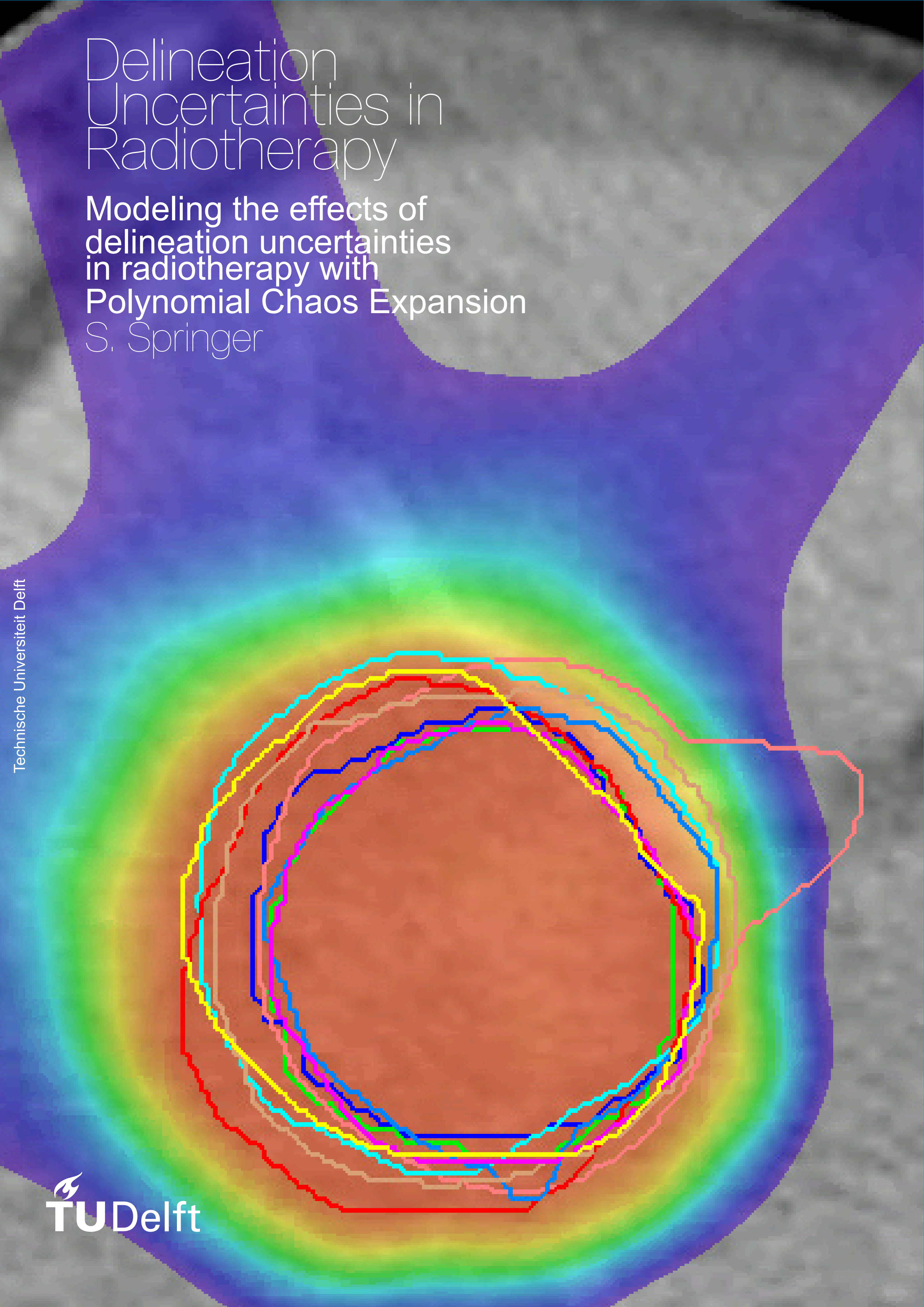


Delineation Uncertainties in Radiotherapy

Modeling the effects of delineation uncertainties in radiotherapy with Polynomial Chaos Expansion

S. Springer

Technische Universiteit Delft

Delineation Uncertainties in Radiotherapy

Modeling the effects of delineation uncertainties in radiotherapy with Polynomial Chaos Expansion

by

S. Springer

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday March 30, 2021 at 2:00 PM.

Student number:	4398920	
Project duration:	August 31, 2020 – March 30, 2021	
Thesis committee:	Dr. Z. Perkó,	TU Delft, supervisor
	Dr. S. J. M. Habraken,	Erasmus MC, supervisor
	Dr. ir. M. C. Goorden,	TU Delft
	Dr. ir. D. Lathouwers,	TU Delft

Acknowledgements

I would like to say a big thank you my supervisors, Zoltán and Steven, for all their support and guidance during my project. Also a special thanks to András for answering all my technical questions related to Matterhorn.

S. Springer
Delft, March 2021

Abstract

The delineation of the regions of interest (ROIs) plays an important role in the radiotherapy workflow. The ROIs include the gross tumor volume (GTV), the clinical target volume (CTV), the planning target volume (PTV) and organs at risk (OARs). There are however uncertainties related to the delineation of the ROIs due to inter-observer variability between radiation oncologists caused by e.g., a lack of consensus on anatomic definition or a lack of contrast in medical images. The largest inter-observer variability has been found in esophageal cancers, head and neck cancers, and lung cancers. To prevent inter-observer variability, auto-contouring software can be used to delineate the ROIs. There are however also uncertainties in the delineations made by auto-contouring software.

The purpose of this study was to perform an accurate evaluation of the dosimetric effect of delineation uncertainties. To do so, first a method to characterize the delineation uncertainties had to be found. The delineations uncertainties were characterized with principal component analysis (PCA). By performing PCA on a set of delineations, the eigenmodes which represent the variations between the delineations can be found and new random delineations based on the eigenmodes can be formed. To determine the dosimetric impact of the delineation uncertainties, two cases were investigated: (1) the dosimetric impact of delineation uncertainties for a fixed dose distribution; and (2) the dosimetric impact of delineation uncertainties for a dose distribution reoptimized for every possible realization of a delineation. For the fixed dose distribution, Polynomial Chaos Expansion (PCE) was used as a meta-model for the dose volume histogram (DVH) of the target, and for the reoptimized dose distribution PCE served as a meta-model for the total dose distribution and the DVHs of the target and other ROIs.

The delineation uncertainties of two data sets were analyzed: (1) 12 manual delineations of the GTV of a hepatocellular carcinoma patient; and (2) 90 auto-contours of the CTV and brainstem of a head and neck patient. The variation of the manual delineations of the GTV could be accurately described by 5 eigenmodes, while 45 eigenmodes were needed to describe the variations of the brainstem auto-contours. The delineation uncertainties in the auto-contours of the CTV could not be characterized by PCA due to the shape of the CTV.

The dosimetric effect of the delineation uncertainties of the manual delineations for a fixed dose distribution was investigated for both an intensity modulated proton therapy (IMPT) plan and a volumetric-modulated arc therapy (VMAT) plan for 10,000 random delineations of the CTV. For this patient the CTV was equal to the GTV. In the VMAT plan, the CTV received sufficient dose for all delineations, but the PTV was underdosed in 17.1% of the delineations. In the IMPT plan, the CTV was underdosed in 69.0% of the delineations. This percentage of underdosed CTV delineations in the IMPT plan was reduced when the plan was made robustly. The results for the fixed IMPT dose distribution should however be verified with a more accurate PCE model.

A proof of principle for the PCE as a meta-model for the reoptimized dose distribution and DVHs of the target and other ROIs was shown. However, a more accurate PCE model with a higher polynomial order would be needed to analyze the effects of the reoptimization on the dose delivered to the ROIs.

It has been shown that the dosimetric impact of delineation uncertainties can be modelled using PCE. This is a first step towards systematically and quantitatively taking into account delineation uncertainties in radiotherapy treatment planning. In future research the analysis of the dosimetric impact of the uncertainties can e.g., be used in adaptive radiotherapy in which auto-contouring is used. By knowing the dosimetric impact of the uncertainty, treatment plans could be optimized such that a structure receives its target dose with a certain probability or critical areas where the dosimetric impact of the delineation uncertainties is large could be flagged such that these areas are checked before treatment.

Contents

List of Figures

List of Tables

1

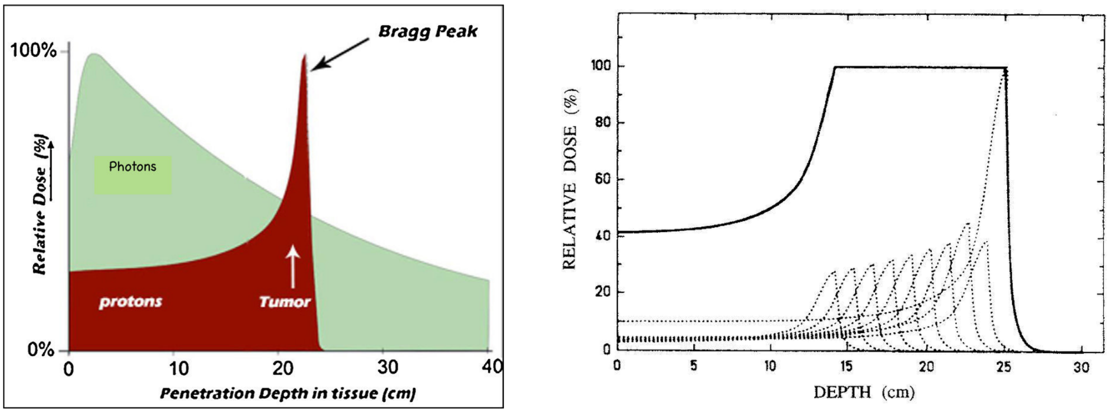
Introduction

In 2020, it is estimated that 19.3 million people worldwide were diagnosed with cancer and there were nearly 10 million cancer deaths. The number of cancer cases is expected to rise to 28.4 million by 2040, which is a 47% increase [39].
The ways to treat cancer include chemotherapy, surgery and radiotherapy. Radiotherapy uses ionizing radiation to kill the malignant cells in the body. The ionizing radiation can either come from a source located inside the body, which is used in brachytherapy, or the ionizing radiation is delivered from outside the body, which is done in external beam radiotherapy. The radiation delivered to the body damages the DNA in the cells in the tumor, while sparing the surrounding tissue as much as possible. The goal of radiotherapy is to cure the patient while minimizing the complications related to treatment. The amount of radiation delivered to a volume is expressed in the unit Gray (Gy). There are two main types of external beam radiation therapy, namely photon therapy and proton therapy.

1.1. Proton and Photon Therapy

Photon therapy utilizes high energy photons which are delivered by a linear accelerator. The photons interact with the molecules in the body and deposit their energy. A photon beams first has a build-up region in the body, after which the energy deposited by the photons decreases exponentially with increasing depth. To achieve a high dose to the tumor and limit the dose to surrounding tissue, it is therefore necessary to use multiple photon beams from different directions.
One way of improving the dose distribution in the body within photon therapy is with intensity-modulated radiation therapy (IMRT). In IMRT, the photon beams are shaped more conformally to the 3D shape of the tumor using multi-leaf collimators and by varying the intensity of the photon beams. This way the dose can be more focused on the tumor and the dose to surrounding tissue is thus reduced. One type of IMRT is volumetric-modulated arc therapy (VMAT). VMAT uses multi-leaf collimators and photon beam intensity modulation to shape the photon beam to the tumor while the gantry is continuously turning around the patient. The advantage of VMAT compared to conventional IMRT is that the patient can be treated with a wider range of beam angles in a shorter time [2].

Another form of external beam radiation therapy is proton therapy. In proton therapy, high energy protons deposit energy in the body. In contrast to photons, protons deposit most of their energy in a small region called the Bragg peak. The depth dose profiles of a photon beam and a proton beam travelling through tissue are displayed in Figure 1.1a.
With the Bragg peak, proton therapy has the big advantage that the high dose region can be centered around the tumor while the dose to surrounding tissue is low. This characteristic of proton therapy makes it ideal for treating tumors in regions near organs at risk (OARs). The location of the Bragg peak in the body can be adjusted by adjusting the proton beam energy. The depth range of the Bragg peak can be widened by using multiple proton beams with different energies. By combining these proton beams, a wider high dose region is created which is called the spread out Bragg peak (SOBP). An illustration of the SOBP is displayed in Figure 1.1b. One type of proton therapy is intensity modulated proton therapy (IMPT). IMPT uses proton pencil beams to deliver dose to spots in the tumor.



(a) The depth-dose profile of photons and protons travelling through tissue. For photons, there is a short build-up region where the deposited dose increases as a function of the depth in the tissue, after which the deposited dose falls exponentially as the photons travel farther through the tissue. For protons, the dose is mainly deposited in the Bragg peak. Figure reproduced from [12].

(b) An illustration of the depth-dose profile of the spread out Bragg peak (SOBP) depicted by the solid line. The dashed lines represent the depth-dose profile of the individual proton beams, which combined together result in the SOBP. As can be seen, the SOBP resulting from combining the proton beams has a larger region where a high dose is delivered than the Bragg peak of the individual proton beams. Figure reproduced from [8].

Figure 1.1: An illustration of (a) the depth-dose profile of photons and protons in tissue and (b) the depth dose profile of the spread out Bragg peak

These spots are located at different depths and lateral locations in the tumor. The proton pencil beam originates from the accelerator and the energy, magnetic deflection and weights of the beam are continuously altered to deliver the desired dose to the spots [21]. In Figure 1.2 an example of the difference between the dose distribution of an IMPT and an IMRT treatment plan is shown. As can be seen, the high dose region is located around the target in both the IMPT plan and the IMRT plan. However, in the IMPT plan the low dose and middle high dose regions are also conformal to the target, while in the VMAT plan the low dose and middle high dose regions are spread out over a larger region in the head.

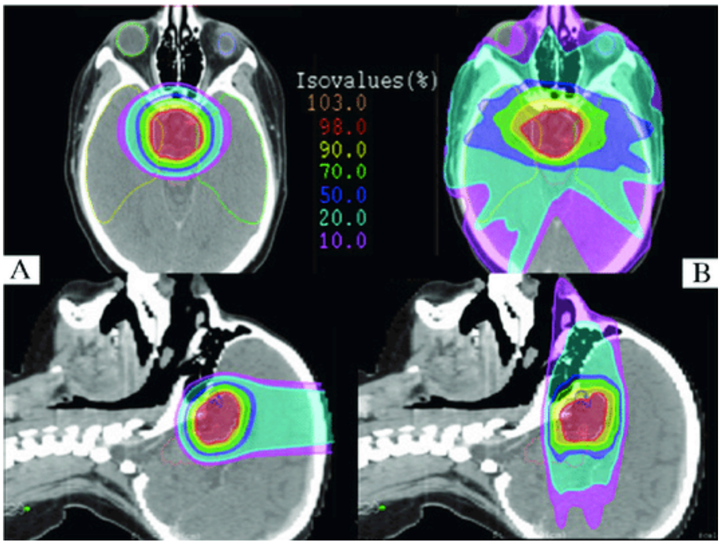


Figure 1.2: In (A), the dose distribution of an IMPT plan to treat craniopharyngioma is displayed, and in (B) an IMRT plan for the same treatment can be seen. The IMPT plan delivers less dose to the structures surrounding the tumor than the IMRT plan. Furthermore, the integral dose to the brain tissue is decreased with IMPT. Figure reproduced from [31].

1.2. Radiotherapy Workflow

In Figure 1.3 the typical workflow of a radiotherapy is demonstrated. The workflow consists of a pre-radiotherapy workup where a tumor is diagnosed in the patient. Next the radiotherapy treatment is prepared by obtaining a CT-scan or an MRI-scan of the patient and delineating the regions of interest (ROIs) in the patient. The ROIs include the gross tumor volume (GTV), the clinical target volume (CTV), the planning target volume (PTV) and the OARs. The GTV is the part of the tumor which is visible on the planning image of the patient. Because there are also microscopic tumor extensions present which cannot be seen on the planning image, a CTV is defined. The CTV is the GTV plus an additional margin to account for the microscopic tumor extensions. To account for uncertainties in the setup of the patient and other possible motions of the CTV, an additional margin is added to the CTV. The resulting volume is the PTV. The PTV is however not suitable for IMPT treatments and also cannot be used for range uncertainties in proton therapy. Therefore robust treatment planning is used in IMPT to ensure that the prescribed dose is delivered to the CTV [43]. Finally, the OARs are organs near the CTV which could be damaged by irradiation.

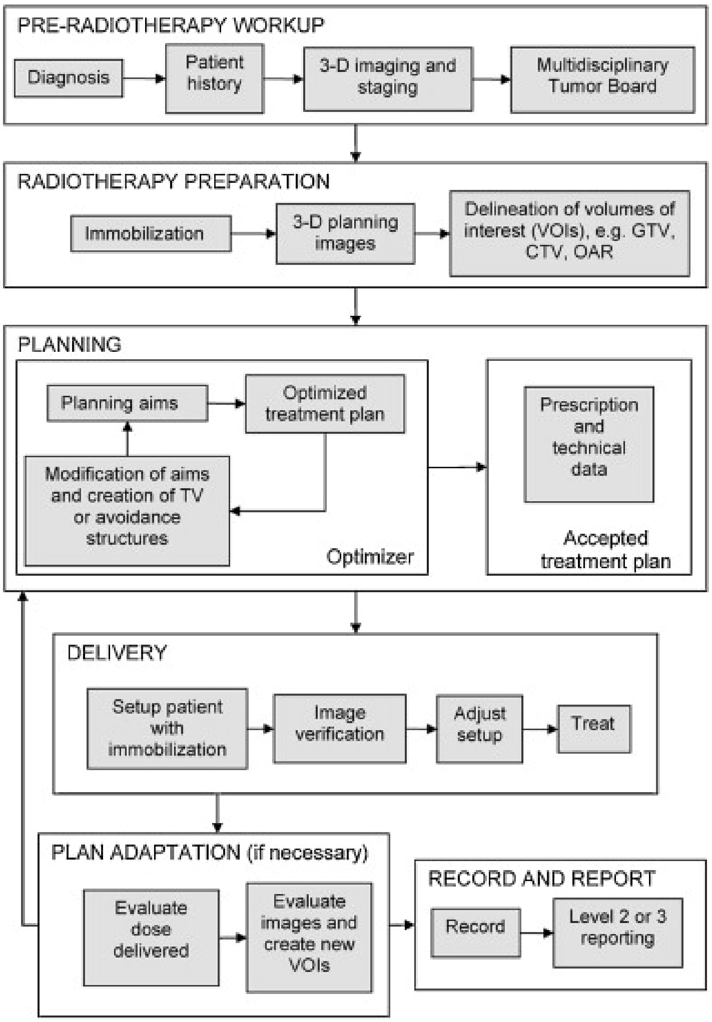


Figure 1.3: The workflow of a typical radiotherapy treatment. Figure reproduced from [14].

Next, a treatment plan is made which meets the planning aims of the treatment. To do so, the prescribed dose to the ROIs is specified and an iterative process takes place to obtain the radiation beams which achieve the prescribed dose to the ROIs. Subsequently, the radiotherapy treatment takes place where the patient is treated with the optimized treatment plan.

1.3. Delineation Uncertainties in Radiotherapy

In the radiotherapy workflow, the delineation of the ROIs is one of the most important steps, as it determines which areas in the body will receive a high dose (GTV, CTV and PTV) and which regions in the body should receive a dose which is as small as possible (OARs). If the delineation of the target volume is not satisfactory, this can lead to an inadequate coverage of the target volume and consequently lead to a lower dose delivered to the tumor than planned, which can lead to recurrences. Inconsistencies in the target volume delineation can thus possibly have large implications. The variations in target delineation can be due to a number of factors. One of the factors is inter-observer variability, which means that the delineations by different observers will be different. The inter-observer variability can be caused by e.g., a lack of consensus in the anatomic definition of the ROI or a lack of contrast in a medical image. The inter-observer variability has been found to depend on the tumor site. The largest inter-observer variabilities have been reported in esophageal cancers, head and neck cancers, and lung cancers [35]. For example, in one study the ratio between the largest and smallest GTV volume delineated of an esophageal tumor delineated by 48 radiation oncologists was found to be 6 [40]. In Figure 1.4 the delineations by 11 radiation oncologists of the GTV on a CT of a patient with lung cancer is shown. As can be seen, the differences between the delineations are very significant.
It has also been found that there is uncertainty in the delineation of OARs. *Schikk et al.* found that for delineations of the bladder and prostate by 6 radiation oncologists the ratio of the largest delineated volume to the smallest delineated volume was approximately 4 [32].

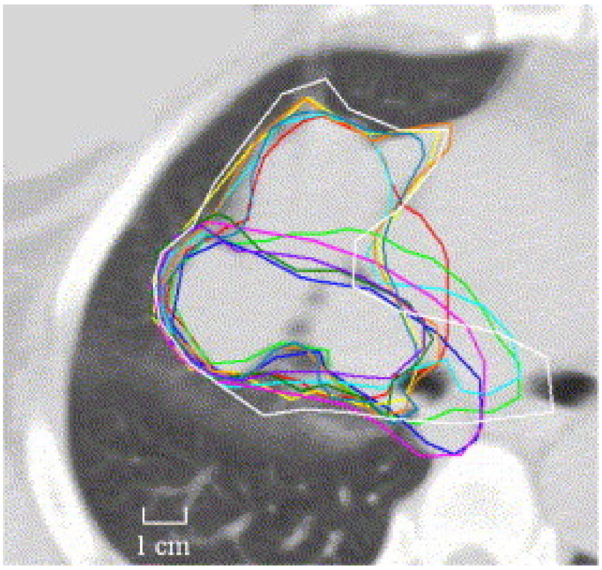


Figure 1.4: Delineations of 11 radiation oncologists of the GTV on a CT of a patient with lung cancer. Figure reproduced from [38].

Another factor that causes variation in the target delineation can be intra-observer variability, thus that the same observer will not delineate the exact same volume when presented with the same image at different points in time. This has however been found to be a smaller source of uncertainty than inter-observer variability [22].
To decrease human errors, auto-contouring software can be used to automatically delineate the ROIs. There are several auto-contouring techniques which are used in auto-contouring software, including atlas-based and deep learning techniques. However, these auto-contouring techniques have also been found to be have inherent delineation uncertainties [18].

With the advent of increasingly precise radiotherapy treatments, like IMPT and IMRT, which have a high dose fall-off at the edge of the delineations, uncertainties in the ROI delineation can have a large impact on the dose delivered to the ROI. For instance, if a CTV delineated by a radiation oncologist is smaller than the true CTV and the radiation oncologist's CTV is used for the treatment planning, the dose to the true CTV will probably be too low. It is therefore essential that the delineation uncertainties are well characterized and taken into account in radiotherapy treatment planning. Furthermore, the development of more automated radiotherapy techniques like adaptive radiotherapy rely on auto-contouring software to delineate the ROIs. To ensure a successful treatment, it is essential to know the uncertainties related to the auto-contouring software.

There have been several studies which have quantified the effect that delineation uncertainties have on the dose delivered to the ROIs. For instance, in a study by *Rasch et al.* two radiation oncologists delineated the CTVs on 9 CTs of patients with paranasal tumors and the PTVs were made by expanding the CTVs by 5 mm. If a treatment plan was made on the PTV of one of the radiation oncologists, the same treatment plan would always lead to an underdosage in the PTV delineated by the other radiation oncologist [30]. In another study, *Aliotta et al.* generated many delineations by randomly perturbing manual delineations of several OARs and determined whether the maximum allowed dose in the OARs was exceeded in any of the randomly perturbed delineations. For the spinal cord and esophagus it was found that the maximum dose could exceed the maximum allowed dose in several of the perturbed delineations [1].

The studies done so far on evaluating the dosimetric impact of delineation uncertainties either only evaluate delineations on the same image by a small number of observers, which does not give an accurate probability of whether a target dose is met, or the perturbations of one reference delineation are made randomly, which possibly does not lead to realistic delineations. To accurately determine the probability of reaching a target dose, a large number of realistic delineations would be needed.

1.4. Research Purpose

The purpose of this study is to perform an accurate evaluation of the dosimetric effect of delineation uncertainties. There are several goals related to this purpose. Firstly, the delineation uncertainty should be characterized and a method to generate random delineations based on this characterization should be made. Next, using the characterization of the uncertainties, it is investigated whether Polynomial Chaos Expansion (PCE) can be used to model dosimetric quantities as a function of the characterized uncertainties. It is researched whether PCE can serve as a meta-model for dosimetric quantities in the case that the dose distribution stays fixed and in the case that the dose distribution is reoptimized for each realization of a new delineation.

This project is performed at both the Medical Physics & Technology research group, which is part of the Radiation Science & Technology department of the Applied Physics faculty of Delft University of Technology, and the Radiotherapy Medical Physics and Technology research group of the Erasmus Medical Center in Rotterdam.

1.5. Structure

The rest of the thesis follows the following structure. In Chapter 2 the theory relevant to this research will be discussed, which starts with theory on the characterization of delineation uncertainties. Then the idea behind principal component analysis and its applications to delineations will be set out, after which auto-contouring and treatment planning are discussed. Finally the theory of Polynomial Chaos Expansion and its application to serve as a model for dosimetric quantities. In Chapter 3, the methods and materials used in this thesis are set out, which starts with the description of the data used in the research, followed by how the uncertainty models for both inter-observer variability and auto-contouring are determined. Finally, it is shown how the polynomial chaos model in constructed and validated as a meta-model for dosimetric quantities for both a fixed dose distribution and a reoptimized dose distribution. In Chapter 4 the results are presented, which include the uncertainty quantification and characterization of both manual delineations and auto-contours. Then the PCE validation results are shown, followed by the dosimetric effects that delineation uncertainties have. In Chapter 5 the results are discussed, after which the conclusions of the research are given in Chapter 6.

2

Theory

In this chapter the theory used in this research is detailed. First the theory of delineation uncertainties is discussed in Section 2.1, after which the idea behind principal component analysis is explained along with its application to delineations in Section 2.2. Then the applications of auto-contouring and one auto-contouring method is presented in Section 2.3, followed by a description of the steps in treatment planning and the treatment planning system used in this thesis (iCycle) in Section 2.4. Finally, the theory of Polynomial Chaos Expansion and its application to delineation uncertainties are presented in Section 2.5.

2.1. Delineation Uncertainties

In this section the theory of delineation uncertainties is explained. First the different ways of characterizing delineation uncertainties is set out along with the methods to determine the gold-standard delineation. Finally, conventional ways of compensating for delineation uncertainties are discussed.

2.1.1. Characterization of Delineation Uncertainties

It is difficult to compare different studies regarding delineation uncertainties as a variety of parameters is used to compare delineations. These parameters can either describe the distribution of the delineations or they describe the concordance, i.e. the agreement, between the delineations.
Parameters which can be used to describe the distribution of the delineations are e.g.,:

- Range (Lists the minimum delineated volume and the maximum delineated volume),
- Mean (gold-standard) delineation and (local) standard deviation,
- Ratio between largest and smallest delineated volume.

Parameters to describe the concordance between delineations are e.g.,:

- Dice similarity coefficient $\left(\frac{2(A\cap B)}{A+B}\right)$,
- Jaccard conformity index $\left(\frac{A\cap B}{A\cup B}\right)$,
- Geographical miss index $\left(\frac{B-(A\cap B)}{B}\right)$.

$A \cap B$ denotes the intersection between A and B and $A \cup B$ denotes the union between A and B. In the concordance measures, A is the delineation made by the radiation oncologist and B is the gold-standard delineation. The gold-standard is discussed in Section 2.1.2.

The ratio between largest and smallest delineated volume is the easiest comparison measure to apply as there is no need for a reference standard.
The dice similarity coefficient (DSC) is the most widely used standard to compare delineations. The

DSC, similar to other concordance parameters, has the disadvantage that it is not sensitive to differences at the edge of the delineations if these differences have a small impact on the volume of the delineation. Measures like the standard deviation on the other hand are not sensitive to the volume of the delineation.
The DSC and standard deviation both require a reference delineation from which they are measured. This reference delineation is also known as the gold-standard delineation.

2.1.2. Gold-Standard Delineation

When looking at different delineations on the same CT, the problem of the lack of a ground truth arises. Lacking the ground truth, there are several methods to find a gold-standard delineation which replaces the ground truth. The gold-standard delineation can be determined by:

- The delineation based on a consensus among multiple radiation oncologists.
- A delineation determined by the simultaneous truth and performance level estimation (STAPLE) algorithm. The STAPLE algorithm determines a probabilistic estimate of the gold-standard delineation based on the optimal combination of all delineations [45].
- A probability map which indicates the probability that a certain voxel belongs to a delineated contour. The gold-standard is then chosen as the contour which contains all voxels which were inside the delineations of at least a given percentage of the delineations [30]. An example of the use of a probability map to determine the gold-standard is shown in Figure 2.1.

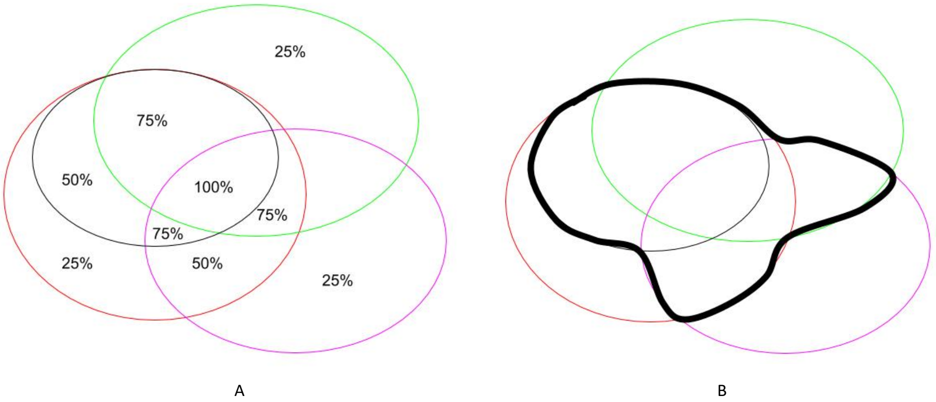


Figure 2.1: An example of the use of a probability map to form a gold-standard delineation. In (A), four delineations are displayed along with the probabilities that a certain area is encompassed by a delineation. (B) displays the gold-standard delineation which is obtained by setting a threshold of 50% on the probability map.

The use of a probability map to determine the gold-standard has some drawbacks. A threshold of 50%, which is a commonly used threshold, is possibly not reliable if there is a small number of delineations. The consensus between the delineations can also vary depending on the delineated organ. This suggests that the consensus threshold should be dependent on the organ type and the number of delineations [11].

2.1.3. Conventional Methods of Compensating for Delineation Uncertainties

Delineation uncertainties are most often compensated for by adding a margin around the CTV according to a margin recipe. In the margin recipe developed by *van Herk et al.* [44] the delineation uncertainty is considered as a systematic error with a standard deviation of $\vec{\Sigma}_d$. $\vec{\Sigma}_d$ is a vector because the size of the delineation uncertainty can be different in the x, y and z direction. $\vec{\Sigma}_d$ is then combined with the standard deviation $\vec{\Sigma}_s$ of the setup error and the standard deviation $\vec{\Sigma}_m$ of the organ motion to form the combined standard deviation $\vec{\Sigma}$. The x-component of $\vec{\Sigma}$ is given by:

$$\Sigma_x = \sqrt{\Sigma_{m,x}^2 + \Sigma_{s,x}^2 + \Sigma_{d,x}^2}, \tag{2.1}$$

where $\Sigma_{m,x}$, $\Sigma_{s,x}$ and $\Sigma_{d,x}$ are the x-components of $\vec{\Sigma}_m$, $\vec{\Sigma}_s$ and $\vec{\Sigma}_d$, respectively. The y- and z-components are of a similar form. To ensure that 90% of the patient population receives at least a minimum dose of 95% of the nominal dose to the CTV, the CTV is expanded by using a rolling ball algorithm, where the shape of the 'ball' is an ellipsoid of which the lengths of the principal semi-axes in the x, y and z direction are the x, y and z components of $2.5\vec{\Sigma}$, respectively. [44].

The above method concerns the expansion of the CTV to the PTV. It thus does not concern delineation uncertainties of e.g., OARs. The uncertainties in the delineation of the OAR however also have an important role in the treatment planning, as the goal is to minimize the dose delivered to the OARs. Furthermore, the above method does not work for IMPT where no margins are used around the CTV.

2.2. Principal Component Analysis (PCA)

In this section the principle of Principal Component Analysis (PCA) will be explained, along with its application to delineations.

2.2.1. Idea Behind PCA

Principal Component Analysis, from now on 'PCA', is a tool used in data analysis that computes the so called prinicipal components of a data matrix $\mathbf{X}$. These principal components can be used to reduce the dimensionality of $\mathbf{X}$, while preserving as much of the data's variability as possible. The representation of the data in this lower dimension can show the structures which explain the data-set.

For PCA, the data matrix $\mathbf{X}$ should be arranged in such a way that all N rows contain the measurements and the M columns contain the variables, i.e., $\mathbf{X}$ is an $N \times M$ matrix.

The N principal components of the data are vectors in N-dimensional space along which the variance of the projection of the data points on that principal component is maximized. Each principal component is orthogonal to the other principal components and the first principal component has the largest projected variance of the data-points, the second principal component the second largest projected variance, and so forth. In Figure 2.2 a 2D example is given which visualizes the two principal components which are centered around the mean of the data points. It can be seen that the projected variance of the data is largest on the first principal component and the second principal component is orthogonal to the first principal component.

In Figure 2.3 an example of the use of PCA for dimensionality reduction for a 3D data set is demonstrated. In Figure 2.3a the 3D data set is visualized, with the corresponding principal components in Figure 2.3b. By projecting the data points onto the first two principal components, the dimension of the data set can be reduced from three to two dimensions, as can be seen in Figure 2.3c.

2.2.2. PCA Applied to Delineations

To apply PCA to delineations, an adaptation of the point-distribution model [7] is used. In a point distribution model, an object is represented by points on its surface. PCA is used to analyze the statistics of the movement of these points across the different delineations, resulting in the average position of the points on the surface and the main modes of variation of the points. These modes of variation are called eigenmodes, and are the same as the principal components of the data. Therefore, the principal components will from now on be referred to as eigenmodes in the rest of the thesis. In this section the application of the point-distribution model to delineations is shown.

Parameterization of delineation geometries

Given N delineations on the same CT scan, the geometry of the i^{th} delineation ($i = 1, ..., N$) is parameterized by the surface shape vector $\vec{p}_i$, which is given by:

$$\vec{p}_i = (x_1(i), ..., x_M(i)), \tag{2.2}$$

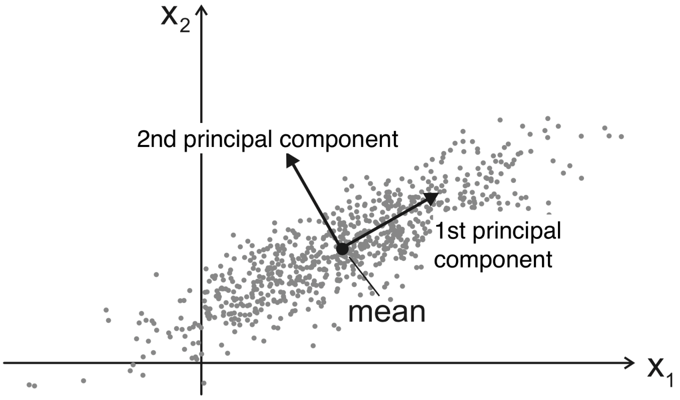


Figure 2.2: Example of principal component analysis applied to a 2D data set. It can be seen that the projected variance of the data is largest on the first principal component and the second principal component is orthogonal to the first principal component. Figure adapted from [37].

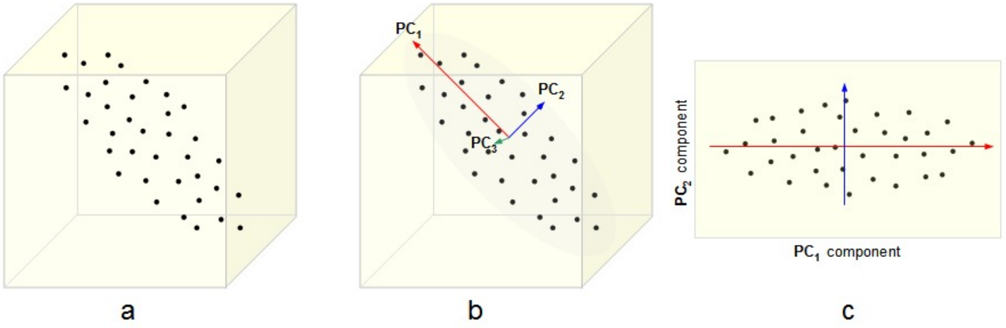


Figure 2.3: Example of principal component analysis to reduce the dimensionality of a 3D data set. In (a) the data set is visualized and the corresponding principal components of the data are visualized in (b). (c) shows the projection of the data set onto the first two principal components. Figure reproduced from [29].

where x_j $(j = 1, ..., M)$ denotes the distance from the center of mass of the gold-standard delineation to the j^{th} surface point. The surface points for each delineation are obtained by determining the intersection point between the surface of the delineation and a line which originates in the center of mass of the gold-standard delineation and travels in the direction of polar angle θ $(\theta = 0°, 0.1°, ..., 179.9°, 180°)$ and azimuthal angle ψ $(\psi = 0°, 1°, ..., 359°, 360°)$. When there is a 0.1 degree increment in the polar angle and a 1 degree increment in azimuthal angle between measurements of the surface point distance, the length M of each surface shape vector $\vec{p}_i$ is 1801×361 = 650161 elements.

There exists the possibility that a line which originates in the center of mass of the gold-standard contour intersects the delineation surface multiple times. This can for instance happen when the delineation forms a concave 3D surface. In this case, the first intersection point of the line with the 3D surface is taken as the input for the surface shape vector $\vec{p}_i$. This will result in points on the 3D surface not being characterized by $\vec{p}_i$. It is therefore important to validate that the surface shape vectors $\vec{p}_i$ characterize the 3D delineation surface well enough. An example of a 2D concave surface with multiple intersections is shown in Figure 2.4.

Decomposition of delineation geometries into eigenmodes

To perform PCA on the parameterized delineation geometries, it is assumed that the surface shape vectors $\vec{p}_i$ are correlated and the data is thus suitable for dimensionality reduction. This is a safe assumption to make, as delineations are located around the same position and will mainly only vary in shape. This variability in shape can be described by much less than M dimensions of the surface shape vector $\vec{p}_i$. Using PCA, the correlated deformations of the delineations are found and the corresponding eigenmodes which describe the deformations are found. To perform PCA, the data stored in the sur-

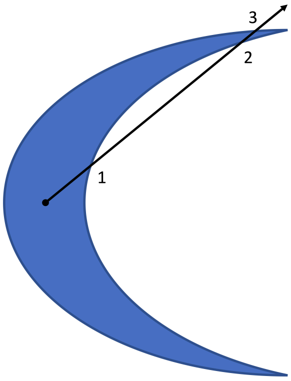


Figure 2.4: Example of multiple intersections on a 2D concave surface. The line originating from the center of mass crosses the border of the surface 3 times. In this case only the first intersection point would be used as input for the surface shape vector.

face shape vectors $\vec{p}_i$ needs to be altered.

From the surface shape vectors, a mean surface shape vector $\vec{\mu}$ is made. The mean shape vector is given by:

$$\vec{\mu} = \frac{1}{N}\sum_{i=1}^{N}\vec{p}_i \tag{2.3}$$

The mean surface shape vector $\vec{\mu}$ is subtracted from the surface shape vectors for each delineation and then stacked into one surface shape matrix $\mathbf{P}$ as follows:

$$\mathbf{P} = \begin{pmatrix} \vec{p}_1 - \vec{\mu} \\ \vec{p}_2 - \vec{\mu} \\ \vdots \\ \vec{p}_N - \vec{\mu} \end{pmatrix} \in \mathbb{R}^{N\times M} \tag{2.4}$$

The goal of principal component analysis is to find a collection of orthogonal eigenmodes $\vec{v}$ such that the variance of the dataset projected onto the direction of each eigenmode $\vec{v}_l$ $(l = 1, ..., L_{\text{total}})$ is maximized, where L_{total} is the total number of eigenmodes. These eigenmodes are equal to the eigenvectors of the covariance matrix $\mathbf{C}$, which is defined as:

$$\mathbf{C} = \frac{1}{N-1}\mathbf{P}^T\mathbf{P} \in \mathbb{R}^{M\times M} \tag{2.5}$$

The covariance matrix represents the geometric variability between the delineations, and the eigenmodes describe the deformation characteristics of the ensemble of delineations, under the assumption of a Gaussian distribution. The eigenvalues corresponding to the eigenmodes rank the eigenmodes with respect to how much they represent the variability between the delineations. The eigenmodes with the largest eigenvalues represent the eigenmodes which span the space where the largest variations between the delineations occur.

A new space is spanned by the eigenmodes in which the spread of data along the direction of each eigenmode $\vec{v}_l$ turns into a 1D Gaussian distribution. The variance σ_l^2 of the data projected on eigenmode $\vec{v}_l$ is given by:

$$\sigma_l^2 = \lambda_l, \tag{2.6}$$

where λ_l is the eigenvalue corresponding to eigenmode $\vec{v}_l$

Singular Value Decomposition

The eigenmodes are equal to the eigenvectors of the covariance matrix $\mathbf{C}$, but it is unnecessary to determine all eigenvectors because there are only N non-zero eigenvalues, as the covariance matrix is calculated from N surface shape vectors, and typically $N << M$. This means that the eigenvectors with an eigenvalue of zero are not necessary to calculate. Therefore, it is more computationally efficient to determine the eigenmodes using singular value decomposition (SVD) of the surface shape matrix $\mathbf{P}$. SVD decomposes the $N \times M$ surface shape matrix $\mathbf{P}$ into three matrices:

$$\mathbf{P} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \tag{2.7}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices and $\mathbf{S}$ is a diagonal matrix. The columns of $\mathbf{U}$ are the left singular vectors and the vectors of $\mathbf{V}$ are the right singular vectors. The scalar values on the diagonal of $\mathbf{S}$ are the singular values s_l corresponding to the columns $\vec{u}_l$ of U and columns $\vec{v}_l$ of V. From the SVD of $\mathbf{P}$ it follows that:

$$\mathbf{P}^T\mathbf{P} = (\mathbf{U}\mathbf{S}\mathbf{V}^T)^T(\mathbf{U}\mathbf{S}\mathbf{V}^T) = (\mathbf{V}\mathbf{S}\mathbf{U}^T)(\mathbf{U}\mathbf{S}\mathbf{V}^T) \tag{2.8}$$

$\mathbf{U}$ is an orthonormal matrix, so $\mathbf{U}^T\mathbf{U} = \mathbf{I}$. From this it follows that:

$$\mathbf{P}^T\mathbf{P} = \mathbf{V}\mathbf{S}^2\mathbf{V}^T = (N-1)\mathbf{C} \tag{2.9}$$

The covariance matrix $\mathbf{C}$ can thus be rewritten as

$$\mathbf{C} = \frac{1}{N-1}\mathbf{V}\mathbf{S}^2\mathbf{V}^T, \tag{2.10}$$

which is basically an eigendecomposition of $\mathbf{C}$. The columns $\vec{v}_l$ of $\mathbf{V}$ are the eigenvectors of $\mathbf{C}$ and the diagonal elements s_l^2 of $\mathbf{S}^2$ are the corresponding eigenvalues λ_l. It can be seen that the eigenvalues λ of the covariance matrix are equal to the square of the singular values of $\mathbf{P}$ divided by $(N-1)$, i.e.,:

$$\lambda_l = \frac{s_l^2}{N-1} \tag{2.11}$$

It is thus possible to calculate the eigenvectors of the covariance matrix, i.e., the eigenmodes of deformation of the delineations, and the corresponding eigenvalues by determining the SVD of the surface shape matrix $\mathbf{P}$. Within principal component analysis, the original data is transformed to a new coordinate system with new coordinate axes along the directions of the eigenmodes. The importance of each eigenmode is represented by the magnitude of the corresponding singular value. The larger the singular value corresponding to an eigenmode, the larger the variance of the data when it is projected onto the corresponding new coordinate axis. Usually, the magnitude of the singular values corresponding to the eigenmodes decreases quickly and the data can be explained well by projecting it onto a coordinate system made up of the most dominant eigenmodes which explain most of the variance of the data.

In the case of the delineation geometries, the number of surface points M will be much larger than the number of delineations N. Consequently, the last $M-N$ columns of $\mathbf{V}$ have no significance. The corresponding singular values in $\mathbf{S}$ will also be zero. It is therefore more efficient to calculate the economy-sized SVD of the surface shape matrix $\mathbf{P}$, which computes only the first N columns of $\mathbf{V}$, resulting in an $M \times N$ matrix.

Construction of delineation geometries with eigenmodes

The developed PCA model can be used to recreate the original delineation geometries as described in the surface shape matrix $\mathbf{P}$, but can also be used to create new delineation geometries. Using L eigenmodes, a reconstructed shape matrix $\mathbf{P}^L_{\text{rec}}$ can be created using the following formula:

$$\mathbf{P}^L_{\text{rec}} = \mathbf{U}_L\mathbf{S}_L\mathbf{V}_L^T, \tag{2.12}$$

where $\mathbf{U}_L$ is an $N \times L$ matrix containing the first L columns of $\mathbf{U}$, $\mathbf{S}_L$ is an $L \times L$ matrix containing the upper left part of $\mathbf{S}$ and $\mathbf{V}_L$ is an $M \times L$ containing the first L columns of $\mathbf{V}$. The resulting reconstructed shape matrix $\mathbf{P}^L_{\text{rec}}$ is an $N \times M$ matrix, analogous to the shape matrix $\mathbf{P}$. $\mathbf{P}^L_{\text{rec}}$ now however has a lower rank of L.

The reconstructed surface shape vectors $\vec{p}^L_{i,rec}$ are obtained by adding the mean shape vector $\vec{\mu}$ to the i^{th} row of $\mathbf{P}^L_{\text{rec}}$. The more eigenmodes L that are used to create $\mathbf{P}^L_{\text{rec}}$, the smaller the difference will be between $\mathbf{P}$ and $\mathbf{P}^L_{\text{rec}}$.

To create a random delineation geometry, the random surface shape vector $\vec{p}^L_{\text{random}}$ is calculated by:

$$\vec{p}^L_{\text{random}} = \vec{\mu} + \sum_{l=1}^{L} c_{l,\text{random}} \cdot \vec{v}_l, \tag{2.13}$$

where coefficients $c_{l,\text{random}}$ for eigenmode $\vec{v}_l$ are chosen from a Gaussian distribution with mean $\mu = 0$ and standard deviation $\sigma_l = \sqrt{\lambda_l}$, i.e.,:

$$c_{l,random} \sim \mathcal{N}(0, \lambda_l) \tag{2.14}$$

2.3. Auto-Contouring

In this section the principles of auto-contouring will be set out. First the necessity of auto-contouring in the current advancements of radiotherapy will be explained, followed by the theory behind atlas-based auto-contouring and finally the principles of B-spline deformable image registration will be detailed.

2.3.1. Auto-Contouring in Adaptive Radiotherapy

The manual delineation of a CT by a radiation oncologist is a very time-consuming process which can take several hours per patient. Furthermore, as discussed in the introduction, manual delineations are prone to inter- and intraobserver variabilities which limit accuracy and can have an impact on the dose delivered to the ROIs. Auto-contouring can solve these issues which arise from manual delineations. In auto-contouring the ROIs are automatically identified and delineated by an auto-contouring software in a short time period. This is very important regarding the new advancements in adaptive radiotherapy. In traditional radiotherapy, a CT scan is often acquired one week before the treatment and the treatment plan is based on this single CT scan. The radiation dose is split into fractions, which means that the dose will be delivered to the patient over the course of several days to weeks. However, the anatomy of the patient will likely change before the treatment starts and between the daily fractions. The dose delivered to the ROIs will therefore not be the same as the planned dose. The goal of adaptive radiotherapy is to deliver the dose accurately in the presence of such anatomical changes. Adaptive radiation therapy consists of the following steps:

1. Approximately 1 week before the treatment a planning CT is made which is delineated by a radiation oncologist and a treatment plan is made.

2. On the day of the treatment a daily CT is made which is delineated by an auto-contouring software. The treatment plan is then adapted accordingly to compensate for the observed anatomical variations.

There are two types of adaptive therapy, namely offline- and online adaptive radiation therapy. In offline adaptive radiotherapy, the treatment plan is adapted at certain timepoints during the treatment course (e.g., once a week), while in online adaptive radiotherapy the treatment plan is adapted before every fraction.

Thyrza et al. have found that it is possible to adapt a proton therapy treatment plan in a few minutes using plan-library supported automated replanning [19]. With such fast replanning, it is possible to treat the patient very quickly after the daily CT has been made. To further reduce the time between the daily CT and the treatment, auto-contouring plays an important role.

The field of auto-contouring has grown in a fast pace over the past two decades [6] and there are different types of auto-contouring techniques. In this thesis an auto-contouring software using the atlas-based auto-contouring technique is used.

2.3.2. Atlas-Based Auto-Contouring

There are two types of atlas-based auto-contouring methods, namely single atlas-based auto-contouring and multi-atlas-based auto-contouring. In single atlas-based auto-contouring there is one reference image which has its ROIs already delineated by a radiation oncologist. This delineated reference image is called the atlas and is used as a reference for other images which need to be delineated. To create a delineation on a new image, the auto-contouring software uses deformable image registration (DIR) to find the transformation which optimally maps the reference image to the new image. The reference delineations are then also transformed to the new image using this transformation. Single atlas-based auto-contouring can however be impacted by the anatomical variability that exists between different images of the same patient [13].

Multi-atlas-based auto-contouring uses multiple atlases which represent a variety of anatomical variabilities of the patient. There are then several approaches which can be used to generate the delineation on the new image. First, all atlas images can be transformed using DIR and the atlas delineations are mapped onto the new image. The deformed delineations are then combined to form the delineation produced by the auto-contouring software. This is called multi-atlas fusion. A second option is to select a single atlas image which most closely resembles the new image. Using DIR the atlas image is again transformed to the new image and the atlas delineation is mapped to the new image. This is called single-atlas selection. Finally, it is possible to select a selection of atlases which resemble the new image. The generation of the delineation on the new image is then again done by combining the mapped selected atlas delineations. This is called multi-atlas selection and fusion [33]. A visualization of all these atlas-based auto-contouring methods can be seen in Figure 2.5.

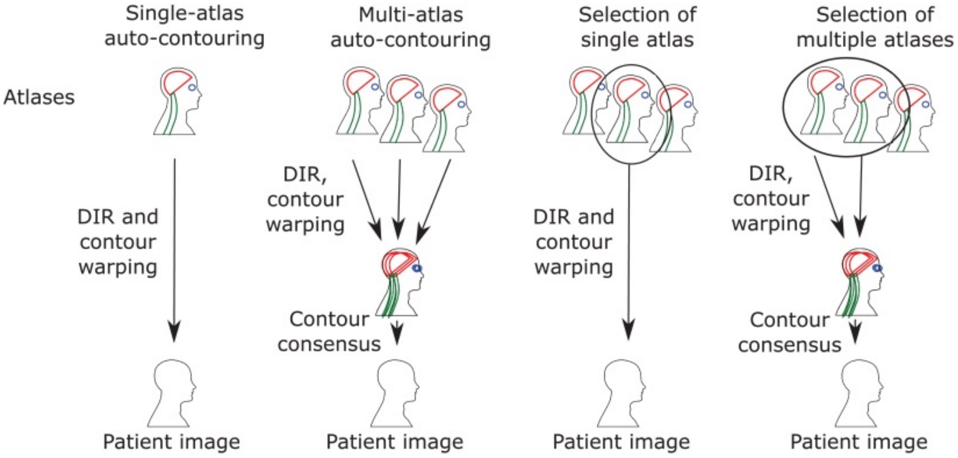


Figure 2.5: Overview of all atlas-based auto-contouring methods. From left to right: single-atlas auto-contouring, multi-atlas fusion auto-contouring, single-atlas selection auto-contouring, multi-atlas selection and fusion auto-contouring. Figure reproduced from [33].

2.3.3. B-spline Deformable Image Registration

DIR plays an important role in auto-contouring. One method of DIR is B-spline DIR. B-spline DIR is a type of nonlinear transformation. In B-spline DIR, every voxel in a fixed image F is mapped to a corresponding voxel in a moving image T. The deformation vector field (DVF) $\vec{\phi}$ is defined on every point in the fixed image and describes how the voxels in the moving image are displaced with respect to their initial position in the fixed image. The DVF is modeled with cubic B-splines.

A sparse set of uniformly distributed control points is superpositioned on the fixed image's voxel grid. These control points divide the voxel grid up into many equally sized tiles. The cubic B-splines are continuous curves which are only defined on a set of control points. The number of control points per dimension that the B-splines are defined on equals $O_{spline} + 1$, where O_{spline} is the order of the B-spline curve. For cubic B-splines, where O_{spline} equals 3, there are thus 4 control points per dimension, resulting in $4^3 = 64$ control points in a 3D image that a cubic B-spline is defined on. This means that the DVF of a voxel in a tile is calculated using the 64 control points in the vicinity of that tile. Each control point can be described by its local coordinates (α, β, γ). Given a voxel at position (x, y, z), the control points (α, β, γ) that are used in the calculation of the DVF in this voxel are given by:

$$\begin{aligned} \alpha &= \left\lfloor \frac{x}{N_x} - 1 + i \right\rfloor \\ \beta &= \left\lfloor \frac{y}{N_y} - 1 + j \right\rfloor \\ \gamma &= \left\lfloor \frac{z}{N_z} - 1 + k \right\rfloor, \end{aligned} \tag{2.15}$$

where N_x, N_y and N_z are the number of voxels in a tile in the x, y and z direction, respectively, and $i, j, k = 0, \ldots, 3$.

A voxel in a tile can also be described in terms of the local coordinates (u, v, w) within a tile. (u, v, w)

are given by:

$$
\begin{aligned}
u &= \frac{x}{N_x} - \left\lfloor \frac{x}{N_x} \right\rfloor \\
v &= \frac{y}{N_y} - \left\lfloor \frac{y}{N_y} \right\rfloor \\
w &= \frac{z}{N_z} - \left\lfloor \frac{z}{N_z} \right\rfloor
\end{aligned}
\tag{2.16}
$$

Using the local coordinates of a voxel in a tile, the B-spline basis functions are defined as:

$$
\beta_i(u) = \begin{cases} \frac{(1-u)^3}{6} & \text{if } i = 0 \\ \frac{3u^3-6u^2+4}{6} & \text{if } i = 1 \\ \frac{-3u^3+3u^2+3u+1}{6} & \text{if } i = 2 \\ \frac{u^3}{6} & \text{if } i = 3, \end{cases}
\tag{2.17}
$$

and similarly for $\beta_j(v)$ and $\beta_k(w)$.
Finally, the components of the DVF $\vec{\phi}$ in point (x, y, z) are given by:

$$
\begin{aligned}
\phi_x(x,y,z) &= \sum_{i=0}^{3}\sum_{j=0}^{3}\sum_{k=0}^{3} \beta_i(u)\beta_j(v)\beta_k(w)P_x(\alpha,\beta,\gamma) \\
\phi_y(x,y,z) &= \sum_{i=0}^{3}\sum_{j=0}^{3}\sum_{k=0}^{3} \beta_i(u)\beta_j(v)\beta_k(w)P_y(\alpha,\beta,\gamma) \\
\phi_z(x,y,z) &= \sum_{i=0}^{3}\sum_{j=0}^{3}\sum_{k=0}^{3} \beta_i(u)\beta_j(v)\beta_k(w)P_z(\alpha,\beta,\gamma),
\end{aligned}
\tag{2.18}
$$

where $P_x(\alpha,\beta,\gamma)$, $P_y(\alpha,\beta,\gamma)$ and $P_z(\alpha,\beta,\gamma)$ are the B-spline coefficients in control point (α,β,γ).
Using the DVF $\vec{\phi}$, each voxel in the moving image is moved, creating a deformed moving image. To compare the deformed moving image with the fixed image, a cost function is used. A cost function commonly used in B-spline DIR is the sum of squared differences (SSD) cost function [36]. The SSD cost function computes the intensity differences between the deformed moving image and the fixed image and is given by:

$$
C = \frac{1}{N_{\text{vox}}} \sum_{x=1}^{N_{\text{vox},x}} \sum_{y=1}^{N_{\text{vox},y}} \sum_{z=1}^{N_{\text{vox},z}} \left(F(x,y,z) - T(x+\phi_x(x,y,z), y+\phi_y(x,y,z), z+\phi_z(x,y,z))\right)^2,
\tag{2.19}
$$

where N_{vox} is the number of voxels in the moving image T after it has been deformed by the DVF $\vec{\phi}$. $N_{\text{vox},x}$, $N_{\text{vox},y}$ and $N_{\text{vox},z}$ are the number of voxels in the deformed moving image in the x, y and z direction, respectively.
To get the best registration between the deformed moving image and the fixed image, the cost function must be minimized. When the cost function is minimized, the deformed moving image and the fixed image are most similar to each other. To search for the minimum value of the cost function, its derivative is taken with respect to the B-spline coefficients P_x, P_y and P_z. The derivative of the cost function with respect to the B-spline coefficients is also called the cost function gradient. Using a gradient descent method, the cost function gradient will be used to iteratively decrease the cost function by altering the B-spline coefficients until a minimum is reached. Once the cost function has been minimized, the obtained optimal B-spline coefficients are used to form the DVF which can be used for DIR between the moving image and the fixed image.

2.4. Treatment Planning

In this section, first the general principles of radiotherapy treatment planning will be explained, after which the working principles of the treatment planning system iCycle will be set out. Finally, the concept of robust treatment planning will be detailed, along with its implementation in iCycle.

2.4.1. Principles of Treatment Planning

The goal of radiotherapy treatment planning is to find the machine parameters that result in a dose distribution in the body which results in the highest quality of life for the given patient. The main machine parameters that need to be determined are the number of beams, the beam intensities and beam angles. These parameters are chosen such that the dose delivered to the tumor is sufficient, while the dose to surrounding healthy tissue is minimized. Treatment plans are made with a treatment planning system. The CT with the delineated structures is used as input for the treatment planning system, along with the prescribed dose $D^{\text{prescribed}}$ for the tumor and the dose constraints for the surrounding tissue. Based on this input, the treatment planning system determines the required machine parameters to satisfy these treatment aims and also outputs the corresponding dose distribution. The dose distribution is checked by the radiation oncologist and if the dose distribution does not satisfy the treatment aims, the treatment aims are altered, and a new treatment plan is made. This process is repeated until a satisfactory dose distribution is achieved.

In treatment planning, the number of beams and the angles of these beams are often chosen in a trial-and-error procedure or based on a template. Depending on the experience of the planner, choosing beams with a trial-and-error procedure can be a long process. Using a template is less time consuming, but it may lead to a treatment plans which could have been better with different beam angles and number of beams [5].

2.4.2. Treatment Plan Evaluation Tools

Different metrics can be used to evaluate the treatment plan quality. Firstly, a dose volume histogram (DVH) can be used. In a DVH, the 3D dose distribution in a structure is displayed in a 2D graph. In the graph, the relative volume is plotted against the dose. The points on the curve display how much of the volume of a structure receives at least a certain dose.

Another way to evaluate a treatment plan is with the percentile dosage $D_{x\%}$, which represents the minimum dose received by the $x\%$ of the volume which receives the most dose. For instance, $D_{98\%} = 54$ Gy means that the minimum dose received by 98% of the volume equals 54 Gy.

A more clinically significant treatment plan evaluation tool is the normal tissue complication probability (NTCP) model. The NTCP estimates the risk that an OAR risks radiation-induced complications. In this thesis, the Lyman Kutcher Burman model [10] is used to calculate the NTCP. The NTCP is a function of the generalized equivalent uniform dose (gEUD). Given an inhomogeneous dose distribution delivered to an OAR, an homogeneous irradiation with the gEUD will lead to the same effect on the OAR. The gEUD is given by:

$$\text{gEUD} = \left(\sum_j \nu_j \cdot D_j^{1/n} \right)^n , \tag{2.20}$$

where ν_j is the volume fraction of the OAR which receives dose D_j. n is the volume effect parameter and depends on the irradiated organ. If the organ has a serial architecture, e.g., the bowel, n is low. In organs with a serial architecture complications can already occur when a small region of the organ is irradiated. Organs with a parallel architecture, e.g., the lung, have a high n. Parallel architectures may receive a high dose in a small region without any complications.

The NTCP is given by:

$$\text{NTCP} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx, \tag{2.21}$$

where t is given by:

$$t = \frac{\text{gEUD} - TD_{50}}{m \cdot TD_{50}} \tag{2.22}$$

In Equation 2.22, TD_{50} is the tolerance dose for a 50% probability of complications in the organ and m is a fit parameter. TD_{50} and m are organ-dependent, just like n.

2.4.3. iCycle: Automated Multicriteria Treatment Planning System

In this thesis the fully automated treatment planning system iCycle is used. iCycle uses multicriteria optimization to choose the best beam angles and beam intensities in the case of IMRT [5], and the best pencil beam locations and corresponding intensities in the case of IMPT [46]. To do so, iCycle uses a wish-list. The wish-list contains prioritized objectives and constraints which are defined by the physician. The objectives are given a certain priority. An objective with a higher priority will be given more importance to be reached within the constraints than an objective with a lower priority. An objective could for instance be to deliver a prescribed dose $D^{\text{prescribed}}$ to the CTV and a constraint could be a maximum dose that is allowed to be delivered to a surrounding OAR.

To select the optimal beam angles in IMRT, iCycle first generates a plan with no beams. Then beams from a set of candidate beam directions are added to the plan. These candidate beam directions are chosen such that there would be no collisions between the patient couch and the gantry. The determination of the beam directions to add to the plan is an iterative process. In iteration i, every single non-chosen candidate beam direction is one-by-one individually added to the existing beam directions from iteration $i-1$ and the beam intensities for this beam configuration are optimized with a radiotherapy optimization problem. The beam direction which gives the most favorable dose distribution is then added to the previous $i-1$ beam directions to give a beam arrangement with i beams.

For the pencil beam optimization in IMPT, every iteration i a number of pencil beams are randomly selected from a fine grid and added to the existing pencil beams from iteration $i-1$. During the optimization process the optimal weight of each pencil beam is found. The pencil beams with a low weight after the optimization are excluded from the set of pencil beams, resulting in an optimal set of pencil beams for iteration i.

The optimal beam direction or pencil beams are found by using the $2p\epsilon c$ method, which results in a Pareto-optimal radiotherapy plan [4]. A Pareto-optimal plan is a plan in which an objective cannot be minimized more without negatively affecting the other objectives. The imposed constraints are strictly met in the Pareto-optimal plan and the objectives are optimized while taking into account their priorities. This optimization is achieved in two phases. In the first phase, each objective is minimized while respecting the constraints. The highest prioritized objective is minimized first, followed by the lower prioritized objectives. When an objective has been minimized, this objective is turned into a constraint. This then serves as a constraint for the lower prioritized objectives. Defining these new constraints assures that the minimization of objectives with a lower priority will not negatively affect the obtained values of the higher priority objectives. Consequentially, more constraints are imposed for the minimization of lower objectives. At the end of the first phase, each objective has reached its goal value or will be higher than its goal value due to the imposed constraints. If there are objectives which could have been minimized beyond their goal value in the first phase, they will be minimized as much as possible in the second phase. This is again done in the order of decreasing objective priority.

The resulting Pareto-optimal plan is the i^{th} treatment plan. In IMRT, the beam direction in the i^{th} treatment plan is in turn used to determine the $(i+1)^{\text{th}}$ beam direction, while in IMPT the pencil beams in the i^{th} treatment plan are added to the random selection of pencil beams to determine the $(i+1)^{\text{th}}$ treatment plan. This iterative process continues until the addition of more beams or pencil beams does not lead to a better clinically significant plan. A Pareto-optimal plan can then be selected which has the best trade-off between the quality of the plan and the number of beams.

2.4.4. Robust Treatment Planning

As discussed in the introduction, the concept of the PTV is applicable to photon therapy, but not applicable to IMPT. In photon therapy the dose distribution stays largely the same when there is a change in the geometry of the patient. Therefore, a movement of the CTV within the PTV will lead to an approximately equal dose delivered to the CTV. However, in the case of IMPT the dose distribution does not stay the same when the patient geometry changes. This is because density differences along the beam can alter the position of the Bragg peak. Furthermore, the range uncertainty of the proton beams can lead to an under- or overshoot of the CTV, which can cause overdosage or underdosage in the CTV [43]. To account for the uncertainties in proton therapy, robust treatment planning is used. In robust treatment planning, systematic and random setup errors are considered for, along with range errors. iCycle uses objective-wise minimax optimization to perform robust treatment planning. In objective-wise minimax optimization, the dose distribution is considered in a scenario with no errors (nominal scenario) and a number of error scenarios. In these scenarios, the worst case value of the objectives and constraints is optimized. To account for the range error along the beam direction, a range robustness (RR) setting is used, which is given as a percentage. For the random and systematic setup errors in the x, y and z direction, a setup robustness (SR) setting is used, given in millimeters. In iCycle a total of 19 scenarios are considered during the objective-wise minimax optimization.

2.5. Polynomial Chaos Expansion (PCE)

In this section the theory behind Polynomial Chaos Expansion (PCE) will be introduced. Polynomial Chaos (PC) techniques are schemes which approximate a model output as a polynomial function of the model input parameters and are part of a group of spectral techniques of which the aim is to use a Fourier series like expansion to reproduce the solution of a stochastic problem. PCE is a PC function which forms a meta-model of the problem it is designed for and it can be used to describe the stochastic nature of the output of the problem [28].

In PCE, the output of the problem is represented by a Fourier series like expansion using multi-dimensional polynomials, also called PC basis vectors. The spectral expansion of the output R as a function of the uncertain input $\vec{\xi}$ is expressed as:

$$R(\vec{\xi}) = \sum_{k=0}^{\infty} r_k \Psi_k(\vec{\xi}), \tag{2.23}$$

where Ψ_k are the PC basis vectors and r_k are the expansion coefficients. The uncertain input $\vec{\xi}$ is given by:

$$\vec{\xi} = (\xi_1, \xi_2, ..., \xi_S), \tag{2.24}$$

which is a vector with length S, where each element ξ_j corresponds to a random variable with a probability density function (PDF) $p_{\xi_j}(\xi_j)$. In this research it is assumed that the random variables are independent and the joint PDF is simply given by:

$$p_{\vec{\xi}}(\vec{\xi}) = \prod_{j=1}^{S} p_{\xi_j}(\xi_j) \tag{2.25}$$

2.5.1. PC Basis Vectors

The PC basis vectors are constructed via the tensorization of uni-variate polynomials. Each uni-variate polynomial depends on a random variable ξ_j. Based on the distribution of the uncertain input $\vec{\xi}$, the polynomial type is provided by the Wiener-Askay scheme, which gives the most optimal polynomial type for the most common uncertain input distributions [47]. This research focuses on a Gaussian distributed random input, so probabilists' Hermite polynomials He_o are the most suitable. Here o denotes the order of the polynomial. The first five orders of the Hermite polynomials He_o are displayed in Table 2.1.

Table 2.1: The first five polynomials of the probabilists' Hermite polynomial family

Order	Polynomial
$o = 0$	$He_0(\xi) = 1$
$o = 1$	$He_1(\xi) = \xi$
$o = 2$	$He_2(\xi) = \xi^2 - 1$
$o = 3$	$He_3(\xi) = \xi^3 - 3\xi$
$o = 4$	$He_4(\xi) = \xi^4 - 6\xi^2 + 3$

With the choice of the probabilists' Hermite polynomials, all PC basis vectors are of the form

$$\Psi_k(\vec{\xi}) = \prod_{j=1}^{S} He_{\gamma_{k,j}}(\xi_j), \tag{2.26}$$

where $\vec{\gamma}_{k,j} = (\gamma_{k,1}, ..., \gamma_{k,S})$, which is a multi-index to differentiate between the different orders of the polynomials. The index $\vec{\gamma}$ is used to determine which orders of polynomials are included in the PCE.

In the full domain of the random variables, indicated by $\mathcal{D}$, the PC basis vectors defined in 2.26 are orthogonal, i.e.:

$$\langle \Psi_k, \Psi_l \rangle = \int_{\mathcal{D}} \Psi_k(\vec{\xi})\Psi_l(\vec{\xi})p_{\vec{\xi}}(\vec{\xi})\, d\vec{\xi} = h_k^2 \delta_{k,l}, \tag{2.27}$$

where $\langle . \rangle$ is the inner product in $\mathcal{D}$, $\delta_{k,l}$ is the Kronecker delta and h_k is the norm of the k-th basis vector. The O^{th} order full PC basis set $\Gamma(O)$ is given by:

$$\Gamma(O) = \left\{ \prod_{j=1}^{S} He_{\gamma_{k,j}}(\xi_j) : \sum_{j=1}^{S} \gamma_{k,j} \leq O \right\} = \left\{ \Psi_k(\vec{\xi}) : \sum_{j=1}^{S} \gamma_{k,j} \leq O \right\} \tag{2.28}$$

The number of basis vectors for an O^{th} order PC basis set is given by:

$$P + 1 = \frac{(S+O)!}{S!O!} \tag{2.29}$$

The PCE in Equation 2.23 can then be truncated to $P + 1$ PC basis vectors, i.e.:

$$R(\vec{\xi}) \approx \sum_{k=0}^{P} r_k \Psi_k(\vec{\xi}), \tag{2.30}$$

with $\Psi_k(\vec{\xi}) \in \Gamma(O) \forall k$.
For example, in a 2D problem with input parameters $\vec{\xi} = (\xi_1, \xi_2)$, where ξ_1 and ξ_2 both are normally distributed, and $O = 2$, the full PC basis set would contain 6 basis vectors according to Equation 2.29. These 6 basis vectors along with the corresponding multi-indices are displayed in Table 2.2.

Table 2.2: The PC basis vectors with the corresponding multi-indices for a 2D problem with a 2nd order full basis set

Order	Multi-indices	PC basis vectors
$o = 0$	$\vec{\gamma}_0 = (0, 0)$	$\Psi_0(\xi) = He_0(\xi_1)He_0(\xi_2) = 1$
$o = 1$	$\vec{\gamma}_1 = (1, 0)$	$\Psi_1(\xi) = He_1(\xi_1)He_0(\xi_2) = \xi_1$
	$\vec{\gamma}_2 = (0, 1)$	$\Psi_2(\xi) = He_0(\xi_1)He_1(\xi_2) = \xi_2$
$o = 2$	$\vec{\gamma}_3 = (2, 0)$	$\Psi_3(\xi) = He_2(\xi_1)He_0(\xi_2) = \xi_1^2 - 1$
	$\vec{\gamma}_4 = (1, 1)$	$\Psi_4(\xi) = He_1(\xi_1)He_1(\xi_2) = \xi_1\xi_2$
	$\vec{\gamma}_5 = (0, 2)$	$\Psi_5(\xi) = He_0(\xi_1)He_2(\xi_2) = \xi_2^2 - 1$

2.5.2. Hyperbolic Trim

In practice, it is often not required to use all PC basis vectors in the full PC basis set. This is because responses are generally dominated by only a small number of important parameters and low order interactions, as stated by the ”sparsity of effects” principle [25]. As a consequence, PC basis vectors which represent higher order interactions can be left out of the PC basis set. This process of removing PC basis vectors from the full PC basis set is called hyperbolic trimming. Hyperbolic trimming is done using the q-quasi-norm for multi-indices $\vec{\gamma}_{k,j}$, which is defined as [3]

$$||\vec{\gamma}_k||_q = \left(\sum_{j=1}^{S} \gamma_{k,j}^q \right)^{1/q}, \tag{2.31}$$

where q is the hyperbolic trim factor which ranges between the values 0 and 1. Only the vectors which satisfy $||\vec{\gamma}_k||_q \leq O$ are included in the PC basis set. When $q = 1$, it can be seen that the full PC basis set is obtained. However, for decreasing values of q, more and more multivariate PC basis vectors, which depend on multiple terms in the input, will be cut out. The univariate PC basis vectors, depending on one value in the input, will however never be cut out. This leads to a PC basis set containing less PC basis vectors which are mainly a function of the dominant univariate dependencies and low order interactions [27].

2.5.3. Expansion Coefficients

Now the expansion coefficients r_k can be determined. Because the $\Psi_k(\vec{\xi})$ polynomial basis functions are orthogonal to each other with respect to the joint probability density function $p(\vec{\xi})$, the expansion coefficients r_k can be calculated with:

$$\begin{aligned} r_k &= \frac{\langle R(\vec{\xi}), \Psi_k(\vec{\xi})\rangle}{\langle \Psi_k(\vec{\xi}), \Psi_k(\vec{\xi})\rangle} \\ &= \frac{\int_{\mathcal{D}} R(\vec{\xi})\Psi_k(\vec{\xi})p_{\vec{\xi}}(\vec{\xi})d\vec{\xi}}{\langle \Psi_k(\vec{\xi}), \Psi_k(\vec{\xi})\rangle} \\ &= \frac{1}{h_k^2}\int_{\mathcal{D}} R(\vec{\xi})\prod_{j=1}^{S} He_{\gamma_{k,j}}(\xi_j)p_{\xi_j}(\xi_j)\, d\vec{\xi} \end{aligned} \tag{2.32}$$

This method of calculating the expansion coefficients is called spectral projection. Equation 2.32 can however not be readily solved as the numerator contains the unknown dependence of the response on the input parameters $R(\vec{\xi})$. To solve this problem, the multi-dimensional integral is approximated by a cubature formula. This cubature formula is constructed from one-dimensional quadratures using tensorization. The quadrature formula for a general function $f(\xi_j)$, which depends on a single variable ξ_j, is given by:

$$I^{(1)}f = \int_a^b f(\xi_j)p_{\xi_j}(\xi_j)\, d\xi_j \approx Q_{lev}^{(1)}f = \sum_{i=1}^{n_{lev}} f(\xi_{j,lev}^{(i)})w_{lev}^{(i)}, \tag{2.33}$$

where $\xi_{j,lev}^{(i)} \in [a,b]$ and $w_{lev}^{(i)} \in \mathbb{R}$ are quadrature points and weights which are predefined by $p_{\xi_j}(\xi_j)$ and the chosen quadrature rule. The level index lev indicates the accuracy of the quadrature. A higher level index gives a more precise approximation of the integral but this also gives a higher number of function evaluations n_{lev}.

There are many different quadrature rules which can be distinguished by their accuracy and by their nestedness. Quadratures in rules with full nestedness contain all points of the lower level quadratures too, while quadratures in rules with no nestedness contain no points of the lower level quadratures.

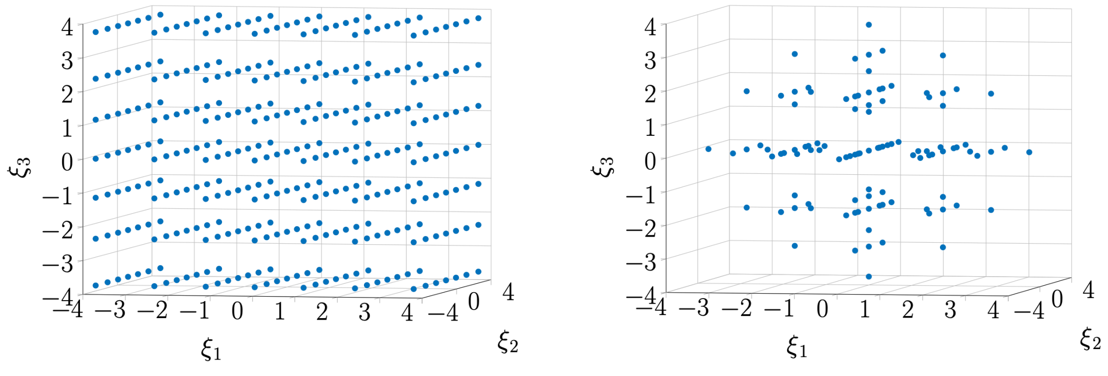
In this research the Gauss-Hermite quadrature rule is used. The Gauss-Hermite quadrature rule can integrate accurately up to order $2n_{lev}-1$ and it has a low nestedness. The number of value evaluations n_{lev} is chosen to be equal to to $2\cdot lev-1$, so the order of polynomials that the Gauss-Hermite quadrature rule is accurate to equals $2(2\cdot lev-1)-1 = 4\cdot lev-3$.

The one-dimensional quadratures are used to form the cubature by tensorization. An S dimensional cubature calculated using a full tensorization looks as follows:

$$\begin{aligned} I^{(S)}f &= \int_{\mathcal{D}} f(\vec{\xi})p_{\vec{\xi}}(\vec{\xi})\, d\vec{\xi} \approx Q_{\vec{lev}}^{(S)}f = (Q_{lev_1}^{(1)} \otimes Q_{lev_2}^{(1)} \otimes \cdots \otimes Q_{lev_S}^{(1)})f \\ &= \sum_{i_1=1}^{n_{lev_1}}\sum_{i_2=1}^{n_{lev_2}}\cdots\sum_{i_S=1}^{n_{lev_S}} f(\xi_{1,lev_1}^{(i_1)}, \xi_{2,lev_2}^{(i_2)}, \cdots, \xi_{S,lev_S}^{(i_S)})w_{lev_1}^{(i_1)}w_{lev_2}^{(i_2)}\cdots w_{lev_S}^{(i_S)} \\ &= \sum_{i}^{n} f(\xi^{(i)})w^{(i)} \end{aligned} \tag{2.34}$$

In the final line of Equation 2.34 there is a sum over all possible combinations of quadrature points. This will lead to an exponentially growing number of required needed evaluations of $f(\vec{\xi})$ as the dimension of the problem increases. The number of function evaluations needed for an S-dimensional problem is equal to $\prod_{j=1}^{S} n_{lev_j}$.

In Figure 2.6a an example of a full 4th level Gauss-Hermite grid in 3 dimensions can be seen. A 4th level grid contains 7 points which means that a total of $7^3 = 343$ function evaluations need to performed for the 3 dimensions. To decrease the number of function evaluations needed, and thus decrease the computation time, Smolyak sparse grids are used. The idea of Smolyak sparse grids is



(a) Cubature points for the full grid

(b) Cubature points for the Smolyak sparse grid

Figure 2.6: Full and Smolyak sparse 4$^{\text{th}}$ level Gauss-Hermite grid in 3 dimensions. Figures reproduced from [27].

that there are elements calculated with the quadrature rules which are more important for representing multidimensional functions than others [20]. To produce the Smolyak sparse grid, a small number of the most important cubature points are selected. In order to select the cubature points to include in the Smolyak sparse grid, the difference formula of the quadratures is used:

$$\Delta^{(1)}_{lev} f = Q^{(1)}_{lev} f - Q^{(1)}_{lev-1} f, \tag{2.35}$$

with $Q^{(1)}_0 f = 0$. With Equation 2.35 the cubature tensorization (Equation 2.34) can be rewritten as:

$$Q^{(S)}_{\vec{lev}} f = \sum_{l_1=1}^{lev_1} \sum_{l_2=1}^{lev_2} \cdots \sum_{l_S=1}^{lev_S} (\Delta^{(1)}_{l_1} \otimes \Delta^{(1)}_{l_2} \otimes \cdots \otimes \Delta^{(1)}_{l_S}) f, \tag{2.36}$$

where $\vec{lev}$ is a vector which contains the different quadrature levels used along different directions. By introducing a multi-index $\vec{l}$ which distinguishes the different grids, the cubature tensorization turns into:

$$Q^{(S)}_{\vec{lev}} f = \sum_{\vec{l} \in \mathcal{J}(lev)} \Delta^{(S)}_{\vec{l}} f, \tag{2.37}$$

with $\mathcal{J}(lev)$ being the set of included multi-indices depends on the level lev. For the Smolyak sparse grid, the multi-indices included in $\mathcal{J}$ are given by:

$$\mathcal{J}^{Smolyak}(lev) = \left\{ \vec{l} : \sum_{j=1}^{S} l_j \leq lev + S - 1 \right\} \tag{2.38}$$

The consequence of the definition of the multi-indices in equation 2.38 is that the maximum integration level in any direction is equal to lev and the maximum dimension of the included grids is $lev - 1$. The use of a Smolyak drastically decreases the number of function evaluations needed. In Figure 2.6b it can be seen that a sparse grid has significantly less cubature points than a full grid. In this particular case of a 4th level grid, only 105 function evaluations need to take place compared to 343 in a full grid.

In this thesis the extended Smolyak sparse grids will be used, which is a method to improve the integration accuracy with a small number of extra function evaluations. In the extended Smolyak sparse grids, the integration level of the one-dimensional grids of the highest order with $l_j = lev + S - 1$, where j indicates one of the dimensions, is increased. This is done by increasing l_j with lev_{extra}, i.e.,: $l_j = lev + S - 1 + lev_{extra}$. The multi-indices in $\mathcal{J}$ are then given by:

$$\mathcal{J}^{Extended\ Smolyak}(lev) = \mathcal{J}^{Smolyak}(lev) \cup \{\vec{l} : l_j = lev + S - 1 + lev_{extra}\} \setminus \{\vec{l} : l_j = lev + S - 1\} \tag{2.39}$$

The replacement of the level lev with $lev + lev_{extra}$ leads to only $2 \cdot lev_{extra}$ extra calculations per dimension, while significantly improving the performance of the PCE. Furthermore, higher order univariate polynomials can be used in the PCE with the addition of lev_{extra} [27]. In the rest of the thesis, the extra level index lev_{extra} will be referred to as Extra Levels (EL).

The polynomial chaos coefficients from Equation 2.32 can now be calculated using the extended Smolyak sparse grids with:

$$r_k = \frac{1}{h_k^2} \sum_{\vec{l} \in \mathcal{J}^{Extended\ Smolyak}(lev)} \Delta_{\vec{l}}^{(S)}(R\Psi_k) = \frac{1}{h_k^2} \sum_{i}^{n} R(\vec{\xi}^{(i)}) \Psi_k(\vec{\xi}^{(i)}) w^{(i)}, \tag{2.40}$$

where the function f in Equation 2.32 has been replaced by the product of the response and the PC basis function. With Equation 2.40 only a limited number of responses $R(\vec{\xi}^{i})$ are needed to calculate the PC coefficients.

In the rest of this thesis the grid level lev will be called the Grid Level (GL) and the polynomial order O will be referred to as PO. The GL and PO must be chosen such that a desired integration accuracy is reached. Once the desired accuracy has been reached, the cubature points $\vec{\xi}^{(i)}$ and the weights $w^{(i)}$ need to be determined which belong to the grids given by Equation 2.39 for the given GL. With this info, Equation 2.40 can be solved to determine the PC coefficients, which in turn are used in the truncated response function (Equation 2.30) in combination with the corresponding PC basis vectors.

2.5.4. Multi-Dimensional Polynomial Regression

Another way to build the PCE is with multi-dimensional polynomial regression. In multi-dimensional polynomial regression, the expansion coefficients r_k are calculated with a least square problem. Given a set of $N_{samples}$ samples of the uncertain input $\vec{\xi}^{(i)}$, with the corresponding responses $R(\vec{\xi}^{(i)})$ and PC basis functions $\Psi_k(\vec{\xi}^{(i)})$, the expansion coefficients are obtained by solving the least square problem:

$$\vec{r} = \arg\min_{\tilde{\vec{r}}} \sum_{i=1}^{N_{samples}} \left(R(\vec{\xi}^{(i)}) - \sum_{k=0}^{P} r_k \Psi_k(\vec{\xi}^{(i)}) \right)^2 , \tag{2.41}$$

where $\vec{r} = (r_0, \cdots, r_P)$, The solution to the least square problem, given that $N_{samples} \geq P$, is given by:

$$\vec{r} = \left(\mathbf{A}^T \mathbf{A}\right)^{-1} \mathbf{A}^T \begin{pmatrix} R(\vec{\xi}^{(1)}) \\ \vdots \\ R(\vec{\xi}^{(N_{samples})}) \end{pmatrix}, \tag{2.42}$$

where $\mathbf{A} \in \mathbb{R}^{N_{samples} \times P}$, with $A_{i,k} = \Psi_k(\xi^{(i)})$. The points where the uncertain input $\vec{\xi}$ is sampled can be chosen either randomly according to the distribution of the elements in the uncertain input or can be equal to the quadrature points which are described in Section 2.5.3. To obtain a robust solution and to prevent overfitting, $N_{samples} \approx 2(P+1)$ has been found to be a sufficient number of sample points [17]. Once the expansion coefficients $\vec{r}$ have been determined, they can be used in the truncated response function (Equation 2.30) along with the corresponding PC basis vectors to form the PCE.

2.5.5. PCE Applied to Delineation Uncertainties

In this thesis PCE will be used to model dose distributions (exact output R) as a function of the random eigenmode coefficients $c_{l,\text{random}}$ of the PCA model (uncertain input). The uncertain input vector will thus look like:

$$\vec{\xi} = (c_{1,\text{random}}, c_{2,\text{random}}, \cdots, c_{L,\text{random}}) \tag{2.43}$$

All eigenmode coefficients $c_{l,\text{random}}$ are Gaussian distributed with mean 0 and standard deviation $\sqrt{\lambda_l}$.

3

Methods & Materials

In this chapter the methods used in this thesis are shown. First the raw data of both a manual delineation data set and an auto-contouring data set are shown in Section 3.1, along with the wish-lists to make a radiotherapy plan with these data sets. Then the construction of the delineation uncertainty model is discussed for both the manual delineations (Section 3.2) and the auto-contours (Section 3.3). Subsequently the construction and validation methods of the PCE as a meta-model for dosimetric quantities of random delineations in a fixed dose distribution are shown in Section 3.4 and finally the construction and validation methods of the PCE as a meta-model for the reoptimized dose distribution are shown in Section 3.5.

3.1. Raw Patient Data

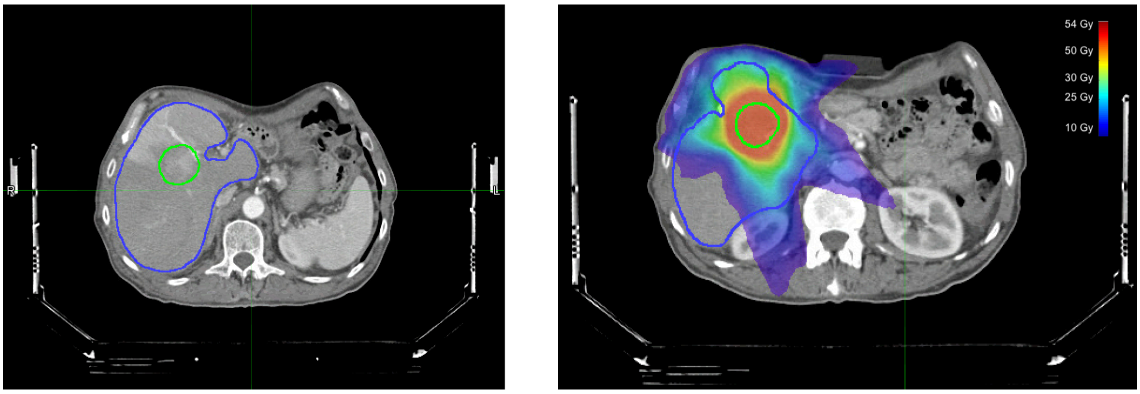
In this thesis two data sets were used. One data set was used to study the uncertainties in manual delineations and the other data set was used to study the uncertainties in auto-contouring.

3.1.1. Manual Delineation Data Set

The first data set was obtained from a study performed by *Habraken et al.* [15]. In this study, 12 participating centers were provided with a CT scan of the same patient with hepatocellular carcinoma. The resolution of the CT was $0.9766 \times 0.9766 \times 2.5000$ mm. The participating centers submitted delineations of the GTV and the liver on the CT scan. For this data set the CTV was equal to the GTV. One of the delineation sets was marked as the planning delineation set and a VMAT plan was made with this these delineations, which also included delineations of the stomach, kidneys, esophagus, bowel, galbladder, heart and spinal cord. The CT with the planning delineations of the CTV and liver is displayed in Figure 3.1a. The prescribed dose delivered to the PTV (CTV plus a 5 mm margin) with this plan was 54 Gy in six fractions. The dose distribution of this plan can be seen in Figure 3.1b. For the CTV, the dose constraint was that the dose received by 98% of the volume had to be at least the prescribed dose on the PTV ($D_{98\%} \geq 54$ Gy). Furthermore, for the PTV the dose constraint was that 95% of the volume had to receive at least 54 Gy ($D_{95\%} \geq 54$ Gy).
In this thesis an IMPT plan was also made for the above patient. To create an IMPT plan for this patient, a wish-list was constructed, which can be seen in Table 3.1. Both non-robust and robust treatment plans were made. In the case of the robust treatment plans, the RR equalled 2%, the SR was set to 1, 2, 3, 4 or 5 mm and only the constraint on the minimum dose in the CTV was optimized robustly.

3.1.2. Auto-Contouring Data Set

The second data set was obtained from a patient at the Erasmus Medical Center in Rotterdam, The Netherlands. This data set consisted of one planning CT and six repeat CTs of the head and neck region. All CTs had a resolution of $0.9766 \times 0.9766 \times 2.0000$ mm. The structures on the planning CT were manually delineated by a radiation oncologist. The most important delineated structures are the CTV along with the OARs near the CTV. The OARs near the CTV are the brainstem and spinal cord. The CTV was separated into a high dose CTV and a low dose CTV. The prescribed dose to the high



(a) Planning delineations of the CTV (green) and the liver (blue)

(b) Dose distribution of the VMAT plan planned on planning delineations

Figure 3.1: Planning delineations of the CTV and liver with the corresponding dose distribution in a 54 Gy VMAT plan for the hepatocellular carcinoma patient.

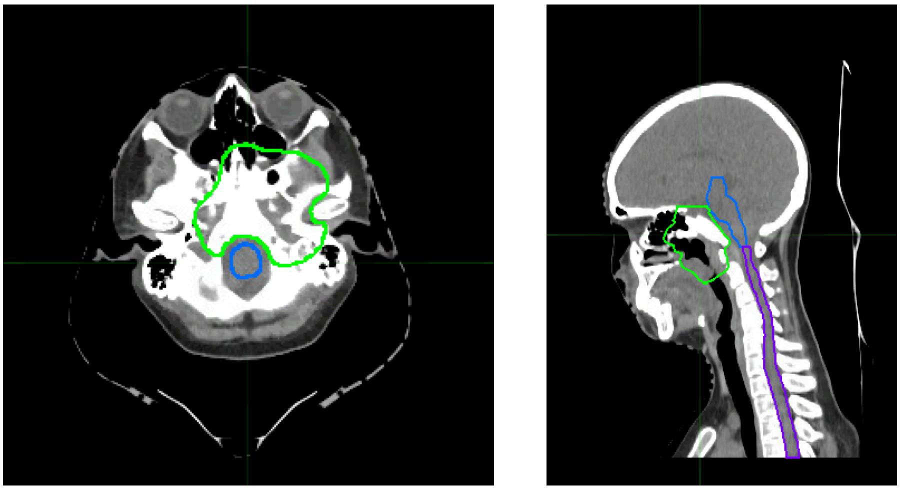
Table 3.1: The wish-list to generate treatment plans in iCycle for the hepatocellular carcinoma patient. The column 'Type' indicates whether the goal is to minimize/maximize the maximum/minimum dose (Linear), or to minimize/maximize the mean dose (Mean). 'Ring' indicates a ring around the structure. The structure 'MU' (monitor units) is added to push down the dose overall dose in the patient.

Priority	Structure	Min/Max	Type	Goal
Constraint	CTV	Maximize maximum	Linear	54 Gy
1	CTV	Minimize maximum	Linear	60 Gy
2	CTV-ring$_{0\text{-}5mm}$	Minimize maximum	Linear	0.90*60 Gy
3	CTV-ring$_{5\text{-}10mm}$	Minimize maximum	Linear	0.85*60 Gy
4	CTV-ring$_{10\text{-}15mm}$	Minimize maximum	Linear	0.75*60 Gy
5	CTV-ring$_{15\text{-}25mm}$	Minimize maximum	Linear	0.50*60 Gy
6	Liver excluding CTV	Minimize minimum	Linear	1 Gy
7	Liver excluding CTV	Minimize maximum	Mean	1 Gy
8	Stomach	Minimize maximum	Linear	1 Gy
9	Bowel	Minimize maximum	Linear	1 Gy
10	Esophagus	Minimize maximum	Linear	1 Gy
11	Spinal cord	Minimize maximum	Linear	1 Gy
12	Right kidney	Minimize maximum	Mean	1 Gy
13	Heart	Minimize maximum	Linear	1 Gy
14	MU	Minimize maximum	Linear	1

dose CTV was 70 Gy and the prescribed dose to the low dose CTV was 54 Gy. The dose contraint to the CTVs was that 98% of the CTV volume had to receive at least 95% of the prescription dose ($D_{98\%} \geq 0.95 \cdot D^{\text{prescribed}}$). The delineations of the high dose CTV, brainstem and spinal cord on the planning CT are displayed in Figure 3.2. The delineations on the planning CT were registered to the six repeat CTs using B-spline DIR and the resulting delineations on the repeat CTs were checked and modified if necessary. The wish-list to make an IMPT treatment plan for this patient is shown in Table 3.2. The dose distribution of the IMPT plan made with iCycle on the planning CT with the manual delineations can be seen in Figure 3.3.

3.2. Delineation Uncertainty Model for Inter-Observer Variability

In this section the formation of the delineation uncertainty model for inter-observer variability between the manual delineations of the GTV in the liver is set out, along with metrics used to determine the accuracy of the model.



(a) Transverse view of the planning CT with delineations of the high dose CTV (green) and brainstem (blue)

(b) Sagittal view of the planning CT with delineations of the high dose CTV (green), brainstem (blue) and spinal cord (purple)

Figure 3.2: Planning CT with delineations of the high dose CTV, brainstem and spinal cord.

3.2.1. Building the Uncertainty Model

PCA was used to develop a delineation uncertainty model of the manual delineations of the GTV in the liver. To do so, the delineation information contained in the DICOM files needed to be preprocessed. The DICOM files of the delineations of the GTV were imported into Matlab [23] using the function *dicomrt2matlab* [26]. Using this function, the delineation information of the manual delineations was converted from DICOM files to Matlab structures. These structures contained the delineation information of the GTV in the form of the coordinates of the delineation on each CT slice and in the form of a 3D binary map which was equal to 1 inside the GTV and equal to 0 outside the GTV. To convert the slice-per-slice delineations into a 3D structure, the Matlab function *boundary* was used, which forms an exterior surface around the slice-per-slice delineations. The 3D structures of the GTV delineations by the different observers were used to form a probability map of the collection of delineations. The probability map was created by adding the 3D binary masks of the delineations made by the different observers and normalizing it to the number of delineations. A threshold of 50% was then used to form the gold standard delineation.

From the gold-standard delineation, the eigenmodes v_l and the corresponding eigenvalues λ_l were calculated as described in Section 2.2. To calculate the distance from the center of mass of the gold-standard delineation to the surface of each delineation, the function *Triangle/Ray Intersection* was used [42].

3.2.2. PCA Model Accuracy

To evaluate the accuracy of the PCA model of the delineations with respect to the number of eigenmodes, two metrics were used. Firstly, the number of eigenmodes needed to explain 90% of the variance of the data was determined. Secondly, the reconstruction error of the PCA model as a function of the number of eigenmodes was determined. The reconstructed delineations with L eigenmodes were compared to the real delineations. The local absolute reconstruction error $\vec{d}^{L}_{abs,i}$ for delineation i reconstructed with L eigenmodes is given by:

$$\vec{d}^{L}_{abs,i} = |\vec{p}_i - \vec{p}^{L}_{i,rec}|, \tag{3.1}$$

where $\vec{p}_i - \vec{p}^{L}_{i,rec}$ indicates an element-wise subtraction and |...| indicates the absolute value.

The local absolute reconstruction error was measured to determine the minimum number of eigenmodes necessary to give an accurate reconstruction.

Table 3.2: The wish-list to generate treatment plans in iCycle for the head and neck patient. The column 'Type' indicates whether the goal is to minimize/maximize the maximum/minimum dose (Linear), or to minimize/maximize the mean dose (Mean). 'Expanded' indicates an expansion of the structure, 'shrunk' indicates a shrinkage of the structure and 'ring' indicates a ring around the structure. The structure 'MU' (monitor units) is added to push down the dose overall dose in the patient.

Priority	Structure	Min/Max	Type	Goal
Constraint	Brainstem	Minimize maximum	Linear	54 Gy
Constraint	Spinal cord	Minimize maximum	Linear	50 Gy
Constraint	Optic nerves	Minimize maximum	Linear	55 Gy
Constraint	Optic chiasm	Minimize maximum	Linear	55 Gy
Constraint	Lenses	Minimize maximum	Linear	10 Gy
Constraint	High dose CTV	Maximize minimum	Linear	0.97*70 Gy
Constraint	Low dose CTV-expanded$_{10\text{mm}}$	Maximize minimum	Linear	0.98*54 Gy
Constraint	Low dose CTV-expanded$_{10\text{mm}}$	Maximize minimum	Linear	0.98*54 Gy
Constraint	High dose CTV	Minimize maximum	Linear	1.3*70 Gy
Constraint	Low dose CTV-expanded$_{10\text{mm}}$	Minimize maximum	Linear	1.3*54 Gy
Constraint	Low dose CTV-shrunk$_{10\text{mm}}$	Minimize maximum	Linear	1.3*54 Gy
1	High dose CTV ∩ OARs	Maximize minimum	Linear	0.97*70 Gy
1	High dose CTV ∩ OARs	Maximize minimum	Mean	0.96*70 Gy
2	High dose CTV	Minimize maximum	Linear	1.06*70 Gy
2	Low dose CTV-expanded$_{10\text{mm}}$	Minimize maximum	Linear	1.06*54 Gy
2	Low dose CTV-shrunk$_{10\text{mm}}$	Minimize maximum	Linear	1.06*54 Gy
3	(High dose CTV ∪ low dose CTV)-ring$_{0\text{-}10\text{mm}}$	Minimize maximum	Linear	1.06*54 Gy
3	(High dose CTV ∪ low dose CTV)-ring$_{10\text{-}15\text{mm}}$	Minimize maximum	Linear	0.90*54 Gy
4	MU per spot	Minimize maximum	Linear	35
5	Brainstem	Minimize maximum	Linear	30 Gy
5	Spinal cord	Minimize maximum	Linear	30 Gy
6	Cochlea	Minimize maximum	Linear	40 Gy
7	Optic nerves	Minimize maximum	Linear	30 Gy
7	Optic chiasm	Minimize maximum	Linear	30 Gy
8	Eyes	Minimize maximum	Mean	30 Gy
8	Lenses	Minimize maximum	Linear	5 Gy
9	Brain	Minimize maximum	Linear	30 Gy
10	Parotids	Minimize maximum	Mean	1 Gy
11	Submandibular glands	Minimize maximum	Mean	1 Gy
12	Musc constrict	Minimize maximum	Mean	1 Gy
13	Oral cavity	Minimize maximum	Mean	1 Gy
15	Esophagus	Minimize maximum	Mean	1 Gy
15	Larynx	Minimize maximum	Mean	1 Gy
15	Glottis	Minimize maximum	Mean	1 Gy
15	Cricopharyngeus	Minimize maximum	Mean	1 Gy
16	(High dose CTV ∪ low dose CTV)-ring$_{0\text{-}10\text{mm}}$	Minimize maximum	Mean	1 Gy
16	(High dose CTV ∪ low dose CTV)-ring$_{10\text{-}15\text{mm}}$	Minimize maximum	Mean	1 Gy
16	(High dose CTV ∪ low dose CTV)-ring$_{15\text{-}25\text{mm}}$	Minimize maximum	Mean	1 Gy
17	MU	Minimize maximum	Linear	1

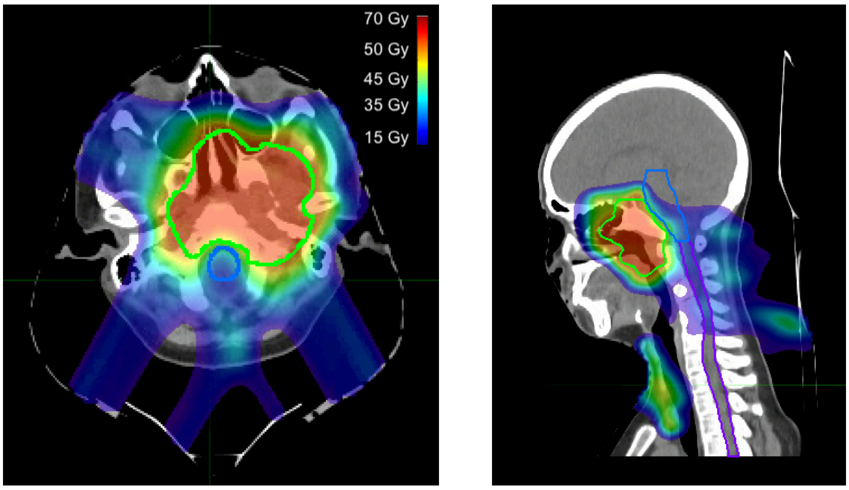
3.2.3. Generating 2D Delineations from 3D Shapes

The goal of PCA was to create random delineations. The delineations created with PCA are however in 3D, while real delineations are made slice per slice in 2D. Therefore, the 3D random delineations were projected back onto the 2D slices of the CT scan and saved as Cartesian coordinates on that slice. For a CT scan with K slices the following steps were taken:

- The random surface shape vector $\vec{p}^{L}_{\text{random}}$ was converted into Cartesian coordinate vectors $\vec{x}_{\text{random}}$, $\vec{y}_{\text{random}}$ and $\vec{z}_{\text{random}}$.
- For the CT slice $(k = 1, ..., K)$ at height $h_{\text{slice}}(k)$, it is determined which values in the vector $\vec{z}_{\text{random}}$ are between 0.3 mm smaller and 0.3 mm bigger than $h_{\text{slice}}(k)$. The indices of these values in $\vec{z}_{\text{random}}$ were stored in the vector $\vec{a}_{\text{slices}}(k)$.
- The vectors $\vec{x}_{\text{slice}}(k)$ and $\vec{y}_{\text{slice}}(k)$ contained the values of $\vec{x}_{random}$ and $\vec{y}_{random}$ at the indices $\vec{a}_{\text{slices}}(k)$.
- The function *boundary* was used to form a boundary around $\vec{x}_{\text{slice}}(k)$ and $\vec{y}_{\text{slice}}(k)$. The points in this boundary were saved in the vectors $\vec{x}_{\text{slice,boundary}}(k)$ and $\vec{y}_{\text{slice,boundary}}$.

$\vec{x}_{\text{slice,boundary}}(k)$ and $\vec{y}_{\text{slice,boundary}}(k)$ could now be saved as the coordinates of the random delineation at height $h_{\text{slice}}(k)$.

To avoid the generation of unrealistic delineations, the absolute value of random coefficients $c_{l,random}$



(a) Transverse view of the dose distribution on the planning CT with delineations of the high dose CTV (green) and brainstem (blue)

(b) Sagittal view of the dose distribution on the planning CT with delineations of the high dose CTV (green), brainstem (blue) and spinal cord (purple)

Figure 3.3: IMPT therapy dose distribution on the planning CT made with iCycle.

was limited $2.5 \cdot \sqrt{\lambda_l}$. In Figure 3.4 an example of an unrealistic delineation is shown. The 2D slice of the delineation is at the top of the delineation where it splits up into separate regions. As can be seen, the delineation comprises of three different regions at this height. However, the function *boundary* treats all the points as one region. Therefore the generation of multiple regions is avoided in this research. It was found that this was avoided when $|c_{l,random}| \leq 2.5 \cdot \lambda_l$.

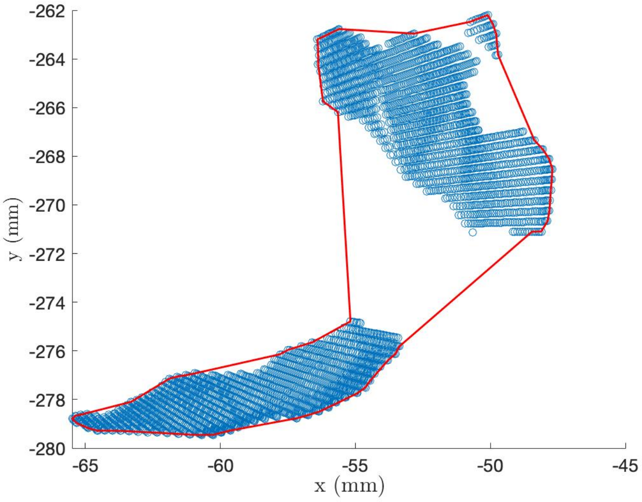


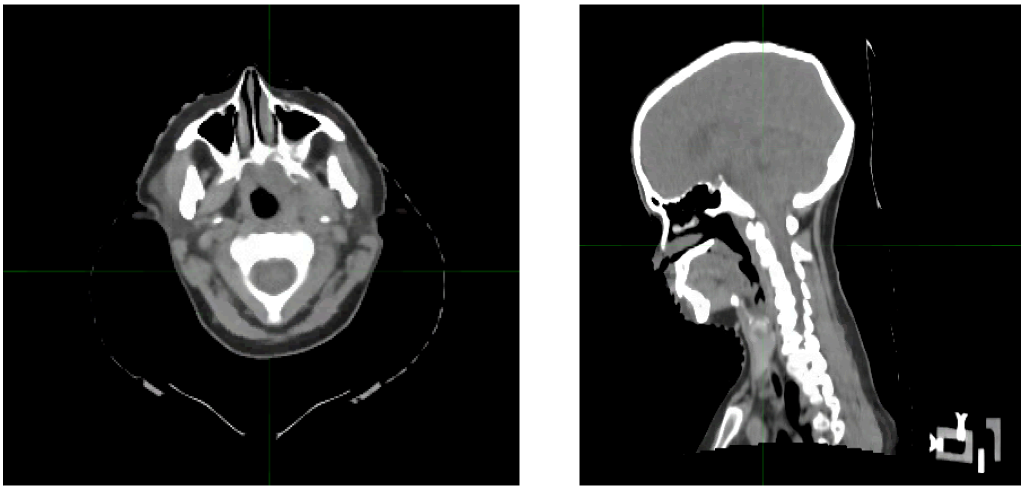
Figure 3.4: An example of the formation of an unrealistic delineation. The blue scatter points are the points given by $(\vec{x}_{\text{slice}}, \vec{y}_{\text{slice}})$ and the red line is the 2D delineation formed by the function $boundary$. The points $(\vec{x}_{\text{slice}}, \vec{y}_{\text{slice}})$ form three separate regions, but the 2D delineation only consists of one region.

3.3. Delineation Uncertainty Model for Auto-Contouring

The next goal was to make an uncertainty model of an auto-contouring software. In this thesis the commercially available auto-contouring software MIM (MIM Version 6.9.3, MIMVista Corp, Cleveland, Ohio) is used to create auto-contours. To create the auto-contours, MIM utilized atlas-based auto-contouring.

3.3.1. Generation of Synthetic CTs

To make an uncertainty model of the auto-contouring software MIM, first a large number of synthetic CTs were generated. To do so, the planning CT was registered to the six repeat CTs using B-spline DIR in RTStudio, i.e., the planning CT was the moving image and the repeat CTs were the fixed images. This resulted in six DVFs. From each DVF 15 randomly scaled DVFs were generated. These 15 randomly scaled DVFs were formed by sampling 15 scaling factors from a normal distribution with mean 0.5 and standard deviation 0.3 and multiplying the B-spline coefficients P_x, P_y and P_z of the DVF (Equation 2.18) with each scaling factor. Doing this for all six DVFs resulted in a total of 6×15=90 random DVFs. In RTStudio, the planning CT was then transformed with each of the 90 random DVFs to form 90 synthetic CTs. The scaling factors were sampled from a normal distribution with mean 0.5 because this would lead to synthetic CTs which are the most dissimilar to the planning CT and the repeat CTs, while still generating realistic CTs. If the sampled scaling factor was negative, this would still lead to a realistic synthetic CT. An example of a synthetic CT can be seen in Figure 3.5.



(a) Transverse view of the synthetic CT (b) Sagittal view of the synthetic CT.

Figure 3.5: One of the 90 synthetic CTs. The synthetic CT was made by deforming the planning CT (figure 3.2) with a random DVF.

3.3.2. Auto-contouring in MIM

To make auto-contours in MIM, first an atlas had to be made. This atlas consisted of the delineated planning CT and the delineated repeat CTs. The 90 synthetic CTs were imported into MIM. Within MIM, an automatized workflow was established which delineated the input CTs and saved the resulting delineations. It was chosen to use single-atlas selection auto-contouring, along with the default settings in MIM. In Figure 3.6 the workflow can be seen. Auto-contours were made on all 90 synthetic CTs using this workflow. The auto-contours were then all exported to RTStudio, where they were registered back to the planning CT using the inverse random DVFs. This resulted in 90 delineation sets on the planning CT.

3.3.3. Building and Validating the Uncertainty Model

The 90 delineation sets on the planning CT were used to form an uncertainty model of the auto-contouring software using PCA. The uncertainty model was made for the high dose CTV and brainstem delineations. The uncertainty model was made and validated using the same approaches as described in Section 3.2. One difference however was that the high dose CTV consisted of three separated regions. To perform PCA on the high dose CTV and the brainstem together, a gold-standard of each region of the CTV and of the brainstem were determined along with the corresponding centers of mass. From the centers of mass of the high dose CTV, three sets of surface shape vectors $\vec{p}_{i,j,\mathrm{CTV}}$ and three

Figure 3.6: Overview of the workflow in MIM. In the row *Atlas Segmentation*, the settings for the auto-contouring are displayed. It was chosen to use one match to run the auto-contouring, which is equivalent to single-atlas selection auto-contouring. The other settings are default auto-contouring settings in MIM. In the row *Save RTstruct*, the name and the save location of the auto-contour are defined.

mean shape vectors $\vec{\mu}_{j,\text{CTV}}$ were calculated for the high dose CTV, where $j \in (1, 2, 3)$ indicates the regions of the high dose CTV. From the center of mass of the brainstem, a set of surface shape vectors $\vec{p}_{i,\text{brainstem}}$ and a mean surface shape vector $\vec{\mu}_{\text{brainstem}}$ were calculated. The surface shape matrix $\mathbf{P}$ was then made as follows:

$$\mathbf{P} = \begin{pmatrix} \vec{p}_{1,1,\text{CTV}} - \vec{\mu}_{1,\text{CTV}} \\ \vdots \\ \vec{p}_{N,1,\text{CTV}} - \vec{\mu}_{1,\text{CTV}} \\ \vec{p}_{1,2,\text{CTV}} - \vec{\mu}_{2,\text{CTV}} \\ \vdots \\ \vec{p}_{N,2,\text{CTV}} - \vec{\mu}_{2,\text{CTV}} \\ \vec{p}_{1,3,\text{CTV}} - \vec{\mu}_{3,\text{CTV}} \\ \vdots \\ \vec{p}_{N,3,\text{CTV}} - \vec{\mu}_{3,\text{CTV}} \\ \vec{p}_{1,\text{brainstem}} - \vec{\mu}_{\text{brainstem}} \\ \vdots \\ \vec{p}_{N,\text{brainstem}} - \vec{\mu}_{\text{brainstem}} \end{pmatrix} \in \mathbb{R}^{4N \times M} \tag{3.2}$$

SVD was performed on the surface shape matrix in Equation 3.2, which resulted in eigenmodes v_l with $4N$ elements. Elements (1,...,N) of each eigenmode v_l were the eigenmode $v_{l,\text{CTV},1}$ of the first region of the high dose CTV, elements (N+1,...,2N) of v_l were the eigenmode $v_{l,\text{CTV},2}$ of the second region, elements (2N+1,...,3N) of v_l were the eigenmode $v_{l,\text{CTV},3}$ of the third region and elements (3N+1,...,4N) of v_l were the eigenmode $v_{l,\text{brainstem}}$ of the brainstem.

3.4. PCE Model for a Fixed Dose Distribution

Previous work by *van der Voort et al.* [44] has shown that PCE can be used as a meta-model for the dose distribution as a function of setup and range uncertainties in IMPT. Using PCE, it is possible to simulate many dose distributions for different setup and range error scenarios in a very short amount of time and to thus perform a statistically significant analysis on the effect of setup and range uncertainties. In this thesis PCE was first used to model the DVH of a delineated structure for a fixed dose distribution.

3.4.1. Construction of PCE

The PCE was constructed using the openGPC package in Matlab [28] [27]. The openGPC can build the PCE using sparse grids with a fixed PC basis set and can also do this adaptively.

The inputs of the openGPC code were the grid level GL, the number of extra levels EL, the polynomial order PO and the polynomial type, the included eigenmodes $(\vec{v}_1 \cdots \vec{v}_L)$ of the delineated structure, the mean μ and standard deviation σ of coefficients $(c_{1,random}, \cdots, c_{L,random})$ corresponding to the included eigenmodes and the fixed dose distribution location. The mean μ of the coefficients $(c_{1,random}, \cdots, c_{L,random})$ equalled 0 and the standard deviation σ equalled $(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_L})$

The next step in the openGPC code is to generate the cubature points needed to approximate the integrals to calculate the PC coefficients. In this thesis, the number of cubature points was reduced by using extended Smolyak Sparse grids.

Next, the PCE object is generated which contains the polynomial basis vectors with a maximum polynomial order of PO. Furthermore, a hyperbolic trim is performed on the polynomial basis vector set to exclude basis vectors which represent higher order interactions. The hyperbolic trim factor is calculated

within openGPC. If spectral projection is used, the hyperbolic trim q is chosen such that the value of the expansion coefficients can still be accurately determined. If multi-dimensional polynomial regression is used, no hyperbolic trim is applied, but it is checked that $P < \frac{N_{samples}}{2}$.
Subsequently, the response for all cubature points is calculated. Each cubature point corresponds to a certain set of coefficients $(c_{1,random}, \cdots, c_{L,random})$ and an external code, the 'black box', computes the DVHs for the delineations constructed with the coefficients corresponding to each cubature point. The black box is a function which has as input the coefficients $(c_{1,random}, \cdots, c_{L,random})$, the eigenmodes $(\vec{v}_1, \cdots, \vec{v}_L)$ and the fixed dose distribution. Within the black box, first a surface shape vector $\vec{p}^{L}_{random}$ is constructed as described by Equation 2.13. Then the random delineation is projected onto 2D slices as described in Section 3.2.3. The random delineation is then stored as a DICOM file and the DVH of the random delineation for the fixed dose distribution is calculated. This calculation is done using the Matterhorn Python library from the Erasmus MC. The DVH calculated by the black box is then fed back to the openGPC code.
Using the responses obtained from the black box, the PC coefficients are calculated using either spectral projection or multi-dimensional polynomial regression and saved in a PCE-object, along with the basis vectors. The PCE can then be checked for accuracy.

3.4.2. PCE Validation

To determine which grid order and polynomial order were necessary to generate an accurate PCE to model the DVH as a function of the eigenmode coefficients, the exact DVH was compared to the DVH generated by the PCE in different validation scenarios. Furthermore, the difference in the $D_{98\%}$ (CTV) or the $D_{95\%}$ (PTV) generated by the PCE was compared to the exact value. This evaluation was done with two metrics. Firstly the mean absolute difference $\Delta D_{98\%,mean}$ (CTV) or $\Delta D_{95\%,mean}$ (PTV) between the PCE generated value and the exact value in all validation scenarios was determined. Secondly, the 90th percentile value of the absolute difference $\Delta D_{98\%,90}$ (CTV) or $\Delta D_{95\%,90}$ (PTV) for all validation scenarios was determined. Finally, the mean difference $\Delta D_{0\%,mean}$ in the maximum dose and the mean difference $\Delta D_{100\%,mean}$ in the minimum dose were calculated.
100 validation scenarios were chosen in which the coefficients $(c_{1,random}, \cdots, c_{L,random})$ were sampled from a Gaussian distribution with mean μ equal to 0 and standard deviation σ equal to $(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_L})$. However, if the absolute value of $c_{l,random}$ was larger than $2.5 \cdot \lambda_l$, the validation scenario was not used. This was to avoid the generation of unrealistic random delineations, as described in Section 3.2.3.

3.4.3. Characterizing Delineation Uncertainty Effects for a Fixed Dose Distribution

Once the right GL and PO had been chosen, PCE could be used to model the DVH of a delineated structure as a function of the eigenmode coefficients $(c_{1,random}, \cdots, c_{L,random})$ and thus characterize the effects of the delineation uncertainties. The effect of the delineation uncertainties was characterized in two ways.
Firstly, the under-dosage probability of the targets was investigated. This was done by creating 10,000 DVHs of a target using the PCE and determining in how many of these cases the dose was under the constraint dose of $D_{95\%} \geq 0.95 \cdot D^{\text{prescribed}}$ for the PTV and $D_{98\%} \geq 0.95 \cdot D^{\text{prescribed}}$ for the CTV.
Secondly, DVH distributions for the delineated volumes were created. The DVH distributions were created by again sampling 10,000 DVHs using the PCE and visualizing the confidence intervals in which the sampled DVHs lie.
For the above characterization methods the values of $c_{l,random}$ for each realization of the DVH were checked and if the absolute value of $c_{l,random}$ was larger than $2.5 \cdot \lambda_l$, the DVH was not used.

3.5. PCE Model for a Reoptimized Dose Distribution

PCE was also used to model the DVHs of delineated structures and the total dose distribution when a treatment plan was reoptimized to a random realization of a delineated target.

3.5.1. Construction of PCE

The PCE was constructed in a similar fashion to the method described in Section 3.4. However, now the input consisted of the grid level GL, the polynomial order PO and the polynomial type, the included eigenmodes $(\vec{v}_1, \cdots, \vec{v}_L)$ of the delineated structure, the mean μ and standard deviation σ of coefficients $(c_{1,random}, \cdots, c_{L,random})$ corresponding to the included eigenmodes, and the locations of the planning CT, the planning delineations and the wish-list.

Furthermore, before the responses of all the cubature points were calculated, first a dose mask was generated. To generate the dose mask, a number of worst case scenarios, where the random coefficients $(c_{1,random}, \cdots, c_{L,random})$ are large, are fed to the black box and the response is evaluated.

For this PCE model, the random coefficients $(c_{1,random}, \cdots, c_{L,random})$, the eigenmodes $(\vec{v}_1, \cdots, \vec{v}_L)$ and the locations of the planning CT, the planning delineations and the wish-list serve as input for the black box. Within the black box, first a random delineation is formed of the structure of interest using the random coefficients $(c_{1,random}, \cdots, c_{L,random})$. The random delineation is then saved along with the other delineations needed for planning. Then iCycle makes a treatment plan using the saved delineations, the planning CT and the wish-list. Once iCycle has made the treatment plan, the black box returns the dose distribution per voxel on the planning CT, the DVH of the random delineation made with coefficients $(c_{1,random}, \cdots, c_{L,random})$ and the DVHs of any other desired delineated ROIs.

If in any of the worst case scenarios the dose in a voxel is lower than a cut-off dose $D_{\text{cut-off}}$, the voxel is excluded from the PCE. $D_{\text{cut-off}}$ was set to 0.01 Gy. The dose mask is a binary vector which indicates which voxels are above $D_{\text{cut-off}}$ and thus included in the PCE.

3.5.2. PCE Validation

To validate the constructed PCE, 9 validation scenarios were used to determine the GL and PO necessary to generate an accurate PCE. The size of the coefficients was also limited to $2.5 \cdot \lambda_l$ to avoid the generation of unrealistic delineations.

To validate the PCE model, firstly the DVHs of the random delineations were compared to the exact DVHs, as also described in Section 3.4.2. In this case $\Delta D_{98\%}$ and $\Delta D_{2\%}$ were determined for every validation scenario.

Secondly, the total dose distribution on the CT in the validation scenarios was compared to the exact dose distribution using gamma evaluation. In gamma evaluation, a calculated dose distribution D_c and a reference dose distribution D_m are compared by considering the spatial displacement and the dose difference between the two dose distributions. For each reference point $\vec{r}_m$ in D_m, the dose/distance variable $\Gamma(\vec{r}_m, \vec{r}_c)$ is given by:

$$\Gamma(\vec{r}_m, \vec{r}_c) = \sqrt{\frac{|\vec{r}_m - \vec{r}_c|^2}{(\Delta r)^2} + \frac{|D_m(\vec{r}_m) - D_c(\vec{r}_c)|^2}{(\Delta D)^2}}, \tag{3.3}$$

where Δr is the distance-to-agreement and ΔD is the acceptable dose deviation. For each point $\vec{r}_m$ in the reference dose distribution the value of $\gamma(\vec{r}_m)$ is determined, which is given by:

$$\gamma(\vec{r}_m) = \min\{\Gamma(\vec{r}_m, \vec{r}_c)\}, \quad \forall\{\vec{r}_c\} \tag{3.4}$$

If the value of $\gamma(\vec{r}_m)$ is smaller than 1, the point $\vec{r}_m$ is accepted and if $\gamma(\vec{r}_m)$ is larger than 1, the point $\gamma(\vec{r}_m)$ is rejected.

For each validation scenario, $\gamma(\vec{r}_m)$ was calculated for every voxel in the exact dose distribution with respect to the dose distribution obtained with the PCE model. In this thesis Δr was set to 1 mm and ΔD was set to 0.1 Gy.

Furthermore, the dose distribution generated by the PCE and the exact dose distribution were compared voxel-wise. This was done by calculating the ΔD_2 of the dose difference between the exact dose and the PCE generated dose. ΔD_2 is the minimum dose difference of the 2% of voxels which have the largest dose difference.

3.5.3. Characterizing Delineation Uncertainty Effects for a Reoptimized Dose Distribution

Once the right GL and PO had been chosen, the PCE was used to model the DVH of the CTV in the same way as described in Section 3.4.3. For the hepatocellular carcinoma patient a DVH distribution of the healthy liver was also made and the the NTCP of the healthy liver was calculated in all scenarios. The healthy liver was defined as the liver delineation excluding the CTV delineation. The NTCP was calculated with the Luman Kutcher Burman model, as desribed in Section 2.4.2. The NTCP parameters were based on the model obtained by *Dawson et al.*, which describes the dose-volume tolerance for radiation-induced liver disease [9]. The volume effect n equalled 0.9700, the tolerance dose TD_{50} equalled 39.8 Gy and the fit parameter m equalled 0.12.

4

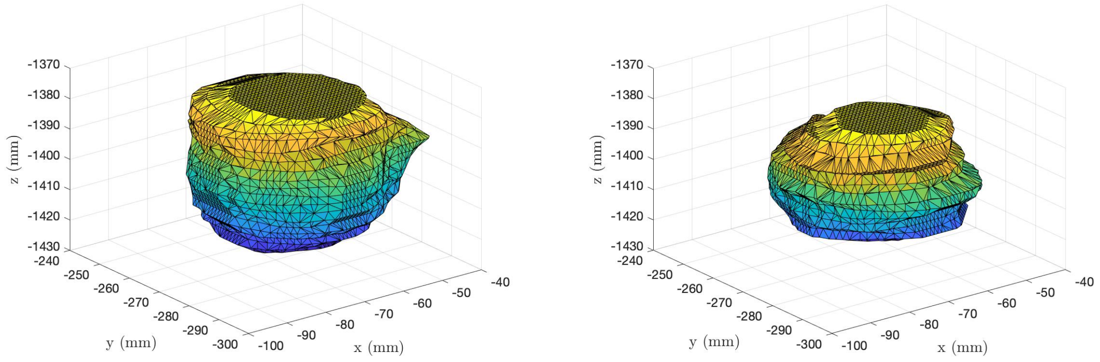
Results

In this chapter the results of the research are presented. First the results of the delineation quantification and characterization for both the manual delineations and auto-contours are shown in Section 4.1. Thereafter the use of PCE as a meta model for the DVH for a fixed dose distribution and both the DVH and the dose distribution for a reoptimized dose distribution is validated in Section 4.2. Finally, the dosimetric effects of delineation uncertainties for a fixed dose distribution (Section 4.3) and a reoptimized dose distribution (Section 4.4) are shown.

4.1. Delineation Uncertainty Quantification and Characterization

4.1.1. Manual Delineations

The 3D delineations of two out of the 12 manual delineations of the GTV in the liver are displayed in Figure 4.1. The probability map for all manual delineations of the GTV on two different CT slices can be seen in Figure 4.2. The gold-standard delineation was obtained by thresholding the probability map at 0.5. The gold-standard delineation is displayed in Figure 4.3.



(a) Delineation 1

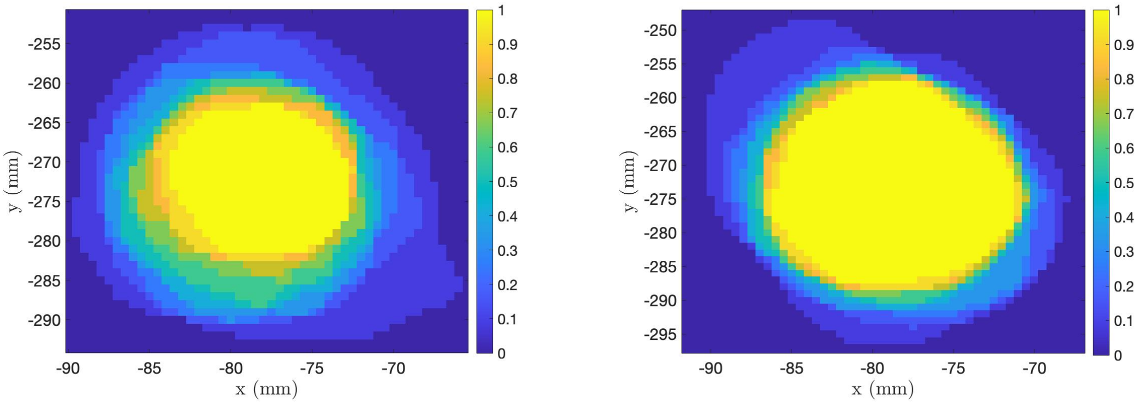
(b) Delineation 2

Figure 4.1: Two of the 12 manual delineations of the GTV by the radiation oncologists.

The DSCs of all the delineations with respect to the gold-standard delineation are displayed in Table 4.1. The mean DSC is 0.8780 with a standard deviation of 0.0542. The ratio between the largest delineated volume and the smallest delineated volume equals 1.8088. Next PCA was performed on the

Table 4.1: The DSCs for all manual delineations of the GTV with respect to the gold-standard delineation.

Delineation number	1	2	3	4	5	6	7	8	9	10	11	12
DSC	0.7475	0.8364	0.8541	0.8902	0.9083	0.9155	0.8921	0.8749	0.8385	0.9352	0.8925	0.9507



(a) Probability map at $z = -1387.5$ mm (b) Probability map at $z = -1411.5$ mm

Figure 4.2: Probability map of the manual delineations of the GTV at $z = -1387.5$ mm and $z = -1411.5$ mm.

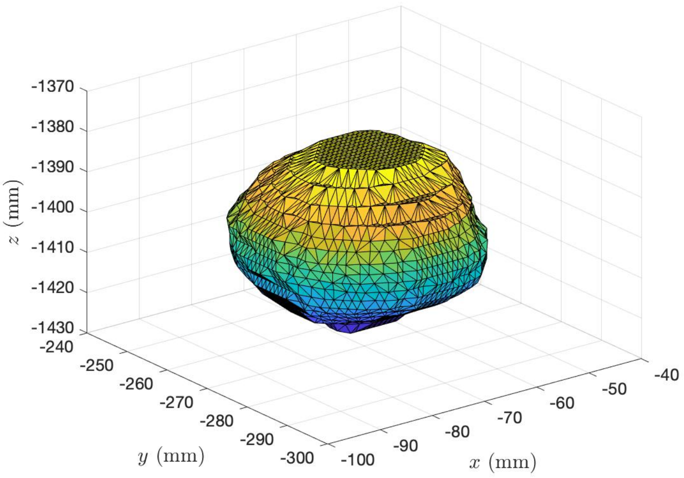


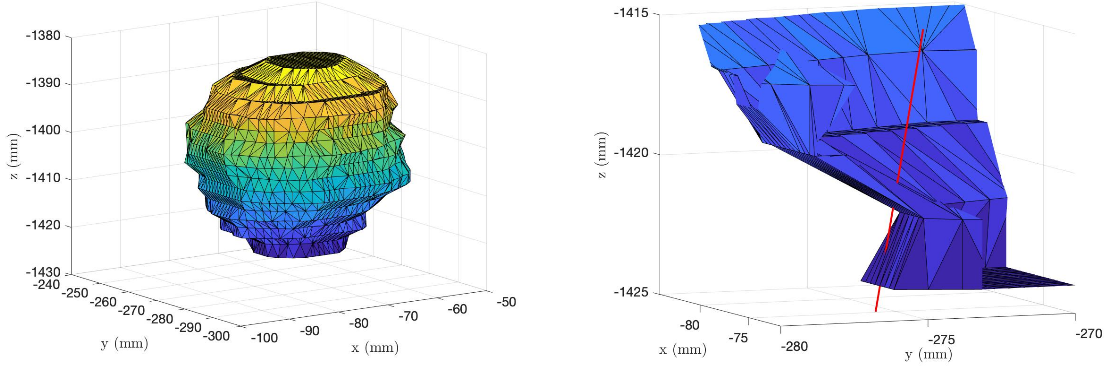
Figure 4.3: Gold-standard delineation of the manual delineations of the GTV.

set of delineations, as described in Section 3.2. First the surface shape vectors for each delineation were determined. The total length of each surface shape vector $\vec{p}_i$ was 650161. In Table 4.2 the number of multiple intersections in the determination of the surface shape vectors for each delineation is shown. The average number of multiple intersections equals 818 per delineation and the maximum is 5035, which corresponds to 0.13% and 0.77% of the number of elements in each surface shape vector, respectively. The time to build the surface shape vectors for all delineations was approximately 3 hours. The delineation with 5035 multiple intersections and an example of a multiple intersection with this delineation are shown in Figure 4.4. The high number of multiple intersections is partially due to a concavity at the bottom of this delineation.

Table 4.2: The number of multiple intersections in the determination of the surface shape vectors $\vec{p}_i$ for the manual GTV delineations.

Delineation number	1	2	3	4	5	6	7	8	9	10	11	12
Number of multiple intersections	652	165	423	308	298	5035	654	1704	144	224	175	32

The eigenmodes of the set of delineations were determined using PCA. In Figure 4.5a, the fraction of the variance of the data that each eigenmode explains is shown. As can be seen, most of the variance of the data is explained by a small number eigenmodes, which means that the differences between

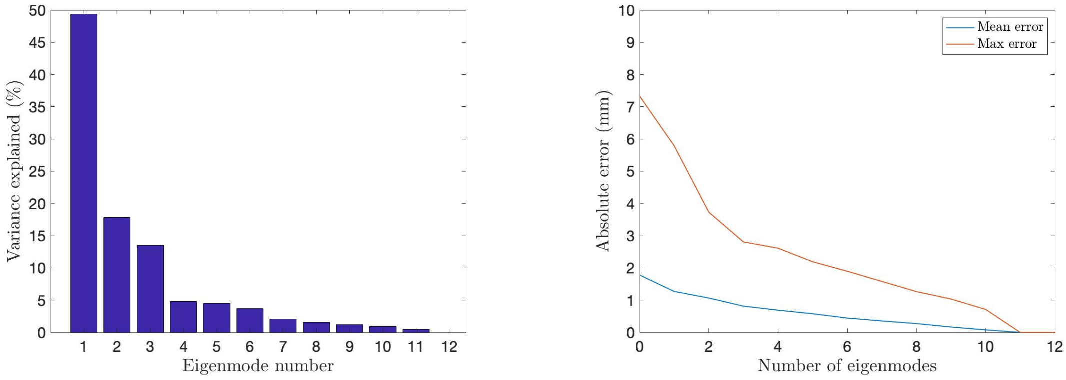


(a) Delineation 6

(b) Multiple intersections of line with the surface of the delineation

Figure 4.4: (a): Manual delineation 6 of the GTV. (b): An example of the line originating in the center of mass of the gold-standard GTV delineation and intersecting the surface of delineation 6 multiple times.

the different delineations are concentrated in a small number of correlated deformations. The first 5 eigenmodes explain 90.0 % of the variance. In Figure 4.6, a visualization of the effect of the two most important eigenmodes on the delineation made with the mean surface shape vector $\vec{\mu}$ (Equation 2.3) is shown. The delineation made with the mean surface shape vector $\vec{\mu}$ will be referred to as the mean delineation in the rest of this thesis. It can be seen that the eigenmodes deform the mean delineation by several millimeters.



(a)

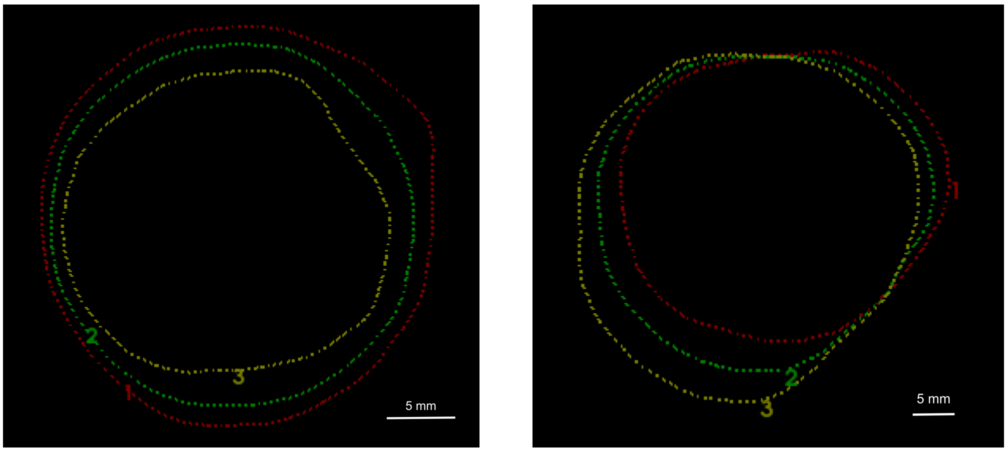
(b)

Figure 4.5: (a): Fraction of the variance of the manual GTV delineation data that each eigenmode explains. (b): The mean error along with the maximum reconstruction error for the manual delineations as a function of the number of eigenmodes.

Figure 4.5b shows the mean and maximum reconstruction accuracy of the original delineation geometries as a function of the number of dominant eigenmodes used in the reconstruction across all delineations, as described in Equation 3.1. As expected, the maximum error and mean error decrease as more eigenmodes are used to reconstruct the data and the mean and maximum error go to 0 mm when all 12 eigenmodes are used. When using 4 eigenmodes the mean reconstruction error is already less than 1 mm. It was thus decided to use 4 eigenmodes to form the random delineations of the GTV in the PCE model.

4.1.2. Auto-Contours

The delineation uncertainty of the auto-contouring software MIM was quantified for both the brainstem and the high dose CTV delineations. First the delineation uncertainty of the high dose CTV is quantified



(a) The effect of eigenmode $l = 1$ on the mean delineation at $z = -1389$ mm

(b) The effect of eigenmode $l = 2$ on the mean delineation at $z = -1416.5$ mm

Figure 4.6: Visualization of the effect of the first (a) and second (b) eigenmode on the mean GTV delineation. Delineation 1 (red) is obtained from the surface shape vector $\vec{p} = \vec{\mu} - \sqrt{\lambda_l} \cdot \vec{v}_l$, delineation 2 (green) is obtained from $\vec{p} = \vec{\mu}$, and delineation 3 (yellow) is obtained from $\vec{p} = \vec{\mu} + \sqrt{\lambda_l} \cdot \vec{v}_l$

and characterized, after which the same is done for the brainstem.

High Dose CTV Delineations

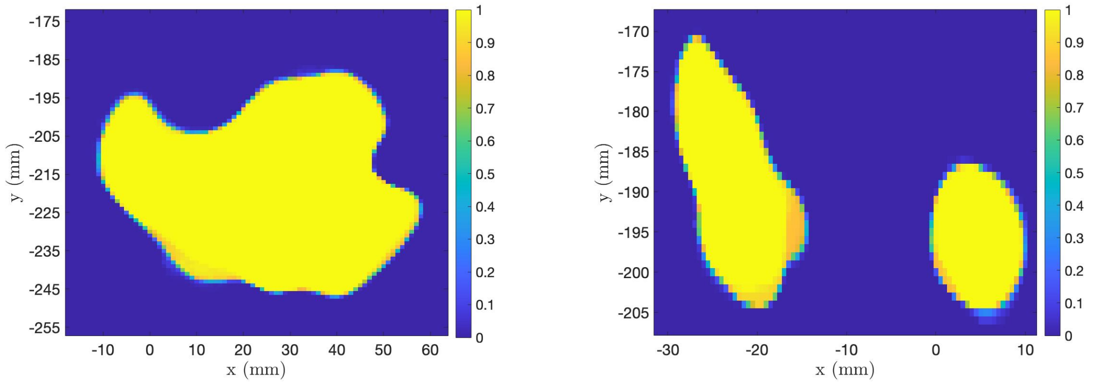
The probability map of the high dose CTV at three different heights is shown in Figure 4.7. It can be seen that the delineations can differ by a couple of millimeters at some spots. This is for instance visible on the probability map at $z = -490$ mm, where there is a part of the left region which is only covered by approximately 80% of the delineatinons. In Figure 4.8 the gold-standard delineation of the high dose CTV is shown. The mean DSC of all delineations with respect to the gold-standard is 0.9670 with a standard deviation of 0.0075. The ratio between the maximum delineated volume and the smallest delineated volume is 1.0245.

The high dose CTV consists of 3 separate regions, which means that 3 surface shape vectors were made per delineation, as described in Section 3.3.3. To reduce the computation time, the polar angle θ was increased with 1° increments when determining the surface shape vectors, opposed to the 0.1° increments for the manual delineations. This resulted in surface shape vectors with a length of $361 \times 181 = 65341$ elements. The computation time to construct all surface shape vectors for all delineations of the high dose CTV was approximately 10 hours. The average number of multiple intersections for the top region of the high dose CTV is 2261, which is 3.5% of the number of elements in the surface shape vector. The maximum number of multiple intersections equals 2825, which is 4.3% of the number of elements in the surface shape vector. For the lower region of the high dose CTV the mean number of multiple intersections is 1094 (1.7%), with a maximum of 1491 multiple intersections (2.3%). Finally, for the middle region the mean number of intersections equals 114 (0.2%), with a maximum of 440 (0.7%). The high number of multiple intersections in the top part of the high dose CTV is due to its concave shape. In Figure 4.9 one delineation of the top part of the high dose CTV is shown along with a line originating in the center of mass of the gold-standard and crossing the surface of this delineation multiple times.

Because of the high number of multiple intersections in the high dose CT, the shape of the individual delineations could not be properly established using surface shape vectors. Therefore it was decided not to perform PCA on the high dose CTV delineations, as this would not give an accurate characterization of the variation between the delineations.

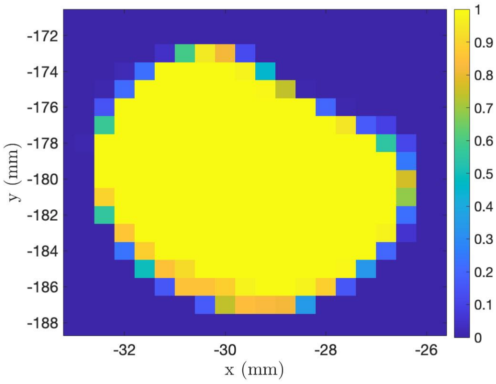
Brainstem Delineations

In Figure 4.10 the probability map of the brainstem delineations is shown on two CT slices. It can be seen that the variation among the delineations is minimal, as the largest part of the delineations totally overlap and there are only small variations of around 1 millimeter with respect to the region where all



(a) Probability map at $z = -430$ mm.

(b) Probability map at $z = -490$ mm.



(c) Probability map at $z = -540$ mm.

Figure 4.7: Probability map of the delineations of the high dose CTV at $z = -430$ mm, $z = -490$ mm and $z = -540$ mm.

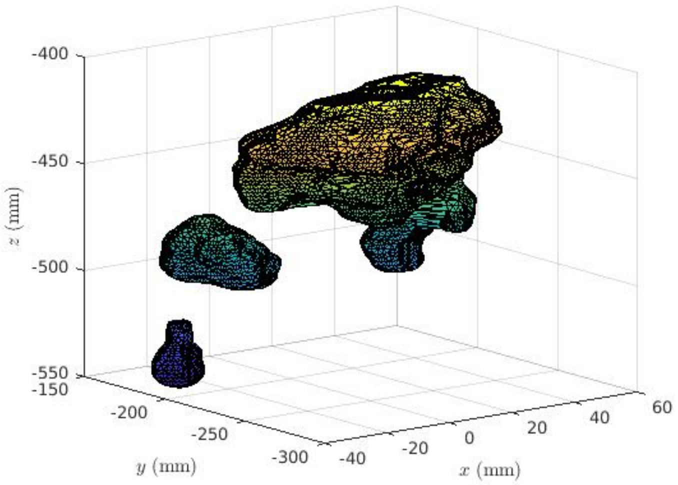
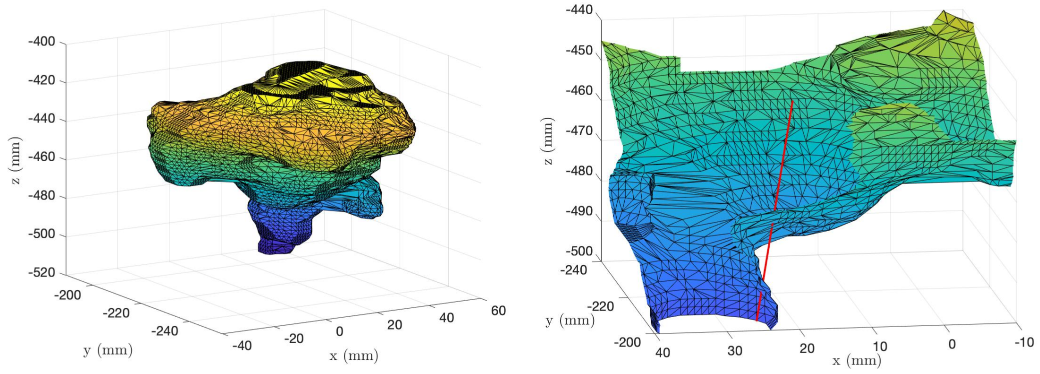


Figure 4.8: Gold-standard delineation of the auto-contours of the high dose CTV.

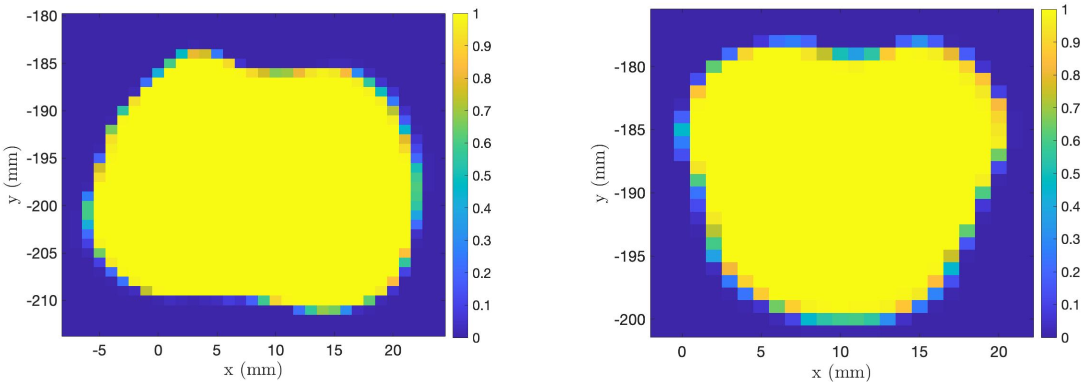
delineations overlap. The gold-standard delineation obtained from thresholding the probability map at 0.5 can be seen in Figure 4.11. The mean DSC of the delineations with respect to the gold-standard delineation is 0.9680 with a standard deviation of 0.0071. Furthermore, the ratio of the maximum delineated volume to the minimum delineated volume equals 1.0784.



(a) Delineation 20

(b) Multiple intersections of line with the surface of the delineation

Figure 4.9: (a): The top region of delineation 20 of the high dose CTV. (b): An example of the line originating in the center of mass of the gold-standard delineation and crossing the surface of delineation 20 multiple times.



(a) Probability map at $z = -408$ mm

(b) Probability map at $z = -428$ mm

Figure 4.10: Probability map of the delineations of the brainstem at $z = -408$ mm and $z = -428$ mm.

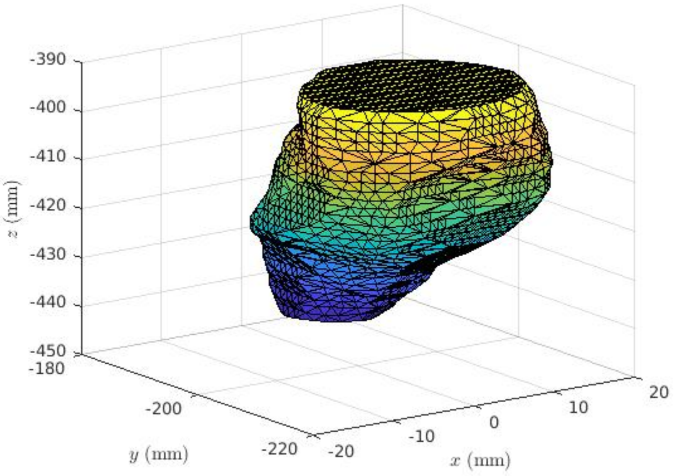
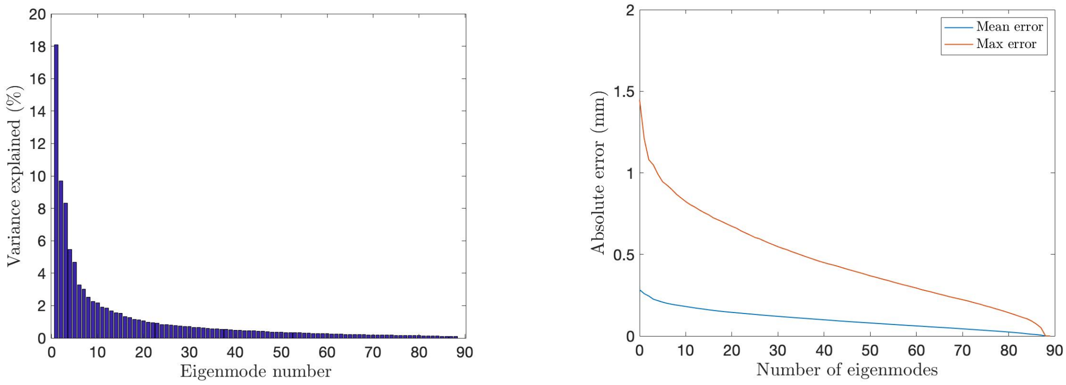


Figure 4.11: Gold-standard delineation of the auto-contours of the brainstem.

The time to compute all surface shape vectors was approximately 3 hours. The average number of multiple intersections per delineation is 84 and the maximum number equals 480, which correspond to 0.13% and 0.73% of the total number of elements in the surface shape vectors, respectively.
From the surface shape vectors the eigenmodes were determined using PCA. Because the surface shape vectors of the high dose CTV could not be determined accurately, PCA was only performed on the surface shape matrix of the brainstem, and not on the combined surface shape matrix which is shown in Equation 3.2.
The fraction of the variance explained by each eigenmode is shown in Figure 4.12a. The variance of the data is not dominated by a small number of eigenmodes as 45 eigenmodes are needed to explain 90.2% of the variance of the data. The effect of the two most dominant eigenmodes on the mean delineation is shown in Figure 4.13. It can be seen that these eigenmodes have very little effect on the mean delineation.
In Figure 4.12b the reconstruction error as a function of the number of dominant eigenmodes is shown. When 0 eigenmodes are used for the reconstruction, which means that the error is being calculated with respect to the mean delineation, the mean error is already below 0.5 mm and the maximum error is below 1.5 mm. As the number of eigenmodes increases, the mean error and maximum error tend towards 0 mm.



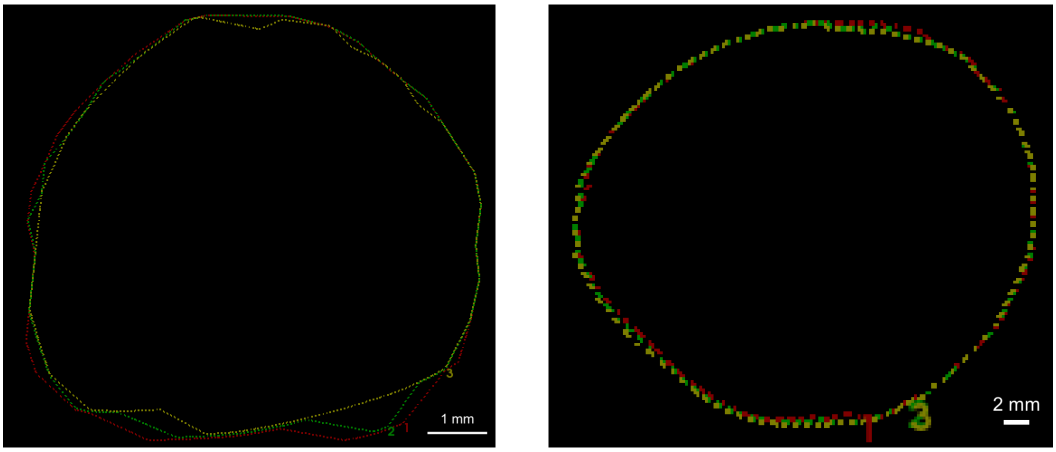
(a)

(b)

Figure 4.12: (a): Fraction of the variance of the brainstem delineation data that each eigenmode explains.
(b): The mean error along with the maximum reconstruction error for the brainstem delineations as a function of the number of eigenmodes.

4.2. PCE Validation

In this section PCE is validated as a meta-model for two responses as a function of the random eigenmode coefficients.
First the use of the PCE as a model for the DVH of a CTV and PTV for a fixed dose distribution is validated. The PCE model was built to model the DVH of the CTV and PTV of the hepatocellular carcinoma patient using 4 eigenmodes, and thus resulting in the uncertain input vector $\vec{\xi} = (c_{1,random}, c_{2,random}, c_{3,random}, c_{4,random})$. The PCE model was made for both the VMAT plan and the IMPT plan. It was decided not to build a PCE model for the head and neck patient with the auto-contours because the number of eigenmodes needed to explain the variations between the brainstem auto-contours was too large to build a PCE model with. Building a PCE model with 45 eigenmodes would require a large amount of computation time and memory and would not be feasible. The fixed dose distribution for the IMPT plan with which the PCE was built is shown in Figure 4.14. This dose distribution was obtained by making a treatment plan with iCycle with the wish-list given in Table 3.1. The CTV delineation used in planning was equal to the gold-standard GTV delineation determined in Section 4.1.1.



(a) The effect of eigenmode $l = 1$ on the mean delineation at $z = -446$ mm

(b) The effect of eigenmode $l = 2$ on the mean delineation at $z = -390$ mm

Figure 4.13: Visualization of the effect of the first (a) and second (b) eigenmode on the mean brainstem delineation. Delineation 1 (red) is obtained from the surface shape vector $\vec{p} = \vec{\mu} - \sqrt{\lambda_l} \cdot \vec{v}_l$, delineation 2 (green) is obtained from $\vec{p} = \vec{\mu}$, and delineation 3 (yellow) is obtained from $\vec{p} = \vec{\mu} + \sqrt{\lambda_l} \cdot \vec{v}_l$

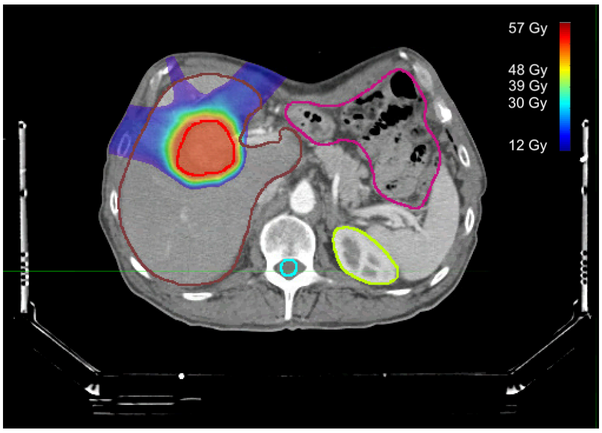


Figure 4.14: IMPT dose distribution on the gold-standard delineation of the CTV. The delineations of the CTV (red), liver (brown), bowel (purple), spinal cord (light blue) and left kidney (light green) are also shown.

The PCE is also validated as a model for the reoptimized dose distribution for different realizations of the CTV delineations for the hepatocellular carcinoma patient using 2 eigenmodes. The uncertain input vector was $\vec{\xi} = (c_{1,random}, c_{2,random})$. This model was only made for the IMPT plan.

4.2.1. PCE Model for a Fixed Dose Distribution

To model the effect of the delineation uncertainties for a fixed dose distribution with the PCE, first the grid level and polynomial order had to be established with which the PCE accurately recreated the exact DVHs. To do so, the DVHs created by the PCE were visually compared with the exact DVHs and the values of the $D_{98\%}$ (for the CTV delineations) and the $D_{95\%}$ (for the PTV delineations) obtained from the PCE model were compared to the exact values in 100 validation scenarios. The results of the comparison are shown in Table 4.3, along with the time to build the PCE with the given PCE settings. All these PCEs were built using spectral projection. Furthermore, in Table 4.4 the mean difference in the maximum dose $\Delta D_{0\%,mean}$ and the mean difference in the minimum dose $\Delta D_{100\%,mean}$ is shown. It can be seen that for the VMAT plan, $\Delta D_{95\%,mean}$ and $\Delta D_{98\%,mean}$ for the PTV and CTV, respectively, are relatively low when a grid level of 4 with an extra level is used in combination with a polynomial order of 5. However, for the IMPT plan the $\Delta D_{98\%,mean}$ remains high up to a grid level of 5 with an extra level and a polynomial order of 6.

As can be seen in Table 4.4, $\Delta D_{100\%,mean}$ and $\Delta D_{0\%,mean}$ are relatively large for all PCE settings for all structures. This is because the $D_{100\%}$ and $D_{0\%}$ are very sensitive parameters which only depend on the dose in one voxel and thus these parameters are hard for the PCE to model. However, because in this thesis only the $D_{95\%}$ and the $D_{98\%}$ are used to determine whether a delineation meets a planning constraint, the inaccuracies in the $D_{100\%}$ and $D_{0\%}$ are not deemed significant.

Table 4.3: The mean and 90th percentile of the difference across the 100 validation scenarios between the PCE calculated $D_{98\%}$ and $D_{95\%}$ and the exact values for various PCE settings. Here the number following GL is the grid level, the number following EL is the number of extra levels and the number following PO is the polynomial order. The computation time for the PCE construction with the given PCE settings is also given.

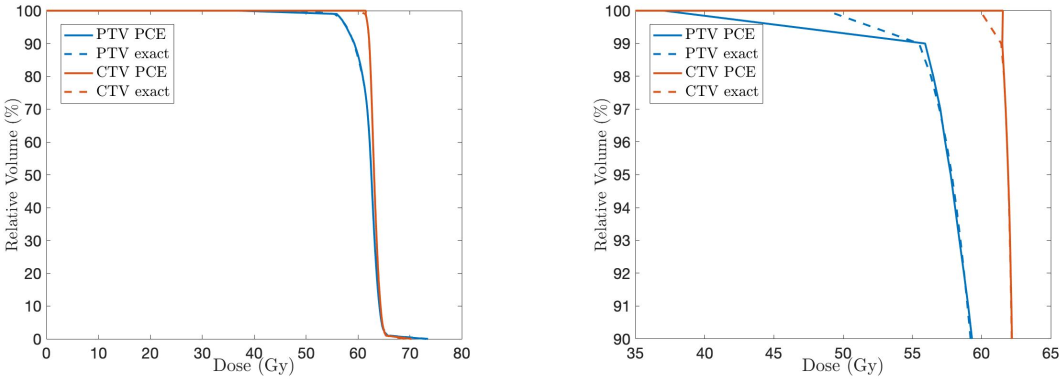
	CTV VMAT plan		PTV VMAT plan		CTV IMPT plan		Computation time
	$\Delta D_{98\%,mean}$	$\Delta D_{98\%,90}$	$\Delta D_{95\%,mean}$	$\Delta D_{95\%,90}$	$\Delta D_{98\%,mean}$	$\Delta D_{98\%,90}$	
GL3EL1PO4	0.23 Gy	0.61 Gy	0.20 Gy	0.35 Gy	0.64 Gy	1.31 Gy	10 min
GL4EL1PO5	0.09 Gy	0.20 Gy	0.13 Gy	0.28 Gy	0.39 Gy	0.95 Gy	15 min
GL5EL1PO6	0.09 Gy	0.19 Gy	0.13 Gy	0.33 Gy	0.37 Gy	0.71 Gy	45 min

Table 4.4: The mean and 90th percentile difference across the 100 validation scenarios between the PCE calculated $D_{100\%}$ and $D_{0\%}$ and the exact values for various PCE settings. Here the number following GL is the grid level, the number following EL is the number of extra levels and the number following PO is the polynomial order.

	CTV VMAT plan		PTV VMAT plan		CTV IMPT plan	
	$\Delta D_{100\%,mean}$	$\Delta D_{0\%,mean}$	$\Delta D_{100\%,mean}$	$\Delta D_{0\%,mean}$	$\Delta D_{100\%,mean}$	$\Delta D_{0\%,mean}$
GL3EL1PO4	1.13 Gy	1.19 Gy	6.09 Gy	0.61 Gy	3.12 Gy	0.08 Gy
GL4EL1PO5	1.16 Gy	1.20 Gy	5.68 Gy	0.80 Gy	1.69 Gy	0.15 Gy
GL5EL1PO6	1.17 Gy	1.49 Gy	6.20 Gy	0.75 Gy	2.18 Gy	0.25 Gy

For the CTV and PTV in the VMAT plan, the PCE model built with 4 grid levels, 1 extra level and a polynomial order of 5 was chosen to model the DVH of these structures on a fixed dose distribution. In the construction of the PCE there were 209 function evaluations, including 8 function evaluations because of the extra level. This PCE consisted of 73 PC basis vectors after a hyperbolic trim with $q = 0.8614$.

In Figure 4.15 the DVHs of the mean delineations of the CTV and PTV, where all coefficients $c_{l,random}$ equal 0, for the fixed dose distribution are shown. As can be seen, the DVHs overlap very well, except at 100% relative volume. This is due to the inaccuracy of the PCE determining the $D_{100\%}$, as discussed earlier.

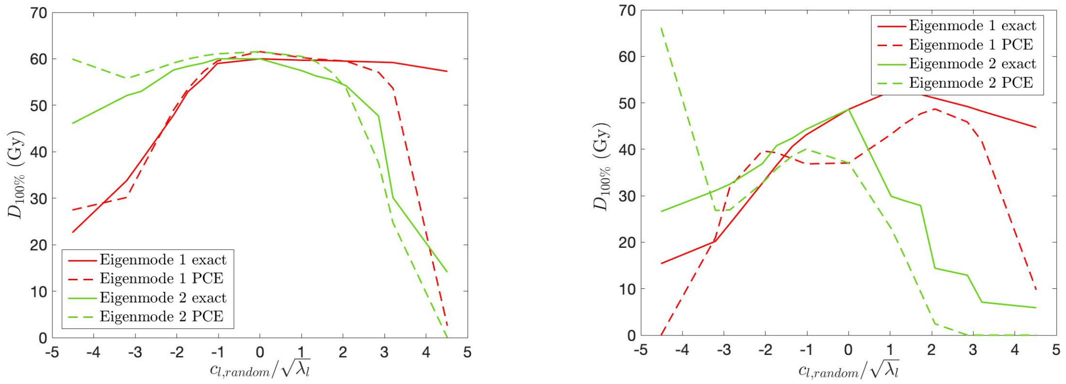


(a) An overview of the whole dose range of the DVH.

(b) A zoomed in view of the DVH.

Figure 4.15: The exact DVH and the PCE generated DVH of the mean CTV and PTV delineations in the fixed VMAT plan. The DVHs overlap everywhere, except at 100% relative volume. This is due to an inaccuracy of the PCE in determining the $D_{100\%}$.

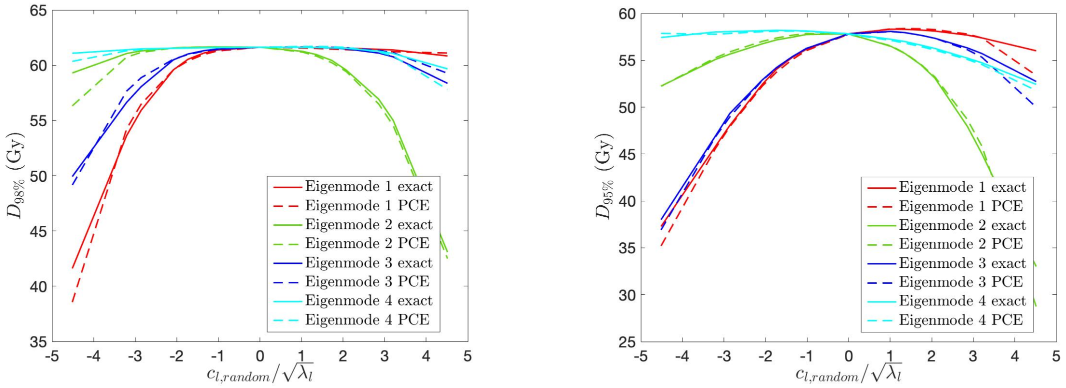
The inaccuracy of the PCE in modeling the $D_{100\%}$ is also shown when looking at the dependence of the $D_{100\%}$ on the individual eigenmodes. In Figure 4.16 the $D_{100\%}$ of the CTV and PTV in the VMAT plan is plotted as a function of the random coefficients $c_{l,random}$ for eigenmode 1 and eigenmode 2. The exact $D_{100\%}$ clearly does not overlap with the $D_{100\%}$ calculated by the PCE. This is because the exact $D_{100\%}$ does not change smoothly when the random coefficients change and there are sudden big jumps in dose when the random coefficients change. This is hard for the PCE to reconstruct. However, when looking at the dependence of both the $D_{98\%}$ for the CTV and the $D_{95\%}$ for the PTV on the random coefficients (Figure 4.17), it can be seen that the exact $D_{98\%}$ and $D_{95\%}$ overlap with the values calculated by the PCE in the region where realistic delineations are made ($-2.5 \leq \left(c_{l,random}/\sqrt{\lambda_l}\right) \leq 2.5$).



(a) The dependence of the $D_{100\%}$ of the CTV on the random coefficients of the first two eigenmodes.

(b) The dependence of the $D_{100\%}$ of the PTV on the random coefficients of the first two eigenmodes.

Figure 4.16: The dependence of the $D_{100\%}$ of the CTV and PTV on the first 2 eigenmodes in the fixed VMAT plan calculated by the PCE along with the exact dependence of $D_{100\%}$ on the first 2 eigenmodes.



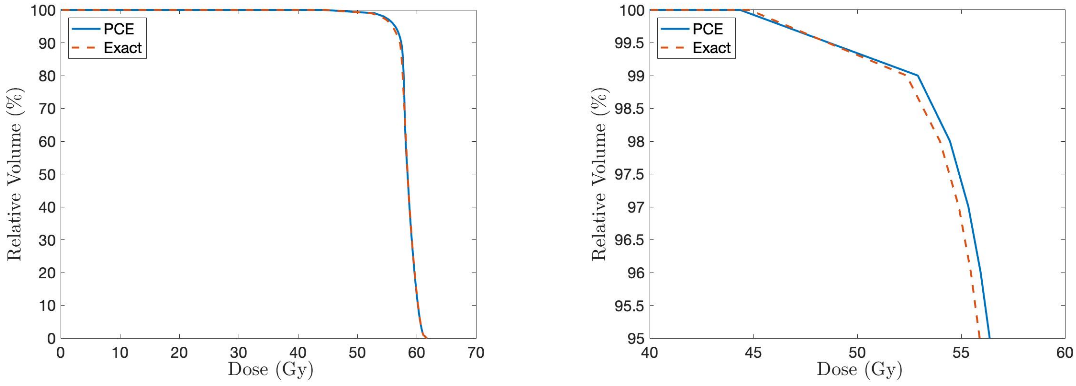
(a) The dependence of the $D_{98\%}$ of the CTV on the random coefficients of all eigenmodes.

(b) The dependence of the $D_{95\%}$ of the PTV on the random coefficients of all eigenmodes.

Figure 4.17: The dependence of the $D_{98\%}$ of the CTV and the $D_{95\%}$ of PTV on all eigenmodes in the VMAT plan calculated by the PCE along with the exact dependence of $D_{98\%}$ and $D_{95\%}$ on all eigenmodes.

For the CTV in the IMPT plan, the PCE models shown in Table 4.3 were deemed too inaccurate. Therefore, a PCE model was built using a grid level of 6, 1 extra level and a polynomial order of 7. Furthermore, this PCE model was built using regression, as opposed to spectral projection. In the construction of the PCE there were 2009 function evaluations, including 8 function evaluations because of the extra level. The PCE consisted of 330 PC basis vectors and there was no hyperbolic trim applied. The construction time of this PCE was approximately 2.5 hours.

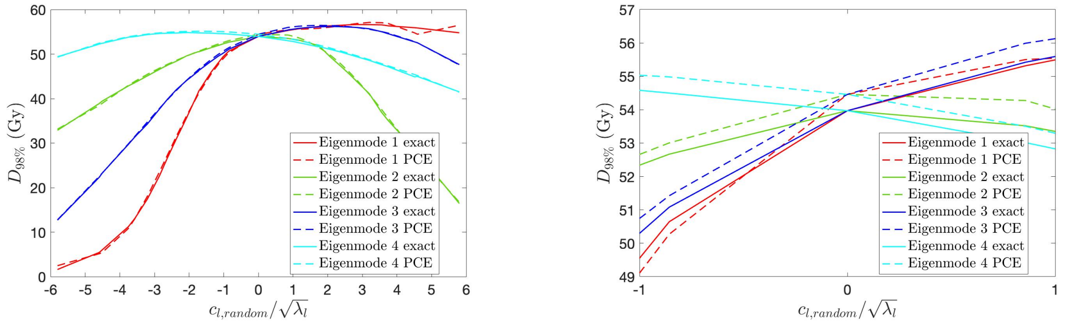
In the 100 validation scenarios, $\Delta D_{98\%,mean}$ was 0.34 Gy and $\Delta D_{98\%,90}$ was 0.61 Gy. In Figure 4.18 the DVH of the mean CTV is shown. The dependence of the $D_{98\%}$ on the individual random coefficients for each eigenmode is shown in Figure 4.19. The PCE calculated values of the $D_{98\%}$ overlap well with the exact $D_{98\%}$ values in the region where realistic delineations are made ($-2.5 \leq \left(c_{l,random}/\sqrt{\lambda_l}\right) \leq 2.5$), but there is some discrepancy between the curves near $c_{l,random}/\sqrt{\lambda_l} = 0$, as can be seen in Figure 4.19b.



(a) An overview of the whole dose range of the DVH.

(b) A zoomed in view of the DVH.

Figure 4.18: The exact DVH and the PCE generated DVH of the mean CTV delineation in the fixed IMPT plan. The DVHs overlap well, but there is some difference between the curves at the dose fall off point.



(a) An overview of the whole range of $c_{l,random}$ used in building the PCE.

(b) A zoomed in view with $-1 \leq (c_{l,random}/\lambda_l) \leq 1$.

Figure 4.19: The dependence of the $D_{98\%}$ of the CTV on all eigenmodes in fixed IMPT plan calculated by the PCE along with the exact dependence of $D_{98\%}$ on all eigenmodes.

4.2.2. PCE Model for a Reoptimized Dose Distribution

To validate the use of the PCE as a meta-model for the reoptimized dose distribution for different realizations of the CTV for the IMPT plan and find the desired grid level and polynomial order, a gamma evaluation between the exact dose distribution and the PCE generated dose distribution was performed for 9 different validation scenarios, as described in Section 3.5.2. Furthermore, the ΔD_2 was determined between the exact and PCE generated dose distribution in the validation scenarios. The percentage of voxels that passed the gamma evaluation and the ΔD_2 for each validation scenario for various PCE settings are shown in Table 4.5. All these PCEs were made using spectral projection. It can be seen that the number of accepted voxels is rather low for all PCE settings and the ΔD_2 is high. The percentage of accepted voxels would ideally be close to 100%. The computation time for the PCE built with settings GL2EL1PO3 was approximately 24 hours and for the settings GL3EL1PO4 it was approximately 140 hours. The long computation times are caused by the time that it takes iCycle to create a treatment plan, which is around 1.5 hours. The treatment plans were made sequentially, which meant that only one treatment plan was made at a time. Furthermore, the computation time was increased due to other jobs which were being performed on the server. The absolute difference between the exact $D_{98\%}$ and the PCE calculated $D_{98\%}$ for varying PCE setting is shown in Table 4.6, along with the absolute difference in the $D_{2\%}$. It can be seen that $\Delta D_{98\%}$ is always very low, but $\Delta D_{2\%}$ is quite high for all validation scenarios.

This is probably because there is only a constraint on the minimum dose delivered to the CTV, so the $D_{98\%}$ stays near the same value across all validation scenarios. However, there is no constraint on the maximum dose delivered to the CTV, so the $D_{2\%}$ can vary more between the different validation scenarios.

It was chosen to construct the PCE with a grid level of 3, 1 extra level and a polynomial order of 4. The number of function evaluations was 21, including 4 function evaluations due to the extra level. The PCE consisted of 12 basis vectors after a hyperbolic trim with $q = 0.995$.

Table 4.5: The percentage of accepted voxels and the ΔD_2 in the reoptimized IMPT plans in all the validation scenarios for various PCE settings. The number following GL is the grid level, the number following EL is the number of extra levels and the number following PO is the polynomial order. In the validation scenario column, the scenario is given by $(c_{1,random}, c_{2,random})$.

	GL2EL1PO3		GL3EL1PO4	
Validation scenario	Accepted voxels	ΔD_2	Accepted voxels	ΔD_2
(0,0)	77.51%	2.44 Gy	76.80%	3.07 Gy
$(1 \cdot \sqrt{\lambda_1}, 0)$	68.39%	4.70 Gy	68.80%	6.32 Gy
$(0, 1 \cdot \sqrt{\lambda_2})$	77.29%	3.38 Gy	69.94%	6.19 Gy
$(2 \cdot \sqrt{\lambda_1}, 0)$	70.69%	4.45 Gy	66.46%	5.83 Gy
$(0, 2 \cdot \sqrt{\lambda_2})$	72.34%	4.46 Gy	62.05%	6.72 Gy
$(1 \cdot \sqrt{\lambda_1}, 1 \cdot \sqrt{\lambda_2})$	73.57%	4.12 Gy	66.26%	7.27 Gy
$(1 \cdot \sqrt{\lambda_1}, 2 \cdot \sqrt{\lambda_2})$	66.81%	6.05 Gy	68.54%	5.30 Gy
$(2 \cdot \sqrt{\lambda_1}, 1 \cdot \sqrt{\lambda_2})$	68.34%	5.79 Gy	68.20%	5.11 Gy
$(2 \cdot \sqrt{\lambda_1}, 2 \cdot \sqrt{\lambda_2})$	61.19%	6.31 Gy	68.15%	4.56 Gy

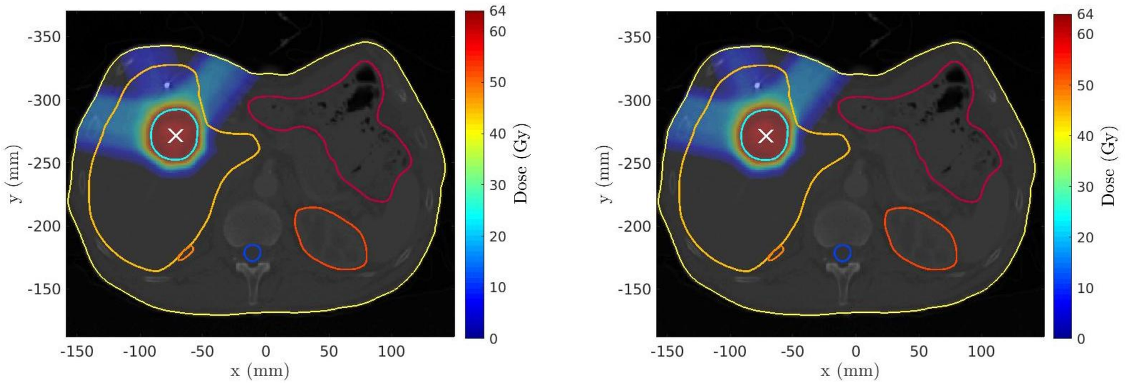
Table 4.6: The difference between the exact and PCE calculated $D_{98\%}$ and $D_{2\%}$ of the CTV in the reoptimized IMPT plans in all validation scenarios for various PCE settings.The number following GL is the grid level, the number following EL is the number of extra levels and the number following PO is the polynomial order. In the validation scenario column, the scenario is given by $(c_{1,random}, c_{2,random})$.

	GL2EL1PO3		GL3EL1PO4	
Validation scenario	$\Delta D_{98\%}$	$\Delta D_{2\%}$	$\Delta D_{98\%}$	$\Delta D_{2\%}$
(0,0)	0.4522 Gy	0.2966 Gy	0.4283 Gy	0.0348 Gy
$(1 \cdot \sqrt{\lambda_1}, 0)$	0.0776 Gy	0.4675 Gy	0.0046 Gy	0.6680 Gy
$(0, 1 \cdot \sqrt{\lambda_2})$	0.0269 Gy	0.0432 Gy	0.0963 Gy	0.6329 Gy
$(2 \cdot \sqrt{\lambda_1}, 0)$	0.2272 Gy	0.1350 Gy	0.1319 Gy	0.0506 Gy
$(0, 2 \cdot \sqrt{\lambda_2})$	0.0382 Gy	0.1879 Gy	0.0021 Gy	0.4857 Gy
$(1 \cdot \sqrt{\lambda_1}, 1 \cdot \sqrt{\lambda_2})$	0.0736 Gy	0.4974 Gy	0.0463 Gy	0.9474 Gy
$(1 \cdot \sqrt{\lambda_1}, 2 \cdot \sqrt{\lambda_2})$	0.0809 Gy	0.8471 Gy	0.0137 Gy	0.6094 Gy
$(2 \cdot \sqrt{\lambda_1}, 1 \cdot \sqrt{\lambda_2})$	0.0295 Gy	0.0196 Gy	0.0597 Gy	0.0640 Gy
$(2 \cdot \sqrt{\lambda_1}, 2 \cdot \sqrt{\lambda_2})$	0.0342 Gy	0.9505 Gy	0.0031 Gy	0.0506 Gy

For the nominal case where $c_{1,random} = 0$ and $c_{2,random} = 0$, the exact dose distribution along with the dose distribution calculated with the PCE are shown in Figure 4.20. The two dose distributions look similar.

The DVHs of the CTV and liver in the exact nominal dose distribution and the dose distribution generated by the PCE are shown in Figure 4.21a. The DVHs based on the PCE were made in 2 ways. For the first approach, the DVH was calculated directly by the PCE. For the second approach, the DVH was based on the dose distribution generated by the PCE. Both approaches yield DVHs which overlap with each other in the nominal case. When comparing the exact DVHs with the DVH calculated with the PCE, the DVHs of the CTV overlap well. However, the exact DVH and the PCE calculated DVH of the liver do not overlap.

In Figure 4.21b the DVHs of the CTV and liver in the validation scenario where $c_{1,random} = 1 \cdot \sqrt{\lambda_1}$ and $c_{2,random} = 1 \cdot \sqrt{\lambda_2}$ are shown. Here it can be seen that the exact DVH and the PCE generated DVHs of the CTV do not overlap near the maximum dose in the CTV, but they do overlap until and a bit past

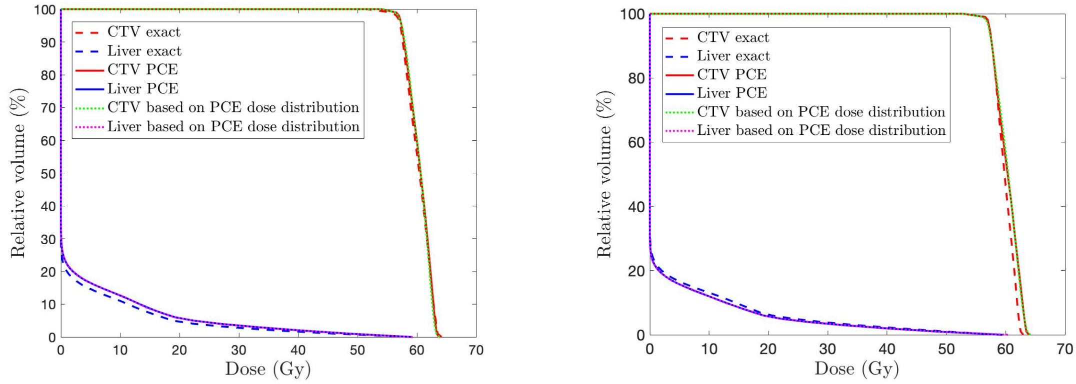


(a) Exact dose distribution

(b) PCE generated dose distribution

Figure 4.20: The exact dose distribution (a) and the dose distribution calculated by the PCE (b) in the nominal scenario in the reoptimized IMPT plan. The delineations of the CTV (light blue), liver (yellow), spinal cord (blue), kidneys (orange and red), and bowel (purple) are also shown.

the dose fall-off point.

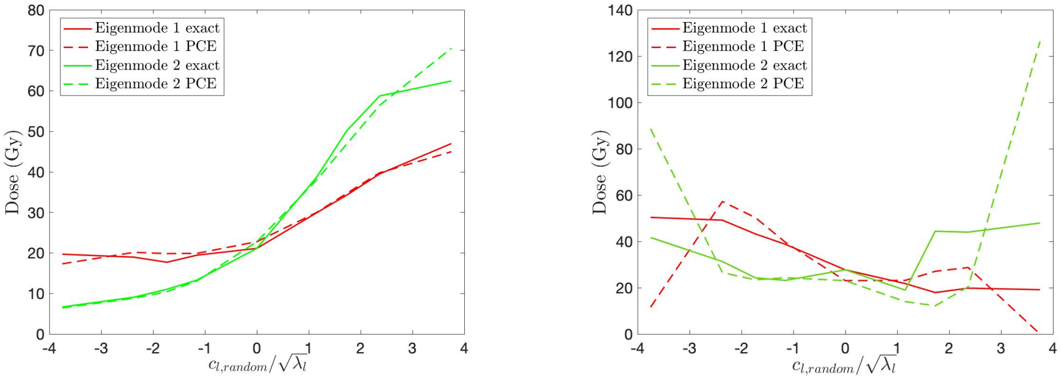


(a) Nominal scenario

(b) Validation scenario where $c_{1,random} = 1 \cdot \sqrt{\lambda_1}$ and $c_{2,random} = 1 \cdot \sqrt{\lambda_2}$

Figure 4.21: The exact DVH and the DVH calculated by the PCE in the nominal scenario and a validation scenario in the reoptimized IMPT plan. For the CTV and liver, two DVHs based on the PCE are plotted: one of a direct calculation of the DVH by the PCE, and one which is based on the PCE dose distribution.

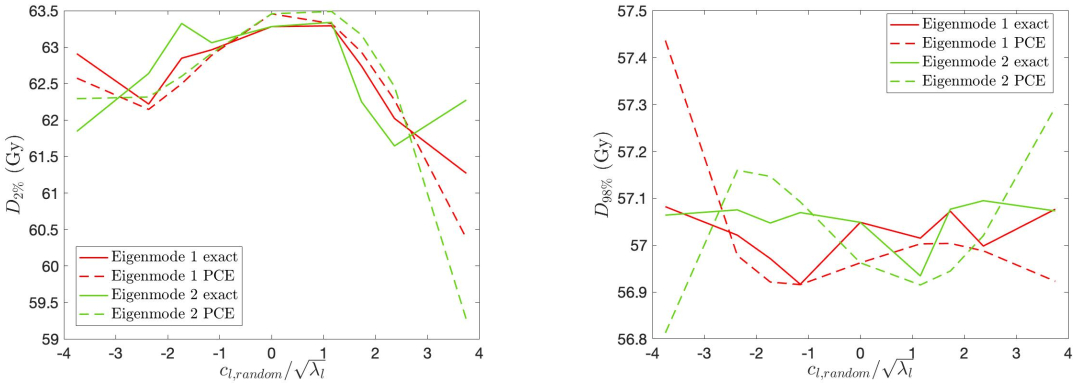
The dependence of the dose in the voxels close to the edge of the mean CTV delineation on the random coefficients is shown in Figure 4.22. When the random coefficients change, which corresponds to a change in shape of the CTV, the dose in these voxels can increase or decrease rapidly. In voxel 113847 the dose calculated with the PCE follows the exact dose well as a function of $c_{l,random}$. However, in voxel 203089 the dose calculated by the PCE does not exactly follow the exact dose, especially as a function of only $c_{2,random}$. This could be because the exact dose makes a big jump (from 20 Gy to 40 Gy) when $c_{2,random}$ goes from $1 \cdot \sqrt{\lambda_2}$ to $2 \cdot \sqrt{\lambda_2}$. If a higher polynomial order is used to build the PCE, the dose calculated by the PCE could possibly be able to follow this sudden jump better.
The dependence of both the $D_{98\%}$ and the $D_{2\%}$ of the CTV on the random coefficients is shown in Figure 4.23. As expected, the $D_{98\%}$ does not vary much for different realizations of the CTV delineation. The exact $D_{2\%}$ and $D_{98\%}$ do not seem to vary smoothly as a function of the random coefficients, so it is hard for the PCE to model these parameters. However, a PCE with a higher polynomial order could possibly replicate the $D_{2\%}$ and $D_{98\%}$ better.



(a) Voxel 113847

(b) Voxel 203089

Figure 4.22: The dependence of the dose in two voxels on the random coefficients for both the exact dose distribution and the dose distribution generated by the PCE in the reoptimized IMPT plans. The voxels are situated near the edge of the mean CTV delineation.



(a) $D_{2\%}$

(b) $D_{98\%}$

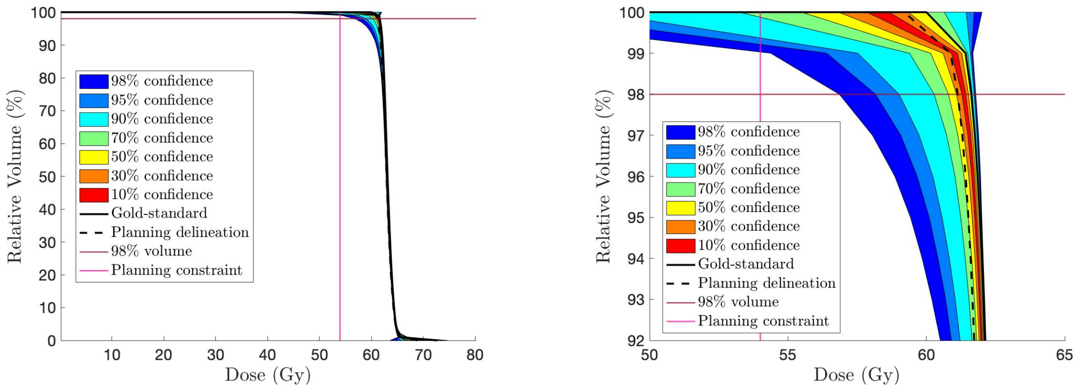
Figure 4.23: The dependence of both the exact and PCE calculated $D_{2\%}$ and $D_{98\%}$ of the CTV on the random coefficients in the reoptimized IMPT plans.

4.3. Dosimetric Effect of Delineation Uncertainties for a Fixed Dose Distribution

4.3.1. VMAT Plan

For the fixed VMAT plan, the DVH distribution of 10,000 randomly generated delineations of the CTV is shown in Figure 4.24. A histogram of the corresponding $D_{98\%}$ values is shown in Figure 4.25. Only 0.2% of the delineations do not meet the dose constraint ($D_{98\%} \geq 54$ Gy). This signifies that the 5 mm margin between the CTV and the PTV is enough to account for the delineation uncertainties in the CTV. There are some artefacts in the DVHs made by the PCE, as can be seen in the DVH distribution. In some of the DVHs, the $D_{100\%}$ is larger than the $D_{99\%}$ and in other cases the maximum dose is smaller than the $D_{1\%}$. This is because the PCE model is not accurate in the determination of the $D_{100\%}$ and the $D_{0\%}$, as discussed in Section 4.2.1.

The DVH distribution of 10,000 random PTV delineations was also made. The PTV was equal to the CTV plus a 5 mm margin. The DVH distribution is shown in Figure 4.26. The histogram of the $D_{95\%}$ values of the generated DVHs is displayed in Figure 4.27. In 17.1% of the random delineations the $D_{95\%}$ value is under the 54 Gy dose constraint. The mean $D_{95\%}$ for the random delineations equals 56.0 Gy and the average underdosage for the delineations which did not meet the target dose equals 2.1 Gy.



(a) An overview of the whole dose range of the DVH.

(b) A zoomed in view of the DVH distribution.

Figure 4.24: The DVH distribution of 10,000 randomly generated CTV delineations in the fixed VMAT plan. The colors indicate the confidence levels for a DVH to be in. The DVH of the gold-standard delineation and the delineation used for planning are also plotted. The intersection of the planning constraint and the 98% volume line indicates the $D_{98\%}$ dose constraint.

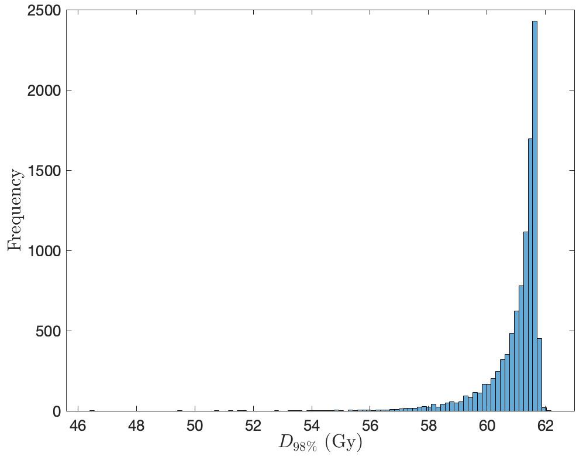
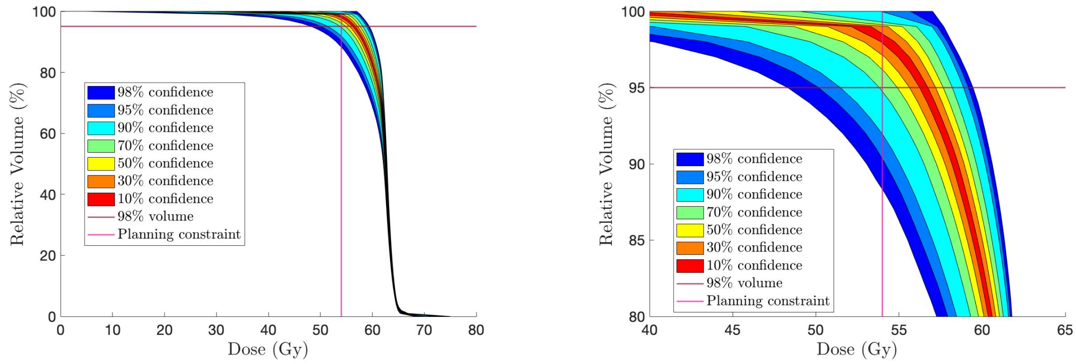


Figure 4.25: Histogram of the $D_{98\%}$ values for the 10,000 random delineations of the CTV in the fixed VMAT plan.



(a) An overview of the whole dose range of the DVH.

(b) A zoomed in view of the DVH distribution.

Figure 4.26: The DVH distribution of 10,000 randomly generated PTV delineations in the fixed VMAT plan. The colors indicate the confidence levels for a DVH to be in. The intersection of the planning constraint and the 95% volume line indicates the $D_{95\%}$ dose constraint.

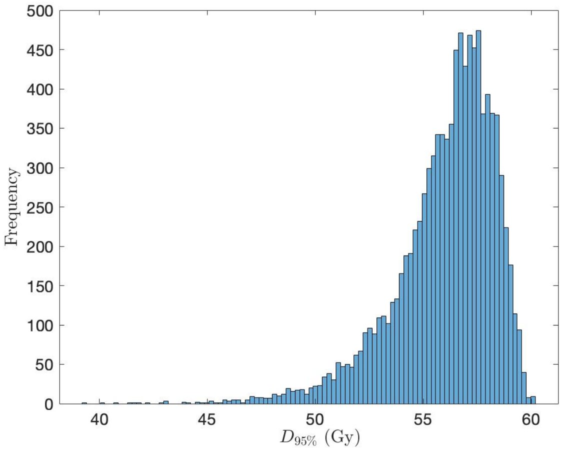
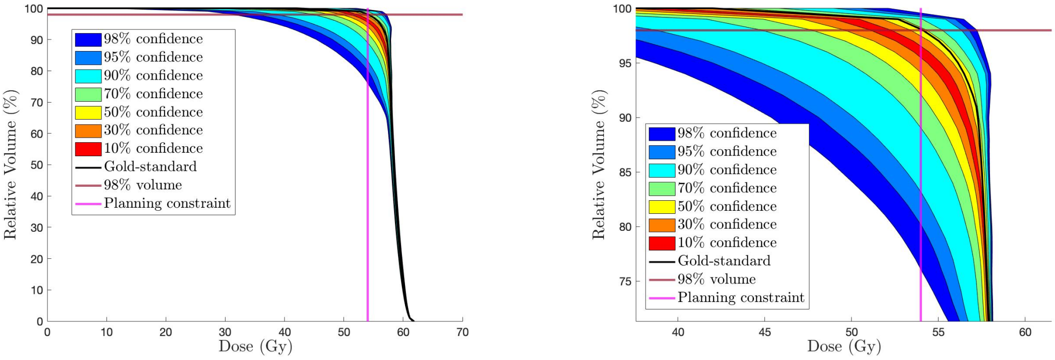


Figure 4.27: Histogram of the $D_{95\%}$ values for the 10,000 random delineations of the PTV in the fixed VMAT plan.

4.3.2. IMPT Plan

For the IMPT plan, the DVH distribution of 10,000 randomly generated CTV delineations is shown in Figure 4.28. It can be seen that for the gold-standard CTV delineation the $D_{98\%}$ dose constraint is exactly met, but in a majority of the cases the $D_{98\%}$ is under 54 Gy. This is reinforced by the histogram of the $D_{98\%}$ for all 10,000 random delineations displayed in Figure 4.29. The $D_{98\%}$ is under the dose constraint for 69.0% of the random delineations and the mean $D_{98\%}$ equals 50.4 Gy. Of the random delineations which do not meet the dose constraint, the average $D_{98\%}$ is 5.8 Gy under the target dose. The fact that the dose constraint is not met in a majority of the cases is due to the fact that the dose distribution is highly conformal to the gold-standard delineation. A delineation which is deformed with respect to the gold-standard will likely have part of its volume in a low dose region, which will make the $D_{98\%}$ drop. It should be noted that the DVHs calculated with the PCE for the fixed dose distribution in the IMPT plan are inaccurate, as discussed in Section 4.2.1. It should therefore be investigated whether the PCE model can be made more accurate to verify these results.



(a) An overview of the whole dose range of the DVH.

(b) A zoomed in view of the DVH distribution.

Figure 4.28: The DVH distribution of 10,000 randomly generated CTV delineations in the fixed IMPT plan. The colors indicate the confidence levels for a DVH to be in. The DVH of the gold-standard delineation used for planning is also plotted. The intersection of the planning constraint and the 98% volume line indicates the $D_{98\%}$ dose constraint.

The effect of adding robustness settings to the IMPT treatment plan was also investigated. To this end, the RR setting was set to 2% and the SR setting varied from 1 mm to 5 mm, in 1 mm increments. For each of these robustness settings a new PCE model was built using spectral projection with a grid level of 4, 1 extra level and a polynomial order of 5. The $\Delta D_{98\%,mean}$ and $\Delta D_{98\%,90}$ between the exact DVH and the PCE calculated DVH for all robustness settings in 100 validation scenarios are shown in Table 4.7. It can be seen that the $\Delta D_{98\%}$ is relatively large for all robustness settings, but does decrease as

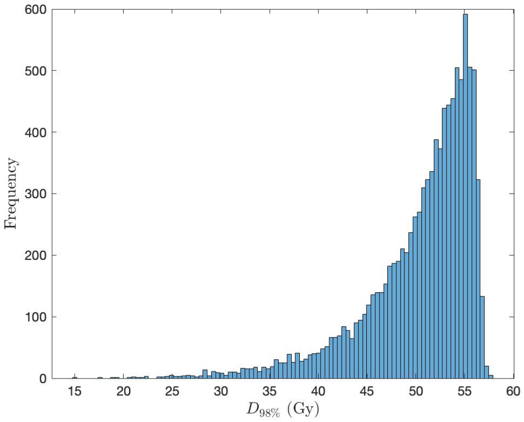


Figure 4.29: Histogram of the $D_{98\%}$ values for the 10,000 random delineations of the CTV in the fixed IMPT plan.

SR increases.

Table 4.7: The mean and 90th percentile difference across the 100 validation scenarios between the PCE calculated $D_{98\%}$ and the exact values for 5 different robustness settings.

Robustness settings	$\Delta D_{98\%,mean}$	$\Delta D_{98\%,90}$
RR=2%, SR=1 mm	0.3382 Gy	0.8549 Gy
RR=2%, SR=2 mm	0.2682 Gy	0.6366 Gy
RR=2%, SR=3 mm	0.1909 Gy	0.3881 Gy
RR=2%, SR=4 mm	0.1700 Gy	0.3806 Gy
RR=2%, SR=5 mm	0.1397 Gy	0.2824 Gy

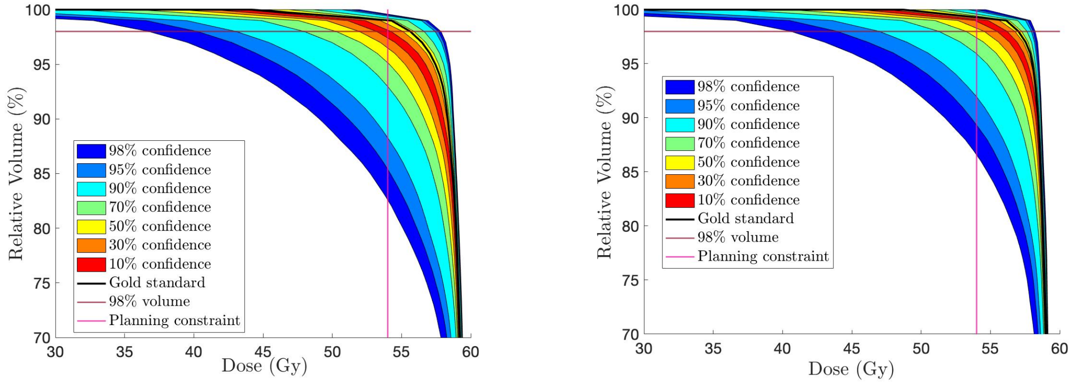
The resulting DVH distributions of 10,000 random CTV delineations in the robust IMPT plans are shown in Figure 4.30. It can be seen that the number of delineations that passes the CTV dose constraint increases as the SR setting is increased. This is also as expected as the area of high dose becomes larger when the SR increases and therefore a deformation of the gold-standard delineation will not lead to a lower $D_{98\%}$ in all cases. The fraction of delineations under the target dose, the mean $D_{98\%}$ and the average underdosage for the delineations that do not meet the dose constraint are summarized in Table 4.8 for the various robustness settings.
Similar to the non-robust IMPT plan, the PCE model is not entirely accurate for the robust IMPT plans. These results should therefore also be verified with a more accurate PCE model.

Table 4.8: The fraction of delineations under the target dose ($D_{98\%}$ = 54 Gy), the mean $D_{98\%}$ and the average underdosage for the delineations which do not receive the target dose for 10,000 random CTV delineations in 5 fixed IMPT treatment plans made with 5 different robustness settings.

Robustness settings	Fraction of delineations under target dose	Mean $D_{98\%}$	Average underdosage
RR = 2%, SR = 1 mm	52.6 %	52.4 Gy	4.8 Gy
RR = 2%, SR = 2 mm	30.3 %	54.6 Gy	4 Gy
RR = 2%, SR = 3 mm	17.9 %	55.8 Gy	3.3 Gy
RR = 2%, SR = 4 mm	9.7 %	56.7 Gy	3.0 Gy
RR = 2%, SR = 5 mm	4.0 %	57.3 Gy	2.5 Gy

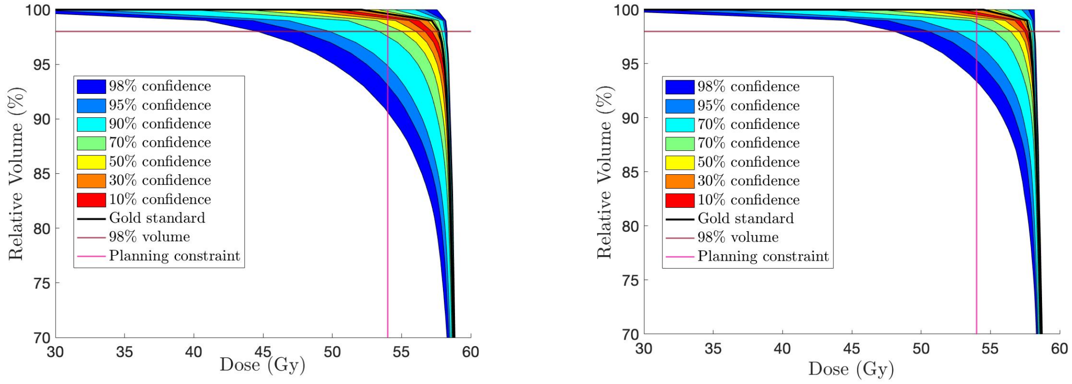
4.4. Dosimetric Effect of Delineation Uncertainties for a Reoptimized Dose Distribution

For the CTV in the reoptimized IMPT plan, the DVH distribution of 10,000 randomly generated delineations is shown in Figure 4.31. The DVHs were generated directly by the PCE. The mean $D_{98\%}$ for



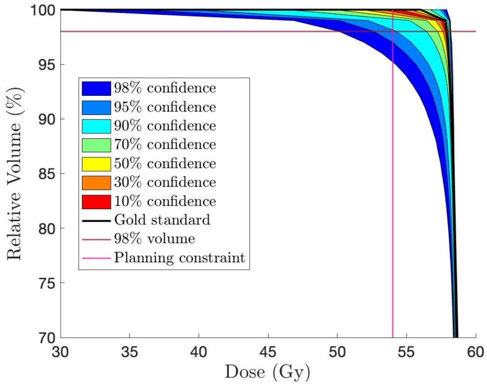
(a) RR = 2%, SR = 1 mm.

(b) RR = 2%, SR = 2 mm.



(c) RR = 2%, SR = 3 mm.

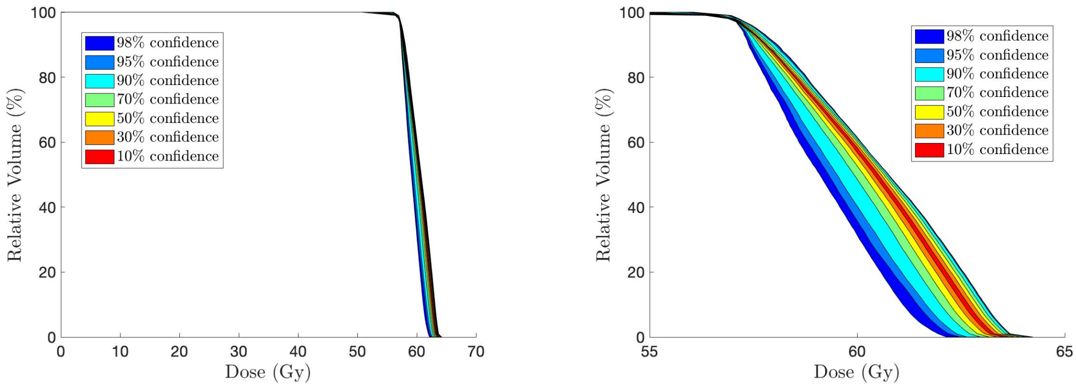
(d) RR = 2%, SR = 4 mm.



(e) RR = 2%, SR = 5 mm.

Figure 4.30: The DVH distributions of 10,000 randomly generated CTV delineations in 5 fixed robust IMPT plans. The colors indicate the confidence levels for a DVH to be in. The DVH of the gold-standard delineation used for planning is also plotted. The intersection of the planning constraint and the 98% volume line indicates the $D_{98\%}$ dose constraint.

the 10,000 delineations equals 56.99 Gy and all delineations meet the dose constraint ($D_{98\%} \geq 54$ Gy). This is to be expected as a minimum dose constraint to the CTV was set during the treatment plan optimization. However, it can be seen that the maximum dose in the CTV can vary. The maximum dose of the 10,000 random delineations varies between 62.4 Gy and 64.2 Gy. The maximum dose can vary because there is no constraint set on the maximum dose. It should be noted that the PCE model is not accurate with the used polynomial order of 4, as discussed in Section 4.2.2. Therefore these results are not entirely accurate.



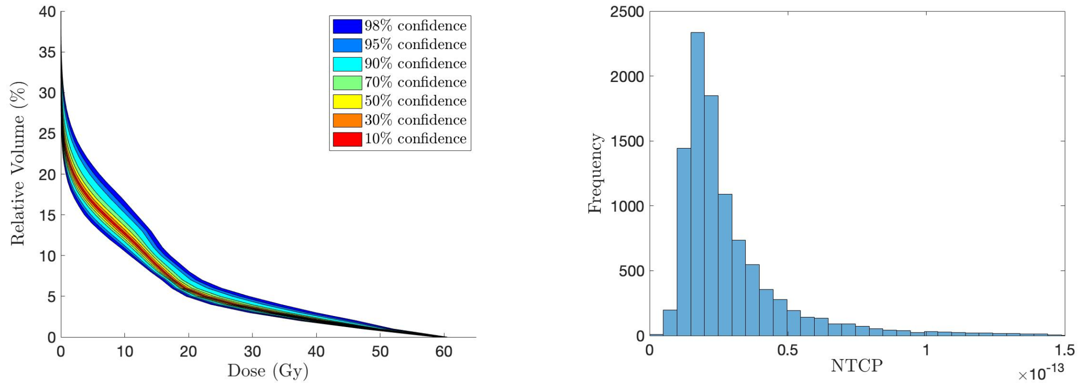
(a) An overview of the whole dose range of the DVH distribution.

(b) A zoomed in view of the DVH distribution.

Figure 4.31: The DVH distribution of 10,000 randomly generated CTV delineations in the reoptimized IMPT plans.

Next the effect of the reoptimized dose distribution on the DVH of the healthy liver was investigated by creating a DVH distribution of the liver in 10,000 reoptimized IMPT plans. The DVH distribution is shown in Figure 4.32a. It can be seen that the different reoptimized IMPT plans result in different doses delivered to liver and thus different DVHs. These results are however not entirely accurate because the PCE generated DVHs do not entirely overlap with the exact DVHs, as seen in Section 4.2.2. Therefore these results should be investigated more thoroughly when a more accurate PCE model has been built. The clinical effect that these differences in the DVH have on the liver can be quantified by the NTCP with the parameters given in Section 3.5.3. The NTCP gives the probability of radiation induced liver disease. The NTCP values of the liver in the 10,000 reoptimized IMPT plans are shown in the histogram in Figure 4.32b. The exact NTCP in the 9 validation scenarios was also calculated. In Table 4.9 the exact NTCP in each validation scenario is shown.

The NTCP of the liver is very low in all 10,000 reoptimized IMPT plans, meaning that the effect of the dose on the liver is minimal. However, it should be noted that NTCP values are dependent on the liver DVHs generated by the PCE which were not accurate, so the NTCP values are also not accurate. The exact NTCP values calculated in all validation scenarios are also very small and in the same order of magnitude as the PCE calculated NTCP values in the 10,000 reoptimized IMPT plans.



(a)

(b)

Figure 4.32: (a): The DVH distribution of the liver in 10,000 reoptimized IMPT plans. (b): Histogram of the NTCP values of the liver for the 10,000 DVHs in (a).

Table 4.9: The exact NTCP of the liver in all validation scenarios in the reoptimized IMPT plans. In the validation scenario column, the scenario is given by $(c_{1,random}, c_{2,random})$.

Validation scenario	NTCP
$(0, 0)$	$8.91 \cdot 10^{-15}$
$(1 \cdot \sqrt{\lambda_1}, 0)$	$2.44 \cdot 10^{-14}$
$(0, 1 \cdot \sqrt{\lambda_2})$	$3.66 \cdot 10^{-14}$
$(2 \cdot \sqrt{\lambda_1}, 0)$	$1.65 \cdot 10^{-14}$
$(0, 2 \cdot \sqrt{\lambda_2})$	$7.02 \cdot 10^{-14}$
$(1 \cdot \sqrt{\lambda_1}, 1 \cdot \sqrt{\lambda_2})$	$2.88 \cdot 10^{-14}$
$(1 \cdot \sqrt{\lambda_1}, 2 \cdot \sqrt{\lambda_2})$	$5.94 \cdot 10^{-14}$
$(2 \cdot \sqrt{\lambda_1}, 1 \cdot \sqrt{\lambda_2})$	$2.70 \cdot 10^{-14}$
$(2 \cdot \sqrt{\lambda_1}, 2 \cdot \sqrt{\lambda_2})$	$5.29 \cdot 10^{-14}$

5

Discussion

In this chapter the results obtained in this thesis are discussed. First the use of PCA as a method to characterize the delineation uncertainties for manual delineations and auto-contours is discussed in Section 5.1, along with the limitations of PCA to characterize delineation uncertainties and other potential methods to model delineation uncertainties. Then the use of PCE as a meta-model for the DVH and dose distribution for the fixed dose distribution and the reoptimized dose distribution is discussed in Section 5.2. Finally, the dosimetric effects that delineation uncertainties have for a fixed dose distribution and for a reoptimized dose distribution are examined in Section 5.3.

5.1. Principal Component Analysis

5.1.1. Inter-Observer Variability

PCA was used to characterize the delineation uncertainty of a GTV in the liver delineated by different observers. As far as we know, this is the first time that PCA has been used to describe delineation variability. PCA has been used before to characterize e.g., the interfractional motion of the prostate, rectum and bladder during radiotherapy treatment [37].

The delineations of the GTV were shown to vary quite significantly as could be seen in the probability map. The ratio of the maximum and minimum delineated volume and the DSCs also showed a big variation between the delineations. 5 eigenmodes were needed to explain 90% the variability in the shapes of the 12 delineations of the GTV and the dimensionality of the problem could be greatly reduced with these eigenmodes.

It should however be noted that the delineations of 1 GTV of 1 patient were analyzed in this study. It could therefore be that for other ROIs the shape variations between delineations by different radiation oncologists are less or more correlated than found in this study and a different number of eigenmodes would be needed to explain the variations. The GTV is an ROI which does not require much interpretation by the radiation oncologist, as only the visible part of the tumor on the medical image needs to be delineated. The GTV in this thesis was also equal to the CTV, so the variations in the CTV were equal to the variations in the GTV. However, for most tumors the CTV is not equal to the GTV, as the CTV encompasses microscopic tumor extensions around the GTV which are not visible on the medical image. In this case it is expected that the variations in the CTV delineations would be very different from the variations of the GTV found in this thesis because there is more uncertainty in which parts of the medical image belong to the CTV. The variations between CTV delineations can be reduced by using a fixed tumor specific GTV to CTV margin which is determined from pathological studies [16]. Further analysis of CTV delineation uncertainties due to the presence of microscopic tumor extensions is however outside the scope of this thesis.

Another aspect to take into account in the use of PCA to characterize delineation uncertainties is that the shapes of tumors from patient to patient can vary, which means that the PCA model to characterize the uncertainties in the delineations of the GTV for one patient is not directly transferable to another patient.

5.1.2. Auto-Contouring Variability

PCA was also used to characterize the delineation uncertainty of the brainstem and the high dose CTV of a head and neck patient. The method of creating synthetic CTs from random DVFs worked well as this ensured that only realistic synthetic CTs were formed. In MIM it was chosen to use single-atlas based auto-contouring. This choice was made because this is the atlas-based auto-contouring method which has the biggest inaccuracies, as discussed in Section 2.3.2. The auto-contours registered back to the planning CT could have variations with respect to each other due to one or both of the following reasons:

- An inaccuracy in the DVF formed by registering the planning CT to one of the repeat CTs.
- An inaccuracy in the auto-contours on the synthetic CTs made with single-atlas based auto-contouring.

It was found that 45 eigenmodes were needed to explain 90% of the variance between the auto-contours of the brainstem. This is a very large number and it is not convenient to use 45 random coefficients as an input for a PCE model or any other model. The large number of eigenmodes shows that many variations between the different auto-contours are highly uncorrelated. In the probability map which shows the overlap of the different auto-contours it could already be seen that the differences between the 90 auto-contours were at most a 1 millimeter shift in any direction. These differences in the auto-contours seem to be arbitrary as these shifts occur in every single direction along the contour. These random shifts could be a consequence of an inaccuracy in the registration of the planning CT, an inaccuracy in the auto-contouring software or an inaccuracy in the registration of the auto-contours back to the planning CT. In any case, the uncertainty in the delineation of the brainstem is not convenient to model with PCA. Therefore a different method to characterize the uncertainty in the brainstem delineations would need to be found which is dependent on less variables.

For the high dose CTV delineations it was not possible to build a PCA model due to the number of double crossings during the determination of the surface shape vector. More details on this limitation are discussed in Section 5.1.4. From the probability map it could however be seen that there is more variation in the auto-contours of the high dose CTV than in the auto-contours of the brainstem. This is possibly because the tumor is not as visible on the CT as the OARs and therefore the delineations by the auto-contouring software will vary more.

5.1.3. PCA Performance

The number of eigenmodes needed to explain 90% of the variability of the delineations was significantly different for the manual delineations and the auto-contours. However, when looking at the relative number of principal components with respect to the number of delineations, the dimension of the variation space was reduced by 58% for the manual delineations, while for the auto-contours of the brainstem the dimension of the variation space was reduced by 50%. The dimensionality reduction is thus relatively similar for both the manual delineations and the auto-contours.

The eigenmodes for the auto-contours and the manual delineations represent very different variations between the delineations. In the case of the auto-contours, the eigenmodes represent the variations due to inaccuracies in the auto-contouring software or image registration, as discussed in Section 5.1.2. These variations are very small and can be seen as noise which is largely uncorrelated, which leads to a large number of eigenmodes needed to explain the variations. For the manual delineations, the eigenmodes represent the variations in the delineations caused by the fact that different radiation oncologists delineate the same structure differently. It can for instance be that one radiation oncologist systematically delineates a structure slightly larger or more shifted to one direction than other radiation oncologists. Such variations can also be seen in the eigenmodes in Figure 4.6, where the first two eigenmodes of the manual GTV delineations led to a change of shape and a change of size of the mean delineation. These variations are more correlated than the auto-contouring variations and can be described by fewer eigenmodes.

The ability to create realistic random delineations with the eigenmodes from PCA is limited by the data used as input for PCA, as the eigenmodes only represent variations that are present in the input. It is therefore necessary to make sure that the input delineations contain as much of the realistic variations as possible and therefore a large number of input delineations is desired. However, the input delineations can contain redundancy which means that there can be an ideal number of input delineations which represents the variability between the delineations. *Yu et al.* [48] performed a study looking at the ideal number of input DVFs needed as input for the PCA to represent all variations present in the DVFs. The input DVFs were obtained by registering a planning CT to a number of repeat CTs. To find the ideal number of DVFs needed, the number of eigenmodes needed to explain 90% of the variations in the data was studied as a function of the number of input DVFs. When the number of input DVFs was not enough, the number of eigenmodes needed to explain 90% of the variations increased when the number of input DVFs increased. However, when the number of input DVFs was sufficient, the number of eigenmodes stayed approximately the same, which meant that the new input DVFs did not add any new variations. It was found that 26 DVFs were needed to capture all possible variations.

A similar analysis could be performed with the input delineations in this thesis. It could for instance be that that the 12 manual delineations of the GTV by the radiation oncologists used in this study do not capture all possible variations that exist due to inter-observer variability between radiation oncologists and that more delineations would be needed create an accurate uncertainty model.

5.1.4. Limitations

There are several limitations of the use of PCA to characterize delineation uncertainties. The first limitation is that multiple intersections can occur during the determination of the surface shape vector. This effect was particularly visible in the determination of the surface shape vector for the high dose CTV for the head and neck patient, where in one of the regions 3.5% of the elements of the surface vector were inaccurate due to multiple intersections. This high number multiple intersections was due to a concavity in the shape of the delineation, which is the case for the top region of the high dose CTV. Because only the first intersection point is taken into account in the surface shape vector, there will be parts of the surface which are not characterized by the surface shape vector in the case of multiple intersections. Therefore this method of delineation characterization is only particularly suited to convex shapes.

The second limitation of the use of PCA lies in the process of obtaining the random delineations on a 2D CT slice. The random surface shape vector produces a cloud of scatter points which form the surface of the 3D random delineation. To obtain the delineations on a 2D CT slice, the 3D scatter points just above and below the height of the CT slice are projected onto the 2D CT slice. However, these points do not truly lie on the 2D slice, which means that the delineation on the 2D slice does not exactly follow the 3D delineation shape. This problem could be alleviated by reducing the increments with which polar angle increased when determining the surface shape vector. This way more points of the 3D random delineation will lie closer to or on the 2D slice, resulting in a more accurate 2D delineation. However, this solution would come at the price of an increased computation time and memory use.

Another limitation encountered in this research uncertainties comes forth from the aforementioned projection of the scatter points onto the 2D CT slices. The function *boundary* in Matlab is used to form a boundary around the scatter points on the 2D slice. This boundary is the 2D delineation. However, the function *boundary* does not take multiple regions of the delineation on the 2D slice into account. Therefore, the 2D delineation is formed around both regions and encloses parts of the CT which are not part of the true 3D delineation. This is however more a limitation of the function used in Matlab than a limitation of PCA applied to delineations.

5.1.5. Other Delineation Uncertainty Characterization Methods

There are other methods which can potentially be used to characterize delineation uncertainties and do not suffer from the limitations discussed in Section 5.1.4. One of these methods would be the use of a Variational Autoencoder (VAE). The VAE is a deep learning generative model which is able to generate new data based on training with real data. A VAE consists of two parts: an encoder and a decoder. In the encoder the dimensionality of the data is reduced from the initial space to the latent space. In the case of delineation geometries, the input to the encoder could be the a 3D matrix with the size of the CT, where the matrix elements equal 1 inside the delineation and 0 outside the delineation. When the dimensionality of the input is reduced from the input space to the latent space, the goal is to minimize the amount of information that is lost. From the latent space the decoder then reconstructs the original input data. The encoder and decoder pair are modified during training such that the input data and the reconstructed input data are as similar as possible. Within the latent space, each input is encoded as a Gaussian distribution. The latent space is regularized such that the Gaussian distributions of each input are overlapping to ensure there is continuity in the latent space. Once the model has been trained it can be used to generate new data. This is done by sampling a point from latent space and decoding this point with the decoder.

The VAE could be used in combination with the PCE to create random delineations, as was done with PCA in this thesis. The uncertain input vector of the PCE would then consist of a point which is sampled from the overlapping Gaussian distributions in latent space.

5.2. Polynomial Chaos Expansion

5.2.1. PCE as a Meta-Model for the DVH for a Fixed Dose Distribution

PCE was used to model the DVH of different realizations of a delineation on a fixed dose distribution. For the VMAT plan, the PCE was accurate in determining the DVH of the CTV and PTV, with a mean error $\Delta D_{98\%,mean}$ of 0.09 Gy for the CTV and a mean error $\Delta D_{95\%,mean}$ of 0.13 Gy in the validation scenarios. Furthermore, in 90% of the validation scenarios the error in the $D_{98\%}$ was less than 0.19 Gy for the CTV and the error in the $D_{95\%}$ for the PTV was less than 0.33 Gy in 90% of the validation scenarios.

For the CTV in the IMPT plan, the error was however larger, with $\Delta D_{98\%,mean}$ equal to 0.34 Gy and $\Delta D_{98\%,90}$ equal to 0.61 Gy. The discrepancy between the exact $D_{98\%}$ and the PCE calculated $D_{98\%}$ is particularly present when $c_{l,random}$ is near 0, as can be seen in Figure 4.19. This could be caused by the range of $c_{l,random}$ that the PCE was built with. The range of $c_{l,random}$ in building the PCE was from approximately $-6 \cdot \sqrt{\lambda_l}$ to $6 \cdot \sqrt{\lambda_l}$. This is outside the range that realistic delineations are formed and therefore the expansion coefficients r_k are calculated using potential unrealistic values of $D_{98\%}$. Because the polynomials are then fit to incorrect training points, the accuracy of the PCE calculated $D_{98\%}$ when $c_{l,random}$ is near 0 could decrease.

5.2.2. PCE as a Meta-Model for the DVH and Dose Distribution for a Reoptimized Dose Distribution

A proof of principle of the use of PCE to serve as a meta-model for the dose distribution and DVHs of the CTV and liver for a random realization of the CTV was given. To make the PCE more accurate, a higher polynomial order would be needed in the construction of the PCE. This was however not achievable in this research as it took approximately 6 days to construct the PCE with a polynomial order of 4. The computation time could be decreased by running multiple iCycle optimizations simultaneously when building the PCE model.

It is surprising that the number of accepted voxels was higher for the PCE model built with a polynomial order of 3 than for the PCE model with a polynomial order of 4 in many of the validation scenarios. This could be because the PCE models were only validated in a small number of scenarios and that the model with polynomial order of 4 does perform better in other scenarios which were not checked.

5.3. Dosimetric Effect of Delineation Uncertaties

5.3.1. Fixed Dose Distribution

VMAT plan

The effect of the manual delineation uncertainties on the CTV in the VMAT plan was found to be small, as only 0.1% of the random delineations did not receive the target dose. It can thus be concluded that on this treatment plan made with the planning delineation, the PTV made with a 5 mm margin with respect to the planning delineation is large enough to account for other possible realizations of the CTV. This means that the 5 mm PTV margin is enough to account for the delineation uncertainties in this case. However, the PTV is also meant to deal with e.g., setup uncertainties, which were not taken into account in this study. For the PTV, it was found that in 17.1% of the realizations of the PTV the target dose was not met, which is a significant amount. An underdosed PTV means that the CTV could also not receive enough dose. It was found that basically all realizations of the CTV would receive enough dose, but if the delineation uncertainty is combined with setup uncertainties, the potential underdosage of the PTV could lead to an underdosage of the CTV. Therefore the effect of delineation uncertainties in combination with setup uncertainties would need to be investigated.

IMPT plan

The effect of delineation uncertainties in the IMPT plan was significant, with 69% of the realizations of the CTV not receiving the target dose. This is a consequence of the very steep dose fall-off around the gold-standard delineation which was used for planning. If a part of a delineation falls outside the high-dose region, the $D_{98\%}$ drops drastically. This shows that delineation uncertainties can be a big problem in proton therapy treatments. Robust treatment plans were shown to increase the number of CTV delineations which meet the target dose, with 96.0% of the delineations meeting the target dose when the RR is set to 2% and the SR is set to 5 mm.

It should be noted that the PCE did have relatively large errors in determining the $D_{98\%}$ of the CTV in the IMPT plan, as $\Delta D_{98\%,mean}$ was 0.34 Gy in the validation scenarios. It should therefore be investigated whether a more accurate PCE model can be made.

If the above results are found to be accurate, then it seems possible that robustness settings can be used to mitigate delineation uncertainties in IMPT. This makes sense because the high dose region becomes larger when robustness settings are used and a deformed delineation will thus probably still be positioned in the high dose region. These robustness settings are however meant to compensate for range and setup uncertainties, and not for delineation uncertainties. Delineation uncertainties are fundamentally different from range and setup uncertainties, as in the case of range and setup uncertainties the delineation stays the same and the larger dose cloud assures that the CTV will receive sufficient dose in the presence of setup or range errors. However, the dose in this dose cloud is not homogeneous dose and is not static. The shape and dose distribution of the dose cloud will change if there is e.g., a setup shift. The amount of change in the shape and dose distribution of the dose cloud are dependent on e.g., the change of the density of the tissue that the proton beams travel through due to this setup shift. Due to this changing nature of the dose cloud when a setup shift occurs, it is not assured that a random realization of the CTV, that receives sufficient dose when there is no setup shift, will still receive sufficient dose in the presence of a setup shift. It would therefore be necessary to investigate whether delineation uncertainties can still be accounted for with robust treatment planning while simultaneously taking into account setup and range uncertainties.

5.3.2. Reoptimized Dose Distribution

For the dose distribution which was reoptimized for every random CTV delineation, it was found that the dose constraint ($D_{98\%} \geq 54$ Gy) was always met. This was as expected as a new reoptimized plan for each delineation was made which should satisfy the planning constraints. However, the dose along the rest of the DVH did vary between the different realizations of the delineations. A PCE model with a higher accuracy would however be needed to study this effect more thoroughly.

The PCE model for the reoptimized dose distribution was also able to model the DVH of the liver, and from the DVH the NTCP could be calculated. It could be seen that the DVHs in the different reoptimized dose distributions differed from each other. The exact effect on the DVH of the liver can however only be investigated with a more accurate PCE model. Nonetheless, from the NTCP values in the validation scenarios it can be concluded that the dosimetric effect on the liver was minimal.

The use of the PCE as a meta-model for the reoptimized dose distribution could be an important tool in the future of adaptive radiotherapy. In adaptive radiotherapy, as discussed in Section 2.3.1, a daily CT is made of the patient which is delineated. Current research in adaptive radiotherapy aims at creating the auto-contours and a treatment plan within several minutes after the repeat CT so that the patient can be treated immediately before any anatomical changes occur [24]. These auto-contours are used to adapt the existing treatment plan. It is however important to take the uncertainty in the auto-contours into account and know the dosimetric effect of the uncertainty on the target and on the OARs. This dosimetric effect can be studied using PCE as a meta-model for the reoptimized dose distribution. Based on the dosimetric effect on the target and the OARs, it can be decided whether adapting an existing treatment plan to the daily CT is sufficient, or that a new reoptimized plan needs to be made on the daily CT.

6

Conclusions & Recommendations

In this chapter the conclusions of this thesis will be given in Section 6.1 along with recommendations for future research in Section 6.2. Finally, in Section 6.3 an outlook will be given on the application of the methods developed in this thesis to future clinical practice.

6.1. Conclusions

In this work principal component analysis (PCA) was used to characterize the delineation uncertainties for manual delineations of the GTV of a hepatocellular carcinoma patient and the auto-contours of the CTV and brainstem of a head and neck patient. The uncertainty in the auto-contours of the CTV could not be characterized by PCA due to the concave shape of the CTV. 45 eigenmodes were needed to explain the variation in the auto-contours of the brainstem and the uncertainties in the manual delineations of the GTV could be explained by 5 eigenmodes. Using these eigenmodes it was possible to create random realizations of the delineations. The characterization of delineation uncertainties with PCA has its limitations and is not applicable to all delineation shapes. Therefore a different method of characterization is needed in future research.

Polynomial Chaos Expansion (PCE) was shown to be capable of modeling the dosimetric effects of the manual delineation uncertainties. For the manual delineations, the CTV was equal to the GTV and the PTV was equal to the CTV plus a 5 mm margin. On a fixed VMAT dose distribution it was found that uncertainties in the GTV delineations did not lead to an underdosage of the CTV. However, the uncertainties did lead to an underdosage of the PTV in 17.1% of the random delineations. For a fixed IMPT dose distribution, the CTV was underdosed in 69.0 % of the random delineations. This percentage went down when the IMPT treatment plan was made robustly. The results on the fixed IMPT plan however need to be verified with a more accurate PCE model.

PCE was also used to model the reoptimized dose distribution and the DVH of the CTV and liver for random realizations of the CTV delineation. However, a more accurate PCE model with a higher polynomial order would be needed to analyze the dosimetric effect to the CTV and surrounding OARs due to the reoptimization.

PCE has been found to be a powerful tool in modeling the dosimetric effect of delineation uncertainties on a fixed dose distribution and a proof of principle has been given for the PCE as a model of the reoptimized dose distribution. This is a first step towards systematically and quantitatively taking into account delineation uncertainties in radiotherapy treatment planning.

6.2. Recommendations

For future research on the effect of delineation uncertainties on treatment planning, there are several areas that can be improved or researched further.

Firstly, a different method of characterizing the delineation uncertainty should be developed which can be applied to more delineation shapes. This could be a machine learning approach like the VAE, as described in Section 5.1.5.

The results in this thesis have shown that the dosimetric impact of delineation uncertainties of the CTV and PTV on a fixed dose distribution can be significant. For a photon therapy plan it could be investigated whether the current CTV to PTV margin is sufficient to account for delineation uncertainties in combination with setup uncertainties, or that the CTV to PTV margin needs to be altered. PCE is a powerful tool to perform this analysis with. For proton therapy plans it was shown that robust planning is able to compensate for delineation uncertainties. However, the PCE model lacked accuracy for the proton therapy plan. It therefore should be investigated whether the PCE model can be made more accurate. If an accurate PCE model is obtained, a next step would be to study whether robust treatment planning is able to account for the delineation uncertainties in combination with setup and range uncertainties and which robustness settings would be needed to do so.

Another aspect that can be researched further is the PCE model as a meta-model for the reoptimized dose distribution. The model developed in this thesis was not accurate enough to perform an analysis on the dosimetric impact that the reoptimization has, but if a more accurate PCE with a higher polynomial order is built, an analysis of the dose delivered to the target and OARs can be performed. This analysis can also include the impact on the NTCP of an OAR. In this research it was found that the NTCP of the liver was very low in all scenarios. However, in other more sophisticated cases where an OAR receives more dose, an analysis of the DVH and the NTCP of an OAR with the PCE model for the reoptimized dose distribution can give a very good insight into the variability of these metrics and if the OAR is at risk of being overdosed.

The impact of auto-contouring uncertainties in adaptive radiation therapy can also be investigated further using the PCE model for the reoptimized dose distribution. Using this PCE model, critical structures on which the delineation uncertainties have a large dosimetric impact can be identified. The identification of critical structures can take place by modeling the DVH distribution of all structures in many realizations of the delineation and determining which structures have the possibility of being underdosed (in the case of a target) or overdosed (in the case of an OAR).

6.3. Outlook

Hopefully, with further research and improvements, the methodology researched in this can be implemented in future clinical practice to account for delineation uncertainties in radiotherapy treatment planning. For instance, treatment plans could be optimized such that the dose to the ROIs is adequate with a certain probability given the delineation uncertainties of the ROIs. By taking the characterized uncertainties into account during the treatment planning, the use of margins around a target to account for delineation uncertainties would be unnecessary. A similar technique is used in robust treatment planning to account for setup and range uncertainties in IMPT. However, robust treatment planning only considers a limited number of error scenarios. A more promising approach is probabilistic treatment planning. In probabilistic treatment planning many more scenarios are taken into account along with the probability of the scenarios [41]. Applied to delineation uncertainties, the scenarios would be the different possible realizations of the delineation. Furthermore, developments are being made in the use of functional MRI to quantify the tumor cell density in the target [34], which can help in providing probabilistic definitions of the target.

The developed methods can also be of importance in adaptive radiotherapy which uses auto-contouring. Adaptive radiotherapy can use probabilistic treatment planning based on the characterized uncertainties, but critical structures or areas where the dosimetric impact of delineation uncertainties is high can also be identified. The critical areas where the dosimetric effects of delineation uncertainties are large can be flagged such that they can be manually assessed by a radiation oncologist, ensuring that the delineation is correct and the dose contraints will be met during treatment.

Finally, in adaptive radiotherapy the developed methods in this thesis can also be used to determine whether adapting an existing treatment plan is sufficient for a successful treatment, or that a new reoptimized treatment plan needs to be made, as also discussed in Section 5.3.2.

Bibliography

[1] Eric Aliotta, Hamidreza Nourzadeh, and Jeffrey Siebers. “Quantifying the dosimetric impact of organ-at-risk delineation variability in head and neck radiation therapy in the context of patient setup uncertainty”. In: *Physics in Medicine & Biology* 64.13 (2019), p. 135020.

[2] James L Bedford and Alan P Warrington. “Commissioning of volumetric modulated arc therapy (VMAT)”. In: *International Journal of Radiation Oncology* Biology* Physics* 73.2 (2009), pp. 537–545.

[3] Géraud Blatman and Bruno Sudret. “Adaptive sparse polynomial chaos expansion based on least angle regression”. In: *Journal of computational Physics* 230.6 (2011), pp. 2345–2367.

[4] Sebastiaan Breedveld, Pascal RM Storchi, and Ben JM Heijmen. “The equivalence of multicriteria methods for radiotherapy plan optimization”. In: *Physics in Medicine & Biology* 54.23 (2009), p. 7199.

[5] Sebastiaan Breedveld et al. “iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans”. In: *Medical physics* 39.2 (2012), pp. 951–963.

[6] Carlos E. Cardenas et al. “Advances in Auto-Segmentation”. In: *Seminars in Radiation Oncology* 29.3 (2019). Adaptive Radiotherapy and Automation, pp. 185–197. ISSN: 1053-4296. DOI: `https://doi.org/10.1016/j.semradonc.2019.02.001`. URL: `http://www.sciencedirect.com/science/article/pii/S1053429619300104`.

[7] Timothy F Cootes et al. “The use of active shape models for locating structures in medical images”. In: *Biennial International Conference on Information Processing in Medical Imaging*. Springer. 1993, pp. 33–47.

[8] Daniele Davino. *Theory, Design and Tests on a Prototype Module of a Compact Linear Accelerator for Hadrontherapy*. Feb. 2016. DOI: `10.13140/RG.2.1.3010.2803`.

[9] Laura A Dawson et al. “Analysis of radiation-induced liver disease using the Lyman NTCP model”. In: *International Journal of Radiation Oncology* Biology* Physics* 53.4 (2002), pp. 810–821.

[10] Joseph O. Deasy. “Comments on the use of the Lyman-Kutcher-Burman model to describe tissue response to nonuniform irradiation.” In: *International journal of radiation oncology, biology, physics* 47.5 (2000), pp. 1458–1460.

[11] Matthew Deeley et al. “Comparison of manual and automatic segmentation methods for brain structures in the presence of space-occupying lesions: a multi-expert study”. In: *Physics in Medicine & Biology* 56.14 (2011), p. 4557.

[12] Omar Desouky and Guangming Zhou. “Biophysical and Radiobiological Aspects of Heavy Charged Particles”. In: *Journal of Taibah University for Science* 10 (Apr. 2015). DOI: `10.1016/j.jtusci.2015.02.014`.

[13] Nhat Trung Doan, Jonathan Orban de Xivry, and Benoit Macq. “Effect of inter-subject variation on the accuracy of atlas-based segmentation applied to human brain structures”. In: *Medical Imaging 2010: Image Processing*. Vol. 7623. International Society for Optics and Photonics. 2010, 76231S.

[14] Vincent Grégoire and Thomas R. Mackie. “State of the art on dose prescription, reporting and recording in Intensity-Modulated Radiation Therapy (ICRU report No. 83)”. In: *Cancer/Radiothérapie* 15.6-7 (2011), pp. 555–559.

[15] Steven J.M. Habraken et al. “Tumor and OAR delineation variation–Dice coefficient versus dose assessed with automated planning”. In: *Radiotherapy and Oncology*. Vol. 127. ELSEVIER IRELAND LTD ELSEVIER HOUSE, BROOKVALE PLAZA, EAST PARK SHANNON, CO … 2018, S1182–S1183.

[16] Christian Rønn Hansen et al. “Consequences of introducing geometric GTV to CTV margin expansion in DAHANCA contouring guidelines for head and neck radiotherapy”. In: *Radiotherapy and oncology* 126.1 (2018), pp. 43–47.

[17] Serhat Hosder, Robert Walters, and Michael Balch. “Efficient sampling for non-intrusive polynomial chaos applications with multiple uncertain input variables”. In: *48th AIAA/ASME/ASCE/AHS /ASC Structures, Structural Dynamics, and Materials Conference*. 2007, p. 1939.

[18] Dominique P Huyskens et al. “A qualitative and a quantitative analysis of an auto-segmentation module for prostate cancer”. In: *Radiotherapy and Oncology* 90.3 (2009), pp. 337–345.

[19] Thyrza Z Jagt et al. “Plan-library supported automated replanning for online-adaptive intensity-modulated proton therapy of cervical cancer”. In: *Acta Oncologica* 58.10 (2019), pp. 1440–1445.

[20] Kenneth L Judd et al. “Smolyak method for solving dynamic economic models: Lagrange interpolation, anisotropic grid and adaptive domain”. In: *Journal of Economic Dynamics and Control* 44 (2014), pp. 92–123.

[21] Hanne M. Kooy and Clemens Grassberger. “Intensity modulated proton therapy”. In: *The British journal of radiology* 88.1051 (2015), p. 20150195.

[22] Alexander V Louie et al. “Inter-observer and intra-observer reliability for lung cancer target volume delineation in the 4D-CT era”. In: *Radiotherapy and Oncology* 95.2 (2010), pp. 166–171.

[23] MATLAB. *9.7.0.1190202 (R2017a)*. Natick, Massachusetts: The MathWorks Inc., 2017.

[24] Michael Matter et al. “Intensity modulated proton therapy plan generation in under ten seconds”. In: *Acta oncologica* 58.10 (2019), pp. 1435–1439.

[25] Douglas C Montgomery. *Design and analysis of experiments*. John Wiley & Sons, 2017.

[26] Jens Olesen. *DICOM-RT to Matlab*. `https://github.com/ulrikls/dicomrt2matlab`. 2019.

[27] Zoltán Perkó et al. “Fast and accurate sensitivity analysis of IMPT treatment plans using Polynomial Chaos Expansion”. In: *Physics in Medicine & Biology* 61.12 (2016), p. 4646.

[28] Zoltán Perkó et al. “Grid and basis adaptive polynomial chaos techniques for sensitivity and uncertainty analysis”. In: *Journal of Computational Physics* 260 (2014), pp. 54–84.

[29] Ken Pu. *SVD and PCA*. 2010. URL: `https://sites.google.com/site/csci5740g/home/topics-lectures/svdandpca` (visited on 01/21/2021).

[30] Coen RN Rasch et al. “Decreased 3D observer variation with matched CT-MRI, for target delineation in Nasopharynx cancer”. In: *Radiation Oncology* 5.1 (2010), p. 21.

[31] Barbara Rombi et al. “Proton radiotherapy for pediatric tumors: Review of first clinical results”. In: *Italian journal of pediatrics* 40 (Sept. 2014), p. 74. DOI: `10.1186/s13052-014-0074-6`.

[32] Karlissa Schick et al. “An assessment of OAR delineation by the radiation therapist”. In: *Radiography* 17.3 (2011), pp. 183–187.

[33] Bas Schipaanboord et al. “Can Atlas-Based Auto-Segmentation Ever Be Perfect? Insights From Extreme Value Theory”. In: *IEEE Transactions on Medical Imaging* 38.1 (2019), pp. 99–106. DOI: `10.1109/TMI.2018.2856464`.

[34] Kathleen M Schmainda. “Diffusion-weighted MRI as a biomarker for treatment response in glioma”. In: *CNS oncology* 1.2 (2012), pp. 169–180.

[35] Barbara Segedin and Primoz Petric. “Uncertainties in target volume delineation in radiotherapy–are they relevant and what can we do about them?” In: *Radiology and oncology* 50.3 (2016), pp. 254–262.

[36] James Shackleford, Nagarajan Kandasamy, and G Sharp. “Accelerating MI-based B-spline registration using CUDA enabled GPUs”. In: *Int. Conf. Med. Image Comput. Comput. Assist. Interv.* 2012.

[37] Matthias Söhn et al. “Modelling individual geometric variation based on dominant eigenmodes of organ deformation: implementation and evaluation”. In: *Physics in Medicine & Biology* 50.24 (2005), p. 5893.

[38] Roel JHM Steenbakkers et al. “Observer variation in target volume delineation of lung cancer related to radiation oncologist–computer interaction: a ‘Big Brother’evaluation”. In: *Radiotherapy and Oncology* 77.2 (2005), pp. 182–190.

[39] Hyuna Sung et al. “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries”. In: *CA: a cancer journal for clinicians* (2021).

[40] Patricia Tai et al. “Variability of target volume delineation in cervical esophageal cancer”. In: *International Journal of Radiation Oncology* Biology* Physics* 42.2 (1998), pp. 277–288.

[41] David Tilly et al. “Probabilistic optimization of dose coverage in radiotherapy”. In: *Physics and Imaging in Radiation Oncology* 10 (2019), pp. 1–6.

[42] Jaroslaw Tuszynski. *Triangle/Ray Intersection*. `https://nl.mathworks.com/matlabcentral/fileexchange/33073-triangle-ray-intersection`. 2021.

[43] Jan Unkelbach and Harald Paganetti. “Robust proton treatment planning: physical and biological optimization”. In: *Seminars in radiation oncology*. Vol. 28. 2. Elsevier. 2018, pp. 88–96.

[44] Marcel Van Herk et al. “The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy”. In: *International Journal of Radiation Oncology* Biology* Physics* 47.4 (2000), pp. 1121–1135.

[45] Simon K Warfield, Kelly H Zou, and William M Wells. “Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation”. In: *IEEE transactions on medical imaging* 23.7 (2004), pp. 903–921.

[46] Steven van de Water et al. “Improved efficiency of multi-criteria IMPT treatment planning using iterative resampling of randomly placed pencil beams”. In: *Physics in Medicine & Biology* 58.19 (2013), p. 6969.

[47] Dongbin Xiu and George Em Karniadakis. “The Wiener–Askey polynomial chaos for stochastic differential equations”. In: *SIAM journal on scientific computing* 24.2 (2002), pp. 619–644.

[48] Henry Yu et al. “Learning anatomy changes from patient populations to create artificial CT images for voxel-level validation of deformable image registration”. In: *Journal of applied clinical medical physics* 17.1 (2016), pp. 246–258.