Margin recipes for stereotactic body radiation therapy for prostate cancer treatments

Master thesis

Wens Kong

Applied Physics

Supervisors:

Dr. Z. Perkó, TU Delft, RST, MP &T

Dr. S.J.M. Habraken, EMC

Prof. Dr. B.J.M. Heijmen, EMC

02-05-2019

Acknowledgements

I am very grateful to have found a field within physics that I find so beautiful and this master thesis opportunity certainly was the cherry on top. I would like to thank my direct supervisors for their commitment and involvement. Zoltán, thank you for your guidance and supervision. Your impeccable insight in problems never failed to amaze me and your patience in explaining concepts was very admirable. Steven, thank you for always making time for me at the end of the afternoon. Your advice helped me to see clinical relevance and practical implications of my work. Ben, thank you for always helping me see the bigger picture.

Furthermore, I would like to thank Linda for helping me with very practical challenges that I have encountered working with scripts and data and Sebastiaan for helping me getting calculations done.

Contents

Abstract

Radiotherapy is one of the primary modalities for treating cancer of the prostate. The most common radiotherapy technique for treating prostate cancer is photon external beam radiotherapy. In the last two decades low values for the α/β ratio for prostate cancer have been reported. The α/β for the organs at risk (OAR) surrounding the prostate is postulated to be larger. The low α/β ratio for the prostate and the high α/β ratio for the surround organs suggest that hypofractionated treatments with large fraction size should spare the OARs better while delivering an isoeffective dose to the prostate compared to conventionally fractionated radiotherapy.

For hypofractionated treatments the propagation of the systematic and random error in treatment dose parameters is less known and needs to be quantified. No commonly used margin recipe exist that gives an indication what the needed margin is given a systematic and random error. The validity of the well known linear van Herk recipe to hypofractionated treatment is clearly debatable as it assumes a large number of treatment fractions.

In this work Polynomial Chaos Expansion was used to model dose distributions. Polynomial chaos methods approximate the dose by sets of polynomials that are functions of the errors involved in radiotherapy. This meta-modelling approach makes the costly evaluation of the clinical dose engine obsolete and allows fast computation of dose distributions. Polynomial Chaos Expansion have been shown to model the dose distribution delivered by VMAT and Cyberknife accurately in prostate cancer treatment.

With the use of the meta-models fractionation effects and effects of setup errors on the planning target volume, clinical target volume and three organs at risk have been studied for treatment plans of Cyberknife prostate cancer patients. The treatment plans prescribe a hypofractionated treatment in 4 fractions of 9.5 Gy that spare the urethra. The near minimum dose D_{98} to the prostate and the near maximum dose D_2 to the urethra, bladder and rectum have been determined under various setup errors for various fractionation regimes ranging from hypofractionation to conventional hyperfractionation.

It was found that the 98th percentile of the near maximum dose to the organs at risk increases for increasing systematic and random error and for decreasing fraction number. For large fraction numbers an increase in the 2nd percentile of the near minimum dose to the prostate has been observed for increasing random error for systematic errors that are not too large. The increase in the second percentile of the near minimum dose to the CTV is believed to be caused by dose blurring effects in the urethra that give rise to high urethra doses.

Polynomial Chaos Expansion has been used to construct margin recipes for various fractionation regimes. The margin recipes dictate what setup errors ensure that at least 98% of the simulated population receives for at least 98% of the CTV the full prescription dose for a CTV-PTV margin M_{PTV} of $3, 4$ or 5 mm. The margin recipes were found to be highly non linear and strongly dependent on the fraction number. The margin recipes are given by:

$$\Sigma(\sigma, M_{PTV}) = \frac{P_1(M_{PTV}) * \sigma + P_2(M_{PTV})}{\sigma^2 + P_3(M_{PTV}) * \sigma + P_4(M_{PTV})} \tag{1}$$

where P_1, P_2, P_3 and P_4 are polynomials given in Equations 36 to 39 with coefficients tabulated in Table 7. The margin recipes have been tested on ten patients and were shown to be valid for eight out of ten patients.

Finally, the constructed margin recipes are compared to the linear simplified van Herk recipe. It was found that the van Herk recipe is not valid for the investigated hypofractionated urethra sparing dose distribution.

1 Introduction

1.1 Radiation Therapy for prostate cancer

Prostate cancer is the development of cancer in the prostate, a gland in the male reproductive system. It is the most common cancer among men. In the Netherlands the majority of prostate cancer diagnosis are among men aged 65 to 79 years old [1]. In many cases thanks to modern technology it can be treated successfully, the five year survival rate is 90%, the mortality increases with increasing age of diagnosis [2]. Localised prostate tumours can be treated surgically (radical prostatectomy), with external beam radiation therapy, with brachytherapy or left untreated as it is a slow developing type of cancer.

Radiotherapy is one of the primary modalities for treating cancer of the prostate. The most common radiotherapy technique for treating prostate cancer is external beam radiotherapy, the latter will be the focus of this work. External beam therapy is is a radiation therapy modality that locally treats a patient. Before a patient can be treated many preparatory steps have to be undertaken. After the diagnosis of the cancer the tumour has to be biomedically imaged, delineated and a treatment plan for the treatment has to be made. The delivered radiation dose is planned carefully to optimise the success of the treatment.

1.2 Radiotherapy treatment planning and dealing with uncertainties

The goal of radiotherapy treatment planning is to design a beam configuration which will deliver a homogeneous dose to the specified planning target volume (PTV), ensuring that normal tissue receives a reasonably low dose and that critical organs receive less than their tolerance doses [3]. Optimisation in beam energy, beam arrangement, beam compensation or conformal methods have to be performed. Conformal methods shape the radiation dose distribution to the tumour morphology which allows it to deliver higher doses of radiation than standard dose conventional external beam RT [4].

Unfortunately, one also has to deal with uncertainties in radiotherapy, these are taken into consideration in making a treatment plan. Correctly dealing with uncertainties of RT is a crucial aspect in successful treatment outcome. Usually a dose of radiation is divided into several, smaller doses over a period of several days, to minimise toxic effects on healthy normal tissues [5]. A typical hyperfractionated radiation dose is divided into 37 units delivered every weekday [6]. Dose escalation by hypofractionation, i.e. delivering higher doses of radiation in fewer fractions, in prostate cancer has been increasingly popular as a result of relative new findings in the radiobiology of the prostate [7].

In the delivery of a fractionated dose two types of uncertainties can be distinguished, treatment execution uncertainties of a fraction, often called random variations, and treatment preparation uncertainties, often called systematic variations. Random variations vary from day to day whereas systematic uncertainties are systematic for a single radiotherapy course of a single patient, but they are stochastic over a group of patients [5].

The conventional way to deal with these uncertainties is to apply margins to clinically defined volumes to be irradiated. A well-established van Herk margin recipe [8] can provide the margin that needs to be used for given errors in patient positioning. However, this recipe has been derived for traditional hyperfractionated RT treatment regimens, typically consisting of 35-45 treatment fractions. For hypofractionated treatment no margin recipe has been established.

1.3 Goals of this research

Previous work [9, 10] was performed for external intensity modulated proton therapy (IMPT). For proton therapy both the errors in patient setup and proton range could result in no treatment to parts of the tumour as proton dose distributions are highly conformal due to the shape of the proton Bragg peak. Therefore, the robustness of treatment plans is crucial in successfully treating patients in IMPT. Conventional methods in photon therapy to cope with uncertainties cannot be applied to proton therapy due to the differences in dose deposition, therefore, robust treatment planning was introduced.

Treatment plans obtained by robust treatment planning include worst case scenarios of setup errors in the plan optimisation. For each iteration of the optimisation, nine different dose distributions are computed. These dose distribution each correspond to the dose for positive and negative setup uncertainties along the three patient axis, i.e. anteroposterior, lateral and superior-inferior, for positive and negative range uncertainty and the nominal dose distribution [11]. In previous work [9, 12] robustness recipes were derived that indicated the required setup robustness setting that should be used for certain systematic and random errors.

Setup robustness analysis was performed using a method called Polynomial Chaos Expansion (PCE), it was proven to be accurate and fast [10]. Polynomial chaos constructs a meta-model of the delivered dose by replacing the exact model with a set of polynomials that allow fast evaluations. During the course of this master thesis this method will be extended to Stereotactic Body Radiation therapy (SBRT), a RT modality that utilises a low number of fractions. For treatments conducted using a linear accelerator range errors are less relevant as photon depth dose distributions do not show steep dose fall off. However, SBRT is characterized by sharp dose gradients and high doses per fraction, thus small geometric errors in a single fraction can affect the dose considerably.

This work has two main focuses. First the effects of fractionation on the delivered dose are investigated and second, the required margins that deal with certain setup errors are derived. The latter will be achieved by constructing a margin recipe, which prescribes the required margins for known values of the random and systematic variations.

1.4 Structure of this work

As basis for the motivation of this research an introduction in key concepts in radiobiology is presented in Chapter 2. It provides the rationale behind fractionated radiotherapy, introduces the errors involved in radiotherapy and a way to deal with these errors and it introduces two RT modalities that are relevant to this work. Chapter 3 touches shortly on probability theory and discusses the theory behind Polynomial Chaos Expansion and steps involved in its construction. Treatment planning and meta-modelling treatment plans is discussed shortly after in Chapter 4. The steps involved in constructing Polynomial Chaos Expansion meta-models for RT are explained more thoroughly and ways to compare the resulting meta-model to the clinical treatment plan are introduced. Chapter 5 described how the obtained meta-models are used to study fractionation effects and how margin recipes are derived. The results are presented in Chapter 6, the analysis on fractionation effects on the dose in the prostate and adjacent critical organs is presented first and thereafter the constructed margin recipes are shown. Finally, in Chapter 7 the results are discussed and suggestions for future research are discussed.

2 Radiotherapy

Radiotherapy is the clinical use of ionising radiation as part of palliative or curative treatment of malignant tissues. It is used for various types of cancer, the subspecialty concerned with prescribing radiotherapy doses is called radiation oncology. Both photon and particle beams can be used to kill malignant cells. Radiation could arise from external beams and from internally implanted radio sources. In the scope of this research, only external photon beams are considered for prostate cancer (PCa) treatment.

Often the prescribed radiation dose is given in multiple treatments, so called *fractions*, of low doses. Conventionally the dose is hyperfractionated with a prescription of 38 to 40 fractions of 2 Gy, but more recently published research has found strong indications that a hypofractionation scheme of less than five fractions might be beneficial in PCa treatment. This section provides basic insights behind the rationale of fractionated radiotherapy [6].

2.1 Radiobiology

Photons can induce biological damage on tissues in a direct and indirect way. Direct interaction of photons with tissues causes ionisation and excitation of its constituent atoms. This could lead to a chain of physical and chemical interactions that eventually cause biological damage. In the indirect way incident photons deposit their energy via interactions with electrons due to the photoelectric effect, Compton scattering, and pair production. The electrons deposit energy on the target while passing through it [13].

Energy deposition to the tissues' molecules could result in free radicals. Water is the most prevalent molecule within the cell as about 80% of a cell is composed of water. Most of the free radicals are produced by the radiolysis of water. Free radicals are unstable as they possess unpaired valence electron(s). These highly reactive radicals can diffuse through the cell causing damage to parts of the cell. All components of the cell will be damaged in this way: proteins, enzymes, membrane components but damage to a cell's DNA has the most impact to the viability of a cell. Direct evidence that DNA damage is a critical event for cell viability has been established by experiments where short range Auger-emitting isotopes were substituted into DNA. It has been shown that these isotopes were far more toxic than when the same type and amount of radioactivity was substituted in other parts of the cell [14].

2.1.1 Cell survival curves

The radiation dose is a determining factor of success in clinical RT. In principle any tumour can be controlled if the dose is sufficiently high, but these high doses are not possible to give as it would be very toxic or lethal for the patient. A dose that is too low on the other hand may not cure the patient at all.

From published in-vitro data for different irradiated tumour cell lines surviving fractions are known. A curve can be fitted to these data points, these curves are known as cell survival curves. A typical survival curves is shown in Figure 1. This curve is a linear quadratic curve for the surviving fraction SF:

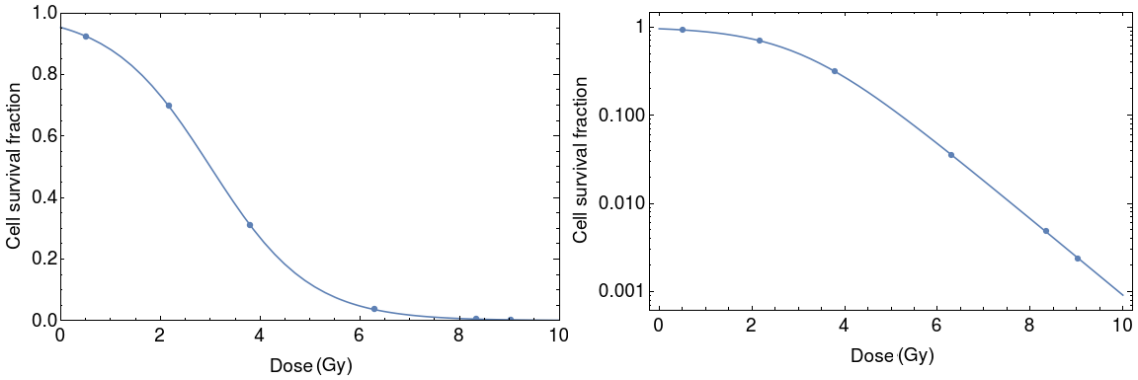
$$SF = e^{-\alpha D - \beta D^2} \tag{2}$$

where the radiation dose is denoted by D and α and β are fit parameters. Equation 2 is known as the linear quadratic (LQ) model. A linear alpha component and a quadratic beta component can be distinguished. The alpha and beta parameter can be calculated for different cell lines.

2.1.2 The 5Rs of radiotherapy

The biological factors that influence the response of normal and benign/malignant tissues to fractionated radiotherapy can be summarised as the *5Rs of radiotherapy*:

- **Repair:** cellular recovery hours post exposure.
- **Reassortment:** progression of the cell cycle.



(a) Linear axis

(b) Logarithmic axis

Figure 1: An example of a typical cell survival curve on linear axis (a) and logarithmic axis (b).

- **Repopulation:** proliferation of tumour cells that survived.
- **Reoxygenation:** increase of oxygen in cells that were hypoxic during treatment. and survived
- **Radiosensitivity:** of a certain type of tissue to a certain fractionation scheme

These factors will be discussed in Sections 2.1.2.1 and 2.1.2.2.

2.1.2.1 Repair, repopulation and reassortment of normal tissue

Different classifications of radiation damage are distinguished:

- Lethal damage, which is irreversible, irreparable and leads to cell death
- Sublethal damage, which can be repaired in hours unless additional sublethal damage is added that eventually leads to lethal damage;
- Potentially lethal damage, which can be repaired under certain conditions.

Cells that are exposed to low doses of radiation experience repairable sublethal damage. Most tissue repair occurs in about 3 hours and up to 24 hours after receiving radiation. For various tissue types, the radiation dose to produce damage and the timing of the expression of damage varies greatly.

Through repair of sublethal damage between dose fractions and repopulation of cells normal tissue is believed to have a therapeutic advantage over tumour cells in hyperfractionated RT regimes. A balance must be achieved between the response of tumour and early and late responding normal tissues such that a fraction dose spares late reacting tissues and the time between doses allow for regeneration of early responding tissues. It must also be noted that tumour cells may also show intra fraction repopulation [15].

2.1.2.2 Reoxygenation and radiosensitivity

Experiments have shown that if oxygen is present during radiation exposure the detrimental effects are enhanced. The levels of oxygen differ among cells within tissue due to local differences in blood flow and pressure. Cells that have comparatively high oxygen levels are called oxic whereas cells that have comparatively low oxygen levels, are called hypoxic.

The oxygen effect during radiation exposure can be quantified through the oxygen enhancement ratio (OER), which is simply the ratio of radiation doses for hypoxic cells compared to oxic cells for the same biological effect. It was found that oxygen does not need to be present during the irradiation to sensitise but could be added shortly afterwards [16].

Irradiation of a tumour will inevitably kill more oxic than hypoxic cells. After just a single large radiation dose, the hypoxic fraction may approach 100% as oxic cells are more radiosensitive leaving

parts of the tumour thus untreated. Reoxygenation refers to the process by which these hypoxic cells become better supplied with oxygen. Fractionated treatments allow cells to reoxygenate, cells that were hypoxic at the time of one radiation-dose fraction may be oxic by the time of a subsequent dose fraction. For inter-fraction times that are sufficiently long the presence of hypoxia may be compensated for [17].

2.1.3 Fractionation and biological effective dose

Intrinsic radiosensitivity differences for tissue types rates to the different components of that tissue. Tissues can be roughly divided into two categories, early and late responding tissues. Early-responding tissues show the effects of radiation damage within a few weeks of being irradiated, examples include skin, intestines and bone marrow. Late-responding tissues expresses the response to radiation damage in months to years post exposure, examples include lung, kidney, and spinal cord. The early effects arise due to damage to cells that have a short functional life span. Damage to late responding tissues may often be the result of damage to connective tissues, especially blood vessels. Repair of sublethal damage is greater for late responding tissues, the repopulation of cells is greater for early responding tissues.

The differences can be quantified via the α/β ratio. For fractionated doses Equation 2 can be rewritten as

$$SF_{kd} = e^{-\alpha BED} \tag{3}$$

where k denotes the fraction number such that the total dose is given by $D = k * d$. By doing this, a new quantity is introduced:

$$BED = D(1 + \frac{d}{\alpha/\beta}) \tag{4}$$

called the biological effective dose (BED) [18].

All fractionated treatments with equal BED have equal biological effect. The α/β ratio determines the fractionation sensitivity of the irradiated tissue.

Late-responding normal tissues show greater changes in sensitivity in response to a change in dose per fraction than early-responding tissues. Early responding tissues are more sensitive to fractionation which corresponds to a lower α/β ratio [15]. Most tumours have an α/β ratio larger than for normal tissue. The practical implication of this is that in cases where a large volume of normal tissue will be exposed to radiation, small doses per fraction are used to reduce normal tissue complications. If the α/β ratio of the tumour is equal or even lower than the α/β ratio of surrounding normal tissue, larger doses per fraction are more beneficial [19].

2.1.4 Fractionation of radiotherapy for prostate cancer treatment

In the last two decades comparatively low values for the α/β ratio for prostate cancer have been reported [7]. The low α/β values suggest a greater sensitivity to increasing fraction size. The α/β for the organs at risk (OAR) surrounding the prostate is postulated to be larger. These findings together with the linear quadratic formalism suggest an improvement of the therapeutic ratio can be achieved by larger fraction doses. Hypofractionated treatments with large fraction size should therefore theoretically spare the OARs better while delivering an isoeffective dose to the prostate compared to conventional RT [20].

2.2 Target definition and errors inherent to radiotherapy

Sophisticated advances in computational and biomedical engineering make it possible to highly conform doses. It is of high importance for the patients' health that toxic doses are exclusively delivered at the tumour site, sparing surrounding organs. Definition of the tumour and adjacent organs at risk is therefore an essential part of treatment planning process to ensure that beam properties such as size, number, angle and weighting are optimised.

The International Commission on Radiation Units and Measurements (ICRU) 62 report stipulates standard protocols for recording and reporting radiotherapy treatments [21]. Clinical volume definition and related uncertainties will be discussed in Section 2.2.1.

2.2.1 Target definition

The volumes defined in the ICRU 62 report are summarised in Figure 2. A brief description of the structures can be found in this subsection.

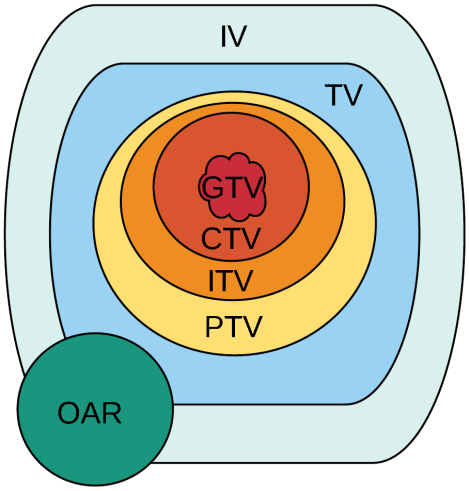


Figure 2: Target definition as defined by ICRU 62 [21].

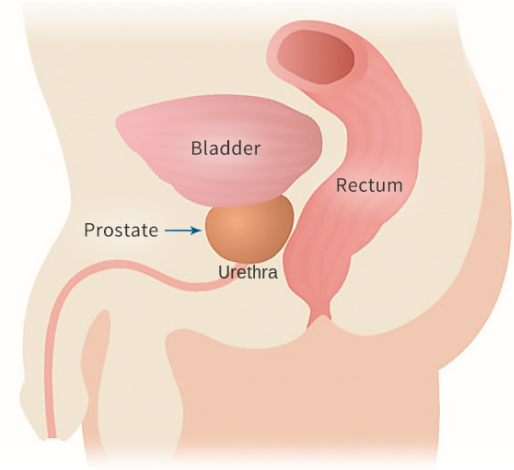


Figure 3: A schematic overview of the OARS involved in PCa treatment

Gross tumour volume (GTV)
The GTV encompasses the gross, palpable, visible or clinically demonstrable location and extent of the malignant growth. The GTV is delineated by the oncologist after thorough examination including 2D or 3D imaging.

Clinical target volume (CTV)
The CTV consist of the GTV and a margin that takes into account suspected subclinical malignant disease. The additional margin accounts for possible microscopic extension of the primary tumour or regional lymph node spread. Since it is not possible to determine the degree of microscopic spread around tumour non-invasively, the CTV may not be defined separately but considered when defining the planning target volume (PTV).

Internal target volume (ITV)
The ITV consists of the CTV plus an internal margin that takes into account the variations in the size and position of the CTV relative to the patient's bony anatomy. These variations could be due to organ motions and bladder or bowel/rectal filling.

Planning target volume (PTV)
The PTV is the geometrical volume that encompasses the CTV with an additional safety margin to compensate for the different types of variations and uncertainties of the dose distribution relative to the CTV. The PTV is constructed by adding a concentric margin, determined from uncertainty analysis, around the GTV/CTV. The margin is a compromise between ensuring sufficient CTV coverage and limiting toxic doses to the OARs. These margins are different across types of tumour and tumour sites and could for certain tumour sites be anisotropic.

Organ at risk
Organs at risk are organs adjacent to the PTV. These organs, as they do not contain malignant cells, should ideally receive zero dose. In practice OARs are spared as much as possible during treatment planning but receive dose. Damage to these organs may lead to substantial toxicity and morbidity. The most important OARs for prostate tumours are the bladder, rectum and urethra, the critical organs that are involved in treating PCa patients, are depicted in Figure 3.

Many OARs have a defined tolerance dose that should not be exceeded to avoid late radiation morbidity. These tolerance levels are taken into account during treatment planning. Some OARs demonstrate a volume effect, meaning that an increasing loss of function is observed for an increasing irradiated organ volume. Other organs, often organs that have a serial structure, for example the spinal cord, should not receive a high dose even for a small volume as these organs could demonstrate serious functional loss if one of its sub-volumes is damaged.

Treated volume (TV) and Irradiated Volume (IV)
The TV and is the volume of tissue enclosed by an isodose surface selected and specified by the clinician as being appropriate to achieve the aim of treatment. The TV should not be significantly larger than the PTV as treatment planning on the TV will then ensure PTV coverage whilst sparing surrounding OARs. The IV encompasses all tissue volume that receives a radiation dose that is considered significant in relation to normal tissue tolerance [3].

2.2.2 Errors inherent to radiotherapy

Conformal radiotherapy has been widely used for treatment of the prostate. Identifying and quantifying uncertainties in RT is a crucial in steering the treatment outcome as target volumes are defined by the magnitude of the uncertainties. We distinguish systematic and random errors in the setup of the patient. Gross errors that can be caught with standard quality assurance protocols are not considered in this work.

Delineation of the tumour and the ICRU volumes is a crucial part of treatment planning. Although defining these volumes is performed based on high resolution images (Magnetic resonance imaging, computer tomography, etc.), uncertainties inherent to biology, patient setup and the treatment delivery site exists. For some tumours the tumour boundary may not be clear and the GTV may be hard to define. Some tumours peripheries may be poorly defined because of diffuse infiltration or because the tumour has a similar radiographic density compared to the surrounding normal tissue making it hard to resolve.

Deviations at the treatment delivery site could arise from geometric errors, these

(a) Dose blurring by the random error (b) Dose shifting by the systematic error

Figure 4: This figure shows the difference in effect the different setup errors have. The random error blurs the dose whereas the systematic error causes an isocenter offset between the CTV and the dose distribution.

could include deviations in laser alignment, data transfer from treatment planning to the linear accelerator, couch position, image resolution, margin expansion algorithms, multi leaf collimator (MLC) position and sequence and collimator angle accuracy.

Systematic patient setup errors could arise from displacement of (internal) markers with respect to the planning images or treatment execution on a different couch over the course of a treatment. The systematic error is assumed to be constant over a treatment course and therefore patient specific. Over a population of patients the systematic error is assumed to be Gaussian distributed.

The random error is a deviation that can vary in direction and magnitude for each delivered treatment fraction. Random errors could arise from organ motion and patient setup variations for every fraction. The prostate is a moderately fixed organ, its position is dependent on bladder and rectal filling and could differ for every fraction [22].

Random and systematic errors affect the dose in different ways. Random errors 'blur' the dose distribution, making the edges less sharp. The blurring of the dose distribution by the random error can be quite accurately described as a convolution of the dose distribution with the probability distribution function of the random error [23]. Systematic errors simply cause a shift of the cumulative dose distribution relative to the target, as illustrated in Figure 4.

The standard deviation of the systematic error and the random error is denoted by vectors. The systematic error is denoted by $\vec{\Sigma} = (\Sigma_x, \Sigma_y, \Sigma_z)$ and the standard deviation of the random error is denoted by $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. The subscripts x, y, and z represent the standard deviation along the three axes of the Cartesian coordinate system.

2.2.3 Margin recipe for hyperfractionated RT

Van Herk et al. [8] published an article in which the different effects of the systematic and random error on the target dose were analytically described to derive a margin. The recipe states that required margin is dependent on the standard deviations of the systematic and random error and the penumbra width σ_p, the latter is defined as the distance between the 95% and 50% isodose surface of the planned dose distribution. The margin recipe for isotropic margins reads:

$$M_{PTV} = \mu\Sigma + \nu\sigma - \nu\sigma_p \tag{5}$$

where μ and ν are numerical constants that depend on the dimension of the dose distribution and patient confidence levels. In this recipe the standard deviation of the systematic and random error are vectors allowing a non-isotropic margin that can handle non-isotropic geometric uncertainties and. Van Herk et al. also provide a simplification to this recipe that reads:

$$M_{PTV} = \alpha\Sigma + \gamma\sigma' \tag{6}$$

$$\sigma' = \sqrt{\sigma_m^2 + \sigma_m^2} \tag{7}$$

where the combination standard deviation σ' is a combination of the standard deviation of all random variations including both setup and organ motion. Assuming that the dose distribution is three-dimensional, with the clinically acceptable criterion set at a minimum dose to the CTV of 95% of the prescribed dose for 90% of the patient and that $\sigma_p = 3.2$ mm the margin recipe is simplified to:

$$M_{PTV} = 2.5\Sigma + 0.7\sigma' \tag{8}$$

which is approximately valid for random errors up to 5 mm. The simplicity of this margin recipe comes at a cost in terms of its validity. Simplifications and assumption were introduced in the derivation of this recipe, these include:

1. The patient population was assumed to have the same standard deviation for the setup errors
2. The dose is delivered in many fractions making the average random error zero.
3. The irradiated target was assumed to be spherically symmetric
4. Rotations and shape variations of the tumour were neglected
5. Errors were assumed to be isotropic
6. Different sources of errors were taken statistically independent
7. The considered errors were assumed to be distributed normally

2.3 Hypofractionated RT: Stereotactic body radiotherapy

The relevant RT modality in this work is Stereotactic body radiotherapy. SBRT is a radiotherapy modality in which low number of fractions is used (and therefore a much higher dose per fraction) compared to conventional RT. Multiple external photon beams of various intensities aimed from different angles deliver dose to the tumour site.

The goal is to conform the dose to the tumour morphology as much as possible, this could be done with the aid of compensating wedges, collimators, pencil beams etc., which allows higher doses to be delivered at once. The term stereotactic refers to three-dimensional localisation of a particular point in space by a unique set of coordinates that relate to a fixed, external reference frame.

Two SBRT modalities that deliver doses in hypofractionated schemes are considered in this work: Volumetric Modulated Arc Therapy (VMAT) and Cyberknife.

2.3.1 Volumetric Modulated Arc Therapy

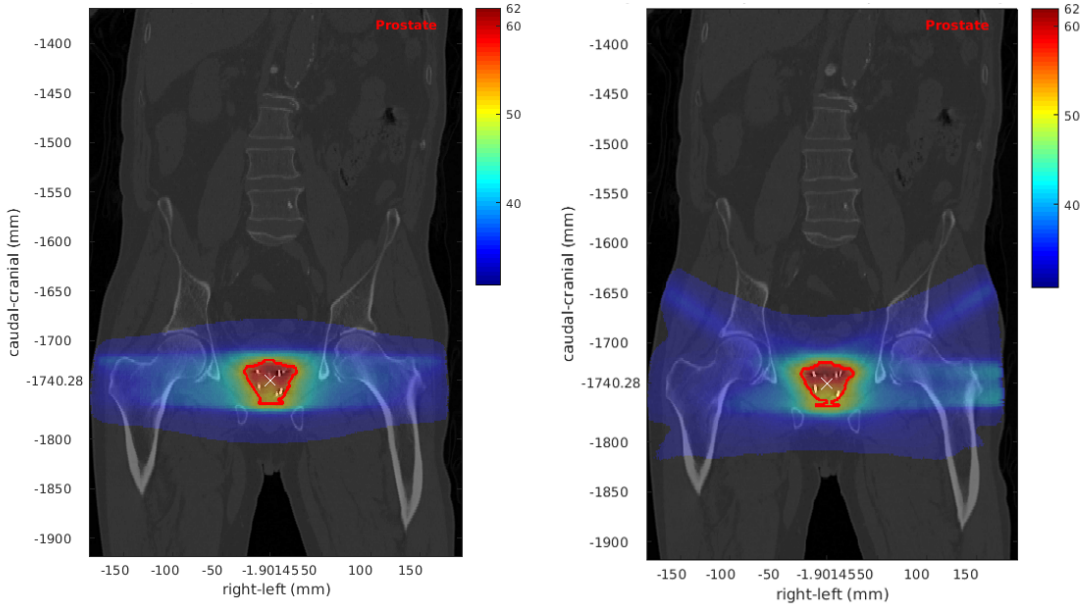
Over the past few decades radiotherapy delivery has been subjected to leaps of improvement in terms of accuracy. Sophisticated imaging techniques have led to improved target definition and delineation, precise linear accelerators for clinical use have been developed, advances in treatment planning systems have reduced normal tissue dose and many beam shaping techniques have been introduced.

These advances in technology gave rise to intensity modulated radiotherapy (IMRT) techniques in which variable intensity across multiple radiation beams leads to the construction of highly conformal dose distributions. In IMRT multiple beams are used, the radiation fluence per beam is delivered by multiple beamlets that each have an individual intensity which allows higher target volume conformity than conventional radiotherapy, particularly in volumes with complex concave shapes. A downside of IMRT is that it can come with an increase in the amount of low dose radiation to the rest of the body compared with conventional conformal RT plans.

Arc based therapies, including volumetric modulated arc therapy (VMAT), have been of interests to tackle this problem. Arc therapy is essentially an alternative form of IMRT, where radiation is delivered from a continuously rotating radiation source. Patients can be treated from a large range of beam angles which allows highly conformal dose distributions, a reduction in treatment delivery time and possibly a reduction of the integral radiation dose to the rest of the body compared depending on the number of beams that IMRT utilises [24], meaning OARs sparing. VMAT utilises conventional linear accelerators but by varying the gantry rotation speed, treatment aperture shape via movement of multi leaf collimator leaves and the dose rate, conformal distributions can be delivered [25].

2.3.2 Cyberknife

The Cyberknife (developed by Accuray) is a frameless stereotactic radiosurgery system, meaning that it does not require patients to be fixated by a frame. The Cyberknife characterised by image guided manipulation of a high-energy linear accelerator by a robotic arm. A Control loop between the imaging and beam delivery systems allows the Cyberknife beams to follow a moving target real-time within a patient. This makes rigid fixation of the patient obsolete.



(a) VMAT treatment plan dose distribution with $D_{mean}(Prostate) = 48.7$ Gy, $D_{mean}(PTV_{3mm}) = 48$gray.

(b) Cyberknife treatment plan dose distribution with $D_{mean}(Prostate) = 48.7$ Gy, $D_{mean}(PTV_{3mm}) = 45.4$ Gy.

Figure 5: This figure shows a comparison between a VMAT and a Cyberknife plan for a PCa patient.

Continuous image guidance makes the Cyberknife system capable to real-time check the target position and adjust the treatment accordingly if the target moves during a treatment. Frameless target positioning utilises the patient's bony anatomy or implanted markers but plans are made based on planning images. Real-time imaging during treatment is provided by X-ray devices. The images are registered to translate the position of the treatment site to the coordinate frame of the linear accelerator to assure accurate targeting.

The high degree of mobility of the robotic arm allows for non isocentric and non coplanar dose delivery resulting in conformity in the delivered dose distributions. Targets of many shapes can be conformed using this technique without the use of multiple overlapping isocenters [26].

Figure 5 shows a comparison between a coplanar VMAT plan and a non coplanar Cyberknife plan for a particular patient. A larger normal tissue volume is irradiated if the patient is treated with Cyberknife.

3 Uncertainties and uncertainty propagation

In Chapter 3 fractionation, different errors inherent to radiotherapy and conventional ways to deal with these errors were introduced. In this section Polynomial Chaos will be introduced, which will be used to quantify the effects of uncertainties in RT. First some background in probability theory will be discussed to better understand the stochastic nature of uncertainties.

3.1 Introduction to probability theory

In probability theory the set, Θ, of all possible outcomes of a particular experiment is called the sample space of the experiment. Any collection of possible outcomes of an experiment, meaning any subset of $\theta \in \Theta$, is called an event.

Sampling i.e. conducting the experiment may lead to different outcomes or some outcomes may repeat. This "frequency of occurrence" of an outcome can be thought of as a probability. For each event θ in the sample space Θ a number between zero and one can be associated with θ that will be called the probability of θ, denoted by $p(\theta)$, meaning for each $\theta \in \Theta$ it gives the probability that θ occurs [27].

3.1.1 Random variables

In most cases it is more convenient to deal with a quantity of interest than the entire sample space. In probability theory situations can be described for which precise values of variables are unknown, but the variables are expected to be distributed in some way. These variables are referred to as a random variables. It is a function that maps from a sample space Θ into the real numbers. The likelihood that the random variable will have a certain value is called the probability.

Variables that can take on an uncountably infinite number of possible outcomes are called continuous random variables, these are described by their probability density function (PDF). The mean and variance of a random variable can be calculated using the PDF. The mean μ_X and variance σ_X^2 for a random variable X and its probability density $P(x)$ are given in Equation 9 and 10

$$\mu_X = \int_{-\infty}^{\infty} xP(x)dx \tag{9}$$

$$\sigma_X^2 = \int_{-\infty}^{\infty} P(x)(x-\mu_X)^2 dx \tag{10}$$

$$\tag{11}$$

Probability models often involve more than one random variable. A random vector consisting of several random variables can be defined similarly to its univariate counterpart. An n-dimensional random vector is a function that maps from a sample space into and N-dimensional Euclidean space $\mathbb{R}^n$. The random vector is described by a multivariate joint probability density function [28].

3.1.2 Normal distribution

A very popular continuous probability distribution for random real valued continuous variables is the normal or Gaussian distribution, which is a symmetric bell

shaped distribution making it very suitable for population models. It has the advantage that it is very tractable analytically. Complete information about the exact shape and location of the distribution are given by its only two parameters, namely the mean and the variance. The PDF of the normal distribution is given by:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} \tag{12}$$

A variable with a Gaussian distribution is notated as $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$.

Among the many uses of the normal distribution, an important one is its use as an approximation to other distributions (which can partially be justified by the Central Limit Theorem) [28].

3.2 Spectral methods for stochastic quantities

A well established method to quantify the effects of uncertainties on a system are Monte Carlo methods. These methods rely on large amounts of pseudo-random sampling, by the law of large numbers properties it can estimate the expectation value and variance of the response of a system. When the number of sampling points is sufficiently large the expectation value and variance obtained from Monte Carlo sampling are close to the actual mean and variance of the system.

Samples are selected randomly from the sample space, the probability density function that depends on the random vector dictates the probability of a sample point being chosen to obtain response of the system. This sampling should be performed on the entire sample space to determine the global variability to uncertainty. This simplicity comes at a cost of having a slow convergence rate with the number of realisations (M). The convergence of variance estimates behaves as $\frac{1}{\sqrt{M}}$.

Spectral methods are based on a different approach, these methods are based on constructing the dependence of the stochastic quantity realisations on the random vector. The idea is to write the dependence of a system on the random input variables as a sum of certain basis vectors and then to choose the coefficients in the sum in order to satisfy the system's response. The main feature of the spectral methods is to take various orthogonal systems of global functions as trial functions for different problems. Different trial functions lead to different spectral approximations [29].

3.3 Polynomial chaos expansion

Polynomial Chaos Expansion (PCE) is a spectral modelling approach to approximate stochastic model outputs. It is used to describe the output of a model in terms of its mean value, variance, etc. [30]. By doing so the goal is to investigate the variability of certain model parameters as a function of the uncertain variables characterising the modelled phenomenon. The input variables are assumed to be second order random variables, meaning that they have a finite variance.

3.3.1 Input variables

The goal is to model stochastic processes. To meet this end the responses of the modelled phenomenon are investigated as function of the uncertain parameters characterizing it, the uncertain parameters correspond to distinct sources of uncertainties.

The variability of the response is therefore assumed to be depended on the input variables that characterises the modelled phenomenon. The number of uncertain inputs directly determines the dimensionality of the model and they are assumed to belong to a certain probability density function.

We define Θ to be a sample space containing all possible random events $\theta \in \Theta$ and $\{\xi_i(\theta)\}_{i=1}^{d}$ to be the set of d random variables that can also be described by a random vector. In this work the random variables are assumed to be independent normalised Gaussian variables described by their Gaussian probability density function. The joint probability density function (PDF) $p(\vec{\xi})$ is in this case given in Equation 13 which is a product of all PDFs.

$$p(\vec{\xi}) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(\xi_i-\mu_i)^2}{2\sigma_i^2}} \tag{13}$$

The expansion of a stochastic response of interest $R(\theta)$ consists of basis vectors Ψ_k and expansion coefficients r_k:

$$R(\vec{\xi}(\theta)) = \sum_{k=1}^{\infty} r_k \Psi_k(\vec{\xi}), \qquad \vec{\xi} = (\xi_1, \xi_2, ..., \xi_d) \tag{14}$$

where conventionally $\Psi_0 = 1$ and d is the number of random inputs that are considered. In this work the stochastic response is real valued, it maps from the sample space into the real numbers. The responses are assumed to belong to the L_2 space of functions for which the integral of the square of the absolute value is finite.

As the number of input variables must be finite, the sum in Equation 14 is truncated at P . From Equation 14 it can be seen that PCE expansion resembles ordinary Fourier expansion.

3.3.2 Polynomial chaos basis vectors

The basis vectors in PCE are polynomials from the subspace of polynomials in all possible combinations of the random variables $\{\xi_i(\theta)\}_{i=1}^{d}$ up to a degree p. The choice of the polynomials is directly related to the probability distribution of each random input variable. For each input variable a univariate polynomial is selected based on the distribution of the input variable. They are chosen according to the PDF of the input random variable as specified in the Askey scheme [31]. Following this scheme for Gaussian distributed input variables, *probabilists' Hermite polynomials* are deemed most suitable for fast convergence. The probabilists' Hermite polynomials are given by:

$$He_n(\xi) = (-1)^n e^{\frac{\xi^2}{2}} \frac{d^n}{d\xi^n} e^{-\frac{\xi^2}{2}} \tag{15}$$

The Hermite polynomials $He_n(\xi)$ are a set of orthogonal polynomials over the domain $(-\infty, \infty)$, the first polynomials are illustrated in figure 6.

The space spanned by these polynomials is a subset of the L_2 space, called the *p-th Homogenous Chaos*. All polynomials are mutually orthogonal with respect to the joint PDF: [30]:

$$\langle \Psi_k, \Psi_l \rangle = \int \Psi_k(\vec{\xi}) \Psi_l(\vec{\xi}) p(\vec{\xi}) = \delta_{k,l} \langle \Psi_k, \Psi_l \rangle \qquad = h_k^2 \delta_{k,l} \tag{16}$$

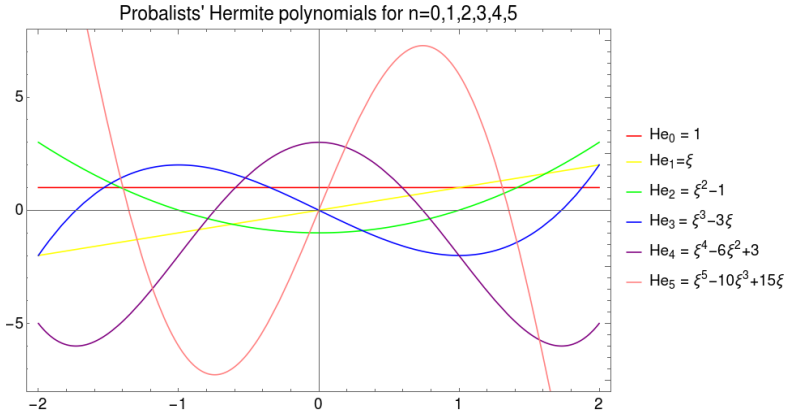


Figure 6: This figure shows the first six Hermite polynomials i.e.He_n for $n = 0, 1, 2, 3, 4, 5$

with h_i being the norm of Ψ_i. The PC basis vectors must form a Hilbert space such that the convergence of equation 14 is full filled.

Finally, the PC basis vectors Ψ_k are constructed by tensorisation of the univariate polynomials. Each PC basis vector can be characterized by a multi-index $\{\gamma_{k,j}\}_{j=1}^{N}$ which differentiate between the different polynomial families corresponding to the different random variables of different orders:

$$\Psi_k(\vec{\xi}) = \prod_{j=1}^{N} \psi_{j,\gamma_{k,j}}(\xi_j) \tag{17}$$

where the first index corresponds to the different random variables and the second index $\gamma_{k,j}$ corresponds to different orders in each variable. A traditional choice is to include all multi-dimensional polynomials having a combined order of at most PO. The number of basis vectors in the expansion in Equation 14 can be expressed by

$$P + 1 = \frac{(N + PO)!}{N!PO!} \tag{18}$$

To illustrate this, suppose $PO = 2$ and $N = 2$. For PO=2 there would be 6 basis vectors according to Equation 18. The multi indices of all six basis vectors would be $(0, 0)$, $(1, 0)$, $(0, 1)$, $(1, 1)$, $(2, 0)$ and $(0, 2)$. The number on the left and right in the bracket represent the order of the included univariate polynomial in that direction.

In practice many of the higher order multidimensional PC basis vectors can be excluded without affecting the accuracy of the PCE as not all input parameters are equally important and not all interactions between them must be included to reach desired accuracy. *The sparsity of effects principle* [32] dictates that responses are usually dominated by a few important parameters and most high-order interactions are negligible.

Full PCE basis sets can be avoided to reduce computational cost and memory without affecting the accuracy too much. Numerous ways have been introduced to select the basis vectors to be included in a sparse PCE [27].

3.3.3 Calculation of the PCE coefficients

Next the expansion coefficients r_k must be determined. By exploiting the orthogonality of the PC basis, the expansion coefficients can be determined using spectral

projection. This projection of Equation 14 with substitution of the definition of the L_2 inner product is given by [33, 34, 27]:

$$r_k = \frac{\langle R(\vec{\xi}), \Psi_k(\vec{\xi})\rangle}{\langle \Psi_k(\vec{\xi}), \Psi_k(\vec{\xi})\rangle} = \frac{\int_{\mathfrak{D}(\Theta)} R(\vec{\xi})\Psi_k(\vec{\xi})p_{\vec{\xi}}(\vec{\xi})d\vec{\xi}}{\langle \Psi_k(\vec{\xi}), \Psi_k(\vec{\xi})\rangle} \tag{19}$$

where the numerator is a multi-dimensional integral over the domain $\mathfrak{D}(\Theta)$ of the random variables describing the θ sample space. This integral in Equation 19 must be approximated as it contains the unknown dependence of the response $R(\vec{\xi})$ on the input parameters.

3.3.3.1 Quadratures

The integral is approximated numerically using a quadrature rule. A quadrature rule is a numerical approximation of the definite integral of a function. Many different quadrature rules are available in literature.

The quadrature approximation $Q^{(1)}$ of a one dimensional integral is simply a weighted summation over function evaluations with the subscript $^{(1)}$ denoting single dimensionality. The function $f(\xi)$ is evaluated on predefined integration points ξ_j, called quadrature points, weighted with predefined weighting coefficients w_j as expressed in Equation 20 [29].

$$\int_b^a p_\xi(\xi)f(\xi)d\xi \approx Q^{(1)}f = \sum_{j=1}^{n_{lev}} w_j f(\xi_j) \tag{20}$$

The quadrature points and weights depend on the chosen quadrature rule and the probability function p_ξ which also determines the accuracy of the approximation. With increasing quadrature level lev, more and more quadrature points are taken into consideration in approximating the integral, making the approximation more accurate. The level corresponds to a quadrature rule with $n_{lev} = 2 \cdot lev - 1$ quadrature points, the subscript $^{(j)}$ denotes the jth quadrature point and weight contained in the quadrature.

Many different quadrature rules are available in literature. In this work Gauss quadratures were used as the weighting coefficients and quadrature points can be chosen such that they are exact for Hermite polynomials up to order $2 \cdot n_{lev} - 1$ [35].

3.3.3.2 Cubatures

To approximate the multi-dimensional integral in Equation 19 a cubature formula can be used, which is constructed by tensorisation from one-dimensional quadratures. In each direction a quadrature rule of level lev is used which results in a summation over all possible combinations of quadrature points. If the independent random variables have different distributions, different quadratures rules can be used along the different integration directions [29]. For an N_{dim} integral the tensorisation is given by Equation 21

$$Q_{\underline{lev}}^{N_{dim}} f = \big(Q_{lev_1}^{(1)} \otimes ... \otimes Q_{lev_{N_{dim}}}^{(1)}\big) f \tag{21}$$

$$= \sum_{j_1=1}^{nlev_1} ... \sum_{j_{N_{dim}}=1}^{nlev_{N_{dim}}} f\big(\xi_{1,lev_1}^{(j_1)}, ..., \xi_{1,lev_{N_{dim}}}^{(j_{N_{dim}})}\big) w_{lev_1}^{j_1} ... w_{lev_{N_{dim}}}^{j_{N_{dim}}} \tag{22}$$

where $\underline{lev}$ is a multi-index indicating the different quadrature levels that are used for each direction of integration.

In each direction *dir* a quadrature rule of level lev_{dir} is used which results in a summation over all possible combination of quadrature points, that span a multidimensional grid of points, as expressed in Equation 21. The order of this grid GO, indicates the level of the quadrature points that that span the grid. The higher the grid order, the higher the level of quadratures constituting it and the higher the accuracy. Including quadratures with a higher level in the cubature also means that more function evaluations have to be performed.

3.3.4 Smolyak sparse grids

From Equation 21 one can observe that the total number of cubature points is $\prod_{dir=1}^{N_{dim}} lev_{dir}$. Hence the needed evaluations of the integrand grows exponentially with the dimension of the problem, this is the *curse of dimensionality*. This problem can be alleviated by leaving out high order integration grids that correspond to high order interaction of input variables [27]. If the high order interactions do not play an important role in the modelled phenomenon, a sparse grid that neglects high order interaction will result roughly the same accuracy as the full grid.

In this work *extended Smolyak sparse grids* will be used [36]. The Smolyak sparse grids are based on cubatures that use difference formulas of quadratures instead of the original quadratures in the construction. This difference is also a quadrature rule.

$$\Delta_{lev}^{(1)} f = Q_{lev}^{(1)} f - Q_{lev-1}^{(1)} f \tag{23}$$

$$Q_0^{(1)} \equiv 0 \tag{24}$$

The quadrature rules can be expressed as the sum of the difference formulas

$$Q_{lev}^{(1)} f = \sum_{l=1}^{lev} \Delta_l^{(1)} f \tag{25}$$

using Equation 25 the cubature formula in Equation 21 can be rewritten as

$$Q_{\underline{lev}}^{(N_{dim})} f = \sum_{l_1=1}^{lev_1} ... \sum_{l_{N_{dim}}=1}^{lev_{N_{dim}}} \left(\Delta_{l_1=1} \otimes ... \otimes \Delta_{l_{N_{dim}}}^{(1)}\right) f \tag{26}$$

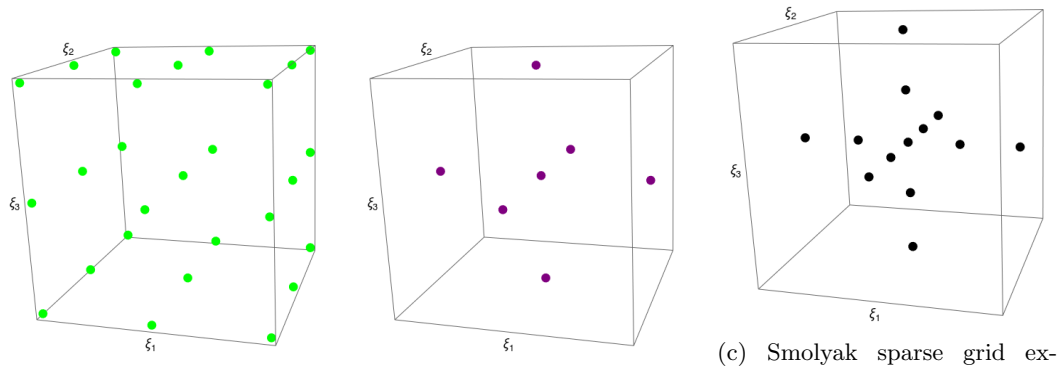
where $\underline{lev}$ again denotes different quadrature levels along different directions and l the different grids.

Sparse grids can significantly reduce the computational costs by having to evaluate the integrand at significantly less points. Using the sparse grids, the number of evaluation points and subsequently the computationally cost is greatly reduced.

3.3.5 Extended sparse grids

The sparse grids obtained in Section 3.3.4 can be extended to achieve higher accuracy achieved without adding too much extra computational burden. The extension of the grid is done by extending only the single dimensional grid, a higher quadrature level will then be used for this direction leaving the multi-dimensional grids unchanged.

Extended grids compared to conventional Smolyak sparse grids offer higher accuracy along the single dimensions. Higher order univariate polynomials can be included by doing this whilst the increase in calculation burden is minimal, as an extension of EL levels adds $2 \cdot N_{dim} \cdot EL$ extra quadrature points only.



(a) Full grid (b) Smolyak sparse grid (c) Smolyak sparse grid extended with one level

Figure 7: The comparison between a full grid, sparse grid and an extended sparse grid for a second order grid [36].

3.4 Hyperbolic trimming

In this work basis vectors that represent high order interactions are cut out a priori. When the dimensionality of the polynomial basis vectors is bigger than the grid order, the PCE coefficients can not be accurately determined. A *hyperbolic trimming* is applied in this work, only basis vectors that satisfy the quasi norm given in equation 27 are included. For decreasing q values higher order cross terms are penalized more, therefore more of them will get cut out. In the case $q = 1$ the full PC basis set will be used.

$$||\underline{\gamma_k}||_q \equiv \Big(\sum_{j=1}^{N} \gamma_{k,j}^q\Big)^{\frac{1}{q}} \leq GO \qquad q \in (0, 1] \tag{27}$$

The maximum grid order and polynomial order included in the PCE basis vectors affect the accuracy of the PCE. As the use of one extra extended level is shown to improve the accuracy at a computationally cheap cost [27], this will be used. The maximum polynomial order PO will be chosen such that it satisfies:

$$PO = GO + EL = GO + 1 \tag{28}$$

to ensure that the expansion coefficients can be determined accurately.

4 Treatment planning and robustness analysis

This section will first give a short introduction in treatment planning of the dose distributions that were modelled, then some background in quantitative comparison of dose distributions using different metrics and finally it will explain how PCE was used to compute the effects of uncertainties on different dose metrics and the construction of margin recipes.

4.1 Treatment planning

Treatment planning is a multidisciplinary process in which a team consisting of both medical clinical experts (i.e. oncologists, therapist etc.) and medical physicists plan an adequate radiotherapy treatment. In this thesis all treatment plans were planned for external beam therapy for PCa patients. Treatment planning aims at planning a dose distribution that delivers a prescribed dose in one or multiple fractions to the CTV that maximizes the tumour control probability whilst taking into consideration normal tissue complication.

Treatment planning is patient specific and gives a prescription of the to be delivered treatment and its parameters. During treatment planning the beam setup is determined, the number of beams, beam angles, isocenters and intensity/weight are determined through optimisation. All treatment plans were image guided plans, i.e. CT and MRI data were available, planned in an inverse manner.

In inverse planning the clinical volumes are delineated in image data and prescription doses and maximum tolerance levels are set. Through optimisation a treatment plan is constructed which best fits all input criteria. This could be done manually but also automatically. All plans in this project were generated automatically using Erasmus-iCycle: a novel algorithm for integrated, multicriterial optimization of beam angles, and intensity modulated radiotherapy (IMRT) profiles developed at Erasmus Medical Center [37].

4.1.1 Erasmus-iCycle

A multicriterial plan optimization in terms of beam orientations selection and beam profile optimisation is based on a prescription called wish-list. Wish-lists are constructed by the physician and contain pre-defined hard constraints and prioritised optimisation objectives. The higher an objective priority is, the higher the probability that the corresponding objective will be achieved during the optimisation process. The wish list for PCa patients in this work is shown in Figure 8.

Beam directions are selected from an input set of candidate directions, these sets could be coplanar or non-coplanar. Beams are added sequentially in an iterative procedure. Each iteration loop starts with the selection of an orientation to be added to the treatment plan. All candidate beam directions that have not been selected yet are evaluated one-by-one by solving for each of them an optimisation problem for beam arrangement consisting of the candidate plus the previously selected beam directions. The orientation with the most favourable scores is selected as the beam direction to be added for that particular iteration.

Then a pareto-optimal IMRT plan is generated for the beam setup of that iteration that includes all selected beam directions up to that iteration. For the next selection, all not yet selected candidate directions are temporarily added to the

	Structure	Min/Max	Type	Goal	Limit	Sufficient	Priority	Weight	Parameters	Active
4	Rectum_DummyStructure	Minimize (maximum) ⇩	linear	38			Constraint	1		Yes
5	UrethraPlan	Minimize (maximum) ⇩	EUD	39			Constraint	1	3	Yes
6	UrethraPlan	Minimize (maximum) ⇩	linear	50			Constraint	1		Yes
7	Bladder_DummyStructure	Minimize (maximum) ⇩	linear	41.8			Constraint	1		Yes
8	Bladder	Minimize (maximum) ⇩	linear	39.5			Constraint	1		Yes
9	Bladder_DummyStructure2	Minimize (maximum) ⇩	linear	41.8			Constraint	1		Yes
10	PS	Minimize (maximum) ⇩	linear	1.5			Constraint	1		Yes
11	PTV4Planning	Minimize (maximum) ⇩	LTCP	0.1084		0.1084	1	1	37 0.9	Yes
12	PTV4Planning	Minimize (maximum) ⇩	LTCP	2.2		2.2	2	1	57 0.07	Yes
13	Rectum	Minimize (maximum) ⇩	LTCP	0			4	1	28 -0.3	Yes
14	Prostate	Maximize (minimum) ⇧	linear	34		34	3	1		Yes
15	Bladder	Minimize (maximum) ⇩	LTCP	0			5	1	34 -0.1	Yes
16	PTV	Maximize (minimum) ⇧	linear	30			8	1		Yes
17	Rectum	Minimize (maximum) ⇩	mean	0			6	1		Yes
18	Urethra	Minimize (maximum) ⇩	mean	0			9	1		Yes
19	Bladder	Minimize (maximum) ⇩	mean	0			7	1		Yes
20	from3cm-toexternal-2cm	Minimize (maximum) ⇩	linear	15			10	1		Yes
21	Right Femoral Head	Minimize (maximum) ⇩	linear	24			12	1		Yes
22	Left Femoral Head	Minimize (maximum) ⇩	linear	24			11	1		Yes

Figure 8: This figure shows the wishlist that was used to optimise the plans for a PCa patient treated with Cyberknife.

plan, an optimisation problem, derived from the Lagrangian that is obtained with the pareto-optimal plan, is solved. This iteration continuous until addition of beams does no longer result in significant plan quality improvement [37].

4.1.2 Treatment plans for SBRT prostate patients

For malignancies in the prostate it is the usual to treat the whole organ making the entire prostate the CTV. The primary tumour is hard to resolve from surrounding tissue. The PTV is obtained from the CTV by expanding the CTV isotropically with a certain margin.

In this project we looked at two types of treatment plans: coplanar VMAT treatments and non-coplanar Cyberknife treatments. For the coplanar VMAT plans only treatment plans with 3 mm CTV to PTV margins were considered. For the Cyberknife plans chosen margins were 3, 4 and 5 mm. Treatment plans for the Cyberknife were optimised using the same wish list for CTV to PTV margins of 3, 4 and 5 mm. The Cyberknife treatment plans were optimised for a CTV to PTV margin of 3 mm, the same beam configuration was used for all other margins.

The organs at risk surrounding the prostate are the bladder which lies below the prostate, the urethra which runs through the prostate and the rectal wall behind the prostate. Serious life quality limiting complications could come to expression if these OARs are damaged, therefore both the delivered dose to the prostate and the OARs will be of interest for this project.

4.2 PCE meta-modelling of dose distributions

Polynomial Chaos Expansion will be used to quantify effects of uncertainties on the delivered dose in SBRT. It will be used to model the dose distribution, which will be the stochastic quantity of interest R, as a function of the uncertainties which characterise it, the uncertain input parameters. The uncertain input in this work will be the setup errors.

4.2.1 PCE construction

To investigate the effect of margin size and fractionation on several dose parameters, PCEs were constructed for treatment plans with different CTV to PTV margins. This was done by using the Polynomial Chaos Expansion For Radiotherapy (PCFORT) algorithm written by Zoltán Perkó . It constructs PCE for a dose distribution in a patient using non intrusive spectral projection and Polynomial Chaos Expansion. The construction of the PCE by PCFORT can be described by the following consecutive steps:

1. Inputs: these include standard deviations of all random input variables and the chosen polynomial (PO) and grid order (GO) and the q value for hyperbolic trimming.
2. Construction of the cubatures
3. Construction of the basis vectors
4. Construction of a dose mask: determine the voxels included in the PCE
5. Calculation of the PCE coefficients
6. Construction of the PCE

4.2.1.1 Initialisation of a PCE

The first steps in constructing a PCE model is choosing the family of polynomials for each stochastic input variable and choosing an appropriate polynomial and grid order, quadrature level extension and hyperbolic trim.

Choice of polynomial order A higher polynomial order means that more complex dose distributions can be modelled, but it also means that a higher grid order is needed to determine the expansion coefficients accurately. This extra accuracy comes at a price, higher order PCEs take longer to sample and construct. Choosing the appropriate polynomial order is done by comparing the clinical dose distribution to the dose distributions resulting from PCEs of different orders. Dose distribution comparison will be discussed in Section 4.3.

A trade off between calculation time, cost and accuracy will determine the optimal order. The maximum polynomial order PO will be given in the input settings for the chosen grid order plus extended level such that $PO = GO + 1$. After the determination of the polynomial order that provides the desired accuracy is determined, the PCE can be used to sample dose distributions.

Collapsing input variables Any shift of the patient during every treatment day with respect to the planning images is a combination of the systematic error and random error of that day. The input variables in this work are the errors in patient position in three dimensions, these are assumed to be Gaussian distributed and are hence characterised by their standard deviations, which is used in the construction. The standard deviation for the systematic and random error are assumed to be isotropic. The real dose that is delivered is depending on the combination in each direction, therefore in each of the three directions the systematic and random setup errors can be substituted by a combined setup error.

For a setup error in direction i with the standard deviations Σ_i and σ_i and expectation values μ_i and M_i, the new combined setup variation would be $\Delta_i \sim \mathcal{N}(\mu_i + M_i, \Sigma_i^2 + \sigma_i^2)$. This reduces the number of input standard deviations from six ($\vec{\Sigma} = (\Sigma_x, \Sigma_y, \Sigma_z)$ and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$) to three. The resulting PCE is a "compressed" PC model depending on $\vec{\xi} \sim (\Delta_x, \Delta_y, \Delta_z)$, which is computationally cheaper as the number of cubature points decreases drastically.

Cubature construction and PCE object initialisation Next, cubatures are generated which will be needed for the numerical approximation of the integral in Equation 19 to obtain the expansion coefficients. For different random variables the best quadrature rules corresponding to their probability density function will be used in the respective directions. As this work only deals with Gaussian distributed variables, Gauss quadratures of maximum level GO and an extra level EL will be used in every direction.

Multi indices of all sparse grids that make up the final cubature are generated. Cubatures are formed by tensorisation of the quadratures, the final cubatures used for integration are formed by summing over the sparse grids.

In the last step of the initialisation a PCE object as defined by a PCE class is initialised. In this step basis vectors are constructed from all multidimensional polynomials up to a maximum degree PO and a hyperbolic trim is performed. Using an initial guess, in this work $q = 0.861$, the q value is calculated in accordance to Equation 27 such that only basis vectors remain for which the integration scheme in the multi index can be supposed to be accurate which will cut out the basis vectors representing high order interactions. Details about the PCE that were used as input are also saved in this object, the resulting PCE objects has the properties:

- A structure that contains the details of the PCE
- PCE basis which is a matrix of multi indices representing the PCE basis vectors
- A multidimensional orthogonal polynomial object that represent the basis vectors
- An initialisation of the PCE coefficients matrix
- The polynomial type of each direction

4.2.1.2 Dose mask and the calculation of PCE coefficients.

To model the clinical dose distribution a PCE must be constructed for every voxel. The patients' CT scans contained about 30 to 55 million voxels. It would be computationally unfeasible to construct a PCE for all voxels as determining the expansion coefficients is very costly. To alleviate the amount PCEs, voxels are only included if the voxels receive a dose exceeding a certain threshold $D_{cut-off}$ for the nominal setup and setup errors of magnitude $3\sigma^*$, where $\sigma_i^* = \sqrt{\sigma_i^2 + \Sigma_i^2}$ for $i \in (x, y, z)$, in positive and negative directions of a 3D Cartesian coordinate system, the six setup errors and the nominal setup are tabulated in Table 1.

The six setup errors are simulated by changing the isocenter of the different beams that deliver dose. setup variation up to $3\sigma^*$ corresponds to the errors that lie within three standard deviations of the mean and within 99.73% confidence. In

$\hat{x}$	$\hat{y}$	$\hat{z}$
0	0	0
$3\sigma_x^*$	0	0
$-3\sigma_x^*$	0	0
0	$3\sigma_y^*$	0
0	$-3\sigma_y^*$	0
0	0	$3\sigma_z^*$
0	0	$-3\sigma_z^*$

Table 1: This table shows all scenarios of isocenter variations that are evaluated for the dose mask construction where $\sigma_i^* = \sqrt{\sigma_i^2 + \Sigma_i^2}$ for $i \in (x, y, z)$.

this project the threshold dose was chosen to be $D_{cut-off} = 0.1$ Gy. Roughly 12 million voxels are included in the dose mask, this number depends on the standard deviations used in building the PCE.

For all voxels included in the dose mask the dose is evaluated for the cubature points to obtain the PCE coefficients. These points belong to a space spanned by the random input variables and represent therefore in the case of patient setup errors as input variable certain perturbations in setup expressed as shifts in the dose distribution with respect to the patient. The dose corresponding to these perturbed patient setups is calculated using the dose engine of the treatment planning system. Finally, the coefficients are saved in the PCE object and the PCE object is saved.

4.3 Validation of PCE dose distributions

In order to investigate the effects of uncertainties in SBRT using PCE modeling, one must first construct a model that mimics the dose distribution and ensures that the discrepancy between the PCE meta-model and the original dose distribution is sufficiently small. This section is dedicated to the construction of the PCE meta model based on the original clinical dose distribution and the validation of the PCE meta model to be used for further analysis. This is done by comparing it to the clinical dose distribution. PCE models were validated with different tools that can compare dose distributions. These tools are discussed below.

4.3.1 Dose difference

A very straightforward way of comparing dose distributions is by testing the dose difference between the distributions. The numerical dose difference is simply computed voxel by voxel and can be presented by a histogram or table. It can be computed for the nominal plan or for certain patient displacements. A limitation of the dose-difference test is that it becomes overly sensitive in steep dose gradient regions as small spatial differences can cause large absolute dose differences in these regions. The dose differences can also be presented in a relative cumulative frequency, the fraction of voxels that exhibit at least a dose difference of ΔD can be determined easily from these diagrams.

Instead of considering all included voxels, one could also compare the dose received by a single voxel under displacement in three directions for the exact and PCE dose distribution. These graphs can give a sense of the quality of the PCE for non nominal cases.

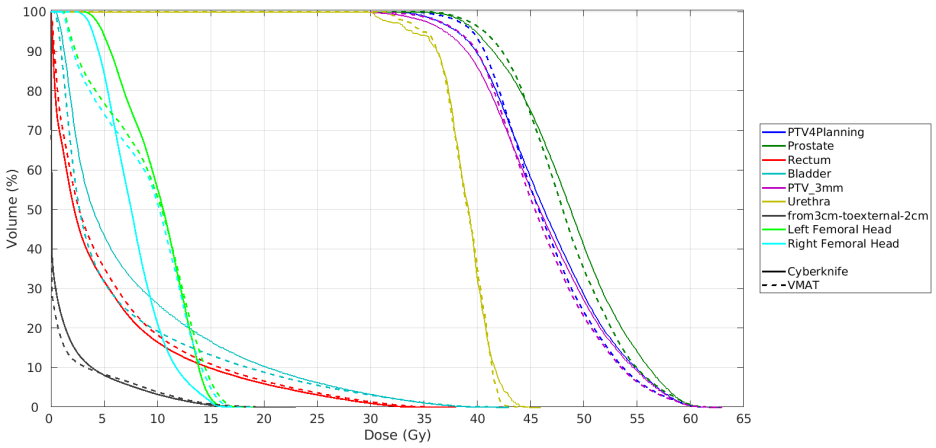


Figure 9: Dose volume histograms showing two dose volume histograms: the dose volume histogram for a Cyberknife treatment in solid lined and for a VMAT treatment in dotted lines. Multiple structures/organs are included.

4.3.2 Dose volume histograms

A Dose Volume Histograms (DVH) is a great tool for interpreting 3D treatment plans. The 3D Dose distributions as calculated by 3D treatment planning systems can be difficult to interpret when displayed as isodose curves on several planes. DVHs summarise the delivered 3D dose distribution of a treatment plan within a defined volume of interest. A (cumulative) DVH shows the volume receiving a dose greater than, or equal to, a given dose, against dose. The volume accumulates starting at the highest dose bin continuing towards zero dose, eventually reaching 100% of the total volume. For a certain patient comparison between DVHs can give a quick insight in the median dose D_{50} within a structure and other dose metrics such as the dose that at least 98% of a volume receives D_{98}, known as the near minimum dose and the dose that at least 2% of a volume receives D_2, known as the near maximum dose. A DVH comparison for a VMAT and Cyberknife plan is shown in Figure 9.

DVHs can be used to investigate the quality of a treatment plan as it shows whether the dose is adequate and uniform throughout the target volume an OARs, however they do not display positional information. DVHs will be used to test the PCE dose distribution against the exact dose distribution.

4.3.3 Gamma evaluations

The sensitivity of dose difference analysis to steep dose gradients led to the development of the distance-to-agreement test. This test evaluates dose distributions independently for each reference point. For a specific reference point, the evaluated dose distribution is searched to locate the nearest point with the same dose value. Unlike the dose difference method, it is not overly sensitive in steep-dose gradient regions. However, in shallow dose gradient regions, a large distance-to-agreement value may be computed even for relatively small dose differences as shallow dose gradient regions typically are larger than steep dose gradient regions. For this reason, the distance-to-agreement could exceed the agreed clinically acceptable criterion. A

commonly used method to compare dose distributions in radiotherapy, the Gamma evaluation, incorporates both the dose and distance criteria [38].

Gamma evaluation provides a numerical quality index in comparing an *evaluated* dose with a *reference* dose. It can quantify the disagreement in the dose regions that fail and the acceptance in dose regions that pass. Passing criteria for the dose difference and distance-to-agreement must be selected beforehand. The Gamma value is computed voxel by voxel in accordance to equation 29.

$$\gamma(\vec{r}_r) = min\{\Gamma(\vec{r}_e, \vec{r}_r)\}, \forall\{\vec{r}_e\} \tag{29}$$

$$\Gamma(\vec{r}_e, \vec{r}_r) = \sqrt{\frac{r^2(\vec{r}_e, \vec{r}_r)}{\Delta d^2} + \frac{\delta(\vec{r}_e, \vec{r}_r)}{\Delta D^2}} \tag{30}$$

where the symbols are defined as:

- $\gamma(\vec{r}_r)$ the minimised general Γ for the set of points belonging to the reference distribution
- Generalized Γ function, computed for all evaluated positions $\vec{r}_e$ and reference positions $\vec{r}_r$
- $r(\vec{r}_e, \vec{r}_r) = |\vec{r}_e - \vec{r}_r|$ Spatial distance between evaluated and reference dose points
- $\Delta\delta$ The distance to agreement criterion
- $\delta(\vec{r}_e, \vec{r}_r)$ Difference between evaluated dose $D_e(\vec{r}_e)$ at position $\vec{r}_e$ and reference dose $D_r(\vec{r}_r$ at $\vec{r}_r$
- ΔD Dose difference criterion

This gamma calculation is done for each voxel in the reference distribution. The passing-fail criterion for this method will be:

$$\begin{aligned} \gamma(\vec{r}_r) \leq 1, &\qquad \text{voxel passes} \\ \gamma(\vec{r}_r) > 1, &\qquad \text{voxel fails} \end{aligned}$$

meaning that the comparison of two dose distributions using a gamma evaluation will determine for each voxel whether it has passed or failed.

By using the passing-fail criterion the results of a Gamma evaluation can easily be visualised in a plot. The worst differences that still passes are either two voxels that have the same position but a dose difference of exact the chosen dose difference criterion $\delta(\vec{r}_e, \vec{r}_r) = \Delta D$ or two voxels that receive the same dose but are exact the distance to agreement criterion apart$|\vec{r}_e - \vec{r}_r| = \Delta\delta$. In this work the evaluated and reference dose distributions will be the PCE dose and the dose according to the dose engine.

5 Margin recipe construction using Polynomial Chaos Theory

The previous chapter gave an overview in how PCE meta models are constructed. This chapter discusses the use of polynomial chaos expansion to compute dose metrics and derive margin recipes for fractionated treatments.

5.1 Modelling fractionation effects

The PCEs that are constructed based on a certain treatment plan are PCEs that depend only on the combined setup error in each direction as the systematic and random errors are collapsed in each direction. From this "compressed" PCE the different effects that systematic and random errors have on a dose distribution can still be investigated as it is possible to analytically derive the PCE depending on the original six variables. One could derive from this compressed PCE a PCE that depends on only the random setup variations for a specific value of systematic error. This specific value must be chosen from the Gaussian distribution with the same standard deviation for the systematic setup as given in the input settings during construction. The PCE that one obtains could be considered as a patient's PCE. As discussed earlier the systematic error is patient specific and constant throughout a treatment.

Suppose a PCE that is constructed for certain values of Σ and σ, making the combined error
$sigma_c^* = \sqrt{\Sigma^2 + \sigma^2}$. A synthetic patient population of $N_{patients}$ that receive treatments in FN fractions can be simulated. For this population $N_{patients}$ realisations of the systematic errors are drawn from a Gaussian distribution $\mathcal{N}(0, \Sigma^2)$ and for each systematic error FN random errors are drawn from a Gaussian distribution $\mathcal{N}(0, \sigma^2)$. To simulate the treatment PCEs are built for the drawn systematic setup error and evaluated for the FN different random errors, the responses of these PCEs are averaged over these FN fractions to yield the final dose distribution for the fractionated treatment. By doing this population percentiles of treatment parameters such as:

- D_α: dose that at least $\alpha\%$ of the voxels receive
- V_β^O: the fraction of voxels belonging to organ O that receive at least $\beta\%$ of the prescribed dose

can be determined. In this work the near minimum dose $D_{98\%}$ to the PTV and prostate, near maximum dose $D_{2\%}$ to the OARs and dose coverage V_{100} of the prostate will be investigated. The near minimum and maximum doses are chosen as dose metrics instead of the absolute extrema because the latter would only hold for a few voxels, which would be meaningless on a scale of millions of voxels.

By doing this for different number of fractions, fractionation effects on the desired dose metrics for organs or volumes of interest can be studied. Suppose the fractionation schemes of interest are $\{FN1, FN2, ..., FNmax\}$, an efficient way to do this is to calculate the PCE response of $FNmax$ random errors for each systematic error. By averaging over FN PCE responses per systematic error, the PCE response of an FN fractionated treatment can be obtained for all fractionation schemes of interest at once. Figure 10 summarises this process.

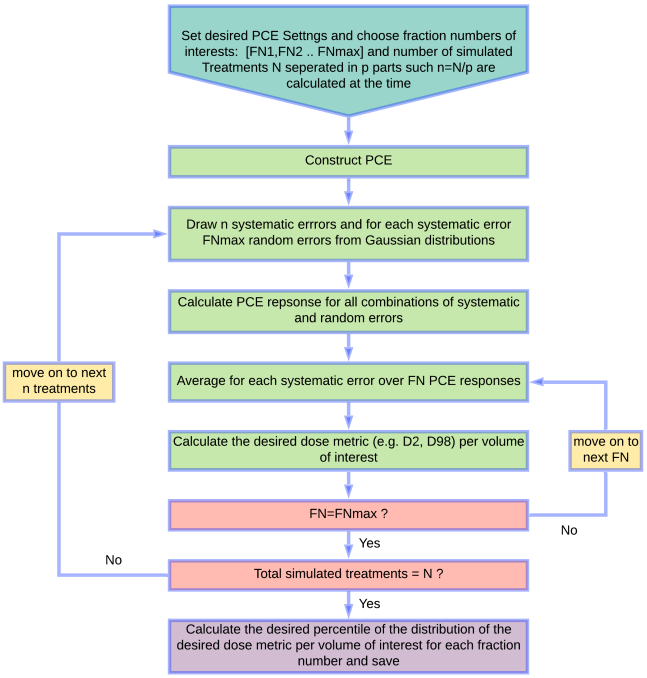


Figure 10: A flow chart that summarises how fractionation effects can be modelled for an entire patient population using a single PCE. The purple block can be appended if one is interested in values at certain percentiles of the patient population or omitted if one is interested in the full distribution over the entire population.

5.2 Constructing margin recipes

PCE for radiotherapy can be used for the construction of a margin recipe. A margin recipe describes the needed CTV to PTV margin M_{PTV} that is needed to deal with certain setup errors. Using PCE methods it is possible to simulate large populations of patients to construct these recipes at an affordable computational cost.

First a criterion has to be defined regarding what is clinically acceptable as treatment outcome in terms of a CTV dose parameter. Within this work a treatment was considered successful if the probability P of getting the full prescribed dose for at least 98% of the CTV is greater than or equal to 98%:

$$P\left(V_{100}^{CTV} \geq 98\%\right) \geq 98\% \tag{31}$$

For the margin recipes it is of more interest when this criterion is violated as it gives the maximum systematic and random error a certain margin can handle. Hence, the criterion, i.e. at least 98% of the simulated patient population receives at least the full prescribed dose in 98% of the CTV, that was used in determining what combination of Σ and σ are clinically acceptable, was:

$$97.95\% \leq P\left(V_{100}^{CTV} \geq 98\%\right) \leq 98.10\% \tag{32}$$

meaning that for a given margin the combination of Σ and σ are determined such that the probability of successful target coverage ($V_{100}^{CTV} \geq 98\%$) is in $[97.95, 98.10]$.

5.2.1 Sampling PCE responses for margin recipes

For a certain PTV-CTV margin an initial guess on what values for Σ and σ would meet the criterion in Equation 32 for the lowest fraction number FN_{min} of interest was made. PCEs were based on treatment plans that have this value for M_{PTV} with the initially guessed values for Σ and σ. After the construction of a PCE, the CTV dose distribution, of 10^5 fractionated treatments was obtained and multiplied by a scaling factor that will be explained in Section 5.2.1.1. The CTV dose for each of the simulated treatments was calculated using a different systematic setup error and FN random errors representing the different fractions sampled using the Gaussian distributions with mean zero and standard deviations Σ and σ. From the CTV dose distribution the probability $P\left(V_{100}^{CTV} \geq 98\%\right)$ was calculated.

If the probability did not satisfy Equation 32, a second and sometimes a third guess was made based on the value of the $P\left(V_{100}^{CTV} \geq 98\%\right)$ corresponding to the initial guess. If the probability of successful treatments is less than 97.95%, the systematic error for a certain random error is apparently too big and was decreased for the next iteration. When the probability is more than 98.1% the systematic error is increased. A new PCE is then constructed for the same plan using the new guesses for the systematic error and treatments are simulated and evaluated again. These steps are repeated until the passing criterion is achieved. Figure 11 summarises these steps.

5.2.1.1 Scaling factor

The margin recipe is obtained from sampling PCE responses that are models for the dose that a single patient receives under certain setup errors and thus patient specific. To illustrate why this is suppose patient 1 has a PTV coverage of 99% in the nominal case ($V_{100}^{PTV} = 99\%$) and patient 2 has a PTV coverage of 95% in the nominal case. The treatment plans are optimised to satisfy a certain PTV constraint that will lead to successful CTV coverage. Sampling what combinations of setup errors still result in a clinically acceptable CTV coverage would be different for these patients. Due to the intrinsically better PTV coverage for patient 1, larger setup errors for patient 1 would satisfy Equation 32 than for patient 2. A recipe based on patient 1 will therefore allow greater setup errors than a recipe based on patient 2 and validating a margin recipe based on patient 1 for patient 2 would give negative results.

In order to make the recipe valid for an entire cohort of PCa patients, some measure has to be undertaken to ensure the plan quality is the same for the patient the recipe is based on and patients the recipe is validated for. In the scope of this work the PCE response for the dose for all simulated treatments is scaled with a fixed scaling factor. This scaling factor is chosen such that the PTV coverage in the nominal scenario satisfies Equation 33 exactly.

$$V_{100}^{PTV} = 95\% \tag{33}$$

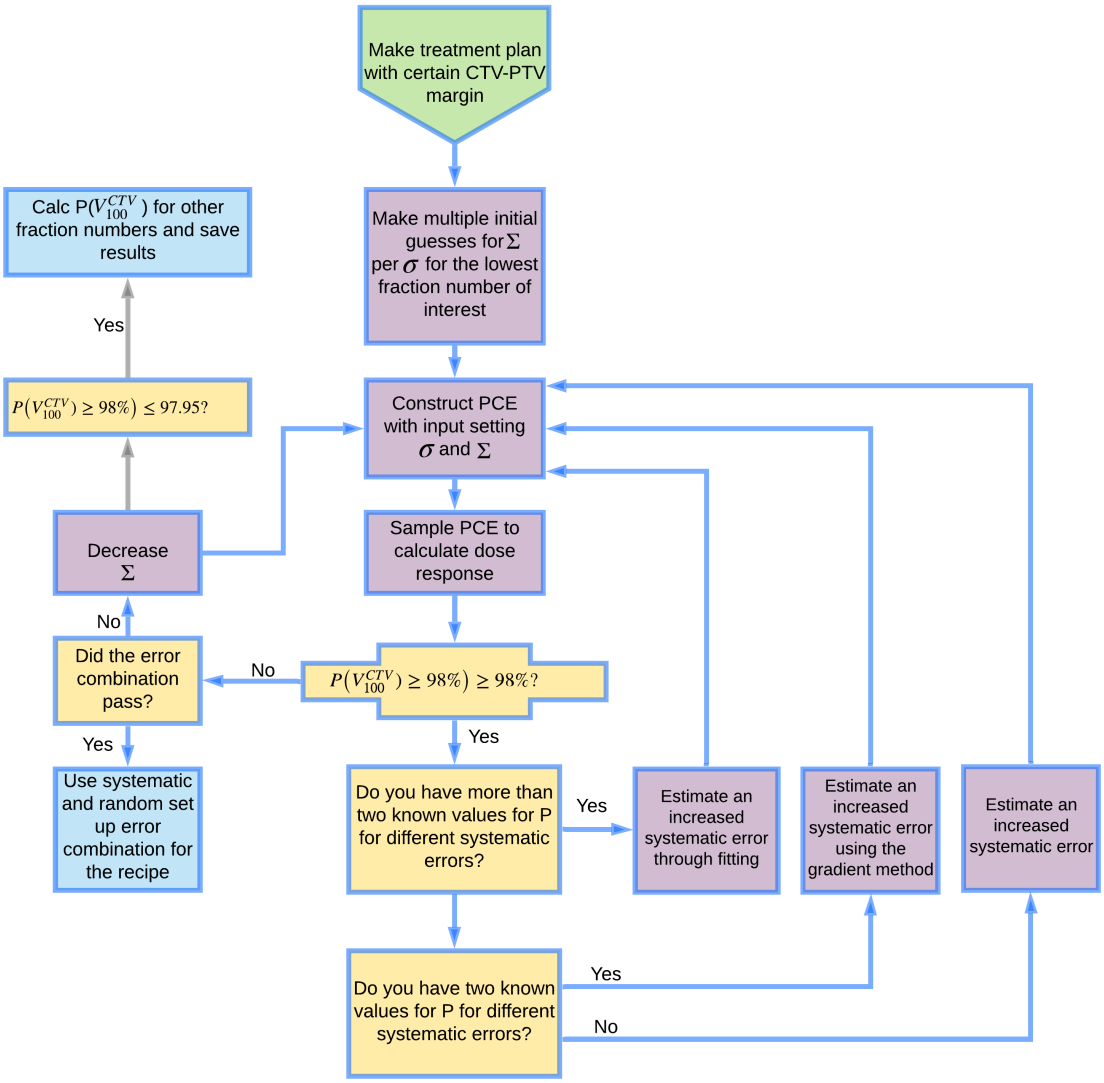


Figure 11: This figure summarises the iterative manner of finding error combinations that yield the desired passing probability

5.2.2 Estimating Σ and σ for the next iteration

A smart estimation for the next value of Σ that hopefully does satisfy Equation 32 decreases the number of needed iterations and therefore computational costs. After the first iteration cycle an estimation must be made based on $P\big(V_{100}^{CTV} \geq 98\%\big)$. This could be done in multiple ways.

If there is only a single value known for $P\big(V_{100}^{CTV} \geq 98\%\big)$, the value for the standard deviation for the systematic error for the next iteration round is simply based on $P\big(V_{100}^{CTV} \geq 98\%\big)$. The increment or decrement in Σ is based on how much this probability deviates from the criterion in Equation 32, this is tabulated in Table 2. If there are at least 2 values for $P\big(V_{100}^{CTV} \geq 98\%\big)$ known for a certain random error, the systematic error for the next iteration can be determined by a first order gradient method.

Condition for $P\big(V_{100}^{CTV}\big) \geq 98\%\big)$	$\Delta\Sigma_{Next}$
$P \leq 97\%$ & $\Sigma \geq 0.2$	-0.2
$P \leq 97.95\%$ & $\Sigma > 0.1$	-0.1
$P \leq 97.95\%$ & $\Sigma \geq 0.05$	-0.05
$P \leq 97.95\%$ & $\Sigma \geq 0.1$	-0.1
$P = 100\%$	+0.5
$98.10 \geq P \leq 98.9\%$	+0.05
$98.10 \geq P \leq 99.10\%$	+0.1
$P \geq 99.10\%$	+0.2

Table 2: Choice of the systematic error for the next iteration if there is only a single value for $P\big(V_{100}^{CTV} \geq 98\%\big)$ known. The next systematic error is given by $\Sigma + \Delta\Sigma_{Next}$ based on the value for $P\big(V_{100}^{CTV} \geq 98\%\big)$ that is denoted as P. The conditions for P are checked in priority from the top to bottom.

The fastest way to construct a recipe is to estimate multiple values for the systematic error for an array of random errors. Based on the $P\big(V_{100}^{CTV} \geq 98\%\big)$ values of these error combination a good estimation for the systematic error can be made by fitting the data to these points. For example, one could start with the error combinations as tabulated in Table 3, these combinations could be educated guesses based on a previously sampled fraction number or margin.

σ	Σ_1	Σ_2	Σ_3
0	1.20	1.40	1.60
0.50	0.50	0.80	1.00
0.80	0.40	0.60	0.80

Table 3: This table shows an example array of systematic setup errors that one could try for a certain random error as an initial approach to converge to the desired combination of Σ and σ that will satisfy Equation 32

Based on for example the $P\big(V_{100}^{CTV} \geq 98\%\big)$ values for the third row , a prediction for Σ that will give the desired probability can be made by fitting a polynomial of the form $P\big(V_{100}^{CTV} \geq 98\%\big) = a + b \cdot \Sigma + c \cdot \Sigma^2$ to the data, this is illustrated in Figure 12. By solving this fit for $p = 98\%$, the value for the systematic error was predicted to be $\Sigma = 0.51$, which turned out to exactly meet the passing criterion. The advantage of using this method is that it requires less iteration points and therefore it decreases the number of PCEs that need to be constructed. This is probably because it can consider a non linear dependence as at least three points will be used for the fit.

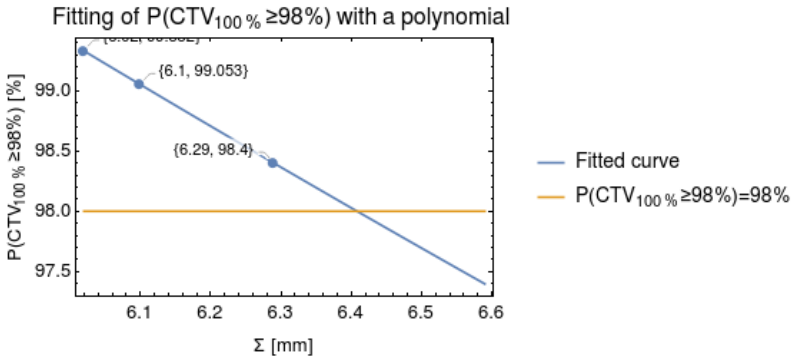


Figure 12: An example fit of different systematic errors and their passing probabilities.

5.2.3 Estimating Σ and σ for higher fraction numbers

Suppose we are interested in treatments that are fractionated in 2, 3, 4 and 10 fractions. Due to averaging effects a treatment that is fractionated in more fractions is able to handle larger errors. Now suppose we build a PCE for a certain combination of setup errors for a treatment fractionated in only 2 fractions and the $P\left(V_{100}^{CTV} \geq 98\%\right)$ value turns out to be 2%. Although this combination did not meet the desired passing probability, it might do for another higher fraction number.

If $P\left(V_{100}^{CTV} \geq 98\%\right) \leq 97.5\%$ the PCE was reused for the remaining higher fraction numbers to calculate $P\left(V_{100}^{CTV} \geq 98\%\right)$. An efficient way of doing this is by constructing a scenario matrix for the largest fraction number of interest FN_{max}, this matrix will contain 10^5 values for the systematic setup errors obtained from a Gaussian distribution with standard deviation Σ and for each systematic error FN_{max} random errors from a Gaussian distribution with standard deviation σ. The responses of in total $10^5 * FN_{max}$ treatments are averaged over all fraction numbers of interest up to FN_{max}, this means that the $P\left(V_{100}^{CTV} \geq 98\%\right)$ for all fraction numbers of interest up to FN_{max} can be simulated at once. When the results are obtained, they can be saved and the PCE is no longer needed.

These results can be used to make better initial estimations for the standard deviations a plan with the same margin can handle if the treatment is delivered in more fractions for future use. This is indicated in Figure 11 with gray arrows.

6 Results

6.1 Choice of maximum included polynomial order

For the PCE to mimic the dose distribution in a way that doses can be recalculated for different patient setup up variations using only the PCE dose distributions, sufficiently high order polynomials must be included in the expansion. The higher the polynomial order that is included, the more accurate but computationally costly the PCE construction and sampling becomes. A higher polynomial order requires higher level quadratures to determine the coefficients accurately. PCE dose distributions were constructed with polynomials up to 6th order. These PCEs were compared to the clinical dose distribution, results of this comparison are presented in this subsection.

PCEs were constructed for the 3 mm VMAT plans for $\Sigma = \sigma = 3$ mm for max polynomial order $PO \in \{2, 3, 4, 5, 6\}$. For the Cyberknife plans the same polynomial orders were investigated. For each polynomial order the minimal required grid order GO plus extra quadrature level EL that results in accurate PCE coefficients was used:

$$PO = GO + EL = GO + 1 \tag{34}$$

To differentiate between PCEs of different grid and polynomial orders, the following notation is adapted:

$$GO\tau EL\nu PO\phi$$

where τ denotes the used grid order, ν denotes the grid extension, which will be 1 throughout this work, and ϕ denotes the maximum included polynomial order.

An overview of the construction time and the needed evaluations of the clinical dose engine for each PCE is given in Table 4.

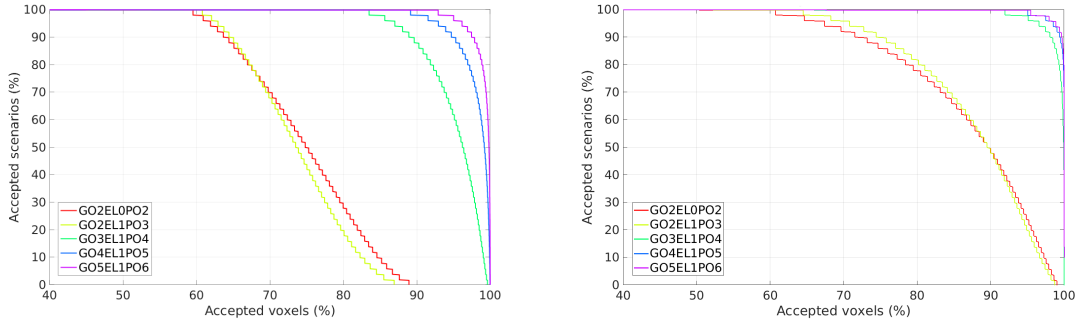
POELGO	**RT modality**	**Construction time** (hh:mm)	**# Evaluations**
GO2EL0PO2	VMAT	00:23	7
GO2EL1PO3	VMAT	00:52	13
GO3EL1PO4	VMAT	01:32	37
GO4EL1PO5	VMAT	03:12	111
GO5EL1PO6	VMAT	09:29	303
GO2EL0PO2	cyberknife	00:35	7
GO2EL1PO3	cyberknife	00:50	13
GO3EL1PO4	cyberknife	01:22	37
GO4EL1PO5	cyberknife	03:38	111
GO5EL1PO6	cyberknife	15:06	303

Table 4: The construction time for PCEs with different maximum included polynomial orders. The used standard deviations for VMAT and Cyberknife were $\Sigma = \sigma = 3$ mm and $\Sigma = \sigma = 2$ mm respectively. One can see that the construction time scales about linearly with the number of needed clinical dose evaluations that are needed to calculate the coefficients as that is the most computationally expensive and time consuming step in PCE construction. The construction time between the Cyberknife PCEs of order $GO4EL1PO5$ and $GO5EL1PO6$ did not scale linearly as not enough memory was reserved, only 90 gb were reserved while 130 gb was used.

6.1.1 Gamma Evaluation

Gamma Evaluations to compare the different PCEs constructed for several selected patients to the clinical real dose distribution were performed for 50 combined setup

variations that lie on the surface of the ellipsoid that is constructed by 99% of the combined errors. Seven patients' VMAT treatment plans and four Cyberknife treatment plans were analysed. Since all input variables are assumed to be Gaussian, these 50 combined set up variations lay in the upper 99th percentile of the Chi-square distribution with three degrees of freedom. The dose difference ΔD and distance to agreement criterion Δd in all gamma evaluations were chosen to be $\Delta D = 0.1Gy$ and $\Delta\delta = 1$ mm.

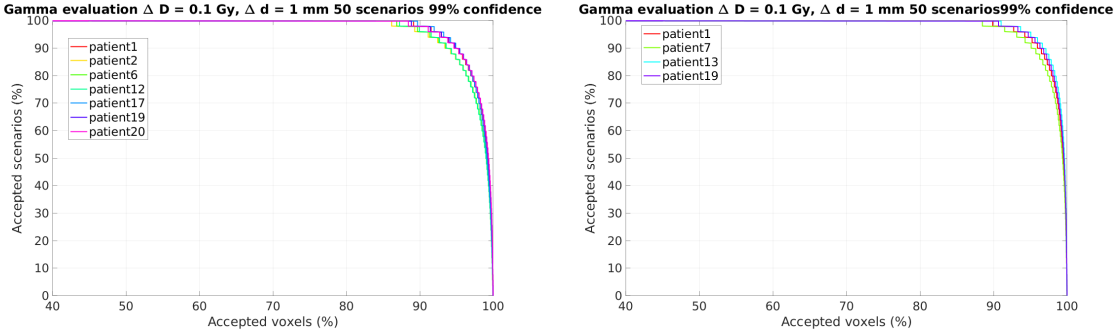


(a) VMAT (b) Cyberknife

Figure 13: Γ – evaluation comparison between different grid and polynomial orders for a PCE VMAT dose (a) and a PCE cyberknife dose (b).

From Figure 13 it can be seen that for increasing maximum included grid and polynomial order, the PCE models the clinical dose distribution better and better. A large improvement is made going from $GO = 2, PO = 3$ to $GO = 3, PO = 4$ whereas the improvement between $GO = 4, PO = 5$ to $GO = 5, PO = 6$ is more subtle.

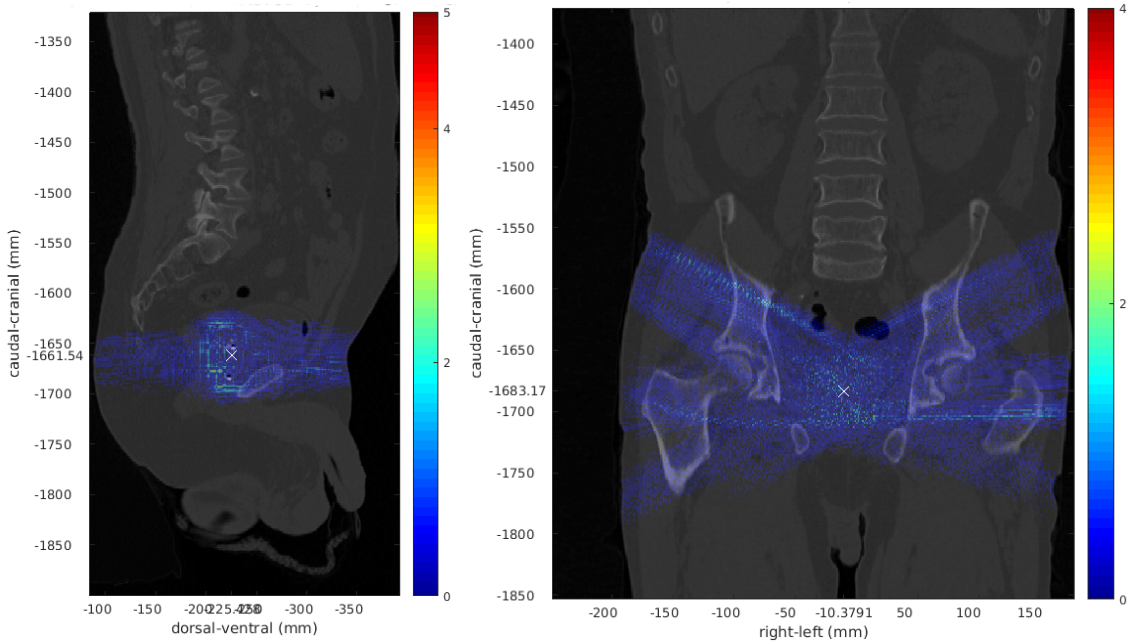
For PCEs of order $GO4EL1PO5$ gamma evaluations have been performed for multiple patients to verify whether this order gives the desired accuracy across patients, these results for VMAT and Cyberknife dose distributions are depicted in Figure 14. The results from the gamma evaluation show that this grid order is comparably accurate among the evaluated patients.



(a) VMAT (b) Cyberknife

Figure 14: Intra patient Γ – evaluation comparison between the clinical true dose and the PCE dose for VMAT (a) and Cyberknife (b), where the polynomial and grid order were GO4EL1PO5.

To investigate where the PCE dose distribution lacks accuracy within the patient, the dose difference between the clinical and PCE dose distribution have been plotted. This was done for the scenario with the least amount of accepted voxels out of the 50



(a) VMAT (b) Cyberknife

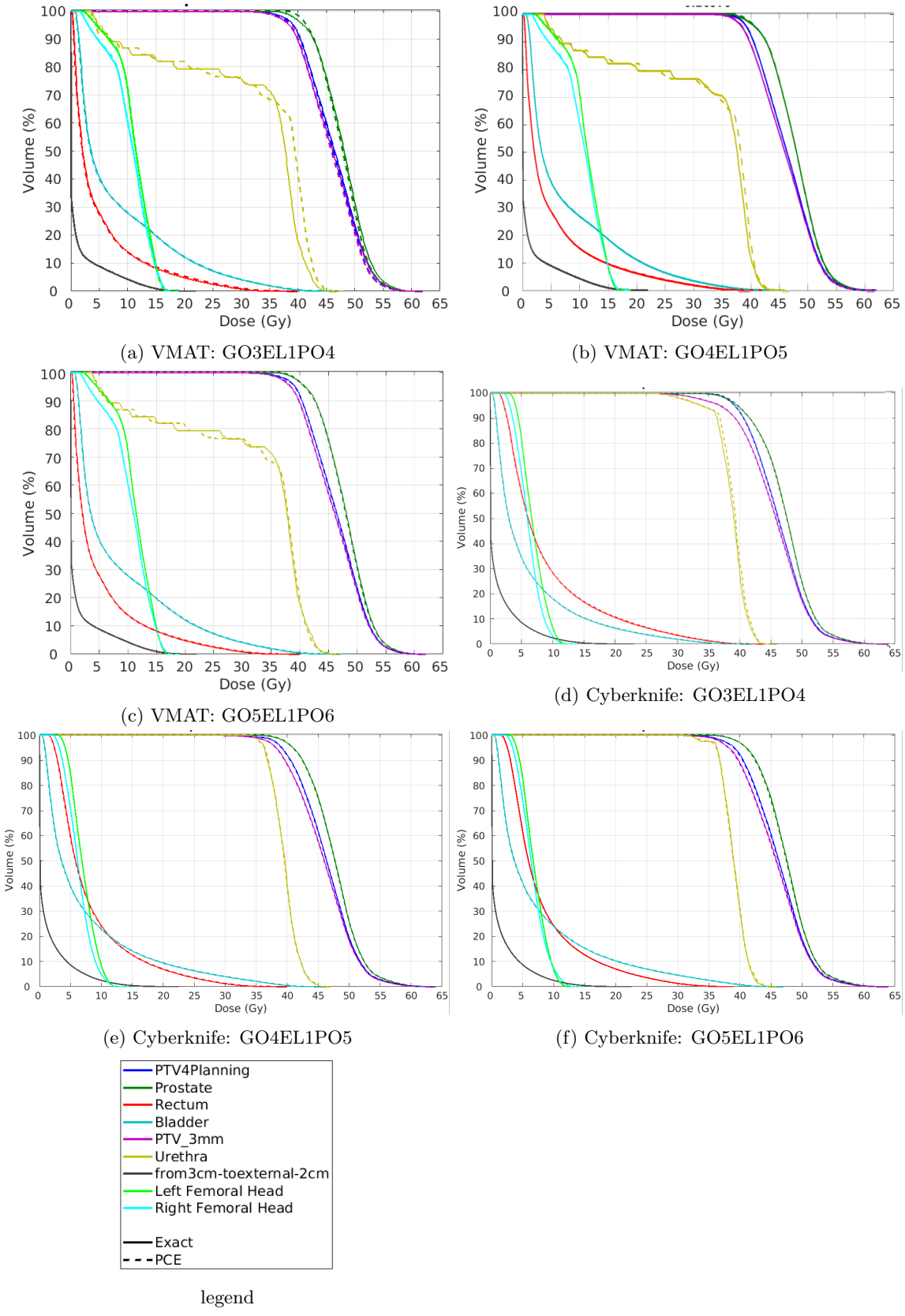
Figure 15: The absolute dose difference between PCE and dose engine for the scenario with the least number of accepted voxels 99% confidence ellipsoid of the gamma evaluations for VMAT (a) and a PCE cyberknife (b). The polynomial and grid order were GO4EL1PO5.

combined set up variations for patient 1. The absolute dose difference for the worst performing slice within the patient for the worst performing scenario of the gamma evaluations is depicted in Figure 15. As seen in Figure 15 there is no particular structure/organ where the PCE dose performs poor. The largest dose differences in the VMAT dose was observed on the rectum side of the prostate, where the dose distribution is steep. The largest dose difference for the Cyberknife dose are for certain voxels, about evenly distributed within the PTV and for certain beams.

6.1.2 Dose Volume Histograms

Dose volume histograms could also be used to compare between the PCE dose distribution and the clinical dose distribution. DVHs were constructed for 3 mm VMAT plans and for 3 mm Cyberknife plans. This was done for all polynomial orders that were used in the gamma evaluations for the same patient. DVHs were compared in absence of patient set-up errors and for the worst performing scenarios in the gamma evaluations.

The DVHs of the worst performing scenario for the VMAT treatment plans for maximum included polynomial order four, five and six are shown in Figure 16. As expected, the DVHs overlap more for increasing maximum included polynomial order, meaning the PCE dose distribution models the clinical dose distribution better and better. The largest difference in the DVHs for both RT modalities are observed for the Urethra. The urethra is situated at around the center of the prostate and is a very small organ. Due to its small size, deviations in the received dose lead to largest discrepancies in the DVH with respect to the clinical true dose DVH, though this difference becomes negligible for $GO4EL1PO5$ and $GO5EL1PO6$.



(a) VMAT: GO3EL1PO4

(b) VMAT: GO4EL1PO5

(c) VMAT: GO5EL1PO6

(d) Cyberknife: GO3EL1PO4

(e) Cyberknife: GO4EL1PO5

(f) Cyberknife: GO5EL1PO6

legend

Figure 16: DVH comparison between the VMAT and Cyberknife true dose and the corresponding PCE dose for different polynomial and grid orders. The comparison is made for the scenario with the least number of accepted voxels from the 50 scenarios on the 99% confidence ellipsoid.

6.1.3 Dose comparison for a single voxel under spatial translation

For a certain voxel in the prostate the PCE dose has been calculated under simulated spatial translation of the patient in six directions, namely left, right, anterior, posterior, superior and inferior, i.e. along the positive and negative x, y, z axes. The translated dose has been calculated for all polynomial orders included in the gamma evaluations.

This dose represents the dose that a certain voxel within the CTV would receive under a certain patient setup error (which has a systematic and random component). A comparison can be made between the voxel dose according to the clinical dose engine and the PCE dose under that respective shift to see whether the PCE behaves as desired. This was done for multiple randomly selected voxels within the CTV, for a certain voxel results are shown in Figure 17.

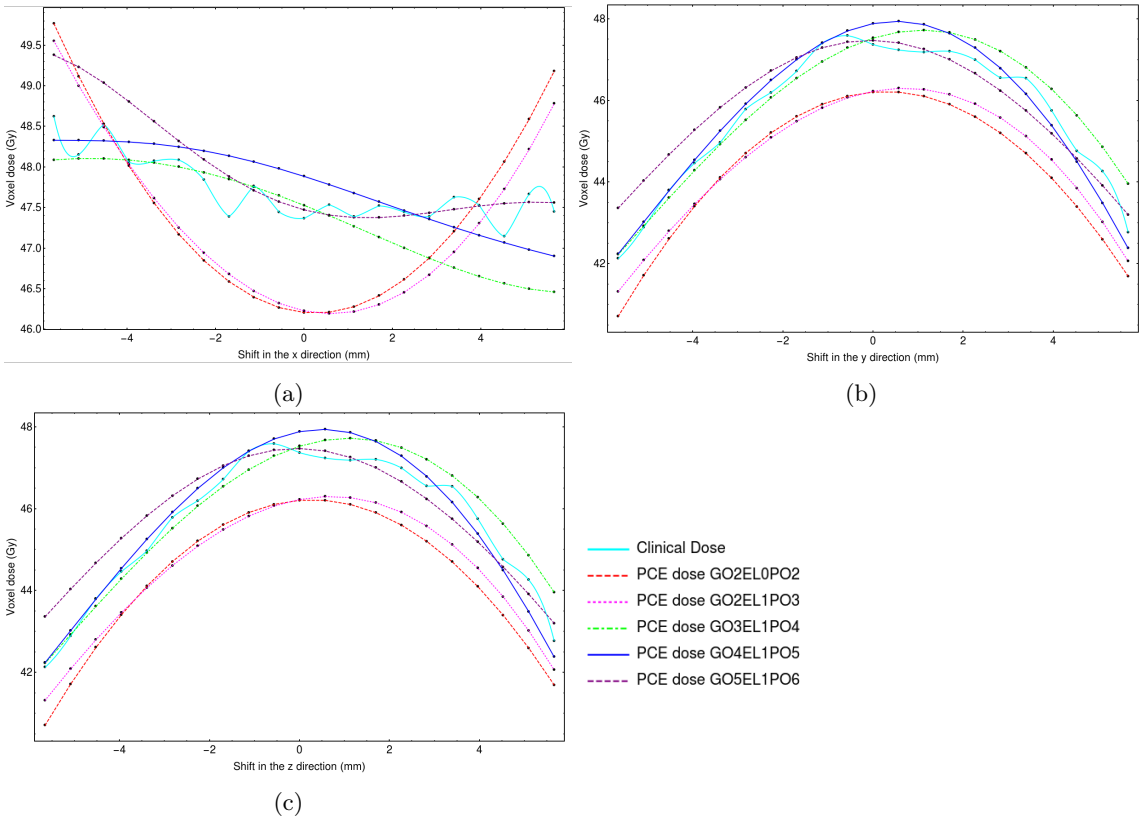


Figure 17: Voxel dose according to the clinical dose engine in yellow, and PCEs that include different polynomial degrees for voxel translation along the x-axis (a), y-axis (b) and z-axis(c).

For increasing polynomial order, the response of the PCE dose under spatial translation becomes better, a leap in improvement is observed going from $PO = 3$ to $PO = 4$ in agreement with the results from the gamma evaluations. The points for the nominal scenario (i.e. no shift) do not overlap as the expected squared difference for all scenarios included in the construction of the PCE is minimised in the construction of the PCE.

For both RT modalities it can be seen in Figure 13 and 16 that PCEs of order $GO4EL1PO5$ and $GO5EL1PO6$ have a very similar performance. The number of dose engine calculations and therefore the construction time does increase drastically for $GO5EL1PO6$ as seen in Table 4, these PCEs are also more than twice as large

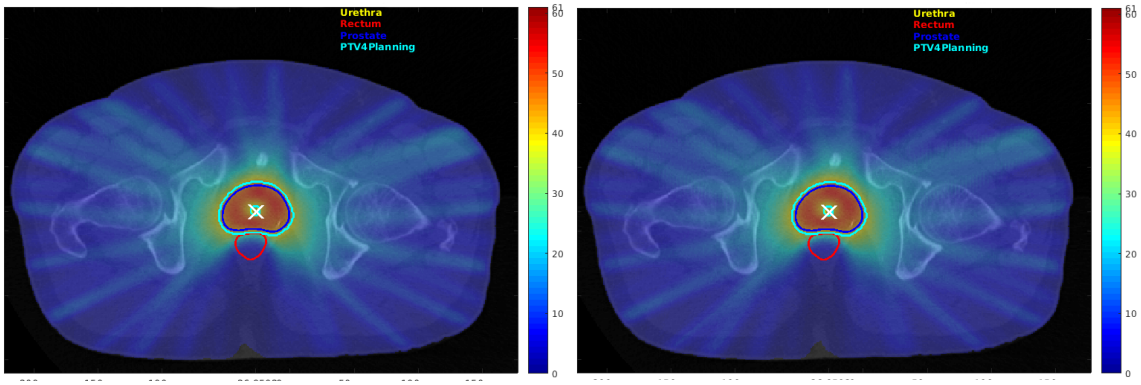
Patient	1	2	6	12	17	19	20
Rate	0.9809	0.9760	0.9766	0.9772	0.9806	0.9811	0.9807

Table 5: The passing rates for the VMAT plans of the scenario with the least number of accepted voxels on the 99% confidence sphere. The gamma evaluation was performed on the dose engine dose and a PCE of order $GO4EL1PO5$.

in terms of storage memory, namely roughly $10GB$ for $GO4EL1PO5$ and $25GB$ for $GO5EL1PO6$. Based on these numbers and validation results the maximum included orders for the PCE is chosen to be $GO4EL1PO5$.

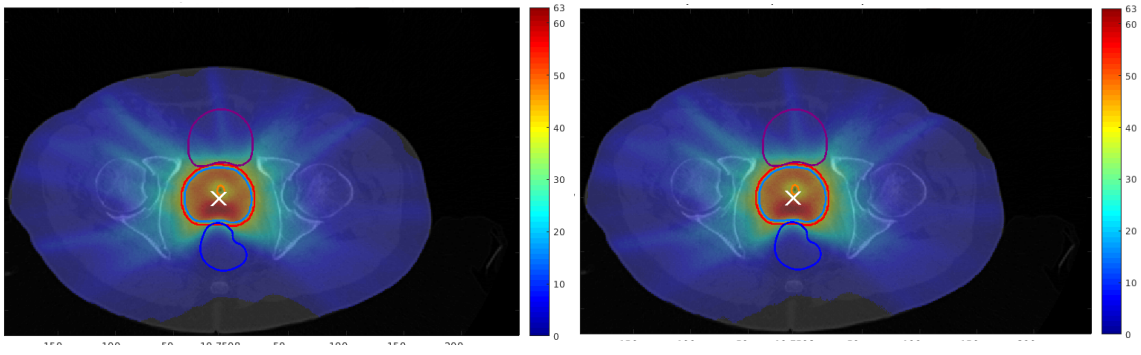
The gamma passing rates for the scenario with the least number of accepted voxels on the 99% confidence sphere for all patients considered in the gamma evaluation performed on VMAT dose distributions is tabulated in Table 5. For the chosen order $GO4EL1PO5$ the passing rate of the worst scenario was around 98%.

A side by side comparison of the nominal case between the PCE dose distribution and clinical dose distribution is depicted in Figure 18 and 19. More gamma Evaluations and DVH comparisons were performed for five other patients, these show similar results.



(a) PCE dose distrubution nominal scenario (b) Treatment plan dose distribution

Figure 18: The PCE dose distribution in absence of patient errors and the clinical treatment plan dose distribution for a coplanar VMAT treatment.



(a) PCE dose distrubution nominal scenario (b) Treatment plan dose distribution

Figure 19: The PCE dose distribution in absence of patient errors and the clinical treatment plan dose distribution for a non coplanar cyberknife treatment.

6.2 Sampling patients' PCEs

By building different patients' PCEs as described in Section 5.1 one could simulate fractionated treatments for a chosen number of fractions for a large patient population with certain setup and random errors. The expectation value for near minimum dose that the prostate receives, the near maximum doses for several OARs and the expectation value for the CTV coverage was investigated for different treatment plans. By doing this for different number of fractions, the effects of fractionation on the received dose can be quantified.

The fractionation analysis was performed on both VMAT and Cyberknife treatment plans. Because the results for both RT modalities was so similar, only the results for Cyberknife are shown in this section, as for the margin recipes construction shown in Section 6.3.1 Cyberknife plans were used.

6.2.1 CTV coverage

The averaged percentages of voxels within the CTV that receive at least the full prescribed dose was sampled for several treatment plans for both RT modalities for different combination of systematic and random setup errors. These expectation values for a systematic error of 0.5 mm against the random error σ are depicted in Figure 20 for 10.000 fractionated treatments. The effects of fractionation on the CTV coverage can be clearly observed from Figure 20, the higher the fraction number, the higher the expected CTV coverage for that particular combination of systematic and random set-up error.

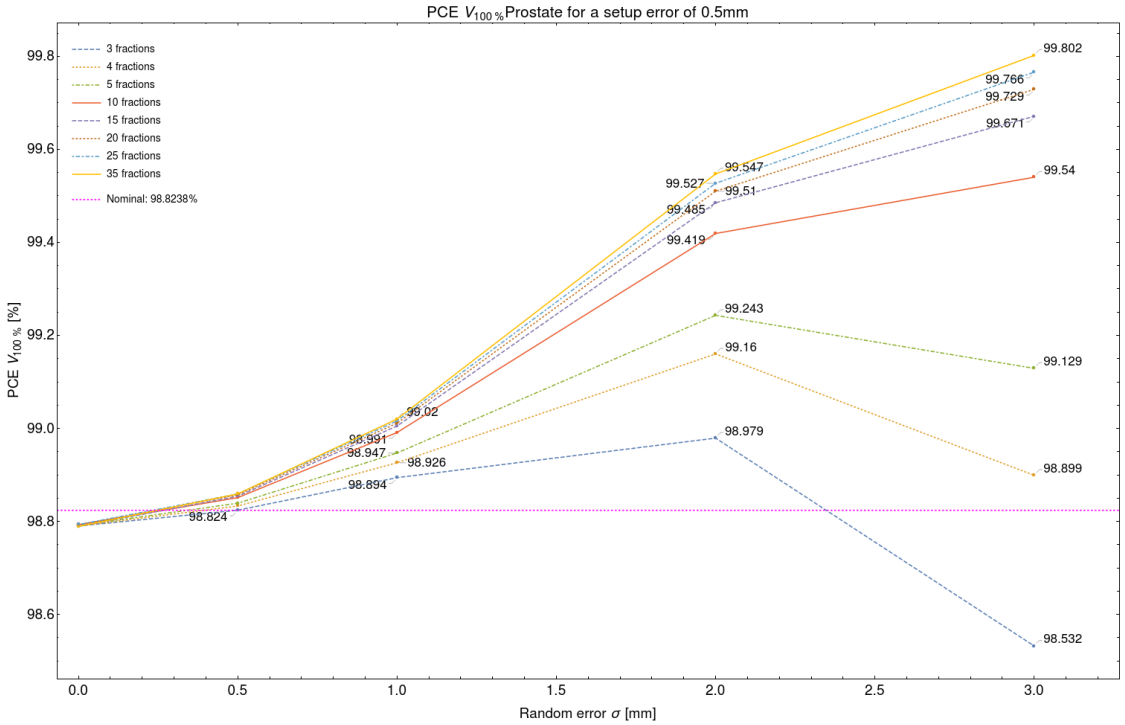


Figure 20: The averaged value of the V_{100} of the CTV in a Cyberknife plan as function of σ for different fraction numbers for $\Sigma = 0.5$ mm.

Another observation from Figure 20 is that the expected CTV coverage increases for small patient shifts with respect to the nominal clinical plan for an increasing

random set-up error. This increase was observed for random errors up to 3 mm for $\Sigma = 0, 0.5$ mm and 1 mm, though for $\Sigma = 2$ mm this increase was observed up to a random error of 2 mm. For large number of fractions in the case $\sigma > 1$ mm and $\Sigma \leq 1$ mm the coverage was higher than for the nominal scenario. The expected CTV coverage decreases for an increasing systematic error at a certain value for the random error. This phenomenon has been observed for both RT modalities.

6.2.2 Near minimum prostate dose

The near minimum prostate dose has been determined for a systematic error of $0, 1$ and 1 mm and a random error between 0 and 3 mm. It was observed that for small systematic set up errors the minimum dose increases for increasing random error and that this effect becomes stronger for an increasing fraction number.

To check whether the increased D_{98} is an artefact of the PCE dose distribution, dose calculations for different combinations of (Σ, σ) have been performed with the clinical dose engine. The D_{98} of the true dose was also found to be higher than for the planned nominal dose for certain combinations of setup errors. The results for $(0.5, 1)$ are depicted in Figure 21. What can also be observed from this figure is that the D_{98} values of the PCE are in good agreement with the true dose.

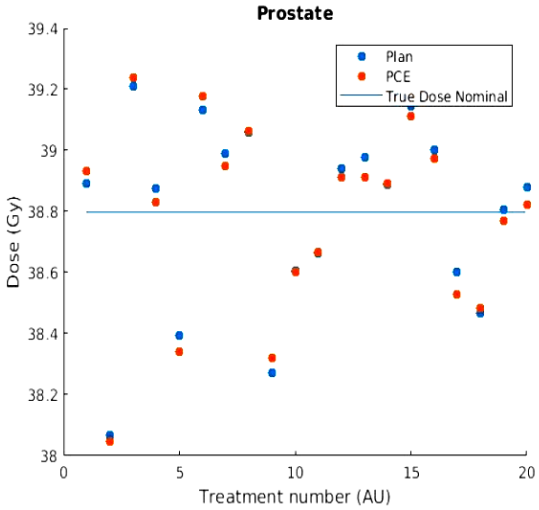
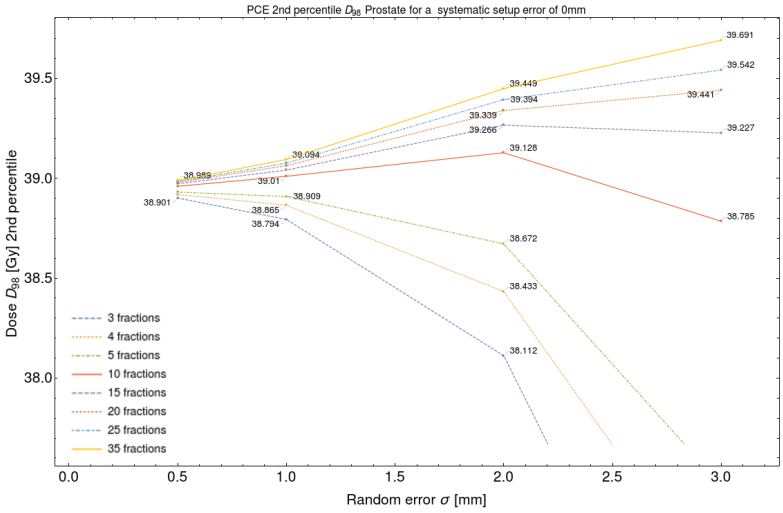


Figure 21: D_{98} comparison for 20 patient shifts from Gaussian distributions with $(\Sigma, \sigma) = (0.5, 1)$

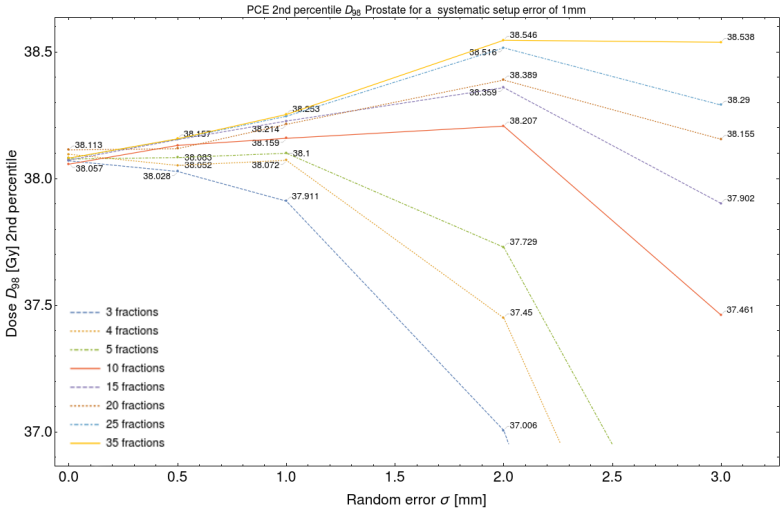
Figures 22a to 22c show the effect of fractionation on the 2nd percentile of the prostate near minimum dose. If we look at the second percentile of the $D98$ of the prostate, an increase with the random error is also observed for treatments fractionated in more than 10 fractions. This effect was observed for a systematic setup error of $0, 0.50,$ and 1 mm and weakens with increasing systematic error. For $\Sigma = 1$ mm this effect breaks down for 10 fractions at $\sigma > 1$ mm and for even higher fraction numbers it breaks down at $\sigma > 1$ mm. For $\Sigma = 2$ mm this effect is not observed any more, though the plot in Figure 22c is noisy due to sampling noise.

Another observation from Figures 22a to 22c is that the second percentile of the $D98$ of the prostate decreases for increasing systematic error given a certain random error.

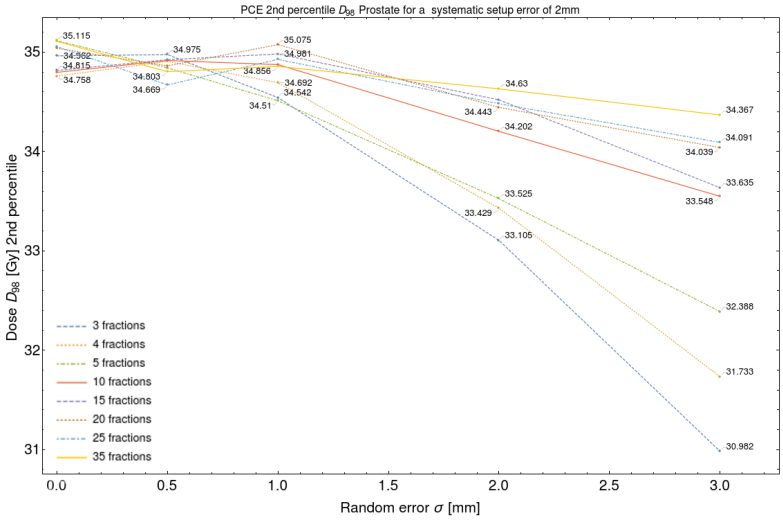
For a certain value of σ the second percentile of the D_{98} decreases for increasing Σ. To really see the dependence on Σ and σ more thoroughly, surface plots were



(a) The 2nd percentile of the D_{98} for the prostate as function of the random error in absence of a systematic error.



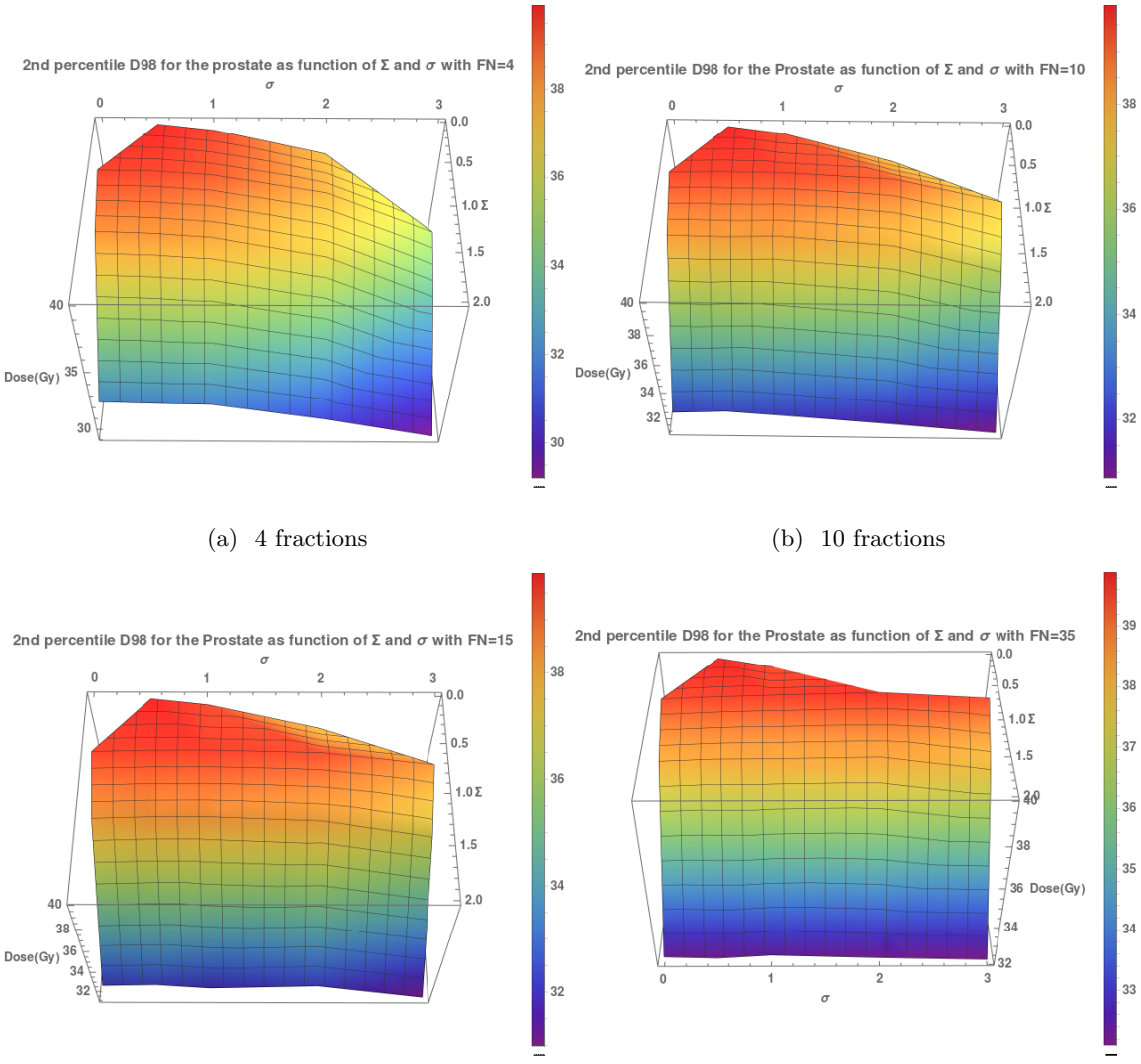
(b) The 2nd percentile of the D_{98} for the prostate as function of the random error for a systematic error of 1 mm.



(c) The 2nd percentile of the D_{98} for the prostate as function of the random error for a systematic error of 2 mm.

Figure 22: The second percentile of the simulated poppulation's D_{98} distribution as function of σ for different fraction numbers for $\Sigma = 0, 1$ and 2 mm.

created for different fraction numbers where the second percentile of the D_{98} was plotted as function of Σ and σ. Plots for 4, 10, 15 and 35 fractions are depicted in Figure 23.



(a) 4 fractions

(b) 10 fractions

(c) 15 fractions

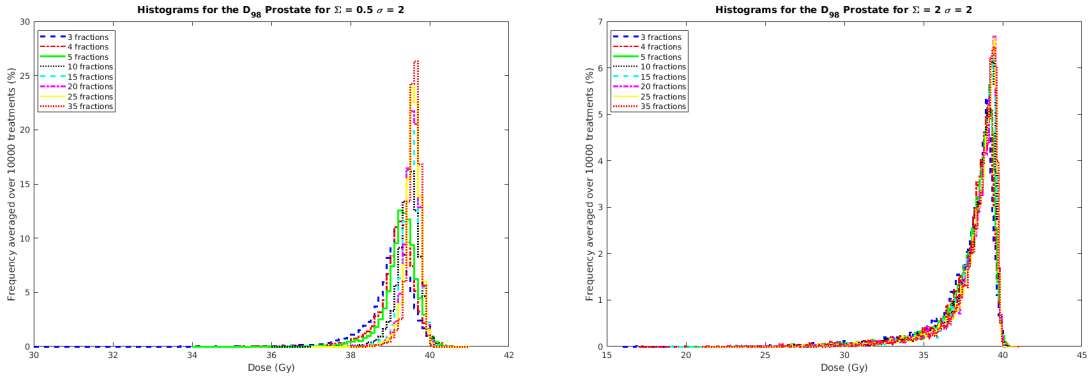
(d) 35 fractions

Figure 23: The second percentile of the prostate D_{98} distributions for 4 fractions (a) 10 fractions (b) 15 fractions(c) and 35 fractions (d) as function of Σ and σ

From the surface plots in Figure 23 we see that the second percentile of the D_{98} decreases faster with increasing systematic error than it does for increasing random error. In the case of 4 fractions no increase in the second percentile of the D_{98} is observed for increasing values of the setup errors. In the case of 10 fractions an increase in the second percentile of the D_{98} is observed for random errors between 1 and 2 mm with a systematic error between 0.7 and 0.8 mm with respect the case of 4 fractions. In the case of 15 and 35 fractions the area of the increased second percentile of the D_{98} extends to a random error of 3 mm. From Figure 23 it is clear that the increase in the second percentile of the D_{98} with respect to the case of four fractions, becomes stronger with increasing fraction number.

Histograms of the distribution of the D_{98} were also generated for all combinations

of systematic and random set up errors that are displayed in Figure 23, two of these histograms are depicted in Figure 24. The width of the distributions widens for increasing random error and increasing systematic error. The distribution of the D_{98} for the PTV and CTV/prostate was found to be asymmetric as its tail extends to lower doses.



(a) Histogram D_{98} prostate for $\Sigma = 0.5$ mm and $\sigma = 2$ mm

(b) Histogram D_{98} prostate for $\Sigma = 2$ mm and $\sigma = 2$ mm

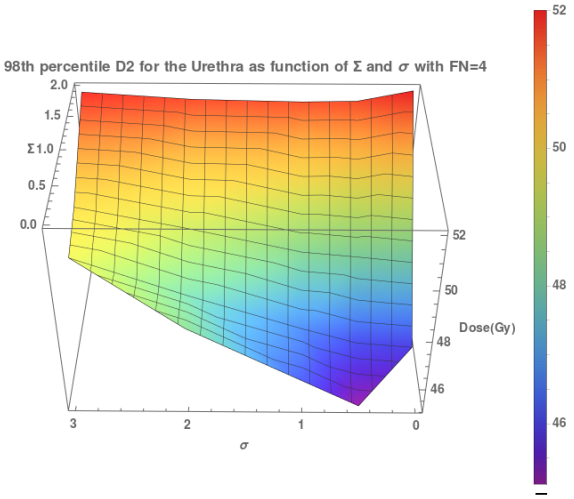
Figure 24: A comparison of the distribution of the D_{98} of a simulated population that receives 3, 4, 5, 10, 15, 20, 25 and 35 fractions. The systematic setup error is 2 mm combined with a random error of 0.5 mm (a) and 2 mm (b). Note that the axes are scaled differently.

For large values of the systematic error, the widening of the D_{98} distribution with increasing fraction number becomes negligible compared to the widening effect of the increased systematic error.

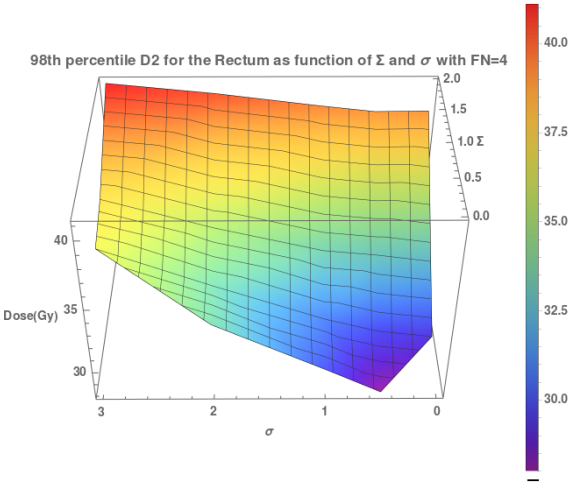
6.2.3 Near maximum dose

Patients' PCEs were also sampled to investigate the effect of fractionation on the distribution of the D_2 of the rectum, urethra and bladder. Especially for serially structured organs like the urethra it is important to have high control over the maximum dose it receives. Because the high doses are in the OARs are of interest, the 98th percentile of the distribution of the D_2 was calculated.

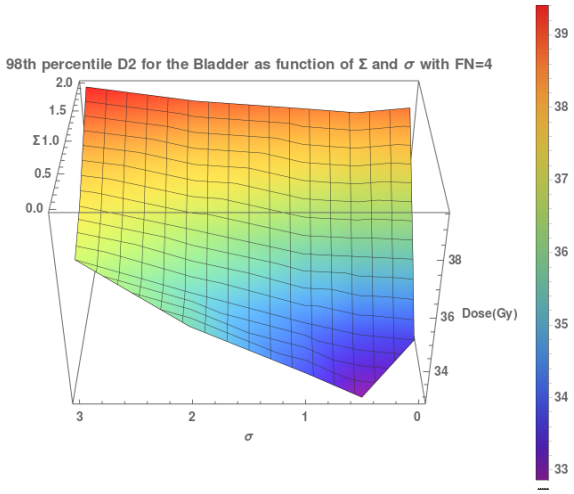
For the simulated treatments with a certain fraction number the 98th percentile of the near maximum dose D_{98} for the organs at risk has been plotted as function of Σ and σ. This resulting surface gives the dependence of the near maximum dose as function of the different set up error. For a treatment given in four fractions these plots are depicted in Figure 25. For a fixed systematic error $\Sigma = 0.5$mm the results are depicted in Figures 26a to 26c.



(a) Urethra D_2 98th percentile in case of 4 fractions



(b) Rectum D_2 98th percentile in case of 4 fractions



(c) Bladder D_2 98th percentile in case of 4 fractions

Figure 25: The 98th percentile of the D_2 distributions for the urethra (a), rectum(b) and the bladder (c) in a treatment that is fractionated in four fractions as function of the systematic and random error.

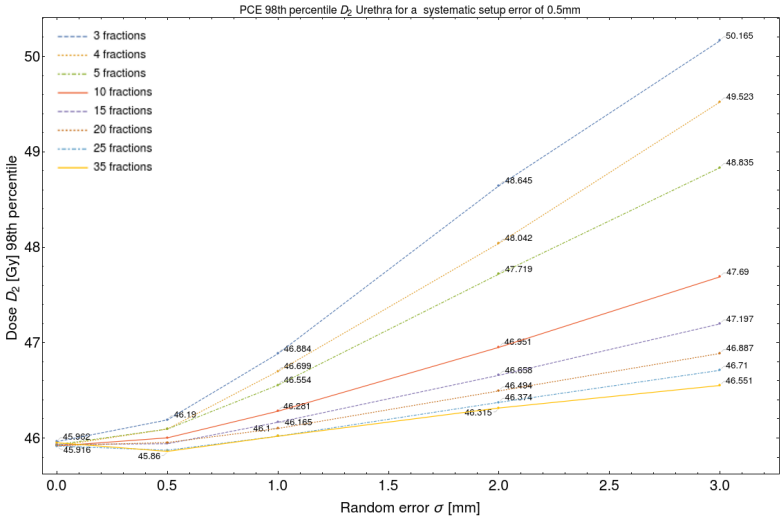
From Figure 25 it can be observed that the 98th percentile of the high dose in the OARs increases with increasing random error and systematic error. The increase is stronger for increasing systematic error than for increasing random error. The high dose in the rectum was found to increase strongly for increasing systematic error. In the urethra for a systematic error of 2mm the 98th percentile of the high dose stays about constant for random errors between 0 and 3mm.

The 98th percentiles of the D_2 have also been plotted for a fixed systematic error of 0.5 mm and varying random error for fractionated treatments in 3, 4, 5, 10, 15, 20, 25 and 35 fractions, these plots are depicted in Figures 26a to 26c. From Figures 26a to 26c it can be seen that the 98th percentiles of the D_2 decrease for increasing fraction number for all investigated combinations of the systematic and random error. The difference becomes less pronounced for increasing fraction number.

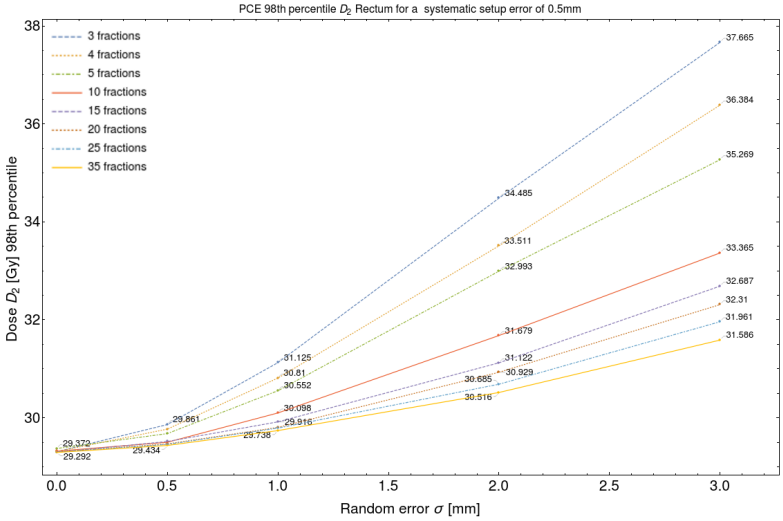
The D_2 distribution has been plotted for all investigated fraction numbers all the combinations of (Σ, σ) that were investigated, in Figure 27 the results of the combinations $(0.5, 2)$ and $(2, 2)$ are shown. From Figure 27 we observe that the mean value for the near maximum dose for the organs at risk is about independent of the fraction number but, the spread of the D_2 distributions widens for increasing fraction number. For increasing σ and a fixed value of Σ the distribution was also found to widen. For $\Sigma > 2$ the distributions are wider than for smaller Σ but the widening effect with increasing fraction number becomes negligible.

The shape of the distributions of 3,4 and 5 fractions and of 15, 20, 25 and 35 were found to be very similar. Especially for the urethra a shape difference of the distributions is observed going from 5 fractions to > 10 fractions.

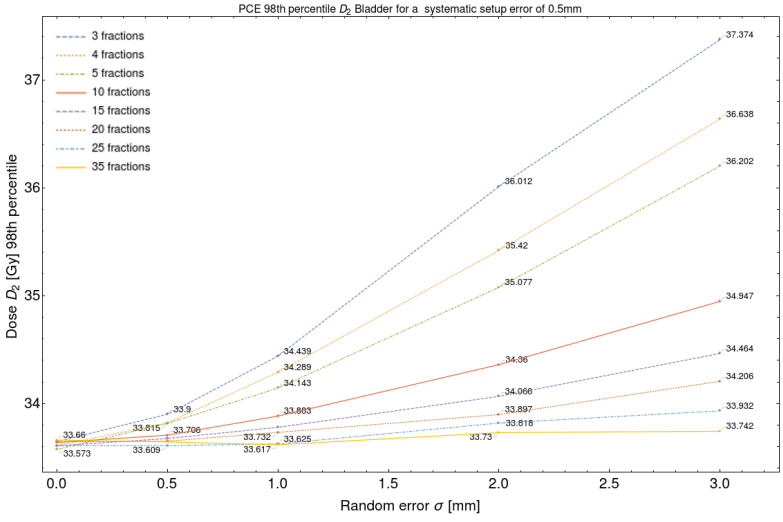
The distribution of the D_2 for the urethra is asymmetric in the sense that its tails extent to high doses. The distribution of the D_2 for the rectum and bladder was found to be (nearly) symmetric.



(a) 98th percentile of the urethra D_2



(b) 98th percentile of the rectum D_2



(c) 98th percentile of the bladder D_2

Figure 26: A comparsion between the 98th percentile of the D_2 in the urethra (a), rectum (b) and bladder (c) for treatments fractionated in 3, 4, 5, 10, 15, 20, 25 and 35 fractions for a fixed systematic error of 0.5 mm against the random error.

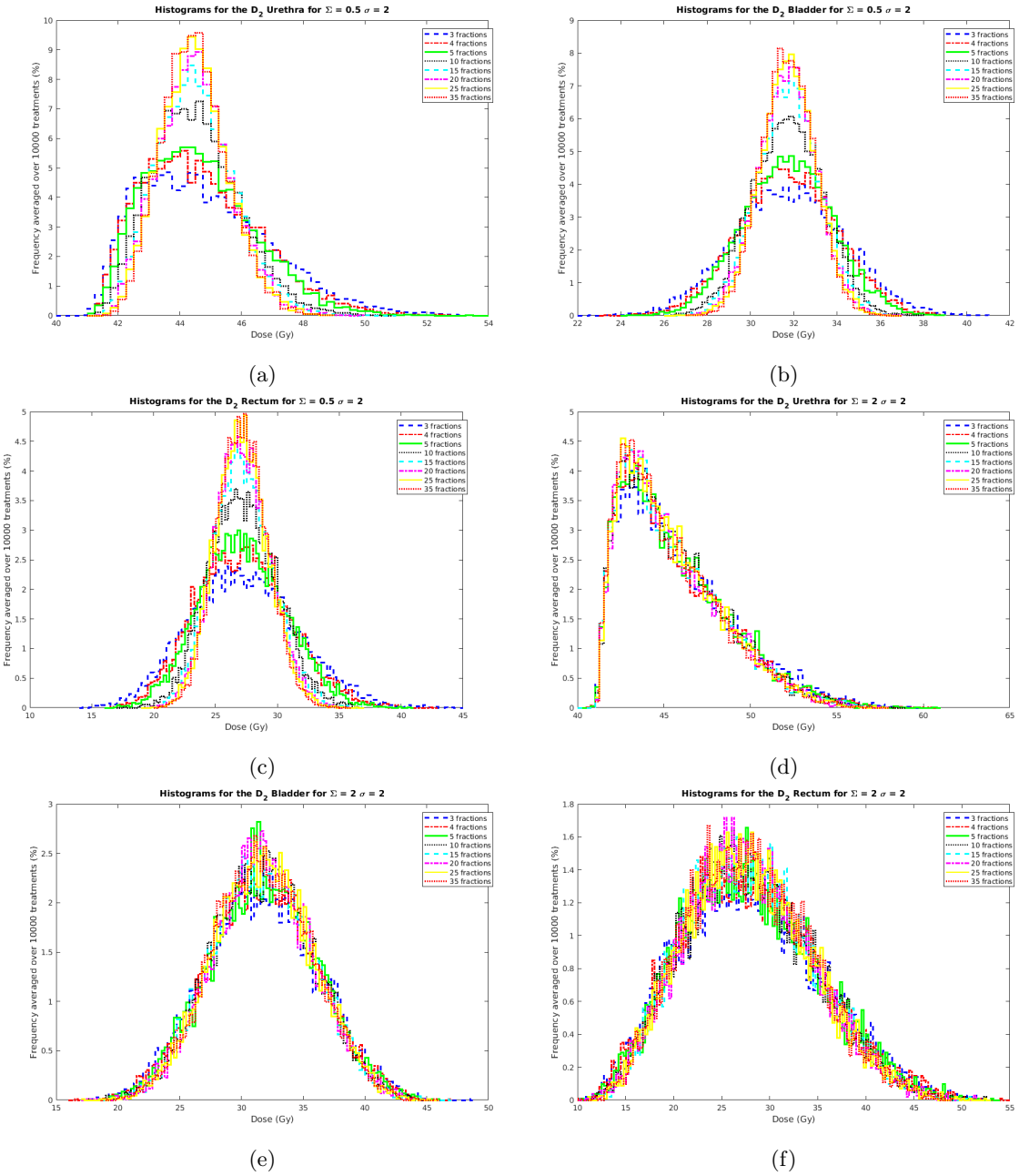


Figure 27: The distribution of the D_{98} of the prostate and the D_2 of the urethra, bladder and rectum for the standard deviations (Σ, σ) $(0.5, 2)$ and $(2, 2)$

6.3 Margin recipes

Recipes were constructed as explained in Section 5.2 for Cyberknife plans with 3 mm, 4 mm and 5 mm isotropic CTV to PTV margin for fraction numbers $2, 3, 4, 10, 30$ and 39. Unfortunately no margin recipe was derived for VMAT as treatment plan optimisation (explained in Section 4.1.1) did not converge for larger margins than 3 mm for the desired constraints and objectives. For a given fraction number the margin recipes indicate the required margin that is needed to have the probability for the CTV coverage at $P(V_{100}^{CTV} \geq 98\%) \geq 97.95\%$. The margin recipes were validated for 10 other patients, the results of the recipe validation are shown in Section 6.3.2. This chapter ends with a comparison between the margin recipes for the given fraction numbers and the linearised van Herk recipe in Section 6.3.3.

6.3.1 Margin recipes fits

The points that meet the chosen criterion (Equation 32) are fitted to the function in Equation 35. The fitting function was chosen to be a rational relation between the CTV-PTV margin, systematic setup error and the random setup error. This function was the same function as used in the work of C. Ter Haar [12] for the setup robustness recipes as the points appeared to follow more or less the same trend. It contains a polynomials of degree one and two in σ and coefficients P_1, P_2, P_3 and P_4 that are third order polynomials of the CTV to PTV margin M_{PTV}. The polynomials P_1, P_2, P_3 and P_4 are given by Equations 36 to 39.

$$\Sigma(\sigma, M_{PTV}) = \frac{P_1 * \sigma + P_2}{\sigma^2 + P_3 * \sigma + P_4} \tag{35}$$

$$P_1 = a * M_{PTV}^3 + b * M_{PTV}^2 + c * M_{PTV} + d \tag{36}$$
$$P_2 = e * M_{PTV}^3 + f * M_{PTV}^2 + g * M_{PTV} + h \tag{37}$$
$$P_3 = i * M_{PTV}^3 + j * M_{PTV}^2 + k * M_{PTV} + l \tag{38}$$
$$P_4 = m * M_{PTV}^3 + n * M_{PTV}^2 + o * M_{PTV} + p \tag{39}$$

The fits for all investigated fraction numbers are depicted in Figure 28, the corresponding fit parameters together with their standard error, t-statistic and p-value are given in Table 6a to 6f. The corrected total degree of freedom (DF) and reduced chi squared statistic χ^2_{Red} of the fits are tabulated in Table 7.

From Figure 28 we see that as the fraction number increases, the data is fits worse to Equation 35. This seems especially true for 39 fractions and $M_{PTV} =$ 3 and 4 mm. From Table 7 we can observe that the t-statistic increases for increasing fraction number and that the p-value decreases for increasing fraction number.

For all margin recipes a couple of parameters could be set to 0 and still result in a good fit. For the fits performed for 2 and 3 fractions very large standard errors in the fit parameters are observed, the small values for the t-statistics and large values for the p-value suggest the data being over fitted. Omitting up to three of the parameters that have a very small absolute t-statistic was found to still result in a good fit. For 2 fractions one could for example omit parameters c, e and f which still results in a good fit and a slight increase in the χ^2_{Red} statistic. For 3 fractions a, b, d and e could be omitted, by doing this the χ^2_{Red} becomes $1.51923e$-3.

For the fit performed for 4 fractions the standard error and p-values for the estimated fit parameters have become smaller with respect to the 2 and 3 fractions case. It was observed that omitting 5 parameters still results in good fit. It is hard to tell what parameter should be neglected, looking up from a table for the t value of a two sided distribution at 95% confidence one finds $t = 2.110$. By omitting the parameters e, g, h, i and k the function still fits well, though the values of the t statistic are much larger than 2.110. After omitting these variables, the χ^2_{Red} statistic becomes 1.51923e-3, which is still very small. Omitting other combinations of equal number of parameters also works.

Moving on to 10, 30 and 39 fractions a decrease in the p-values is seen in Table 7 but if one parameter gets removed some P-values for other parameters increases. Omitting three to four fit parameters still result in a good fit if chosen wisely.
Ideally one would fit all data to one function $\Sigma(M_{PTV}, FN, \sigma)$ that is a function of the margin M_{PTV}, fraction number FN and the standard deviation of the random error σ but no functional has been found yet to fit the data accurately. For the three investigated values of M_{PTV} the margin required per fraction number have been plotted together in Figure 29a to 29c, these plots have been obtained by combining all fits belonging to the same margin from Figure 28.

	Value	Standard Error	t-Statistic	P-Value
a	6.55549	18.8496	0.347779	0.786927
b	-59.053	139.012	-0.424806	0.744266
c	143.498	188.752	0.760245	0.586181
d	-77.3502	175.971	-0.439561	0.736351
e	3.37695	26.6311	0.126805	0.919702
f	-62.1651	196.297	-0.316689	0.804751
g	349.431	266.439	1.31149	0.414725
h	-578.796	248.412	-2.32998	0.258094
i	-77.4709	17.6116	-4.39885	0.142306
j	950.5	129.915	7.31629	0.0864781
k	-3812.53	176.444	-21.6076	0.0294417
l	4972.17	164.49	30.2277	0.0210531
m	19.8117	28.9792	0.68365	0.618239
n	-267.986	213.717	-1.25392	0.42858
o	1185.01	290.19	4.08356	0.152889
p	-1675.92	270.541	-6.19468	0.10189

(a) 2 fractions

	Value	Standard Error	t-Statistic	P-Value
a	-0.0644863	2.94771	-0.0218768	0.98392
b	-2.12085	24.4454	-0.0867587	0.93633
c	16.0602	41.2722	0.389127	0.723159
d	-35.9705	37.2636	-0.965299	0.405601
e	0.218043	3.64867	0.0597597	0.956105
f	2.13917	29.8537	0.071655	0.947386
g	-14.7069	49.729	-0.29574	0.786717
h	29.1512	44.9717	0.648212	0.563039
i	1.0556	4.56005	0.231488	0.831826
j	-15.0055	38.0361	-0.394507	0.719579
k	69.5477	64.5857	1.07683	0.360438
l	-116.416	58.2734	-1.99776	0.139629
m	-2.0781	6.08806	-0.34134	0.755357
n	26.382	50.5774	0.521616	0.637999
o	-105.226	85.5046	-1.23065	0.306146
p	152.284	77.1881	1.97289	0.143043

(b) 3 fractions

	Value	Standard Error	t-Statistic	P-Value
a	-4.51544	0.764291	-5.90801	0.0274743
b	44.9938	5.2594	8.55493	0.0133898
c	-151.719	6.98977	-21.7059	0.00211575
d	169.022	6.53362	25.8696	0.0014909
e	-7.72964	2.16385	-3.57217	0.0702139
f	125.566	14.3878	8.72725	0.0128764
g	-581.538	18.2263	-31.9066	0.000980844
h	827.757	17.1824	48.1747	0.000430607
i	17.0342	0.645502	26.3891	0.0014329
j	-209.595	4.72897	-44.3215	0.000508675
k	834.267	6.84065	121.957	0.0000672264
l	-1082.13	6.30718	-171.571	0.0000339694
m	16.7386	1.7792	9.40791	0.0111104
n	-181.342	12.3579	-14.6742	0.00461187
o	659.777	16.5956	39.756	0.000632094
p	-791.245	15.4857	-51.0954	0.000382814

(c) 4 fractions

	Value	Standard Error	t-Statistic	P-Value
a	0.29507	0.419507	0.703372	0.504539
b	-2.95904	3.0835	-0.959636	0.369204
c	3.40927	4.21285	0.809254	0.445001
d	5.6994	3.92295	1.45283	0.189579
e	0.141591	1.55841	0.0908557	0.930152
f	0.955382	11.3894	0.0838834	0.935498
g	14.0384	15.4488	0.908705	0.393715
h	-46.9262	14.4031	-3.25806	0.0139015
i	13.9208	0.460237	30.2469	1.113538e-8
j	-163.897	3.41714	-47.9634	4.480258e10
k	625.902	4.73834	132.093	3.759006e-13
l	-785.286	4.40139	-178.418	4.585748e-14
m	30.4375	1.62496	18.7313	3.069752e-7
n	-372.613	12.0008	-31.0489	9.282821e-9
o	1505.82	16.4934	91.2985	4.981889e-12
p	-1972.19	15.3434	-128.537	4.549955e-13

(d) 10 fractions

	Value	Standard Error	t-Statistic	P-Value
a	-0.133594	0.106257	-1.25727	0.229224
b	0.534658	0.759855	0.703632	0.493199
c	-1.10723	1.05113	-1.05337	0.310014
d	-1.41113	0.97615	-1.44561	0.170298
e	0.229324	0.554937	0.413244	0.685688
f	6.83968	3.86687	1.76879	0.098701
g	-41.9584	5.17404	-8.1094	1.168764e-6
h	72.0294	4.83154	14.9082	5.520126e-10
i	1.38368	0.0968151	14.292	9.632448e-10
j	-16.6116	0.727092	-22.8466	1.758269e-12
k	64.2008	1.07493	59.7258	2.928147e-18
l	-89.9781	0.987993	-91.0716	8.088319e-21
m	3.02422	0.468477	6.45542	0.0000150904
n	-33.4954	3.42111	-9.79081	1.21299e-7
o	130.051	4.87469	26.6789	2.0996930e-13
p	-146.454	4.50559	-32.505	1.377848e-14

(e) 30 fractions

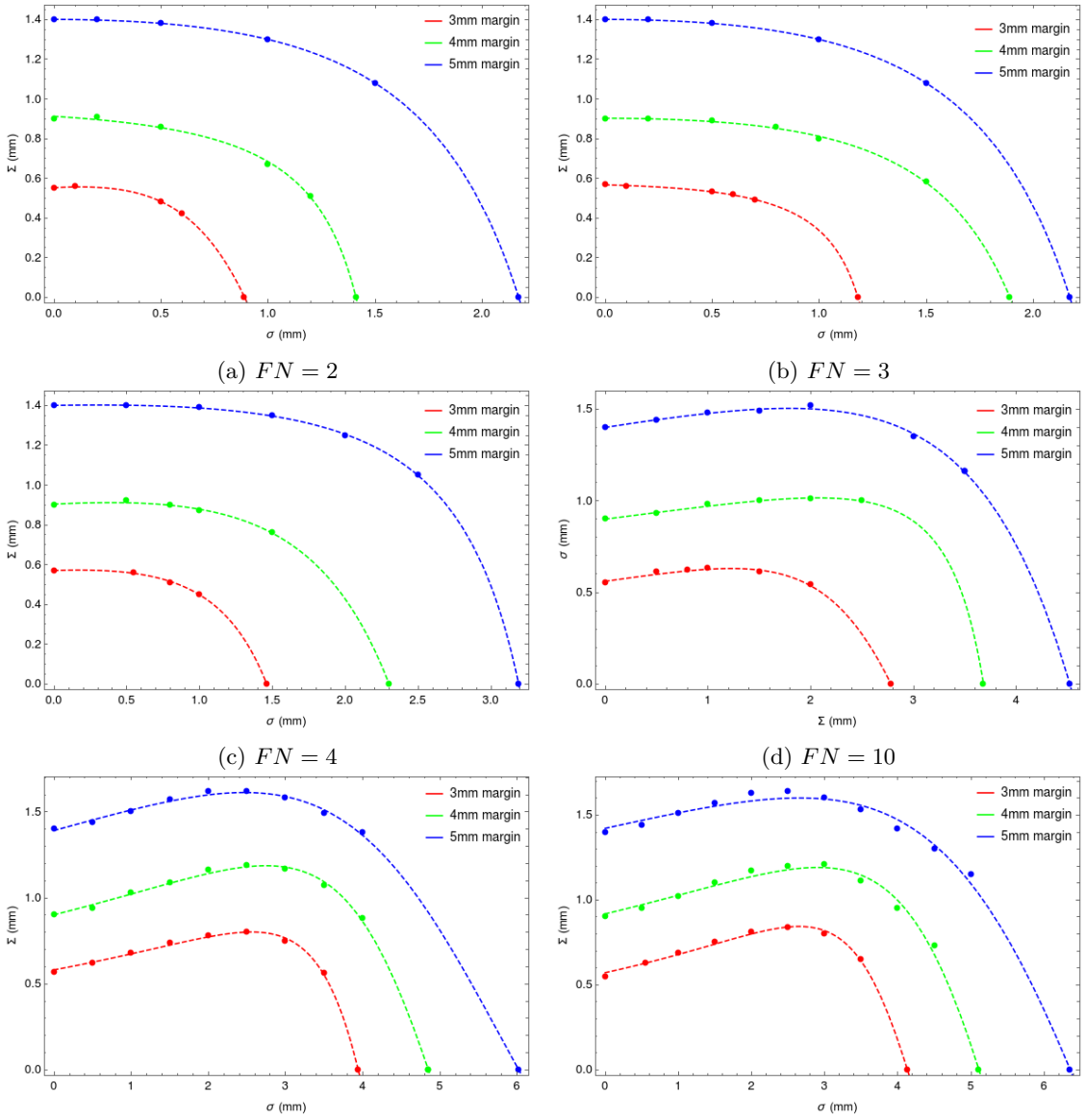
	Value	Standard Error	t-Statistic	P-Value
a	-2.89843	0.236273	-12.2673	1.494306e-9
b	32.8263	1.65907	19.7859	1.129564e-12
c	-126.128	2.23099	-56.5348	7.4580157e-20
d	158.189	2.08171	75.9898	6.681822e-22
e	-9.99652	1.3264	-7.53657	1.192142e-6
f	137.268	9.1043	15.0772	7.074208e-11
g	-569.731	11.9137	-47.8214	1.069612e-18
h	756.154	11.1683	67.7054	4.213610e-21
i	-11.1596	0.194444	-57.3922	5.868314e-20
j	133.771	1.43468	93.2409	2.548717e-23
k	-525.971	2.06128	-255.167	2.606467e-30
l	666.413	1.90288	350.212	1.645614e-32
m	-159.593	1.03916	-153.579	8.760496e-27
n	1920.78	7.43159	258.461	2.122976e-30
o	-7525.54	10.2517	-734.077	1.186039e-37
p	9620.41	9.52513	1010.	7.191991e-40

(f) 39 fractions

Table 6: Tables (a) to (f) show the estimated values for the fit parameters together with their standard error, t-statistic and P-value for each fraction number individually.

FN	2	3	4	10	30	39
χ^2_{Red}	6.72972e-4	9.7863e-5	1.36326e-4	1.90028e-4	3.13934e-4	1.433754e-3
DF	16	18	17	22	29	31

Table 7: The tabulated values of the corrected total degrees of freedom DF and the reduced chi squared statistics corresponding with the different fits depicted in Figure 28

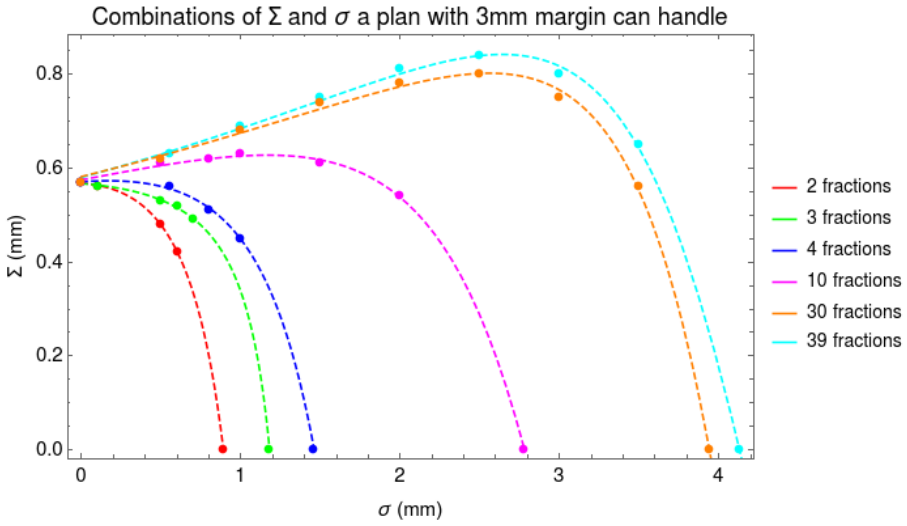


(a) $FN = 2$ (b) $FN = 3$

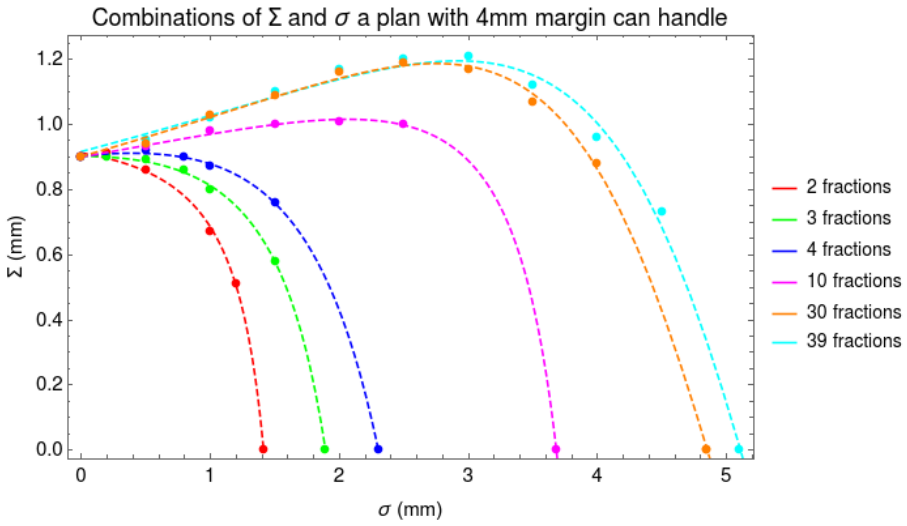
(c) $FN = 4$ (d) $FN = 10$

(e) $FN = 30$ (f) $FN = 39$

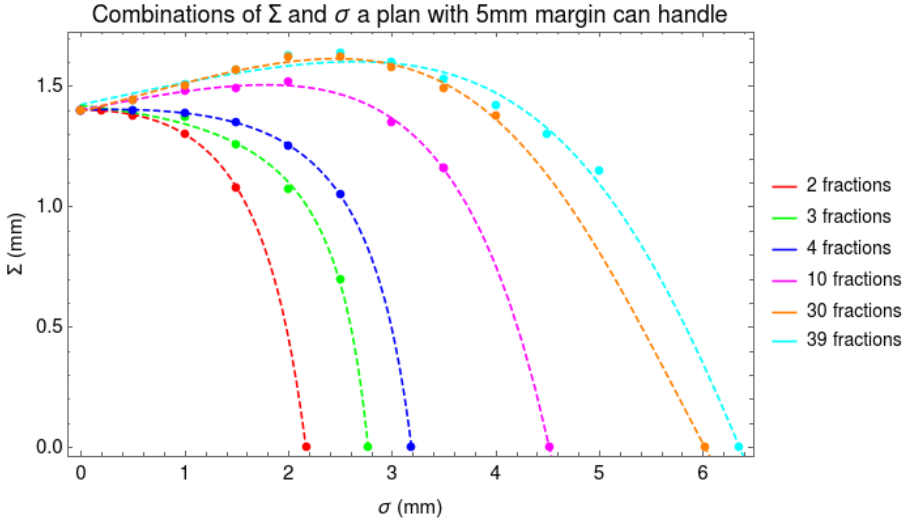
Figure 28: The margin recipes fitted for different fractions (a) to (f). The corrected total degrees of freedom DF and the reduced chi squared statistics are tabulated in Table 7, the parameter values are tabulated in Table 7



(a) $M_{PTV} = 3$ mm



(b) $M_{PTV} = 4$ mm



(c) $M_{PTV} = 5$ mm

Figure 29: Points that were used to construct the margin recipes in Figure 28 collected per margin on a single plot. Curves are drawn to connect the points.

6.3.1.1 Patient dose distribution under spatial translation

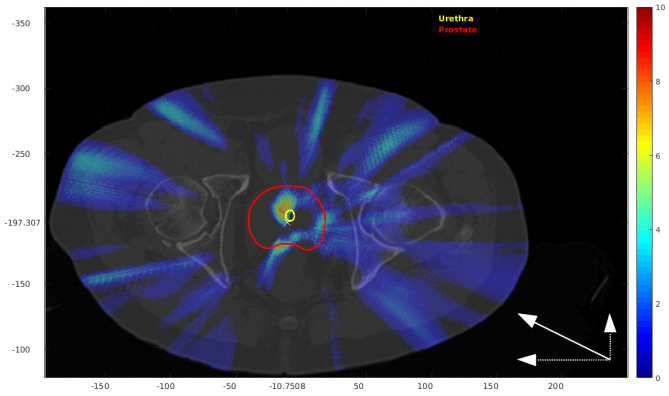
From the margin recipes it can be observed that for treatments fractionated in more than 10 fractions, the CTV coverage increases meaning that the dose the CTV receives increases. This phenomenon is accordance with the increase that was observed for the expected PTV coverage and the 2nd percentile of the $D98$ as depicted in Figure 20, 22 and 23.

To investigate where dose difference arises between a case where an increase in the CTV dose is observed with increasing setup errors, the dose difference has been calculated for a certain scenario that results in a higher CTV dose, this dose difference is depicted in Figure 30. The dose difference was calculated with use of the clinical dose engine. For this particular fraction the total setup error of the dose distribution was:

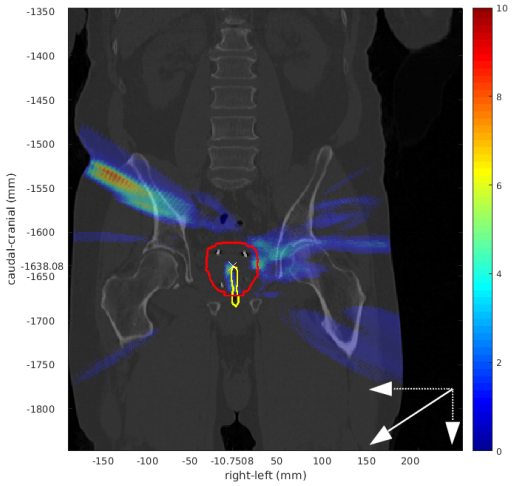
$$-1.1122\hat{x} - 0.62675\hat{y} - 0.9842\hat{z} \tag{40}$$

where the unit vectors in Equation 40 denote the displacement of the isocenters of the beams along the right-left, dorsal-ventral and caudal-cranial direction for $\hat{x}$, $\hat{y}$, $\hat{z}$ respectively. The dose difference was computed by subtracting the nominal dose from the dose in the translated scenario. The median dose difference in the prostate was found to be 0.017 Gy.

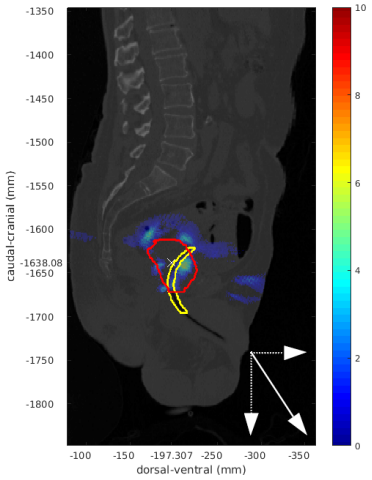
The dose difference for voxels with a positive dose difference is plotted in Figure 30 in the transverse, saggital and coronal plane on top of the patient's CT. The largest dose differences In Figure 30 within the prostate are observed around the urethra. Especially in the transverse and saggital plane it appears that the urethra receives a much higher dose under translation of the dose distribution with respect to the nominal scenario and that this increase in the dose is not compensated for by lower dose on the edge of the prostate. The latter suspicion could be confirmed, in Figure 18 and 19 it can be seen that the voxel dose just outside the edge of the of the prostate is equal to or larger than in the center of the urethra. Another large discrepancy in the voxel dose Figure 30 can be observed around the rectum, where the dose gradient is large.



(a) Transverse plane



(b) Coronal plane



(c) Saggital plane

Figure 30: Dose difference plots for a patient under spatial translation that resulted in a higher CTV coverage depicted on the transverse plane (a) coronal plane (b) and saggital plane (c).The arrows indicate the dose distribution shift within that plane, the delineation of the prostate an urethra are also depicted, these organ positions belong to the nominal plan.

6.3.2 Validation of the margin recipes

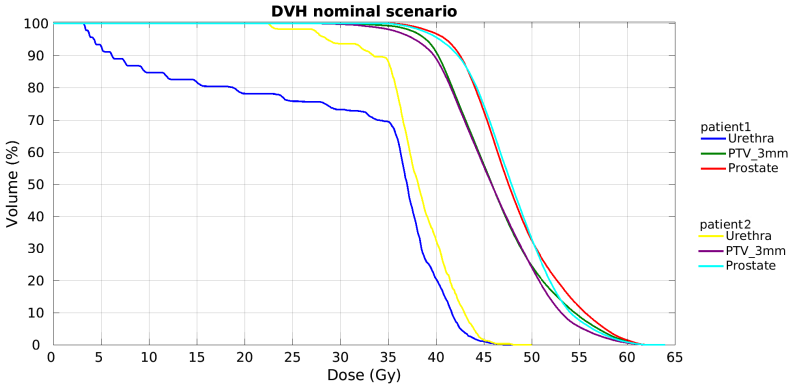
After the construction of the margin recipe based on patient 1, the validity of it must be investigated for all other patients. Ten other patients were randomly selected for this purpose. To compensate for intra patient variations in the plan quality, patient treatment plans are rescaled.

6.3.2.1 Rescaling patient plans

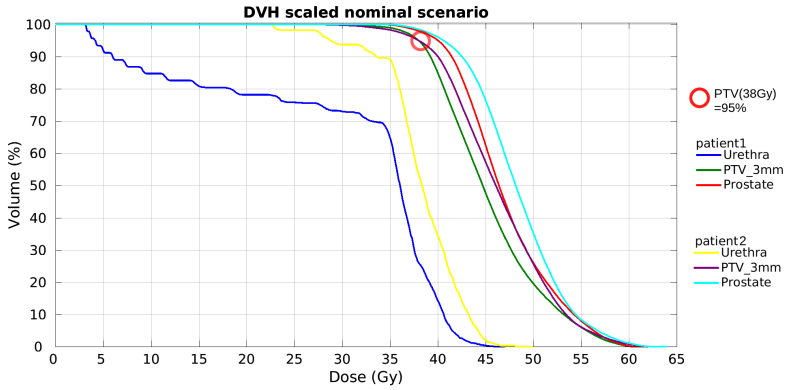
For the validation of the constructed margin recipes 10 other patients' treatment plans were randomly selected. These treatment plans are scaled in the same way as for the construction of the margin recipes, i.e. $V_{100}^{PTV} = 95\%$.

A comparison of the DVH between patient 1 and patient 2 for the PTV, prostate and urethra of the nominal plan with and without scaling is shown in Figure 31.

Although the treatment plans are scaled the same way, the DVHs are very different due to intra patient anatomical variations. In nominal scenario we see that a larger volume of the PTV and prostate of patient 1 receives a high dose$\geq$ $52Gy$ than for patient 2. If the plans are rescaled this coverage difference for the high doses vanishes, the treatment plan of patient 2 after rescaling a difference can be observed in the 38 Gy to 54 Gy region in the prostate and PTV.



(a)



(b)

Figure 31: DVH comparison of the recipe patient, patient 1, and one of the validation patients, named patient 2 of the nominal treatment plan (a) and the rescaled plan (b). Scaling was performed in accordance to Equation 33, the red circle indicates the reference point of rescaling.

All treatment plans have a low dose region in the CTV that prevents over dosage to the urethra. For the validation part it is important to investigate what percentage of the prostate belongs to the urethra and what percentage of the prostate is under dosed in the nominal planned dose. As the clinically acceptable criterion is chosen to be $P\left(V_{100}^{CTV} \geq 98\%\right) \geq 97.95\%$ the under dosed part to the prostate that spares the urethra should not be greater than 2%. To this purpose a voxel count in absolute and relative number have been performed for all validation patients, the results have been tabulated in Table 8.

From the first column of Table 8 we observe that there is a variation in prostate size among patients. The largest prostate, which belongs to patient 1, is more than twice the size of the smallest prostate, which belongs to patient 11. Another observation is that there is a difference in plan quality among the patients in terms of the CTV coverage, as can be seen in the fourth column of Table 8. The percentage of the CTV that is under dosed in the nominal treatment plan ranges from about

	CTV	$urethra_{Encl}$	% $urethra_{Encl}$	$CTV < 38Gy$	% $CTV <$ 38Gy	% $CTV_{Scaled} <$ 38Gy
Patient 1	72361	1384	1.9126	861	1.1899	0.9356
Patient 2	43028	1066	2.4775	717	1.6664	0.7228
Patient 3	74006	1295	1.7499	815	1.1013	0.5351
Patient 4	34294	961	2.8022	284	0.82813	0
Patient 5	52537	1249	2.3774	749	1.4257	0.0952
Patient 6	46611	1010	2.1669	453	0.97187	0
Patient 7	49557	1177	2.375	419	0.84549	0
Patient 8	56113	1400	2.495	581	1.0354	0
Patient 9	53529	1040	1.9429	531	0.99199	0
Patient 10	37396	1138	3.0431	464	1.2408	0.0053
Patient 11	29543	1188	4.0213	324	1.0967	0

Table 8: Absolute and relative voxel count of the prostate, the part of the urethra that is enclosed by the prostate and the parts of the prostate that receive less than the prescribed dose of the treatment plan. The first column shows the absolute voxel count of the CTV , which is the prostate. The second column show the absolute voxel count of the $urethra_{Encl}$ which is the part of the urethra that is contained within the CTV. The third columns shows what percentage of the CTV is occupied by the $urethra_{Encl}$. The fourth column shows the prostate voxel count for parts that receive less than the prescription dose of 38 Gy. The fifth column shows what percentage of the prostate gets less than the prescription dose and the sixth column shows what percentage of the prostate gets less than the prescription dose in the rescaled nominal case.

0.8% to 1.7%.

After rescaling the under dosed voxel of the CTV ranges from 0 to about 1%. Patient 1, the recipe patient, has the highest percentage of under dosed voxels in the nominal plan. Patient 3 and patient 4 have the highest under dosed fraction of the CTV among the validation patient after rescaling.

6.3.2.2 Validation

Margin recipes were validated by simulating 10^5 simulated treatments under set up variations that a treatment plan can handle according the margin recipes. For all fraction numbers and margins at least one validation point has been chosen, these points are tabulated in Table 9. For $M_{PTV} = 4$ and 5 mm a fixed random error was chosen, for $M_{PTV} = 3$ mm points used to construct the margin recipe were reused to save time. For each validation point the probability that $P\left(V_{100}^{CTV}\right) \geq 98\%$ is satisfied and the second percentiles of the D_{98} for each patient are tabulated in Table 10 to 15.

	$M_{PTV} = 3$ mm		$M_{PTV} = 4$ mm		$M_{PTV} = 5$ mm	
Fractions	Σ	σ	Σ	σ	Σ	σ
2	0.4	0.52	0.86	0.5	1.39	0.35
3	0.5	0.53	0.89	0.5	1.39	0.35
4	0.55	0.6	0.91	0.5	1.4	0.35
10	0.62	0.9	0.94	0.5	1.43	0.35
30	0.67	1	0.96	0.5	1.44	0.35
39	0.71	1.2	0.97	0.5	1.47	0.35

Table 9: All chosen validation points per margin and fraction number. The points for $M_{PTV} = 3$ mm were chosen to be combinations that are used for the recipe construction to save computational time, for the other margins a random error was chosen and the corresponding systematic error was calculated using the recipes.

	2 Fractions	3 Fractions	4 Fractions	10 Fractions	30 Fractions	39 Fractions
Patient 1	98.879	98.335	98.123	98.098	98.345	98.136
Patient 2	99.768	99.668	99.506	99.424	99.445	99.307
Patient 3	99.911	99.891	99.841	99.77	99.714	99.603
Patient 4	99.347	99.071	98.662	98.332	98.372	97.821
Patient 5	99.579	99.329	99.082	98.902	98.902	98.711
Patient 6	99.922	99.921	99.873	99.841	99.858	99.816
Patient 7	99.959	99.955	99.905	99.912	99.917	99.866
Patient 8	99.845	99.719	99.601	99.467	99.503	99.35
Patient 9	99.947	99.947	99.929	99.911	99.908	99.844
Patient 10	99.875	99.773	99.724	99.637	99.672	99.608
Patient 11	99.587	99.431	99.284	99.256	99.352	99.181

Table 10: Passing probabilities for the chosen validation points on the margin recipes for $M_{PTV} =$ 3 mm

	2 Fractions	3 Fractions	4 Fractions	10 Fractions	30 Fractions	39 Fractions
Patient 1	98.085	98.034	97.951	97.988	97.988	97.828
Patient 2	99.483	99.499	99.463	99.427	99.464	99.345
Patient 3	99.759	99.742	99.758	99.734	99.736	99.683
Patient 4	99.414	99.38	99.335	99.355	99.355	99.23
Patient 5	99.589	99.523	99.564	99.594	99.594	99.49
Patient 6	99.823	99.773	99.775	99.781	99.769	99.745
Patient 7	99.935	99.951	99.945	99.942	99.942	99.942
Patient 8	99.652	99.701	99.673	99.626	99.626	99.607
Patient 9	99.921	99.923	99.929	99.926	99.925	99.906
Patient 10	99.741	99.781	99.787	99.768	99.767	99.765
Patient 11	99.659	99.688	99.662	99.641	99.671	99.597

Table 11: Passing probabilities for the chosen validation points on the margin recipes for $M_{PTV} =$ 4 mm where $\sigma = 0.50$.

	2 Fractions	3 Fractions	4 Fractions	10 Fractions	30 Fractions	39 Fractions
Patient 1	97.922	98.092	98.033	97.881	97.87	97.758
Patient 2	94.448	94.54	94.545	94.213	94.215	93.687
Patient 3	96.01	96.214	96.139	95.802	95.643	95.19
Patient 4	99.076	99.153	99.128	99.018	99.026	98.912
Patient 5	99.301	99.298	99.347	99.277	99.217	99.097
Patient 6	99.548	99.604	99.578	99.486	99.49	99.377
Patient 7	99.804	99.841	99.823	99.772	99.774	99.733
Patient 8	99.457	99.45	99.4	99.377	99.282	99.207
Patient 9	99.795	99.821	99.81	99.776	99.786	99.714
Patient 10	99.571	99.597	99.593	99.517	99.484	99.439
Patient 11	99.228	99.274	99.23	99.157	99.171	99.017

Table 12: Passing probabilities for the chosen validation points on the margin recipes for $M_{PTV} =$ 5 mm where $\sigma = 0.35$.

	2 Fractions	3 Fractions	4 Fractions	10 Fractions	30 Fractions	39 Fractions
Patient 1	37.66	37.609	37.595	37.602	37.633	37.628
Patient 2	37.938	37.888	37.849	37.846	37.89	37.882
Patient 3	38.275	38.21	38.137	38.105	38.088	38.049
Patient 4	37.927	37.848	37.755	37.7	37.702	37.611
Patient 5	37.816	37.735	37.697	37.685	37.704	37.675
Patient 6	38.269	38.227	38.181	38.204	38.226	38.193
Patient 7	38.245	38.186	38.156	38.158	38.212	38.25
Patient 8	38.076	37.985	37.929	37.891	37.906	37.882
Patient 9	38.435	38.345	38.31	38.291	38.295	38.263
Patient 10	38.044	37.995	37.957	37.956	38.005	38.023
Patient 11	38.027	37.953	37.929	37.932	37.972	37.925

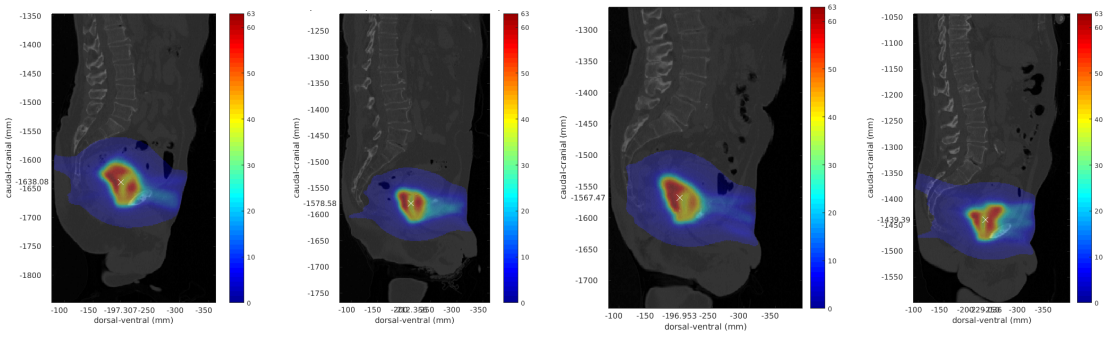
Table 13: Second percentile of the D_{98} for the chosen validation points on the margin recipes for $M_{PTV} = 3$ mm

	2 Fractions	3 Fractions	4 Fractions	10 Fractions	30 Fractions	39 Fractions
Patient 1	37.521	37.526	37.507	37.514	37.506	37.485
Patient 2	38.188	38.215	38.188	38.177	37.137	38.109
Patient 3	38.529	38.523	38.503	38.493	38.463	38.447
Patient 4	38.214	38.202	38.178	38.16	38.128	38.107
Patient 5	38.118	38.089	38.112	38.123	38.086	38.077
Patient 6	38.556	38.549	38.556	38.54	38.532	38.504
Patient 7	39.781	39.796	39.768	39.766	39.723	39.711
Patient 8	38.377	38.416	38.383	38.349	38.329	38.293
Patient 9	39.181	39.169	39.173	39.136	39.117	39.105
Patient 10	38.508	38.515	38.512	38.507	38.468	38.48
Patient 11	38.752	38.807	38.792	38.798	38.794	38.751

Table 14: Second percentile of the D_{98} for the chosen validation points on the margin recipes for $M_{PTV} = 4$ mm.

	2 Fractions	3 Fractions	4 Fractions	10 Fractions	30 Fractions	39 Fractions
Patient 1	37.411	37.465	37.46	37.395	37.397	37.357
Patient 2	36.237	36.274	36.281	36.143	36.101	35.896
Patient 3	36.736	36.849	36.782	36.706	36.629	36.511
Patient 4	38.281	38.4	38.391	38.289	38.22	38.085
Patient 5	38.362	38.397	38.363	38.336	38.255	38.118
Patient 6	39.115	39.13	39.124	39.01	38.965	38.807
Patient 7	40.466	40.575	40.513	40.293	40.311	40.074
Patient 8	38.95	38.999	38.867	38.816	38.71	38.581
Patient 9	39.768	39.885	39.799	39.698	39.681	39.56
Patient 10	39.225	39.3	39.226	39.168	39.054	38.897
Patient 11	38.836	38.893	38.79	38.705	38.789	38.471

Table 15: Second percentile of the D_{98} for the chosen validation points on the margin recipes for $M_{PTV} = 5$ mm



(a) Patient 1 (b) Patient 2 (c) Patient 3 (d) Patient 7

Figure 32: Dose distributions in the saggital plane for patient 1, the recipe patient (a), two poor performing patients in the recipe validation for $M_{PTV} = 5$ mm (c) (d) and the overall best performing patient (d).

For $M_{PTV} = 3$ mm and $M_{PTV} = 4$ mm all patients have better performance than the recipe patient, patient 1. For $M_{PTV} = 5$ mm patient 2 and patient 3 have worse performance than patient 1. The second percentiles of the near minimum dose are given in Table 10 to 12. For many treatment plans the second percentile of the D_{98} is very close to the prescribed dose of $38Gy$.

It appears that for $M_{PTV} = 3$ mm and $M_{PTV} = 4$ mm all patients have better performance than the recipe patient, patient 1. For $M_{PTV} = 5$ mm patient 2 and 3 have worse performance than patient 1, patient 7 has the best overall performance. The worst performing fraction number for patient 3 with $M_{PTV} = 5$ mm is 39 fractions which is still 94.4% of the prescribed dose.The poor performance of patient 2 and 3 for $M_{PTV} = 5$ mm could be due to anatomical differences, the dose distributions have been plotted for patient 1, 2, 3 and 7 in the sagittal plane for comparison in Figure 32. The figures have other dimensions as the dimension of the CT scans differs among patients.

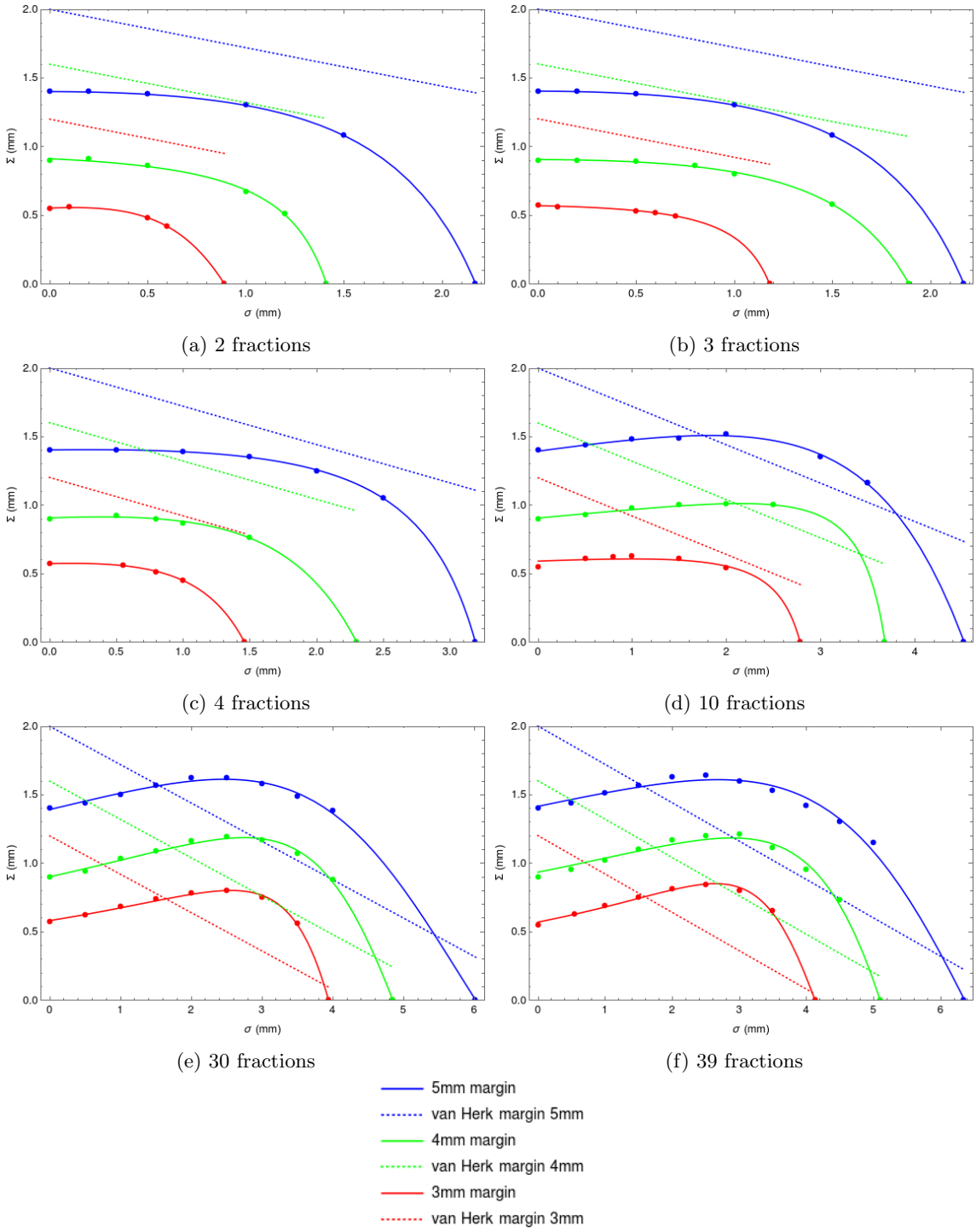
From this figure it is clear that the doses to the prostate is very different from patient to patient. The exact position of the urethra and rectum largely determine the shape of the dose distribution. For patient 3 the urethra sparing lower dose region is more to the front than for other patients. For patient 17 it is in the middle enclosed by two high dose regions. For all four patients a sharp dorsal dose fall off is observed, this dose fall off spares the rectum.

6.3.3 Comparison

in Section 2.2.3 the very well known van Herk margin recipe and its assumptions were briefly discussed. A comparison between the simplified van Herk margin recipe in Equation 8 and the margin recipes derived in this work is shown in Figure 33.

One must notice that the criterion for what is clinically acceptable that was used in this work differs from what van Herk et al. [8] defined. In this work at least 98% of the CTV should receive the full prescribed dose in at least 98% of the patients whereas in the derivation of the margin recipe in Equation 8 the entire CTV should receive at least 95% of the prescribed dose for at least 90% of the patients. In this work due to the urethra sparing dose distribution, the criterion chosen by van Herk is hard to satisfy as the dose to the urethra can be less than 95% of the prescription dose for certain voxels.

Due to the very inhomogeneous prostate dose, a value for the penumbra width is hard to define. Because of this reason the margin recipes are plotted together with the most used margin recipe for 3D dose distributions.



(a) 2 fractions

(b) 3 fractions

(c) 4 fractions

(d) 10 fractions

(e) 30 fractions

(f) 39 fractions

Figure 33: A comparison between the margin recipes that are constructed in this work and the simplified van Herk recipe $M_{PTV} = 2.5\Sigma + 0.7\sigma'$

Least agreement is expected for $\sigma = 0$ mm as no dose blurring occurs in absence of the random error. For $M_{PTV} = 3$ mm according to the recipes constructed in this work the maximum systematic error that is allowed in absence of the random error is described by $\Sigma = 0.55$ mm, according to the simplified van Herk recipe $\Sigma = 1.2$ mm would be allowed. For $M_{PTV} = 4$ and 5 mm according to the recipes constructed in this work the maximum systematic error that is allowed in absence of the random error is described by $\Sigma = 0.90$ and 1.40 mm, according to the simplified van Herk recipe $\Sigma = 1.6$ and 2 mm would be allowed. For $M_{PTV} = 3$ mm the simplified van Herk recipe allows a value for Σ that is more than two times as what we deem acceptable in this work, for $M_{PTV} = 4$ and 5 mm the discrepancy is roughly a factor of 1.4.

Another interesting feature of Figure 33 is the change of discrepancy between the recipes constructed in this work and the simplified van Herk recipe for different fraction numbers. The largest differences are observed for the severely hypofractionated regimes as the margin recipe curves for 3, 4 and 5 fractions and the n Herk lines never intersect. For all investigated values for the systematic and random set up error, van Herk's simplified recipe overestimates the errors a certain margin can handle grossly.

As the fraction number becomes larger the van Herk recipe approaches the margin recipe curves more and more. For $M_{PTV} = 3$ mm two intersections are observed between the van Herk recipe and the recipe constructed in this work for 30 and 39 fractions. Between the intersection points the margin recipe constructed in this work allow larger errors to be handled. As the fraction number becomes larger the van Herk recipe approaches the margin recipe curves more and more. For $M_{PTV} = 4$ and 5 mm two intersections are observed between the van Herk recipe and the recipe constructed in this work for 10, 30 and 39 fractions. Between the intersection points the margin recipe constructed in this work again allow larger errors to be handled. The intersections (Σ, σ) (in mm) for $M_{PTV} = 3$ mm lie around $(0.75, 1.6)$ and $(4, 0.07)$, for $M_{PTV} = 4$ mm they lie around $(1.1, 1.7)$ and $(4.6, 0.3)$ and for $M_{PTV} = 5$ mm they lie around $(1.6, 1.6)$ and $(5.5, 0.5)$.

7 Discussion and conclusion

7.1 Fractionation effects

Fractionation effects have been studied for all organs of interest in prostate cancer treatment. For the prostate it was found that both the second percentile of the D_{98} and the the average of the V_{100}^{CTV} increases for increasing fraction number for both VMAT and Cyberknife doses. This result is expected since when a dose is given in a larger number of fractions, the impact of a single fraction that misses the target decreases due to averaging effects.

An increase in the expectation value of the V_{100}^{CTV} was observed for fraction numbers $10, 15, 20, 25$ and 35 for a systematic error of $0, 0.51$ and 2 mm for increasing random error. This increase was observed for random errors up to $\sigma = 3$ mm, though for $\Sigma = 2$ mm this increase was observed up to $\sigma = 2$ mm. This result is not expected, one would expect the CTV coverage to decrease for increasing random setup errors as the dose on the edge of the CTV blurs out making the penumbra wider for increasing random error.

If we look at the second percentile of the D_{98} of the prostate, an increase with the random error is also observed for treatments fractionated in more than 10 fractions. This effect weakens with increasing systematic error. One could conclude that when the systematic error is not too large, dose blurring actually adds dose to the CTV.

The increase in CTV coverage with increasing setup errors was investigated. In Figure 30 the largest dose differences in the prostate are observed around the urethra. Figures 18 and 19 show that the voxel dose just outside the edge of the prostate is equal to or larger than in the center of the urethra. It could be that the urethra receives a much higher dose under translation of the dose distribution with respect to the nominal scenario and that the increase in the CTV dose due to the increased urethra dose is not compensated for by lower dose on the edge of the prostate.

Another large discrepancy in the voxel dose Figure 30a can be observed around the rectum, which is situated at around $(-10.8, -170)$. As the dose fall off on posterior side of the prostate is very steep, small shifts can cause large dose differences. For this particular shift, the lower dose around the rectum site of the prostate shifts into the higher dose region causing this discrepancy.

For the organs at risk one observes that the 98th percentile of the high dose increases with increasing random error and decreases with increasing fraction number. This is as one would expect as many OAR boundaries face the high PTV voxel doses and dose blurring on the edge causes the dose to the OAR to increase. This blurring to higher doses can also be observed from the symmetry of the D_2 distribution in Figure 27 of the OARs that are adjacent to the prostate as the tails of the D_2 distribution of the urethra and rectum have tails that extent to higher doses.

7.2 Margin Recipes

Margin recipes were created by fitting the functional given in Equation 35 to points that satisfy the criterion in Equation 32. The fit parameters and their standard error, t-statistic and p-value are given in Table 6. After the construction of the margin recipe, it has been validated for other patient treatment plans.

7.2.1 Over fitting of the data

The margin recipes are found to be highly non linear, this is likely due to the very conformal dose distribution. As can be seen from Figure 28 this function can capture the sudden steep fall off behaviour very well, but it must be noticed that the number of data points the function fits to is barely larger than the number of fit parameters for the lower fraction numbers. A natural question that arises is whether the data is being over fitted, especially when one looks at the very small values for the reduced Chi square statistic in Table 7. For all margin recipes a couple of parameters could be set to 0 and still result in a good fit.

All data could be fitted to the function in Equation 35 even if fit parameters are omitted, though not the same parameters for all fraction numbers. The value for the Chi square statistic did stay very small, i.e. in the order of magnitude 10^{-3}, suggesting that the the data points are being over fitted.

For 10, 30 and 39 fractions the increase in CTV coverage with increasing setup errors is noticeable in the recipes as the curves have a positive slope. The chosen fit function fits the data worse as the number of fraction increases. This can also be seen in the extremely small values for the p-value in Table 7, this indicated that the chosen fit function might not be suitable for high fraction numbers.

Optically the fit function fits the data reasonably making the margin recipes suitable to look up the needed margin from Figure 28 given the magnitude of the setup errors and fraction numbers. The margin recipe was validated for 10 patients using the numerical function, which will be discussed in Section 7.2.2.

7.2.2 Margin recipe validation

For the margin recipe validation 10 other treatment plans were selected and rescaled. The validation points are tabulated in Table 9. The probability that Equation 32 is satisfied of the validation patients is tabulated in Table 10 to 12.

7.2.2.1 Rescaling

Although patients are rescaled the same way, anatomical variations that cause differences in plan quality appear to be hard to nullify. The percentage of underdosed voxels still varied among patients, as can be seen in Table 8. The DVHs also did not become more similar after rescaling, as can be seen in Figure 31. Perhaps scaling to a similar D_β within the PTV or CTV would be more suitable to account for this difference in volume that receive a low dose.

7.2.2.2 Validation

From Table 8 one can see that patient 1, the recipe patient, has the highest percentage of under dosed voxels in the nominal plan. This could explain why many validation patients have higher probability of $V_{100}^{CTV} \geq 98\%$.

Patient 3 and patient 4 are the worst performing patients among the validation patients but also have the highest under dosed fraction of the CTV. It could be that for larger errors the effects of anatomical differences are increasingly important. Patient 2 has a comparatively small CTV, certain errors for patient 2 are relatively larger than for patient 1. It could be that due to this size difference and its under dose percentage that these patients have a worse performance. For patient 3 the

poor performance could be due to the position of the urethra within the prostate, as can be seen in Figure 32, the urethra is positioned more off center than for other patients. Most of the CTV dose is deposited behind the urethra. This dose gradient without a second high dose in front of the urethra makes the CTV very prone to shifting out of the high dose region.

It appears that the margin recipes work well for the validation patients, although for a systematic setup error characterised by $\Sigma = 1.4$ mm, which is large for SBRT, the recipe breaks down for two patients. The recipe for the validation points for $M_{PTV} = 3$ and 4 mm the lowest passing probability is 97.828%. The margin recipes appear to be valid if the setup errors are small, for large setup errors anatomical variations may cause the validity to break down.

7.2.3 Comparison with van Herk margin recipe

Some of the assumptions made during the derivation of the Van Herk recipe are hardly valid for the RT modality and dose distributions in this work. The assumption that the dose is fractionated in many fractions obviously is not the case in hypofrationated treatments. On top of the many fractions assumptions, the spherical irradiated target assumption also breaks down. Although one could in principle approximate the prostate as a sphere, the planned prostate dose can hardly be taken spherical.

In this work the planned dose distribution in the prostate is far from a spherical isodose due to the dose constraint for the urethra, thus a large difference in the dose distribution arises within the prostate. Some assumptions that were made during the derivation of the van Herk recipe have also been assumed in this work. Rotations of the prostate were neglected, setup errors were also assumed to be distributed normally and statistically independent and errors were taken isotropic.

From Figure 33 it becomes clear that the simplified van Herk margin recipe does not hold for SBRT. The linear van Herk recipe should not be used for hypofractionated treatments as it overestimates the allowed setup errors grossly for 2, 3 and 4 fractions.

The margin recipes presented in this work show non linear behaviour, due to the high degree of non linearity, linearisation is hardly justified. The suggestion is that margin recipes for very inhomogeneous conformal CTV dose distributions should not be linear. The linear van Herk recipe should not be used when the dose to the CTV is highly conformal and inhomogeneous.

7.3 Conclusion

In this work polynomial chaos expansion was used to model dose distributions. PCE methods have been shown to model the dose distribution delivered by VMAT and Cyberknife accurately in prostate cancer treatment. With the use of PCE the near minimum dose to the prostate D_{98} and the near maximum dose D_2 to the urethra, bladder and rectum have been determined under various setup errors for various fractionation regimes ranging from hypofractionation to conventional hyperfractionation. An increase in the 98th percentile of the near maximum dose to the organs at risk has been observed for increasing systematic and random error and for decreasing fraction number. For large fraction numbers an increase in the 2nd percentile of the near minimum dose to the prostate has been observed for increasing

random error if the systematic error is less or equal to 1.

Margin recipes for SBRT have been derived for various fractionation regimes, these recipes prescribe under what setup errors a CTV-PTV margin of $3, 4$ or 5 mm should be used for a given fraction number to reach the goal of ensuring $V_{100}^{CTV} \geq 98\%$ for at least 98% of the simulated population. The margin recipes were found to be highly non linear. The recipes are given by:

$$\Sigma(\sigma, M_{PTV}) = \frac{P_1 * \sigma + P_2}{\sigma^2 + P_3 * \sigma + P_4} \tag{41}$$

with coefficients P_1, P_2, P_3 and P_4, that are third order polynomials of the CTV to PTV margin M_{PTV}. The polynomials P_1, P_2, P_3 and P_4 are given by Equation 36 to 39 and the values of their coefficients are Table 7.

The validity of the margin recipes has been tested for 10 patients. It was found that although patient dose distributions were all scaled to $V_{100}^{CTV} = 95\%$ differences in CTV coverage still exists. The recipes were found to be valid for almost all patients, except for $M_{PTV} = 5$ mm, where the validity breaks down for two patients perhaps because the chosen validation points has a systematic error that is too large. For the worst performing validation point for the worst performing patient the probability $V_{100}^{CTV} \geq 98\%$ was greater than 94%.

Finally, the recipes were compared to the simplified van Herk recipe. It has been demonstrated that van Herk's simplified recipe overestimates the error a certain margin can handle for severely hypofractionated highly conformal and inhomogeneous CTV dose distributions.

7.4 Future research

PCE extension PCE offers a fast evaluation of the dose, this could be used for more than setup errors alone. One could extent the PCEs in this work with three extra dimensions to incorporate prostate rotation. Each beam is described by its isocenter in three coordinates and a beam angle in three angles. By making six dimensional PCEs one could simultaneously simulate patient translation and prostate rotation. Two assumptions are made in this case.

First the assumption that is made is that dose resulting from rotation of the entire patient is equal to the dose that results from prostate rotations only. This approximation could be a simple but effective as photon dose deposition shows a shallow gradient within the patient. Secondly it is assumed that the prostate does not rotate during a fraction. By incorporating the beam angles as variable, one could do all analysis performed in this work for certain prostate rotations as well.

Verification CTV coverage increase The surprising finding that the CTV coverage increases with increasing random error can be investigated further. One could try to construct PCEs for a highly conformal prostate dose that does not spare the urethra and see if the increase in CTV coverage is still observed. This could help confirming that the behaviour of the CTV coverage is caused by the 'hole' in the dose distribution. Another method to verify this effect is by comparing it to the results from Monte Carlo sampling.

Margin recipe construction The margin recipes in this work are constructed with the goal to ensure that at least 98% of the simulated population receives for at least

98% of the CTV the full prescription dose. This does not constrain the remaining 2% to receive a certain dose. In the most extreme case setup errors that would cause 2% of the CTV to receive no dose whilst 98% receives at least the prescription dose to be considered acceptable. A criterion that would be based on the minimum dose would be interesting

Scaling factor Recipes could also be constructed with a different scaling factor. In this work it was shown that scaling to the same PTV_{100} does not nullify the difference in plan quality in terms of CTV coverage. One could for example scale to the same minimum dose for a certain fraction or volume within the PTV, making the urethra sparing more equal among patients.

Bibliography

[1] Dutch national cancer registration. `https://www.volksgezondheidenzorg.info/onderwerp/prostaatkanker/cijfers-context/huidige-situatie`. Accessed: 2019-02-05.

[2] Claudia Allemani, Tomohiro Matsuda, Veronica Di Carlo, Rhea Harewood, Melissa Matz, Maja Nikšić, Audrey Bonaventure, Mikhail Valkov, Christopher J Johnson, Jacques Estève, et al. Global surveillance of trends in cancer survival 2000–14 (concord-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *The Lancet*, 391(10125):1023–1075, 2018.

[3] Philip Mayles, Alan E Nahum, and Jean-Claude Rosenwald. *Handbook of radiotherapy physics: theory and practice*. CRC Press, 2007.

[4] MichaelJ Zelefsky, StevenA Leibel, PaulB Gaudin, GeraldJ Kutcher, NeilE Fleshner, ES Venkatramen, VictorE Reuter, WilliamR Fair, C Clifton Ling, and Zvi Fuks. Dose escalation with three-dimensional conformal radiation therapy affects the outcome in prostate cancer. *International Journal of Radiation Oncology* Biology* Physics*, 41(3):491–500, 1998.

[5] Joep C Stroom and Ben JM Heijmen. Geometrical uncertainties, radiotherapy planning margins, and the icru-62 report. *Radiotherapy and oncology*, 64(1):75–83, 2002.

[6] David Dearnaley, Isabel Syndikus, Helen Mossop, Vincent Khoo, Alison Birtle, David Bloomfield, John Graham, Peter Kirkbride, John Logue, Zafar Malik, et al. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 chhip trial. *The Lancet Oncology*, 17(8):1047–1060, 2016.

[7] Jack Fowler, Rick Chappell, and Mark Ritter. Is α/β for prostate tumors really low? *International Journal of Radiation Oncology* Biology* Physics*, 50(4):1021–1031, 2001.

[8] Marcel Van Herk. Errors and margins in radiotherapy. In *Seminars in radiation oncology*, volume 14, pages 52–64. Elsevier, 2004.

[9] Sebastian Van Der Voort, Steven van de Water, Zoltán Perkó, Ben Heijmen, Danny Lathouwers, and Mischa Hoogeman. Robustness recipes for minimax robust optimization in intensity modulated proton therapy for oropharyngeal cancer patients. *International Journal of Radiation Oncology* Biology* Physics*, 95(1):163–170, 2016.

[10] Zoltán Perkó, Sebastian R van der Voort, Steven van de Water, Charlotte MH Hartman, Mischa Hoogeman, and Danny Lathouwers. Fast and accurate sensitivity analysis of impt treatment plans using polynomial chaos expansion. *Physics in Medicine & Biology*, 61(12):4646, 2016.

[11] Wei Liu, Xiaodong Zhang, Yupeng Li, and Radhe Mohan. Robust optimization of intensity modulated proton therapy. *Medical physics*, 39(2):1079–1091, 2012.

[12] Carlijn Ter Haar. Polynomial chaos expansion as a tool to construct robustness recipes for proton therapy. *Thesis Master of Science, TU Delft*, 2018.

[13] Teiji Nishio. Physics. In *Stereotactic Body Radiation Therapy*, pages 27–43. Springer, 2015.

[14] Kurt G Hofer, Clifton R Harris, and J Marshall Smith. Radiotoxicity of intracellular 67ga, 125i and 3h: nuclear versus cytoplasmic radiation effects in murine l1210 leukaemia. *International Journal of Radiation Biology and Related Studies in Physics, Chemistry and Medicine*, 28(3):225–241, 1975.

[15] Eric J Hall, Amato J Giaccia, et al. *Radiobiology for the Radiologist*, volume 6. Lippincott Williams & Wilkins Philadelphia, 2006.

[16] JM Thoday and John Read. Effect of oxygen on the frequency of chromosome aberrations produced by x-rays. *Nature*, 160(4070):608, 1947.

[17] Robert F Kallman. The phenomenon of reoxygenation and its implications for fractionated radiotherapy. *Radiology*, 105(1):135–142, 1972.

[18] Ervin B Podgorsak et al. Radiation oncology physics. *Vienna: International Atomic Energy Agency*, pages 123–271, 2005.

[19] Brian D Kavanagh, Moyed Miften, and Rachel A Rabinovitch. Advances in treatment techniques: stereotactic body radiation therapy and the spread of hypofractionation. *The Cancer Journal*, 17(3):177–181, 2011.

[20] Linus C Benjamin, Alison C Tree, and David P Dearnaley. The role of hypofractionated radiotherapy in prostate cancer. *Current oncology reports*, 19(4):30, 2017.

[21] T Landberg, J Chavaudra, J Dobbs, J-P Gerard, G Hanks, J-C Horiot, K-A Johansson, T Möller, J Purdy, N Suntharalingam, et al. Report 62. *Journal of the ICRU*, (1):NP–NP, 1999.

[22] Jane Barret. On target: ensuring geometric accuracy in radiotherapy, 2008.

[23] Marcel Van Herk. Errors and margins in radiotherapy. *Seminars in radiation oncology*, 14:52–64, 2004.

[24] Enzhuo M Quan, Xiaoqiang Li, Yupeng Li, Xiaochun Wang, Rajat J Kudchadker, Jennifer L Johnson, Deborah A Kuban, Andrew K Lee, and Xiaodong Zhang. A comprehensive comparison of imrt and vmat plan quality for prostate cancer treatment. *International Journal of Radiation Oncology* Biology* Physics*, 83(4):1169–1178, 2012.

[25] M Teoh, CH Clark, K Wood, S Whitaker, and A Nisbet. Volumetric modulated arc therapy: a review of current literature and clinical use in practice. *The British journal of radiology*, 84(1007):967–996, 2011.

[26] Pantaleo Romanelli, David W Schaal, and John R Adler. Image-guided radiosurgical ablation of intra-and extra-cranial lesions. *Technology in cancer research & treatment*, 5(4):421–428, 2006.

[27] Zoltán Perkó, Luca Gilli, Danny Lathouwers, and Jan Leen Kloosterman. Grid and basis adaptive polynomial chaos techniques for sensitivity and uncertainty analysis. *Journal of Computational Physics*, 260:54–84, 2014.

[28] George Casella and Roger L Berger. *Statistical inference*, volume 2. Duxbury Pacific Grove, CA, 2002.

[29] Olivier Le Maître and Omar M Knio. *Spectral methods for uncertainty quantification: with applications to computational fluid dynamics*. Springer Science & Business Media, 2010.

[30] Thierry Crestaux, Olivier Le Maıˆtre, and Jean-Marc Martinez. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 94(7):1161–1172, 2009.

[31] Dongbin Xiu and George Em Karniadakis. The wiener–askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing*, 24(2):619–644, 2002.

[32] Douglas C. Montgomery. *Design and Analysis of Experiments*. John Wiley & Sons, Inc., USA, 2006.

[33] K Beddek, S Clenet, O Moreau, V Costan, Y Le Menach, and A Benabou. Adaptive method for non-intrusive spectral projection—application on a stochastic eddy current ndt problem. *IEEE Transactions on Magnetics*, 48(2):759–762, 2012.

[34] Géraud Blatman and Bruno Sudret. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics*, 25(2):183–197, 2010.

[35] Eric Ziegel. Numerical recipes: The art of scientific computing, 1987.

[36] Sergei Abramovich Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. In *Doklady Akademii Nauk*, volume 148, pages 1042–1045. Russian Academy of Sciences, 1963.

[37] Sebastiaan Breedveld, Pascal RM Storchi, Peter WJ Voet, and Ben JM Heijmen. icycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar imrt plans. *Medical physics*, 39(2):951–963, 2012.

[38] Daniel A Low, William B Harms, Sasa Mutic, and James A Purdy. A technique for the quantitative evaluation of dose distributions. *Medical physics*, 25(5):656–661, 1998.