Degree Project in Technology
First cycle, 15 credits

Increasing the Efficiency of CyberKnife Cancer Treatments by Faster Robot Traversal Paths

THEODOR HAGSTRÖM

Stockholm, Sweden 2023
Abstract

Cancer remains a significant global challenge, constituting one of the leading causes of death worldwide. With an aging population, the demand for cancer treatments is increasing. Nevertheless, due to technological advancements, cancer mortality rates are declining. This study contributes to these advancements, focusing specifically on radiation therapy, a crucial technology widely used today.

Since the invention of radiation therapy, there has been significant research and progress in the field. One such advancement is the CyberKnife® system (Accuray Incorporated, Sunnyvale, CA, USA) - a fully robotic radiotherapy device that enables precise patient treatments. Its flexibility allows for the delivery of high-quality plans, but treatment times can be quite long, leading to adverse effects for both patients and healthcare providers.

This thesis introduces algorithms aimed at reducing the robot traversal time of the CyberKnife technology. These algorithms are incorporated into an existing optimization framework for treatment planning, with their effectiveness evaluated across various patient cases.

Significant reductions in treatment times for some patient cases were observed, while maintaining satisfactory plan quality, primarily due to more efficient traversal paths for the CyberKnife robot. The increased efficiency of the robot can also be leveraged to create treatment plans with more irradiation directions, increasing the treatment quality in some cases.
Sammanfattning


Denna uppsats introducerar algoritmer som syftar till att minska traverseringstiden för CyberKnife-roboten. Dessa algoritmer integreras i ett befintligt optimeringsramverk för behandlingsplanering, med deras effektivitet utvärderad baserat på olika patientfall.

Betydande minskningar av behandlingstiderna observerades för vissa patientfall, samtidigt som tillfredsställande plankvalitet behölls, främst med anledning av mer effektiva traverseringsvägar för CyberKnife-roboten. Denna effektivisering möjliggör också skapandet av behandlingsplaner med fler strålriktningar, vilket förbättrade behandlingskvaliteten i vissa fall.
Acknowledgements

First and foremost, I would like to express my profound gratitude to Björn Andersson at RaySearch Laboratories, who has been my guiding light throughout the course of this project. His insightful feedback, encouragement, and expert supervision have been invaluable to this work.

I am also deeply thankful to Albin Fredriksson and RaySearch Laboratories for their support, providing the necessary resources and an enabling environment to conduct this research.

Furthermore, I wish to extend my thanks to Anders Forsgren at KTH, whose expertise and guidance have greatly benefited this work.

My appreciation also goes to Hans Lööf at KTH for his valuable insights into the subject of industrial engineering.

Lastly, I want to thank my family for their unconditional support in all my adventures.
Till Patrik
## Contents

1 Introduction .................................................. 1

2 Radiation therapy ............................................ 4
   2.1 Delivery techniques and technological advancement ..... 4
   2.2 CyberKnife .................................................. 6
   2.3 Patient modeling ........................................... 8
   2.4 Dosage ...................................................... 9
   2.5 Dose-volume histogram .................................... 9
   2.6 Treatment time ............................................ 10

3 Optimization in radiation therapy ......................... 11
   3.1 Mathematical optimization ................................ 11
   3.2 Optimization variables .................................. 11
   3.3 Objective function ....................................... 11
   3.4 Constraints ................................................ 13
      3.4.1 Cardinality constraint ............................... 14

4 Methods ....................................................... 17
   4.1 Node filtration optimization .............................. 17
      4.1.1 Neighbor approach .................................... 18
      4.1.2 Greedy TSP approach ................................ 19
   4.2 Post-filtration optimization .............................. 19
   4.3 Time constraint ........................................... 20

5 Results ....................................................... 22
   5.1 Filtration algorithm ...................................... 22
   5.2 Post-filtration algorithm ................................. 30
   5.3 Time constraint ........................................... 30

6 Discussion ................................................... 37
   6.1 Algorithms ................................................ 37
      6.1.1 Filtration algorithm ................................... 37
      6.1.2 Post-filtration algorithm ............................ 38
      6.1.3 Time constraint ....................................... 38
   6.2 Implications .............................................. 39
   6.3 Further studies ........................................... 39
Notation and abbreviations

Table 1: Common notation used in this paper.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>Vector of optimization variables</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>Vector of beam fluences</td>
</tr>
<tr>
<td>$d$</td>
<td>Vector of dose distribution</td>
</tr>
<tr>
<td>$f$</td>
<td>Objective function</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Weighting factor</td>
</tr>
<tr>
<td>$(x)_+$</td>
<td>$\max{0, x}$</td>
</tr>
<tr>
<td>$\cdot$</td>
<td>Cardinality of a set</td>
</tr>
<tr>
<td>$G$</td>
<td>Group of bixels in a beam</td>
</tr>
<tr>
<td>$H$</td>
<td>Heaviside step function</td>
</tr>
<tr>
<td>$|\cdot|_\infty$</td>
<td>Maximum norm</td>
</tr>
<tr>
<td>$\mathcal{V}$</td>
<td>Set of voxels in an ROI</td>
</tr>
<tr>
<td>$\Delta v_i$</td>
<td>Relative volume of voxel $i$</td>
</tr>
<tr>
<td>$\mathbb{V}$</td>
<td>A set of nodes</td>
</tr>
<tr>
<td>$n_i$</td>
<td>Node $i$</td>
</tr>
<tr>
<td>$T$</td>
<td>Robot traversal time</td>
</tr>
</tbody>
</table>

Table 2: Common abbreviations used in this paper.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINAC</td>
<td>Linear Accelerator</td>
</tr>
<tr>
<td>MLC</td>
<td>Multileaf Collimator</td>
</tr>
<tr>
<td>3D-CRT</td>
<td>3D Conformal Radiotherapy</td>
</tr>
<tr>
<td>IMRT</td>
<td>Intensity-Modulated Radiation Therapy</td>
</tr>
<tr>
<td>SMLC</td>
<td>Static Multileaf Collimation</td>
</tr>
<tr>
<td>ROI</td>
<td>Region Of Interest</td>
</tr>
<tr>
<td>PTV</td>
<td>Planning Target Volume</td>
</tr>
<tr>
<td>OAR</td>
<td>Organs At Risk</td>
</tr>
<tr>
<td>DVH</td>
<td>Dose-Volume Histogram</td>
</tr>
<tr>
<td>FMO</td>
<td>Fluence Map Optimization</td>
</tr>
<tr>
<td>SQP</td>
<td>Sequential Quadratic Programming</td>
</tr>
<tr>
<td>TSP</td>
<td>Traveling Salesman Problem</td>
</tr>
</tbody>
</table>
1 Introduction

Cancer is after heart disease the leading cause of death in Sweden [1]. One in every third Swede is projected to be diagnosed with cancer once in their life [2]. Mortality rates are however going down, being reduced by 27% from 2001 and 2020 in the United States. The U.S. Department of Health and Human Services aims to reduce the mortality rate with another 15% to 2030 [3]. As cancer poses a great future challenge, cancer treatment will play an instrumental role.

The amount of cancer cases has increased steadily for the last 50 years and this will likely continue as a consequence of both population growth and an increased lifetime expectancy [1]. This will not only lead to increased costs but will also necessitate either more healthcare capacity or more efficient treatment plans.

There are many different types of cancer treatments, such as surgery, chemotherapy and radiation therapy. The latter is very cost-efficient and is used worldwide. Around half of all cancer patient require at least one course of radiation therapy during the treatment of their disease [4].

Radiation therapy is a cancer treatment that uses high-energy particles or waves, such as X-rays, gamma rays, electrons or protons to kill cancer cells. The radiation damages the DNA of the cancer cells, making them unable to divide and grow, and eventually leading to their death. The radiation is typically delivered using a linear accelerator (LINAC) or a radioactive source, and the dose and delivery method are carefully planned. The treatment can be given externally, where the radiation is directed at the tumor from outside the body, or internally, where a radioactive source is placed inside the body near the tumor.

Most radiation treatments deliver the dose externally using LINACs [5][6]. The technology in the field has evolved steadily since its conception in the 1950s, with a device called CyberKnife® (Accuray Incorporated, Sunnyvale, CA, USA) being introduced in 1994 [7]. CyberKnife is a type of radiosurgery system that uses a robotic arm to irradiate cancerous tissues with
sub-millimeter precision. The system uses advanced imaging techniques to track the movement of the tumor in real-time, allowing for precise targeting of the radiation beam [8].

An ever-present problem in radiation therapy is the trade-off between delivering sufficient radiation to the tumor and minimizing the radiation damage to the surrounding tissue. This, combined with the many degrees of freedom, makes mathematical optimization an important tool in this discipline. An objective function to be optimized is constructed so that it correlates with the clinical goals. Optimization algorithms are then used to solve the formulated problem, generating suitable treatment plans.

This thesis investigates how the treatment time for plans using CyberKnife technology can be reduced. Such a reduction would not only be beneficial for healthcare providers, but would also provide less taxing treatments for patients. Faster delivery times also entail the possibility of using more degrees of freedom, potentially increasing the treatment quality.

The research questions to be considered are:

How can the patient treatment time of CyberKnife be reduced while still maintaining satisfactory treatment coverage? Can increased delivery efficiency also be used to improve plan quality?

There has been some previous research done regarding time reductions for CyberKnife treatments. CyberArc, an optimization algorithm introduced by Kearney et al., uses continuous beam delivery instead of the conventional static step-and-shoot delivery resulting in drastic time reductions [9]. While promising, it remains a theoretical study as of writing since the CyberKnife system is not yet capable of beam delivery during robot movement. Another approach, suggested by van der Water et al., achieves time reductions by decreasing the number of irradiation directions required [10]. The time for treatment planning however rises drastically with the algorithm requiring several hours to complete. Bedford et al. also look at finding beam directions associated with shorter treatment times, succeeding by employing an evolutionary algorithm for beam selection [11]. This algorithm however also
requires hours to finish. Another, less computationally heavy approach is thus requested.

The thesis is conducted in collaboration with RaySearch Laboratories AB (Stockholm, Sweden), a company that develops software for radiation therapy. Most of the work is carried out in Raysearch’s treatment planning system RayStation, which is a software application that generates radiation therapy treatments. Introduced algorithms are implemented in the C++ source code for RayStation.
2 Radiation therapy

In the late 19th century, French physicist Antoine Becquerel found that certain minerals emit radiation, marking the discovery of radioactivity [12]. The first use of radiation for cancer treatment was in 1896, shortly after German physicist Wilhelm Roentgen discovered X-rays [13]. Since then, radiation therapy has become an essential tool in the treatment of cancer, with ongoing advancements in technology improving the efficacy and precision of treatments. This section aims to introduce the basics of radiation therapy and detail some of the technological advancements in the field.

Radiation therapy aims to deliver high-energy radiation to cancerous tissues, with the intent to kill cancer cells or halt their growth. The energy used to destroy cancer cells can be delivered in the form of high-energy X-ray photons or charged particles, such as electrons, protons, or carbon ions. These can penetrate the tumor tissue to deliver the radiation dose. The treatment plan is designed to deliver a sufficient dose to the cancerous tissue while minimizing damage to healthy tissue surrounding the cancerous area. Ultimately, the goal is to improve the patient’s overall survival chance and quality of life by shrinking or eliminating the tumor, relieving symptoms, and preventing the cancer from recurring.

2.1 Delivery techniques and technological advancement

The most common way to deliver external radiation therapy is using photons and different delivery techniques have been developed throughout the years [14]. In external radiation therapy, the patient is typically irradiated from various angles. This method is used to concentrate a high dose of radiation in the tumor, while minimizing the dose reaching the surrounding healthy tissue. The patient is placed on a couch, allowing a gantry – which holds the linear accelerator – to rotate around them, as illustrated in Figure 1.
Another pivotal technology in radiation therapy is the multileaf collimator (MLC). This device is composed of leaves - movable metal components designed to block part of the radiation beam. This provides greater control over the radiation that reaches the patient, as illustrated in Figure 2. In which way the LINAC, gantry and MLC are used result in different radiotherapy techniques.

Figure 2: MLC leaves shaping the beam around the outline of a tumor (in red) from different angles. Each leaf is marked with its respective number. The bright square details the end of the so-called jaws, which are collimation devices that further stop undesired radiation. From RayStation.
The invention of computed tomography (CT) in 1972 made 3D imaging and 3D dose calculation available, allowing for greater precision [17]. Subsequently, 3D conformal radiotherapy (3D-CRT) was developed. With a 3D image detailing the structure of the tumor, the patient is irradiated from several different angles by rotating the gantry. The beam is conformed to the projection of the treated volume using the MLC technology in order to reduce unnecessary radiation.

*Intensity-modulated radiation therapy* (IMRT) evolved as a generalization of 3D-CRT. By irradiating with several MLC shapes from the same beam configuration, the profile of the radiation can be modulated. Each shape is referred to as a beam segment. While IMRT allows for more satisfactory dose distributions, it is important to note that it is also more costly due to increased treatment times and the complexity of beam delivery [18]. Longer treatment times affect both the healthcare provider and the patient. Clinics will require more man-hours, and their throughput decreases, resulting in extended waiting lists. The patient is also more likely to move during longer treatments, which may lead to decreased treatment results.

There are different techniques for delivering IMRT treatments. One example is *static multileaf collimation* (SMLC). Such a plan consists of using different beams at different gantry angles. Each beam has a set of segments that are defined by their respective MLC configurations. An exposition into IMRT and different delivery techniques can be found in [19].

### 2.2 CyberKnife

The CyberKnife system is a fully robotic radiotherapy device, featuring a LINAC mounted on a robot as illustrated in Figure 3. Invented in the 1990s, it is the first and only device of its kind. The robot moves around the patient and is able to deliver radiation with high precision. Although CyberKnife technology is not currently available in Sweden as of writing, hundreds of facilities in the U.S. and Europe are using this advanced treatment method [8].
The robot does not move freely in 3D-space but instead traverses through a predetermined set of nodes, illustrated in Figure 4. Each node may be viewed similarly as a gantry angle, and the principles of SMLC as described earlier hold the same. Given a set of nodes to irradiate from, it is likely that the robot will need to traverse through nodes from which it does not irradiate. A node from which the robot emits radiation is henceforth referred to as a dose node, with non-dose nodes being referred to as dummy nodes. Each node has a set of neighbors to which the robot can traverse directly. The time to traverse between two nodes varies, but is usually around 5-15 seconds. Treatments use different node sets depending on if the tumor is in the patients head or body, with the head set consisting of around 200 nodes (of which it is possible to irradiate from around 150) and the body set of around 100 nodes (of which it is possible to irradiate from around 80).
Figure 4: The CyberKnife workspace. The device can irradiate only from each dose node, illustrated as orange beams in the figure [21].

2.3 Patient modeling

In order to generate a radiotherapy plan, a 3D image of the patient is required. This patient image is discretized into $10^6 - 10^7$ three-dimensional cuboids called voxels.

These voxels are divided into two categories, regions of interest (ROI) and external tissue. The regions of interest are then further divided into two categories. First are the target volumes, which are subdivided into

- The gross target volume (GTV), which consists of the observable tumor.
- The clinical target volume (CTV), which expands the GTV in order to account for microscopic spread not discernible in the patient image.
- The planning target volume (PTV), which includes the CTV and expands it to account for uncertainties in the treatment delivery and patient anatomy.

The other ROI subdivision is referred to as organs at risk (OAR). These are organs located close to the tumor that ideally should receive as little radiation exposure as possible.
2.4 Dosage

The dose that the patient receives in different ROIs is central in radiotherapy planning. Dose is measured in the unit Gray (Gy) and is the energy absorbed per unit mass, according to $1 \text{ Gy} = 1 \text{ J/kg}$. It can be viewed as a scalar field $d : \mathbb{R}^3 \to \mathbb{R}_{\geq 0}$ since it is nonnegative. The idea of discretizing the patient is also applied to the cross-sections of the beams, which are discretized into a 2D grid of around 1000 so-called *bixels*. Collecting the beam fluences (radiation per unit area, $\text{J/cm}^2$) in each bixel into the vector $\varphi$ and the dose in each voxel into the vector $d$ gives the dose distribution by the linear relationship

$$d = d(\varphi) = P\varphi$$

where $P$ is the *dose deposition matrix* which can be computed based on patient geometry. Although not directly controllable in photon IMRT, the beam fluences can be found from a fluence map by collecting the MLC leaf positions and the respective segmentation weightings in a vector $x$, i.e. $\varphi = \varphi(x)$. The machine parameters $x$ are thus mapped to the final dose distribution according to

$$x \mapsto \varphi \mapsto d.$$

2.5 Dose-volume histogram

Dose-volume histograms (DVHs) are a graphical representation of the distribution of radiation dose in a given volume of tissue. DVHs are typically generated by calculating the cumulative dose received by a specified volume of tissue, plotted as a function of that volume. For example, a DVH may show the percentage of normal tissue receiving a specific dose level or the percentage of the tumor receiving a minimum dose level. An example DVH is illustrated in Figure 5.
Figure 5: Dose-volume histogram. A point \((x, y)\) on a graph shows that \(y\)% of an ROI receives at least \(x\) units of Gy. From RayStation.

DVHs can provide important information to help optimize treatment plans. The graphical visualization allows for the entirety of the 3D dose to be evaluated by simply inspecting the graph. DVHs compress the relevant information by removing the spatial component of where in the ROI that dose is delivered. This significantly simplifies evaluation and comparison of treatment plans.

### 2.6 Treatment time

The total treatment time of a given plan can be decomposed into several parts. Some notable components are

- The robot traversal time (henceforth denoted \(T\)). This represents the time required for the robot to visit the specified dose nodes.
- The time for the MLC leaves to move between segments.
- Beam on time (the time for the LINAC to deliver the radiation).
- The time for patient imaging.

This thesis focuses on reducing the robot traversal time \(T\).
3 Optimization in radiation therapy

3.1 Mathematical optimization

In mathematical optimization, the considered problem is of the form

\[
\min_{x \in \mathbb{R}^n} \ f(x)
\]

\[
\text{s.t. } g_i(x) \leq 0, \ i = 1, \ldots, m_1 \quad (P)
\]

\[
h_i(x) = 0, \ i = 1, \ldots, m_2
\]

where \( x \in \mathbb{R}^n \) are the optimization variables and \( f : \mathbb{R}^n \to \mathbb{R} \) is the objective function. The set of functions \( g_i : \mathbb{R}^n \to \mathbb{R} \) are called inequality constraints and the set of functions \( h_i : \mathbb{R}^n \to \mathbb{R} \) are called equality constraints. These together define a feasible set \( \mathcal{F} \subseteq \mathbb{R}^n \) consisting of the points \( x \in \mathbb{R}^n \) that satisfy all constraints. The following sections introduce the specifics in how \( \mathcal{P} \) is constructed in radiation therapy.

3.2 Optimization variables

The optimization variables \( x \) consists of the machine parameters (the weights and corresponding MLC leaf positions of all segments). Using the fluence map \( \varphi \), the machine parameters are mapped to beam fluences. The fluences are then mapped to the dose \( d \) using the dose deposition matrix \( P \), see Section 2.4 Dosage. The relationship obtained is thus

\[
d = d(\varphi(x)) = P\varphi(x).
\]

3.3 Objective function

The mathematical objective function \( f \) of consideration maps the dose distribution \( d \) to a scalar. A lower value of \( f \) corresponds to a higher treatment plan quality. \( f \) is a weighted composition of penalty functions \( f_i \) according to

\[
f(d) = \sum_{i \in \kappa} w_i f_i(d)
\]

where \( \kappa \) is an index set corresponding to patient-specific needs and \( w_i \) are some scalars used for weighting. Some different types of objective functions \( f_i \) are detailed below.
The aim of these is to for each voxel $i$ in an ROI’s voxel set $\mathcal{V}$ quadratically penalize deviations from a reference dose $d_{\text{ref}}$. Each penalty is normalized by the relative volume of the voxel $\Delta v_i$ as well as the reference dose so that all voxels are equally important until the weights $w_i$ are introduced.

**Uniform dose function**

A uniform dose optimization function aims to penalize both under- and overdosage. This is formulated as

$$\sum_{i \in \mathcal{V}} \Delta v_i \left( \frac{d_i - d_{\text{ref}}}{d_{\text{ref}}} \right)^2.$$ 

**Minimum and maximum dose functions**

A minimum dose optimization function aims to penalize underdosage. This is formulated as

$$\sum_{i \in \mathcal{V}} \Delta v_i \left( \frac{d_{\text{ref}} - d_i}{d_{\text{ref}}} \right)^2$$

while a maximum dose function instead penalizes overdosage according to

$$\sum_{i \in \mathcal{V}} \Delta v_i \left( \frac{d_i - d_{\text{ref}}}{d_{\text{ref}}} \right)^2$$

where $(x)_+ = \max(x, 0)$.

**Minimum and maximum DVH functions**

DVH functions are another common type of objective function, with basis in the DVH-curve (see Section 2.5 Dose-volume histogram). These objectives usually specify a threshold of how much dose that reaches a specified amount of a certain region ("No more than 40% of the liver should receive more than 30 Gy") [22]. Letting $\hat{d}$ be the dose that reaches the specified volume level $\hat{v}$ gives the maximum DVH function according to

$$\sum_{i \in \mathcal{V}} \Delta v_i \left( \frac{d_i - d_{\text{ref}}}{d_{\text{ref}}} \right)^2 \cdot H(d_i - \hat{d})$$

where $H(x) = 1_{\{x > 0\}}$ is the Heaviside step function. The definition of the minimum DVH function is analogous.
3.4 Constraints

The constraints imposed on $\mathcal{P}$ can be categorized into delivery constraints and planning constraints. Delivery constraints consist of limitations in the specific delivery technique used, such as the movement restrictions of the MLC. This results in a feasible set $\mathcal{X}$ for the machine parameters $x$.

Planning constraints consist of clinical goals that are non-negotiable, i.e. cannot be violated. All functions described in the previous subsections may instead be treated as constraints, which increases their importance. This is represented by $g(d) \leq 0$, where the exact composition of $g$ depends on the patient case and the respective clinical goals.

Since neither beam fluence nor dose can be negative, the additional constraint of non-negativity is imposed on the segmentation weightings in $x$. This is included implicitly in the constraint $x \in \mathcal{X}$. The usual problem considered in SMLC can thus be formulated as

$$\min_{x \in \mathcal{X}} \quad f(d)$$

s.t. $g(d) \leq 0$  \hspace{1cm} (P)

$$d = d(\varphi(x)).$$

Note that the optimization variables are the machine parameters (see Section 3.2 Optimization variables) and that the gantry angles are predetermined based on patient geometry. This problem is generally hard to solve since the mapping $\varphi(x)$ is nonlinear and nonconvex. The optimization problem is divided into two phases in order to be solved. Firstly, the fluences are treated as directly controllable in what is called fluence map optimization (FMO). This problem is formulated as

$$\min_{\varphi \in \phi} \quad f(d)$$

s.t. $g(d) \leq 0$  \hspace{1cm} (FMO)

$$d = d(\varphi)$$

where $\phi$ is the set of possible fluences. The resulting fluence $\hat{\varphi}$ obtained by the FMO is then converted to a deliverable plan of machine parameters by leaf sequencing. The fluence map conversion in general results in a larger
value of $f$ due to the inability of the discrete MLC leafs to reproduce the continuous fluence distribution. A sequential quadratic programming (SQP) algorithm is used to solve the optimization problems. An exposition into SQP can be found in [23].

3.4.1 Cardinality constraint

Instead of determining gantry angles (or nodes in the case of CyberKnife treatments) before the optimization is started, they may be included as optimization variables. Thus, only the amount of beam configurations (gantry angles or nodes, with the latter being used henceforth) to use must be specified. To incorporate this into the problem, the optimization starts with the full set of nodes. Then, a so called cardinality constraint is introduced. Let $G$ be a grouping of the beam fluence bixels in $\varphi$ such that one group corresponds to all the bixels of the beam from one node. The constraint then takes the form

$$\left\{ G : \sum_{\varphi_i \in G} \varphi_i > 0 \right\} \leq \aleph$$

where $\cdot$ denotes the cardinality of a set and $\aleph$ is a positive whole number specifying the amount of nodes. This imposes a restriction on the amount of non-zero beam fluences, i.e. the maximum amount of nodes $\aleph$. The new FMO problem is thus on the form

$$\min_{\varphi \in \phi} f(d)$$

s.t. $g(d) \leq 0$

$$d = d(\varphi)$$

$$\left\{ G : \sum_{\varphi_i \in G} \varphi_i > 0 \right\} \leq \aleph.$$

FMO-CC

Cardinality constraints are generally neither convex nor continuous, making the problem more complex. The constraint is handled heuristically in the FMO by firstly rewriting it using the Heaviside step function according to

$$\left\{ G : \sum_{\varphi_i \in G} \varphi_i > 0 \right\} = \sum_{G} H \left( \sum_{\varphi_i \in G} \varphi_i \right).$$
The fact that the Heaviside step function has a zero derivative everywhere except at zero, where it is undefined, makes this expression ill-suited for gradient-based optimization methods such as SQP. A smooth approximation is needed to facilitate its inclusion in the optimization problem. The approximating function used in this context is the logistic sigmoid, defined as

$$
\zeta_\varepsilon(x) = \frac{1}{1 + e^{-x/\varepsilon}}
$$

where $\varepsilon$ is a smoothness parameter. A smaller $\varepsilon$ corresponds to a better approximation, but is harder for the optimization algorithm to handle. The sigmoid function is illustrated in Figure 6.

![Approximation of the Heaviside step function using a logistic sigmoid function, with different values for $\varepsilon$. Courtesy of [22].](image)

Since beam fluences are non-negative, the sigmoid function is scaled and shifted so that its output is in the interval $[0,1]$. The constraint is therefore approximated as

$$
\sum_{G} \left( 2 \cdot \zeta_\varepsilon \left( \sum_{\varphi_i \in G} \varphi_i \right) - 1 \right) \leq N
$$

which is thus a sum of smooth functions. This is lastly incorporated into a Lagrange relaxed problem FMO-LR as
\[
\min_{\mathbf{\varphi} \in \Phi} \ f(\mathbf{d}) + \lambda \cdot \left( \sum_G \left( 2 \cdot \zeta_\varepsilon \left( \sum_{\mathbf{\varphi}_i \in G} \mathbf{\varphi}_i \right) - 1 \right) \right) - \Re \]  

s.t.  
\[
\mathbf{g}(\mathbf{d}) \leq 0 \\
\mathbf{d} = \mathbf{d}(\mathbf{\varphi})
\]

where \(\lambda\) is some positive penalty term. To reduce the full node set, several filtrations are performed during the FMO. In each filtration, the value \(2 \cdot \zeta_\varepsilon \left( \sum_{\mathbf{\varphi}_i \in G} \mathbf{\varphi}_i \right) - 1\) is calculated for each group \(G\). The groups with the lowest values are removed since beams with less fluence are assumed to contribute less to the objective function. This may be stated as in each filtration solving  
\[
\min_{\mathbf{n}_i \in \mathbb{V}} \ 2 \cdot \zeta_\varepsilon \left( \sum_{\mathbf{\varphi}_i \in \mathcal{G}(\mathbf{n}_i)} \mathbf{\varphi}_i \right) - 1 \quad (I)
\]

repeatedly, where \(\mathbb{V}\) is the remaining set of nodes and the function \(\mathcal{G}\) maps each node to its corresponding group of bixel fluences. Around 10-15 nodes are removed each filtration and the FMO is then restarted. This is repeated until the constraint is satisfied.

Once the final set of dose nodes \(\mathbb{V}\) has been determined, the problem arises of how to find the traversal path which visits all the nodes and minimizes the traversal time. This is an example of the classic Traveling Salesman Problem (TSP), which seeks to find the shortest possible route that visits some given cities at least once. To solve this problem, a heuristic TSP algorithm is used. The algorithm is a version of the Lin-Kernighan-Helsgaun-2 algorithm that starts with a randomly generated tour and iteratively improves it by making small modifications to the tour’s edges. It uses a combination of edge swapping and a local search procedure to improve the quality of the solution and is widely used for solving TSPs. A detailed explanation of the algorithm can be found in [24].
4 Methods

The basis of this section is in the study of the problem FMO-CC together with the associated traversal time \( T(\varphi) \). A weighted optimization problem FMO-T taking into account the traversal time \( T \) can be formulated as

\[
\min_{\varphi \in \Phi} \alpha \cdot f(d) + (1 - \alpha) \cdot T(\varphi) \\
\text{s.t.} \quad g(d) \leq 0 \\
\quad d = d(\varphi) \\
\quad |\{G : \sum_{\varphi_i \in G} \varphi_i > 0\}| \leq \mathbb{N}
\]

(FMO-T)

for some weight \( \alpha \in [0, 1] \). The task of reducing the treatment time is tackled by algorithms using the structure of CyberKnife’s node network. These aim to find sets of dose nodes associated with shorter treatment times, all the while keeping the dose distribution satisfactory. Algorithms introduced are used during the filtration in the FMO.

4.1 Node filtration optimization

As detailed in Section 3.4.1 Cardinality constraint, dose nodes are selected by starting with the entire set and sequentially removing nodes. The problem

\[
\min_{n_i \in V} 2 \cdot \zeta \left( \sum_{\varphi_i \in \mathcal{G}(n_i)} \varphi_i \right) - 1
\]

is solved each filtration. In order to heuristically solve FMO-T, the problem \( \mathcal{I} \) is extended to

\[
\min_{n_i \in V} \alpha \cdot \left( 2 \cdot \zeta \left( \sum_{\varphi_i \in \mathcal{G}(n_i)} \varphi_i \right) - 1 \right) + (1 - \alpha) \cdot \tau(n_i)
\]

(\( \mathcal{I}_T \))

where the function \( \tau \) to each node \( n_i \) assigns a weight that means to correlate with the traversal time of the final node set.

Two different ways to model \( \tau \) are developed. The first model heuristically assigns a weight to each dose node based on its importance in the network.
This is henceforth referred to as the *neighbor approach*, since a node and its neighbors are considered.

The second model calculates $T(\mathbb{V} \setminus n_i)$ directly by computing the TSP on the generated set $\mathbb{V} \setminus n_i$ for each $n_i \in \mathbb{V}$ in every iteration. This is a greedy approach that assigns a weight to each dose node based on how much removing that node reduces the time required to traverse the current set. This will henceforth be referred to as the *greedy TSP* approach.

### 4.1.1 Neighbor approach

The neighbor approach is a heuristic method that assigns a weight to each dose node based on its importance in the network. It is thus a sort of centrality measure which is a commonly used concept in network analysis. An exposition into other such measures can be found in [25]. The algorithm rewards nodes that:

1. Have many neighbors, i.e. can access many other nodes directly.
2. Have low traversal times to other nodes.

Every dose node $n_i$ is assigned a weight $w_i$. Let $t_{ij}$ denote the traversal time from node $n_i$ to node $n_j$. $t_{ij}$ is set equal to $-1$ if traversal between node $n_i$ and $n_j$ directly is impossible. Let

$$\mathcal{N}_i = \{ j : t_{ij} \neq -1 \}$$

be the *neighbor set* containing the indices of all the neighbors to node $n_i$. Each dose node is then assigned the weight

$$w_i = \sum_{j \in \mathcal{N}_i} \frac{1}{t_{ij}}.$$

This gives nodes with many neighbors and low traversal times a high weight. Letting $\mathbf{w}$ be the vector of weights, a scaling is finally performed according to

$$\mathbf{w}^* = \frac{\mathbf{w}}{\| \mathbf{w} \|_{\infty}}$$

where $\| \mathbf{w} \|_{\infty}$ is the max-norm. This ensures that $w_i \in [0, 1]$. A lower weight corresponds to a dose node which is more desirable to remove. The weights
are recalculated after node removals to only take into account the remaining nodes.

This approach is very computationally effective. It requires only to loop through the matrix of traversal times of size around $10^2$, which is done all but instantaneously.

### 4.1.2 Greedy TSP approach

The second approach to model $\tau$ is a weighting system based on solving several TSPs. The reference time $T_0 = T(\mathcal{V})$ to traverse the entire remaining set of dose nodes $\mathcal{V}$ is in each filtration computed. Then, for each dose node $n_i$, the time $T_i = T(\mathcal{V} \setminus n_i)$ to traverse the entire set without that node is computed. The weight $t_i$ is lastly assigned to each node according to

$$t_i = T_i - T_0$$

and the vector $t$ of weights is normalized using the max-norm

$$t^* = \frac{t}{\|t\|_\infty}.$$

The weights thus satisfy $t_i \in [-1, 0]$ where a lower weight once again corresponds to a dose node which is more desirable to remove.

This approach is quite computationally demanding. As an example, consider a set of 150 dose nodes to be reduced to 50 by removing 10 per iteration. This would require 1050 TSPs to be solved in total. Using parallel computing and the heuristic TSP solver detailed in Section 3.4.1 Cardinality constraint makes the computational time not instantaneous, but manageable.

### 4.2 Post filtration optimization

Given a final set of dose nodes $\mathcal{V}$ (of the predetermined size), this section develops a method to fine tune this set with respect to time and objective value. This is done during the FMO, but post filtration.

The method adopted is a local-search method that aims to find better solutions that are similar to the original. For each node $n_i$, its immediate
neighbors are considered as potential alternatives. The immediate neighborhood $D_i$ of a node $n_i$ consists of the nodes $n_j$ such that:

1. $n_j$ is not already in $\mathbb{V}$.
2. $n_j$ is a dose node.
3. $n_j$ is a neighbor to and sufficiently close to $n_i$, i.e. $t_{ij} < t_{max}$ for some $t_{max}$. The value used for $t_{max}$ is 5 seconds.

A dose node $n_i \in \mathbb{V}$ is tested by evaluating

$$\Delta F_i(n_j) = F((\mathbb{V} \setminus n_i) \cup n_j) - F(\mathbb{V})$$
$$\Delta T_i(n_j) = T((\mathbb{V} \setminus n_i) \cup n_j) - T(\mathbb{V}).$$

where $F(\mathbb{V})$ is the optimal solution to the problem FMO restricted to the node set $\mathbb{V}$. A negative value of both $\Delta F_i(n_j)$ and $\Delta T_i(n_j)$ means that switching $n_i$ for $n_j$ results in both a smaller objective value and a shorter traversal path. The algorithm implemented is greedy in that it immediately switches $n_i$ for $n_j$ if both quantities above are negative. It then continues and does the same thing on the new set $\mathbb{V}' = (\mathbb{V} \setminus n_i) \cup n_j$.

In order to evaluate $\Delta F_i(n_j)$, the value $F((\mathbb{V} \setminus n_i) \cup n_j)$ needs to be calculated. This requires restarting the FMO. A phase after the node filtration but before the fluence map conversion is thus added, allowing the optimizer to run the fluence optimization and evaluate the different $\Delta F_i(n_j)$.

### 4.3 Time constraint

Instead of extending the problem FMO-CC into the weighted optimization problem FMO-T as in previous subsections, it may instead be extended by considering the treatment time as a constraint. This however poses several problems. Since the treatment time is not directly controllable, a decomposition of its components (see Section 2.6 Treatment time) is required. The interdependencies between these quantities makes it even more complicated. Instead, only the traversal time is introduced as a constraint, since it can be directly measured during the optimization. The need for specifying the amount of dose nodes to include in the plan is thus eliminated, since this is handled by the traversal time constraint. This new problem is stated as
\[
\begin{align*}
\min_{\varphi \in \phi} & \quad f(d) \\
\text{s.t.} & \quad g(d) \leq 0 \\
& \quad d = d(\varphi) \\
& \quad T(\varphi) \leq T_0.
\end{align*}
\] (FMO-TC)

This is solved in the same manner as the problem FMO-CC, by reducing the full node set node by node. When a node set \( \mathcal{V} \) satisfying \( T(\mathcal{V}) \leq T_0 \) is reached the FMO optimization terminates. To accomplish this, several TSPs are solved during the node reduction.

Introducing the problem FMO-TC further allows for the evaluation of the algorithms presented in Section 4.1 Node filtration optimization. Instead of using these algorithms to simply reduce the traversal time (at a likely cost of plan quality) for a predetermined amount of nodes, they may allow for more nodes to be used while still keeping the traversal time below \( T_0 \). The increased amount of nodes has the possibility to compensate for the worsened objective value that results from optimizing on the traversal time. This would entail the possibility of using the node reduction algorithms not only to reduce the treatment time, but also to generate higher quality plans.
5 Results

This section details the results of the thesis. The introduced optimization algorithms are tested on three different patient cases. Each patient case already contains a finished treatment plan which is considered as the reference plan. All predetermined variables are kept the same as in the reference plan, and the plan generated by the new algorithm is compared to the reference plan.

The three different patient cases used for testing are:

1. Patient I: Brain tumor. The objective function mainly aims to deliver the prescribed dose to the tumor in a conformal manner. Minimum dose of 4000 cGy and maximum dose of 5900 cGy in the PTV.

2. Patient II: Brain tumor. Similar objective as Patient I. Minimum dose of 2500 cGy and maximum dose of 3150 cGy in the PTV.

3. Patient III: Liver tumor. Minimum dose of 4500 cGy and maximum dose of 5600 cGy in the PTV. Notable OARs are the small intestine and a kidney, which are represented in the objective with maximum DVH functions.

5.1 Filtration algorithm

The filtration algorithm consists of using the time weighting system during node filtration. Although both the neighbor approach and the greedy TSP approach performed well during testing (and none performed consistently better than the other), the latter is chosen when presenting the data below. This is done for three reasons. The first reason is that the greedy TSP approach explicitly computes the actual traversal time and awards weights based on this. Having a direct connection to the end goal of a shorter traversal time is an upside for this approach compared to the neighbor approach which only measures this quantity implicitly. The second reason is that the TSP approach is similar to what is done in the post-filtration optimization. Lastly, the amount of plans to generate was fairly small, so the computational disadvantage of the TSP approach wasn’t inhibiting.
For each patient, the reference plan is compared to the new, time-optimized plan. Relevant data from the comparison is included, which covers:

- Objective value, which gives a rough indication of the plan quality. Note that a smaller objective value correspond to increased plan quality.

- Treatment time, which is the total time in seconds required to deliver the plan.

- Traversal time, which is the time in seconds required for the robot to traverse the set of dose nodes.

- Average traversal time, which is the average time in seconds to traverse two nodes in the set.

- The amount of dummy nodes, which are nodes from which the robot does not irradiate, but has to traverse through.

DVH-curves, the dose distribution as well as the robot traversal path for the respective plans are included as figures.

The weight $\alpha = 0.5$ in $\text{FMO-T}$ and $I_T$ is used in all cases, since this gave a substantial effect during testing. Other weights were also successfully tried, although not presented here due to limitation in scope.
Patient I

Table 3: Different plans for Patient I. 40 nodes and 50 segments used.

<table>
<thead>
<tr>
<th></th>
<th>Reference plan</th>
<th>Time-optimized plan</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective value</td>
<td>0.0173</td>
<td>0.0192</td>
<td>+11%</td>
</tr>
<tr>
<td>Treatment time (s)</td>
<td>1153</td>
<td>844</td>
<td>-27%</td>
</tr>
<tr>
<td>Traversal time (s)</td>
<td>625</td>
<td>333</td>
<td>-47%</td>
</tr>
<tr>
<td>Average traversal time (s)</td>
<td>7.8</td>
<td>5.8</td>
<td>-26%</td>
</tr>
<tr>
<td>Dummy nodes</td>
<td>41</td>
<td>18</td>
<td>-56%</td>
</tr>
</tbody>
</table>

Figure 7: DVH-curves for the PTV and the whole brain for Patient I. The continuous line is the reference plan and the dashed line is the new plan.
Figure 8: Dose distribution for Patient I. Centered at the maximal value for the reference plan. PTV outlined in black.

Figure 9: Robot traversal path for Patient I. Dose nodes are colored orange and dummy nodes are colored red.
Patient II

Table 4: Different plans for Patient II. 45 nodes and 60 segments used.

<table>
<thead>
<tr>
<th></th>
<th>Reference plan</th>
<th>Time-optimized plan</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective value</td>
<td>0.0098</td>
<td>0.0106</td>
<td>+ 8%</td>
</tr>
<tr>
<td>Treatment time (s)</td>
<td>1173</td>
<td>898</td>
<td>-23%</td>
</tr>
<tr>
<td>Traversal time (s)</td>
<td>630</td>
<td>328</td>
<td>-48%</td>
</tr>
<tr>
<td>Average traversal time</td>
<td>7.4</td>
<td>6.0</td>
<td>-19%</td>
</tr>
<tr>
<td>Dummy nodes</td>
<td>41</td>
<td>11</td>
<td>-73%</td>
</tr>
</tbody>
</table>

(a) PTV. Objective: dose between 2500 and 3150 cGy.

(b) Whole brain.

Figure 10: DVH-curves for the PTV and the whole brain for Patient II. The continuous line is the reference plan and the dashed line is the new plan.
Figure 11: Dose distribution for Patient II. Centered at the maximal value for the reference plan. PTV outlined in black.

Figure 12: Robot traversal path for Patient II. Dose nodes are colored orange and dummy nodes are colored red.
Patient III

Table 5: Different plans for Patient III. 40 nodes and 70 segments used.

<table>
<thead>
<tr>
<th></th>
<th>Reference plan</th>
<th>Time-optimized plan</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective value</td>
<td>0.0552</td>
<td>0.0641</td>
<td>+16%</td>
</tr>
<tr>
<td>Treatment time (s)</td>
<td>2004</td>
<td>1783</td>
<td>-11%</td>
</tr>
<tr>
<td>Traversal time (s)</td>
<td>402</td>
<td>283</td>
<td>-30%</td>
</tr>
<tr>
<td>Average traversal time (s)</td>
<td>7.3</td>
<td>6.0</td>
<td>-18%</td>
</tr>
<tr>
<td>Dummy nodes</td>
<td>16</td>
<td>8</td>
<td>-50%</td>
</tr>
</tbody>
</table>

(a) PTV. Objective: dose between 4500 and 5600 cGy.

(b) Small intestine (green) and kidney (brown).

Figure 13: DVH-curves for the PTV and some OARs. The continuous line is the reference plan and the dashed line is the new plan.
Figure 14: Dose distribution for Patient III. Centered at the maximal value for the reference plan. PTV outlined in black, small intestine in green and kidney in brown.

Figure 15: Robot traversal path for Patient III. Dose nodes are colored orange and dummy nodes are colored red.
5.2 Post-filtration algorithm

The post-filtration algorithm is run on the different patient cases, both on the reference plan and the time-optimized plan. It is allowed to run until convergence, i.e. until no local perturbation results in an improvement.

For all plans, the algorithm converges quickly, having to try around 5-20 different node configurations and finding a better solution only once or twice. The improvement when finding a better node configuration is also negligible, resulting in a reduction of less than 1% for the objective function and a couple of seconds for the traversal time. The fast convergence and small differences however show that the node configurations generated beforehand (by both the reference plan and the time-optimized plan) are very close to being locally optimal.

5.3 Time constraint

To test the possibility of using the filtration algorithm in the problem FMO-TC in order to generate higher quality plans, a similar analysis as above is done. Using the reference plan for the different patient cases gives a reference traversal time $T_{\text{ref}}$, which is used as the time constraint in FMO-TC (i.e. $T_0 = T_{\text{ref}}$) for each respective patient. The problem

$$\min_{\varphi \in \phi} f(d)$$

s.t. $g(d) \leq 0$

$d = d(\varphi)$

$T(\varphi) \leq T_{\text{ref}}$

is thus solved, with the TSP approach used during node filtration. Although the algorithm terminates immediately when $T(\mathcal{V}) \leq T_{\text{ref}}$, the results show a larger than expected difference between $T(\mathcal{V})$ and $T_{\text{ref}}$. This is a consequence of a different TSP solver being used when the plan is finished than the one used during optimization.
Patient I

Table 6: Different plans for Patient I. $T_{\text{ref}} = 625$.

<table>
<thead>
<tr>
<th></th>
<th>Reference plan</th>
<th>Time-constrained plan</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose nodes</td>
<td>40</td>
<td>75</td>
<td>+88%</td>
</tr>
<tr>
<td>Segments</td>
<td>50</td>
<td>75</td>
<td>+50%</td>
</tr>
<tr>
<td>Objective value</td>
<td>0.0173</td>
<td>0.0137</td>
<td>-21%</td>
</tr>
<tr>
<td>Treatment time (s)</td>
<td>1153</td>
<td>1251</td>
<td>+8%</td>
</tr>
<tr>
<td>Traversal time (s)</td>
<td>625</td>
<td>581</td>
<td>-7%</td>
</tr>
<tr>
<td>Average traversal time (s)</td>
<td>7.8</td>
<td>5.9</td>
<td>-24%</td>
</tr>
<tr>
<td>Dummy nodes</td>
<td>41</td>
<td>24</td>
<td>-41%</td>
</tr>
</tbody>
</table>

(a) PTV. Objective: dose between 4000 and 5900 cGy.

(b) Whole brain.

Figure 16: DVH-curves for the PTV and the whole brain for Patient I. The continuous line is the reference plan and the dashed line is the new plan.
Figure 17: Dose distribution for Patient I. Centered at the maximal value for the reference plan. PTV outlined in black.

Figure 18: Robot traversal path for Patient I. Dose nodes are colored orange and dummy nodes are colored red.
Patient II

Table 7: Different plans for Patient II. $T_{\text{ref}} = 630$.

<table>
<thead>
<tr>
<th></th>
<th>Reference plan</th>
<th>Time-constrained plan</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose nodes</td>
<td>45</td>
<td>81</td>
<td>+80%</td>
</tr>
<tr>
<td>Segments</td>
<td>60</td>
<td>81</td>
<td>+35%</td>
</tr>
<tr>
<td>Objective value</td>
<td>0.0098</td>
<td>0.0043</td>
<td>-56%</td>
</tr>
<tr>
<td>Treatment time (s)</td>
<td>1173</td>
<td>1233</td>
<td>+5%</td>
</tr>
<tr>
<td>Traversal time (s)</td>
<td>630</td>
<td>594</td>
<td>-6%</td>
</tr>
<tr>
<td>Average traversal time (s)</td>
<td>7.4</td>
<td>5.9</td>
<td>-20%</td>
</tr>
<tr>
<td>Dummy nodes</td>
<td>41</td>
<td>20</td>
<td>-51%</td>
</tr>
</tbody>
</table>

(a) PTV. Objective: dose between 2500 and 3150 cGy.

(b) Whole brain.

Figure 19: DVH-curves for the PTV and the whole brain for Patient II. The continuous line is the reference plan and the dashed line is the new plan.
Figure 20: Dose distribution for Patient II. Centered at the maximal value for the reference plan. PTV outlined in black.

Figure 21: Robot traversal path for Patient II. Dose nodes are colored orange and dummy nodes are colored red.
Patient III

Table 8: Different plans for Patient III. $T_{\text{ref}} = 402$.

<table>
<thead>
<tr>
<th></th>
<th>Reference plan</th>
<th>Time-constrained plan</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose nodes</td>
<td>40</td>
<td>60</td>
<td>+50%</td>
</tr>
<tr>
<td>Segments</td>
<td>70</td>
<td>70</td>
<td>0%</td>
</tr>
<tr>
<td>Objective value</td>
<td>0.0552</td>
<td>0.0599</td>
<td>+9%</td>
</tr>
<tr>
<td>Treatment time (s)</td>
<td>2004</td>
<td>1864</td>
<td>-7%</td>
</tr>
<tr>
<td>Traversal time (s)</td>
<td>402</td>
<td>375</td>
<td>-7%</td>
</tr>
<tr>
<td>Average traversal time (s)</td>
<td>7.3</td>
<td>5.8</td>
<td>-21%</td>
</tr>
<tr>
<td>Dummy nodes</td>
<td>16</td>
<td>6</td>
<td>-63%</td>
</tr>
</tbody>
</table>

(a) PTV. Objective: dose between 4500 and 5600 cGy.

(b) Small intestine (green) and kidney (brown).

Figure 22: DVH-curves for the PTV and some OARs. The continuous line is the reference plan and the dashed line is the new plan.
Figure 23: Dose distribution for Patient III. Centered at the maximal value for the reference plan. PTV outlined in black, small intestine in green and kidney in brown.

Figure 24: Robot traversal path for Patient III. Dose nodes are colored orange and dummy nodes are colored red.
6 Discussion

6.1 Algorithms

6.1.1 Filtration algorithm

Based on the three patient cases, the introduced algorithm seems to succeed in reducing the treatment time for each respective plan. An even more drastic reduction is observed in the quantity that is directly optimized upon, the traversal time $T$, being reduced by 47%, 48% and 30% respectively.

This time reduction seems to mainly be a results of a reduced amount of dummy nodes. As the amount of dosage nodes are kept constant, this reduction in dummy nodes leads to a decrease in the total amount of nodes to be visited. Plotting the traversal paths further illustrates this effect, showing simpler traversal paths. The reduction in the average traversal time is also substantial in all three cases, which is another factor in the time reduction.

The change in plan quality depends on the patient case. Since the amount of nodes is kept constant, the time-optimized plan expectantly leads to increases in the objective value. For Patient I, the dose in the PTV is slightly greater than the reference plan, albeit still far from the maximum of 5900 cGy. The dose in the PTV for the other two patients is very similar to the reference plan, with the distribution of dose in the OARs for Patient III differing somewhat. All in all, the plan quality is kept fairly similar.

Although the results only show the latter approach, both the neighbor approach and the greedy TSP approach perform well in testing. The former would instead be preferable if the size of the problem grew considerably, for example if the size of the node set or the amount of filtrations increased substantially. How the different methods would generalize to different, more (or less) complex node sets remains to be seen. However, the fact that the TSP approach fairly directly measures the quantity to be reduced makes a strong case for the algorithm.
6.1.2 Post-filtration algorithm

The post-filtration algorithm gives only a negligible effect on the objective value and the treatment time. While the gains seen in these cases were small, they indicate that the solution obtained in the FMO is close to being locally optimal with respect to objective value and traversal time. This shows the necessity of using appropriate optimization algorithms during the node filtration.

6.1.3 Time constraint

The time constraint was fairly successful in keeping the traversal time, as well as the treatment time, similar between the plans. A small difference is observed in traversal times between plans, which is the result of different TSP solvers being used during and post optimization. This is however not a big problem, since the time constraint should be viewed more as a guidepost than a non-negotiable constraint.

Using the filtration algorithm during the FMO allows for a large increase in the amount of nodes included in the plan, increasing with 88%, 80% and 50% respectively. This leads to a significant drop in objective value in two out of the three cases. For Patient I, more dose is delivered to the PTV leading to a decrease in the minimum dose objective. The objective value for Patient II decreases drastically, mostly as a result of less tissue receiving more than the prescribed maximum of 3150 cGy. Patient III saw a slight increase in objective value as a result of exceeding the threshold of 5600 cGy in the PTV.

To conclude the results, the filtration algorithm shows much promise. It drastically reduces the traversal (and treatment) time in all cases, and was also used to improve the plan quality in two out of three cases. Incorporating the time aspect into the optimization thus seems appropriate even if only plan quality is important. Interesting to note is that the liver case (Patient III) responds the worst in both analyses. The likely explanation is that since the node set for body treatments is significantly smaller (and distributed over a smaller spatial region) than that for head treatments, there is less to gain from optimizing on traversal paths.
6.2 Implications

The implications of higher quality plans are clear, but the effects of reduced treatment times are also many. Firstly, it benefits healthcare providers, allowing them to increase their throughput and treat more patients during the same time span. Such efficiency improvements are always positive considering the insatiable, and growing, demand for healthcare.

Time reductions also have positive effects on patients. One such is the reduced risk for patient movement during the treatment, which might compromise the plan quality. Another is reduced patient discomfort. On a broader level, increased healthcare capacity further benefits patients. Shorter waiting lists lead to reducing waiting times, during which the patient anatomy may change and reduce the quality of the individual treatment plan designed on now outdated patient information. Greater healthcare capacity leads to improved public health.

Lastly, more efficient CyberKnife treatments might increase the interest in the technology. Making the technology more affordable by shortening treatment times might incentivize more healthcare providers to invest in this technology instead of less sophisticated treatment techniques.

6.3 Further studies

Although this thesis details that optimizing on the robot traversal path for the CyberKnife system shows promise in some patient cases, an evaluation of more patient cases is required to validate the conclusions reached in this thesis.

Additional studies into the other components in the treatment time would also be interesting. This thesis only considered the robot traversal path, but investigating ways to for example reduce the time for patient imaging or for the MLC leaf movements would further enhance the results in this thesis. If also a further framework to incorporate these quantities into the time constraint were developed, it would entail the much sought after possibility to directly specify the treatment time in the treatment planning system.
References


