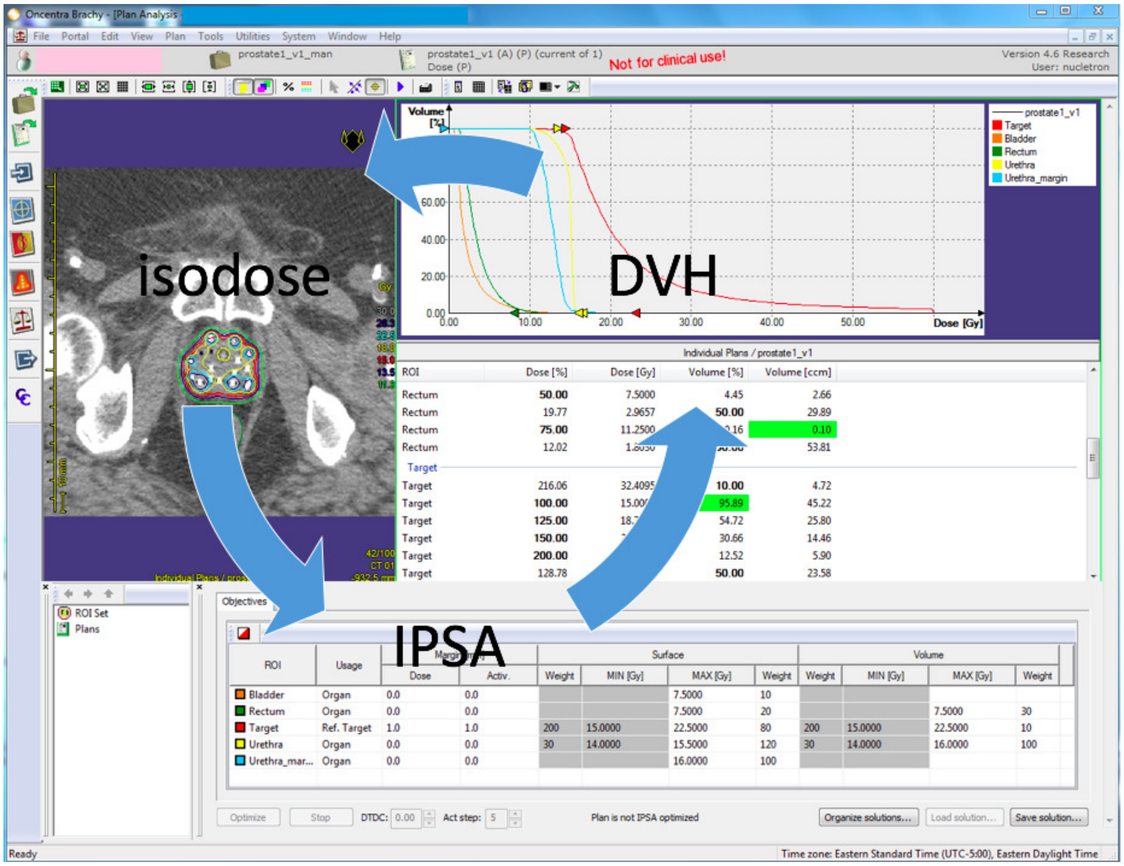
Organ	Surface				Volume			
	w_{min}	D_{min}(Gy)	D_{max}(Gy)	w_{max}	w_{min}	D_{min}(Gy)	D_{max}(Gy)	w_{max}
Target	200	15	22.5	80	200	15	22.5	10
Urethra	30	14	16	100	30	14	16	100
Bladder	0	0	7.5	90	-	-	-	-
Rectum	0	0	7.5	30	0	0	7.5	30

3.4.4 Multi-criteria optimization planning

To improve the efficiency of the standard planning approach, it is desirable to avoid manual modifications of the class solution. To achieve this, the MCO planning algorithm was implemented based on a research version of IPSA [55], which is similar to the one implemented in Oncentra Brachy, and allows us to use the algorithm as we see fit and integrate it within other pieces of code. In Figure 3.2, the formulations of objective functions for both the standard planning and the MCO planning are illustrated.

Here, T refers to Target, U refers to Urethra, B refers to Bladder, and R refers to Rectum. The standard objective function f_S is the summation of all individual objective functions (f_T, f_U, f_B, f_R) which are determined by the dose values x at the dose calculation points (dots in Figure 3.2).

As for the MCO objective function f_{MCO}, the concept of the hidden weights (w_{TU}, w_B, w_R) was introduced. The non-negative hidden weights are organ specific, and are used to represent the user's preferences between treating the target and sparing the OARs. The MCO objective function f_{MCO} is a 'hidden weighted' sum of the individual objective functions.



Individual Plans / prostate1_v1

ROI	Dose [%]	Dose [Gy]	Volume [%]	Volume [ccm]
Rectum	**50.00**	7.5000	4.45	2.66
Rectum	19.77	2.9657	**50.00**	29.89
Rectum	**75.00**	11.2500	[illegible]	0.10
Rectum	12.02	[illegible]	[illegible]	53.81
Target				
Target	216.06	32.4095	**10.00**	4.72
Target	**100.00**	[illegible]	95.89	45.22
Target	**125.00**	[illegible]	54.72	25.80
Target	**150.00**	[illegible]	30.66	14.46
Target	**200.00**	[illegible]	12.52	5.90
Target	128.78	[illegible]	**50.00**	23.58

ROI	Usage	Margin [mm]		Surface				Volume			
		Dose	Activ.	Weight	MIN [Gy]	MAX [Gy]	Weight	Weight	MIN [Gy]	MAX [Gy]	Weight
🟧 Bladder	Organ	0.0	0.0			7.5000	10				
🟩 Rectum	Organ	0.0	0.0			7.5000	20			7.5000	30
🟥 Target	Ref. Target	1.0	1.0	200	15.0000	22.5000	80	200	15.0000	22.5000	10
🟨 Urethra	Organ	0.0	0.0	30	14.0000	15.5000	120	30	14.0000	16.0000	100
🟦 Urethra_mar...	Organ	0.0	0.0			16.0000	100				

DTDC: 0.00 Act step: 5

Plan is not IPSA optimized

Time zone: Eastern Standard Time (UTC-5:00), Eastern Daylight Time

FIGURE 3.1 – The re-planning procedure with Oncentra Prostate where the treatment planner manually adjusts the class solution.

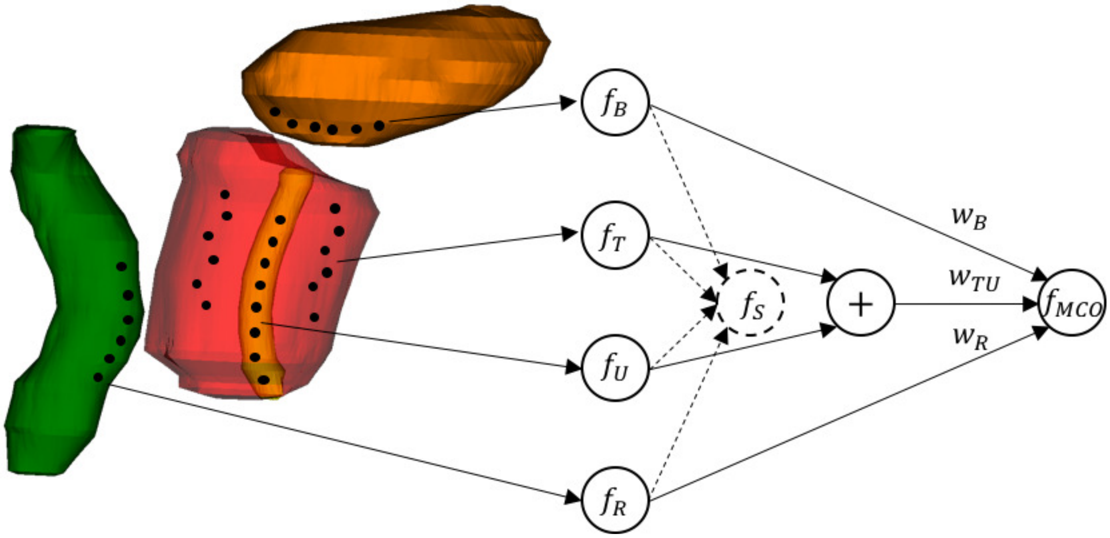


FIGURE 3.2 – The formulations of the standard objective function and the MCO objective function: the standard objective function f_S (dashed lines) is the summation of the individual objective functions, and the MCO objective function f_{MCO} (solid lines) is a 'hidden weighted' sum of the individual objective functions.

Two assumptions were made to construct f_{MCO}. First, it assumes that clinically acceptable trade-offs will be found around the population-based class solution by leveraging hidden weights. Furthermore, the clinical class solution used for prostate and urethra always results in a well-protected urethra. Over the 140 training cases, the mean V_{125} and D_{10} results are 0.08% and 105.7%, respectively (both well below RTOG recommended values). As a result, there is no motivation for this specific protocol to leave the urethra weight as a free parameter. Instead, it assumes that it is sufficient to let it vary from its class solution value by the same magnitude as the target, with the added benefit of a reduction of the search space (4 to 3 dimensions).

The mechanism (Figure 3.3) of the MCO planning can be divided into 4 steps: (1) anchor plan computation, (2) weight adjustment, (3) alternative plan computation and (4) plan selection. The details of each step will be described below.

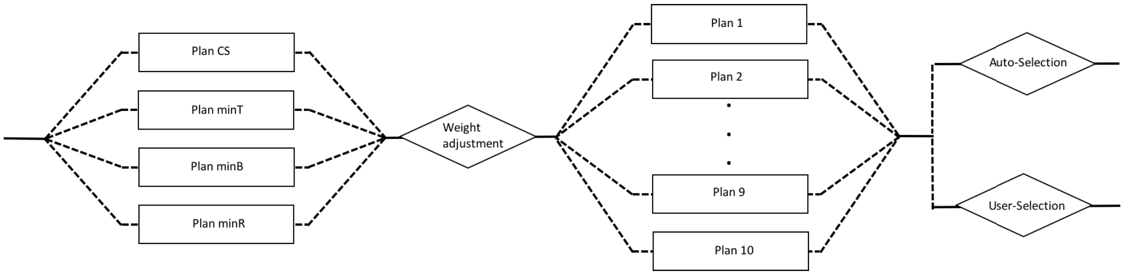


FIGURE 3.3 – The MCO planning mechanism: four anchor plans computation, weight adjustment, ten alternative plans computations, and plan selection (from left to right).

3.4.4.1 Anchor plan computation

This step consists of four anchor plans computations. Usually, an anchor plan is defined as a Pareto optimal solution obtained by minimizing one single objective [135]. The definition of anchor plans was extended by adding the class solution as an anchor plan:

Anchor plan CS:

class solution (CS) or equally minimizing all organs.

$(w_{TU}, w_B, w_R) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$

Anchor plan minT:

only minimizing the target volume (minT).

$(w_{TU}, w_B, w_R) = (1, 0, 0)$

Anchor plan minB:

only minimizing the bladder volume (minB).

$(w_{TU}, w_B, w_R) = (0.1, 0.9, 0)$

Anchor plan minR:

only minimizing the rectum volume (minR).

$$(w_{TU}, w_B, w_R) = (0.1, 0, 0.9)$$

Ideally for anchor plan minB or anchor plan minR, the value of w_{TU} will be set to zero and only the bladder volume or the rectum volume will be optimized. However, in order to keep the optimization from being trapped at an invalid solution (zero dwell times for all dwell positions), an arbitrary and small value of 0.1 is selected to represent an optimization that almost completely favors the bladder or the rectum.

3.4.4.2 Weight adjustment

In the companion paper, organ-specific dosimetric indicators (di) were introduced, and their values can be used to predict the weights in the clinically relevant region of the Pareto surface. Dosimetric indicators are determined by anchor plans (Equations (3.2a) - (3.2c)), and their values indicate the level of difficulty to reach a clinical plan from a planning template (class solution).

$$di_{TU} = \frac{V_{100,RTOG} - V_{100,CS}}{(V_{100,minT} - V_{100,CS})/(w_{TU,minT} - w_{TU,CS})} \tag{3.2a}$$

$$di_{B} = \frac{D_{1ccB,RTOG} - D_{1ccB,CS}}{(D_{1ccB,minR} - D_{1ccB,CS})/(w_{B,minR} - w_{B,CS})} \tag{3.2b}$$

$$di_{R} = \frac{D_{1ccR,RTOG} - D_{1ccR,CS}}{(D_{1ccR,minB} - D_{1ccR,CS})/(w_{R,minB} - w_{R,CS})} \tag{3.2c}$$

In the definition of target-urethra dosimetric indicators di_{TU} (Equation (3.2a)), for instance, variables $V_{100,RTOG}, V_{100,CS}, V_{100,minT}$ are the minimal target V_{100} values required by RTOG, resulted in anchor plan CS, and resulted in anchor plan minT, respectively. A value of 95% (a common criteria in clinic) was assigned to $V_{100,RTOG}$. Variables $w_{TU,minT}, w_{TU,CS}$ (defined in Section 3.4.4.1) are the target-urethra hidden weights in anchor plan minT and in anchor plan CS, respectively.

The numerator of di_{TU} reflects the difference between $V_{100,RTOG}$ (expectation) and $V_{100,CS}$ (starting point), and the denominator of di_{TU} reflects the change of V_{100} per unit of hidden weight w_{TU}(or $\triangle V_{100}/\triangle w_{TU}$). As a result, the value of di_{TU} indicates the level of difficulty to obtain a 95% of target V_{100} value from the class solution. The same approach was also used for OARs, using D_{1cc} for the bladder and the rectum (Equations (3.2b) and (3.2c)).

One of the major results in the companion study is the regression curves (along with their corresponding equations in Figure 3.4) that illustrate the relationship between dosimetric indicators and the hidden weights of the clinically relevant solution space. Therefore, by inputing the results of dosimetric indicators (or anchor plans) into the regression curves, the hidden weight requirements to reach a RTOG valid solution can be estimated.

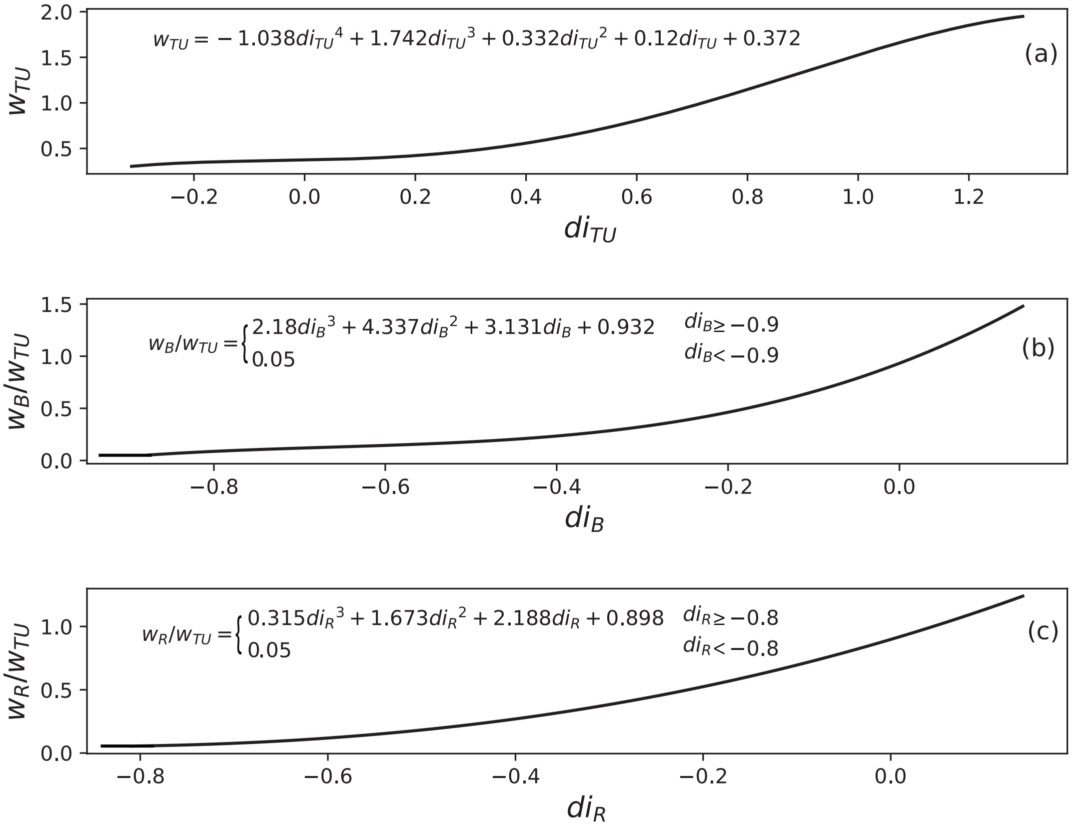


FIGURE 3.4 – Regression curves to predict the RTOG valid region on the Pareto surface based on dosimetric indicators for (a) prostate, (b) bladder and (c) rectum (hidden weight larger than 1 is not necessary achievable, and will be scaled during optimization.).

3.4.4.3 Alternative plan computation

In Section 3.4.4.2, the requirements of hidden weights to reach the RTOG valid region can be estimated, e.g., $w_{TU} \geq a$, $w_B/w_{TU} \geq b$, $w_R/w_{TU} \geq c$ (or approximated as: $w_{TU} \geq a$, $w_B \geq a \times b$, $w_R \geq a \times c$, where a, b, c refer to three constants that are calculated by three equations in Figure 3.4).

In addition to anchor plans, a set of alternative plans that variously favor different organs were created. However, there are two problems when constructing a high quality plan dataset: (1) the requirements are only estimations, the dataset may fail to include any RTOG valid plans, if the dataset size is too small. (2) if $a + a \times b + a \times c > 1$, it could be difficult or impossible to obtain a RTOG valid plan, because these requirements could not be met simultaneously.

For the first problem, 10 alternative plans are generated as a balance between the cost of computational resources and the quality of the plan dataset based on our observations, due to the fact that dual six-core CPUs are able to compute up to 12 plans in parallel.

As for the second problem, it is necessary to rescale a, b, c, so that $a + a \times c + a \times c$ is always below 1. A comparison between a scaling factor of $0.8/(a + a \times c + a \times c)$ and a scaling factor of $0.9/(a + a \times c + a \times c)$ suggested that the scaling factor of $0.8/(a + a \times c + a \times c)$ could better guarantee a RTOG valid plan.

In this way, the rescaled requirements of hidden weights can guide the plan operators to an estimated RTOG valid region on the Pareto surface. Furthermore, by assigning an additional value of 0.2 (i.e., $1 - 0.8$) to the hidden weights w_{TU}, w_B, w_R according to a mathematical combination rule, 10 alternative plans that variously favor different organs were created near the estimated RTOG valid region.

3.4.4.4 Plan selection

After the computations detailed in Sections 3.4.4.1, 3.4.4.2, 3.4.4.3, a plan dataset consisting of 4 anchor plans and 10 alternative plans was created. There are two options to select the final plan: user-selection and auto-selection.

With the option of user-selection, this step will be left to treatment planners according to the clinical practice. A python script was implemented to facilitate the navigation to the final plan among the plan dataset. With the option of auto-selection, a plan can be automatically selected for the user from the plan dataset. The principles of selecting the final plan are described as follows:

- If there is at least one RTOG valid plan in the plan dataset, choose the one with a maximal target V_{100} value.
- If there are no RTOG valid plan in the plan dataset, choose the one with a minimal target V_{100} value (but still $\geq 90\%$), while V_{75} values of the bladder and the rectum $\leq 1\,\text{cm}^3$.

Figure 3.5 illustrates an example of the four steps of the MCO planning. In this example, four anchor plans (stars) were computed, and the results of dosimetric indicators were calculated (Equations 3.2). The requirements (dashed line) of hidden weights to reach the RTOG valid region were estimated based on the results of dosimetric indicators (Figure 3.4). Alternative plans (solid dots) were computed within the range defined by the anchor plans and finally one plan (open squares) was selected from the plan dataset.

3.4.5 Comparisons

For each case, a physician-approved plan obtained using the standard planning was referred as a standard plan. As a retrospective study, the dosimetric parameters results of this plan were

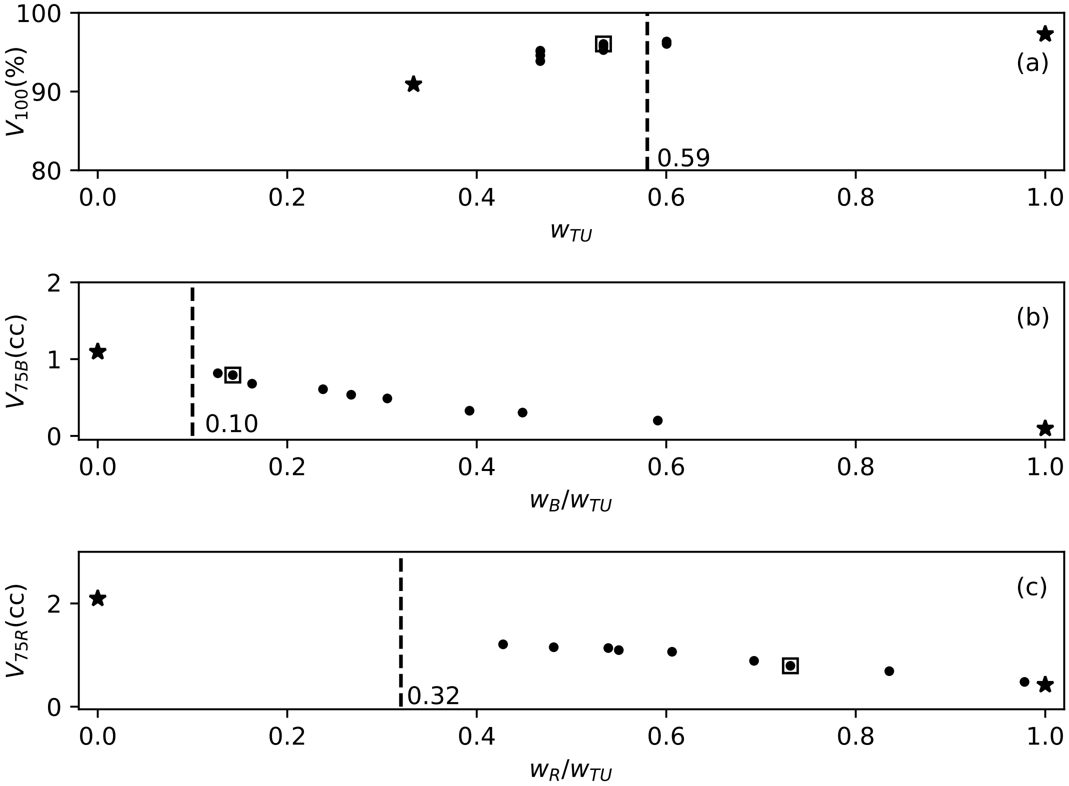


FIGURE 3.5 – An example to illustrate the MCO planning: anchor plan computation (stars) in Section 3.4.4.1, weight adjustment (dashed lines) in Section 3.4.4.2, alternative plan computation (solid dots) in Section 3.4.4.3, and plan selection (open squares) in Section 3.4.4.4.

retrieved using `3D Slicer` [155] from the RS and RD DICOM images that were exported from Oncentra Prostate (Elekta Brachy, Veenendaal, The Netherlands). On the other hand, a plan obtained using the MCO planning under the auto-selection option was referred as a MCO plan.

A set of dosimetric comparisons as well as a planning time comparison were conducted for 236 prostate cases between the standard planning and the MCO planning.

Dosimetric comparisons between two planning approaches include a histogram that illustrated the distributions of dosimetric parameters, and a box-and-whisker diagram that provided the statistic information (such as the median, the upper quartile and the lower quartile) as a supplement.

3.4.6 Computational implementations

The planning time was measured on an Intel Xeon X5650 processor (12M Cache, 2.66 GHz). To achieve a high efficiency, a multithreading scheme `openMP` [156] was implemented for both the anchor plans (Section 3.4.4.1) and the alternative plans computations (Section 3.4.4.3). After acceleration using `openMP`, the MCO planning time was measured on dual Intel Xeon X5650 processors.

In addition, the number of IPSA iterations to compute a plan is 50 000 because the results of the companion study indicated that this number was a clinically pragmatic choice and that the improvement of further iterations (200 000 iterations) was limited.

3.5 Results

3.5.1 The interface of the user-selection option

The procedure of user-selection is interactive, a real-time comparison of both dose-volume-histograms (DVHs) and dosimetric parameters is displayed as the user drags the ID slider (left in the Figure 3.6). A final plan can be selected from the plan dataset until the user is satisfied with the dosimetric parameters results.

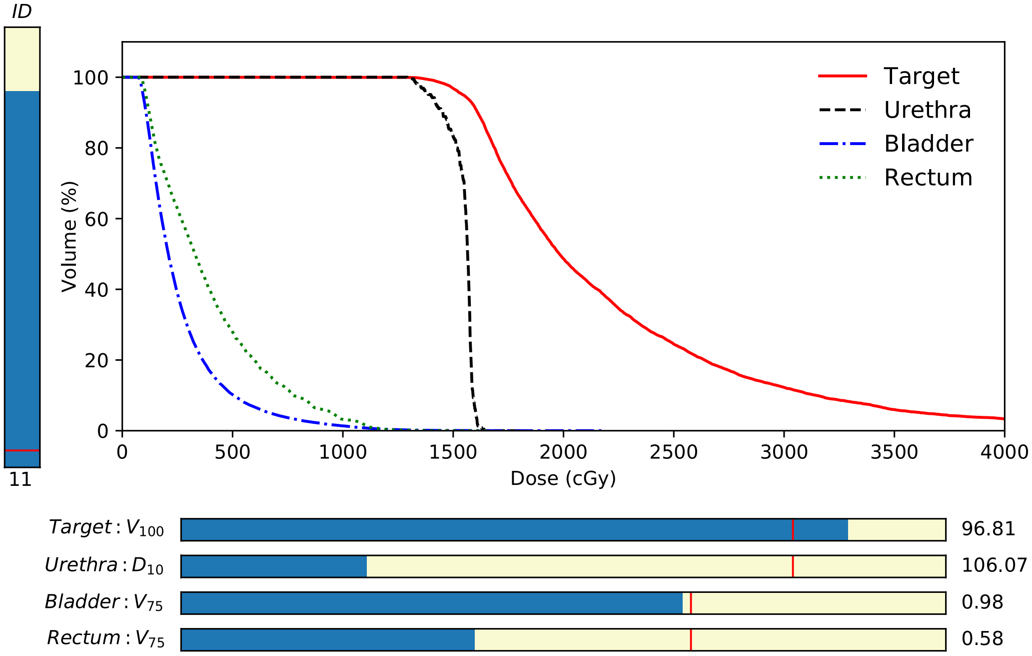


FIGURE 3.6 – An interactive interface that helps user navigate a final plan under the user-selection option (left: the ID slider, middle: the DVH result, bottom: the dosimetric parameter slider, vertical lines in the dosimetric parameter slider indicate the corresponding clinical criteria).

3.5.2 Dosimetric results

In Figure 3.7, the histogram on the left compares the frequency of the dosimetric parameters between two planning methods, and the box plot on the right visualizes the descriptive statistics, such as the median, the upper quartile and the lower quartile.

The proportion of plans with a target V_{100} value of at least 95% or more was 55.5% in the standard plans, and 68.6% in the MCO plans.

As for OARs, a lower mean urethral D_{10} value -2.25%) was observed in the MCO plans, and all plans complied with the RTOG criteria of the urethra ($D_{10} > 118\%$).

The proportion of plans that violated the bladder RTOG criteria ($V_{75} > 1\,\mathrm{cm}^3$) was 15.25% and 2.54% in the standard and MCO plans respectively.

Finally, 2.54% of the standard plans and 2.97% of the MCO plans violated the rectum RTOG criteria ($V_{75} > 1\,\mathrm{cm}^3$).

Overall, the probability of clinically (RTOG, using V_{100} at 90% or more) acceptable plan was 97.46% for the MCO plans without any manual interventions, and was 83.05% for the standard plans with manual intervention.

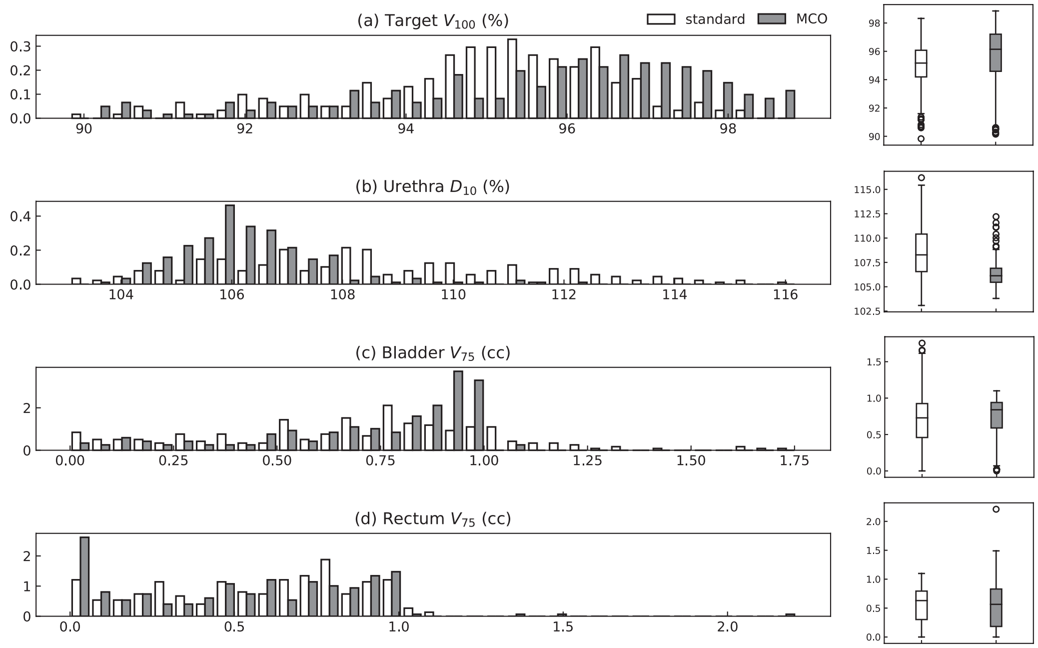


FIGURE 3.7 – A further comparison of dosimetric parameters results between the standard plans and the MCO plans: a histogram (left) that compares the frequency of dosimetric parameters values, and a box-and-whisker plot (right) that compares the descriptive statistics informations of dosimetric parameters.

3.5.3 Performance

The average standard planning time (a single plan with dose optimization and DVH computation) was around 15 s. The average MCO planning time (14 plans) was 3.9 min before acceleration, and was 41 s after acceleration using `openMP`. The time for manual review of the isodose curves is not included in the numbers quoted above.

3.6 Discussion

MCO planning is capable to improve the number of RTOG valid plans based on the following results of two planning methods: (1) a comparable target coverage and urethra protection (all plans but one (standard) meet the RTOG criteria), (2) a comparable rectum protection (plans with a $>1\,\mathrm{cm}^3$ V_{75}: 2.97% (2.54%) for MCO (standard). (3) a significant improvement in the bladder protection of the MCO plans (plans with a $> 1\,\mathrm{cm}^3$ V_{75} value: 2.54% (15.25%) for MCO (standard).

The presented work has shown that it is possible to remove the manual weight adjustments using the resulting regression models, and that the computing efficiency in planning can be potentially improved with the MCO planning. Multiple plans with various trade-offs generated within a clinically acceptable time were presented to the physician to choose from. In this way, the MCO planning enabled a true patient-specific inverse planning algorithm. The proposed algorithm can be implemented in a clinical TPS as an adjunct treatment planning tool that is able to rapidly provide optimal plans by providing already generated alternative plans, even to inexperienced planners. To the best of our knowledge, this is the first study to establish an advanced inverse planning algorithm in HDR brachytherapy for prostate cases.

It is worth noting that the regression models were estimated from a relatively small training set (140), however the results were validated with a relatively large set (236). This is particularly important when high quality and consistent data are limited.

In future studies, the philosophy of the proposed MCO planning algorithm may be tested on other HDR brachytherapy sites such as gynecologic and breast. The implementation of the MCO planning into current treatment planning system may also be desirable.

3.7 Conclusion

A patient-specific brachytherapy inverse planning algorithm called multi-criteria optimization planning is presented in this two-part study. With this algorithm, high quality treatment plans are generated avoiding the need for user intervention in the standard manual iterative planning approach, and the planning time spent in the operating room could therefore be reduced.

3.8 Acknowledgements

This work was supported in part by the National Sciences and Engineering Research Council of Canada (NSERC) via the NSERC-Elekta Industrial Research Chair (Grant number: 484144-15). The authors acknowledge a scholarship from the Chinese Scholarship Council, and acknowledge partial support by the CREATE Medical Physics Research Training Network grant of the Natural Sciences and Engineering Research Council (Grant number: 432290). Computations were made on the supercomputer Briarée from Université Laval, managed by Calcul Québec and Compute Canada. The operation of this supercomputer is funded by the Canada Foundation for Innovation (CFI), the ministère de l'Économie, de la science et de l'innovation du Québec (MESI) and the Fonds de recherche du Québec - Nature et technologies (FRQ-NT). The authors acknowledge the supports from their colleagues, especially Frédéric Lacroix, Éric Poulin, Nicolas Varfalvy, Marie-Claude Lavallée for answering the questions and providing the clinical class solution, Yunzhi Ma for the brainstorming discussions, and Paul Edimo for the data collection.

Chapitre 4

A GPU-based multi-criteria optimization algorithm for HDR brachytherapy

Cédric Bélanger[1,2*], Songye Cui[1,2*], Yunzhi Ma[2], Philippe Després[1,2], J. Adam M. Cunha[3], Luc Beaulieu[1,2]

1. Department of Physics, Engineering Physics and Optics and Cancer Research Center, Université Laval, Quebec City, QC, G1V 0A6, Canada
2. Department of Radiation Oncology and Research Center of CHU de Québec - Université Laval, Quebec City, QC, G1R 2J6, Canada
3. Radiation Oncology, University of California, San Francisco, CA, 94115, USA

* These authors contributed equally to this work.

4.1 Résumé

Actuellement, dans la planification de la curiethérapie HDR, un réglage manuel d'une fonction objectif est nécessaire pour obtenir des plans valides spécifiques à chaque cas. Cette étude vise à faciliter ce processus en proposant un algorithme de planification inverse spécifique au patient pour la curiethérapie HDR de la prostate : optimisation multicritère basée sur GPU (gMCO).

Deux moteurs d'optimisation basés sur GPU, comprenant le recuit simulé (gSA) et un optimiseur quasi-Newton (gL-BFGS), ont été mis en œuvre pour calculer plusieurs plans en parallèle. Après évaluation de l'équivalence et des performances de calcul de ces deux moteurs d'optimisation, un moteur d'optimisation préféré a été sélectionné pour l'algorithme gMCO. Cinquante-soixante-deux cas de HDR de prostate précédemment traités ont été divisés en ensemble de validation (100) et en ensemble de test (462). Dans l'ensemble de validation, le nombre de plans optimaux de Pareto permettant d'obtenir la meilleure qualité de plan a été déterminé pour l'algorithme gMCO. Dans l'ensemble de tests, les plans de gMCO ont été comparés aux plans cliniques approuvés par les médecins.

Nos résultats indiquent que le processus d'optimisation est équivalent entre gL-BFGS et gSA et que les performances de calcul du gL-BFGS sont jusqu'à 67 fois plus rapides que celles de gSA. Sur 462 cas, le nombre de plans valides sur le plan clinique était de 428 (92.6%) pour les plans cliniques et de 461 (99.8%) pour les plans gMCO. Le nombre de plans valides avec une couverture cible de V_{100} supérieure à 95% était de 288 (62.3%) pour les plans cliniques et de 414 (89.6%) pour les plans gMCO. Le temps de planification moyen était de 9.4 s pour que l'algorithme gMCO génère des plans optimaux de Pareto.

En conclusion, le gL-BFGS est capable de calculer des milliers de plans de traitement équivalents au SA en très peu de temps. Optimisé par gL-BFGS, un algorithme d'optimisation multicritères ultra-rapide et robuste a été mis en œuvre pour la curiethérapie HDR de la prostate. Des pools de plans avec divers compromis peuvent être créés avec cet algorithme. Une comparaison à grande échelle avec les plans cliniques approuvés par les médecins a montré que la qualité du plan de traitement pouvait être améliorée et que le temps de planification pouvait être considérablement réduit avec l'algorithme gMCO proposé.

4.2 Abstract

Currently in HDR brachytherapy planning, a manual fine-tuning of an objective function is necessary to obtain case-specific valid plans. This study intends to facilitate this process by proposing a patient-specific inverse planning algorithm for HDR prostate brachytherapy: GPU-based multi-criteria optimization (gMCO).

Two GPU-based optimization engines including simulated annealing (gSA) and a quasi-Newton optimizer (gL-BFGS) were implemented to compute multiple plans in parallel. After evaluating the equivalence and the computation performance of these two optimization engines, one preferred optimization engine was selected for the gMCO algorithm. Five hundred sixty-two previously treated prostate HDR cases were divided into validation set (100) and test set (462). In the validation set, the number of Pareto optimal plans to achieve the best plan quality was determined for the gMCO algorithm. In the test set, gMCO plans were compared with the physician-approved clinical plans.

Our results indicated that the optimization process is equivalent between gL-BFGS and gSA, and that the computational performance of gL-BFGS is up to 67 times faster than gSA. Over 462 cases, the number of clinically valid plans was 428 (92.6%) for clinical plans and 461 (99.8%) for gMCO plans. The number of valid plans with target V_{100} coverage greater than 95% was 288 (62.3%) for clinical plans and 414 (89.6%) for gMCO plans. The mean planning time was 9.4 s for the gMCO algorithm to generate 1000 Pareto optimal plans.

In conclusion, gL-BFGS is able to compute thousands of SA equivalent treatment plans within a short time frame. Powered by gL-BFGS, an ultra-fast and robust multi-criteria optimization algorithm was implemented for HDR prostate brachytherapy. Plan pools with various trade-offs can be created with this algorithm. A large-scale comparison against physician approved clinical plans showed that treatment plan quality could be improved and planning time could be significantly reduced with the proposed gMCO algorithm.

keywords: brachytherapy, prostate cancer, patient-specific, treatment planning, optimization, GPU

4.3 Introduction

About 52.3% of non-skin cancer patients receive radiation therapy during the course of their illness [3–5]. The most common radiation therapy treatment particle type used is photon, which can be delivered either externally from a medical linear accelerator (External Beam Radiation Therapy - EBRT) or internally from an inserted small radioactive source (brachytherapy, high dose rate (HDR) or low dose rate (LDR)).

Dose prescriptions in modern radiation treatment planning contain both tumor and healthy organ objectives. These objectives are often conflicting and can be generalized as: treating the tumor with high radiation dose and sparing the healthy organs with low radiation doses. Computerized treatment planning systems were used to formulate clinical prescriptions into a mathematical optimization problem, and to find treatment plans that well presented these prescriptions with treatment facilities.

However, most available algorithms are not inherently patient-specific in a sense that manual re-plannings are usually inevitable to find a clinically acceptable plan for each patient. As a result, the planning procedure can be time consuming and the planning output is planner dependant [95, 157, 158].

Several patient-specific inverse planning algorithms such as knowledge-based planning (KBP), auto-planning (AP) and multi-criteria optimization (MCO) have been proposed in EBRT. In KBP, one plan is created for a new case by searching in a prior physician-approved plan dataset based on the geometric features [95, 158–160]. In AP, a clinical plan can be obtained by interactively and automatically adapting objectives, constraints and dose shaping contours [115]. In MCO, a plan pool is constructed by generating plans with various trade-offs on Pareto surfaces [102, 146]. Similar studies can also be found in brachytherapy [64, 124, 154, 161, 162].

Our prior studies [64, 162] showed that a patient-specific treatment plan can be created without any user interventions in HDR prostate brachytherapy. However, the optimization engine of these studies was stochastic, and was implemented on CPU hardware [64, 162]. As a result, the algorithm inevitably involved an intensive computation (41 s), which may restrain its application in clinical practice because the patient is under general anesthesia in the operating room waiting for the treatment to be delivered.

The capability of graphics processing unit (GPU) architecture in reducing calculation time in medical physics were reviewed in [49, 163, 164]. The purpose of this study is to propose an ultra-fast patient-specific inverse planning algorithm on GPU for HDR brachytherapy.

4.4 Methods and Materials

This section begins with a detailed description of experimental setups including patient selection, mathematical formulations and computational specifications. Next, two inverse planning optimization engines were implemented on GPU architecture to calculate multiple plans in parallel and to populate the Pareto surfaces. Powered by the preferred optimization engine, a GPU-based multi-criteria optimization algorithm (gMCO) which is able to automatically generate clinical plans was proposed to eliminate the re-planning problem in HDR brachytherapy. In the end, a comprehensive comparison, including dosimetric performance as well as planning time, between clinical plans and gMCO plans was made.

4.4.1 Experimental setup

4.4.1.1 Patient selection

An anonymous dataset that contains 562 prostate cancer patients who received an HDR brachytherapy treatment as a boost to EBRT from 13/04/2011 to 06/07/2016 at our institution was studied. This dataset incorporates the cases studied in prior works [64, 162, 165]. Among the dataset, 100 random cases (validation set) were used to determine the number of Pareto optimal plans with the gMCO algorithm, and 462 random cases (test set) were used in the performance evaluation of the gMCO generated plans. After inserting 16-18 plastic catheters into the prostate under a transrectal ultrasound guidance, the anatomy of these patients was obtained from CT scans. Organ structures (prostate, urethra, bladder and rectum) were delineated and were imported into a commercial treatment planning system (Elekta Oncentra Brachy IPSA, Veenendaal, The Netherlands). The prescription was 15 Gy in a single fraction to the prostate. Plans were delivered using a Flexitron afterloader (Elekta Brachy, Veenendaal, The Netherlands) with an Ir-192 radioactive source.

The dwell positions were extracted from DICOM-RT files of clinical plans, and the mean number of active dwell positions (N_{act}) used for the optimization was 171 (range: 102-385). The dose calculation points were generated based on the description in Section . The mean number of dose calculation points (N_{pnt}) used for the optimization was 5913 (range: 2753-15 998), and the mean number of dose calculation points used for the dose-volume histogram (DVH) computations was 31 039 (range: 11 451-66 089).

4.4.1.2 Quadratic objective function formulation

Inverse Planning Simulated Annealing (IPSA) [55] was used as a dose optimization engine in our prior studies [64, 162]. In IPSA, piecewise linear objective functions were solved with simulated annealing [55], a stochastic optimizer. These objective functions were constructed with a population based planning template called a class solution [64, 162].

In order to implement an efficient optimizer, one option is to replace the stochastic optimizer with a gradient-based optimizer. Therefore, it may be necessary to replace the IPSA linear piecewise objective functions with piecewise quadratic objective functions, so that the first derivative (gradient) of the objective function is continuous. Quadratic objective functions are usually solved with gradient-based optimizers in radiation therapy [56, 57, 92, 166].

The dose at the i^{th} dose calculation point in the j^{th} organ, denoted by d_{ij}, is described in Equation (4.1)

$$d_{ij} = \sum_{l=1}^{N_{\text{act}}} \dot{d}_{ijl} t_l \tag{4.1}$$

where $\dot{d}_{ijl}$ is the dose rate contribution of the l^{th} dwell position to the i^{th} dose calculation point in the j^{th} organ, and t_l is the dwell time of the l^{th} dwell position. In order to avoid negative dwell times, new decision variables called dwell weight ($x_l = t_l^{1/2}$) were introduced as in [56, 57]. With this substitution, the dwell times are always non-negative ($t_l = x_l^2$).

The piecewise quadratic objective function f_{ij} at the i^{th} dose calculation point of the j^{th} organ is given in Equation (4.2)

$$f_{ij}(d_{ij}) = \begin{cases} w_{\min} \cdot (D_{\min} - d_{ij})^2 & d_{ij} < D_{\min} \\ 0 & D_{\min} \le d_{ij} \le D_{\max} \\ w_{\max} \cdot (d_{ij} - D_{\max})^2 & d_{ij} > D_{\max} \, . \end{cases} \tag{4.2}$$

Variables $D_{\min}$ and $D_{\max}$ are the underdose limit and the overdose limit respectively, and variables $w_{\min}$ and $w_{\max}$ are the corresponding weights. The corresponding gradient function g_{ij} of Equation (4.2) is described in Equation (4.3)

$$g_{ij}(x_l) = \frac{\partial f_{ij}}{\partial x_l} = \begin{cases} 4 \cdot \dot{d}_{ijl} \cdot x_l \cdot w_{\min} \cdot (d_{ij} - D_{\min}) & d_{ij} < D_{\min} \\ 0 & D_{\min} \le d_{ij} \le D_{\max} \\ 4 \cdot \dot{d}_{ijl} \cdot x_l \cdot w_{\max} \cdot (d_{ij} - D_{\max}) & d_{ij} > D_{\max} \, . \end{cases} \tag{4.3}$$

The single joint MCO objective function to be minimized is defined as a weighted sum in Equation (4.4)

$$F = \sum_{j=1}^{N_{\text{O}}} w_j \cdot \frac{1}{N_{\text{pnt},j}} \sum_{i=1}^{N_{\text{pnt},j}} f_{ij}(d_{ij}) \tag{4.4}$$

where N_{O} is the number of organs, $N_{\text{pnt},j}$ is the number of dose calculation points in the j^{th} organ. w_j is a hidden weight applied to the objectives (surface and volume) of the j^{th} organ to introduce trade-off in the solution space around the population-based starting point as in [64, 162]. The hidden weights are always non negative and their sum is one (because of the weighted sum method).

The original class solution designed for the piecewise linear objective functions [64, 162] will no longer be appropriate to construct the new quadratic objective functions, and so a new one must be designed (Table 4.1).

TABLE 4.1 – The class solution to formulate quadratic objective functions (Equation (4.2)) for 15 Gy prostate boost HDR treatment (Surface: surface dose calculation points, Volume: volume dose calculation points).

Organ	Surface				Volume			
	$w_{\min}$	$D_{\min}$(Gy)	$D_{\max}$(Gy)	$w_{\max}$	$w_{\min}$	$D_{\min}$(Gy)	$D_{\max}$(Gy)	$w_{\max}$
Target	200	15	22.5	80	200	15	22.5	1
Urethra	30	14	16	160	30	14	16	160
Bladder	0	0	7.5	60	0	0	7.5	60
Rectum	0	0	7.5	15	0	0	7.5	15

4.4.1.3 Computational specifications

The CPU algorithm was written in *C++*, compiled with *g++* (7.3.0) and executed on a six-core Intel® Xeon® CPU (E5-2620 v3 @ 2.40 GHz). The GPU algorithms were written in CUDA C, compiled with *nvcc* (CUDA toolkit 10.0.130) and executed on an NVIDIA Titan X (Pascal) GPU.

4.4.2 GPU-based efficient optimization engines

Previous studies showed that it is feasible to find clinically acceptable treatment plans after exploring Pareto surfaces with MCO approaches [102, 162]. However, constructing Pareto surfaces could be inefficient, if performed sequentially.

4.4.2.1 IPSA on GPU

A traditional CPU-based inverse planning algorithm such as IPSA (or cSA) [55] can be divided into several serial computing steps (Figure 4.1). In each step, the same operation is repeated over a large dataset. For example, the following five steps are essential in cSA:

1. Initialization and dose rate matrix calculation (repeated for: N_{pnt} dose calculation points $\times$ N_{act} dwell positions),
2. dwell time updates (repeated for: N_{act} dwell positions),

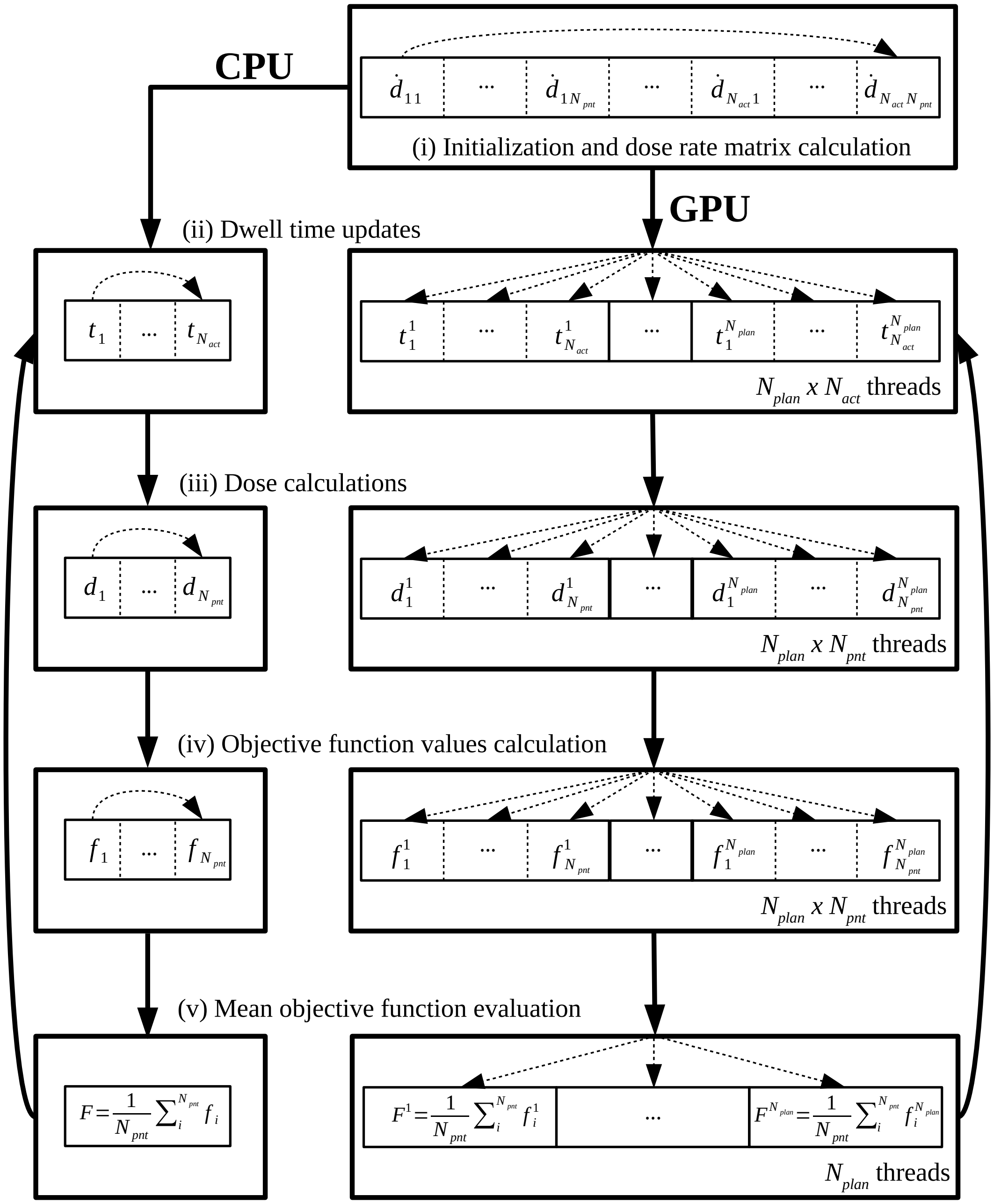


FIGURE 4.1 – Illustration of the iterative procedure to optimize one treatment plan on CPU and N_{plan} plans on GPU. In each CPU or GPU iteration, the steps (ii)-(v) are executed sequentially. In each step on the CPU, the operations are executed sequentially in a loop. In each step on the GPU, the operations are executed in parallel on different threads for N_{plan} plans. (The superscript indicates the plan number on GPU).

3. dose calculations based on Equation (4.1) (repeated for: N_{pnt} dose calculation points),
4. objective function values calculation based on Equation (4.2) (repeated for: N_{pnt} dose calculation points),
5. mean objective function evaluation based on Equation (4.4) (repeated for: one accumulation over N_{pnt} dose calculation points).

To obtain an optimal solution or a treatment plan, steps (ii)-(v) are iteratively repeated in cSA. Furthermore, in order to explore Pareto surfaces by computing N_{plan} treatment plans, it is usually necessary to repeat the aforementioned steps N_{plan} times.

To increase the efficiency of MCO approaches, GPU-based IPSA (or gSA) was implemented on GPU architecture to compute treatment plans with various trade-offs in parallel. Two strategies were applied to achieve this purpose.

First, the serial operations computed in each step in cSA were adapted to run in parallel on GPU, so the operations within each step can be executed simultaneously on different threads (Figure 4.1). Note that in each step on GPU, the computational burden is N_{plan} times larger than in the CPU implementation (N_{plan} plans on GPU vs. one plan on CPU in Figure 4.1). However, a performance gain can be achieved with the GPU implementation, as the huge burden of updating the values for all plans in each step is processed in parallel on different threads. To obtain N_{plan} optimal solutions or N_{plan} treatment plan with the proposed implementation, it is necessary to iteratively repeat steps (ii)-(v) in gSA.

Second, as frequent data transfers between CPU and GPU will slow down the computation, data transfer only occurs twice in gSA: once when preparing the data used for the optimization (CPU to GPU), once more when saving the dosimetric results onto the disk after the optimization (GPU to CPU).

4.4.2.2 Deterministic optimizer

In Section 4.4.2.1, a stochastic optimizer was implemented on CPU and on GPU. To further improve the computational performance, a deterministic optimizer (Limited-memory Broyden-Fletcher-Goldfarb-Shanno, L-BFGS) [167–169] was introduced to replace the stochastic optimizer. There are two reasons to choose this quasi-Newton optimizer, (1) BFGS and its variants are widely studied in brachytherapy [56, 57], and (2) L-BFGS is widely used in clinic after being integrated in Hybrid Inverse Planning Optimization (HIPO) (Elekta Brachy, Veenendaal, The Netherlands) [88].

So far, four optimization engines were implemented: cSA, gSA (simulated annealing on CPU and on GPU), cL-BFGS and gL-BFGS (L-BFGS on CPU and on GPU). The description of

L-BFGS implementation on CPU and on GPU is omitted in this study, due to the similarity with the context and Figure 4.1 in Section 4.4.2.1.

4.4.2.3 Equivalence between the four optimizers

The equivalence between the four optimization engines was evaluated based on the same objective function (class solution in Table 4.1) as tested over the validation set. For cSA, gSA, and cL-BFGS, one plan using uniform 5 s initial dwell times as a starting point was generated. For gL-BFGS, 1000 degenerated plans were calculated to evaluate the convergence of different starting points (randomly distributed between 0 and 10 s). The stopping criteria for cSA and gSA was specified by the number of iterations. The stopping criteria for cL-BFGS and gL-BFGS was specified by the parameter ϵ (based on the relative variation of the objective function [166]). To measure the equivalence between the four optimizers, 1 000 000 iterations and $\epsilon = 10^{-7}$ were used as the stopping criteria because no significant improvements in the objective function were observed.

4.4.2.4 Pareto surfaces characterization with gSA and gL-BFGS

Planning efficiency is a key factor when designing an inverse planning algorithm. For SA, a clinically useful stopping criteria (50 000 iterations) can be used to reach Pareto surfaces [162]. For gradient-based method, it is also desirable to find a stopping criteria that can well approximate the Pareto surfaces.

By computing solutions in parallel with various combinations of hidden weights, Pareto surfaces can be populated either with gSA and with gL-BFGS. Such solutions were Pareto optimal, or non-dominated, if no solution that improves any individual objective value without worsening at least one of the other individual objective values exists. A clinically useful stopping criteria was determined for gL-BFGS to approximate the Pareto surfaces, after examining the effect of different stopping criteria (ranging from $\epsilon = 10^{-7}$ to $\epsilon = 10^{-2}$) based on the fraction of non-dominated solutions and the speedup factor of the optimization time for all 100 validation cases.

4.4.2.5 Computational performance under clinically useful scenarios

The benefits of the proposed GPU implementation over a traditional CPU implementation of inverse planning algorithms were explored. Based on the clinically useful stopping criteria, the computational performance of cSA, gSA, cL-BFGS and gL-BFGS were measured against the number of generated plans.

4.4.3 Patient-specific multi-criteria optimization algorithm

Usually, plans obtained with a population-based planning template are not always directly acceptable, and manual weights adjustments are required to obtain a patient-specific deliverable plan. After reviewing the definition of acceptable plans, a GPU-based multi-criteria optimization algorithm (gMCO) powered with gL-BFGS was proposed to eliminate the procedure of manual weights adjustments.

4.4.3.1 Plan evaluation

The schedules of dose fractionation and the evaluation criteria of HDR prostate brachytherapy plans may vary between centers [41]. According to the Radiation Therapy Oncology Group (RTOG) 0924 protocol [18], RTOG acceptable plans (or valid solutions) can be summarized as follows:

— Prostate/Target coverage constraint: $V_{100} \geq 90\%$ of the volume.
— Urethra constraint: $D_{10} < 118\%$ of the prescription dose.
— Bladder constraint: $V_{75} < 1$ cc.
— Rectum constraint: $V_{75} < 1$ cc.

Note:

(1) V_{x} refers to the absolute volume that receives x% of the prescription dose, and D_{x} refers to the percent of the prescription dose that covers x% of the volume.

(2) In this study, a more stringent set of criteria was introduced. It is designated by the RTOG+ symbol and is the same as the RTOG criteria set except that it specifies a higher target coverage requirement of 95% for the V_{100}. This is usually attainable in the clinic without sacrificing the OAR protection.

4.4.3.2 gMCO algorithm

Compared with our previous studies [64, 162], there are three main differences in gMCO: (1) the trade-off between target and urethra is now explored, (2) the Pareto surfaces are widely explored with a large number of plans, as no prior knowledge of the RTOG+ valid solution space is involved, and (3) the validation cases were used to determine the number of parallel plans (from 1 to 10 000) needed to achieve high RTOG and RTOG+ acceptance rates with random hidden weights. In gMCO, the parallel plan computations were executed with gL-BFGS.

4.4.4 Comparison between clinical plans and gMCO plans

A plan pool was created with the gMCO algorithm. One plan was selected from the plan pool and was referred to as the gMCO plan.

The criteria used for plan selection are, in descending order of priority: RTOG+ valid plan, RTOG valid plan, RTOG invalid plan (violates at least one criteria). If multiple RTOG or RTOG+ valid plans existed, the one with a highest target V_{100} was selected. If multiple RTOG invalid plans existed, the one with the lowest bladder and rectum V_{75} (while not violating the criteria for target and urethra) was selected.

4.4.4.1 Dosimetric performance

The dosimetric results of clinical plans were retrieved from Oncentra Brachy (Elekta Brachy, Veenendaal, The Netherlands). Dosimetric comparisons between clinical plans and gMCO plans were analyzed for 462 test cases. The overall result was examined based on RTOG and RTOG+ acceptance rates (the criteria of all organs were met). The acceptance rate (i.e. target V_{100}, urethra D_{10}, bladder V_{75}, and rectum V_{75}) for each organ was also reported.

4.4.4.2 Planning time

The planning time consists of the time taken for dose calculation points creation, dose rate matrix calculation, optimization, and DVH calculation on GPU. The calculation time of each portion was recorded for gMCO plans. The total planning time was compared between clinical plans and gMCO plans.

4.5 Results

4.5.1 GPU-based optimization engines

4.5.1.1 Equivalence between the four optimizers

The optimization processes of the four optimizers for one random validation case are illustrated in Figure 4.2. From this figure, (1) gL-BFGS plans obtained with different initial dwell times converge to the SA objective function value, (2) no significant differences (within 0.02%) in objective function values resulted from the four optimizers were observed. Over all 100 validation cases, similar results were observed because the final objective function values of the four algorithms were in agreement within 0.2%.

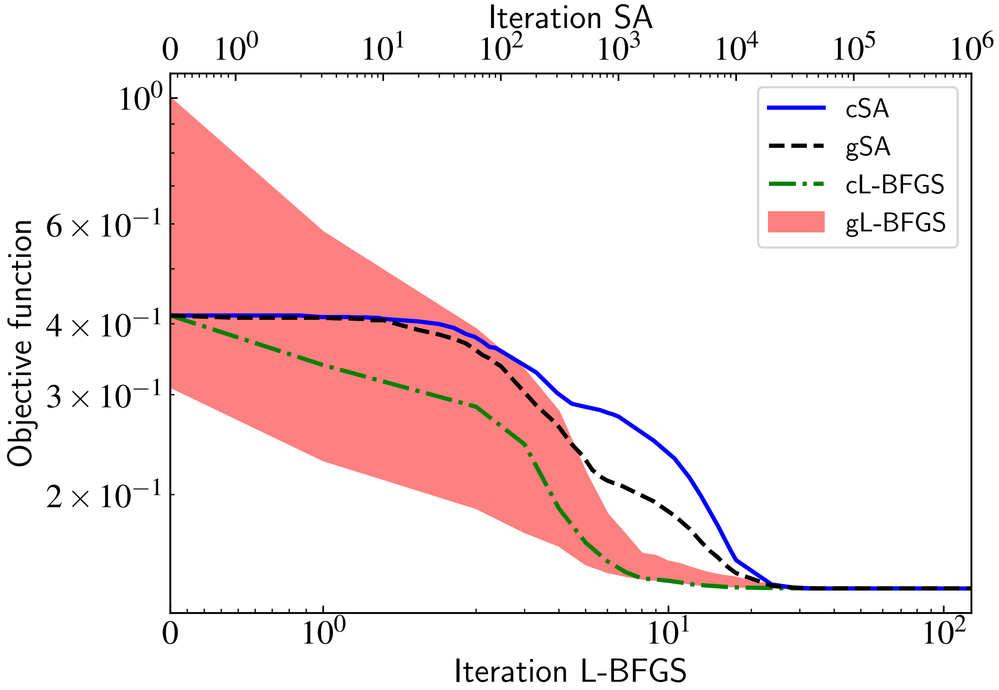


FIGURE 4.2 – Illustration of cSA, gSA, cL-BFGS and gL-BFGS objective function values against the number of iterations for one random validation case. (The difference between CPU and GPU random number generators accounts for the different trajectories for cSA and gSA).

4.5.1.2 Pareto surfaces characterization with gSA and gL-BFGS

To characterize Pareto surfaces, 100 000 different solutions were generated with gSA and gL-BFGS (1000 solutions/case for all 100 validation cases). For gSA, the mean fraction of non-dominated solutions was 99.6% under 50 000 iterations.

For gL-BFGS, the results in figure 4.3a indicate that the fraction of non-dominated solutions decreased (from 100% to 89.3%) as the stopping criteria increased (from $\epsilon = 10^{-7}$ to $\epsilon = 10^{-2}$). On the other hand, the speedup factor in the optimization time increased (from 1 to 10) as the stopping criteria increased (from $\epsilon = 10^{-7}$ to $\epsilon = 10^{-2}$). It should be noted that over 99.3% of the solutions obtained with a larger stopping criteria ($\epsilon = 10^{-3}$) are Pareto optimal solutions. Given that reaching optimality and a reasonable calculation time are important criteria for clinical applicability, the results in Figure 4.3a suggest that there could be a time advantage in using a larger stopping criteria ($\epsilon = 10^{-3}$).

Furthermore, a single 2D Pareto surface characterization with gSA and gL-BFGS is shown in Figure 4.3b. The results suggest that no significant difference in Pareto surfaces approximations is observed with GPU-based optimization engines under clinically useful stopping criteria and under more strict stopping criteria as specified in Section 4.4.2.3. From these results, $\epsilon = 10^{-3}$

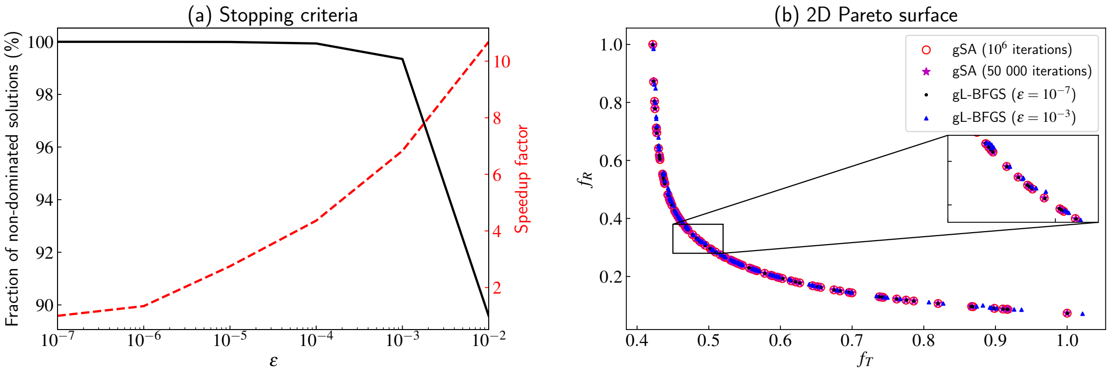


FIGURE 4.3 – (a) Effect of the stopping criteria on the fraction of non-dominated solutions in the Pareto front characterized with gL-BFGS (black solid line) and the speedup factor of the optimization time (red dashed line). The speedup factor are normalized to the values obtained with a stopping criteria of $\epsilon = 10^{-7}$. (b) A comparison of 2D Pareto surface approximations with gSA and gL-BFGS optimization engines for a random case. (f_T is denoted for target individual objective function and f_R is denoted for rectum individual objective function).

is used as the stopping criteria in gL-BFGS afterwards.

4.5.1.3 Computational performance under clinically useful scenarios

Under clinically useful scenarios, the optimization time of cSA, gSA, cL-BFGS and gL-BFGS are shown in Figure 4.4a. From the results, the time of all four engines increased as the number of plans increased. For 1000 plans, the mean optimization time was 9.2 s/plan (cSA), 60 ms/plan (gSA), 1 s/plan (cL-BFGS), and 0.9 ms/plan (gL-BFGS). In other words, compared with the cSA result, cL-BFGS can achieve a speedup factor up to 9, gSA can achieve a speedup factor of up to 176, and gL-BFGS can achieve a speedup factor of up to 10 990.

Figure 4.4b shows that the mean GPU memory usage increased with the number of plans for the GPU algorithms, and that the increase rate becomes significantly large when the number of plans reaches approximately 1000.

4.5.2 Multi-criteria optimization algorithm

As the hidden weights were randomly generated in gMCO algorithm, the RTOG and RTOG+ acceptance rates were measured multiple times with different random hidden weight vectors in Equation (4.4). In Figure 4.5, the RTOG+ acceptance rate increases (from 17% to 85%) and the spread of the acceptance rate distributions decreases with the number of plans. However, a number of 1000 plans was selected as the best compromised between optimization time (which

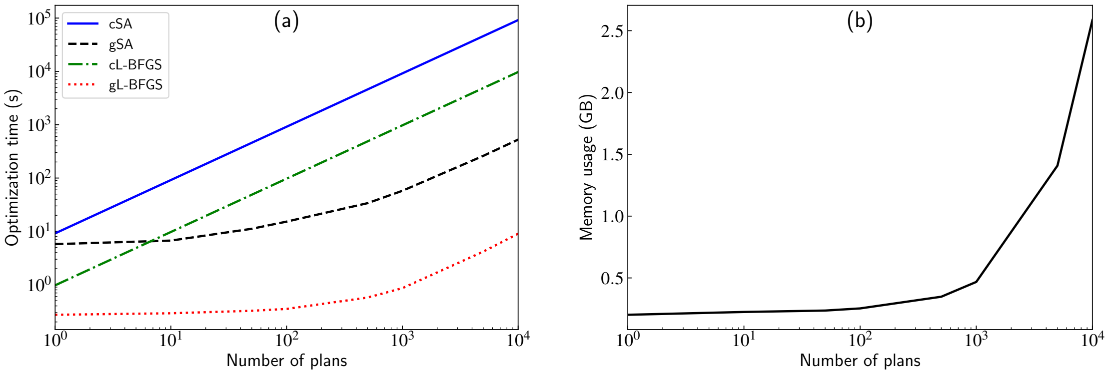


FIGURE 4.4 – Computational performance against the number of plans for cSA, gSA, cL-BFGS and gL-BFGS under clinically useful scenarios (cSA and gSA: 1000 iterations, cL-BFGS and gL-BFGS: $\epsilon = 10^{-3}$): (a) The mean optimization time, (b) The mean GPU memory usage of gL-BFGS (the result of gSA was ignored, for its similarity to the gL-BFGS one).

increases after 1000 plans, see Figure 4.4a) and the RTOG+ acceptance rate (which does not increase significantly after 1000 plans) for gMCO algorithm.

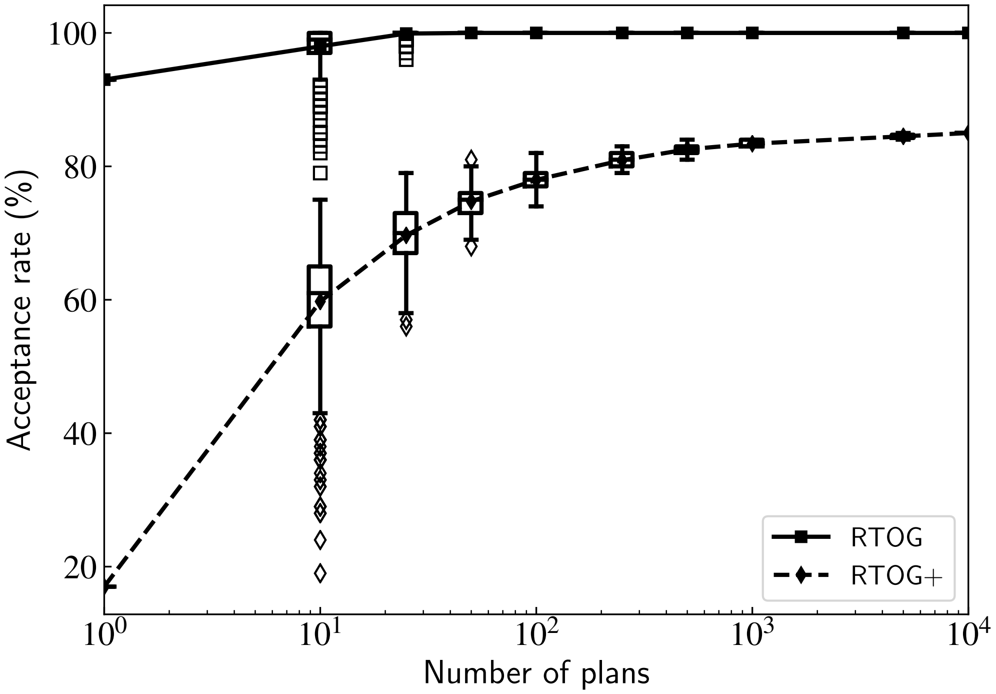


FIGURE 4.5 – The effect of the number of plans on RTOG and RTOG+ acceptance rates for gMCO (including the spread of the distributions in the boxes).

4.5.3 Comparison between clinical plans and gMCO plans

4.5.3.1 Dosimetric performance

The dosimetric comparison between clinical plans and gMCO plans is illustrated in Figure 4.6. These results suggest that the mean target coverage was higher for gMCO plans (97.2%) than for clinical plans (95.3%). The mean urethra D_{10} was significantly higher for gMCO plans (115.7%) than for clinical plans (109.1%). The mean bladder V_{75} was 0.53 cc for clinical plans, and 0.78 cc for gMCO plans. For rectum sparing, the mean rectum V_{75} was 0.56 cc for clinical plans, and 0.52 cc for gMCO plans.

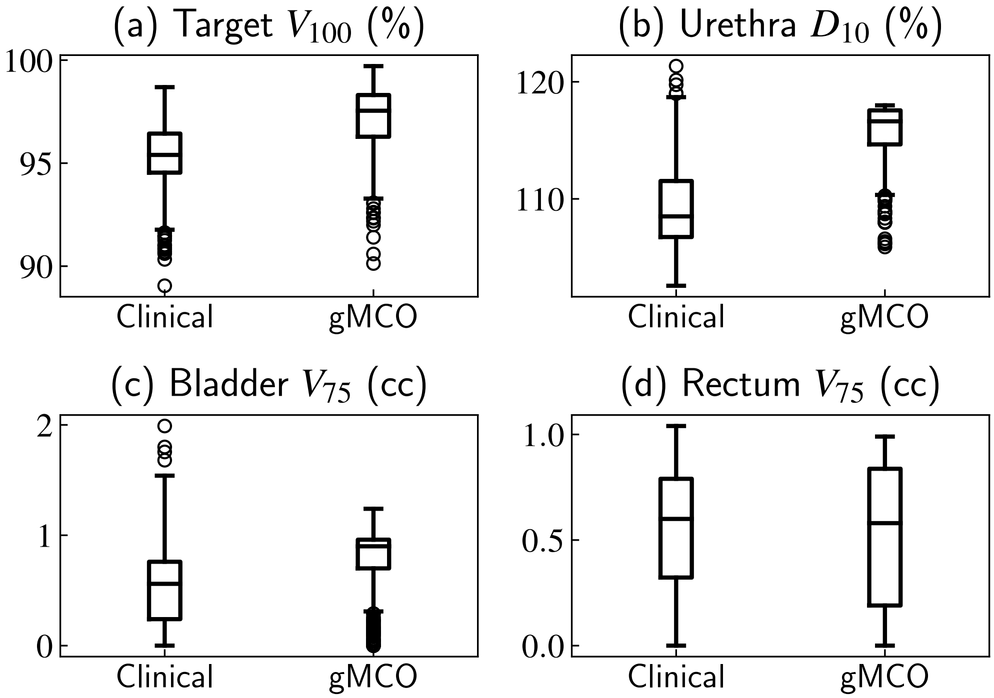


FIGURE 4.6 – Dosimetric comparison between IPSA physician-approved plans and gMCO plans over the test cohort: (a) Target V_{100}, (b) Urethra D_{10}, (c) Bladder V_{75}, and (d) Rectum V_{75}.

The acceptance rate results are summarized in Table 4.2. For overall dosimetric performances, the number of RTOG valid plans was 428 (92.6%) for clinical plans, and 461 (99.8%) for gMCO plans. The number of RTOG+ valid plans was 288 (62.3%) for clinical plans, and 414 (89.6%) for gMCO plans.

The number of plans with a target coverage greater than 95% was 296 (64.1%) for clinical plans, and 414 (89.6%) for gMCO plans. The number of plans that exceeded the urethra sparing constraint was 7 for clinical plans, and 0 for gMCO plans. The number of plans that exceeded the bladder sparing constraint was 22 for clinical plans, and 1 for gMCO plans. The

Table 4.2 – RTOG and RTOG+ acceptance rates (%) for clinically approved plans and gMCO plans over 462 test cases.

	RTOG					**RTOG+**		Time
	Target	Bladder	Rectum	Urethra	All	Target	All	
Clinical	99.8	95.2	98.7	98.5	92.6	64.1	62.3	mins
gMCO	100.0	99.8	100.0	100.0	99.8	89.6	89.6	9.4 s

number of plans that exceeded the rectum sparing constraint was 6 for clinical plans, and 0 for gMCO plans. In addition, the mean number of RTOG valid plans was 617/1000 (61.7%), and the mean number of RTOG+ valid plans was 268/1000 (26.8%) for the gMCO plan pool.

As a supplement to the general comparisons described above, one example case was chosen to illustrate the advantage of gMCO in terms of the results of DVHs and isodose curves in Figure 4.7.

4.5.3.2 Planning time

The time to create a plan is of the order of a few minutes in clinic, including manual tweaking of the objective function and/or dwell times. On the other hand, the mean planning time was 9.4 s for gMCO to generate 1000 optimal plans. Among these numbers, the mean dose calculation points creation time was 7.4 s, which represents 79% of the mean planning time. The mean optimization time was 0.8 s (8.5% of the mean total planning time). Dose rate matrix calculation and DVH calculation on GPU contribute to the rest of the mean planning time. In addition, automatically plan selection from the plan pools was performed in batch for 462 cases, and the corresponding time was negligible (4.2 s for plan selection for 462 cases).

4.6 Discussion

Our recent studies [64, 162] showed that it is possible to obtain a RTOG valid plan without any user interventions. In order to further increase the planning efficiency, four optimization engines were implemented and compared. Our results indicated that (1) gSA and gL-BFGS can speedup the optimization time by two or three orders of magnitude compared to their CPU implementation (Figure 4.4a), (2) L-BFGS is equivalent to simulated annealing, and is not trapped in local minima (Figure 4.2), (3) gL-BFGS is able to compute 10 000 plans within 9 s (optimization time in Figure 4.4a), and (4) the multi-GPU approach is not necessary, considering the fact that the mean GPU memory usage to generate 10 000 plans was 2.6 GB out of 12 GB (Figure 4.4b).

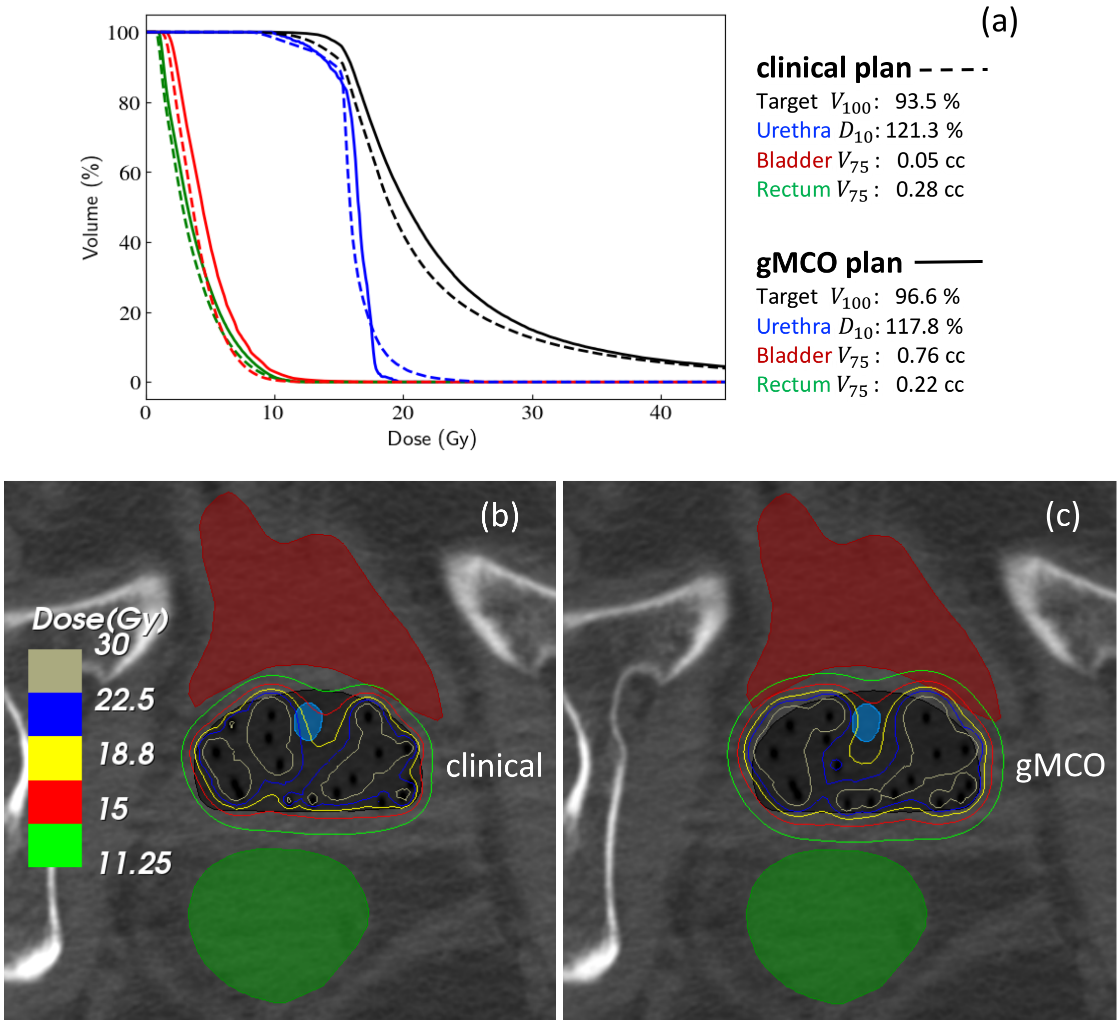


FIGURE 4.7 – A comparison between the clinical plan and the gMCO plan for one example case: (a) DVHs and dosimetric parameters, (b) and (c) Isodose curves (target in black, urethra in blue, bladder in red, and rectum in green.).

A new patient-specific approach called GPU-based MCO (gMCO) was proposed as an upgrade of our prior studies [64, 162]. gMCO can increase the RTOG acceptance rate from 97.5% [64] to 99.8%, and can decrease the planning time from 1 h (300 plans) [162], to 41 s (14 relevant plans) [64], to 9.4 s (1000 plans). Compared with the IPSA physician approved plans, gMCO can increase the RTOG+ acceptance rate by 27.3%, eliminating around 10 manual tweaking needed to achieve the observed clinical level based on the results presented in Figure 4.5. For example, a RTOG invalid plan (urethra D_{10} above 118%) can be escalated to a RTOG+ valid plan by using gMCO. This has been made possible by relaxing the bladder V_{75} dose (still below 1 cc), while still meeting all requirements for target, urethra and rectum dose parameters as shown in Figure 4.7a. Such information can also be seen from the isodose curves in Figure 4.7b-c. Note that in this study, the trade-off involved in the automatic selection scenario is based on selecting the highest target V_{100} while satisfying all the other RTOG criteria (Figure 4.6). However, a high quality gMCO plan pool is available for the user to pick a plan that best suits the patient-specific conditions.

KBP and MCO are widely used patient-specific inverse planning algorithms. In KBP, clinical plans were used to extract the regression models based on geometric features. However, clinical plans are user-dependent [51], and may be inconsistent between centers [52]. On the other hand, gMCO is independent of these issues. In MCO, even though interpolations between calculated plans were usually used to achieve a high planning efficiency, ultra-fast planning remains a challenge since no parallelization scheme was implemented. In this study, it only takes 9.4 s to generate a high quality plan pool with gMCO. However, it is admitted that these comparisons were made by ignoring that the dwell times optimization in HDR brachytherapy is a relatively small scale problem compared to the fluence map optimization in EBRT.

Note that for the objective function considered in this work, the solution space is convex and it would be easy to dismiss SA in favor of the more computationally efficient gL-BFGS algorithm. While this objective function is popular in the field, other types of objective function might have more complex solution spaces. Therefore, having a robust, albeit slower, MCO algorithm based on SA remains an essential tool.

We anticipate that the approach proposed in this study will be implemented in clinical systems as an adjunct tool. In future work, the application of gL-BFGS as well as gMCO to other HDR brachytherapy sites will be investigated.

4.7 Conclusion

Two GPU-based optimization engines were designed to calculate multiple plans in parallel. With the preferred engine, an ultra-fast patient-specific planning tool that is able to generate a high quality plan without any user interventions was proposed. After a validation over a large-scale patient cohort, both plan quality and planning efficiency can be significantly improved compared with the traditional planning in clinic.

Acknowledgement

This work was supported in part by the National Sciences and Engineering Research Council of Canada (NSERC) via the NSERC-Elekta Industrial Research Chair Grant (#484144-15), via the NSERC Discovery Grants (#355493 and #435510), and via the CREATE Medical Physics Research Training Network Grant (#432290). The authors acknowledge a scholarship from the Chinese Scholarship Council and partial support by the Canada Foundation for Innovation

(#CFI30889). The authors acknowledge the supports from their colleagues, especially Louis Archambault, Paul Edimo, Andrea Frezza, Charles Joachim-Paquet, Frédéric Lacroix, Marie-Claude Lavallée, Ghyslain Leclerc, Loïc Paradis-Laperrière, Éric Poulin, and Nicolas Varfalvy.

Conclusion

The development of traditional inverse planning algorithms used nowadays dates back to over two decades ago. In brachytherapy, one well-known example is Inverse Planning Simulated Annealing (IPSA), the beta version of which was integrated into PLATO-BPS (Nucletron, Veenendaal, The Netherlands) in 2000 [82, 170].

Traditional inverse planning algorithms play an important role in radiation oncology because they can propose solutions to the multi-dimensional problems with constraints oncologist prescribed. There is a gap between traditional and operator-free inverse planning algorithms: inevitable user interventions to adjust the input planning parameters (iterative re-planning). Bridging this gap could narrow the variation in treatment plan quality and in planning efficiency across different cancer centers, and which may eventually benefit all the patients who receive radiation therapy.

This study has implemented three algorithms to enable an operator-free inverse planning in HDR brachytherapy. Several high performance computing techniques have been applied to achieve a desirable plan generation time and a large patient cohort has been used to validate the proposed algorithms.

Accomplished work and review of objectives

Three multi-criteria optimization (MCO) algorithms were proposed in this thesis: a traditional one, a knowledge-based one and a GPU-based one. The objectives of this thesis are fulfilled through the implementation of these MCO algorithms, and will be discussed as follows.

Manuscript#1: a traditional MCO algorithm

In Chapter 2, a traditional MCO algorithm was implemented to construct Pareto surfaces for one hundred and forty prostate cases treated with HDR brachytherapy. For each case, IPSA was used to compute Pareto optimal plans to approach its Pareto surface. This MCO algorithm is traditional because Pareto surfaces were approximated in the same way as the method proposed in 2006 for external beam radiation therapy (EBRT) [102].

The implementation of this MCO algorithm fulfills the first objective of this thesis: an MCO algorithm was implemented on top of a traditional dose optimization engine to automatically generate a set of treatment plans with various trade-offs. This portion of the study is critical because it suggests that the clinically acceptable region can be found around the class solution, and that it is possible to implement an operator-free inverse planning algorithm for HDR brachytherapy. The average computing time, however, is about 1 h per case, which could restrain its application in clinic, therefore, the rest of this study aims to achieve a high computing efficiency.

Manuscript#2: a knowledge-based MCO algorithm (kMCO)

In Chapter 3, a knowledge-based MCO algorithm was proposed to achieve a high computational performance. Two updates were implemented into the traditional MCO algorithm: a set of regression models were established to predict the clinically acceptable region on the Pareto surface, and ten alternative plans alongside the predicted region were computed in parallel to reduce the plan generation time.

A validation over 236 prostate cases suggest that the kMCO algorithm can lead to a large improvement in the planning efficiency (from 1 h with the traditional MCO algorithm to 41 s). In addition, the kMCO algorithm is able to improve the RTOG acceptance rate compared to the current treatment planning algorithm (IPSA).

The implementation of the kMCO algorithm fulfills the second objective of this thesis: prior knowledge and multi-core techniques were incorporated into an MCO algorithm, then kMCO plans were compared with the IPSA physician-approved plans. This portion of the study is important because it shows that a fast, operator-free inverse planning algorithm is possible in HDR brachytherapy.

Manuscript#3: a GPU-based MCO algorithm (gMCO)

In Chapter 4, a GPU-based MCO algorithm was proposed to further improve the computational performance. In traditional MCO and kMCO algorithms, simulated annealing (stochastic optimizer) was used to compute treatment plans on the CPU architecture. In the gMCO algorithm, a gradient-based optimizer (deterministic) on the GPU architecture was implemented to compute multiple plans in parallel.

A validation using 462 prostate cases shows that the average gMCO planning time (9.4 s) is lower than that obtained using the kMCO algorithm (41 s) and the clinical IPSA implementation (a few minutes). The gMCO algorithm can also improve the RTOG acceptance rate compared to the kMCO and physician-approved plans (Table 5.1).

The implementation of the gMCO algorithm shows that the gradient-based optimizer and the GPU architecture (the third and forth objectives of this study) can improve the algorithm currently used in clinic. This portion of the study makes it possible to implement a fast, operator-free MCO inverse planning algorithm in HDR brachytherapy.

Table 5.1 – A comparison between clinically approved plans, kMCO plans and gMCO plans. RTOG acceptance rate: target V_{100}>90%, bladder and rectum V_{75}<1 cc, urethra D_{10}<118%. RTOG+ acceptance rate: target V_{100}>95% is the only change).

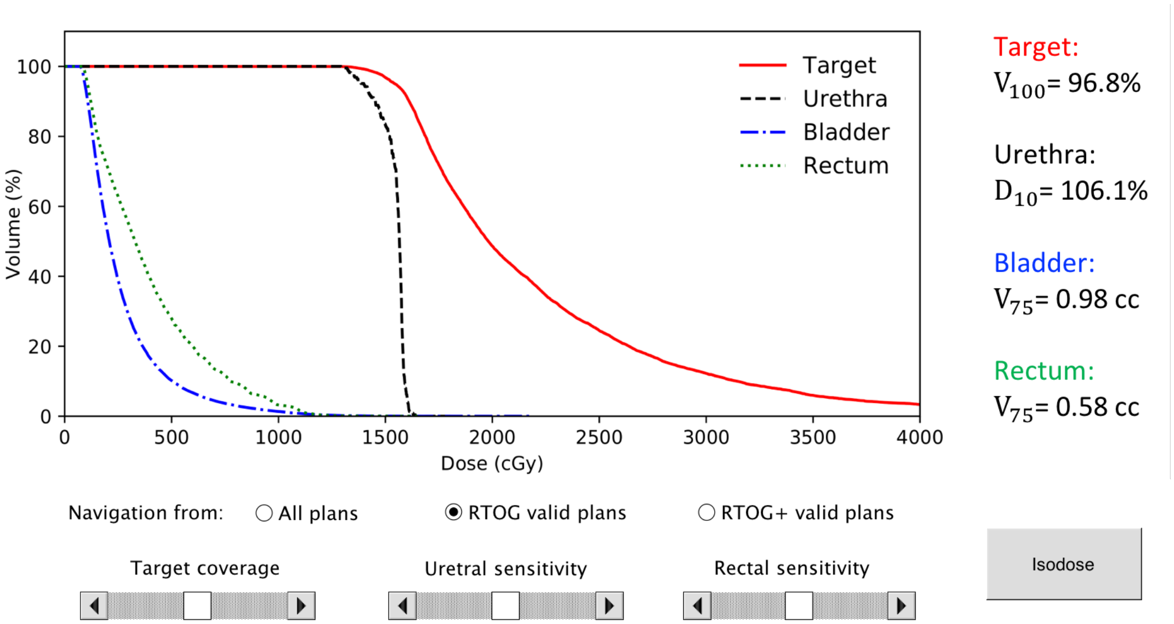
	# of cases	# of plans	RTOG (%)	RTOG+ (%)	Time
Clinical	462	1	92.6	62.3	a few minutes
kMCO	236	14	94.9	68.6	41 s
gMCO	462	1000	99.8	89.6	9.4 s

Perspectives

This thesis has explored different approaches to implement multi-criteria optimization algorithms in HDR prostate brachytherapy, and eventually leads to an ultra-fast, automated inverse planning algorithm (publication [66] inserted in Chapter 4). As a retrospective study, the clinical plans were re-planned with our MCO algorithm, and improved RTOG and RTOG+ acceptance rates could be observed.

From a clinical perspective, an efficient and intuitive plan navigation tool is as important as the MCO algorithm itself for real-time intra-operative use. An efficient plan navigation tool should allow oncologists to specify the navigation range (within RTOG valid plans or RTOG+ valid plans) of the MCO generated treatment plans. An intuitive plan navigation tool should align with the expectations from oncologists, especially when making a decision between treating the target and sparing one or multiple organs at risk (such as urethra/rectum due to their sensitivities). We envision that one possible implementation of such a tool can be seen in Figure 5.1, where plans could be selected efficiently and intuitively.

In order to conduct a clinical validation of the MCO algorithm, a protocol is needed and could be described as follows. With a traditional inverse planning algorithm, a standard plan is created under cooperation between an oncologist and a medical physicist. After using our MCO proposed algorithm, an MCO plan is created by an oncologist. The planning time of both planning approaches would be recorded, and the reports of both plans are sent for an independent and blind evaluation (scoring from 1 to 10). The comparison results would be reported, and the feedback would be collected from the oncologist who works with the MCO



Target:
V_{100}= 96.8%

Urethra:
D_{10}= 106.1%

Bladder:
V_{75}= 0.98 cc

Rectum:
V_{75}= 0.58 cc

Navigation from: ○ All plans ● RTOG valid plans ○ RTOG+ valid plans

FIGURE 5.1 – A plan navigation tool that could be used in clinical implementation.

algorithm to improve user experience and treatment outcome. Ideally, the clinical validation should be multi-institutional, to account for various fractionation schemes and plan evaluation criteria.

The clinical validation should be extended to other tumor sites, particularly interstitial GYN and breast, to test the capabilities of the MCO algorithm. To adapt the algorithm to other tumor sites, it is important to communicate with the clinical medical physicists to identify the major trade-offs, the fractionation scheme (or the prescription dose per fraction), as well as the evaluation criteria. In addition, designing a class solution and determining the size of the plan pool over a validation set (as in [66], inserted in Chapter 4) will be critical to obtain a desirable dosimetric result. In order to conduct a robust validation with an adequate statistical power, it would be preferable to have a large scale patient cohort, which could be a major challenge.

In the future, the following steps are likely to improve the optimization platform presented in this study. First, in order to avoid hot or cold spots, it is desirable to have an option to enable a smooth source movement, through controlling the relative uniformity of dwell times [74]. In addition, reducing the number of catheters may decrease the toxicity while maintaining the plan quality, thus it may be interesting to combine a catheter optimization algorithm (for example, the Centroidal Voronoi Tessellations (CVT) algorithm [171]) with the presented inverse planning algorithm. Finally, incorporating radiobiological models into current algorithms would benefit the patients, because the physical dose-volume information used to formulate the objective function and evaluate an optimized plan may not always reflect the radiation response of biological tissues. Radiobiological models reviewed in the following references [9, 172, 173] can be considered to facilitate the report and the comparison of radiobiological responses in brachytherapy.

It is worthwhile to mention that the following two aspects regarding automation could further simplify the current treatment planning procedure. First, an automated and real-time catheter reconstruction through electromagnetic (EM) tracking systems may open the possibility of dynamical (re)planning [22]. In other words, it will be possible to feed the catheter trajectory into the inverse planning algorithms in real-time, which will likely shorten the treatment course and improve the plan quality. Finally, an image segmentation algorithm that automates current manually organ delineations is desirable. Such algorithms are able to shorten the treatment course, and will play an important role in real-time treatment planning systems.

Bibliographie

[1] R. Lozano et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859):2095–2128, 2012.

[2] National Cancer Institute. Cancer treatment. `https://www.cancer.gov/about-cancer/treatment`. Accessed: 2018-12-23.

[3] V.T. DeVita, T.S. Lawrence, and S.A. Rosenberg. *DeVita, Hellman, and Rosenberg extquotelefts cancer: Principles and practice of oncology: Tenth edition.* Wolters Kluwer Health Adis (ESP), 2015. ISBN 9781451192940.

[4] G. Delaney, S. Jacob, C. Featherstone, and M. Barton. The role of radiotherapy in cancer treatment. *Cancer*, 104(6):1129–1137, 2005.

[5] D.E. Citrin. Recent developments in radiotherapy. *New England Journal of Medicine*, 377(11):1065–1075, 2017.

[6] H. Becquerel and P. Curie. Action physiologique des rayons du radium. *Compt. Rend. Acad. Sci*, 132:1289–1291, 1901.

[7] ICRU. *Dose and volume specification for reporting intracavitary therapy in gynecology.* International Commission on Radiation Units and Measurements, 1985.

[8] R. Nath et al. Guidelines by the AAPM and GEC-ESTRO on the use of innovative brachytherapy devices and applications: Report of Task Group 167. *Medical Physics*, 43(6Part1):3178–3205, 2016.

[9] J. Venselaar, A.S. Meigooni, D. Baltas, and P.J. Hoskin. *Comprehensive brachytherapy: physical and clinical aspects.* Taylor and Francis, 2012.

[10] C.V. Guthier. *Development of a real-time inverse planning system for radiation therapy based on compressed sensing.* PhD thesis, Heidelberg University, 2015.

[11] M.J. Rivard, J. Venselaar, and L. Beaulieu. The evolution of brachytherapy treatment planning. *Medical Physics*, 36(6):2136–2153, 2009.

[12] J.T. Bushberg, J.A. Seibert, E.M. Leidholdt Jr., and J.M. Boone. *The essential physics of medical imaging, 3rd edition.* Lippincott Williams and Wilkins , 2012.

[13] A. Hynes, D.A. Scott, A. Man, D.L. Singer, M.G. Sowa, and K.Z. Liu. Molecular mapping of periodontal tissues using infrared microspectroscopy. *BMC Medical Imaging*, 5(1):2, 2005.

[14] C. M'enard et al. MRI-guided HDR prostate brachytherapy in standard 1.5 T scanner. *International Journal of Radiation Oncology* Biology* Physics*, 59(5):1414–1423, 2004.

[15] J. Pouliot, Y. Kim, E. Lessard, I.C. Hsu, D.B. Vigneron, and J. Kurhanewicz. Inverse planning for HDR prostate brachytherapy used to boost dominant intraprostatic lesions defined by magnetic resonance spectroscopy imaging. *International Journal of Radiation Oncology* Biology* Physics*, 59(4):1196–1207, 2004.

[16] C.D. Kubicky et al. Inverse planning simulated annealing for magnetic resonance imaging-based intracavitary high-dose-rate brachytherapy for cervical cancer. *Brachytherapy*, 7(3):242–247, 2008.

[17] ICRU. *Prescribing, recording, and reporting photon beam therapy.* International Commission on Radiation Units and Measurements, 1994.

[18] RTOG. *Androgen deprivation therapy and high dose radiotherapy with or without whole pelvic radiotherapy in unfavourable intermediate or favourable high risk prostate cancer: A Phase III randomised trial (RTOG 0924).* Radiation Therapy Oncology Group, 2016.

[19] RTOG. *A Phase III randomized study Of hypofractionated 3D-CRT/IMRT versus Cconventionally fractionated 3D-CRT/IMRT in patients with favorable-risk prostate cancer (RTOG 0415).* Radiation Therapy Oncology Group, 2014.

[20] RTOG. *Initial results of a Phase III randomized study of high dose 3DCRT/IMRT versus standard dose 3D-CRT/IMRT in patients treated for localized prostate cancer (RTOG 0126).* Radiation Therapy Oncology Group, 2014.

[21] Elekta. *Online Help for Oncentra® Brachy 4.3*, 2016.

[22] E. Poulin, E. Racine, D. Binnekamp, and L. Beaulieu. Fast, automatic, and accurate catheter reconstruction in HDR brachytherapy using an electromagnetic 3D tracking system. *Medical Physics*, 42(3):1227–1232, 2015.

[23] L. Beaulieu, E. Racine, D.Y. Han, É. Vigneault, I.C.J. Hsu, and A. Cunha. Real-time EM-tracking based treatment platform for high dose rate prostate brachytherapy: end-to-end validation and clinical workflows. *Brachytherapy*, 15:S38–S39, 2016.

[24] E. Lessard. *Development and clinical introduction of an inverse planning dose optimization by simulated annealing (IPSA) for high dose rate brachytherapy.* PhD thesis, Université Laval, 2004.

[25] M.A. Imaging and R.O. Core. Joint AAPM/IROC Houston registry of brachytherapy sources meeting the AAPM dosimetric prerequisites. `http://rpc.mdanderson.org/rpc/`. Accessed: 2019-04-08.

[26] 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. *Medical Physics*, 22(2):209–234, 1995.

[27] M.J. Rivard et al. Update of AAPM Task Group No. 43 Report: A revised AAPM protocol for brachytherapy dose calculations. *Medical Physics*, 31(3):633–674, 2004.

[28] M.J. Rivard et al. Supplement to the 2004 update of the AAPM Task Group No. 43 Report. *Medical Physics*, 34(6):2187–2205, 2007.

[29] CLRP. The Carleton Laboratory for Radiotherapy Physics (CLRP) TG-43 Parameter Database for Brachytherapy. `http://physics.carleton.ca/clrp/seed_database`. Accessed: 2018-12-23.

[30] J. Perez-Calatayud et al. Dose calculation for photon-emitting brachytherapy sources with average energy higher than 50 keV: Report of the AAPM and ESTRO. *Medical physics*, 39(5):2904–2929, 2012.

[31] L.A. DeWerd et al. A dosimetric uncertainty analysis for photon-emitting brachytherapy sources: Report of AAPM Task Group No. 138 and GEC-ESTRO. *Medical physics*, 38(2):782, 2011.

[32] L. Beaulieu et al. Report of the Task Group 186 on model-based dose calculation methods in brachytherapy beyond the TG-43 formalism: Current status and recommendations for clinical implementation. *Medical Physics*, 39(10):6208–6236, 2012.

[33] A. Ahnesjö. Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media. *Medical Physics*, 16(4):577–592, 1989.

[34] T.R. Mackie, J.W. Scrimger, and J.J. Battista. A convolution method of calculating dose for 15-MV x rays. *Medical Physics*, 12(2):188–196, 1985.

[35] D.W.O. Rogers. Fifty years of Monte Carlo simulations for medical physics. *Physics in Medicine and Biology*, 51(13):R287–R301, 2006.

[36] S. Hissoiny, B. Ozell, H. Bouchard, and P. Després. GPUMCD: A new GPU-oriented Monte Carlo dose calculation platform. *Medical Physics*, 38(2):754–764, 2011.

[37] K.A. Gifford, J.L. Horton, T.A. Wareing, G. Failla, and F. Mourtada. Comparison of a finite-element multigroup discrete-ordinates code with Monte Carlo for radiotherapy calculations. *Physics in Medicine and Biology*, 51(9):2253–2265, 2006.

[38] O.N. Vassiliev, T.A. Wareing, J. McGhee, G. Failla, M.R. Salehpour, and F. Mourtada. Validation of a new grid-based Boltzmann equation solver for dose calculation in radiotherapy with photon beams. *Physics in Medicine and Biology*, 55(3):581–598, 2010.

[39] M. D'Amours, J. Pouliot, A. Dagnault, F. Verhaegen, and L. Beaulieu. Patient-specific Monte Carlo-based dose-kernel approach for inverse planning in afterloading brachytherapy. *International Journal of Radiation Oncology* Biology* Physics*, 81(5):1582–1589, 2011.

[40] F.M. Khan, J.P. Gibbons, and P.W. Sperduto. *Khan's Treatment Planning in Radiation Oncology*. Lippincott Williams and Wilkins, 2016.

[41] Y. Yamada et al. American Brachytherapy Society consensus guidelines for high-dose-rate prostate brachytherapy. *Brachytherapy*, 11(1):20–32, 2012.

[42] K. Moore, G.C. Kagadis, T.R. McNutt, V. Moiseenko, and S. Mutic. Vision 20/20: Automation and advanced computing in clinical radiation oncology. *Medical Physics*, 41(1):010901–n/a, 2014.

[43] C.E. Wynn-Williams. The use of thyratrons for high speed automatic counting physical phenomena. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 132(819):295–310, 1931.

[44] W.P. Petersen and P. Arbenz. *Introduction to parallel computing*, volume 9. Oxford University Press, 2004.

[45] B. Chapman, G. Jost, and R. van der Pas. *Using OpenMP: portable shared memory parallel programming (scientific and engineering computation)*. The MIT Press, 2007. ISBN 0262533022, 9780262533027.

[46] F. Gebali. *Algorithms and parallel computing*, volume 84. John Wiley and Sons, 2011.

[47] Nvidia. NVIDIA launches the world's first graphics processing unit: GeForce 256. `https://www.nvidia.com/object/IO_20020111_5424.html`. Accessed: 2018-12-23.

[48] T. Soyata. *GPU parallel program development using CUDA*. CRC Press, 2018.

[49] X. Jia, P. Ziegenhein, and S.B. Jiang. GPU-based high-performance computing for radiation therapy. *Physics in Medicine and Biology*, 59(4):R151–R182, 2014.

[50] R. Tsuchiyama, T. Nakamura, T. Iizuka, A. Asahara, S. Miki, and S. Tagawa. The opencl programming book. *Fixstars Corporation*, 63:11, 2010.

[51] I.J. Das, C.W. Cheng, K.L. Chopra, R.K. Mitra, S.P. Srivastava, and E. Glatstein. Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems. *JNCI: Journal of the National Cancer Institute*, 100(5):300–307, 2008.

[52] H.T. Chung, B. Lee, E. Park, J.J. Lu, and P. Xia. Can all centers plan intensity-modulated radiotherapy (IMRT) effectively? An external audit of dosimetric comparisons between three-dimensional conformal radiotherapy and IMRT for adjuvant chemoradiation for

gastric cancer. *International Journal of Radiation Oncology*Biology*Physics*, 71(4):1167–1174, 2008.

[53] I.J. Das, V. Moskvin, and P.A. Johnstone. Analysis of treatment planning time among systems and planners for intensity-modulated radiation therapy. *Journal of the American College of Radiology*, 6(7):514–517, 2009.

[54] J. Nocedal and S. Wright. *Numerical optimization*. Springer Science and Business Media, 2006.

[55] 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(5):773–779, 2001.

[56] N. Milickovic, M. Lahanas, M. Papagiannopoulou, N. Zamboglou, and D. Baltas. Multi-objective anatomy-based dose optimization for HDR-brachytherapy with constraint free deterministic algorithms. *Physics in Medicine and Biology*, 47(13):2263–2280, 2002.

[57] M. Lahanas, D. Baltas, and S. Giannouli. Global convergence analysis of fast multiobjective gradient-based dose optimization algorithms for high-dose-rate brachytherapy. *Physics in Medicine and Biology*, 48(5):599–617, 2003.

[58] I.C.J. Hsu, É. Lessard, V. Weinberg, and J. Pouliot. Comparison of inverse planning simulated annealing and geometrical optimization for prostate high-dose-rate brachytherapy. *Brachytherapy*, 3(3):147–152, 2004.

[59] S. Pokharel, S. Rana, J. Blikenstaff, A. Sadeghi, and B. Prestidge. Evaluation of hybrid inverse planning and optimization (HIPO) algorithm for optimization in real-time, high-dose-rate (HDR) brachytherapy for prostate. *Journal of Applied Clinical Medical Physics*, 14(4):96–107, 2013.

[60] E. Poulin, N. Varfalvy, S. Aubin, and L. Beaulieu. Comparison of dose and catheter optimization algorithms in prostate high-dose-rate brachytherapy. *Brachytherapy*, 15(1):102–111, 2015.

[61] E. Poulin, N. Varfalvy, S. Aubin, and L. Beaulieu. Comparison Between HIPO and CVT as Well as DVHO and IPSA Algorithm for Catheter and Dose Optimization in HDR Brachytherapy. *Brachytherapy*, 14:S67–S68, 2015.

[62] B. Jeroen and E. Wendy. Dose optimization in HDR brachytherapy: A literature review of quantitative models. *Economics and Management*, 2011.

[63] L.D. Boeck, J. Beliën, and W. Egyed. Dose optimization in high-dose-rate brachytherapy: A literature review of quantitative models from 1990 to 2010. *Operations Research for Health Care*, 3(2):80–90, 2014.

[64] S. Cui, P. Després, and L. Beaulieu. A multi-criteria optimization approach for HDR prostate brachytherapy: II. Benchmark against clinical plans. *Physics in Medicine and Biology*, 63(20):205005, 2018.

[65] S. Cui, P. Després, and L. Beaulieu. A knowledge-based multi-criteria optimization (kMCO) algorithm for high dose rate (HDR) brachytherapy. *Medical Physics*, 45(6):2691–2794, 2018.

[66] C. Bélanger, S. Cui, Y. Ma, P. Després, J.A.M. Cunha, and L. Beaulieu. A GPU-based multi-criteria optimization algorithm for HDR brachytherapy. *Physics in Medicine and Biology*, 64(10):105005, 2019.

[67] M. Lahanas, D. Baltas, and N. Zamboglou. A hybrid evolutionary algorithm for multi-objective anatomy-based dose optimization in high-dose-rate brachytherapy. *Physics in Medicine and Biology*, 48(3):399–415, 2003.

[68] M. Lahanas, K. Karouzakis, S. Giannouli, R.F. Mould, and D. Baltas. Inverse planning in brachytherapy: Radium to high dose rate 192 iridium afterloading. *Nowotwory Journal of Oncology*, 54(6):547–554, 2004.

[69] A. Wu, K. Ulin, and E.S. Sternick. A dose homogeneity index for evaluating 192Ir interstitial breast implants. *Medical Physics*, 15(1):104–107, 1988.

[70] D. Baltas et al. A conformal index (COIN) to evaluate implant quality and dose specification in brachytherapy. *International Journal of Radiation Oncology*Biology*Physics*, 40(2):515–524, 1998.

[71] W. Renner, T. O'Connor, and N. Bermudez. An algorithm for generation of implant plans for high-dose-rate irradiators. *Medical physics*, 17(1):35–40, 1990.

[72] D. Baltas et al. Influence of modulation restriction in inverse optimization with HIPO of prostate implants on plan quality: analysis using dosimetric and radiobiological indices. In *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, volume 25/1 of *IFMBE Proceedings*, pages 283–286. Springer Berlin Heidelberg, 2009. ISBN 978-3-642-03472-5.

[73] J.A.M. Cunha, B. Pickett, and J. Pouliot. Inverse planning optimization for hybrid prostate permanent-seed implant brachytherapy plans using two source strengths. *Journal of Applied Clinical Medical Physics*, 11(3):64–77, 2010.

[74] Elekta. *HIPO: hybrid inverse planning optimization for brachytherapy treatment planning*, 2014.

[75] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge university press, 2004.

[76] S. Kirkpatrick et al. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[77] V. Karyotis and M. Khouzani. Malware-propagative Markov random fields. In *Malware Diffusion Models for Wireless Complex Networks*, pages 107–138. Morgan Kaufmann, Boston, 2016.

[78] R.S. Sloboda. Optimization of brachytherapy dose distributions by simulated annealing. *Medical Physics*, 19(4):955–964, 1992.

[79] J. Pouliot, D. Tremblay, J. Roy, and S. Filice. Optimization of permanent 125 I prostate implants using fast simulated annealing. *International Journal of Radiation Oncology* Biology* Physics*, 36(3):711–720, 1996.

[80] É. Lessard, I.C. Hsu, and J. Pouliot. Inverse planning for interstitial gynecologic template brachytherapy: truly anatomy-based planning. *International Journal of Radiation Oncology* Biology* Physics*, 54(4):1243–1251, 2002.

[81] É. Lessard, S.L. Kwa, B. Pickett, M. Roach III, and J. Pouliot. Class solution for inversely planned permanent prostate implants to mimic an experienced dosimetrist. *Medical physics*, 33(8):2773–2782, 2006.

[82] R. Alterovitz, E. Lessard, J. Pouliot, I. Hsu, J.F. O'Brien, and K. Goldberg. Optimization of HDR brachytherapy dose distributions using linear programming with penalty costs. *Medical Physics*, 33(11):4012–4019, 2006.

[83] M. Lahanas, D. Baltas, and N. Zamboglou. Anatomy-based three-dimensional dose optimization in brachytherapy using multiobjective genetic algorithms. *Medical Physics*, 26(9):1904–1918, 1999.

[84] N.B. Milickovic, M. Lahanas, M. Papagiannopoulou, K. Karouzakis, D. Baltas, and N. Zamboglou. Application of multiobjective genetic algorithms in anatomy based dose optimization in brachytherapy and its comparation with deterministic algorithms. In *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, volume 4, pages 3919–3922, 2001.

[85] M. Lahanas, N. Milickovic, D. Baltas, and N. Zamboglou. Application of multiobjective evolutionary algorithms for dose optimization problems in brachytherapy. In *Evolutionary Multi-Criterion Optimization*, volume 1993, chapter Lecture Notes in Computer Science, pages 574–587. Springer Berlin Heidelberg, 2001. ISBN 978-3-540-41745-3.

[86] N. Milickovic, M. Lahanas, D. Baltas, and N. Zamboglou1. Comparison of evolutionary and deterministic multiobjective algorithms for dose optimization in brachytherapy. In *Evolutionary Multi-Criterion Optimization*, volume 1993, chapter Lecture Notes in Computer Science, pages 167–180. Springer Berlin Heidelberg, 2001. ISBN 978-3-540-41745-3.

[87] M. Lahanas, N. Milickovic, M. Papagiannopoulou, D. Baltas, N. Zamboglou, and K. Karouzakis. Application of a hybrid version of NSGA-II for multiobjective dose optimization in brachytherapy. *Evolutionary*, 108:299–304, 2002.

[88] A. Karabis, S. Giannouli, and D. Baltas. HIPO: A hybrid inverse treatment planning optimization algorithm in HDR brachytherapy. *Radiotherapy and Oncology*, 76:S29, 2005.

[89] A. Karabis, P. Belotti, and D. Baltas. Optimization of catheter position and dwell time in prostate HDR brachytherapy using HIPO and linear programming. In *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, volume 25/1, chapter IFMBE Proceedings, pages 612–615. Springer Berlin Heidelberg, 2009. ISBN 978-3-642-03472-5.

[90] P. Trnková et al. New inverse planning technology for image-guided cervical cancer brachytherapy: Description and evaluation within a clinical frame. *Radiotherapy and Oncology*, 93(2):331–340, 2009.

[91] P. Trnková et al. A detailed dosimetric comparison between manual and inverse plans in HDR intracavitary/interstitial cervical cancer brachytherapy. *J Contemp Brachyther*, 2(4):163–170, 2010.

[92] M. Lahanas, E. Schreibmann, and D. Baltas. Multiobjective inverse planning for intensity modulated radiotherapy with constraint-free gradient-based optimization algorithms. *Physics in Medicine and Biology*, 48(17):2843–2871, 2003.

[93] M. Hussein, B.J.M. Heijmen, D. Verellen, and A. Nisbet. Automation in intensity modulated radiotherapy treatment planning—a review of recent innovations. *The British Journal of Radiology*, 91(1092):20180270, 2018.

[94] UCSD. UCSD knowledge-based radiotherapy initiative. `http://ucsd.knowledgebasedrt.org`. Accessed: 2019-04-15.

[95] K. Moore, R.S. Brame, D.A. Low, and S. Mutic. Experience-based quality control of clinical intensity-modulated radiotherapy planning. *International Journal of Radiation Oncology*Biology*Physics*, 81(2):545–551, 2011.

[96] L.M. Appenzoller, J.M. Michalski, W.L. Thorstad, S. Mutic, and K.L. Moore. Predicting dose-volume histograms for organs-at-risk in IMRT planning. *Medical Physics*, 39(12):7446–7461, 2012.

[97] S. Satomi, T. Jun, A. Olsen Lindsey, and L. Moore Kevin. Knowledge based prediction of plan quality metrics in intracranial stereotactic radiosurgery. *Medical Physics*, 42(2):908–917, 2015.

[98] S. Shiraishi and K.L. Moore. Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy. *Medical Physics*, 43(1):378–387, 2016.

[99] Q. Xin. Optimization techniques in diesel engine system design. In *Diesel Engine System Design*, pages 203–296. Woodhead Publishing, 2013.

[100] S. Brisset and F. Gillon. Approaches for multi-objective optimization in the ecodesign of electric systems. In *Eco-Friendly Innovation in Electricity Transmission and Distribution Networks*, pages 83–97. Woodhead Publishing, Oxford, 2015.

[101] S. Lasaulce and H. Tembine. A very short tour of game theory. In *Game Theory and Learning for Wireless Networks*, pages 3–40. Academic Press, Oxford, 2011.

[102] D.L. Craft, T.F. Halabi, H.A. Shih, and T.R. Bortfeld. Approximating convex Pareto surfaces in multiobjective radiotherapy planning. *Medical Physics*, 33(9):3399–3407, 2006.

[103] Y. Yu. Multiobjective decision theory for computational optimization in radiation therapy. *Medical Physics*, 24(9):1445–1454, 1997.

[104] H. Hamacher and K.H. Küfer. Inverse radiation therapy planning — a multiple objective optimization approach. *Discrete Applied Mathematics*, 118(1):145–161, 2002.

[105] K.H. Kufer et al. Intensity-modulated radiotherapy – a large scale multi-criteria programming problem. *OR Spectrum*, 25(2):223–249, 2003.

[106] W. Chen, D. Craft, T.M. Madden, K. Zhang, H.M. Kooy, and G.T. Herman. A fast optimization algorithm for multicriteria intensity modulated proton therapy planning. *Medical Physics*, 37(9):4938–4945, 2010.

[107] D. Craft and M. Monz. Simultaneous navigation of multiple Pareto surfaces, with an application to multicriteria IMRT planning with multiple beam angle configurations. *Medical Physics*, 37(2):736–741, 2010.

[108] D. Craft, D. McQuaid, J. Wala, W. Chen, E. Salari, and T. Bortfeld. Multicriteria VMAT optimization. *Medical Physics*, 39(2):686–696, 2012.

[109] J. Wala, D.L. Craft, J. Paly, and J. Efstathiou. Multi-criteria optimization of IMRT plans in prostate cancer. *International Journal of Radiation Oncology* Biology* Physics*, 81(2):S419–S420, 2011.

[110] J. Wo, D. Craft, T. Bortfeld, and T. Hong. Multicriteria optimization in dose painted intensity modulated radiation therapy treatment planning for anal cancer. *International Journal of Radiation Oncology* Biology* Physics*, 81(2):S815, 2011.

[111] D. Craft, T.S. Hong, H.A. Shih, and T.R. Bortfeld. Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy. *International Journal of Radiation Oncology*Biology*Physics*, 82(1):e83–e90, 2012.

[112] I. Xhaferllari, E. Wong, K. Bzdusek, M. Lock, and J.Z. Chen. Automated IMRT planning with regional optimization using planning scripts. *Journal of Applied Clinical Medical Physics*, 14(1):176–191, 2013.

[113] C. Cotrutz and L. Xing. IMRT dose shaping with regionally variable penalty scheme. *Medical Physics*, 30(4):544–551, 2003.

[114] D. Gintz et al. Initial evaluation of automated treatment planning software. *Journal of Applied Clinical Medical Physics*, 17(3):331–346, 2016.

[115] I. Hazell et al. Automatic planning of head and neck treatment plans. *Journal of Applied Clinical Medical Physics*, 17(1):272–282, 2016.

[116] C.R. Hansen et al. Automatic treatment planning improves the clinical quality of head and neck cancer treatment plans. *Clinical and Translational Radiation Oncology*, 1:2–8, 2016.

[117] C.R. Hansen et al. Automatic treatment planning facilitates fast generation of high-quality treatment plans for esophageal cancer. *Acta Oncologica*, 56(11):1495–1500, 2017.

[118] B. Vanderstraeten, B. Goddeeris, K. Vandecasteele, M. van Eijkeren, C.D. Wagter, and Y. Lievens. Automated instead of manual treatment planning? A plan comparison based on dose-volume statistics and clinical preference. *International Journal of Radiation Oncology*Biology*Physics*, 102(2):443–450, 2018.

[119] S. Breedveld, P.R.M. Storchi, and B.J.M. Heijmen. The equivalence of multi-criteria methods for radiotherapy plan optimization. *Physics in Medicine and Biology*, 54(23):7199–7209, 2009.

[120] S. Breedveld, P.R.M. Storchi, P.W.J. Voet, and B.J.M. Heijmen. iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. *Medical Physics*, 39(2):951–963, 2012.

[121] P.W.J. Voet et al. Toward fully automated multicriterial plan generation: a prospective clinical study. *International Journal of Radiation Oncology*Biology*Physics*, 85(3):866–872, 2013.

[122] A.W.M. Sharfo et al. Comparison of VMAT and IMRT strategies for cervical cancer patients using automated planning. *Radiotherapy and Oncology*, 114(3):395–401, 2015.

[123] A. Nicolae et al. Evaluation of a machine-learning algorithm for treatment planning in prostate low-dose-rate brachytherapy. *International Journal of Radiation Oncology*Biology*Physics*, 97(4):822–829, 2017.

[124] C. Shen et al. Intelligent inverse treatment planning via deep reinforcement learning, a proof-of-principle study in high dose-rate Brachytherapy for cervical cancer. *Physics in Medicine and Biology*, 64(11):115013, 2019.

[125] F. Bachand, A.G. Martin, L. Beaulieu, F. Harel, and E. Vigneault. An eight-year experience of HDR brachytherapy boost for localized prostate cancer: biopsy and PSA outcome. *International Journal of Radiation Oncology*Biology*Physics*, 73(3):679–684, 2009.

[126] M. Fatyga et al. A comparison of HDR brachytherapy and IMRT techniques for dose escalation in prostate cancer: A radiobiological modeling study. *Medical Physics*, 36(9Part1):3995–4006, 2009.

[127] K.A. Sandler et al. Clinical outcomes for patients with Gleason score 10 prostate adenocarcinoma: results from a multi-institutional consortium study. *International Journal of Radiation Oncology*Biology*Physics*, 99(2, Supplement):S132–S133, 2017.

[128] R. Banerjee and M. Kamrava. Brachytherapy in the treatment of cervical cancer: a review. *Int J Womens Health*, 6:555–564, 2014.

[129] V. Laberta. Benefits of automation in radiation oncology. *Oncology Times*, 39(1):8–10, 2017.

[130] R. Marler and J. Arora. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26(6):369–395, 2004.

[131] M. Zarepisheh, A. Uribe-Sanchez, N. Li, X. Jia, and S.B. Jiang. A multicriteria framework with voxel-dependent parameters for radiotherapy treatment plan optimization. *Medical Physics*, 41(4):041705–n/a, 2014.

[132] O. Chernykh. Approximating the Pareto-hull of a convex set by polyhedral sets. *Computational mathematics and mathematical physics*, 35(8):1033–1039, 1995.

[133] H.P. Benson. An outer approximation algorithm for generating all efficient extreme points in the outcome set of a multiple objective linear programming problem. *Journal of Global Optimization*, 13(1):1–24, 1998.

[134] R.S. Solanki, P.A. Appino, and J.L. Cohon. Approximating the noninferior set in multiobjective linear programming problems. *European Journal of Operational Research*, 68(3):356–373, 1993.

[135] G. Rennen, E.R. van Dam, and D. den Hertog. Enhancement of sandwich algorithms for approximating higher-dimensional convex Pareto sets. *INFORMS Journal on Computing*, 23(4):493–517, 2011.

[136] R. Bokrantz and A. Forsgren. An algorithm for approximating convex Pareto surfaces based on dual techniques. *INFORMS Journal on Computing*, 25(2):377–393, 2013.

[137] D. Craft and T. Bortfeld. How many plans are needed in an IMRT multi-objective plan database? *Physics in Medicine and Biology*, 53(11):2785–2796, 2008.

[138] T. Spalke, D. Craft, and T. Bortfeld. Analyzing the main trade-offs in multiobjective radiation therapy treatment planning databases. *Physics in Medicine and Biology*, 54(12):3741–3754, 2009.

[139] S. Cui, P. Després, and L. Beaulieu. An operator-free inverse planning algorithm using a self-adjusting weighting method for high dose rate (HDR) brachytherapy. *Medical Physics*, 44(6):3192–3192, 2017.

[140] P. Edimo, L. Archambault, and L. Beaulieu. A stochastic frontier analysis for enhanced quality control of high-dose-rate brachytherapy treatment plans. *Medical Physics*, 44(6):2996–2996, 2017.

[141] C.B. Barber, D.P. Dobkin, and H. Huhdanpaa. The Quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.*, 22(4):469–483, 1996.

[142] J.O. Deasy. Multiple local minima in radiotherapy optimization problems with dose–volume constraints. *Medical Physics*, 24(7):1157–1161, 1997.

[143] C.V. Guthier, A. Damato, A.N. Viswanathan, J. Hesser, and R.A. Cormack. A fast multi-target inverse treatment planning strategy optimizing dosimetric measures for high-dose-rate (HDR) brachytherapy. *Medical Physics*, 43(6):3808–3808, 2016.

[144] B.W. Stewart and C.P. Wild. *World cancer report 2014*. IARC Press, International Agency for Research on Cancer, 2014.

[145] N.G. Zaorsky et al. The evolution of brachytherapy for prostate cancer. *Nat Rev Urol*, 14(7):415–439, 2017.

[146] K. Teichert, P. Süss, J.I. Serna, M. Monz, K.H. Küfer, and C. Thieke. Comparative analysis of Pareto surfaces in multi-criteria IMRT planning. *Physics in Medicine and Biology*, 56(12):3669–3684, 2011.

[147] V. Chanyavanich, S.K. Das, W.R. Lee, and J.Y. Lo. Knowledge-based IMRT treatment planning for prostate cancer. *Medical Physics*, 38(5):2515–2522, 2011.

[148] D. Good, J. Lo, W.R. Lee, Q.J. Wu, F.F. Yin, and S.K. Das. A knowledge-based approach to improving and homogenizing intensity modulated radiation therapy planning quality among treatment centers: an example application to prostate cancer planning. *International Journal of Radiation Oncology* Biology* Physics*, 87(1):176–181, 2013.

[149] M.E. Babadagli et al. Initial clinical assessment of "center-specific" automated treatment plans for low-dose-rate prostate brachytherapy. *Brachytherapy*, 17(2):476–488, 2018.

[150] C. McIntosh, M. Welch, A. McNiven, D.A. Jaffray, and T.G. Purdie. Fully automated treatment planning for head and neck radiotherapy using a voxel-based dose prediction and dose mimicking method. *Physics in Medicine and Biology*, 62(15):5926–5944, 2017.

[151] M. Mardani, P. Dong, and L. Xing. Deep-learning based prediction of achievable dose for personalizing inverse treatment planning. *International Journal of Radiation Oncology, Biology, Physics*, 2(96):E419–E420, 2016.

[152] M. Monz, K.H. Küfer, T.R. Bortfeld, and C. Thieke. Pareto navigation—algorithmic foundation of interactive multi-criteria IMRT planning. *Physics in Medicine and Biology*, 53(4):985–998, 2008.

[153] H. Ruotsalainen, K. Miettinen, and J.E. Palmgren. Interactive multiobjective optimization for 3D HDR brachytherapy applying IND-NIMBUS. In *New Developments in Multiple Objective and Goal Programming*, pages 117–131. Springer, 2010.

[154] Y. Zhou et al. Automated high-dose rate brachytherapy treatment planning for a single-channel vaginal cylinder applicator. *Physics in Medicine and Biology*, 62(11):4361–4374, 2017.

[155] A. Fedorov et al. 3D Slicer as an image computing platform for the quantitative imaging network. *Magnetic Resonance Imaging*, 30(9):1323–41, 2012.

[156] L. Dagum and R. Menon. OpenMP: an industry standard API for shared-memory programming. *IEEE Computational Science and Engineering*, 5(1):46–55, 1998.

[157] B.E. Nelms et al. Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. *Practical Radiation Oncology*, 2(4):296–305, 2012.

[158] B. Wu et al. Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Medical Physics*, 36(12):5497–5505, 2009.

[159] B. Wu et al. Data-driven approach to generating achievable dose–volume histogram objectives in intensity-modulated radiotherapy planning. *International Journal of Radiation Oncology*Biology*Physics*, 79(4):1241–1247, 2011.

[160] S.F. Petit et al. Increased organ sparing using shape-based treatment plan optimization for intensity modulated radiation therapy of pancreatic adenocarcinoma. *Radiotherapy and Oncology*, 102(1):38–44, 2012.

[161] M. van der Meer, T. Alderliesten, B. Pieters, A. Bel, Y. Niatsetski, and P. Bosman. Better and faster catheter position optimization in HDR brachytherapy for prostate cancer using multi-objective real-valued GOMEA. In *GECCO 2018-Proceedings of the 2018 Genetic and Evolutionary Computation Conference*, pages 1387–1394, 2018.

[162] S. Cui, P. Després, and L. Beaulieu. A multi-criteria optimization approach for HDR prostate brachytherapy: I. Pareto surface approximation. *Physics in Medicine and Biology*, 63(20):205004, 2018.

[163] G. Pratx and L. Xing. GPU computing in medical physics: A review. *Medical Physics*, 38(5):2685–2697, 2011.

[164] P. Després and X. Jia. A review of GPU-based medical image reconstruction. *Physica Medica*, 42:76–92, 2017.

[165] P. Edimo, A. Kroshko, L. Beaulieu, and L. Archambault. A stochastic frontier analysis for enhanced treatment quality of high-dose-rate brachytherapy plans. *Physics in Medicine and Biology*, 64(6):065012, 2019.

[166] C. Men et al. GPU-based ultrafast IMRT plan optimization. *Physics in Medicine and Biology*, 54(21):6565–6573, 2009.

[167] D.C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1):503–528, 1989.

[168] J. Wetzl and O. Taubmann. CudaLBFGS. `https://github.com/jwetzl/CudaLBFGS/`, 2013. Accessed: 2019-3-4.

[169] J. Wetzl, O. Taubmann, S. Haase, T. Kohler, M. Kraus, and J. Hornegger. GPU-accelerated time-of-flight super-resolution for image-guided surgery. In *Bildverarbeitung fur die Medizin 2013*, pages 21–26, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[170] E. B. R. Thomadsen, M.J. Rivard, and W.M. Butler. *Brachytherapy Physics, 2nd Edition*. Medical Physics Pub Corp, 2005.

[171] E. Poulin, C.A.C. Fekete, M. Létourneau, A. Fenster, J. Pouliot, and L. Beaulieu. Adaptation of the CVT algorithm for catheter optimization in high dose rate brachytherapy. *Medical Physics*, 40(11):111724, 2013.

[172] R. Nath et al. AAPM recommendations on dose prescription and reporting methods for permanent interstitial brachytherapy for prostate cancer: Report of Task Group 137. *Medical Physics*, 36(11):5310–5322, 2009.

[173] C. Chargari, E.V. Limbergen, U. Mahantshetty, E. Deutsch, and C. Haie-Méder. Radiobiology of brachytherapy: The historical view based on linear quadratic model and perspectives for optimization. *Cancer/Radiothérapie*, 22(4):312–318, 2018.