A Monte Carlo Inverse Treatment Planning Algorithm for Trajectory-based Volumetric Modulated Arc Therapy with Applications in Stereotactic Radiosurgery, Total Body Irradiation and Patient-specific Quality Assurance

by

Shiqin Su

B.A.Sc., Shanghai University, 2012
M.Sc., Dalhousie University, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Doctor of Philosophy in THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES (Physics)

THE UNIVERSITY OF BRITISH COLUMBIA (Vancouver)

December 2019
© Shiqin Su, 2019
The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

A Monte Carlo inverse treatment planning algorithm for trajectory-based volumetric modulated arc therapy with applications in stereotactic radiosurgery, total body irradiation and patient-specific quality assurance

submitted by Shiqin Su in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics

Examinig Committee:
Tony Popescu, Physics
Supervisor
Cheryl Duzenli, Physics
Supervisory Committee Member
Stefan Reinsberg, Physics
Supervisory Committee Member
Jeremy S Heyl, Physics & Astronomy
University Examiner
Urs Hafeli, Pharmaceutical Sciences
University Examiner

Additional Supervisory Committee Members:
Alex Mackey, Physics
Supervisory Committee Member
Abstract

The main objective of this thesis is to present a full Monte Carlo (MC)-based inverse treatment planning method for trajectory-based volumetric modulated arc therapy (TVMAT). TVMAT uses continuous and simultaneous gantry and couch rotation to reduce the radiation exposure of normal tissues. However, commercial treatment planning systems for conventional medical linear accelerators do not provide dose calculation of such a beam trajectory. It has been shown that a full MC-based optimization can reduce the optimization convergence errors. Previously published approaches to MC-based optimization have not been clinically implemented, and none has been proposed for VMAT or TVMAT so far. In this work, we developed a method that reflects the dynamic multi-leaf collimator (MLC) and gantry-couch trajectory of the actual beam delivery at all stages of the optimization. Dose optimization was performed in a single MC simulation, thereby greatly reducing computation time. We select the initial trajectory (i.e., the range of the gantry, collimator and couch angles) and the initial set of leaf positions, corresponding to a dynamic beam conformal to the target. The MC simulation starts from a phase space scored at the top of the MLC module and uses a beam source that allows simulations of complex continuous beam delivery. We modified a general-purpose MC code system in order to generate four-dimensional dose files that score individual, time-stamped, energy deposition events in the patient anatomy. Consequently, a relation is established between the space and time coordinates of source particles in the phase space and their contribution to energy deposition. A MC-based direct aperture optimization, with a dose-volume constraint based quadratic objective function, is performed using an in-house code, taking into account the continuous movement of the MLC, gantry and couch between adjacent control points. This
method is also applicable to beam delivery with either fixed couch position for VMAT or couch translation for long-field treatments. It is shown that this novel treatment planning algorithm is capable of generating plans that are generally of higher quality than those generated by standard planning systems.
Lay Summary

Radiation therapy uses high-energy radiation beam to kill cancer cells. During the treatment, the patient usually lies on a static treatment couch. This work presents two new treatment modalities, known as trajectory-based treatments, where the couch moves continuously. While couch rotation can avoid the normal tissues in the beam trajectory, couch longitudinal translation can facilitate the treatment of large fields, such as for the entire patient body. Radiation therapy relies on the treatment planning of beam delivery. However, currently, there are no available planning systems for trajectory-based treatments. This work introduces a novel algorithm for treatment planning, which can accurately model the proposed treatment modalities. It is shown that this method has the capability to achieve plans that are equivalent, or superior, to those generated by standard planning systems with static couch, within a comparable time.
Preface

This thesis is completed by the author, Shiqin Su, at BC Cancer – Vancouver. I conducted all the design and testing of the algorithms and the techniques. The validation of the Monte Carlo modelling of the dynamic trajectory motion used in Chapter 3 and Chapter 4 was performed by a previous UBC Ph.D. student, Tony Teke, using radiochromic film measurement. Chapter 5 represents a continuation of work done by a previous UBC M.A.Sc student, Julio Lobo.

Preliminary versions of Chapter 3 have been accepted for presentations:

Su S, Popescu IA. A particle-based Monte Carlo algorithm for VMAT optimization. World Congress on Medical Physics and Biomedical Engineering 2018.


A preliminary version of Chapter 5 has been accepted for presentation:


Publications are in preparation, which will build on the results presented in Chapter 3 and Chapter 5.
Table of Contents

Abstract ......................................................................................................................................... iii

Lay Summary ................................................................................................................................... v

Preface ........................................................................................................................................... vi

Table of Contents ........................................................................................................................ vii

List of Tables ................................................................................................................................ xii

List of Figures ................................................................................................................................ xiii

List of Abbreviations .................................................................................................................. xxi

Acknowledgements ....................................................................................................................... xxv

Chapter 1: Introduction .................................................................................................................. 1

1.1 Cancer Treatment with Radiation .............................................................................................. 1
   1.1.1 External Beam Radiation Therapy ...................................................................................... 2
   1.1.2 Beam Generation in a Linear Accelerator ........................................................................ 3
   1.1.3 Treatment Planning ......................................................................................................... 5
   1.1.4 Dose Calculation Algorithm in Radiation Therapy ........................................................ 8
      1.1.4.1 Pencil Beam Algorithm (PBA) .................................................................................. 10
      1.1.4.2 Anisotropic Analytical Algorithm (AAA) .............................................................. 11
      1.1.4.3 Collapsed Cone Convolution (CCC) .................................................................... 13
      1.1.4.4 Acuros XB (AXB) ................................................................................................ 14
   1.2 Inverse Treatment Planning .................................................................................................. 15
3.1 Introduction .................................................................................................................................. 51

3.2 Methods ...................................................................................................................................... 54

3.2.1 Reading and Writing 4D Dose Data ...................................................................................... 55

3.2.2 Dose-volume Objective Function .......................................................................................... 56

3.2.3 MC-based Inverse Planning Algorithm .................................................................................. 57

3.2.3.1 Linear Sampling of MLC Leaf Changes ........................................................................ 58

3.2.3.2 MU Weight Changes........................................................................................................ 60

3.2.3.3 Progressive Sampling of Gantry and MLC Positions ..................................................... 61

3.3 Results and Discussions ............................................................................................................ 63

3.3.1 Head and Neck Cancer ---- Standard Coplanar VMAT Example .................................. 64

3.3.1.1 Principle of MC-VMAT .................................................................................................. 64

3.3.1.2 MC-VMAT Re-plan ........................................................................................................ 69

3.3.2 Brain Arteriovenous Malformations (AVM) ---- Non-coplanar VMAT Example .......... 73

3.3.3 Cranial Metastases SRT ---- TVMAT Example ....................................................................... 77

3.4 Conclusions ................................................................................................................................. 83

Chapter 4: Linac-based Helical Intensity-Modulated Total Body Irradiation ......................... 85

4.1 Introduction .................................................................................................................................. 85

4.2 Methods ...................................................................................................................................... 88

4.2.1 Beam Geometry and Patient Setup ...................................................................................... 88

4.2.2 Dose Calculations and Optimization .................................................................................. 91

4.2.3 Dose Rate and Treatment Time .......................................................................................... 92

4.3 Results and Discussions ............................................................................................................ 94

4.3.1 A Total Body Irradiation (TBI) Example ............................................................................. 94
Chapter 5: Efficient, All-In-One, Monte Carlo Simulations of Transit EPID Cine-mode
Dose Distributions for Patient-Specific VMAT Quality Assurance

5.1 Introduction

5.2 Methods

5.2.1 Description of EPID Simulation

5.2.2 Simulation of Cine-mode EPID

5.2.3 MC Prediction of Individual EPID Cine-image

5.2.4 Statistical Considerations for MC-EPID

5.3 Results and Discussions

5.3.1 Faster Simulation Using Slab Phantom

5.3.2 EPID Images Transiting Anthropomorphic Phantom

5.3.3 MC-EPID Prediction for Prostate Cancer Example

5.3.4 MC-EPID Prediction for Brain Cancer Example

5.4 Conclusions

Chapter 6: Conclusions and Future Work

6.1 Conclusions

6.2 Future Work

Bibliography

Appendices

Appendix A Preparatory Steps Involved in MC-based Optimization

A.1 Initial Treatment Parameters and Field Setup
A.2 Modelling of Linear Accelerator and Patient ......................................................... 146
A.3 Dose Deposition in ROI ......................................................................................... 148
A.4 Optimization Constraints ...................................................................................... 151
A.5 Statistical Considerations for MC-based Optimization .......................................... 151
A.6 Optimization Time ............................................................................................... 152

Appendix B Additional Examples of MC-based Optimization ......................................... 154
B.1 Dose Calculation in Inhomogeneity ..................................................................... 154
B.2 Additional Head and Neck Cancer/Brain AVM Examples ..................................... 157
B.3 A Lung Cancer Example ....................................................................................... 159
B.4 Prostate Cancer Examples .................................................................................... 163
B.5 A Cranio-spinal Radiation Therapy (CSRT) Example .......................................... 167
List of Tables

Table 1.1  Typical HU ranges of materials.................................................................9
Table 3.1  Optimization objectives for a head and neck cancer example..................65
Table 3.2  Beam trajectory parameters of the non-coplanar MC-VMAT plan, for a brain
AVM example..............................................................................................................73
Table 5.1  The first and last control points of each arc written in the egsinp file, for
patient EPID dose calculation, for the phantom example........................................112
Table 5.2  The properties of the control points in EPID_egsinp, for EPID dose
calculation, for the phantom example........................................................................114
List of Figures

Figure 1.1  Representation of TCP (curve A) and NTCP (curve B) ............................................ 2
Figure 1.2  A diagram of typical components of medical linear accelerator .............................. 3
Figure 1.3  The dose conformity is defined using TV, PIV and TVPIV ....................................... 7
Figure 1.4  The dose homogeneity index is defined in a DVH, where D$_{5\%}$ and D$_{95\%}$ represent the doses to 5% and 95% of the PTV, respectively ....................................................... 8
Figure 2.1  Illustration of the components modules of a typical Monte Carlo modelled linear accelerator head in photon beam mode ................................................................. 40
Figure 3.1  The workflow of MC-based optimization for TVMAT ............................................. 54
Figure 3.2  An illustration of the MC particle removing. The left figure shows the initial conformal plan where all the particle depositing energy are recorded. An attempted leaf change is made in the right figure (dash line of the leaf). The energy contributions of the particles blocked by the new leaf position are removed ........................................................................ 60
Figure 3.3  The optimization time varies from 5–90 minutes for eleven test cases .................... 64
Figure 3.4  DVH comparison between MC and Eclipse AAA dose calculation, for a head and neck cancer example. The MC simulation of the same VMAT plan indicates overestimation of the PTV dose and coverage ............................................................... 66
Figure 3.5  Procedure of the particle removing at an arbitrary segment. The left figure shows the particle distribution of the initial conformal plan while the right figure shows the particle distribution of the plan with optimized MLC leaf positions. The particles blocked by the optimized leaves are removed and replaced with those in background plan ........................................ 67
Figure 3.6  DVH comparison between the final MC calculation (solid lines) and the MC-based optimization (dash lines), for a head and neck cancer example. The ▶ and ◄ marks indicate the upper and lower optimization objectives ................................................................. 68

Figure 3.7  3D gamma map between DVH comparison between the final MC calculation and the MC-based optimization, for a head and neck cancer example. The dose mask is 30% of the prescription dose. The criteria are 3% dose difference and 3 mm DTA. The pass (γ ≤ 1) rate = 89.3% .......................................................................................................................... 69

Figure 3.8  A log plot shows the cost decreasing with the optimization iteration, for a head and neck cancer example. The first ▼ symbol represents the initial cost value and the following ▼ symbols represent the cost at the end of each iteration .............................................................................. 70

Figure 3.9  DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a head and neck cancer example. The ▶ and ◄ marks indicate the upper and lower optimization objectives ......................................................................................... 71

Figure 3.10  3D dose distribution of the MC-VMAT plan, for a head and neck cancer example. The white lines contour the PTV. The dose mask is 30% of the prescription dose .................................. 72

Figure 3.12  DVH comparison between the base plan (solid lines) and the non-coplanar MC-VMAT plan (dash lines), for a brain AVM example. The ▶ and ◄ marks indicate the upper and lower optimization objectives .............................................................................. 75

Figure 3.13  3D dose distribution of the non-coplanar MC-VMAT plan, for a brain AVM example. The white lines contour the PTV. The dose mask is 30% of the prescription dose ....... 76

Figure 3.14  3D gamma map between the final MC calculation and the MC-based optimization, for a brain AVM example. The dose mask is 30% of the prescription dose. The criteria are 3% dose difference and 3 mm DTA. The pass (γ ≤ 1) rate = 99.8% ......... 77
Figure 3.15  The gantry-couch trajectory of the MC-TVMAT plan, for a cranial metastases SRT example (top figure). The blue lines represent the trajectory of the gantry head. The yellow lines represent the incident direction of the photon beam. The white dots contour the patient body and the red dots contour the PTV. The couch angles vs. gantry angles are shown in the bottom figure, where the instant dose rates at each control point are also plotted (bottom figure).

Figure 3.16  A log plot shows the cost decreasing with the optimization iteration, for a cranial metastases SRT example. The first ▽ symbol represents the initial cost value and the following ▽ symbols represent the cost at the end of each iteration.

Figure 3.17  DVH comparison between the MC-TVMAT plan (solid lines) and the MC-calculated non-coplanar AAA-VMAT plan (dash lines), for a cranial metastases SRT example. The ► and ◄ marks indicate the upper and lower optimization objectives.

Figure 3.18  3D dose distribution of the MC-TVMAT plan, for a cranial metastases SRT example. The white lines contour the PTV. The dose mask is 30% of the prescription dose.

Figure 3.19  DVH comparison between the MC-based optimization (solid lines) and the final MC calculation (dash lines), for a cranial metastases SRT example. The ► and ◄ marks indicate the upper and lower optimization objectives.

Figure 3.20  3D gamma map between the final MC calculation and the MC-based optimization, for a cranial metastases SRT example, for a cranial metastases SRT example. The dose mask is 30% of the prescription dose. The criteria are 3% dose difference and 3 mm DTA. The pass (γ ≤ 1) rate = 95.2%.
Figure 4.1  Plot of double helix HTBI trajectory of isocentre position vs. gantry angle. The top, mid and bottom panels show the HTBI trajectory delivered up to the 4th, 11th and 14th field, respectively .................................................................................................................................90

Figure 4.2  Monte Carlo calculated DVH (top panel) and dose distribution (coronal view, bottom left; and sagittal view, bottom right) of HTBI (not optimized) plan using 6 MV photon beam. ..................................................................................................................................................94

Figure 4.3  Monte Carlo calculated DVH (top panel) and dose distribution (coronal view, bottom left; and sagittal view, bottom right) of HTBI (not optimized) plan using 4 MV photon beam. ..................................................................................................................................................96

Figure 4.4  Monte Carlo calculated DVH (top panel) and dose distribution (coronal view, bottom left; and sagittal view, bottom right) of HTBI (not optimized) plan using 10 MV photon beam. ..............................................................................................................................................97

Figure 4.5  Plot shows the dose rates of four randomly selected voxels in the patient phantom. The instant dose rates at each control point are presented as treatment time vs. dose rate and the average dose rates over the entire treatment are shown at the top-right corner. ..............................................................................................................................................98

Figure 4.6  DVH comparison between the HTBI conformal plan (solid lines) and the modulated HTBI plan (dash lines), using double-helix trajectory. The ► and ◄ marks indicate the upper and lower optimization objectives ..............................................................................................................................................99

Figure 5.1  The cross section of the EPID phantom used in MC simulation, where the eleventh slice is modelled as the effective detector layer and selected as ROI .........................................................107

Figure 5.2  A high-resolution phantom with 256×20×192 voxels (left) .................................................................110
Figure 5.3  The MC simulated dose of Einstein face image using a voxelized phantom (left) vs. a slab phantom (right). .................................................................................................................................111

Figure 5.4  The x (left) and y (right) dose profile of MC simulated dose of Einstein face image using a voxelized phantom (red lines) vs. a slab phantom (blue lines). ........................................111

Figure 5.5  A two-arc VMAT verification plan created on an anthropomorphic phantom for a head and neck patient. The plan is simulated as a single field, using source 20 of DOSXYZnrc. ...............................................................................................................................113

Figure 5.6  An example of the *egsphant* created using CT images of anthropomorphic phantom. ........................................................................................................................................113

Figure 5.7  The measured EPID signal (left) vs. the MC-predicted EPID dose (right) integrated to a random selected frame segment. .....................................................................................................................115

Figure 5.8  The measured EPID signal (left) vs. the MC-predicted EPID dose (right) of the entire treatment. .................................................................................................................................................115

Figure 5.10  Comparison between Eclipse (left) and Monte Carlo (right) dose distributions, of a VMAT prostate plan. .................................................................................................................................................117

Figure 5.11  DVH comparison between Monte Carlo and Eclipse (AAA) dose distributions, for selected structures, of a VMAT prostate plan. .................................................................................................................................................117

Figure 5.12  The transit EPID image (left panel) and the MC-predicted EPID dose (right panel) for two randomly selected frames (14 and 32, of 178), of a VMAT prostate plan. ..........118

Figure 5.13  Statistical uncertainty of the MC-EPID transit dose for randomly selected control point (32 of 178). .................................................................................................................................................119

Figure 5.14  Best-fit line of the measured EPID signals *vs.* MC-predicted image dose for randomly selected control point (32 of 178). .................................................................................................................................................120
Figure 5.15  Residual of best-fit line plotted in Figure 5.14. .................................................................121

Figure 5.16  The MC-predicted EPID transit dose for the original (right) and the 2 cm shifted (left) patient position of a randomly selected frame (12 of 147). .................................................................122

Figure 5.17  Gamma evaluation (3% dose and 3mm DTA) between the MC-predicted EPID transit dose for the original (right) and the 2 cm shifted (left) patient position of a random selected frame (12 of 147). The pass ($\gamma \leq 1$) rate = 15.0%. ............................................................................122

Figure 5.18  The MC-predicted EPID transit dose for the original (right) and the 2 cm shifted (left) patient position summed up for all the frames. .................................................................123

Figure 5.19  Gamma evaluation (3% dose and 3mm DTA) between the MC-predicted EPID transit dose for the original (right) and the 2 cm shifted (left) patient position summed up for all the frames. The pass ($\gamma \leq 1$) rate = 12.1%. ............................................................................123

Figure 5.20  Comparison between the MC simulated forward patient dose (top) and backscatter dose from the EPID (bottom). .................................................................................................124

Figure A.1  ROI is defined as the region including PTV, OARs and dose fall-off region. The dose fall-off region is a shell of voxels surrounding the PTV. .................................................................149

Figure A.2  Template edepheader file. The four columns represent the x and y coordinate of incident particles at the z-position of phsp, the MU index of incident particles and the number of ROI voxels the particle deposit energy to. .................................................................150

Figure A.3  Template edepdat file contains energy deposition of the first incident particle written in Figure A.2. The first column represents the index of the ROI voxel and the second column represents the value of energy deposition. .............................................................................150

Figure B.1  The PDD curve of 15 MV and 6 MV photon beam with 10×10 and 4×4 cm² field sizes incident to 30×30×30 cm³ water-lung-water phantom. The blue curves and the
red curves represent the dose by MC simulation and the AAA model, respectively. The green curves represent MC calculation of the beam with the same energy and field size incident on the 30×30×30 cm³ full water phantom.

Figure B.2  The PDD curve of 15 MV and 6 MV photon beam with 10×10 and 4×4 cm² field sizes incident to 30×30×30 cm³ water-bone-water phantom. The blue curves and the red curves represent the dose by MC simulation and the AAA model, respectively. The green curves represent MC calculation of the beam with the same energy and field size incident on the 30×30×30 cm³ full water phantom.

Figure B.3  DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a head and neck cancer example (case b). The ► and ◄ marks indicate the upper and lower optimization objectives.

Figure B.4  DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a brain AVM example (case c). The ► and ◄ marks indicate the upper and lower optimization objectives.

Figure B.5  DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a lung cancer example (case d). The dose is normalized such that the PTV mean dose equals to the prescription dose. The ► and ◄ marks indicate the upper and lower optimization objectives.

Figure B.6  DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a lung cancer example (case d). The dose is normalized such that 100% of the PTV receives 90% isodose. The ► and ◄ marks indicate the upper and lower optimization objectives.
Figure B.7  3D dose distribution of the MC-VMAT plan, for a lung cancer example (case d). The white lines contour the PTV. The dose mask is 30% of the prescription dose. ..........162

Figure B.8  3D gamma map between DVH comparison between the final MC calculation and the MC-based optimization, for a lung cancer example. The dose mask is 30% of the prescription dose. The criteria are 3% dose difference and 3 mm DTA. The pass ($\gamma \leq 1$) rate = 97.4%. ...............................................................163

Figure B.9  DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a prostate cancer example (case e). The $\triangleright$ and ◄ marks indicate the upper and lower optimization objectives. ..............................................................................................164

Figure B.10  DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a prostate cancer example (case f). The $\triangleright$ and ◄ marks indicate the upper and lower optimization objectives. ..............................................................................................165

Figure B.11  3D dose distribution of the MC-VMAT plan, for a prostate cancer example (case e). The white lines contour the PTV. The dose mask is 30% of the prescription dose. .....166

Figure B.12  3D gamma map between DVH comparison between the final MC calculation and the MC-based optimization, for a prostate cancer example (case e). The dose mask is 30% of the prescription dose. The criteria are 3% dose difference and 3 mm DTA. The pass ($\gamma \leq 1$) rate = 91.6%. ....................................................................................167

Figure B.13  Monte Carlo calculated DVH (top panel) and dose distribution (coronal view, bottom left; and sagittal view, bottom right) of a Helical CSRT (optimized) plan using 6 MV photon beam. ........................................................................................................................................168
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>Two-dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>Three-dimensional</td>
</tr>
<tr>
<td>4D</td>
<td>Four-dimensional</td>
</tr>
<tr>
<td>AAA</td>
<td>Anisotropic analytical algorithm</td>
</tr>
<tr>
<td>API</td>
<td>Application programming interface</td>
</tr>
<tr>
<td>a-Si</td>
<td>Amorphous silicon</td>
</tr>
<tr>
<td>AVM</td>
<td>Arteriovenous malformations</td>
</tr>
<tr>
<td>AXB</td>
<td>Acuros XB</td>
</tr>
<tr>
<td>BEV</td>
<td>Beam's eye view</td>
</tr>
<tr>
<td>BTE</td>
<td>Boltzmann transport equation</td>
</tr>
<tr>
<td>CCC</td>
<td>Collapsed cone convolution</td>
</tr>
<tr>
<td>CI</td>
<td>Conformity index</td>
</tr>
<tr>
<td>CM</td>
<td>Component modules</td>
</tr>
<tr>
<td>CPU</td>
<td>Central processing unit</td>
</tr>
<tr>
<td>CSRT</td>
<td>Cranio-spinal radiation therapy</td>
</tr>
<tr>
<td>CT</td>
<td>Computed tomography</td>
</tr>
<tr>
<td>DAO</td>
<td>Direct aperture optimization</td>
</tr>
<tr>
<td>DICOM</td>
<td>Digital image and communications in medicine</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTA</td>
<td>Distance-to-agreement</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>DVH</td>
<td>Dose-volume histograms</td>
</tr>
<tr>
<td>ECUT</td>
<td>Electron cutoff energy</td>
</tr>
<tr>
<td>EPID</td>
<td>Electronic portal imaging devices</td>
</tr>
<tr>
<td>FFF</td>
<td>Flattening filter free</td>
</tr>
<tr>
<td>FO</td>
<td>Fall-off</td>
</tr>
<tr>
<td>FMO</td>
<td>Fluence map optimization</td>
</tr>
<tr>
<td>GPU</td>
<td>Graphics processing unit</td>
</tr>
<tr>
<td>GTV</td>
<td>Gross tumour volume</td>
</tr>
<tr>
<td>HI</td>
<td>Homogeneity index</td>
</tr>
<tr>
<td>HT</td>
<td>Helical tomotherapy</td>
</tr>
<tr>
<td>HTBI</td>
<td>Helical total body irradiation</td>
</tr>
<tr>
<td>HU</td>
<td>Hounsfield units</td>
</tr>
<tr>
<td>IAEA</td>
<td>International atomic energy agency</td>
</tr>
<tr>
<td>ICRP</td>
<td>International commission on radiological protection</td>
</tr>
<tr>
<td>ICRU</td>
<td>International commission on radiation units and measurements</td>
</tr>
<tr>
<td>IGRT</td>
<td>Image-guided radiation therapy</td>
</tr>
<tr>
<td>IMRT</td>
<td>Intensity-modulated radiotherapy</td>
</tr>
<tr>
<td>kV</td>
<td>Kilovoltage</td>
</tr>
<tr>
<td>LBTE</td>
<td>Linear Boltzmann transport equation</td>
</tr>
<tr>
<td>linac</td>
<td>Linear accelerator</td>
</tr>
<tr>
<td>MC</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>MLC</td>
<td>Multi-leaf collimator</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>MU</td>
<td>Monitor unit</td>
</tr>
<tr>
<td>MV</td>
<td>Megavoltage</td>
</tr>
<tr>
<td>NTCP</td>
<td>Normal tissue complication probability</td>
</tr>
<tr>
<td>OAR</td>
<td>Organ at risk</td>
</tr>
<tr>
<td>PBA</td>
<td>Pencil beam algorithm</td>
</tr>
<tr>
<td>PCUT</td>
<td>Photon cutoff energy</td>
</tr>
<tr>
<td>PD</td>
<td>Probability distribution</td>
</tr>
<tr>
<td>PDD</td>
<td>Percentage depth dose</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>phsp</td>
<td>Phase space</td>
</tr>
<tr>
<td>PIV</td>
<td>Prescription isodose volume</td>
</tr>
<tr>
<td>PTV</td>
<td>Planning target volume</td>
</tr>
<tr>
<td>QA</td>
<td>Quality assurance</td>
</tr>
<tr>
<td>ROI</td>
<td>Region-of-interest</td>
</tr>
<tr>
<td>RTOG</td>
<td>Radiation therapy oncology group</td>
</tr>
<tr>
<td>SA</td>
<td>Simulated annealing</td>
</tr>
<tr>
<td>SAD</td>
<td>Source-to-axis distance</td>
</tr>
<tr>
<td>SBRT</td>
<td>Stereotactic body radiotherapy</td>
</tr>
<tr>
<td>SID</td>
<td>Source-to-imager distance</td>
</tr>
<tr>
<td>SRS</td>
<td>Stereotactic radiosurgery</td>
</tr>
<tr>
<td>SRT</td>
<td>Stereotactic radiotherapy</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>SSD</td>
<td>Source-to-surface distance</td>
</tr>
<tr>
<td>TBI</td>
<td>Total body irradiation</td>
</tr>
<tr>
<td>TCP</td>
<td>Tumour control probability</td>
</tr>
<tr>
<td>TERMA</td>
<td>Total energy released per unit mass</td>
</tr>
<tr>
<td>TLD</td>
<td>Thermoluminescent dosimeter</td>
</tr>
<tr>
<td>TPS</td>
<td>Treatment planning system</td>
</tr>
<tr>
<td>TVMAT</td>
<td>Trajectory-based volumetric modulated arc therapy</td>
</tr>
<tr>
<td>VMAT</td>
<td>Volumetric modulated arc therapy</td>
</tr>
<tr>
<td>XML</td>
<td>Extensible markup language</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to first and foremost thank my supervisor, Tony Popescu, for the patient guidance, encouragement and advice he has provided throughout my time as his graduate student. I have been extremely lucky to have a supervisor who created opportunities for me, and who responded to my questions and queries so promptly. He has become a valued friend and collaborator, who helped me work through my weaknesses and taught me many things about the Monte Carlo techniques in radiation therapy.

I would like to thank the Department of Physics & Astronomy, UBC, not only for providing the opportunity that allowed me to undertake this research, but also for giving me the opportunity to attend conferences and meet so many valuable people. I would also like to thank all the members of staff at BC Cancer – Vancouver who helped me. Among all of the physicists, physics assistants, students, and staff at BC Cancer, I was never deprived of mentorship, technical support as my degree moved along, for which I am boundlessly grateful. In particular, I would like to thank Cheryl Duzenli for relentless support of medical physics research and for the training of the next generation of medical physicists. I would like to thank Parmveer Atwal and Julio Lobo for making all the suggestions and exploring new ideas for the projects. I would like to thank Karl Otto for kindly sharing his prototype of VMAT code. I would like to thank Tony Teke for his validation of the Monte Carlo modelling of the dynamic trajectory motion. I would like to thank Byron Wilson and Ermias Gete for sharing their method of developing the optimized trajectory for TVMAT. I would also like to thank Iulian Badragan for his help in creating the best-fit lines and residuals for Monte Carlo-predicted EPID images.
Finally, I must express my gratitude to my family for their love and for talking me through the various hardships that I encountered while I completed this degree. I would also like to thank the friends that I have made in BC for all the good times we spent together.
Chapter 1: Introduction

1.1 Cancer Treatment with Radiation

Cancer is a class of diseases characterized by unregulated cell growth that harms the body. These damaged cells divide uncontrollably and form masses of tissue called tumours. Tumours can grow and interfere with the digestive, nervous, and circulatory systems or they can release hormones that alter body function. While tumours that do not invade neighboring tissues and demonstrate limited growth are generally considered benign, malignant cancer is any type of cancer that spreads to other parts of the body. There are over 200 different types of cancer, and each is classified by the type of cell that is initially affected. According to Canadian Cancer Statistics 2018 [1], an estimated 206,200 new cancer diagnoses and 80,800 deaths from cancer occurred in Canada in 2017. Lung, colorectal, breast, and prostate cancer are the four most commonly diagnosed cancers.

Surgery, chemotherapy and radiation therapy are all proven to effectively treat cancer, and thus extend life or improve quality of life. The type of treatment or the order of treatment will be different for individual patients, depending on the location of the tumour and the stage of the disease at diagnosis. Among all the cancer treatment, 50% of patients may receive radiation therapy before, during, or after surgery [2]. Radiation therapy uses high-energy radiation to kill cancer cells by damaging their deoxyribonucleic acid (DNA), thereby killing them or preventing their replication. Types of radiation used for cancer treatment include x-rays, gamma rays and charged particles. However, radiation therapy can damage normal cells as well as cancer cells. Therefore, treatment must be planned and delivered accurately with regard to both spatial and
dosimetric accuracy in order to minimize side effects. The principle of radiation therapy is usually illustrated by plotting two curves (Figure 1.1), one for the tumour control probability (TCP) (curve A) and the other for the normal tissue complication probability (NTCP) (curve B), where the optimum choice in the treatment of a given tumour is to maximize the TCP and simultaneously minimize the NTCP. The seriousness of the damage to cancerous tissues is correlated to the amount of energy absorbed per unit mass of tissue or absorbed dose. The international standard unit for dose is Gy, where 1 Gy equals to 1 J/kg. The tumour lethal dose required to achieve tumour control depends on the volume and localization of the target, radiosensitivity as well as type of tumour, which also determines the type of treatment chosen.

Figure 1.1  Representation of TCP (curve A) and NTCP (curve B).

1.1.1 External Beam Radiation Therapy

The radiation used for cancer treatment may be produced by a machine outside the body (i.e., external beam radiation therapy), or it may be emitted by radioactive material placed in the body near tumour cells or injected into the bloodstream (i.e., brachytherapy).

In external beam radiation therapy, the patient sits or lies on a couch and an external source of radiation is directed to the tumour volume from outside the body. Photon and electron
beams are by far the most widely used sources for external beam radiotherapy, however a comparatively small number of cancer centres offer radiation therapy using proton sources. Photon and electron radiation therapy are normally delivered by a clinical linear accelerator (linac).

1.1.2 Beam Generation in a Linear Accelerator

The linac is the most commonly used treatment machine for external beam therapy, which uses high-frequency electromagnetic waves to accelerate electrons through a waveguide. The high-energy electron beam itself can be used for treatment, usually of superficial tumours, or it can be made to strike a target to produce x-rays for treating deep tumours. The main operating components of a medical linac are shown in Figure 1.2: 1) injection system; 2) microwave generator; 3) accelerating waveguide; 4) auxiliary system; 5) beam transport system and 6) beam collimation and monitoring system (treatment head). A typical modern high-energy linac provides multiple photon energies (6 and 18 MV) and several electron energies (6, 9, 12 and 16 MeV) for clinical uses.

Figure 1.2 A diagram of typical components of medical linear accelerator.
The injection system is the source of electrons, also referred to as electron gun, containing a heated filament cathode, a perforated grounded anode and a grid. Electrons are emitted from the heated cathode, focused into a pencil beam by a curved focusing electrode and accelerated towards the perforated anode through which they drift to enter the accelerating waveguide. The timing of the injection of electrons into the accelerating waveguide is controlled by voltage pulses, which are applied to the grid and must be synchronized with the pulses applied to the microwave generator (i.e., radiofrequency system). The radiofrequency system produces the microwave radiation used in the accelerating waveguide to accelerate electrons to the desired kinetic energy.

The waveguides are evacuated structures of rectangular or circular cross section used in transmission of microwaves. The propagation of microwaves through waveguides is governed by Maxwell’s equations and boundary conditions at the metallic walls. The accelerating waveguide is evacuated to allow free propagation of electrons. Two basic waveguide designs exist: travelling wave and standing wave. The microwave electric field in a standing wave design changes its amplitude with time rather than its position (as in a travelling design). In the standing wave waveguide structure, every second cavity carries no electric field and thus produces no energy gain for the electron. Therefore, these cavities, serving only as coupling cavities, can be moved out to the side of waveguide structure and effectively shorten the accelerating waveguide by 50%. This configuration has been developed and is used in the Varian linac (Varian Medical Systems, Palo Alto, CA) machine.

As the high-energy electrons emerge from the exit window of the waveguide, they are in the form of a pencil beam of about 3 mm in diameter. In a high-energy linac, the waveguide is usually long and therefore, is placed parallel to the gantry rotation axis, which accelerates
electrons in a direction perpendicular to that required for incidence on the patient. The electrons are therefore steered through the required angle, usually 270° by an achromatic bending magnet.

The linac treatment head contains the required components for production, monitoring and shaping of the clinical photon and electron beams. The most important components in a treatment head are primary collimator, jaws, ionization chamber, multi-leaf collimator (MLC), x-ray target and flattening filter for photon mode, or scattering foil for electron mode. Special cones and applicators are also used to collimate the electron beams. The primary collimator defines a maximum field (40×40 cm² at isocentre), which is then further truncated by two pairs of movable collimators, i.e., jaws. The MLC is a device consisting of many individual ‘leaves’ of a high atomic numbered material, usually tungsten, which can move independently in and out of the radiation beam path to block it. The monitor chambers are included to monitor the beam output (i.e., monitor unit, MU) during the treatment, where one MU corresponds to one cGy in a water phantom under the calibration setup. In the photon beam mode, x-rays are produced when the electrons are incident on a target of a high-atomic number material, such as tungsten and flattened with a flattening filter that attenuates the central part of the raw beams to levels equal to those on the periphery.

1.1.3 Treatment Planning

In radiation therapy, treatment planning is the process that plans the appropriate external beam radiotherapy or internal brachytherapy treatment technique for a patient with cancer, with the goal of delivering a sufficient and appropriate dose to the tumour volume while maximizing the sparing of surrounding and uninvolved organs and tissues. During this process, treatment simulations are used to plan the geometric, radiological, and dosimetric aspects of the therapy using a treatment planning system (TPS), involving the delineation of the planning target volume
(PTV) and organs at risk (OARs) as well as the selection of the appropriate beam energy and arrangements.

Multiple medical imaging modalities may be used in order to develop a detailed three-dimensional (3D) model of the patient, including the tumour volume and surrounding anatomy. Of all the modalities, computed tomography (CT) is most commonly used, because of its relative low cost, but magnetic resonance imaging (MRI) can provide a more detailed image of soft tissue. Thus, MRI is often used in the diagnostic examination of breast tumours and brain tumours. Ultrasound is another attractive alternative imaging technique that offers excellent soft tissue contrast. It is often used in the imaging guidance for abdominal and pelvic cancer sites, such as liver and prostate. On the other hand, positron emission tomography (PET) is a nuclear medicine imaging that uses radioactive tracers to detect malignant cells. When combined with CT, PET-CT can provide both soft tissue discrimination and functional information (e.g., cellular metabolism) and anatomic imaging of the patient.

Medical imaging allows for the accurate identification and delineation of the target volumes and critical structures. The delineation is the prerequisite for 3D treatment planning. Several volumes related to 3D treatment planning have been defined according to International Commission on Radiation Units and Measurements (ICRU) Reports No. 50 [3]. The volumes of interest are gross tumour volume (GTV), clinical target volume (CTV) and planning target volume (PTV). GTV is the gross, palpable/imaged volume of malignant growth, CTV is the GTV plus a margin to account for the possible presence of microscopic malignant disease and PTV is the volume that contains the CTV plus the setup margins.

The radiation beam is shaped to fit the contour of the target from the beam's eye view (BEV) using either MLC to control beam weighting for photon therapy, or an applicator for
electron therapy. The dose is in three dimensions with a dose-calculation algorithm that accounts for the divergence of the beam and heterogeneity in all directions. The plans are usually assessed with the aid of dose-volume histograms (DVH), allowing the clinician to evaluate the uniformity of the dose to the PTV and sparing of normal structures.

Several metrics are used in the evaluation of dose distributions, among which dose conformity (CI), dose homogeneity (HI) and dose fall-off (FO) are the three most used indices. The dose conformity is determined by the Paddick conformity index (CI) [4], which is defined as

\[
CI = \frac{TV_{PIV}}{TV \cdot PIV}
\]  

(1.1)

where TV is the volume of the target, prescription isodose volume (PIV) is the volume receiving the prescribed dose and TV\textsubscript{PIV} is the volume of TV overlapping with PIV (Figure 1.3).

**Figure 1.3** The dose conformity is defined using TV, PIV and TV\textsubscript{PIV}.

The dose homogeneity index describes the variation of dose across the PTV, which is defined as

\[
HI = \frac{D_{5\%PTV}}{D_{95\%PTV}}
\]  

(1.2)

where \(D_{5\%}\) and \(D_{95\%}\) is given by Figure 1.4. The ideal value of HI is 1 and it increases as the plan becomes less homogeneous. The dose fall-off outside the target volume is described by the ratio of the 80% and 20% isodose surfaces (i.e., surfaces that contain points of equal dose).
Figure 1.4  The dose homogeneity index is defined in a DVH, where $D_{5\%}$ and $D_{95\%}$ represent the doses to 5% and 95% of the PTV, respectively.

1.1.4  Dose Calculation Algorithm in Radiation Therapy

The quality of radiation therapy depends on the ability to maximize the TCP while minimizing the NTCP at the same time. Both quantities are directly dependent on the absorbed dose in the PTVs and the OARs, and thus accurate knowledge of the dose distributions within the patient are crucial in radiation therapy. Therefore, the accuracy of patient dose calculations is one of the most important tasks in modern radiation oncology. The application of dose calculation algorithms involves the plan optimization in the treatment planning process, and the retrospective analysis of the correlation between treatment parameters and clinical outcome. ICRU has recommended an overall dose accuracy within 5% [5], considering the uncertainties resulting from patient setup, machine calibration and dose calculation from treatment planning systems.

In radiation therapy, the microscopic structure (e.g., at the level of lung alveoli, bronchi and blood vessels) is not taken into account. For the purpose of dose calculations in millimeter-
sized voxels (typically 2.5 mm), it suffices to consider an average density as defined at the space resolution achievable with a CT scanner. Modern dose calculations are generally based on the calibration of CT number to electron density. The CT number, i.e., Hounsfield Unit (HU), describes the radiodensity of the material and is defined by the linear attenuation coefficient ($\mu$) using:

$$HU = 1000 \times \frac{\mu - \mu_{\text{water}}}{\mu_{\text{water}} - \mu_{\text{air}}}$$  \hspace{1cm} (1.3)

Typical HU ranges are listed in Table 1.1:

<table>
<thead>
<tr>
<th>Materials</th>
<th>HU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air</td>
<td>-1000</td>
</tr>
<tr>
<td>Lung</td>
<td>-700 to -600</td>
</tr>
<tr>
<td>Fat</td>
<td>-120 to -90</td>
</tr>
<tr>
<td>Muscle and soft tissues</td>
<td>20 to 90</td>
</tr>
<tr>
<td>Bone</td>
<td>700 to 3000</td>
</tr>
</tbody>
</table>

Accurate calculation of dose distributions in a heterogeneous medium, such as tumours located in lung or around air cavities, is an intricate task, where a very detailed modelling of the energy transport in the patient is required. The Monte Carlo (MC) simulation is considered the most accurate algorithm for dose calculation, but it requires the greatest processing time. This approach will be discussed in detail in Chapter 2. Apart from the Monte Carlo method, all other methods make different degrees of approximation and simplification, which lead to much faster calculation speed but also result in less accurate dose distributions. One commonly used dose calculation algorithms in TPS is the pencil beam algorithm (PBA) [6]. The PBA is based on convolution techniques, in which the heterogeneity is handled either by an equivalent path length
correction or by scaled kernels. The lateral electron transport is considered in an approximate way.

1.1.4.1 Pencil Beam Algorithm (PBA)

The PBA assumes that any collimated photon beam incident on the patient is actually a conglomeration of many smaller, narrow ‘pencil beams’. Each of these pencil beams has a central axis ray along which it deposits some dose. The PBA calculates the dose from the convolution of the total energy released per unit mass (TERMA) with the kernel. While TERMA represents the energy imparted to the medium by the interactions of primary photons, the kernel represents the energy deposited about a primary photon interaction site. When one infinitely narrow pencil beam of photons hits the patient surface, the beam deposits teardrop shaped distribution of dose (i.e., kernel) in the patient body according to the scattering and absorption processes that the photons and secondary electrons undergo. The dose kernel is then used to calculate the total dose by adding up the dose contribution to each dose voxels from all the surrounding pencil beams. Real patients, however, have heterogeneities due to the various densities of the tissue (i.e., bone, lung, air cavities, muscle, etc.). Difference in densities leads to changes in photon attenuations and dose absorptions. To account for tissue heterogeneities in a patient, an equivalent path length correction based on the CT image dataset is performed. These corrections are applied to the dose kernel for each pencil beam depending on the local density variations that affect that pencil beam.

However, the PBA does not appropriately account for the heterogeneity correction in the lateral direction and thus can result in deviations up to 5% in low-density volumes for 6 MV photon beams [7]. The pencil beam kernels are usually based on MC generated energy distributions of photons impinging on a semi-infinite slab. This could cause errors when dealing
with small irradiated volumes in the lateral direction. The overestimation of phantom scatter in these configurations yields an overestimation of the calculated dose. For small targets imbedded in low-density environments, which is usually the case in lung stereotactic body radiotherapy (SBRT), the PBA calculated doses are significantly different (i.e., an overestimation up to 9%) from those calculated with MC [8-10]. Another group reported that PBA could overestimate the dose by up to 7.3% in the region beyond the high-density medium, and thus metallic immobilization devices should be avoided in the beam path [11].

1.1.4.2 Anisotropic Analytical Algorithm (AAA)

The anisotropic analytical algorithm (AAA) [12, 13] is implemented in the Eclipse TPS (Varian Medical Systems, Palo Alto, CA) to replace the PBA for photon beam dose calculation. The AAA dose calculation model is a 3D pencil beam convolution-superposition algorithm that has separate modelling for primary photons, scattered extra-focal photons, and contaminant electrons. The goal of the AAA is to enhance the accuracy of dose calculations in heterogeneous media. When heterogeneities lie in the volume of tissue immediately surrounding the voxel of interest, the effect of such heterogeneities is not considered in the PBA. The AAA accounts for the tissue heterogeneity by employing the concept of radiological scaling:

$$I(z, \rho) = I(z')$$  \hspace{1cm} (1.4)

where $z'$ is the radiological depth and is defined by the segment length and the density of the region $\rho(t)$:

$$z' = \int_{0}^{z} \frac{\rho(t)}{\rho_{water}} dt$$  \hspace{1cm} (1.5)

Tissue heterogeneities are handled anisotropically with the full 3D neighborhood by the use of lateral photon scatter kernels. The final dose distribution is obtained by superposition of the
doses from the photon and electron convolutions. These in turn are used to construct a phase space (phsp) defining the fluence and energy spectrum of the clinical beam specific to each treatment unit.

The AAA dose calculation model consists of two components, the configuration algorithm and the actual dose calculation algorithm. The configuration algorithm is used to characterize the clinical beam for a specific treatment unit based on type of particle, fluence and energy and store the data in a phsp. The phsp is created by a multiple source model consisting of a primary photon source, an extra focal photon source and an electron contamination source. Configuration of the AAA model is based on MC-determined basic physical parameters that are adapted to measured clinical beam data. The primary photon source models the bremsstrahlung resulting from the accelerated electrons interaction with the target using pre-calculated Monte Carlo methods. A mean energy radial curve takes into account the effects of the flattening filter by describing how the mean energy decreases with increased distance from the central axis.

The AAA model shows a good agreement with Monte Carlo calculations and the measurements and a great reliability to handle small fields in heterogeneous media compared to pencil beam convolution algorithms [14, 15]. Hence, the AAA is preferred over the PBA for lung cancer SBRT and breast cancer radiotherapy as it gives more realistic dose distributions [16, 17]. However, when photon beam traverses a large air gap (up to 15 cm) after passing through the treatment couch or an immobilization device, it may result in dose calculation errors more than 2.5% [18]. This effect is due to a decrease in predicted scattered radiation in AAA. In addition, one group indicated that AAA brought out larger deviations in a phantom with water-lung interfaces, as compared to MC and measurements, especially for small field sizes and high beam energies [19].
1.1.4.3 Collapsed Cone Convolution (CCC)

The adaption of the superposition approaches to 3D heterogeneous volume can be computationally intensive. To improve the efficiency, a refined convolution method was described by Ahnesjö [20], called collapse cone convolution (CCC). By the introduction of a collapsed cone approximation, the space of dose calculation is discretized into a lot of cones. All energy released into coaxial cones of equal solid angle, from volume elements on the cone axis, is rectilinearly transported, attenuated, and deposited in elements on the axis. The CCC algorithm convolves a poly-energetic spectrum to the point-kernel energy distribution and kernel scaling for the heterogeneity is performed during the convolution process. The point-spread kernels are often pre-calculated with MC simulations for discrete energies and combined into poly-energetic point-spread kernels. This algorithm is used in Pinnacle³ TPS (Philips Medical Systems, Netherlands) [21].

Hasenbalg et al. [22] reported that small differences (≤ 2%) between the CCC and VMC++, a fast MC code [23], were evident in the cases with large heterogeneities, such as lung or breast. However, in breast cases, CCC gave PTV doses that were less homogeneous than the VMC++ results. In a phantom study that involved both point dose measurements and planar measurements [24], average point absolute deviations less than 3% were found for the CCC algorithm. The deviations increased up to 5% for small targets completely embedded in very low-density media (0.004 g/cm³). A recent report [25] also showed that CCC had better agreement with measured data in the presence of tissue heterogeneities and modelled dose changes at the tissue-lung interface better than AAA.
1.1.4.4 Acuros XB (AXB)

Varian also developed a MC-equivalent model-based dose calculation algorithm, Acuros XB (AXB) [26], which directly solves the linear Boltzmann transport equation (LBTE). The Boltzmann transport equation (BTE) is the governing equation that describes the macroscopic behavior of ionizing particles as they travel through and interact with matter. By assuming that radiation particles only interact with the matter they are passing through, the BTE can be written in the linearized form, i.e. LBTE. While MC solves the LBTE through random sampling techniques for the transport of radiation in the medium, the AXB provides a direct solution to the LBTE that gives a dose distribution. The detailed description of this LBTE solver can be found in [27] and is reviewed in [28]. The AXB first solves for photon and electron angular fluence, respectively, through discretization in space, angle and energy. Then, the dose is generated by using the macroscopic cross-sections associated with any given photon or electron interactions and the density of the materials. Errors of AXB are primarily systematic and result from discretization of the variables in space, angle, and energy. Therefore, larger steps in the discretization process result in a faster solution, but less accuracy [26].

Many studies have been published to compare AXB with MC simulations or measurements [25, 29-31]. Fogliata et al. [29] showed an average gamma agreement, with a 3%/3mm criteria, of 100%, 86% and 100% for normal lung (HU = -780), light lung (HU = -942) and bone (HU = 1380) settings, respectively, compared to VMC++ results. Consistent results [30] were observed, with gamma passing rates exceeding 99.5% between AXB and COMPASS system (IBA Dosimetry, Germany). In small field dosimetry, the AXB was also found to be superior to the AAA algorithm [31]. Compared with W1 Exradin Scintillator (Standard Imaging, Middleton, WI) measured values, the AXB showed an average dose difference of 3.2% for
percentage depth dose (PDD), 2.7% for a target located in the brain in a head phantom, and 3.5% for a target located in the neck in the same phantom.

1.2 Inverse Treatment Planning

Inverse planning is the foundation of intensity-modulated radiotherapy (IMRT) and its performance critically determines the success of an IMRT treatment. This planning method is further employed in the more advanced radiation therapy technique, i.e., volumetric modulated arc therapy (VMAT). In conventional 3D conformal radiotherapy (3D-CRT), a treatment planner calculates dose first and determines whether the dose distribution is acceptable. A completely reversed process is performed in inverse planning. Dose constraints are specified before the dose calculation, including the dose limits for the tumour and the dose/dose-volume constraints for the surrounding normal tissues. A computer algorithm then determines beam shapes and beam weights that are most likely to meet the prescription requirements designated at the start of the treatment planning process. After numerous modifying iterations, where beam shape and intensity are adjusted, the planning system generates an optimum treatment plan with a configuration best matched to the desired plan. While 3D-CRT is forward planned and reliant on the skills of the treatment planner to decide the number, shape and orientation of the beams, inverse planning, in contrast, only requires specifications of the plan outcome (i.e., prescription dose and normal structure dose limits). The prescriptions are given by the oncologists according to clinical protocols, based on the type and phase of the disease (e.g., the radiation therapy oncology group (RTOG) protocols: https://www.rtog.org/ClinicalTrials/ProtocolTable.aspx).

The field shapes of inverse planning are not defined manually since the number of fields is usually large (11-13 fields for IMRT or 178 control points for VMAT). In addition, inverse
planning is less dependent on the geometric parameters but more on specification of volumes of tumour targets and sensitive structures, as well as their dose-volume constraints. In complex clinical situations that require the use of many beam directions and fields, inverse planning can be the more efficient strategy.

1.2.1 Objective Function

For inverse treatment planning, the clinical objectives are specified mathematically in the form of an objective function (also called cost function). The value of the objective function is a dimensionless measure of the plan quality, in the sense that plans with lower values of this function are those that approach the clinical dose constraints prescribed by the radiation oncologist. The term ‘cost’ is often used to denote this value. A cost of zero would represent a plan that achieves all clinical constraints. Throughout this thesis, it is in this sense that we characterize plans as being of “higher quality” than other plans. Computer optimization techniques are used to determine the beam parameters (e.g., position of the MLC leaves or intensity of the beamlets) that will most closely achieve the desired solution. The objective function reduces the entire treatment plan into a single numerical value that scores the quality of the plan.

At present, most optimization systems use dose-based and/or dose-volume-based criteria. One method commonly used to create dose-based and dose-volume objective functions is based on minimizing the variance of the dose relative to the prescribed dose for the target volumes or dose limits for the OAR. Variance is defined as the sum of the squares of the differences between the calculated dose ($d_{i, \text{calculated}}$) and the prescribed dose ($d_{i, \text{prescribed}}$). Thus, a typical dose-based or dose-volume-based objective function is the sum of the variance terms representing
each voxel $i$ in anatomic structure multiplied with weighting factor $w_i$. This approach is referred to as a quadratic objective function:

$$\text{Cost} = \sum_i w_i (d_{i,\text{calculated}} - d_{i,\text{prescribed}})^2$$

(1.6)

where weighting factor $w_i$ must be non-negative and is chosen to reflect the relative importance of the objective, e.g., 120% for the PTV. These constraints can be used to guarantee adequate dose uniformity in the tumour and provide improvements in dose conformity, but they may also allow small hot and/or cold spots. Each tissue can be assigned a different priority such that it contributes differently to the overall objective function. The priority of each structure can be adjusted by the planner to introduce a desired modification to the dose distribution.

1.2.2 Intensity-Modulated Radiotherapy (IMRT)

IMRT modulates the intensity of the radiation over the field of delivery. The purpose of IMRT is to manipulate the beam intensity so that the dose conforms closely to the tumour or target volume within the patient. IMRT may use multiple radiation beams of varying sizes and varying intensities to irradiate a tumour with precision and accuracy. The radiation intensity of each part of the beam is controlled, and the beam shape may change or multiple beams used throughout each treatment. Thus, IMRT can shape the radiation dose to avoid or reduce exposure of normal tissue and limit the side effects of treatment while delivering a therapeutic dose to the cancer.

The first step is to determine the treatment plan objectives, including prescription dose and constraints. The clinician must specify the required dose to be delivered to the PTVs and the dose limits for the OARs. The next step of treatment planning for IMRT is the placement of the treatment beams, including the choice of the beam energy, number and direction of the beams.
and the isocentre. In a typical IMRT treatment, source-to-axis distance (SAD) setup is used, where the center of the target volume is placed at the machine isocentre. The beam energy is selected according to the location of the tumour. A combination of photon beams is then planned to concentrate the dose in the target volume and spare the tissues surrounding the target as much as possible. Once the beam configuration and plan objectives are established, the optimum intensity distribution for each beam can be determined. This is achieved through an iterative optimization process. Each beam is divided into a number of smaller beamlets and the corresponding dose distributions are computed. The number of small beams or beamlets is often more than one thousand. The optimization process then adjusts the weights of these beams to produce a desired fluence map and the resulting dose distribution is used to calculate an objective function. The effect of a change in the weight of each individual ray or beamlet is then calculated, with the weight increased, decreased, or left the same depending on whether the change would be favorable for the patient. This process is repeated for all beamlets until no further improvement is seen. This method is referred to as beamlet-based inverse planning or fluence map optimization (FMO).

The optimized intensity modulations, however, must be convertible to a deliverable field, given the physical constraints of the beam-limiting device (i.e., MLCs). Therefore, a leaf-sequencing algorithm needs to be applied to translate the each optimized fluence map into a set of deliverable aperture shapes, where constraints imposed by the MLC are accounted for. Furthermore, the plans that require a large number of fields and a large number of MU can lead to long treatment times. The number of MU in an IMRT plan is two to three times higher than a conventional radiotherapy plan. These deficiencies led to some treatment planning systems
directly including the MLC constraints in the optimization process and the development of aperture-based optimization, i.e., direct aperture optimization (DAO) [32, 33].

1.2.3 Direct Aperture Optimization (DAO)

In beamlet-based inverse planning, the optimization provides a distribution of the beamlet weights. After the optimization is complete, each beamlet weight is sequenced into a set of deliverable aperture shapes. By contrast, in aperture-based inverse planning, a set of deliverable apertures are included in the optimization. The optimization parameters include both aperture shapes and weights. The optimized treatment plan does not require leaf sequencing, and thus is ready to deliver and avoids the plan degradation in beamlet-based optimization. The leaf settings and the aperture intensities are optimized simultaneously using a simulated annealing algorithm. A large number of candidate apertures are sampled and either accepted or rejected depending on whether the plan is improved by adding the new aperture. This technique allows the user to specify the maximum number of apertures per beam direction, and hence provides control over the complexity of the treatment delivery. Therefore, IMRT using DAO can provide a highly conformal IMRT plan, and significantly reduce the number of fields and MU. DAO resulted in an average 32% less time to calculate, 42% fewer MU and 35% fewer segments [34].

However, it is more difficult to solve a DAO problem than FMO because the relation between leaf positions and fluence is neither linear nor convex. A nonconvex optimization can possibly lead to multiple local minima and should have effective strategies for dealing with these local minima [35]. Several optimization methods have been proposed to tackle the nonconvex nature of DAO, including simulated annealing, gradient-based methods and multi-criteria optimization.
Simulated annealing (SA) methods [32, 36] allow occasional ascent in the random search process in order to avoid the trap of local minima. The term ‘simulated annealing’ derives from the physical process of heating and then slowly cooling a substance to obtain a strong crystalline structure. To apply SA, random changes to the variables are introduced and solutions with worse objective values are accepted with some probability.

Gradient-based methods [37, 38] use the gradient of the objective function with respect to the optimization parameters (leaf positions and weights) to find an update of the parameters. The referenced methods use a treatment plan generated by FMO and leaf sequencing as the initial point, and then use first order derivative information to determine the direction of the change and second order derivative information to adjust the step size.

Multi-criteria optimizations [39-41] generate a set of relevant treatment plans that are Pareto-optimal, instead of the planner performing an iterative optimization with adjustments to optimization functions and weights. A solution is Pareto-optimal, if no criteria can be improved without worsening at least one of the remaining criteria. Specifically, each Pareto-optimal plan has varying strengths and weaknesses regarding the treatment planning objectives.

1.2.4 Volumetric Modulated Arc Therapy (VMAT)

Despite improvement to the plan quality, IMRT still has some disadvantages. A standard IMRT plan often requires multiple fixed angle radiation beams, which increases the treatment delivery time. The larger number of MU used in IMRT can lead to an increase in the amount of low dose radiation to the rest of the body. The increase in MU and subsequent increase in low dose radiation can result in increased risk of secondary radiation-induced malignancies, which is of particular relevance in paediatric patients or patients with long life expectancies [42, 43].
To overcome these limitations associated with fixed field IMRT, there has been some interest to introduce of arc/rotational therapies. The basic concept of arc therapy is the delivery of radiation from a continuous rotation of the radiation source and allows the patient to be treated from the entire beam angle. As an alternative form of IMRT, arc therapies have the ability to achieve highly conformal dose distributions, while bringing about improvement in treatment delivery efficiency as a result of the reduction in treatment delivery time and the reduction in MU usage. This advantage subsequently reduces the integral radiation dose to the rest of the body. There are two main forms of arc-based therapies: tomotherapy and VMAT. Tomotherapy machines can deliver the radiation beam in a fan-shaped distribution, similar to CT imaging with a continuously rotating radiation source, while the patient is moved through the machine.

VMAT was first introduced in 2007 and was described as a novel radiation technique that allows the simultaneous variation of gantry rotation speed, treatment aperture shape via movement of MLC leaves and dose rate [44]. While the delivery of a typical IMRT plan with five or seven fields requires 5–20 minutes, that of a single arc VMAT treatment fraction can be reduced to 1–1.5 minutes [45]. VMAT creates a progressive resolution sampling of continuous motion delivery by evenly distributed static source positions to generate and deliver plans in a single gantry rotation. This avoids the repetition of arcs used to produce multiple unique apertures for the same geometry span. Each static source position is referred to as a control point. While the MLC leaf position is initialized at discretized control points, the leaf motion between adjacent angles is limited by leaf travel speed and gantry rotation speed. Beginning with coarse sampling, iterative modifications are made to the MLC leaf positions and MU weights of the sampled points. In the initial stages of optimization, large-scale adjustments are made in leaf sequencing and dose rate. The scale of these adjustments reduces as the optimization progresses.
Samples are added progressively throughout the optimization at the midpoint between existing samples until the sampling has reached its final resolution. The amount of MUs and MLC transition motion can vary between sampled control points. These constraints control the efficiency of the delivery. Similar to IMRT, if a change does not violate the restrictions above, the dose distribution and resultant cost function are evaluated. If the total cost is reduced, the change is accepted. The optimization cost function is based on user-defined objective dose-volume constraints for the structures in the plan. Doses are specified as function of structure volume and a relative priority for each constraint is input into the system.

**1.2.4.1 Trajectory-based VMAT (TVMAT)**

VMAT treatments usually utilize a coplanar beam trajectory realized by a fixed couch angle, typically zero degree, despite the fact that non-coplanar beams are routinely used in IMRT for many treatment sites including intracranial lesions, head-and neck tumours, and lung SBRT treatments. Although VMAT is a well-accepted treatment technique in radiotherapy using a coplanar delivery approach, it might be further improved by including dynamic couch table and collimator rotations, leading to the development of trajectory-based volumetric modulated arc therapy (TVMAT). In the IMRT literatures, it has been established that treatment plans using non-coplanar beam angles can yield considerable improvement in organ sparing [46-48]. This is especially the case if critical organs are close to irregularly shaped target volumes and for target volumes that are not centrally located.

During conventional VMAT treatment, the patient is stationary on a static patient support couch while the linac gantry rotates about the patient to treat the cancerous tissue from a variety of incident paths. As a consequence of this fixed patient approach, tissue along these paths between the source of radiation and the target volume can receive unnecessary radiation dose.
The concept of a gantry-couch trajectory was illustrated to overcome this issue [49], in which a gantry-couch space recorded overlap between OARs of exposure and the PTV to indicate the appropriateness of a particular couch-gantry coordinate for treatment.

The use of non-coplanar beams in VMAT can potentially reduce the radiation exposure of sensitive normal tissues by moving them out of the path of the delivered beam. There are various degrees of complexity for this technique: 1) a small number of non-coplanar IMRT beams added to a standard coplanar VMAT plan; 2) a plan consisting of multiple arcs with different couch angles, while the couch angle is fixed during each arc; and 3) a plan utilizing simultaneous gantry and couch rotation while the treatment beam is on. In the third case, the incident beam direction describes a complex trajectory around the patient and is of our interests here. However, in principle the first two approaches can be realized with current linac units, though the support of planning systems to choose non-coplanar arcs or beam directions may be limited.

1.3 Patient-specific Quality Assurance

Modern radiotherapy has very high requirements on systematic accuracy and precision in relation to the total dose, dose distribution and dose fractionation. In order to satisfy these requirements, both systematic and random errors must be kept within a few percentages during the entire process, from the time of prescription to the completion of the radiation treatment. Consequently, the treatment outcome in an individual patient depends on both the professional involvement of staff and execution accuracy of planned procedure. Therefore, the accurate delivery of planned radiation dose is essential, especially in the advent of complex treatment technology. This makes rigorous quality assurance (QA) of the treatment processes very crucial.
International Atomic Energy Agency (IAEA) [50] and International Commission on Radiological Protection (ICRP) [51] reported that lack of adequate QA could lead to misadministrations of radiation dose. The QA can be categorized in two groups: pretreatment (without patient) verification and in vivo dosimetry (with patient). Pretreatment verification can detect a large number of errors in the radiotherapy before starting treatment, but any unexpected change at the time of treatment will be missed. Possible sources of errors can be changes in the linac output, incorrect accessory, MLC failure, patient changes (tumour shrinkage, weight loss, organ motion or anatomical changes since CT images acquired) or even changes in patient setup, including obstruction from immobilization devices. On the other hand, in vivo dosimetry is more likely to detect adverse clinical events (e.g., setup or transcription errors) during the treatment. By using real-time linac log files and monitoring linac beam output, differences between prescribed and delivered dose can be detected.

Typically, the pretreatment QA process involves comparing a dose plan delivered to a regular phantom with the dose calculated by the treatment planning system for the same geometry. However, plans can pass pre-treatment phantom QA, but still have relatively large dose errors in some regions of the patient’s anatomy [52]. Thus, while these pretreatment QA approaches do ensure that no large errors in dose calculation are present, a calculation of the dose to the patient’s anatomy is necessary. Using a secondary dose calculation package, a patient-specific QA can verify the ability of the treatment planning system to calculate the dose accurately for this patient’s plan. A patient specific QA is usually measurement-based, which uses the same fields and number of MUs as applied for the patient irradiation. This requires an additional treatment plan and dose calculation with the CT data of the phantom used for these measurements. Since the beam-patient geometry is different from the beam-phantom geometry,
an additional check of the effect of the patient anatomy on the dose calculation is often performed with an independent MU/dose verification. A dose detector (e.g., ArcCHECK, MatriXX) or a film is placed at a certain depth in the phantom to measure either the dose distribution of each beam separately or the cumulative dose for all beams.

### 1.4 Gamma Evaluation

The gamma evaluation technique [53] can be used to compare two dose distributions. It compares dose differences at specific points, and determines the distance to the nearest point having the same dose value. To perform the comparison between the reference dose at a point \( r \), the test dose at all points \( r' \) within a predefined searching radius \( R \) are evaluated. The gamma index at \( r \) is defined as

\[
\gamma(r) = \min_{(r' \in R)} \left\{ \sqrt{\frac{D^2(r, r')}{\Delta D^2} + \frac{d^2(r, r')}{\Delta d^2}} \right\}
\]  

(1.7)

where \( D(r, r') \) is the dose difference between test and reference doses at point \( r \), \( d(r, r') \) is the spatial distance between test and reference dose points, \( \Delta D \) is the dose difference criterion and \( \Delta d \) is the distance-to-agreement (DTA). The DTA at the comparison point \( r \) is the distance from point \( r \) to the closest point \( r' \) such that \( D(r, r') = 0 \). The dose difference criterion is important in low dose gradient regions, while the DTA criterion yields valuable information in regions having a high dose gradient. Dose difference and DTA criteria can be selected and often values of 3\% (of prescription dose)/3mm are used. The pass-fail criteria \( \gamma \) is calculated for each voxel such that the evaluation passes if \( \gamma \leq 1 \) and fails if \( \gamma > 1 \). Once the \( \gamma \) of all voxels is determined, the pass rate is calculated as the ratio between the number of voxels that passes the evaluation and the total number of voxels.
The gamma evaluation method has the advantage of producing a quantitative measure based on both dose and spatial criteria. However, it has been criticized for being insensitive to dose grid size [54], being less clinically intuitive than more conventional dose comparison methods [52, 54, 56] and having poor sensitivity to clinical dosimetric inaccuracies (i.e., single field gamma analysis is insensitive to dosimetric inaccuracies of the overall plan) [56, 57]. It has been shown that even when QA gamma passing rates are high, many of the larger critical errors in patient dose are still occurring [52].

Currently at BC Cancer – Vancouver, the gamma evaluation method is used for the comparison between dose distributions. Thus, this method has been used for dose comparisons through the thesis, since the focus of this thesis was not the modification of the existing QA program. The gamma method was supplemented by plan analysis based on DVH metrics in all cases presented in this thesis. New concepts, such as maximum allowed dose difference [54] and normalized dose difference [54], will be considered in future.

1.5 Electronic Portal Imaging Devices (EPID)

An effective means to reduce setup errors would be to increase the frequency of treatment verification with portal imaging [58]. Historically, portal imaging was performed through film cassettes, where a sheet of film is sandwiched between a front metal plate and a rear plastic or metal plate. By detecting the incident x-rays, the front plate acts as a build-up layer. In addition, this front layer serves to block scattered secondary radiation incident on the cassette to avoid a loss of contrast. The back plate serves as an electron backscatter material. Along with the overall design of the cassette, the back plate also helps to ensure good contact between the film and the surrounding materials thereby contributing towards the preservation of image quality. Utilizing
port films as control system, however, is time consuming and labor intensive, which could decrease the throughput in a busy radiotherapy clinic. Therefore, the need for an efficient and accurate portal imaging system to intensify verification of radiation therapy arouses the clinic use of electronic portal imaging devices (EPID). Instead of a film, an EPID is placed in the exit beam to produce the image. The images are captured and displayed on a video screen and can be compared with the reference images to determine the placement differences or geometric and dosimetric errors. These include uncertainties in a particular patient set-up, uncertainties in the beam set-up, anatomical and physiological changes of patient. The most commonly applied type of EPID is the amorphous silicon (a-Si) EPID [59], which is able to acquire high-resolution (1024×768) images. The panel of an EPID consists of an x-ray converter, light detector, and an electronic acquisition system for receiving and processing the resulting digital image. The main principle of the a-Si EPID is based on a two-step process. The megavoltage photon beams are first converted into visible light by means of a metal plate and phosphor screen. The generated photons are then absorbed by the underlying a-Si photodiodes, which create electrical signals.

1.5.1 EPID in vivo Transit Dosimetry

The EPID is an excellent in vivo dosimetry system since it allows more precise, quantitative results to be obtained with much less effort than would have been achievable using conventional QA tools [60]. EPIDs can be implemented as a tool for both transit and non-transit dosimetry, i.e., with or without an attenuation object (patient or phantom) in the beam path. In EPID transit dosimetry, the dose is verified either at the level of the EPID or by reconstructing the dose inside a digital representation of the patient, and thus verifies whether the patient is receiving the correct dose during treatment.
Several groups have applied EPID *in vivo* dosimetry in various treatment techniques, including IMRT and VMAT [61, 62]. Van Elmpt et al. [63] developed a system that predicted a transit portal dose image using a through-air portal dose image and the radiological thickness of the path of the beam through the patient. Berry et al. [64] used the same algorithm along with Monte Carlo simulations to simulate the effect of an attenuating object in the beam (transit dosimetry). Peca et al. [65] proved EPID is sensitive to dose discrepancies and detected discordances in seven rectal cancer patients. Mans et al. [66] revealed serious underdosage of the PTV due to a network transfer error based on EPID reconstructed dose. Nijsten et al. [67] developed a model for dose calibration of a-Si EPIDs for 2D transit dosimetry, in which corrections for the ghosting effect, image lag, and beam profile were implemented. In addition, a convolution model was included to correct for the field size dependence and scattering of the EPID response. Persoon et al. [68] used this method to analyze the interfractional trend of the patient dose throughout the course of treatment. Consorti et al. [69] checked isocentre dose and the treatment reproducibility for non-small cell lung cancer and observed paradigmatic discrepancies in three patients.

**1.5.2 EPID Cine-mode**

One interesting topic in EPID image acquisition is cine-mode, where images are continually acquired and saved in real time. Thus, instead of a single integrated image, a continuous acquisition results in a series of images where EPID dose data are captured as a function of time during the entire irradiation.

The EPID cine-mode has a great potential for QA of arc-based radiotherapy techniques [70, 71] due to its high resolution matrix for imaging and the capability of acquiring cine-images with a high frame rate (e.g., up to 30 frames per second for PortalVision aS1000, Varian Medical
The MLC aperture of each control points is verified by acquiring cine-images throughout the whole arc. Snapshots of the MLC apertures at individual control point, specific to a gantry angle, are acquired. Therefore, each snapshot acquired at a specific gantry angle can be compared to the MLC shape generated by the TPS at this specific gantry angle. Piermattei et al. [73] proposed a method for the in vivo determination of the isocentre of dose dynamic conformal arc therapy technique using the transited signal. The method was tested with cylindrical and anthropomorphic phantoms. Berbeco et al. developed a matching technique for respiratory-gated liver radiotherapy treatment verification [74]. The precision of the patient set-up, the placement of the beam-gating window, as well as the residual tumour motion was assessed for each treatment fraction. This method was further introduced to monitor the tumour location during stereotactic body radiotherapy (SBRT) [75]. Arimura et al. [76] considered using EPID cine-images for measurement of displacement vectors of tumour positions in lung radiotherapy by calculating the similarity between a reference portal image and each EPID cine-image. Liu et al. [77] proposed a VMAT QA procedure where the acquired cine-images were converted to dose matrices after profile correction and dose calibration. The portal image prediction was then compared to the original plan.

1.6 Stereotactic Radiotherapy (SRT)

Stereotactic radiotherapy (SRT) is a specialised type of radiation technique that precisely delivers targeted radiation with high dose in fewer fractions compared to traditional therapy. It is usually used in the treatment of functional abnormalities and small, well-defined tumours. SRT usually involves 1-5 treatments over 1-2 weeks, while conventional radiation involves five days a week typically over 5-9 weeks, depending on the diagnosis. Specifically, stereotactic
radiosurgery (SRS) is a form of stereotactic radiation that involves a single fraction and is generally reserved for treatment of lesions within the brain. When used for body tumours, it is also referred to as stereotactic body radiotherapy (SBRT). SRT relies on several technologies including 3D imaging and localization techniques that determine the exact coordinates of the target within the body; systems to immobilize (i.e., a mask or a head frame) and carefully position the patient and maintain the patient position during therapy; highly focused photon beams that converge on a tumour or abnormality; and image-guided radiation therapy (IGRT) which uses medical imaging to confirm the location of a tumour immediately before, and in some cases, during the delivery of radiation.

Although stereotactic radiation involves fewer fractions, the radiation dose delivered during each fraction is much higher than conventional radiation in order to achieve the same biological effect. Since the dose of radiation to that single point is high, the radiation oncologist must very precisely target the location of lesion/tumour. In order to do this, SRT delivers radiation from different angles to focus the radiation at one small point where the beams converge. The use of multiple unique beam angles limits dose to normal tissues as the beams enter and results in a very sharp radiation dose fall-off around the tumour. The sharp fall-off allows for decreased radiation dose to the surrounding normal tissue. In order for the radiation dose to adequately conform around the tumour, the size and shape of the tumour are important. Additionally, because high doses of radiation are used, the location of the tumour may impact whether SBRT is a viable treatment option. For example, SBRT may not be safe if a tumour is bordering or touching a critical, sensitive, normal structure. Due to the strict criteria for size, shape and location of the tumour, SBRT can only be used in select cases.
The biologic mechanism for killing cancer cells is different in stereotactic radiation and conventional radiation. Conventional radiation (e.g., 2 Gy per fraction) damages the DNA of cancer cells to prevent them from replicating and multiplying by causing sublethal DNA damage, which allows normal tissues to repair the DNA damage. In contrast, the high dose (e.g., > 10 Gy per fraction) delivered in stereotactic radiation tends to cause cell death by lethal DNA damage. Therefore, SRS/SBRT is often used to treat cancers that have been considered less responsive to conventional radiation, such as melanoma, renal cell carcinoma and sarcoma. The lung and the brain are the two most common and well-established sites treated with SRT. Multiple, well-conducted, prospective studies have been performed to demonstrate and validate the effectiveness and safety of stereotactic radiation in the brain and lung. Other increasingly common sites are the abdomen (liver, pancreas, and kidney), spine and prostate.

Treatment with SRT techniques can further benefit from the use of non-coplanar beam. Use of a multiple non-coplanar conformal static field treatment technique can reduce the volume of normal tissue receiving high and intermediate doses compared with a single isocentre multiple arc treatment technique, while providing a uniform dose in the target volume [78]. Dong et al. [79] re-planned twelve patients with centrally located or larger lung tumours using non-coplanar SBRT. Compared against the clinical VMAT and static intensity-modulated radiation therapy plans, this method yielded consistently improved tumour coverage and critical organ sparing. Yu et al. [80] observed lower dose of chest wall without increasing lung dose with a fixed couch at 345°. Similarly, Rwigema et al. [81] and Zindler et al. [82] reported improvements in critical organ sparing and tumour control head-and-neck cancer. Another study from Hallemeier CL et al. [83], however, suggested that patients treated with lung SBRT using a coplanar technique had similar outcomes as those treated with a non-coplanar technique. This study followed up 149
patients who had primary or recurrent non-small cell lung cancer or lung oligometastasis for 21 months.
Chapter 2: Monte Carlo (MC) Techniques in Radiation Therapy

MC simulations are a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results, typically running simulations many times over in order to obtain the distribution of an unknown probabilistic entity. The name Monte Carlo comes from the use of random numbers, i.e., similar to outcomes produced by rolling of dice or spinning of a roulette wheel. MC techniques are often used in physical and mathematical problems, particularly when they are intractable by analytic methods. However, MC simulation is often computationally intensive.

In radiation therapy, MC methods can model the physical processes of radiation transport, simulating the extensive random trajectories of individual particles by using a random number generator. With MC method, the known probability distributions (PDs) that govern the physical interaction processes of photons, positrons and electrons in matter can be sampled. The technique is accepted as the most accurate and versatile method for radiation dose calculation and has always been used for the generation of benchmark dose distribution to compare the results of other dose calculation methods [14]. The MC calculations are in good agreement with the thermoluminescent dosimeter (TLD) and film measurements (within the measurement errors and statistical uncertainty) [84]. The method can be further used to evaluate clinical cases where measurements are not possible.

A MC radiation transport algorithm simulates the transport of individual incident particles in a modelled beam as well as all secondary particles that receive kinetic energy during the transport process. Discrete photon interactions, such as photoelectric absorption, coherent or incoherent scatter, pair production, as well as collisional and radiative losses by charged particles
are modelled by sampling well-known PDs using random numbers. A particle history (also known as shower) is the ensemble of events that occur from the beginning of its transport until both the original particle and all progeny are absorbed (or have energies below a specified threshold), or exit the volume of interest. Along the path of the particle, selected physical quantities of interest (e.g., energy absorbed) are stored (or, in Monte Carlo terminology, ‘scored’). After transport of a large number of particles, the average and the standard deviation of these quantities can be calculated. The final dose distribution is acquired by summing the energy deposition in each particle’s history. The accuracy of the MC simulation increases with the number of particle histories.

2.1 Electron Gamma Shower Computer Code (EGSnrc)

EGSnrc [85] is an electron gamma shower software toolkit developed by National Research Council Canada (NRC) to design safe experiments for high energy physics. As a general-purpose package for the Monte Carlo simulation, EGSnrc is capable of modelling the coupled transport of electrons and photons in an arbitrary geometry, with energies ranging from a few keV up to several GeV. Showers are developed by simulating in as much detail as possible the various electromagnetic shower processes. When high-energy electrons and/or photons hit matter, they travel through the material, interacting with atoms and their nuclei in various ways predicted by physics. The path of each particle can be modelled as a random process as collisions with atoms occur with well-defined probability. Each individual interaction can generate many more electrons and photons due to collisions as it travels through matter. The PDs of the processes are used, and thus an EGSnrc simulation mimics real showers with real fluctuations in details. Currently, EGSnrc is considered the most accurate MC code available to calculate
radiation effects in body tissue and determine dosages. The code includes BEAMnrc [86] component for modelling radiotherapy sources and the dose scoring utility DOSXYZnrc [87] to estimate radiation dose in voxelized phantom.

2.2 Radiation Transport Calculations

In EGSnrc, particles are generated according to distributions describing the source of radiation (e.g., the linac target). They travel certain distances, determined by a PD depending on the total interaction cross section, to an interaction site and scatter according to the corresponding cross section, possibly producing new particles that have to be transported as well. This procedure is continued until all particles are absorbed or leave the geometry under consideration. Energy cut-offs are often set for both electron and photon to terminate the history and deposit the energy locally. Quantities of interest can be calculated by averaging over a given set of MC particle histories. The Monte Carlo estimate of quantities of interest is subject to a statistical uncertainty, which depends on the number of particle histories simulated.

2.2.1 Photon Interactions

Photons interact with surrounding matter via four basic processes: incoherent (Compton) scattering with atomic electrons, photoelectric absorption, materialization into an electron/positron pair in the electromagnetic field of the nuclei and surrounding atomic electrons and coherent (Rayleigh) scattering with the molecules (or atoms) of the medium. The first three collision types transfer energy from the photon radiation field to electrons, and one of them dominates depending on energy and the medium in which the transport takes place. The pair production process dominates at high energies, while incoherent scattering is the most important process at energies within megavoltage (MV) ranges.
Compton scattering is the dominant process for MV beams. The photon scatters off an electron in an outer orbit of an atom and the energy of the incident photon $k$ is then shared between the outgoing electron and the scattered photon. This results in a decrease in energy of the photon since part of the energy of the photon is transferred to the recoiling electron. EGSnrc samples the shell number $i$ using the probabilities $Z_i/Z$, where $Z$ is the occupation number of shell. The interaction is rejected if the incident photon energy is smaller than the binding energy $U_i$, otherwise the cross section of the Compton process is modelled using the Klein-Nishina equation. The cross section per electron ($\sigma$) can be gotten by an integration over all photon scattering angles ($\varphi$):

$$\frac{d\sigma}{d\varphi} = (2\pi sin\varphi) \frac{e^4}{m_e^2 c^4} \left( \frac{h\nu'}{h\nu} \right)^2 \left( \frac{h\nu'}{h\nu} + \frac{h\nu'}{h\nu} - sin^2 \varphi \right)$$

(2.1)

The EGSnrc further considers the binding effects and Doppler broadening according to the impulse approximation when calculating the energy $k'$ of a photon scattered at an angle $\theta$ [85], based on the assumption that the potential in which the target electrons move is constant so that their states can be represented by plane waves. The electron being set in motion then has a kinetic energy of $k' - U_i$.

In the photoelectric absorption process, a photon is completely absorbed by an atom and an inner-shell electron is emitted with an energy given by the incident photon energy minus its binding energy. The atom, left in an excited state with a vacancy in the ionized shell, relaxes via the emission of fluorescent photons and recoil electrons. In the latter case, the energy is transferred to another electron, which is ejected from the atom. This ejected electron is called an Auger electron. Immediately following this, an outer shell electron makes a transition to fill the gap in the inner shell. While photoelectric effect dominates in the low kV energy range where
cross section uncertainties are 5–10%, it is almost negligible for MV beams. In EGSnrc, the first step is to sample the atomic number of the element the photon is interacting with, followed by the sampling of the shell with which the interaction takes place. The probability that the photon is absorbed depends on the incident photon energy for shells other than $j$-th shell, which is energy independent. Once the element and its shell absorbing the photon is determined, a photo-electron with the kinetic energy $k-U_j(Z)$ is set up, where $U_j(Z)$ is the binding energy of the selected shell. The angular distribution of the photoelectrons directions is determined from the Sauter distribution [88].

Pair production is the creation of a subatomic particle and its antiparticle from a neutral boson (e.g., an electron and a positron created by a photon. In order for pair production to occur, the incoming energy of the interaction must be above a threshold (i.e., the total rest mass energy of the two particles) in order to create the pair. The quantum numbers of the produced particles must sum to zero, and thus the created particles should have opposite values of each other. In EGSnrc, the user has the option to use a library of differential pair cross sections created at the NRC. When the NRC pair differential cross section tabulations are used, an alias table is employed for picking the positron energy. The electrons and positrons are produced at a fixed polar angle $\theta$ with respect to the direction of the incoming photon. In addition to this, a leading order term of the angular distribution is employed to correct the particle angles.

2.2.2 Electron Interactions

Electrons lose energy as they traverse matter by two basic processes: inelastic collisions with atomic electrons and radiation. The incident electron may undergo a soft collision where a small amount of energy is transferred to the medium, or a hard collision with an orbital electron, which is then ejected from the atom ($\delta$-ray). Radiative energy loss, which occurs in form of
bremsstrahlung and positron annihilation, transfers energy back to photons and leads to the coupling of the electron and photon radiation fields. The bremsstrahlung process is the dominant mechanism of electron energy loss at high energies, while inelastic collisions are more important at low energies.

The expectation value of the rate of energy loss per unit of path length \( l \) by a charged particle with energy \( T \) in a medium is called its linear stopping power. Dividing the linear stopping power by the density \( \rho \) of the absorbing medium results in the mass stopping power \( (dT/\rho dl) \), expressed with of MeV cm\(^2\)/g. The energy losses due to inelastic collision and radiation are characterized by mass collision stopping power \( (dT/\rho dl)_{\text{col}} \) and mass radiative stopping power \( (dT/\rho dl)_{\text{rad}} \), respectively. The rate of energy loss for both collisional and radiative interactions depends on the electron energy and on the electron density of the medium. The collision stopping power is greater for lower electron energy and lower atomic number materials, while the radiative stopping power is approximately proportional to the electron energy and to the square of the atomic number. The mass collision stopping power is given by:

\[
\left( \frac{dT}{\rho dl} \right)_{\text{col}} = 2\pi r_e^2 Z N_A \frac{m_e c^2}{\beta^2} \left[ 2ln \frac{E_K}{I} + \ln \left( 1 + \frac{\tau}{2} \right) + F^\pm (\tau) - \delta \right]
\]  

where \( \tau = E_K/m_e c^2 \), \( \beta = v/c \), \( r_e \) is the classic electron radius, \( I \) is the average ionization potential of water and \( \delta \) is the polarization correction term. \( F^- (\tau) \) is derived from Möller cross section for electron-electron scattering and \( F^+ (\tau) \) is derived from Bhabha cross section for electron-positron scattering.

Incident electrons have high probabilities of experiencing nuclear elastic scattering. In elastic collision with a nucleus, the electron is deflected but does not radiate, nor does it excite the nucleus. The electron loses only insignificant amount of kinetic energy required for
conservation of momentum between the two particles. Though this is not a mechanism for the
transfer of energy to the absorbing medium, it is an important means of deflecting electrons. It is
the principle reason that electrons follow very tortuous paths, especially in high-Z media, and
that electron backscattering increases with atomic number.

Under the assumption of constant collision stopping power, negligible scattering and
negligible kinetic energy being carried out by δ-rays, the energy lost $E$ in collision interactions
by a fluence of $\Phi$ (charged particles/cm$^2$) of kinetic energy $T_0$ passing perpendicularly through a
foil of mass thickness $\rho t$ (g/cm$^2$) can be expressed by [89]

$$E = \Phi\left(\frac{d\tau}{\rho \Delta l}\right)_{col} \rho t \left(\text{MeV cm}^2\right)$$  \hspace{1cm} (2.3)

Therefore, the absorbed dose $D$ in the foil can be calculated by dividing the energy by the mass
per unit area of the foil

$$D = \Phi\left(\frac{d\tau}{\rho \Delta l}\right)_{col} \left(\frac{\text{MeV}}{\text{g}}\right)$$

$$= 1.602 \times 10^{-10} \Phi\left(\frac{d\tau}{\rho \Delta l}\right)_{col} \hspace{1cm} (\text{Gy})$$  \hspace{1cm} (2.4)

where $1.602 \times 10^{-10}$ is a fluence-to-dose conversion factors.

2.3 Modelling of Accelerator Head

The BEAMnrc code of EGSnrc provides a tool to accurately model a particle shower
using MC methods within a defined geometry. Geometries of the accelerator's components are
defined through component modules (CMs), where each part of the accelerator is considered a
single independent CM. In order to create an accurate Monte Carlo beam model of a linac, the
user requires precise geometrical and material information from the vendor. In a Monte Carlo
beam model, all components interacting with a beam must be included for an accurate simulation. For photon beams, the mean energy and focal spot size of the incident electron beam striking the target are key parameters, in addition to the geometrical and material composition of the exit window, target, flattening filter, and primary collimator [90]. An illustration of these components is shown in Figure 2.1.

![Figure 2.1 Illustration of the components modules of a typical Monte Carlo modelled linear accelerator head in photon beam mode.](image)

For MC simulation, the monitor ion chamber is an essential component for the calculation of the absolute dose since it scores the beam output of the linac. A relationship is built by converting the number of initial particles incident to the absolute dose and is achieved by Popescu et al. [91]. In this method, in addition to the forward radiation dose, the backscattered radiation from the secondary collimator collected in the monitor chamber is used in the calculation. This method allows a simple but accurate absolute dose calculation for complex
treatment planning techniques, such as IMRT and VMAT. For each field incident on the phantom from various gantry angles, the total absolute dose in any voxel is given by:

\[ D_{xyz,abs} = D_{xyz} \frac{D_{ch} (10 \times 10)}{D_{ch}} \frac{D_{xyz,abs}^{cal}}{D_{xyz}^{cal}} U \]  

(2.5)

where \( D_{xyz,abs} \) is the total absolute dose (in Gy) deposited in the \( xyz \) voxel; \( D_{xyz} \) is the normalized dose (in Gy per incident particle) deposited in that voxel; \( D_{ch} \) is the normalized dose (in Gy per incident particle) accumulated in the monitor chamber while 10×10 refers to a specific field size; \( cal \) means dose is scored under calibration conditions, where the ratio of \( D_{xyz,abs}^{cal} \) (in Gy per MU) and \( D_{xyz}^{cal} \) (in Gy per incident particle) determines the number of electrons incident on the target for a photon beam calibration run of 1 MU; \( U \) is the number of monitor units. Following the conversion into absolute dose, the dose distributions of all the fields are summed to provide a total plan dose.

For a given treatment, the MLC motion may be static or dynamic. The latter is commonly used in VMAT. Explicit transport is the most accurate method of modelling the MLC since a particle shower is transported through a detailed MLC geometry. In EGSnrc, the collimator has a single layer composed of a user-specified number of leaves. The collimators are defined when all individual leaves are open. Apart from the position in \( z \) direction and the width, the leaf-end or the tongue-and-groove effects are also taken into account for each type of leaf. For example, for the Varian high-definition MLC, the MLC is divided into an upper region and a lower region, split at approximately the mid-thickness point of the MLC. The MLC leaves move in the \( x \) direction and the \( y \) coordinate is perpendicular to the direction of leaf motion. The users must
define leaf cross section data and leaf opening position. A more detailed explanation of the MLC
geometry and required input is provided in the BEAMnrc manual [86].

2.4 Phase Spaces (phsp)

In EGSnrc, particles transported are stored in a phsp with their parameters, such as the
type of the particle, energy, position and direction. Two types of phsp file formats can be read
and scored in EGSnrc: the native EGSnrc format and the IAEA format [92], which is compatible
with all major MC codes. While EGS phsp always sets the incident $z$ position equal to zero, the
IAEA format permits the storage of nonplanar phsps ($z$ position can be defined for each particle).
This feature allows an IAEA phsp to handle 3D phsp data and thus greatly enhances the
flexibility of DOSXYZnrc. A particle's history is defined from the beginning of its transport until
both the original particle and its progeny are absorbed or exit the volume of interest. A phsp file
contains these data for every particle crossing a scoring plane. Phase space files can be used as
output for each scoring plane in an accelerator (e.g., the back surface of the last CM in Figure
2.1) or as input source for a new simulation. A phsp is usually scored for specific treatment unit
and beam energy to increase efficiency, and thus can avoid the re-simulation for treatment unit-
specific part of accelerator.

Creating an accurate beam model in MC requires the geometric and material details of
the linear accelerator. This information is shared by manufacturers under non-disclosure
agreements with institutions conducting research. However, for the TrueBeam linac, Varian no
longer provides the linac head information for the purpose of MC research. Instead, they release
IAEA phsp data for the upper linac components [93]. The phsp plane is located above the jaws.
The phsp file of the 6 MV beam contains $1.1 \times 10^9$ original histories and $5.2 \times 10^7$ particles.
Detailed data are reported in the header file, available in www.myvarian.com/montecarlo. The source parameters of their model are tuned to match the beam data for Eclipse TPS [94].

2.5 Simulation through Patient-specific Phantom

EGSnrc performs the dose calculation in a voxelized phantom and writes the deposited dose in a 3ddose file. A phantom (i.e., egsphant) is constructed using a 3D matrix of voxels where each voxel contains various pre-defined physical density and material assignment (e.g., H2O, soft tissue, lung and bone). There are two ways by which this matrix can be created, by either DOSXYZnrc or ctcreate code. The former way is by defining a set of $x$, $y$, and $z$ boundaries and then assigning densities and material types to these voxels and the second way is used for simulations on patient geometries from the CT image set of patients. While one can easily customize the phantom geometry with DOSXYZnrc, this option does not provide patient anatomy. Therefore, ctcreate is required to obtain a CT phantom from the CT data set. From this set of images, the CT values in HU are interpolated onto a 3D matrix of voxels and converted into equivalent physical densities.

The conversion from CT densities to physical densities is achieved through interpolation of CT value-to-mass density relations forming a CT ramp. This calibration curve for the conversion of CT number to density and CT number to medium have to be supplied by the user. For each CT images, the ctcreate code converts the CT value to the corresponding material and mass density. The data are then written into the egsphant file, along with the number and a list of media present in the patient model and the boundaries of voxels in $x$, $y$ and $z$ direction. However, only the most common material types of patient body and medical devices can be included in this file.
The CT image set is written in Digital Image and Communications in Medicine (DICOM) format and can be exported for MC simulation. The image resolution is usually set to 512×512 pixels, where the spacing of the pixels is determined by the size of the area that is imaged. The resolution of CT slices is determined by the volume of the prescribed region to be imaged and is usually set to 2.5 mm. The raw CT images, however, contain the couch of CT scanner. This structure can lead to incorrect modelling of the phantom if left in the CT image since the material and geometry of this couch is not the same as treatment couch of linac. To eliminate such structures from the CT images, a program is written to assign all pixels outside of the defined contour of patient body to air.

2.6 Synchronized Simulation and Sources 20, 21

EGSnrc uses a source routine (e.g., source 20, 21) to describe the radiation source of particle histories. The type of source can be a point source, a parallel beam or a phase space. Standard BEAMnrc and DOSXYZnrc simulations only model the static setup with a fixed beam incident direction. However, to simulate varying orientation, collimation or isocentre, such as those in VMAT, a four-dimensional (4D) MC model is necessary. To generate a 4D model, source 20 in DOSXYZnrc and source 21 along with the synchronized CM (i.e., SYNCJAWS and SYNCVMLC/SYNCHDMLC) in BEAMnrc developed by Lobo and Popescu [95] are used. The two sources support a continuous gantry motion, collimator rotation, couch rotation and translation in any direction, variable MU rate, arbitrary isocentre motion with respect to the patient, and variable SAD, as well as dynamic MLC motion and dynamic jaw motion. The models allow the users to simulate continuously variable beam configurations and complex treatment geometry and kinematics with respect to the patient. As a result, MC simulation has
the capability of representing as realistically as possible the beam delivery and dose deposition for a VMAT treatment. The source 21 uses a full BEAMnrc simulation as an input, and thus the resulting phsp of source 21 is usually employed as input to source 20. The geometrical parameters controlling the orientation of the source plane relative to the phantom have the same definition as those for sources 2, 8, 9 and 10, where a detailed explanation can be found in DOSXYZnrc Manual [87].

Unique to these sources is the synchronization between the motion in the BEAMnrc shared library (i.e., the motion of the jaws, MLC) and the motion in the DOSXYZnrc geometry. This synchronization is indispensable in order to perform single-run MC simulations of technologies involving dynamic MLC motion and dynamic jaw tracking of the maximum MLC aperture, as well as gantry and collimator rotation. For instance, a set of SYNCJAWS is synchronized with a set of SYNCVMLC/SYNCHDMLC in the accelerator and with the motion of source. The geometry settings of the synchronized CM are specified in an independent file. To determine the weight of each control point in VMAT, the arc distribution is sampled through a randomly sampled dimensionless time scale, or MU scale (i.e., MU index or muIndx in EGSnrc code) between zero and one. The MU index is added to the particle record in the IAEA phsp when sources 20 and 21 are used. (The IAEA format is the only one allowing for this type of variable.) At the beginning of each primary history, another random number, MU RND, is compared to the muIndx and used to determine the incident geometry settings. The MU RND is generated by the first synchronized CM (usually SYNCJAWS) and is passed on to any downstream synchronized CMs, and thus the value of MU RND used by SYNCJAWS for each primary history is the same as that used by synchronized MLCs in the accelerator. This MU
RND is finally passed on to DOSXYZnrc to allow the dynamic motion of the source plane to be synchronized with the accelerator head.

Moreover, these sources allow the user to interpose a shared library, generated using either a BEAM accelerator or a non-EGSnrc code (e.g., particleDMLC for photon beam MC IMRT dose computation [96]), between the source plane and the DOSXYZnrc phantom. Thus, the particle data at this intermediate plane are not actually output but are stored in an internal array for eventual DOSXYZnrc transport through the phantom. The shared libraries are dynamically loaded by DOSXYZnrc at run time to deliver particles for the simulation of patient dose deposition. Consequently, source 20 eliminates the need for intermediate phsps below the MLC, regardless of the plan complexity. A typical shared library applied is synchronized MLC (i.e., SYNCHDMLC or SYNCVMLC). Due to the absence of phsp files below the MLC, a full, continuous VMAT beam delivery can be simulated with source 20 in a single DOSXYZnrc run, with only one 3ddose file being generated. Therefore, simulations of complex continuous beam delivery with sources 20 or 21 require no more computing resources than a single static simulation.

One issue when running source 20 through a shared library is the determination of how many times to recycle incident particles so that the phsp source does not rewind, causing loss of information about particle correlations and subsequent underestimation of dose uncertainties. DOSXYZnrc automatically estimates the number of times to recycle each particle, NRCYCL, based on the number of particles in the phsp source, the number of histories requested, and the charge of particles requested. However, interactions in the shared library geometry change the number of particles reaching the DOSXYZnrc phantom per particle incident from the phsp source. Thus, the original estimate of NRCYCL is incorrect because the beam is blocked by the
moving MLC. To overcome this when a BEAM shared library is used, an initial calibration comprising $1 \times 10^6$ particles run through the shared library geometry only is used to determine the ratio of the number of particles emerging from the bottom of the BEAM library geometry to the number of incident particles (survival ratio). Typical values for the survival ratio are 5% - 10%, depending on the beam modulation. NRCYCL is then recalculated by multiplying the original estimate by $1/$survival ratio. Only the particles that survive passing through the MLC are recycled, as they are transported through the patient CT phantom.

2.7 Monte Carlo-based Inverse Treatment Planning

The widespread clinical implementation of VMAT in the modern cancer treatment has led to a requirement of highly accurate dose calculation for heterogeneities and small fields. For example, the main aim of lung SBRT is to deliver high dose to the target with sub-millimeter positional accuracy and less than 3% dose accuracy along with sharp dose gradient outside the PTV. In the case of small field geometries, the absorbed dose changes rapidly with field size and depth due to the lack of both lateral and longitudinal electronic equilibrium when the field size is smaller than the maximum range of secondary electrons [97]. In addition, introducing a heterogeneity, such as lung (which has low density), inside such small fields makes the dose calculation more difficult to predict for model-based algorithms [98, 99]. Hence, the choice of treatment planning algorithms to predict the dose in small field and low density medium and the detector to verify the planned dose is important to achieve the results. It is proven that model-based algorithms (i.e., AAA) show more deviation with the beam aperture size less than $3 \times 3$ cm$^2$, regardless of delivery techniques [100]. A study was also performed in this thesis to evaluate the difference between dose distribution given by MC calculation and AAA calculation.
on a water-bone-water and a water-lung-water interface phantom (Appendix B.1). Shortcomings of the dose prediction could lead to dosimetric errors, which often would lead to errors in the implementation in a planning system and serious discrepancies between the planned and delivered dose to the patient. The accuracy of dose calculation can be improved if the TPS uses algorithms where reliable modelling are included to track every secondary scattered photon and electron and its further dose deposition in non-equilibrium situation. Taking into account the detailed geometry of the treatment head and the lack of electron equilibrium in small fields due to an accurate electron transport, the MC algorithm can bring remarkable accuracy to the treatment planning. However, the rigorous dose distribution results at the trade-off of longer computation times. This (central processing unit) CPU intensive nature of MC simulation remains as the most significant challenge for the use of MC dose calculation in inverse planning process [84].

The idea of MC-based inverse treatment planning system was first developed by Jeraj et al. in 1999 [101]. In their method, the energy distribution of photons was first specified in the tumour volume performed by either backward or adjoint transport. All particles exiting the patient geometry were scored corresponding to the positions and directions at which particles started in the forward transport stage. This built up an intensity distribution, which was used as the initial input into an optimization algorithm. The dose distribution from each beamlet that made up the fields was then calculated using Monte Carlo transport. A quadratic cost function based on dose criteria was used to evaluate the treatment plan. A fast simulated-annealing optimization [102] was used to determine weighting of each beam element, and subsequently, the dose distribution that best satisfied the criteria and constraints specified by the objective function. They further investigated the statistical uncertainty in MC-based IMRT and concluded
that the noisy calculation used to determine optimum weights could lead to noise convergence error [103]. The feasibility of this study was then demonstrated with a prostate case in realistic patient anatomy by a different group [104].

Despite the effectiveness, direct storage of beamlet intensity and weights distributions for IMRT treatment planning can easily exceed more than ten gigabytes, and is therefore often not feasible. Several strategies can be applied to improve the computation efficiency by using fast approximate MC dose algorithms (e.g., VMC++ code) or reducing the number of voxels at which dose must be computed through. The second approach is most commonly achieved by only including voxels in region-of-interest (ROI) (which however can potentially push hot spots into the rest of the body) or using large voxels, such as 0.5×0.5 cm² at the cost of resolution. Zakarian et al. [105] proposed a method to compress and store the beamlet data for rapid iterative dose computations based on wavelet compression. The beamlet dose distributions were wavelet transformed and compressed by dropping wavelet coefficients below a given threshold value. Dose was then computed using the remaining wavelets. Their overall compression performance was on the order of 100:1.

Dogan et al. [106] proved that combining MC and superposition convolution algorithm reduced the systematic dose prediction and optimization-convergence errors, where their plan optimized with the MC was based on the intensities from deliverable superposition convolution as an initial guess. Bergman et al. [107] assessed the combination of Monte Carlo tissue inhomogeneity modelling and DAO inverse planning and concluded that the introduction of DAO resulted in an increase of the treatment delivery efficiency. An in-house algorithm was developed to subdivide the phsp into 0.25×0.5 cm² beamlets, where each beamlet is projected onto a patient-specific phantom. The MLC leaf positions were mapped to the location of the
beamlet and were optimized using DAO. This method generated an acceptable plan that provides adequate coverage to the PTV. The time required to obtain and run the separation of a single phsp into beamlet phsps of $5 \times 10^6$ particle histories on a single CPU can be achieved in less than eight seconds [108]. The beamlets were then transported separately into the phantom and produce separate dose distributions. Bogner et al. [109] introduced the concept of pre-calculated MC inverse kernel using XVMC in a single simulation. The inverse kernel was generated for each beamlet in all ROIs. The fluence-weight optimization was accomplished by means of an iterative search engine. A final dose calculation was performed as a re-optimization of the segment weights to gain improvement in plan quality. The DAO method was then applied in their later research to overcome the loss of treatment plan quality after segmentation in IMRT [110]. However, while MC-based IMRT has achieved promising results in terms of dosimetric plan quality and computation efficiency, the approaches discussed above have not been applied to VMAT, due to the continuous beam-delivery feature.
Chapter 3: A Monte Carlo Inverse Treatment Planning Algorithm for Trajectory-based VMAT

3.1 Introduction

We present a full MC-based inverse treatment planning method for trajectory-based volumetric modulated arc therapy (TVMAT). TVMAT uses continuous and simultaneous gantry and couch rotation to avoid the OARs in the beam trajectory, and thus it reduces the radiation exposure of normal tissues. The dynamic gantry-couch trajectory allows for more degrees of freedom in the design of a treatment plan, and thus a higher dosimetric quality compared to standard VMAT can be achieved [111-116]. MacDonald et al. [114] quantified the geometric overlap between PTVs and OARs based on their 2D projection from source to a plane at isocentre as a function of gantry and couch angle. A navigable ideal trajectory for the patient specific couch-gantry space was then generated. The optimized treatment trajectories decreased both the mean dose and maximum dose to the OARs by 19% and 12%, respectively, though a slight degradation (an average of 6%) was seen for PTV homogeneity and conformation. Wilson et al. [115] sampled a $4\pi$ (solid angle space) geometry where the couch slowly rotated through $180^\circ$ while the gantry continuously swept within (a maximum of) eight partial arcs. Both MLC and dose rate were optimized through an inverse planning framework. Improvements in dose conformity were observed in ten cranial SRS patients. Langhans et al. [116] introduced an algorithm called NoVo (non-coplanar VMAT optimization), which determined the simultaneous trajectories of the gantry and the couch by first computing beams from every possible direction and then eliminating beams by examining fluence contributions. The NoVo method reduced the
objective function value for the lung, brain and liver cases, compared to coplanar VMAT. However, while the planning of non-coplanar beams with static couch positions is readily available, that of dynamic trajectory is not accessible in commercial treatment planning systems for conventional linacs (e.g., TrueBeam). Currently, no commercial TPS has the capability of calculating the fluence as MLCs sweep from one position to another during the inverse planning (i.e., the optimization of continuous beam trajectory). On the other hand, MC techniques are particularly well suited to modelling dynamic source and MLC movement due to the particle-by-particle calculation framework. This is a major difference compared to fluence source models used in convolution/superposition methods [12, 13, 20, 21] or Acuros XB [26].

VMAT planning uses fast model-based dose calculation algorithms (e.g., AAA) during the optimization process. Nevertheless, the approximations used in dose calculation algorithms utilized in VMAT optimization affect the optimality and accuracy of the treatment plans, introducing dose-prediction errors and optimization-convergence errors. The dose-prediction error is caused by the difference between the calculated dose and the measured dose. The optimization-convergence error is caused by the difference between the deliverable optimized plan and the ideal optimized plan. It has been shown that a full MC-based optimization greatly reduces both kinds of errors [106, 107, 109]. However, although these published methods could effectively reduce the total computation time by generating MC-based beamlets or introducing hybrid dose calculation algorithms, MC-based optimization was still far more computationally intensive per iteration than model-based optimization. Consequently, these previously published approaches to MC-based optimization have not been clinically implemented, and none has been proposed for VMAT or TVMAT so far.
There are several important differences between the method presented in this thesis and those previously published in the literature [113-116]. While the gantry is rotating continuously around the patient, the current commercial dose calculation of this treatment is based on the static beam approximation, where the continuous gantry motion sweeping a whole arc of 360° is represented by e.g., 178 static beams (such that a 2° spacing between adjacent control points can be achieved). However, the dose is calculated based on the static approximation of the dynamic trajectory, where each control point is considered as a static field. As a result, the optimization is applied to the aperture and weight defined for individual control point. A study by Christiansen et al. [117] showed that the integration of continuous aperture fluence calculation in VMAT optimization process greatly reduced the dose errors resulting from discrete aperture approximation, from more than 10% to less than 1%. By contrast, the dose calculation of continuous beam delivery can be achieved using sources 20 and 21 in EGSnrc [95], which support the modelling of trajectory motion and various dose rates.

In this work, we have developed a method that can accurately represent the actual dynamic MLC motion and gantry-couch trajectory of the continuous beam delivery at all stages of the optimization, instead of a static beam calculation at each control point to approximate the continuous motion, as in traditional methods. We take advantage of the particle-by-particle calculation framework of MC and its accurate modelling of the electronic disequilibrium. Dose optimization is performed in a single MC simulation, thereby greatly reducing calculation time. The TVMAT MC-based optimization is also applicable to standard VMAT delivery with fixed couch position. It should be pointed out that while the techniques in this thesis were developed using Varian systems (TrueBeam linac, Eclipse TPS), they are in principle applicable to other vendor solutions, as well.
3.2 Methods

The workflow of the MC-based inverse planning algorithm for VMAT (i.e., MC-VMAT) is illustrated in Figure 3.1. Briefly, the steps involved are as follows. The treatment planner selects the initial trajectory (i.e., the range of the gantry, collimator and couch angles) using the method described in [115] and creates a base plan (i.e., non-optimized initial configuration). In this plan, the initial set of leaf positions corresponds to a dynamic beam conformal to the target. The single MC simulation required starts from an IAEA phsp scored at the top of the SYNCVMLC or SYNCHDMLC modules of BEAMnrc and uses source 20 of DOSXYZnrc [95], which was previously validated by measurements for trajectory-based beam delivery. We modified DOSXYZnrc in order to generate 4D dose files (i.e., ‘edep’ files) that score individual, time-stamped, energy-deposition events in the voxels of the ROI. Consequently, a relation is established between the space and time (i.e., MU index) coordinates of source particles in the phsp and their contribution to energy deposition.

Figure 3.1 The workflow of MC-based optimization for TVMAT.
A direct aperture optimization with a dose-volume based quadratic objective function is performed using an in-house code [44], with the additional improvement that we took into account the continuous movement of the MLC, gantry and couch between adjacent control points. A detailed explanation of the preparatory steps involved is available in Appendix A including the initial beam setup, modelling of the linac and the patient, the selection of the ROI and optimization constraints, and considerations about the statistical uncertainties and the optimization time.

3.2.1 Reading and Writing 4D Dose Data

MC considers the continuous motion of MLC leaves, and thus can reflect the actual beam trajectory. However, current DOSXYZnrc (version V4-2.4.0) does not store intermediate energy depositions of each particle history. Those intermediate energy deposition events are essential to generate 4D dose files. We modified DOSXYZnrc to generate a pair of files (edepheader and edepdat) for each sub-process of a MC parallel simulation. These files score individual, time-stamped, energy deposition events. A template of edepheader and edepdat files are shown in Figure A.2 and Figure A.3. The x and y positions of the source particles in the first patient-specific phsp (i.e., phase space B in Figure 1.2) are recorded, which can be transferred to either DOSXYZnrc for dose calculation or to the subsequent phsp for further simulations. In addition, the random MU index and the number of voxels in ROI (i.e., PTV + OARs + dose fall-off region), in which the particles deposit energy, are also stored in edepheader files. In edepdat files, the value of the energy deposition (edep) and the index of voxel this event taking place are scored.

The saved edepheader and edepdat files are then used to reconstruct both 4D dose distribution and dose uncertainty after the simulation. A dose matrix is established for the 4D
dose distribution, whose time coordinate is related to the cumulative MU. The randomly assigned MU index of the incident particle is partitioned into a bin, where the bin edge is based on the MU index of each control point. Subsequently, the edep data are summed up for each voxel in ROI, in each MU index bin. The 4D matrix is further summed up to create the initial 3D dose distribution by converting the edep to dose using:

\[
dose_i = \frac{1.602 \times 10^{-10}}{rh_\text{on}_i \times amass_i \times a\text{influ}} \times \Sigma \text{edep}
\]  

(3.1)

where \( rh_\text{on}_i \) and \( amass_i \) are the mass density and volume for voxel \( i \), \( a\text{influ} \) is the total number of incident particles from original source.

### 3.2.2 Dose-volume Objective Function

The objective function is designed to include dose prescriptions to PTV and OARs, conformity, and dose fall-off. Doses to OARs are represented as user-assigned dose-volume constraints and relative importance factors. The calculation of the cost is conducted using the following equation:

\[
\text{Cost} = \left[ \frac{1}{2} + \frac{1}{4} \min(CI_{MC} - CI_{goal}, 0) \right]^2 + \frac{1}{4} FOMC^2 \cdot \Sigma_{j=1}^{S} \Sigma_{i=1}^{n(j)} \left( \frac{p(j)}{n(j)} (D_{i,MC} - D_{i,goal}) \right)^2
\]

(3.2)

where \( CI_{MC} \) is the conformity index calculated by Equation 1.1; \( CI_{goal} \) is the optimal conformity index and is set to 1.0; \( n(j) \) is the number of voxels contained in a particular structure \( j \) and \( p(j) \) is the weight of priority for that structure, indicating the importance of the structures (e.g., 120% for the PTV). \( FOMC \) is the calculated dose fall-off and is defined as

\[
FOMC = \frac{V_{FO_{80\%}}}{V_{FO}}
\]

(3.3)
where VFO is the volume of the fall-off region and VFO\textsubscript{80\%} is the VFO region that receives 80% of the prescription dose. This fall-off function is designed specifically for the ROI dose recording method, where limited body dose are stored. This formulation of the cost function allows treatment planner to control the relative importance of CI and FO objectives, while the dose-volume constraints are of unconstrained magnitude. The dose-volume objective function to OAR and PTV is assigned a total weight of 1/2 of the cost, while conformity and dose fall-off are each assigned a weight of 1/4 of the cost of the dose-volume objective function. These weights are assigned so that the combined contribution of the conformity and the dose fall-off objectives cannot exceed that of the dose-volume objectives. The homogeneity index (HI) is also calculated using Equation 1.2 at the end of the optimization to analyze the uniformity of dose distribution in the target volume. Apart from this, a statistical uncertainty constraint independent of the dose-volume constraint is also employed, where the maximum statistical uncertainties allowed for each voxel in ROI during the entire optimization is 2%, according to the AAPM Task Group No. 65 [118] and Task Group No. 157 [119].

3.2.3 MC-based Inverse Planning Algorithm

The sources 20 and 21 support the continuously variable beam configurations. In the most simplified case, a whole arc can be represented by only two control points, providing that all beam geometries can be linearly interpolated. In contrast to the AAA, where each sample only defines a static field, in MC, even with a coarse sampling, the trajectory motion through the whole arc is still properly modelled through the entire range among all the samples. Therefore, the MC-based optimization starts with a relatively small number of samples and progressively increase the number of control points. The beam configuration is first resampled based on the MU index, from 0 to 1. The resample resolution is on the order of 0.0001 (i.e., $\Delta \text{muIndx} =$
The MLC leaf positions are then linearly interpolated according to the resampled MU index, which are used in the following direct aperture optimization.

Two kinds of iteration are introduced in the optimization and need to be clarified here. The first kind of iteration is any change in leaf position or MU weight, while the second kind of iteration is the progressive sampling of control points, where the number of control points increases in each iteration. In each iteration of leaf/weight change the optimizer must randomly select the control point to optimize, the aperture weight or the aperture shape to adjust, the MLC leaf within that field to optimize, and the amount by which the weight or the position of the leaf changed. The individual MLC leaf positions defining the aperture shape and the aperture weights are optimized using simulated-annealing to minimize the dose-volume objective function described in Equation 3.2. If a proposed change does not violate a mechanical or dose rate constraint, the dose distribution and cost function are calculated. The change is accepted if the cost reduced, otherwise it is rejected. In addition to the above stochastic search scheme of leaf/weight, a simple annealing algorithm (with a fixed probability of 1%) is applied to allow the increase of the cost so that the optimization would not be trapped in a local minimum.

At current stage, the final MC-VMAT plans undergo a full forward MC calculation. In this simulation, all the leaf tips and the tongue-and-groove effect are considered, which are simplified during the optimization. The differences for the PTV and the OARs are evaluated between the dose distributions given by the final MC calculation and the MC-based optimization.

### 3.2.3.1 Linear Sampling of MLC Leaf Changes

One of the key features of the standard VMAT algorithm is the mechanism used to sample the dynamic source motion by a finite number of static beams. The instantaneous MLC configuration is defined at each sample, and the MU setting for the sampling interval is assigned
to that static MLC configuration. Clearly, the accuracy of the delivered plan depends on the amount that MLC leaf positions change and the range that the gantry rotates in between each sample. While in standard VMAT planning each control point is considered one sample that defines one beam field, a significant difference in our MC-based optimization is that each sample is defined as the motion bounded by two control points. Therefore, the number of control points ($N_{ctrl}$) equals to the number of samples ($N_{sample}$) plus the number of beam arcs ($N_{arc}$), i.e.,

$$N_{ctrl} = N_{sample} + N_{arc}$$

Consequently, instead of modelling the moving linac source as a series of static source position samples, this MC-VMAT optimization models the actual continuous source motion, where the MLC leaves and gantry move linearly in between each sample.

Samples are included at the beginning and end of the range with evenly distributed samples in between. However, since each sample is represented by two control points, the aperture shape in each control point is no longer independent and is related to two sampled segments (with the exception of first and last control points). Therefore, the change of MLC positions in a certain control point must be applied to both samples. Once all related samples are determined, the MLC positions in between are linearly interpolated based on the MU index. For each $\Delta muIndx$, energy depositions of the particles blocked by interpolated leaf change are removed from the 4D dose matrix (Figure 3.2) and replaced with the dose distribution pre-calculated in the background plan in which all MLC leaves are closed (Appendix A.1). This procedure is referred to as ‘particle removing’. It is worth mentioning that, during the particle removing, all the $edep$ correlated to the blocked source particles are removed, even when these particles scatter to other leaves. The new 3D dose in ROI is then summed up from all 4D dose
matrix. After each iteration of successful leaf change, the dose distribution is renormalized. The dose can be normalized to either the mean PTV dose equal to the prescription dose or a user-defined coverage specific isodose, e.g., 95% volume of PTV is covered by the 98% isodose line.

Figure 3.2  An illustration of the MC particle removing. The left figure shows the initial conformal plan where all the particle depositing energy are recorded. An attempted leaf change is made in the right figure (dash line of the leaf). The energy contributions of the particles blocked by the new leaf position are removed.

3.2.3.2  MU Weight Changes

Similar to the leaf position changes, the MU weight changes are also applied to the sample defined by two control points. The maximum MU rate for a TrueBeam linac is 600 MU/min in the standard 6 MV mode. The initial weights of all samples are set to one and a modification step no larger than 0.3 is allowed in each iteration. While a larger step (e.g., 0.5) can result in a higher probability of being rejected, a smaller step (e.g., 0.1) increases the number of changes attempted. Therefore, 0.3 was found to be a practical choice. Nevertheless, this maximum allowed weight change is only valid for standard beams, and thus it should be re-defined for flattening filter free (FFF) beams, due to their higher MU rate. The modified beam weight is then multiplied by the dose of the selected sample as the input for the cost function.
However, while the beam weights are changed in each iteration, the new MU indices are not updated correspondingly, but recalculated at the end of the optimization. Therefore, the relationship between the gantry angles and the MU indices are not perturbed. The updated MU indices and MU are acquired by

\[ \text{MUsequence} = \text{diff(MUindex)} \times \text{NewWeight} \]  \hspace{1cm} (3.5)

\[ \text{NewMUindex} = \text{cumsum(MUsequence}/\text{sum(MUsequence)}) \]  \hspace{1cm} (3.6)

\[ \text{NewMU} = \text{MU} \times \text{sum(MUsequence)} \]  \hspace{1cm} (3.7)

The changes in MU indices are then taken into account for the calculation of final beam outputs.

### 3.2.3.3 Progressive Sampling of Gantry and MLC Positions

At the start of MC-VMAT, a relatively coarse sampling of the gantry positions is used to model the gantry rotation range. After a few number of successful MLC and/or MU weight changes, an additional set of samples is added to the pool of optimizable gantry positions. The new samples are added midway between two existing samples. In MC-VMAT, the new sample is not added one by one, but added once as a complete set that spans the full range of gantry motion. As a result, in each new iteration of sampling, the number of new samples is the same as the existing samples and the total \( N_{sample} \) is doubled. Each time a new sample is added, the algorithm continued to optimize both the previous beam samples as well as the newly added sample. Samples are continuously added in this way until a desired sampling frequency is attained. For example, for a typical VMAT plan, the \( N_{sample} \) begins at 11 and doubles in each following iteration (i.e., 22, 44, 88 and 176). One advantage of MC-VMAT is that, for the leaf positions and MU weight of the new samples, no additional calculations are required. Since the changes of leaf position and weight are not just applied to the selected control point, but to all
related samples in a wider range, the aperture and MU weight of the new sample can be directly taken from the existing samples. Consequently, uniform sampling intervals are retained during the introduction of new samples. The number of iterations can be preset by the planner and is usually related to five steps of progressive sampling for a typical VMAT plan, which allows 176 samples (approximately 2° spacing between adjacent control points) at the end of the optimization.

A gradient descent algorithm is used as a checkpoint for the progressive sampling. A user-defined threshold, referred to as ‘optimization gradient criterion’, is used to control the progress of the optimization. The local minimum cost value is recorded every 5 seconds and the first order derivatives of the recorded values are calculated. The algorithm automatically enters the next iteration or terminates if all first order derivatives recorded in the last minute are lower than the gradient criterion. This criterion decreases exponentially for each new iteration. While the smaller threshold can result in lower cost value and thus better plan quality, it increases the computation time. The criterion is set to 0.2 in order to compromise between the plan quality and the computation time. Consequently, while the optimal solution found by gradient descent method does not necessarily represent the global minimum, this algorithm can achieve a solution close enough to optimal within time comparable to the commercial TPS.

The MLC leaf position change and MU weight permitted between samples are directly proportional to the gantry rotation range between samples. For instance, the maximum possible leaf change starts at 20 mm and is halved for each following iteration. When a coarse sampling of gantry positions is used, there is a larger gantry range between samples. Therefore, a coarse sampling of gantry positions allows for a larger maximum MLC leaf displacement and MU weight change at any iteration, thereby providing more flexibility in deriving a desired plan.
Conversely, as new samples are introduced, the maximum possible change decreases and the flexibility of the optimization reduces. The overall effect is that VMAT optimization is effectively not restricted at the start but gradually becomes more restricted as new samples are introduced. A similar characteristic related to the plan accuracy is also observed. Due to the small amount of samples at the beginning of optimization, the accuracy of the optimized plan with respect to the plan that will be delivered is relatively poor. As new samples are introduced, the sampling frequency increases, thereby improving the modelling accuracy of dynamic MLC and gantry motion.

3.3 Results and Discussions

The method proposed here is applicable to standard coplanar VMAT, non-coplanar VMAT and trajectory-based VMAT, and has been performed on eleven patients (case a-k) with various tumour sites, including head and neck, lung, prostate and brain cancers. Three examples (case a, g and h) are selected for the following discussion and other cases are presented in Appendix B.2-B.4. The optimization time varied from 5–90 minutes, depending on the number of voxels in the ROI and the complexity of the dose-volume constraints (Figure 3.3).
Overall, the developed MC-VMAT method has the capability of producing plans that meet prescribed clinical dosimetric indices. However, the purpose of this study is to present the principle of MC-VMAT, and thus the discussion here will focus on the optimization strategy and concerns specific to MC method, but not the dose distribution comparisons to the VMAT plan created using Eclipse TPS (AAA-VMAT).

3.3.1 Head and Neck Cancer — Standard Coplanar VMAT Example

3.3.1.1 Principle of MC-VMAT

An arbitrary head and neck cancer example (case a) was selected to show the difference in the DVH between the MC simulation and the AAA model based on the dose calculation of a clinical VMAT plan. The plan was comprised of 228 control points in two arcs using 6 MV
photon beam. The PTV was prescribed with 24 Gy dose. The optimization objectives for this

case are listed in Table 3.1.

Table 3.1 Optimization objectives for a head and neck cancer example.

<table>
<thead>
<tr>
<th>Structures</th>
<th>Optimization Objectives</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Volume (%)</td>
<td>Dose (Gy)</td>
</tr>
<tr>
<td>PTV</td>
<td>0</td>
<td>≥ 26</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>≤ 23</td>
</tr>
<tr>
<td>Spinal Cord</td>
<td>0</td>
<td>≥ 2.4</td>
</tr>
<tr>
<td>Chiasm</td>
<td>0</td>
<td>≥ 4.5</td>
</tr>
<tr>
<td>Brainstem</td>
<td>0</td>
<td>≥ 3.6</td>
</tr>
</tbody>
</table>

Both MC and AAA calculated the dose distribution with a 2.5 mm resolution. While the

dose showed almost identical dose distribution for both OARs (i.e., brainstem and chiasm), the
MC calculation revealed a worse PTV coverage compared to that of the AAA (Figure 3.4). The
MC and the AAA calculated a mean dose of 23.0 Gy and 23.9 Gy in the PTV, respectively. In
addition, the MC simulation indicated that only 77.9% of the volume of PTV was covered by
95% of the prescription dose (i.e., V95%dose = 77.9%), in contrast to the AAA result, where
V95%dose = 93.2%. The comparison suggests that while the AAA has been proven to be a
reasonably accurate and efficient dose calculation model, it may overestimate the target volume
and result in insufficient dose coverage.
The same clinical plan was adopted to show the feasibility of the MC optimization, where the above MC dose calculation result (i.e., final MC calculation) was compared to that calculated by the method described in section 3.2.3.1 (i.e., ‘particle removing’), which employed the same principle as the MC-based optimization. The process started with a plan conformal to the PTV. The energy deposition of the particle histories blocked by the MLC positions in the clinical VMAT plan were then removed and replaced with the energy deposition in the background plan, while those particles that went through the MLC leaves remained contributing to the final dose distribution (Figure 3.5).
Figure 3.5  Procedure of the particle removing at an arbitrary segment. The left figure shows the particle distribution of the initial conformal plan while the right figure shows the particle distribution of the plan with optimized MLC leaf positions. The particles blocked by the optimized leaves are removed and replaced with those in background plan.

Excellent agreement was observed in terms of the PTV dose (Figure 3.6). The dose in the chiasm, however, was slightly different between the two calculation methods. This is likely because the ‘particle removing’ approach simplified the complex geometry of the MLC leaves shape as slab blocks. In addition, while the adoption of the background plan during the optimization procedure can compensate the missing energy deposition due to leaf leakage, the model underestimates the attenuation and the scatter of the photon at the leaf tips, and consequently, overestimates the dose in OARs.
The overestimation of the OAR doses was further investigated through the calculation of the 3D gamma map between the dose distribution given by the forward MC calculation and the ‘particle removing’ method. The criteria of the gamma map were set to 3% of the prescription dose and 3 mm DTA. A dose mask was applied as threshold where the dose lower than 30% isodose were neglected in the gamma evaluation. However, it is worth noting that the gamma map and dose mask here only took into account the dose in ROI region but excluded the dose in the rest part of the body. The gamma map showed a pass rate of 89.3% (i.e., 89.3% of voxels had $\gamma \leq 1$), where the voxels that failed to pass were mostly located at the edge of the field belonging to the OARs (Figure 3.7). The resulting gamma map confirms the above hypothesis that the simplification of modelling of the leaf tips brings about the overestimation of the OAR doses. Therefore, a correction will be applied to overcome the leaf-tips effect and will be the next step.
of this MC inverse planning algorithm project. However, for the work already done here, this
effect remains in all the dose calculation and will potentially cause difference in terms of the
OARs between the final MC calculation and optimization algorithm.

Figure 3.7  3D gamma map between DVH comparison between the final MC calculation and the MC-based
optimization, for a head and neck cancer example. The dose mask is 30% of the prescription dose. The
criteria are 3% dose difference and 3 mm DTA. The pass ($\gamma \leq 1$) rate = 89.3%.

3.3.1.2 MC-VMAT Re-plan

The same patient case was re-planned using the MC-VMAT platform. The purpose of
this example is not to find an optimal solution for this patient case or compete with the AAA-
based plan, but to prove the feasibility of the MC-based inverse planning and the strategy on choosing various optimization parameters. The same amount of the control points, gantry rotation angles and collimator angles were set up in the initial base plan, while the MLC leaf positions were set to conform to the target and the monitor unit required was to be determined. This base plan was simulated using EGSnrc code system. All the energy deposition events happened in the voxels of ROI were recorded in the edepheader and edepdat files. Those data, along with the DICOM plan and structure files were then imported to MC inverse planning algorithm. During the first iteration of the optimization, the dosimetric quality of the plan was significantly improved (Figure 3.8). The total computation time of the optimization was 27.0 minutes, while more than half of the time was spent in the first two iterations. More specifically, the first two iterations should be considered as the most important part of the optimization where major adjustment of the leaf position and beam weight are made. On the other hand, the rest three iterations are designed to make minor adjustments and refinements of the plan.

![Figure 3.8](image.png)

Figure 3.8 A log plot shows the cost decreasing with the optimization iteration, for a head and neck cancer example. The first ▼ symbol represents the initial cost value and the following ▼ symbols represent the cost at the end of each iteration.
Figure 3.9 shows the DVHs of the PTV and the OARs of both the initial base plan and the optimized plan, where the base plan only serves as a reference to show the effect of the optimization. Although the optimization objectives of the PTV were not met, the maximum dose of the PTV decreased from 29.2 Gy to 26.7 Gy and the volume covered by 90% isodose ($V_{95\%dose}$) increased from 80.1% to 89.8%. On the other hand, as for the OARs, the maximum dose in the spinal cord, the brainstem and the chiasm reduced by 14.8% (from 3.4 Gy to 2.1 Gy), 59.8% (from 12.7 Gy to 4.5 Gy) and 70.7% (from 21.9 Gy to 6.4 Gy), respectively.

The optimized leaf positions and beam weights were then written into a DICOM plan file. This new DICOM plan file was imported back for a final MC calculation. The dose distributions of axial, sagittal and coronal planes are shown in Figure 3.10, where the white lines contour the PTV. Apart from the axial plane, which disclosed the under coverage of the target,
the prescribed dose conformed well to the shape of the PTV. However, at current stage of the development of the MC inverse planning algorithm, a final MC simulation is still required. This step can be skipped in the future to improve the efficiency when accurate correction of the leaf-tip effect is applied.

Figure 3.10 3D dose distribution of the MC-VMAT plan, for a head and neck cancer example. The white lines contour the PTV. The dose mask is 30% of the prescription dose.
3.3.2 Brain Arteriovenous Malformations (AVM) ---- Non-coplanar VMAT Example

While the MC-based inverse planning is a valuable tool for the standard VMAT, it also fits very well to the planning of non-coplanar VMAT. A brain AVM patient (case g) was selected for the testing purpose. The plan consisted of 469 control points within five non-coplanar arcs. The gantry angles, the couch angles and the collimator angles are listed in Table 3.2.

Table 3.2 Beam trajectory parameters of the non-coplanar MC-VMAT plan, for a brain AVM example.

<table>
<thead>
<tr>
<th># Arc</th>
<th>Start Gantry Angle (deg)</th>
<th>Stop Gantry Angle (deg)</th>
<th>Couch Angle (deg)</th>
<th>Collimator Angle (deg)</th>
<th>Number of Control Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>179.9</td>
<td>315</td>
<td>315</td>
<td>97</td>
</tr>
<tr>
<td>2</td>
<td>179.9</td>
<td>180.1</td>
<td>0</td>
<td>45</td>
<td>177</td>
</tr>
<tr>
<td>3</td>
<td>180.1</td>
<td>0</td>
<td>45</td>
<td>45</td>
<td>97</td>
</tr>
<tr>
<td>4</td>
<td>250</td>
<td>180.1</td>
<td>90</td>
<td>45</td>
<td>49</td>
</tr>
<tr>
<td>5</td>
<td>180.1</td>
<td>250</td>
<td>90</td>
<td>315</td>
<td>49</td>
</tr>
</tbody>
</table>

The PTV was prescribed with 31 Gy and was constrained such that 0% volume ≤ 32.5 Gy and 95% volume ≥ 30 Gy. The additional difficulty of the non-coplanar MC-VMAT was the transformation of structures in the BEAMnrc coordinates, where the couch rotation must be correctly taken into account. The transformed PTV contours were used to determine the leaf positions that conform to the target. In spite of various couch angles, the coordinates of the ewsphant created in MC algorithm were fixed. More specifically, in MC-VMAT, the transformations of the patient geometry due to the couch rotation were converted to equivalent beam incident directions. As a result, only one MC simulation for the conformal plan and one subsequent optimization of the leaf position/beam weight were performed, where all five arcs were considered as a whole.
Figure 3.11 The gantry-couch trajectory of the non-coplanar MC-VMAT plan, for a brain AVM example. The blue lines represent the trajectory of the gantry head. The yellow lines represent the incident direction of the photon beam. The white dots contour the patient body and the red dots contour the PTV.

The plan modulation was achieved in 7.8 minutes. The optimization for this example implies that while the computation time of the MC-VMAT has a strong relation with the number of voxels (i.e., the size of the ROIs), it is less relative to the complexity of the plan or the number of control points in the plan.
Figure 3.12 DVH comparison between the base plan (solid lines) and the non-coplanar MC-VMAT plan (dash lines), for a brain AVM example. The ▶ and ◄ marks indicate the upper and lower optimization objectives.

The dose sparing of the normal tissue was greatly improved with the optimization procedure, where the maximum dose in brainstem reduced from 24.4 Gy to 14.2 Gy (Figure 3.12). The optimization also improved the target homogeneity and conformity. The conformity index increased from 0.49 to 0.74 and the homogeneity index decreased from 1.61 to 1.04. The maximum dose of the PTV decreased from 33.6 Gy to 32.7 Gy while the minimum dose increased from 27.6 Gy to 28.3 Gy.
Figure 3.13 3D dose distribution of the non-coplanar MC-VMAT plan, for a brain AVM example. The white lines contour the PTV. The dose mask is 30% of the prescription dose.

The forward dose calculation of the optimized plan is shown in Figure 3.13. The dose was compared to the ROI dose calculated by the optimization algorithm using 3D gamma map and a pass rate of 99.8% was observed (Figure 3.14).
3.3.3 Cranial Metastases SRT ---- TVMAT Example

We further investigated the capability of MC-VMAT with more complicated beam delivery technique (i.e., TVMAT) using four cranial SRT patient cases (h-k). On average, the optimizations of all four cases improved the conformity and homogeneity by 26.7% and 23.1% and reduced the mean dose in OARs by 34.7%. One example (case h) is discussed in detail here. In this example, a total of 2324 voxels were selected as ROI, including the PTV, brainstem, chiasm, both left and right optic tracts, and the dose fall-off region. A plan involved five partial
beam arcs and 181 control points, in which the gantry and the couch move simultaneously, was created as the initial base plan. The gantry and couch angles trajectories are shown in Figure 3.15.

Figure 3.15 The gantry-couch trajectory of the MC-TVMAT plan, for a cranial metastases SRT example (top figure). The blue lines represent the trajectory of the gantry head. The yellow lines represent the incident direction of the photon beam. The white dots contour the patient body and the red dots contour the PTV. The couch angles vs. gantry angles are shown in the bottom figure, where the instant dose rates at each control point are also plotted (bottom figure).
The number of particle histories was set to $8 \times 10^8$ to ensure the statistical errors were less than 2% all over the ROI after the optimization based on the ‘particle removing’. The MC simulation for the base plan took 27.1 minutes to finish and stored 2.29 GB $edep$ data. The PTV was prescribed with 24 Gy in 12 fractions. The brainstem was constrained to $0\% \geq 12$ Gy, while chiasm, left and right optic tract were constrained to $0\% \geq 3.6$ Gy. The dose was normalized to the target mean during the optimization. The optimization took 10.3 minutes within five iterations (Figure 3.16). The conformity index increased from 0.25 to 0.56 and the homogeneity index decreased from 1.27 to 1.08. The optimization improved the dose sparing of all three OARs. The maximum and the mean dose decreased by 29.4% and 45.2% in brainstem, 15.2% and 14.7% in chiasm, 18.2% and 17.2% in left optic tract and 30.5% and 21.9% in right optic tract, respectively.

![Figure 3.16](image)

**Figure 3.16** A log plot shows the cost decreasing with the optimization iteration, for a cranial metastases SRT example. The first $\nabla$ symbol represents the initial cost value and the following $\nabla$ symbols represent the cost at the end of each iteration.
Figure 3.17 DVH comparison between the MC-TVMAT plan (solid lines) and the MC-calculated non-coplanar AAA-VMAT plan (dash lines), for a cranial metastases SRT example. The ▶ and ◄ marks indicate the upper and lower optimization objectives.

The MC-TVMAT plan was compared to a non-coplanar AAA-VMAT plan. However, it should be noted that while the latter was planned in Eclipse TPS using AAA, the DVH presented here was calculated by MC simulation (i.e., MC-calculated AAA-VMAT plan). Figure 3.17 suggests that MC-TVMAT plan was dosimetrically superior to the non-coplanar in terms of both the PTV and the OARs. The 3D dose distribution of the MC-based plan is shown in Figure 3.18. The DVHs and the dose distribution were compared between the MC-TVMAT plan and the final MC calculation in Figure 3.19 and Figure 3.20, respectively.
Figure 3.18 3D dose distribution of the MC-TVMAT plan, for a cranial metastases SRT example. The white lines contour the PTV. The dose mask is 30% of the prescription dose.
Figure 3.19 DVH comparison between the MC-based optimization (solid lines) and the final MC calculation (dash lines), for a cranial metastases SRT example. The ▶ and ◄ marks indicate the upper and lower optimization objectives.
3.4 Conclusions

We developed a MC-based inverse treatment planning algorithm for trajectory-based VMAT, capable of meeting all clinical objectives. This method requires a single MC simulation and takes into account the continuous motion of the beam trajectory at all stages of the optimization. Eleven previously treated VMAT patient cases with various sites were re-planned using this approach, including head and neck, lung, prostate and brain cancers. For all the cases, the Monte Carlo inverse planning algorithm is able to achieve plans that meet the prescribed...
dose-volume constraints. It is noted that the simulation time required for MC-based optimization is comparable to the standard simulation. The optimization time, however, depends on the number of the voxels labelled as PTV and OAR, but not the complexity of the dynamic trajectory or the number of control points because of the particle-by-particle calculation framework of Monte Carlo. Compared to the current VMAT planning technique, this last point is of great importance as Monte Carlo is potentially more time efficient as well as more accurate than existing algorithms for modelling dynamic source and MLC motion, especially for beam delivery requiring a large amount of control points. These results show that this thesis presents the first tool capable of performing intensity modulation using Monte Carlo dose calculation for VMAT delivery, both standard coplanar beams and highly non-coplanar beams.
4.1 Introduction

Total body irradiation (TBI) is frequently used in a conditioning regimen for patients undergoing bone marrow transplantation in order to treat various forms of leukemia and lymphoma [120-122]. The two main purposes of the TBI treatment are: 1) to ablate the malignant cancer cells and 2) to suppress the immune system and increase the chance of acceptance of donor marrow. In order to effectively eradicate residual tumours and suppress the immune system in the whole body, TBI requires a uniform dose across the body within ±10% variation and a low dose rate that is no larger than 20 cGy/min [123]. This low dose rate protocol of delivering TBI stems from radiobiological consideration. It was suggested that a higher incidence of interstitial pneumonitis is directly related to high dose rate [124, 125]. There is also a physiological reason, related to blood flow. A low dose rate ensures that circulating leukemia cells receive the required dose [126]. Various techniques are used in TBI to ensure that the entire patient fits within the radiation beam. In BC Cancer – Vancouver, a sweeping-beam Cobalt-60 source (Theratron 780C, Atomic Energy of Canada Ltd., Ottawa, Canada) is used, with a head swivel feature to sweep a treatment field over the patient at an extended source-to-surface distance (SSD). Each fraction is comprised of two fields, with the patient lying supine for one field and prone for the other. This technique, however, would lead to larger dose being delivered to the centre of the patient, which is closest to the source, compared to the extremities in the superior and inferior directions [127]. While achieving a high level of success, traditional TBI does carry risk for early and late toxicity, particularly for lung. Consequently, it is preferable to
have a reduced dose to the lungs with body compensators or lung attenuators to irradiate the patient’s entire body. In our agency, patient-specific lead lung compensators are manufactured and are placed on a thin plastic tray above the patient.

There are certain considerations associated with traditional TBI techniques, including the patient setup and the treatment planning [128]. While the supine/prone positions allow patients to be positioned underneath the treatment gantry, patients have to be positioned twice in one treatment because of couch limitations. The long treatment and setup time can cause discomfort and stress in some patients [129]. In addition, the extended SSD setup brings about additional difficulty to acquire accurate 3D treatment planning data [130]. A manual measurement of the SSD is required for each field and the treatment time is determined as a function of the measured SSD.

Traditional TBI techniques are very taxing for patients and require specialized equipment (i.e., Cobalt-60 units) that are now becoming obsolete, which motivates the development and implementation of sophisticated beam modulation techniques for TBI. The two main alternatives are helical tomotherapy (HT-TBI) and VMAT-TBI. HT-TBI, with which the radiation is delivered slice-by-slice, involves a couch translation while the gantry fully rotates around the patient [131], while VMAT-TBI uses an optimization technique to produce multiple treatment arcs with continuously changing gantry rotation speeds, MLC patterns [132, 133]. Both TBI using HT and TBI emerging VMAT techniques provide the capability of selectively lowering radiation dose to the lungs. This reduces the incidence of symptomatic and life-threatening pneumonitis, and thus allows TBI to be used in cases where compromised lung function would have previously precluded this. Specifically, the planning of HT-TBI involves the determinations of field width, the distance that the couch travels per rotation and the maximum leaf opening.
The HT-TBI treatment required approximately 30 minutes of beam-on time, for a field width of 2.5 cm and the optimization process took from 45-60 minutes [131, 134]. Nevertheless, while clinical successes using HT-TBI have been reported [134, 135], the HT system is not available at BC Cancer – Vancouver. On the other hand, current VMAT-TBI technique involves intense use of planning resources and long treatment times due to the need for multiple isocentres and beam junctions [136, 137]. A report showed that the total time for optimization of a nine-arc VMAT-TBI plan required approximately 21 hours [133]. Compared to the traditional techniques, the differences in organ dose constraints, dose-volume coverage (total marrow versus total body), the timing of treatment delivery and the impact of these differences to the clinical outcomes for emerging VMAT techniques remain unclear. These factors all limit the widespread implementation of these emerging techniques [126, 138].

We present a new linac-based treatment modality for TBI, which allows us 1) to accurately plan the treatment beam thus to limit the dose to organs at risk, such as lungs, without compromising dose to the rest of the body, and 2) to enhance the tolerability of the treatment by eliminating the need for prone patient positioning. Our technique uses a conventional isocentric 6 MV photon beam linac, with a helical beam delivery path defined by simultaneous gantry rotation and longitudinal couch translation, with which no field junctions will be introduced. The technique is similar to HT in that a full 360° of beam directions are available for optimization but is fundamentally different in that the slice-by-slice delivery of HT is inherently less efficient with regard to both radiation output and treatment time. On-board kV imaging may be used to monitor patient position for the entire course of treatment. We use MC dose calculation algorithm, since the dose delivered to bones and lungs is of central interest for TBI. We expect that the optimized treatment delivery, combined with the accurate dose calculation, will lead to a reduction in the
morbidity associated with total body irradiation and consequently to improved quality of life and higher survival rates.

The purpose of this study is to develop an achievable helical pattern trajectory and determine the feasibility of adapting the trajectory-based VMAT optimization algorithm (Chapter 3) to the beam modulation of TBI. Our optimization strategies are flexible enough to allow for various clinical objectives, ranging from uniform TBI to total marrow or lymphatic system irradiation. Our new linac-based technique allows the generation of patient-specific, intensity-modulated TBI plans, with the potential of conformal sparing of organs at risk, for various clinical objectives. Specifically, we expect to achieve a total body dose distribution equivalent to, or better than, the current sweeping beam technique [139], where no attempt at organ sparing is made, and organ sparing equivalent to, or better than, what is achievable using a dedicated HT unit. Apart from providing a technique that will achieve delivery parameters close to our current practice, we will also introduce a planning and delivery framework for clinical research into the impact of dose constraints, MU rate, overall treatment time and beam trajectory on TBI patient outcomes. While the trajectory and optimization algorithm is developed for TBI, this methodology can be extended to other long field (> 40 cm) treatment applications including cranio-spinal radiation therapy (CSRT). This CSRT example will be presented in Appendix B.5.

4.2 Methods

4.2.1 Beam Geometry and Patient Setup

For our HTBI technique, we propose to use a double-helix beam trajectory. This is achieved by performing a series of 360° gantry arcs while the treatment couch is translated longitudinally in both directions. We use a series of alternating clockwise and counter-clockwise
arcs. On the other hand, as the couch has a longitudinal range of 90 cm from the isocentre to the point of collision with the gantry head, longer treatment lengths (> 110 cm) will be accommodated by a 180° couch top rotation, while a 110 cm treatment length is achievable without the need to reposition the patient. Using this method, coverage of the entire body for patients up to a height of 220 cm can be accommodated. Consequently, for longer patients, the HTBI technique requires dose delivery to the torso and legs in sequential treatment segments with a 180° patient rotation performed between segments while the gantry is paused in the upright 0° position. Such trajectory starts at the top of the thigh and moves toward the top of the head. Once the beam trajectory reaches the top of the patient's head, the couch motion reverses direction such that a double pass of the helical trajectory is applied. The gantry pauses at the 0° position (vertically upright) once the top of thigh returns to isocentre and therapists manually rotate the patient on a pivoting couch top support 180° from head to toe. The HTBI delivery then continues from thigh to toes and back. As a result, there will be no beam junction, as the beam trajectory will pick up from where it left off prior to the patient rotation.
Figure 4.1  Plot of double helix HTBI trajectory of isocentre position vs. gantry angle. The top, mid and bottom panels show the HTBI trajectory delivered up to the 4th, 11th and 14th field, respectively.

An illustration of the gantry-couch motion trajectory for longer treatment lengths is shown in Figure 4.1. The gantry performs a series of 360° arcs while the couch translates longitudinally moving the patient smoothly through the beam. In each 360° gantry arc, 97 control points are distributed uniformly with an interval of 3.75° between two adjacent control points. Each control point has its own isocentre and only radiates a selected part of the target. The isocentres are continuously moving with the longitudinal couch position. The number of the arcs \(N_{arcs}\) is determined by the gantry rotation speed and couch translation speed using:

\[
N_{arcs} = \frac{\text{PatientLength}}{360 \times \frac{\text{GantrySpeed}}{\text{CouchSpeed}}} \tag{4.1}
\]
The maximum allowed field size for each control point is determined by:

\[
\text{FieldSize} = \frac{\text{PatientLength}}{\text{Narcs}} \quad (4.2)
\]

In this feasibility study, the PTV is defined as the total body minus the OARs (lungs for example). The MLC leaf positions of each control point are initially set to conform to the target. Constraints are applied to the leaf motion and dose rate to ensure that optimization leads to a deliverable solution. The speed of delivery is adjustable to ensure the patient comfort and to achieve clinically recommended dose rates. On-board imaging modalities (Cone Beam CT, kV and MV imaging) would be used to verify patient position [140].

4.2.2 Dose Calculations and Optimization

We use EGSnrc code system for the dose calculations in this study, as the dose delivered to bones and lungs is of central interest for TBI. We employ the source 20 of DOSXYZnrc to perform MC simulations for moving radiation sources, which allows for gantry and collimator rotation, couch rotation and translation in any direction, arbitrary isocentre motion with respect to the patient, and variable MU rate. Our simulations of complex treatment plans take no more CPU time per particle and require no more computing resources than a single static simulation.

The novel trajectory-based optimization algorithm we developed in Chapter 3 is used to determine MLC leaf positions and beam intensity for the helical VMAT TBI. The preparatory step for the plan optimization is the determination of the dose to each patient CT voxel from every beamlet in the helical trajectory. These data are pre-calculated and stored with a single MC simulation. Specifically, the process is performed once with fully retracted MLC and once with fully closed MLC leaves. Dose associated with each simulated particle in every patient phantom voxel is scored and stored, using the method described in Section 3.2.1 (i.e., \(edep\) data). This
information is then used in the MC-DAO algorithm to minimize a quadratic objective function (Equation 3.2) based on user-defined constraints. This approach is realistic and achievable in terms of computing time as the entire helical trajectory is simulated as a single beam. While MC simulation is the first choice used in this study since the optimal beamlet is an individual incident particle, an Acuros XB dose calculation algorithm (Varian Medical Systems, Palo Alto, CA) is also conceivable.

A 6 MV photon beam is the first choice while other beam energies (e.g., 4 MV and 10 MV) were also tested depending on target coverage. We developed an interchangeable environment (DICOM, data inputs, Application Programming Interface (API) scripting language, and Extensible Markup Language (XML, Developer Mode file format), for transferring planning instructions between Eclipse, the VMAT HTBI optimizer and the TrueBeam linac. Using API scripting these plans can be evaluated in Eclipse (dose/volume analysis and radiation oncologist review) and potentially verified through comparison with an approximated sequence of static Acuros XB calculations. While the patient setup changes from head first supine (HFS) to feet first supine (FFS) for patients longer than 110 cm who requires couch top rotation, the phantom used in MC algorithm is always considered as HFS setup. This setup difference between Eclipse and MC is also taken into account during the plan interchange.

4.2.3 Dose Rate and Treatment Time

The timing of the beam trajectory (depending on the choice of pitch, MU rate and number of longitudinal passes) is of particular interest in this study. Our existing sweeping beam TBI technique delivers the beam in an average patient dose rate of 10 to 20 cGy/min, which is consistent with the current clinical guidelines. Therefore, we aim to achieve the same average overall patient dose rate to mimic our current technique the minimum requirement for the HTBI
technique, though the effective dose rate to a particular leukemic cell depends on the selected trajectory, MU rate and timing. Furthermore, a treatment time of less than 20 minutes (depending on the treatment length) is chosen to match our existing sweeping beam technique. The total treatment time \( T \) is calculated based on the number of arcs and the gantry speed by:

\[
T = N_{arcs} \times CouchSpe
\]  \hspace{1cm} (4.3)

The patient is planned on Varian TrueBeam linac in the developer mode, in which we set the maximum gantry rotation speed to 6 deg/s and maximum couch translation speed to 5 mm/s. This setting allows a maximum longitudinal motion of 30 cm in one 360° rotation, and thus coverage of the 90 cm field length requires 3 minutes in one direction. Using a linac with a maximum reference dose rate of 600 MU/min and a maximum leaf speed of 3 cm/s, this treatment delivery is feasible.
4.3 Results and Discussions

4.3.1 A Total Body Irradiation (TBI) Example

A patient who has a total treatment length of 161.0 cm was selected as the first test case. The prescription dose to the PTV was 12 Gy in 6 fractions over three days. In this example, coverage of 80.5 cm treatment length was required for one direction of the couch travel, which was achieved in fourteen arcs. A couch speed of 4.65 mm/s and a field size of 40×28 cm² were determined to make use of the full range of MLC leaf motion. Fourteen continuously full arc
fields were planned with 97 control points in each field, which allowed 3.75° spacing between adjacent control points. A standard gantry rotation speed of 6 deg/s was applied. As a result, the total beam-on time was 14 minutes. The DICOM plan file was converted into XML file and uploaded for beam delivery in TrueBeam developer mode. A dose calculation grid of 10.0 mm was used in the MC simulation, where 50633 voxels were labelled as ROI (i.e., patient body including lungs). The MC dose calculation, with \(4 \times 10^8\) particle histories incident to the patient body, took 36.4 minutes to finish. It should be pointed out that the number of particle histories was selected to accommodate the memory limitation (16 GB RAM) of the optimization algorithm and was not able to achieve 2% statistical uncertainty (section 3.2.2) in dose distribution. Consequently, for the presented TBI case, the rejection threshold for uncertainty was increased to 5%. The MC dose was written into a DICOM dose file and imported to Eclipse TPS for plan viewing. The dose distribution is shown in Figure 4.2.
The same plans were simulated using 4 MV and 10 MV photon beam and the dose distribution are presented in Figure 4.3 and Figure 4.4. While the 4 MV beam was less penetrating compared to 6 MV beam and required more MU to achieve the prescription dose, the 10 MV beam was more penetrating, but required less MU. However, neither 4 MV nor 10 MV beam provided obvious advantages over the 6 MV beam in terms of the homogeneity and conformity. Therefore, we continued to use 6 MV beam as our first choice.
Both instantaneous dose rate at each control point and the average dose rate over the entire treatment were calculated. Our HTBI technique is able to keep the instant dose rate under 300 cGy/min and the average dose rate under 20 cGy/min (Figure 4.5), which is consistent with the current clinical guidelines.
Figure 4.5  Plot shows the dose rates of four randomly selected voxels in the patient phantom. The instant dose rates at each control point are presented as treatment time vs. dose rate and the average dose rates over the entire treatment are shown at the top-right corner.

The dose distribution was modulated using the current version of MC-based optimization algorithm used for VMAT and TVMAT. The results are shown in Figure 4.6. The entire patient body, including lungs, was constrained within 12 Gy ± 10%. The computation time for the optimization was approximately 90 minutes. However, while the intensity modulation was able to reduce the dose in lungs, it traded off the maximum dose in the target (i.e., patient body).
Several challenges in this example must be pointed out here. First, while the longitudinal translation of the couch permits a treatment length > 40 cm, the maximum width in the lateral direction is still 40 cm. Consequently, the patient arms should be repositioned in future cases. In addition, compared to other standard tumour sites (e.g., prostate, breast), a major difference of TBI is that the OARs are surrounded by the target. It should be noted that the work presented here remains a preliminary. While the test beam trajectory is proven to be feasible, the intensity modulation strategy still requires improvement. In principle the HTBI could achieve the same results as the VMAT-TBI plans. The modulation results could largely depend on the selection of the initial trajectory, the collimator angles, the field size determined at each control point and the constraints of the objectives. Therefore, we will continue this project and further investigate the optimization of this HTBI technique. This may include optimizing the double helical trajectory,
introducing dynamic collimator rotation and developing more advanced MC-based optimization algorithm.

4.4 Conclusions

We have designed a beam trajectory on linac with double-helix pattern for long field (> 40 cm) treatment applications including both total body irradiation and cranio-spinal irradiation (see Appendix B.5). This trajectory involves couch translation (both directions) with gantry rotation, leaf motion, variable dose rate, variable gantry speed, collimator rotation, jaw tracking. Through MC simulations, the test pattern is proven feasible to deliver uniform dose to the targets and will not introduce hot/cold spots due to field junctions. The average dose rate during the entire treatment is kept below 20 cGy/min. The beam trajectory created with in-house code using Matlab can be easily transferred to DICOM and XML files, and thus the delivery of HTBI trajectories can be achieved in developer mode.

The prototype of optimization code enables user-defined organ dose constraints, couch pitch, MU rate, number of longitudinal passes, variable collimator angles and jaw tracking. With the intensity modulation, the dose in OARs can be effectively reduced. Future work required are to improve the dose homogeneity and to progressively develop more complex optimization schemes, involving collimator rotation. In addition, we will begin the validation of the VMAT HTBI technique with performance assessment of linac couch, gantry, MLC and MU rate, separately and in combination. Systematic testing for the developed patterns will also be generated prior to dosimetric validation of actual treatment plans.
Chapter 5: Efficient, All-In-One, Monte Carlo Simulations of Transit EPID Cine-mode Dose Distributions for Patient-Specific VMAT Quality Assurance

5.1 Introduction

Due to the high level of complexity of VMAT delivery, there is an increasing need for more comprehensive QA techniques that can properly assess the performance of the individual VMAT control systems. A truly patient-specific QA protocol should be relevant to the actual patient treatment and be capable of providing *in vivo* dosimetry. Although phantom-based commercially available QA solutions, such as ArcCHECK, MapCheck (Sun Nuclear Corporation, Melbourne, FL), or MatriXX (IBA Dosimetry GmbH, Germany), are widely used for pretreatment verification, they would obviously not be able to catch treatment errors, such as patient mispositioning or errant MLC position. In contrast, patient dose calculation using real-time linac log information and EPID transit dosimetry can provide the information required for accurately assessing the patient-dose accumulation over the entire course of treatment [141, 142].

EPID can be used for both pretreatment and *in vivo* transit dose verification. Pretreatment dose verification allows for detection of errors made during the transfer of treatment parameters between the TPS and the treatment unit, and for detection of malfunctioning equipment, including machine output variations, flatness and symmetry errors, or leaf positioning errors (e.g., due to incorrect MLC calibration). On the other hand, transit dose verification allows for detection of errors related to patient treatment, such as setup errors, organ motion, anatomy changes, or changes related to different patient arm position.
Monte Carlo simulations are ideally suited for treatment QA purposes, besides the proven advantage in terms of dose calculation accuracy. The MC technique allows users to perform independent verification of patient dose in every treatment fraction in an adaptive radiotherapy process, in a framework outside the TPS [143]. To this point, we have introduced the capability of simultaneously providing dose deposition in both the patient CT data set and any type of planar (entrance or exit) detector for VMAT MC-based QA [144]. During a source 20 simulation, the particles exiting the patient can be collected by enabling DOSXYZnrc to optionally output a 4D IAEA phsp [145]. The 4D exiting phsp, stored at the patient CT boundary, takes advantage of the fact that the IAEA format permits storage of nonplanar phsps, and thus allows for the capturing of all particles that exit the patient, regardless of their direction. This scoring geometry can be saved in either the linac (BEAMnrc) or the patient (DOSXYZnrc) coordinate systems. The geometry of this phsp is that of the parallelepiped boundary of the patient CT phantom. In comparison with discrete per-control-point simulations, an advantage of the 4D phsp approach is that the continuity of the beam delivery between control points is considered in the same manner as it occurs for the beam delivery. Even for large control point spacing, the beam delivery and hence the patient dose can be accurately simulated. This phsp is used as input source for further synchronized simulations of EPID dose or images, either integrated or individual for each control point. As a result, comparisons between the 4D MC-EPID prediction and the actual measured EPID data acquired with the patient on treatment become possible.

During a standard DOSXYZnrc simulation, each particle history is tracked using a randomly sampled time-like variable (i.e., MU index ranging from 0 to 1). However, the temporal information is discarded during dose deposition, producing a cumulative distribution.
This means that if we wish to determine dose distribution as a function of time, separate simulations would have to be run for each interval of interest. We previously introduced a MU filter in source 20 that rejected particles outside a user-defined MU index interval [141]. This filter was enabled by including two new optional variables defining the lower and upper limits of a MU index interval in the _egsinp_ file. This feature could be used to calculate integral, incremental, or per control point, dose distributions either in the patient CT phantom, or in the EPID. This method, nevertheless, required individual simulations for each MU filter. Consequently, it was clinically unfeasible to perform MC simulations of transit EPID cine-mode acquisition for a VMAT plan delivered to a patient.

To overcome this limitation we developed a new method capable of providing these transit EPID distributions, per control point, or over any user-defined time interval, in a single simulation. The development of this technique gradually evolved from separate simulations for each control point [143], which was too computationally intensive to be clinically feasible, to the current method [146], which is capable of generating all cine-mode images in a single simulation. It should be noted that the same MC technique that generates all of the EPID images in a single simulation could also be applied to the generation of dose rate maps in the patient CT phantom, per control point, or per any user-defined time interval [147]. With the coupling between the MC image prediction and the MC patient-dose prediction, the MC-EPID method presented here can be used for both pretreatment and transit _in vivo_ dose verification, and thus it can rigorously verify the linac beam delivery. Various types of errors can be detected by comparing measured and MC-predicted EPID transit images, such as errors during information transfer to the delivery system or errors in patient positioning. In addition, by collecting EPID
images acquired during each fraction, we have the capability to evaluate the influence of the change of the anatomy, which can happen during the day-to-day treatment.

The purpose of this study is to present a novel tool within the BEAMnrc/DOSXYZnrc environment for MC simulations of transit EPID dose. We have developed a MC technique to predict transit EPID images, in cine-mode, for patients undergoing VMAT treatments. These predicted images can be compared with the actual EPID cine-mode images acquired while the patient is being treated. This methodology can be adopted as an integral part of a comprehensive patient-specific QA, which is directly relevant to the actual patient treatment and is not being limited to pretreatment verification, as traditional QA methods typically are. In addition, we also investigate the feasibility of this MC-EPID method in detecting the mispositioning errors. Ideally, this tool can be used to monitor treatment delivery in every fraction, during the entire course of treatment. Since anatomy changes may be expected during a long course of treatment, we also propose that cone-beam CT data should be used to update the patient anatomy used in the MC simulations [148]. This will allow, on one hand, to predict EPID images accurately and on the other hand, to accumulate the dose delivered to the patient adaptively. Therefore, this methodology would independently assess whether the original planning constraints for the PTV and OARs have been achieved and the treatment has been delivered precisely.

5.2 Methods

The patient-specific QA protocol is developed based on our current automated Monte Carlo workflow, which has the capability of simultaneously providing dose deposition in both the patient CT data set and any type of planar (entrance or exit) detector for VMAT Monte Carlo QA [143]. The input for the process consists of DICOM files exported from the TPS, including
the DICOM plan and structure files for the creation of *egsinp* file and the DICOM CT files for the patient-specific phantom. While the DICOM dose file is not used during the simulation, the export of this file is also expected to allow the import of the MC dose back into the TPS for in-depth comparisons with the planned dose. Optionally, if the comparisons between the cine-mode EPID images (i.e., cine-images) and MC-predicted EPID dose are desired, the DICOM image files that contain the measured EPID data are also exported.

We modelled the EPID as a planar phantom consisting of 256×20×192 voxels. The phantom has a resolution of 1.563 mm on both $x$-axis and $z$-axis and non-uniform resolution on the $y$-axis. A single simulation is again required in all cases, without the need to build a phantom for each patient-EPID relative position.

### 5.2.1 Description of EPID Simulation

We modified our current MC-based QA script such that an *EPID.egsinp* is automatically created based on the regular *egsinp* (i.e., for patient-dose simulation) when EPID simulation option is requested. The script reads the DICOM image files to acquire the source-to-imager distance (SID). Theoretically, this simulation only requires two control points with MU index 0 and 1 since there is no MLC sequence and gantry rotation needed. Nevertheless, while the EPID planar is static in BEAMnrc coordinate system, the exit phsp does not record the collimator angles for each control point. Therefore, both the number of control points and the collimator angle for each control point in the *EPID.egsinp* are set to the same as the regular *egsinp*. The global electron cutoff energy (ECUT) is set to 0.521 MeV (kinetic energy of 0.01 MeV + rest energy of electron of 0.511 MeV, instead of the standard setting for MV photon beam where ECUT = 0.7 MeV) in order to capture the low-energy electrons contributed to EPID phantom.
and the global photon cutoff energy (PCUT) is set to 0.01 MeV. The selection of ECUT is also related to the lower energy limit of cross-section data set available in EGSnrc [85].

At the end of the patient-dose simulation, a 4D IAEA phsp is output, which collects particles exiting the patient. By employing the SYNC* modules (SYNCVMLC or SYNCHDMLC used in section 3.2), this phsp preserves the full properties of beam particles, and thus it can be used to synchronize the simulations of EPI D dose using the MU index associated with each particle. The MC-EPI D simulation is performed in the BEAMnrc coordinate system where the EPI D position is stationary and is invariant as the gantry rotates around a patient. The time-like dimension is provided by the MU index variable scored for each particle, which permits the computation of EPI D images over arbitrary user-defined gantry angles, including between adjacent control points, partial arcs, or integrated over the entire treatment delivery. With this phsp, it is straightforward to perform a synchronous simulation into an EPI D phantom, since the EPI D is static in this coordinate system, even for multiple-field non-coplanar VMAT plans.

5.2.2 Simulation of Cine-mode EPI D

The cine-mode EPI D acquisition is modelled as a process of 4D dose generation. Therefore, we employ the modified DOSXYZnrc to generate edepheader and edepdat files, in which the eleventh slice of the EPI D phantom, made of Gd₂O₂S, is modelled as the effective detector layer (Figure 5.1). The EPI D exact geometry and composition are provided under non-disclosure agreement. We choose the voxels in this slice as the ROI to score the edep data, which are essential to predict the cine-mode EPI D images. Unlike the MC-based optimization, where we require the position of particle histories in the phsp for optimization, here we only store the MU index and the number of voxels that particle deposited energy for further dose binning.
specifically, with the individual energy deposition of each particle history, we have the capability of binning the dose of any MU interval without the pre-definition of the MU filter. The time-like variables in these files can be binned in a user-defined manner (e.g., consistent with the EPID cine-mode acquisition). Associated with each bin is the amount and location of energy deposited by particles whose MU indices fall within that bin. As a result, only one MC simulation for cine-mode EPID images is demanded to provide predicted EPID images between arbitrary MU intervals.

![Figure 5.1](image.png)

Figure 5.1 The cross section of the EPID phantom used in MC simulation, where the eleventh slice is modelled as the effective detector layer and selected as ROI.

5.2.3 MC Prediction of Individual EPID Cine-image

Ideally, the MC-predicted EPID images should be acquired before the acquisition of the measured EPID images. However, at the current stage of the development, we do not have enough information to predict the time when each EPID image is recorded. The continuous set of EPID cine-mode images are acquired during the treatment beam-on-time, based on a manufacturer-defined frame. While both the MU index of each control point recorded in the DICOM plan file and the EPID images acquisition frame can be considered as time-like variable, they do not share the same time interval. Namely, the exported DICOM image files do not provide a corresponding EPID image for each control point. As a result, the comparisons
between the measured EPID and MC-predicted EPID are based on the image acquisition frame, but not based on the control points. The relationships between the frame and MU index are established through the gantry angle, since each DICOM image file records the gantry angle of the acquired image while the DICOM plan file records the gantry angles corresponding to the control points. The scored $edep$ data are binned according to the period of EPID images acquisition. The binned MC-EPID cine-images can be imported back to the DICOM image files or compared directly to the measured EPID images using gamma evaluation.

### 5.2.4 Statistical Considerations for MC-EPID

In MC simulation, the statistical precision of a scored quantity is proportional to the number of events that contribute to the score, i.e., the number of particle histories ($nhist$) simulated. In order to ensure an acceptable uncertainty for each individual MC-predicted EPID cine-image, the number of particle histories simulated required to compute an EPID image is substantially greater than the number for a patient-dose calculation with an equivalent statistical precision.

With the same incident fluence, the number of events occurring in two volumes containing equivalent materials is proportional to the voxel volume. For example, a simulation with an EPID phantom consisting of $1024 \times 768$ voxels in the detector layer (which is the number of imagers in the actual $400 \times 300$ mm$^2$ EPID) requires sixteen times the particle histories for a $256 \times 192$ voxels layer, for equivalent statistical precision.

An additional factor that affects the $nhist$ required for EPID calculations is the number of frames used in the image acquisition. For the integrated EPID dose calculations, the $nhist$ required for a given statistical precision is independent of the number of frames. Unfortunately, this independence does not carry over for the cine-mode EPID images formed by each frame.
When individual per-frame EPID images are desired, the $nhist$ for each beam must be sufficient to reach the desired statistical precision. This becomes especially burdensome when EPID cine-images are desired for arc-therapy verification, with images at, for example, 178 control points, or continuous cine-images yielding more than 100 independent EPID frames. Therefore, for such 4D cine-mode MC-EPID simulation, the $nhist = 1.5 \times 10^{10}$ was set to acquire the same statistical precision as MC patient-dose simulation (where $nhist = 5 \times 10^8$). This MC-EPID simulation requires 30-60 minutes on our computer cluster, since particles in the 4D exit phsp only has to travel a few centimeters in air before being incident on the EPID phantom and no MLC module is involved. The time required to compute these EPID images could be reduced with multiple fast multicores CPUs or graphics processing units (GPU).

5.3 Results and Discussions

In order to illustrate the versatility of this QA system, we present three cases that have been encountered in our practice. Although they are different in terms of treatment techniques, they are all treated from the same point of view of our MC-based QA system, namely as single-field treatments with multiple variable degrees of freedom.

5.3.1 Faster Simulation Using Slab Phantom

One of the key features of DOSXYZnrc is that it simulates the energy deposition of the particles in each voxel as well as the particle transportation at the boundary with the neighbor voxels. Therefore, the computation time for the simulation largely depends on the size of the phantom, i.e., the number of voxels in the phantom. We simplified the EPID phantom to only 20 voxels (Figure 5.2) as the origin EPID phantom has the same media in each layer of the phantom along the $y$-axis. This largely reduces the amount of dose calculation at the boundary of voxels.
and accordingly the simulation time (by approximately 2/3). Instead of recording the energy deposition in each voxel, we recorded the 3D coordinates each time the energy was deposited, i.e., the \textit{edep} files. These data were voxelized to the original size after the simulation.

![Diagram of 256 x 20 x 192 voxels and 20 slabs]

Figure 5.2  A high-resolution phantom with 256×20×192 voxels (left) can be converted into a phantom with only 20 voxels, where each voxel is represented as a slab (right), since there is only one material per layer.

The image of Einstein’s face was used to prove the feasibility of this method. Comparisons between a standard voxelized phantom and the slab phantom in terms of the dose distribution, dose profile of both $x$ and $y$ direction and gamma map were performed. The criteria of the gamma evaluation was 3% dose difference and 3 mm DTA. The Einstein’s face dose image (Figure 5.3) and the dose profile of both directions (Figure 5.4) suggested that using a slab phantom and a voxelized phantom brought about identical results. Using our current MC cluster system, this method speeded up the simulation by more than three times faster, while it made no trade-off in terms of the dose distribution and was able to maintain the same level of the statistical uncertainty. However, this approach requires having same media in a defined region, and thus has limited application in the MC dose calculation of the patient phantom. As a result, this method can only be applied to a planar phantom, such as EPID.
Figure 5.3 The MC simulated dose of Einstein face image using a voxelized phantom (left) vs. a slab phantom (right).

Figure 5.4 The x (left) and y (right) dose profile of MC simulated dose of Einstein face image using a voxelized phantom (red lines) vs. a slab phantom (blue lines).

5.3.2 EPID Images Transiting Anthropomorphic Phantom

We first employed MC simulation of transit cine-mode EPID (i.e., 4D MC-EPID simulation) on an anthropomorphic phantom. A verification plan for a clinical VMAT head and neck case was created for the purpose. The site was irradiated with a two-arc plan, as shown in Figure 5.5. There were 314 MU for the combined two fields. A single source 20 egsinp file was
required for this simulation, containing 356 control points that defined the two whole arcs of this plan. The field properties are given in Table 5.1. The numbers in each line represent the following variables: the index of the control point; the x, y, z coordinates of the isocentre; the spherical coordinate angles defining the beam direction in the DOSXYZnrc coordinate system, $\theta$ and $\phi$; the collimator rotation angle about the beam axis, $\text{phicol}$; a variable with the role of $\text{dsource}$, which is the linac SAD in the case of IAEA phsp; and the fractional monitor units delivered up to this control point, $\text{muIndx}$. The anthropomorphic phantom was converted to an $\text{egsphant}$ for the MC simulation based on the scanned CT images (Figure 5.6).

<table>
<thead>
<tr>
<th>#CP</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>$\theta$</th>
<th>$\phi$</th>
<th>$\text{phicol}$</th>
<th>$\text{dsource}$</th>
<th>$\text{muIndx}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.500</td>
<td>3.000</td>
<td>-5.500</td>
<td>90.0</td>
<td>89.0</td>
<td>240.0</td>
<td>100.0</td>
<td>0.0000</td>
</tr>
<tr>
<td>178</td>
<td>1.500</td>
<td>3.000</td>
<td>-5.500</td>
<td>90.0</td>
<td>91.0</td>
<td>240.0</td>
<td>100.0</td>
<td>0.4923</td>
</tr>
<tr>
<td>179</td>
<td>1.500</td>
<td>3.000</td>
<td>-5.500</td>
<td>90.0</td>
<td>91.0</td>
<td>300.0</td>
<td>100.0</td>
<td>0.4923</td>
</tr>
<tr>
<td>356</td>
<td>1.500</td>
<td>3.000</td>
<td>-5.500</td>
<td>90.0</td>
<td>89.0</td>
<td>300.0</td>
<td>100.0</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
Figure 5.5  A two-arc VMAT verification plan created on an anthropomorphic phantom for a head and neck patient. The plan is simulated as a single field, using source 20 of DOSXYZnrc.

Figure 5.6  An example of the egphant created using CT images of anthropomorphic phantom.
The EPID was placed at 150 cm from the source (i.e., SID = 150 cm). An *EPID.egsinp* was automatically created by the QA script at the beginning of the simulation, with the field properties given in Table 5.2.

<table>
<thead>
<tr>
<th>#CP</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>θ</th>
<th>φ</th>
<th>φicol</th>
<th>dsource</th>
<th>muIndx</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0</td>
<td>-42.64</td>
<td>0.0</td>
<td>90.0</td>
<td>270.0</td>
<td>0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0</td>
<td>-42.64</td>
<td>0.0</td>
<td>90.0</td>
<td>270.0</td>
<td>0.4923</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.0</td>
<td>-42.64</td>
<td>0.0</td>
<td>90.0</td>
<td>270.0</td>
<td>0.4923</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.0</td>
<td>-42.64</td>
<td>0.0</td>
<td>90.0</td>
<td>270.0</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Compared to the initial *egsinp* setup for the patient phantom simulation shown in Table 5.1, only four control points were required in the *EPID.egsinp*. Each arc was defined by the beginning and ending control points (i.e., 2 arcs × 2 control points) with the same MU index in the initial *egsinp*. However, while the collimator angles *φicol* for each control point was the same as those in the initial *egsinp*, the gantry angles *φ* were static in the EPID simulation.

Figure 5.7 shows the image comparison between the measured EPID image signals and MC-predicted EPID transit dose for a random selected image acquisition frame, while Figure 5.8 shows those summed up from all the frames. For this anthropomorphic example, the MC simulation successfully predicted the radiation accumulated between any specific EPID frames as well as the integrated EPID image. Both radiations at the centre of the field and at the leaf edge were clearly predicted in the MC calculation.
Figure 5.7 The measured EPID signal (left) vs. the MC-predicted EPID dose (right) integrated to a random selected frame segment.

Figure 5.8 The measured EPID signal (left) vs. the MC-predicted EPID dose (right) of the entire treatment.

Although the MC-predicted images and the measured EPID images visually agreed with each other, at this point, we lack the quantitative approaches to compare the two image sets since the MC images are doses in the unit of Gy while the EPID images read the signal of the detector. As a temporary solution, both measured and predicted images were normalized such that the relative intensity/dose at the centre of the field equals to 100% and a 2D gamma evaluation was performed between the relative doses (Figure 5.9). However, while this method could provide a
quantitative comparison between the two image sets, the result is heavily dependent on the selection of the reference point. Therefore, a more comprehensive methodology is needed for the data analysis.

![Gamma Evaluation](image)

**Figure 5.9** Gamma evaluation between measured EPID signal and the MC-predicted EPID dose shown in Figure 5.8. The criteria are 3% dose difference and 3 mm DTA. The pass ($\gamma \leq 1$) rate = 89.3%.

### 5.3.3 MC-EPID Prediction for Prostate Cancer Example

The second case was a prostate cancer patient, where a VMAT plan was created for the patient, with 178 control points and 586.2 MU per fraction. The target was prescribed with 74 Gy in 37 fractions. The dose in the patient phantom and the transit dose in the EPID phantom were acquired within one MC-EPID simulation. Figure 5.10 shows the MC dose calculation of the VMAT plan vs. that of Eclipse on the patient CT, while Figure 5.11 shows a DVH comparison for the PTV and four OARs.
Figure 5.10 Comparison between Eclipse (left) and Monte Carlo (right) dose distributions, of a VMAT prostate plan.

Figure 5.11 DVH comparison between Monte Carlo and Eclipse (AAA) dose distributions, for selected structures, of a VMAT prostate plan.
The actual measured EPID images and the Monte Carlo calculated EPID transit doses for two randomly selected frames (14 and 32, of 178) are shown in Figure 5.12. The statistical uncertainties of the MC-EPID transit dose for all frames were calculated using Equation 6.2 and the uncertainty distribution of the selected frame (i.e., 32 of 178) is presented in Figure 5.13. During this particular frame, $8.4 \times 10^7$ of $1.5 \times 10^{10}$ particle histories were incident to the EPID phantom, which resulted in an average uncertainty of 16.7% in the high dose region. Though an incident
with more particle histories could further decrease the uncertainty, a testing simulation with $2.8 \times 10^8$ of $5 \times 10^{10}$ particle histories indicated that the uncertainty of a given frame was only reduced to 13.2%, while significantly increasing the simulation time to over three hours. Consequently, the $nhist$ was chosen to be $1.5 \times 10^{10}$ as a trade-off between the simulation time and the uncertainty. Fortunately, while the uncertainty for a single frame may be larger than 10%, that of the integrated images decrease with the number of frames summed and can easily meet the 5% uncertainty criteria.

![Figure 5.13 Statistical uncertainty of the MC-EPID transit dose for randomly selected control point (32 of 178).](image)

With the help of Dr. Iulian Badragan (BC Cancer – Fraser Valley), we created the best-fit lines of the EPID signals vs. MC image dose for all the frames. The fitting result of the frame selected in Figure 5.12 is displayed in Figure 5.14, while the residual of this plot is displayed in Figure 5.15. The best-fit lines established relationships between the measured EPID signals and the MC calculated image dose. However, it is still desirable to have a calibrated relationship between the MC results and the actual EPID signals.
Figure 5.14 Best-fit line of the measured EPID signals vs. MC-predicted image dose for randomly selected control point (32 of 178).
5.3.4 MC-EPID Prediction for Brain Cancer Example

The third example is a patient who received radiation to head and neck, where 18 Gy dose was prescribed to the target in 10 fractions. We intentionally created a patient mispositioning error by laterally shifting the patient phantom for 2 cm, in order to assess whether the errors could be detected by our method. Both MC-EPID images for original and shifted patient position were calculated. The predicted dose accumulated in a randomly selected frame (12 of 147) and the integrated dose are shown in (Figure 5.16 and Figure 5.18). This selected frame can be considered as an early stage of the beam delivery.
Figure 5.16 The MC-predicted EPID transit dose for the original (right) and the 2 cm shifted (left) patient position of a randomly selected frame (12 of 147). Dose difference can only be detected from gamma evaluation (Figure 5.17).

Figure 5.17 Gamma evaluation (3% dose and 3mm DTA) between the MC-predicted EPID transit dose for the original (right) and the 2 cm shifted (left) patient position of a random selected frame (12 of 147). The pass ($\gamma \leq 1$) rate = 15.0%.
Figure 5.18 The MC-predicted EPID transit dose for the original (right) and the 2 cm shifted (left) patient position summed up for all the frames. Dose difference can be both visually captured and detected from gamma evaluation (Figure 5.19).

Figure 5.19 Gamma evaluation (3% dose and 3mm DTA) between the MC-predicted EPID transit dose for the original (right) and the 2 cm shifted (left) patient position summed up for all the frames. The pass ($\gamma \leq 1$) rate = 12.1%.

Gamma evaluation was performed to compare between the MC-EPID images of shifted setup and that of the original setup (Figure 5.17 and Figure 5.19). A significant amount of the voxels failed the criteria (i.e., $\gamma > 1$) in both gamma maps. While the dose difference can only be
visually captured from the integrated images (Figure 5.18), but not from the individual-frame image (Figure 5.16), the mispositioning error can be detected by both gamma evaluations (Figure 5.17 and Figure 5.19). The results confirm the feasibility to detect the errors during the treatment using the MC-EPI D method developed here.

The backscatter radiation from the EPI D plane to the patient was also considered in this technique. This was achieved by collecting the exit IAEA phsp of the EPI D phantom as the input source and then the phsp was backscattered to the patient phantom. Figure 5.20 indicates that the amount of the backscatter dose to the patient is negligible, compared to the forward irradiation dose.

![Figure 5.20](image.png)

Figure 5.20 Comparison between the MC simulated forward patient dose (top) and backscatter dose from the EPI D (bottom). A dose mask of 0.2% of prescription dose is applied.

### 5.4 Conclusions

In this work, we modified the DOSXYZnrc code system to generate continuous 4D MC dose data in a single simulation. This method allows for comparisons of the 4D MC-EPI D prediction with the actual EPI D data acquired with the patient on treatment. We demonstrated the feasibility of including transit EPI D cine-mode MC simulations as part of a comprehensive
patient-specific VMAT quality assurance process, during which the patient mispositioning can be readily detected. In addition, it has been proven that by placing the EPI D below the patient, negligible amount of backscatter radiation is observed.

Future work on this project includes the calibration of the MC-EPI D image dose to the EPI D signals, with which quantitative analysis can be performed between the measured and predicted images. In addition, the time interval between each EPI D frame needs to be determined such that the MU index of the EPI D cine-mode images can be acquired before the measurement. With this time interval, a real MC prediction of the EPI D signals can be achieved.
Chapter 6: Conclusions and Future Work

6.1 Conclusions

The main objective of this thesis is to develop a full MC-based inverse planning algorithm for complex radiation therapy technique involving non-coplanar trajectory with continuous couch motion. This work is motivated by the availability of the previously developed source 20 in DOSXYZnrc, which is capable of accurately computing dose distributions for a continuously moving radiation source and all the degrees of freedom of a linac. The intent is to develop a tool that takes into account the actual geometry of the linac and the continuous motion of the MLC leaves during every step of the optimization, while still achievable with realistic computation resources. By using the Monte Carlo code system, this tool could improve the accuracy of dose calculation in regions that required inhomogeneity correction or lack electronic disequilibrium, and thus minimizes the inconsistency between the planned dose and the actual dose delivered to the patient.

The MC-based inverse planning algorithm developed in this thesis is based on the existing condensed history algorithm EGSnrc and the direct aperture optimization with progressive sampling used in VMAT. The author modified the DOSXYZnrc in EGSnrc code system to score the individual energy deposition information of each time-stamped incident particle history. As a result, a 4D dose database can be established within one Monte Carlo simulation, regardless of the complexity of the dynamic trajectory. This database is then read by the optimization algorithm to determine the optimal combination of the leaf sequence and beam weight. During the optimization, the constraints of objective function, the dose rate and
maximum allowed gantry, couch rotation speed and leaf travelling speed are considered. The monitor unit required for beam output is also calculated.

The method proposed in this thesis only stores the energy deposition events in the PTVs and OARs, and thus it avoids large amount of computations during a MC-based inverse planning. However, this may result in a relative higher dose in the rest of the body compared to current planning algorithm. In order to compensate this, the author introduced a dose fall-off region surrounding the PTV in the objective function to represent the dose consideration of the patient body. The other thing unique to the MC-based optimization is the number of histories incident and the statistical uncertainty considerations. While it is clear that the increase of the number of histories reduces the uncertainty, it also results in longer computation time and eventually significantly larger amount of 4D dose data. In addition, the algorithm tends to reduce the number of the particles as the optimization goes on. Therefore, the author determined that $8 \times 10^8$ particle histories should be simulated to reach a reasonable statistical uncertainty, while performing the optimization using practical computation resource.

The increase of planning complexity of non-coplanar VMAT or TVMAT does not bring about additional complexity to the MC inverse planning. Both calculation time for a MC simulation and computation time for optimization are not related to the number of beams, making it more time-efficient as well as more accurate than existing algorithms (e.g., AAA) for modelling dynamic source and MLC motion, especially for complicated planning, such as trajectory-based VMAT. While the calculation time for a MC simulation to achieve a desired level of uncertainty is proportional to the number of simulated particle histories, the computation time to acquire an optimal plan is proportional to the number of voxels in the ROI (i.e., the size of the $edep$ data). The presented site-specific planning studies suggest that the MC-
VMAT/TVMAT has the capability to achieve plans that are equivalent, or superior, to those generated by standard planning systems with static couch, within a comparable time. This is a major advance for the field since there has been a long period research about the development of MC-based treatment planning system.

In chapter 4, a feasibility study of a new treatment method combining couch longitudinal translation and gantry rotation (helical delivery) is presented. The author developed a double-helix trajectory that takes into account the couch translation speed, gantry rotation speed and restrictions, thus allowing continuous gantry rotation and couch translation. The test pattern takes advantage of the continuous motion of couch and gantry, and hence would not result in hot spots at field junctions. A fairly uniform dose distribution (i.e., within ±10% prescription dose) was achieved using the test double helix pattern with 14 fields, covering a 161.0 cm treatment length. This new delivery method is able to keep the average dose rate in the patient body under 20 cGy/min. The trajectory planned using an in-house algorithm can be written either in DICOM format that can be recognized by Eclipse TPS or in XML file that can be delivered in Varian TrueBeam developer mode. The EGSnrc code system is used for the dose calculation as the dose in lungs and bones are of utmost importance to the TBI. The optimization algorithm developed in Chapter 3 was applied to the intensity modulation of TBI dose, where the entire body was selected as the PTV and the lungs were selected as OARs. The current version of the optimization algorithm can reduce the dose in both lungs. This method can also be applied to other treatment involved long treatment field, such as CSRT. It has been demonstrated that a uniform dose can be reached in the spinal cord while low dose delivered to both lungs. These results show that it is possible to use a MC-based algorithm to generate treatment plans that could meet the planning objectives.
Another clinical research project that has benefitted from these software tools is the study of *in vivo* patient-specific VMAT quality assurance using EPID in Chapter 5. For this project, the author used the code developed to simulate the transit EPID signals in cine-mode. Consequently, the EPID image at any time scale can be generated by Monte Carlo. By synchronizing the MU index of control points in the DICOM plan file and the number of frames used to record the EPID signals, the author is able to predict the EPID cine-mode images. Instead of using MU filter developed in our previous publication, the method in this thesis can achieve the same results efficiently using a single Monte Carlo simulation. The EPID phantom simulated has a resolution of 1.563 mm. The EPID part of the simulation uses the exit phsp from the patient-dose simulation and requires $1.5 \times 10^{10}$ particle histories to ensure an acceptable uncertainty. Nevertheless, while the integrated EPID images has uncertainties lower than 5%, the individual EPID image for each frame usually has a higher uncertainty due to the lower number of histories simulated within an individual frame. The high uncertainty can be reduced by summing up the dose from several frames at the cost of losing dosimetric details of the individual frame. It has been proven that this tool has the capability of detecting errors related to patient treatment, such as the patient mispositioning by comparing the MC-predicted EPID images to the measured EPID images. The author also determined that this method would only introduce negligible amount of backscatter from the detector to the patients. The development of this tool can facilitate the procedure of the *in vivo* patient-specific quality assurance and the understanding of the errors possibly encountered during the treatment.
6.2 Future Work

We aim to integrate the developed in-house optimization algorithm into the EGSnrc code system, in order to improve the consistency between the planned dose and final MC dose calculation. This integration can also eliminate the data transmission between the two software.

Improvement of the inverse planning algorithm will also be investigated. This is essential to the intensity modulation of the HTBI delivery. With the particle-by-particle nature of MC dose calculation, it is possible to track the source particles that contribute to the hot/cold spots, thus to optimize the dose by directly modifying the weight of these particles or blocking them with MLC leaves. Another promising solution is to customize the collimator angles, which may improve the plan quality by minimizing the radiation to OARs.

Another area of active development is the validation of a new transit MC-EPID algorithm with extensive clinical data for patients undergoing treatment at our cancer centre. As a next step, we will de-noise MC data for cine-mode and develop a quantitative comparison methodology between the raw EPID images and the MC dose arrays scored in the EPID virtual phantom. The adaptive radiotherapy planning potential of this algorithm will also be explored.
Bibliography


Appendices

Appendix A  Preparatory Steps Involved in MC-based Optimization

A.1 Initial Treatment Parameters and Field Setup

The initial arrangement of the photon beam and field sizes are determined prior to the MC simulation. Beams causing collision between the gantry and the couch or patient are eliminated by simulating all beam orientations using a precise computer-assisted design model of the linear accelerator. The planning parameters (i.e., beam energy, patient setup) and the structure sets (i.e., organ contours) are exported in DICOM format from the treatment planning system. The initial beam arrangement is determined using the method described by Wilson et al. [115], including the gantry rotation range, collimator angle and couch trajectory for each control point. This algorithm provides a solution to find an optimized, continuous trajectory through the most promising beam angles via both dosimetrical and geometrical considerations. The couch angle selection and optimization are not the purpose of this study, and thus a detailed description of this approach will not be given here. Nevertheless, while the use of Wilson’s method determined a non-coplanar VMAT solution, where the gantry and the couch move simultaneously, the dose calculation algorithm of current commercial TPSs only supports the static approximation of dynamic beam motion. Therefore, the dose calculation of such dynamic gantry-couch plans is not available.

To overcome this limitation, we calculate the dose distribution using EGSnrc code system instead, where we write all the beam parameters into a file (i.e., egssinp) that is readable for sources 20 or 21 as input. This initial beam arrangement is further modified to fit the MC-VMAT framework. First, the gantry rotation range ($\Delta \theta$) between two adjacent control points are
resampled so that the final number of optimized samples is chosen to ensure a sampling frequency that provides adequate dosimetric accuracy. The typical value of $\Delta \theta$ is chosen to be approximately $2^\circ$, which allows a maximum of 177 control points within a full arc of $360^\circ$, though shorter ranges are allowed for small partial arcs. The field isocentre is placed at the middle of the PTV. The maximum possible number of control points (e.g., 177 for a typical VMAT plan) are defined and the initial positions of the MLCs are set up for all the control points. The leaf positions are based on the BEV projection of the target plus a uniform 5 mm margin, taking into account the coordinate transformation due to gantry, collimator and couch rotation. These initial MLC positions also define the maximum allowed leaf opening for each field. Instead of a fixed value, the jaw openings are individually calculated for each control point, where the two pairs are set to the maximum opening position of the MLC leaves. The dynamic jaws could reduce normal tissue dose spread without prolonging the treatment time significantly. A 200 MU temporary beam output is set as a template and is uniformly distributed among all the control points. The actual beam output of this initial beam arrangement is calculated after the simulation using the method described by Popescu et al. [91].

In addition to the above plan (referred to as the base plan) where all the leaf positions are conformal to the PTV, a second $egsinp$ plan in which all leaves are closed (while all other beam configuration remaining the same) is created. This plan is used as a background (referred to as the background plan) for the optimization to simulate the situation when incident particles are blocked by the MLC leaves.

A.2 Modelling of Linear Accelerator and Patient

The BEAMnrc is used to simulate a 6 MV photon beam output from a Varian iX/TrueBeam medical linear accelerator. The main components in this clinical accelerator are the
target, primary stationary collimator, monitor chamber, mirror, the secondary collimating adjustable jaws and the MLCs (Figure 2.1). The MC simulation of the patient case is divided into two steps. The photon cutoff energy PCUT is set to 0.01 MeV and the electron cutoff energy ECUT is set to 0.7 MeV. The components above the jaws are all patient-specific, and thus are pre-calculated once for all patients as a source phsp (i.e., phase space A in Figure 2.1). This phsp is treatment unit-specific and requires no simulation here. The phase space A is used as input source for the simulation of the jaws. The phase space B is generated for every secondary collimator jaw setting and is placed directly at the top of the MLC without any air gap. This phsp is used as input for the simulation of the MLC plus the patient phantom section in DOSXYZnrc for both initial base plan and the background plan. The MLC leaves are compiled as a shared library during this part of the simulation. We run simulations in parallel using the Condor High-Throughput Computing software on a MC cluster of eight servers, under the Red Hat Enterprise Linux (release 6.4) operating system. Each server has two physical processors (Intel Xeon CPU E5-2650 0 @ 2.00GHz), with eight cores per CPU and two threads per core. A Monte Carlo job submission can be divided into up to 256 parallel processes. The computation time to simulate $5 \times 10^8$ histories from phase space B of Figure 1.2 to a patient phantom is about 30 minutes (but varies with the size of the phantom).

A voxelized phantom is generated from DICOM format CT slice data. The resolution used in this study is set to $2.5 \times 2.5 \times 2.5$ mm$^3$. The DICOM CT slice data are first processed using in-house software such that CT artifacts and all objects (e.g., CT couch that are extraneous to the body contour) are removed (i.e., Hounsfield Units set to air, to ensure that the upcoming MC dose simulation does not attempt to calculate dose in regions of material that are nonexistent during treatment). The linac treatment couch is then added to the phantom if its position and
contour is given in the structure DICOM file. The couch surface is made of graphite and assigned a density of 0.6 g/cm$^3$, and the couch interior is made of Rohacell 51A foam and assigned a density of 0.052 g/cm$^3$. In order to reduce the simulation time, the phantom is then cropped to the outer contour of patient body plus the treatment couch to avoid the unnecessary calculation of particle transportation in air. An additional cropping option is developed to further crop the phantom according to the field size. The boundaries of this second cropping are defined by the maximum jaw opening in the $z$ direction, with possible field rotation and dose fall-off region considered.

**A.3 Dose Deposition in ROI**

Since we are mostly interested in the dose prescribed to the PTV and normal tissue sparing, we outline the PTVs and OARs as a ROI in the phantom. The voxels occupied by the ROI are determined based on the structure contours of the PTV and OARs, with an additional shell of voxels surrounding the PTVs. While these voxels are neither PTV nor OAR, but part of the body structure, they represent the dose fall-off region, where a sharp fall-off is desired for lower dose in the body. The ROI is defined in Figure A.1, including the PTV, the OARs and the dose fall-off region. The dose in this region is also considered during the optimization to help the dose sparing of the body. The 3D subscript of each voxel is converted to an equivalent one-dimensional index, which is then written into a *roidat* file. This *roidat* file is used as an input file for the MC simulation, which marks the voxels of ROI for DOSXYZnrc code. No data will be recorded if the particle does not contribute dose to ROI.
Figure A.1 ROI is defined as the region including PTV, OARs and dose fall-off region. The dose fall-off region is a shell of voxels surrounding the PTV.

A template edepheader and edepdat file is shown in Figure A.2 and Figure A.3, where e.g., the first line in the edepheader file indicates that a particle history deposits energy in 17 voxels in the ROI. The energies and the voxel indices where the deposition happens are then recorded in the edepdat file.
Figure A.2  Template edepheader file. The four columns represent the x and y coordinate of incident particles at the z-position of phsp, the MU index of incident particles and the number of ROI voxels the particle deposit energy to.

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0544</td>
<td>0.0411</td>
<td>0.4241</td>
<td>17</td>
</tr>
<tr>
<td>0.0544</td>
<td>0.0411</td>
<td>0.9858</td>
<td>8</td>
</tr>
<tr>
<td>0.0544</td>
<td>0.0411</td>
<td>0.1408</td>
<td>16</td>
</tr>
<tr>
<td>0.0544</td>
<td>0.0411</td>
<td>0.7845</td>
<td>6</td>
</tr>
<tr>
<td>0.0544</td>
<td>0.0411</td>
<td>0.0696</td>
<td>20</td>
</tr>
<tr>
<td>0.0544</td>
<td>0.0411</td>
<td>0.9583</td>
<td>9</td>
</tr>
<tr>
<td>0.0544</td>
<td>0.0411</td>
<td>0.8351</td>
<td>15</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.4047</td>
<td>17</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.9214</td>
<td>12</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.3286</td>
<td>13</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.2980</td>
<td>14</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.9645</td>
<td>4</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.3553</td>
<td>15</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.1552</td>
<td>11</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.3259</td>
<td>7</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.6233</td>
<td>11</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.1523</td>
<td>20</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.8956</td>
<td>8</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.5995</td>
<td>8</td>
</tr>
<tr>
<td>-0.5542</td>
<td>-0.0822</td>
<td>0.5006</td>
<td>23</td>
</tr>
</tbody>
</table>

Figure A.3  Template edepdat file contains energy deposition of the first incident particle written in Figure A.2. The first column represents the index of the ROI voxel and the second column represents the value of energy deposition.

<table>
<thead>
<tr>
<th>index</th>
<th>energy deposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>277</td>
<td>0.58525696</td>
</tr>
<tr>
<td>354</td>
<td>0.0126852935</td>
</tr>
<tr>
<td>38</td>
<td>0.1968310872</td>
</tr>
<tr>
<td>2458</td>
<td>0.0515512177</td>
</tr>
<tr>
<td>2547</td>
<td>0.290237581</td>
</tr>
<tr>
<td>187</td>
<td>1.1488879</td>
</tr>
<tr>
<td>241</td>
<td>0.768645659</td>
</tr>
<tr>
<td>194</td>
<td>0.238307522</td>
</tr>
<tr>
<td>186</td>
<td>0.198250575</td>
</tr>
<tr>
<td>195</td>
<td>1.0072892</td>
</tr>
<tr>
<td>85</td>
<td>0.5805997</td>
</tr>
<tr>
<td>233</td>
<td>0.243565834</td>
</tr>
<tr>
<td>94</td>
<td>0.624426281</td>
</tr>
<tr>
<td>598</td>
<td>0.0568842815</td>
</tr>
<tr>
<td>712</td>
<td>0.00910505954</td>
</tr>
<tr>
<td>147</td>
<td>0.786088892</td>
</tr>
<tr>
<td>197</td>
<td>0.133540035</td>
</tr>
</tbody>
</table>
A.4 Optimization Constraints

During the optimization, MLC leaf positions or MU weights are constrained such that the aperture shapes and MU values are physically achievable in practice. For example, overlapping of opposing leaves or negative MU weights are physically impossible and are therefore rejected by the optimization. The maximum protrusion the leaf can travel to the opposite bank is 14.5 cm. The dose rate is also considered in the optimization. While for a 6 MV photon beam, the maximum dose rate \( (\Delta MU/\Delta t)_{max} \) is 600 MU/min, it can be as high as 1400 MU/min for a 6 MV FFF beam. Along with the maximum gantry rotation speed, the maximum dose per degree can be attained. For instance, for a Varian TrueBeam system in 6 MV beam mode, the maximum gantry rotation speed \( (\Delta \theta/\Delta t)_{max} \) is 6 deg/s, and thus the maximum MU rate per degree is 1.67 MU/deg. The beam output is constrained by this value throughout the optimization to ensure that the maximum dose rate is not exceeded. Likewise, the maximum leaf motion speed of the Varian Millennium 120 leaf MLC is \( (\Delta x/\Delta t)_{max} \) is 3.0 cm/s, which constrains the beam delivery to a maximum leaf displacement of 0.5 cm per degree of gantry rotation.

A.5 Statistical Considerations for MC-based Optimization

The results of a MC-based optimization are affected by both systematic uncertainties and statistical uncertainties. While the systematic uncertainties result from the inaccuracy of the beam modelling or using the low resolution phantom, the statistical uncertainties can be reduced by running more particle histories. However, as the algorithm progresses, the dose contributions of a large amount of particle histories are removed since they are blocked by the optimized apertures. This in turn increases the statistical uncertainty of the optimization. Therefore, the uncertainty must be calculated and evaluated in each iteration of the leaf/MU change. For a certain voxel \( i \) in ROI, we first sum up both energy deposition \( (\sum_i edep) \) and square of energy
deposition ($\Sigma_i e_{dep}^2$) of the contributing particle. The uncertainty for this voxel is then calculated using

$$unc_i = \frac{1}{\Sigma_i e_{dep}} \sqrt{ainflu(\Sigma_i e_{dep}^2 - (\Sigma_i e_{dep})^2/ainflu)/(ainflu-1)}$$

(6.1)

The number of histories in a MC simulation for radiotherapy dose calculation is generally greater than $10^8$ (i.e., ainflu >> 1), and thus we can simplify the equation as

$$unc_i = \sqrt{\frac{\Sigma_i e_{dep}^2}{(\Sigma_i e_{dep})^2}}$$

(6.2)

One simple solution to evaluate the statistical uncertainties in optimization is to include them in the cost function. Nevertheless, the optimization process will not be significantly affected by the statistical uncertainties in the dose distributions [149], as other terms in the equation still play a dominant role in searching the global minimum. Therefore, we apply a robust strategy that the maximum statistical uncertainties ($unc_{max}$) allowed during the entire optimization is 2% (5% for HTBI cases). More specifically, any change of leaf position or MU weight that causes the uncertainty of any ROI voxels larger than the threshold will be rejected.

A.6 Optimization Time

Samples are introduced throughout the optimization using a strategy that is a function of the optimization progress, based on the gradient of the cost value. For each successful change in either leaf position or MU weight, the cost value is saved in a cost list. The gradient of this list is calculated every minute, where a negative value of the gradient means the reduction of cost and no positive gradient should be observed (except those accepted by SA). The optimization adds new samples when all calculated gradient in the last one minute are larger than a pre-set threshold or no new changes can be found in the same period of time. The number of successful
iterations between the additions of each new sample follows the form of a decreasing exponential. When there are fewer samples (i.e., the beginning of the optimization), the number of iterations between each sample addition is relatively large. Increasing the sampling frequency increases the number of optimizable gantry positions and restricts the maximum MLC leaf position changes in between consecutive samples. As the number of samples increases, there are exponentially fewer iterations between each new sample. The optimization terminates once all samples are introduced. The schedule describes above is specifically designed to remove the demands of pre-set optimization time by automatically determining the required time for each iteration.
Appendix B  Additional Examples of MC-based Optimization

B.1  Dose Calculation in Inhomogeneity

A 30×30×30 cm³ custom phantom was created in Eclipse TPS with the materials assigned to water. Both slice thickness and the axial resolution were 1.0 mm. An 8 cm slab was introduced into the phantom ranging from 5–13 cm depth. The material of the slab were assigned to lung (-719 HU and 0.26 g/cm³ mass density) and bone (1488 HU and 1.85 g/cm³ mass density), respectively. Meanwhile, a voxel phantom with same geometry and resolution was created in DOSXYZnrc, where ICRPBONE700ICRU and LUNG700ICRU were chosen as the media for bone and lung. Four single field photon beams with different open field size and energy (i.e., 15 MV 10×10 cm², 15 MV 4×4 cm², 6 MV 10×10 cm² and 6 MV 4×4 cm²) from Varian TrueBeam were simulated using both Eclipse TPS and EGSnrc. The SAD was 100 cm and the beam had normal incidence in all cases. The PDD obtained from Monte Carlo (where 1×10⁹ particle of histories simulated) was compared to the doses obtained from the pencil beam algorithm (i.e., AAA) implementing the inhomogeneity correction.
The depth dose curves drawn through the single field dose distribution illustrate the problem encountered when target volumes are located inside or directly beyond an inhomogeneity. Both AAA and Monte Carlo simulation model the perturbation due to the low or high density material. More specifically, for the water-lung-water phantom, the increase of dose after the lung inhomogeneity (except in 15 MV 4×4 cm$^2$) and the dose re-buildup region located at the lung-water interface were observed (Figure B.1). However, while the perturbations due to the inhomogeneity are clearly shown for small field size cases, this effect is not as noticeable as in both 10×10 cm$^2$ field sizes cases in MC. The results suggest an underestimation of dose inside
the inhomogeneity and overestimation of dose beyond the re-buildup region by AAA model.

Assuming a PTV is positioned below the slab, the re-buildup region in this example can cause an under-dosage at the side of the target proximal to the beam source.

On the other hand, the drops in dose were observed in bone inhomogeneity for all cases (Figure B.2). Nevertheless, the AAA model did not properly calculate the dose increase at the water-bone interface and the dose decrease at the bone-water interface. In addition, while AAA predicted identical depth dose compared to MC, it overestimated the dose in bone inhomogeneity.
region by an average of 2.3%. This can potentially result in dosimetric error providing that the bone or air cavities are surrounding the target of the radiation.

B.2 Additional Head and Neck Cancer/Brain AVM Examples

Additional head and neck example (case \textit{b}) and brain AVM example (case \textit{c}) are presented here. The optimization took 27.2 minutes and 10.3 minutes, respectively. The varying computation time of the optimization can be explained by the difference of the size of the ROI, where the case \textit{b} contained 40394 voxels while the case \textit{c} contained 5152 voxels. This relationship between the optimization time and the size of the ROI indicates that it is necessary to only choose structures contributing to the optimization objectives in order to improve the performance of the algorithm.

![Figure B.3 DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a head and neck cancer example (case b). The \textbullet and \textdownarrow marks indicate the upper and lower optimization objectives.](image)

157
The DVHs of the two cases are shown in Figure B.3 and Figure B.4, respectively. In the case \(b\), while the optimization only decreased the maximum dose of the PTV from 28.3 Gy to 26.1 Gy, the \(V_{90\% \text{dose}}\) increased 10.1\%, from 89.9\% to 90.0\%. Although the desired dose coverage (i.e., 95\% of the PTV covered by 98\% isodose) was not met, the optimization did reduce the cost by increasing the target coverage. In addition, considerable improvements in dose statistics have been seen for normal tissue, mainly the right parotid and the laryngopharynx. The mean doses of the two OARs decreased from 19.6 Gy to 7.6 Gy and 15.7 Gy to 8.1 Gy, respectively. On the other hand, in the case \(c\), all the optimization objectives were met except that the \(D_{95\% \text{Vol}}\) did not meet the 98\% isodose. The optimization improved the target conformity and homogeneity by increasing the \(D_{95\% \text{Vol}}\) from 94.2\% to 97.6\%. The maximum dose in left anterior chamber and left retina decreased from 3.0 Gy to 0.5 Gy and from 3.6 Gy to 0.6 Gy.

![Figure B.4 DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a brain AVM example (case c). The ▶ and ◄ marks indicate the upper and lower optimization objectives.](image)
B.3 A Lung Cancer Example

A lung cancer patient (case d) was selected to evaluate the effect of different normalization methods. The MC-VMAT normalized the mean target dose to the prescription dose as default, while other normalization strategies (e.g., based on the isodose coverage) are allowed and can be determined by the users. The PTV was prescribed with 30 Gy dose. Apart from the PTV and dose fall-off region, only the spinal canal and esophagus were also selected as OARs. While both right and left lung were contoured for the planning, no optimization objective was required for the two structures, and thus they were excluded from the ROI. As a result, the number of voxels in ROI reduced from 87942 (with lungs) to 38925 (without lungs). This, in turn, effectively saved the memory and computation time for the optimization by ~50%.

The data read-in time and plan-optimization time were 17.2 minutes and 28.2 minutes, respectively. While the optimization objectives for both OARs were obtained, the coverage and the maximum dose of the PTV were not achieved. The mean dose and maximum dose in spinal canal reduced by 18.7%, from 3.2 Gy to 2.6 Gy and 40.0%, from 25.7 Gy to 15.4 Gy, respectively. However, the maximum dose in esophagus increased from 29.0 Gy to 29.7 Gy. Despite of the slight increment, the dose constraint for esophagus (i.e., 100% volume ≤ 30 Gy) was still met. The increase in maximum esophagus dose was a result of the trade-off between the dose constraints for esophagus and other structures. On the other hand, while the optimization did not achieve the PTV coverage, it increased the D95%Vol from 26.2 Gy to 27.6 Gy and thus improved the homogeneity.
Figure B.5  DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a lung cancer example (case d). The dose is normalized such that the PTV mean dose equals to the prescription dose. The ▶ and ◄ marks indicate the upper and lower optimization objectives.

The patient case was re-planned with the dose normalized such that 100% of the PTV receives 90% of the prescription dose. This normalization was employed at every step of the optimization. The maximum dose and the mean dose in the PTV decreased from 34.1 Gy to 32.8 Gy and from 27.0 Gy to 25.8 Gy, separately (Figure B.6). The dose distributions are shown in Figure B.7. The dose calculated by forward MC was compared to that by MC-VMAT using 3D gamma evaluation (Figure B.8). An excellent agreement between the two dose distributions was observed, where the pass rate was 97.4%.
Figure B.6  DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a lung cancer example (case d). The dose is normalized such that 100% of the PTV receives 90% isodose. The ▶ and ◄ marks indicate the upper and lower optimization objectives.

The comparison of the dose statistics indicates the effectiveness of different normalization strategies. Although two methods proposed here resulted in similar target coverage, the normalization based on the coverage caused higher maximum target dose and worse homogeneity due to a higher initial target dose. In order to maintain the coverage during every stage of the optimization, the algorithm traded off the dose in OARs. Consequently, higher dose in normal tissues were observed. However, while normalization to target mean dose offered better results in DVH compared to the coverage normalization, the latter remains a useful alternative for specific patient where optimization priority lies in the target coverage.
Figure B.7  3D dose distribution of the MC-VMAT plan, for a lung cancer example (case d). The white lines contour the PTV. The dose mask is 30% of the prescription dose.
Figure B.8  3D gamma map between DVH comparison between the final MC calculation and the MC-based optimization, for a lung cancer example. The dose mask is 30% of the prescription dose. The criteria are 3% dose difference and 3 mm DTA. The pass ($\gamma \leq 1$) rate = 97.4%.

B.4 Prostate Cancer Examples

Two prostate cancer examples (case e and f) were selected. The bladder, the rectum, both left and right femoral head, and the PTV were included as ROI. The large number of voxels brought about large $edep$ data (i.e., more than 8 GB), which caused the computer running out of memory as all the data were read by the optimization algorithm in once. A simple solution to this issue was to increase the RAM to 16 GB or more so that the computer was able to assign enough
memory for the optimization. With the large size of the data, the computation time of the optimization notably increased to an average of 44.1 minutes for the presenting prostate cases.

The dose sparing of the normal tissue were significantly improved in both cases and can be seen from the DVHs (Figure B.9 and Figure B.10). However, the improvement in OARs traded off both conformity and homogeneity in the PTV. In case e, the mean dose in bladder, rectum, left and right femoral head reduced 30.8%, 25.3%, 53.9% and 60.3%, respectively.

Figure B.9  DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a prostate cancer example (case e). The ► and ◄ marks indicate the upper and lower optimization objectives.
Figure B.10 DVH comparison between the base plan (solid lines) and the MC-VMAT plan (dash lines), for a prostate cancer example (case f). The ▶ and ◄ marks indicate the upper and lower optimization objectives.

The dose distribution of the forward MC calculation of the plan is shown in Figure B.11 and the dose difference between the optimized plan and the final calculation is shown in Figure B.12 using 3D gamma. The pass rate was 91.6%. A similar dosimetry result was observed for case f (Figure B.10).
Figure B.11 3D dose distribution of the MC-VMAT plan, for a prostate cancer example (case e). The white lines contour the PTV. The dose mask is 30% of the prescription dose.
Figure B.12 3D gamma map between DVH comparison between the final MC calculation and the MC-based optimization, for a prostate cancer example (case e). The dose mask is 30% of the prescription dose. The criteria are 3% dose difference and 3 mm DTA. The pass ($\gamma \leq 1$) rate = 91.6%.

B.5 A Cranio-spinal Radiation Therapy (CSRT) Example

A CSRT patient example was chosen to represent the long-field treatment of small target. For the study here, we only selected the spinal cord as the target, while the optimization of dose in brain will be investigated in the future. The treatment length was 675 mm, and thus 180° couch rotation was not required for this case. Seven full-arcs were planned through the patient body. The target was prescribed with 36 Gy dose. A dose calculation grid of 2.5 mm was used in
the MC simulation and 73814 voxels were labelled as ROI. After 76.4 minutes optimization, the prescription dose in the spinal cord was achieved, with a dose variation of ±10%. The MC-based optimization successfully met the desired constraints for this example. The optimized dose is presented in Figure B.13, where a homogeneous dose distribution is observed in the spinal cord and a low dose has been achieved in both lungs and rest of the body.

Figure B.13 Monte Carlo calculated DVH (top panel) and dose distribution (coronal view, bottom left; and sagittal view, bottom right) of a Helical CSRT (optimized) plan using 6 MV photon beam.