SENSITIVITY AND UNCERTAINTY ANALYSIS IN PROTON THERAPY TREATMENT PLANS USING POLYNOMIAL CHAOS METHODS

C.M.H. Hartman

Master Thesis

Supervisors:
Dr. ir. D. Lathouwers
Dr. M. Hoogeman
Z. Perkó, MSc
S. van de Water, MSc

Delft University of Technology
Faculty of Applied Sciences
Department of Radiation Science and Technology
Section of Nuclear Energy and Radiation Applications

in cooperation with

Erasmus MC Daniel den Hoed Cancer Institute
Department of Radiation Oncology
Unit Medical Physics
June 2014

Abstract

Proton therapy is a novel type of radiotherapy which can deliver superior dose distributions compared to conventional photon radiotherapy. However, setup and range uncertainties can severely compromise the planned dose delivery. Statistical methods can be used for the quantification of uncertainty in the treatment plan, but they are not suitable for clinical use because the computational costs involved are too high. In this thesis project, polynomial chaos is proposed as an alternative method for sensitivity and uncertainty analysis of proton therapy treatment plans. The focus of the project lies on proving that the polynomial chaos method can produce results as accurate as statistical methods, but at considerably lower computational costs.

Treatment plans of two different head-and-neck cancer patients have been created for this thesis project. A treatment plan is evaluated on three criteria: a visual inspection of the full dose distribution, the examination of dose-volume histograms for the tumor and relevant organs and on a list of relevant dose and tumor parameters. The polynomial chaos method is used to produce results for each of these criteria, after which the results are validated against Monte Carlo simulations.

The full dose distribution is built up from the dose distributions in over ten million voxels. Due to computational limitations in Matlab, it is only possible to calculate the full dose distribution with first order polynomials. The dose distributions in several individual voxels have been investigated and first order polynomials were found adequate to accurately describe the dose distribution in half of the voxels. The accuracy with which the expected full dose distribution can be approximated is 0.2 Gy in the presence of only setup uncertainties and 0.4 Gy when range uncertainties are additionally included.

Dose-volume histograms of the expected dose in all investigated organs can be accurately predicted by first order polynomials. Confidence bandwidths around the expected dose can also be adequately approximated by first order polynomials, although post-processing of the results is needed for several organs to avoid unphysical results. Higher order polynomials can reproduce the bandwidths perfectly.

Over 50 dose and tumor parameters have been investigated. Polynomial chaos is able to approximate the uncertainty distribution within acceptable limits with first order polynomials in over 90% of these parameters. The inclusion of higher order polynomials further improves the performance of this method.

For this particular application of polynomial chaos, dose parameters and dose-volume histograms can be calculated directly by choosing them as responses of the system or indirectly by calculating the full dose distribution and deriving relevant parameters from that. The differences between results from the two routes is very small and they can be used alongside of each other. Both routes can produce unphysical results, but this can easily be fixed during post-processing and should be no reason for concern.

It is possible with polynomial chaos expansion to attribute the uncertainty in the output to the uncertainties in the various input variables. In a clinical environment this could be used to focus the alignment of the patient in the direction with the highest contribution to the uncertainty. Variance decomposition of the parameters showed that range uncertainty only has a significant contribution to the variance of the tumor, and not to the uncertainty in any of the investigated organs.

The conclusions from the proof of concept study were found to hold both for robustly and non-robustly generated treatment plans as well as for a second patient case with a completely different shaped tumor. The polynomial chaos method performs slightly better in the case of robust optimization.

Acknowledgements

I would like to thank all my supervisors for their help and guidance during my project. Danny, your ability to always see the bigger picture and your constructive feedback was extremely helpful and helped me get back on track on several occasions. Mischa, thank you for the opportunity to do this MSc thesis at Daniel den Hoed and for your input during the project. Zoltan, I already miss our daily fruit hours. We had many useful discussions on my project, but the gossip and personal conversations we had made fruit hour the highlight of most days during my project. Steven, thank you for tirelessly explaining medical concepts that seemed alien to me at first and for saving me from more than one awkward lunch at the hospital. I would like to thank the whole group of NERA for the fantastic times over the past year, all the wonderful cakes and the many, many beers we drank in the Koepeltje on Fridays. Maurits, thank you for proofreading my thesis and giving me some very useful tips. Finally, I would like to thank Jelle for his support during the past months. You were always able to put a smile on my face even after very long and difficult days in my project.

Contents

Chapter 1

Introduction

The World Health Organization defines cancer as a generic term for diseases involving the rapid creation of abnormal cells that grow beyond their usual boundaries. The collection of abnormal cells is called a neoplasm, more popularly known as a tumor. Even though the term tumor can also designate a benign growth of cells, in this thesis the term is used as a synonym for neoplasm. Tumors can arise when the DNA in a single cell mutates. This mutation is the result of the interaction between a person's genetic factors and his or her environment. Exposure to ultraviolet and ionizing radiation, chemical carcinogens found in asbestos and tobacco for example all increase the probability of certain types of cancer. The tumor, when left untreated, continues to grow at the expense of healthy cells in the vicinity, as nutrients are primarily directed towards the tumor cells. Eventually the abnormal cells can invade adjoining parts of the body or travel through the vascular system and reach other organs or vital structures, a process called metastasis. These new occurrences of cancer, metastases, are the major cause of death from cancer.

In the Netherlands cancer is the primary cause of death for both men and women, responsible for approximately one-third of all deaths equalling 43,666 cases in 2012 ([1],[2]). Furthermore, in the same year over 100,000 people were diagnosed with some type of cancer. The number of cancer diagnoses and deaths has been increasing over the past 20 years as a consequence of the aging population. The number of people currently living with cancer is reflected in the five years prevalence, which is defined as the number of persons who have been diagnosed with cancer within the past five years that are still alive today. It can be seen as a measure of cancer burden. Persons in this group are a mix of people having just been diagnosed, people during treatment or clinical follow-up, or people already cured. For the Netherlands this number has been steadily increasing for the past years, crossing the 300,000 line between 2012 and 2013 [1]. According to this same definition, over 32,6 million people are currently living with cancer worldwide [3].

The type of treatment a person receives after being diagnosed with cancer depends among other things on the size and the location of the tumor. When the cancer is clustered in a specific area, local treatments are applied. If possible, the tumor is physically removed via surgery, often in combination with radiation therapy. Radiation therapy uses ionizing radiation to kill the cancer cells and will be explained further in Section 1.1. When the cancer cells have spread or are likely to spread through the body, systemic treatment methods such as chemotherapy, hormone therapy or immunotherapy are used to treat the patient. The goal of cancer treatment is either to cure the patient by killing all cancerous cells, or prolong life expectancy and reduce the symptoms related to cancer by controlling the tumor, while maintaining the patient's quality of life.

1.1 Radiotherapy

Over half of the patients treated for cancer receives some form of radiotherapy, which is the medical use of radiation to kill or control malignant cells. Ionizing radiation kills cancer cells by damaging the

DNA either directly or indirectly. Indirect damage is done as the radiation ionizes the water in the cell, causing the formation of free radicals which in turn damage the DNA. The radiation can also directly damage the DNA. The most important effects are single or double strand breaks of the DNA, damage to the DNA base or cross links between DNA and a protein for example. In Figure 1.1 various types of radiation damage to the DNA are depicted. Single strand breaks or damage to the base of the DNA can often be repaired within minutes after irradiation by proteins in the cell. Double strand breaks are much harder to repair, since it requires a sister chromatid in the vicinity. When the DNA is not able to repair itself, the cell either lives on as a modified cell or it dies. This holds both for healthy normal cells and cancerous cells. Radiation damage to healthy cells could result in modified cells which in the long term could eventually lead to unwanted side effects or even to the development of secondary tumors. The latter is especially the case for young children with cancer, who have a relatively high life-time risk for secondary tumors due to cancer treatments during childhood.

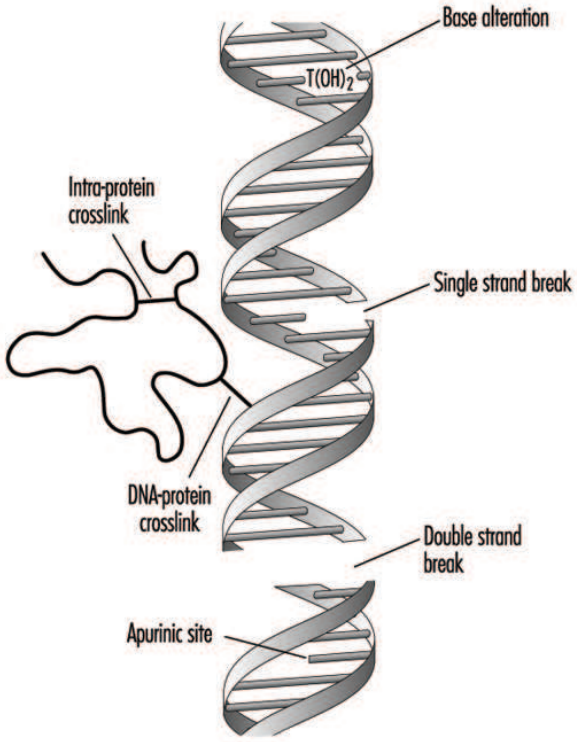


Figure 1.1: Illustration of several types of damage to the DNA induced by radiation. [4]

The amount of radiation a certain tissue receives is called the *dose*, which is the energy absorbed per unit mass, and is measured in Gray ([Gy = J/kg]). An increase in dose leads to a higher percentage of cell death. Fortunately, cancerous tissue is less able to repair the irradiation damage than normal healthy tissue and thus cancer cells have a lower survival rate during treatment. Furthermore, healthy cells are also faster in the repair of the damage than cancer cells. In radiotherapy this property is taken into account in the treatment of patients. Instead of giving patients the complete dose in a single session, the total dose is split into multiple fractions. Patients typically return to the hospital around 30 times to receive parts of their treatment. On average, around 2 Gy of radiation dose is given to the patient per fraction. Between these fractions, the healthy tissue is able to repair a larger part of the damage than cancerous tissue. The dose is administered from outside the body via external beams or from within the patient, in which case a radioactive source is implanted in the body. In Figure 1.2 an example of an external beam treatment facility is shown.

Different types of radiation can be used for radiotherapy. The biological effect of each type of radiation can be evaluated with the relative biological effectiveness (RBE), an empirical value determining the ratio of two types of radiation in terms of their effectiveness to damage biological tissue. The reference value is the damage done by high energy photons, which have an RBE of 1. Charged particles generally have

Figure 1.2: A proton therapy treatment facility [5]. The patient is placed on the couch, which can be shifted to perfectly align with the gantry. Both the couch and the gantry can be rotated during treatment to make dose delivery from different angles possible. Immobilization devices are used to keep the patient fixed in space during both imaging and dose delivery.

a higher RBE, as they create more clustered DNA damage, which is more difficult for the cell to repair. The most commonly used charged particles in external beam radiotherapy are protons, which have an RBE of 1.1, only slightly higher than gamma rays, meaning that the difference in cell kill between protons and photons is small. However, protons have a few characteristics making them especially suitable for radiotherapy.

1.2 Proton therapy

Protons are subatomic particles with a positive electric charge. Because they are charged particles, magnetic fields can be used to control their trajectory. With a mass of roughly 2000 times that of an electron, protons experience little scatter when traveling through a medium, allowing them to travel in a relatively straight line, depositing very little energy along the way. Only at the very end of their range do they interact strongly with their environment, resulting in a strong peak in the energy deposition. This peak is called the Bragg peak and is characteristic for charged particles traveling through a medium. The total dose deposited is determined by the number of protons in the beam. The (finite) range of protons in matter depends both on the initial beam energy and on the ability of the irradiated material to absorb the energy of the protons during interaction, which is defined as the stopping power of the material. By adapting the energy of the incoming beam, the Bragg peak can be shifted in depth. By combining proton beams with different energies, the Bragg peaks can be combined into a 'spread out Bragg peak' (SOBP). This is depicted in red in Figure 1.3, along with the energy deposition profile for a photon and a proton beam.

Figure 1.3 demonstrates the advantage proton therapy has over conventional radiotherapy. First of all, the spread out Bragg peak can be tuned in such a way that the protons lose most of their energy inside the tumor. The tumor therefore receives a higher and more homogeneous dose when compared to conventional radiotherapy. Secondly, behind the tumor virtually no dose is deposited. Furthermore the dose in front of the tumor is considerably lower, with the exception of a slightly higher skin dose for protons (visible in Figure 1.3 at zero depth). The tumor is typically irradiated from different angles, to minimize the burden on tissue located in front of the tumor. Together, these properties allow for a superior conformity in the dose distribution within the patient.

The most advanced technique used in proton therapy for dose delivery is pencil beam scanning. Pencil beams are very narrow unscattered beams of protons which can be aimed at any spot inside the patient. The lateral position of the beam can be altered with the help of magnets and by adapting the energy of the beam the penetration range can be varied. The dose can thus be controlled in three dimensions

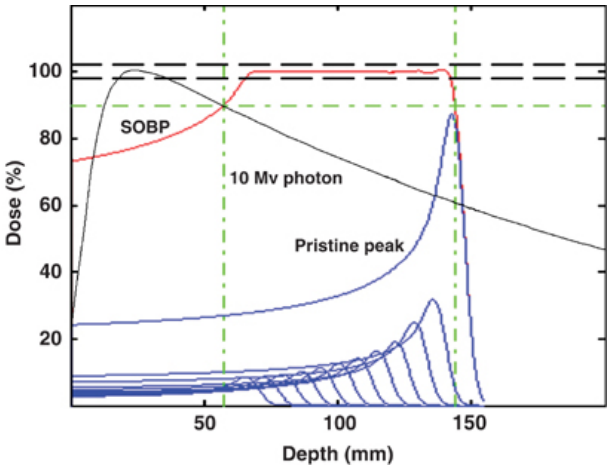


Figure 1.3: Depth-dose curve for a spread-out Bragg peak (red), built up from individual Bragg peaks with different energies (blue). For comparison the depth dose curve for a 10 MV photon beam is shown in black [6].

without the need for field- or tumor-specific hardware. A weight is assigned to each pencil beam, which is proportional to the number of protons in the beam and thus the total amount of deposited dose. These weights are optimized when using intensity modulated proton therapy (IMPT). The individual fields created by the pencil beams can deliver a highly inhomogeneous dose distribution to the target, but the combined fields ensure the desired homogeneous tumor coverage.

For two groups of patients (the more expensive) proton therapy is considered to be a better alternative to conventional photon radiotherapy. The first group consists of patients with tumors in the eye, tumors in the base of the skull or pediatric tumors, for which the advantages of proton therapy over conventional radiotherapy clearly have been demonstrated [7]. The second group of patients eligible for proton therapy are people for whom the more accurate dose delivered by protons or the possibility to deposit higher doses to the tumor are necessary for successful treatment. Examples of cancer in this group are head-and-neck tumors, lung cancer, prostate cancer and breast cancer. The advantages of proton therapy need to be proven in individual studies comparing the proton dose distribution to the dose delivered by conventional radiotherapy.

1.3 Uncertainties in proton therapy treatment

Proton beams are much more sensitive to uncertainties than the photon beams used in conventional radiotherapy, which can possibly compromise the superior dose distribution described in Section 1.2. Small changes in the energy of the beam for example can have a severe impact on the dose distribution within the patient, because the Bragg peak will be located at a different position. If this uncertainty is not incorporated in the treatment plan, important organs close to the tumor might receive a dose above their tolerance limit. Similar uncertainties in conventional radiotherapy have a much smaller effect. In Figure 1.4 the effects of range uncertainty for both a proton and a photon beam are depicted. For photon beams the dose distribution within the patient will stay more or less the same in the presence of uncertainties. In order to ensure adequate tumor coverage, physicians add an extra margin around the tumor in the treatment plan. For proton therapy this margin is not the answer, because the misalignment of individual dose fields can also lead to degradation of the dose conformity inside the tumor [8].

Generally two causes of uncertainties are distinguished. The first type are 'calculation uncertainties'. An analytical dose model is used to approximate the dose, which by definition is a simplified model that cannot take all relevant physical processes into account. One of the inputs for the analytical dose model is the density of the patient along the axis of the pencil beams. The density is calculated from the planning CT scan and is measured in Hounsfield units. Metal implants or implanted markers cause scatter artifacts on the CT scan, resulting in wrongly calculated densities. Furthermore, in the model Hounsfield units have to be converted to proton stopping power and as there is not a one to one correspondence between

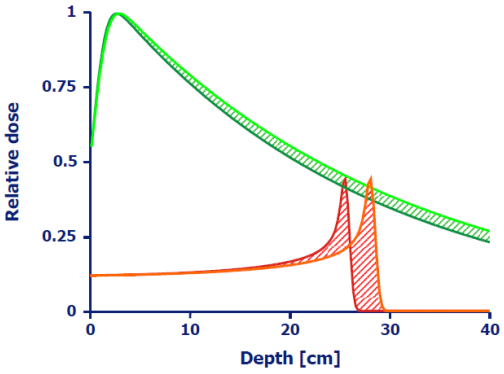


Figure 1.4: Difference between the impact of range uncertainty of a proton and a photon beam. The unperturbed photon beam is depicted in dark green and the proton beam in red. The uncertain beams are represented in light green and orange for photons and protons respectively.

the Hounsfield units and tissue materials, uncertainties up to $\pm 1.5\% - 2.0\%$ in the range of the proton beam are introduced [9].

The second type of uncertainties are 'motion-based uncertainties', which arise from movements of the patient or tumor during the course of treatment. The treatment spans several weeks as the total dose is delivered in fractions. As a result, the patient and tumor alignment with respect to the gantry cannot be perfectly matched every day. Furthermore, the treatment plan is based on an initial planning CT scan, not taking into account anatomical changes during treatment, such as tumor shrinkage or severe weight loss. Finally, processes such as swallowing, respiratory motions or organ motion occurring during irradiation can also have an impact on the accuracy of the dose delivery.

The two types of uncertainties are commonly converted into two types of error. The calculation uncertainties are modeled as an uncertainty in the range of the proton beam. These uncertainties can be modeled either as a relative error, thus a percentage deviation from the intended beam energy, or an absolute error, in which case the entire Bragg peak is shifted in the direction parallel to the beam. In this thesis only relative range uncertainties are considered. Misalignment of the patient or the tumor relative to the gantry are classified as setup uncertainties. Setup uncertainties consist of three components, the x-, y- and z-component, where each component can be assigned their own uncertainty distribution. Anatomical changes are not taken into account in this research.

1.4 Objective and outline of this thesis

The effects of uncertainties can be investigated by means of sensitivity and uncertainty analysis. Traditionally, deterministic methods such as perturbation theory or statistical methods based on Monte Carlo simulations are used for this type of analysis. In the case of proton therapy only statistical methods are available, but producing reliable results in this manner is computationally too expensive to be used in a clinical environment. In this research project, an alternative method is proposed. Polynomial chaos methods are part of a wider family of spectral methods and are assumed capable of producing results as accurate as statistical methods, but with a strong decrease in computational time. The objective of this thesis project can in short be summarized as investigating the applicability of polynomial chaos methods for sensitivity and uncertainty analysis of proton therapy treatment plans in terms of accuracy and computation time.

In order to investigate how well polynomial chaos can be applied to proton therapy treatment plans, the method is used to approximate the expected dose in individual voxels (volume elements), the full dose distribution within the patient, dose-volume histograms for the tumor and relevant organs and various dose and tumor parameters. Furthermore, the influence of robust treatment planning and the consistency between results obtained via two routes of the polynomial chaos method are investigated. Results are

validated with Monte Carlo simulations. Finally, the possibilities of variance decomposition are studied.

The structure of the thesis is as follows. An extensive theoretical framework for sensitivity and uncertainty analysis is given in Chapter 2. The first part of this chapter explains the theory behind sampling based methods, which serve as a reference case in this research. The second part of this chapter focuses on the theory, analysis and implementation of polynomial chaos methods. Chapter 3 presents the tools needed for the application of the polynomial chaos method in proton therapy treatment plans. This chapter addresses the medical point of view, in terms of relevant anatomy and tools needed for treatment plan evaluation, as well as the computational point of view, describing the codes for treatment plan generation in more detail. This chapter concludes with the methods used to validate the obtained polynomial chaos results. Results are presented in Chapter 4. The first sections provide the results of the proof of concept study for the full dose distribution, the dose-volume histograms and a selected list of parameters. Next a qualitative comparison with the surface approximation is made. This is followed by a discussion on the consistency of the method by means of a comparison between two routes via which results can be obtained and an investigation on the influence of robust treatment planning. This chapter concludes with an analysis of the variance decomposition of various relevant parameters. This thesis ends with Chapter 5 with conclusions that can be drawn from the conducted research based on the objective of this research project. Only a relevant selection of figures is given in this chapter, for a broader selection of figures the reader is referred to Appendix A. Finally, recommendations are made for future research regarding the polynomial chaos method in proton therapy treatment planning in order for it to be used in a clinical environment.

This report is written in partial fulfillment of the degree 'Master of Science'. The research project is conducted at the section 'Nuclear Energy and Radiation Applications', faculty of Applied Sciences of Delft University of Technology in cooperation with the medical physics unit, Department of Radiation Oncology of Erasmus Medical Center Cancer Institute in Rotterdam. The research is part of a broader research program for the Holland Particle Therapy Center (Holland PTC), a joint venture between Leiden University Medical Center, Erasmus University Medical Center and Delft University of Technology for the establishment of one of the first proton therapy treatment facilities in the Netherlands.

Chapter 2

Sensitivity and uncertainty analysis

The accuracy of dose delivery in proton therapy is one of the main concerns for this type of irradiation. Setup and range errors can cause significant underdosage to the tumor or overdosage to organs at risk. In order to quantify these errors sensitivity and uncertainty analysis (S&U analysis) is performed. Uncertainty analysis investigates how uncertainty in the input parameters is propagated through the model and becomes apparent in the model output. Sensitivity analysis shows how large the impact of an uncertain input is on the response of the model under consideration, either taken alone or in combination with others.

This chapter provides a theoretical description of two methods for performing S&U analysis. In Sections 2.1 and 2.2 the different sampling based methods used in this research are introduced along with the tools necessary to make a full analysis of all acquired data. Spectral projection methods and in particular polynomial chaos expansion are explained in Section 2.3. The analysis of polynomial chaos expansions is discussed in Section 2.4. This chapter concludes with an explanation of the implementation of the polynomial chaos method in Section 2.5.

2.1 Sampling based methods

Sampling based methods are part of the family of statistical methods [10]. Statistical methods can often be implemented in a straight-forward manner. However, for the method to be sufficiently accurate, a large number of model evaluations is needed. For complex models statistical methods then become too expensive and therefore impracticable. Despite this disadvantage, sampling based methods are commonly seen as the most reliable provided that enough samples are available. In that case, sampling based methods are treated as the reference case against which other methods can be compared.

Sampling based methods generally have the following layout:

1. Define a (subjective) distribution describing the uncertainty distribution of the input parameters.
2. Generate N sample points from the defined uncertainty distribution (based on one of the methods described in Section 2.1.2).
3. Run the full model N times with the generated sample points as inputs.
4. Analyze the model outputs with the tools provided in Section 2.2.
5. Perform sensitivity analysis by assessing the effect individual inputs have on the results.

2.1.1 Defining the uncertainty distribution

In practice input parameters can take on a range of possible values, each with their own probability. This probability distribution is not known unambiguously, but is often provided by a combination of research and expert opinion. This means that the uncertainty distribution assigned to the input parameter is a subjective one, which in turn will also lead to a corresponding subjective uncertainty in the model output or model response of interest.

For the quantification of uncertainty in proton therapy some experimental data is available [11]. All input parameters are assumed to follow a Gaussian distribution, formulated in Equation 2.1.

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.1}$$

Here μ represents the average or expected value of the parameter and σ^2 the variance, which is a measure of the spread around the expected value. The square root of the variance is called the standard deviation σ. The values of μ and σ are chosen in such a way to best describe experimental data. A graphic representation of the Gaussian distribution is shown in Figure 2.1.

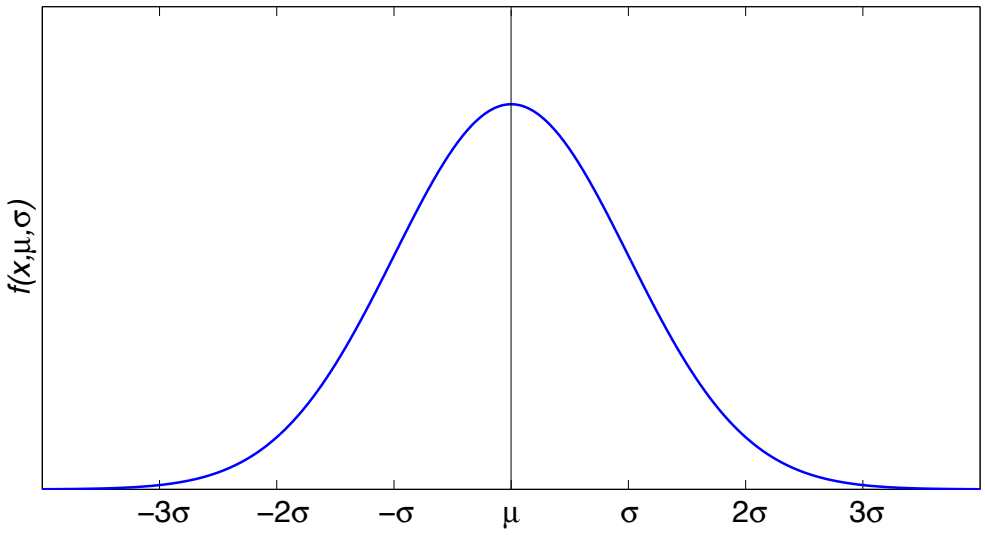


Figure 2.1: Shape of the Gaussian distribution, also referred to as the normal distribution. Unknown distributions are often assumed to follow this curve.

2.1.2 Generating samples

Random sampling

The most widely used form of sampling is random sampling, because of its easy applicability to any system and because it provides unbiased estimates for the mean and variance of the response. The greatest drawback of this method is that it can take a large amount of samples before a response converges. Especially for complex systems, with many input parameters, this can become computationally very expensive, but it is often seen as the only applicable method.

In this method N samples are selected randomly from within the sample space, with the probability density function dictating the probability of a sample point being chosen. Each sample point is selected independently of all other points. There is no guarantee that every part of the sample space is covered. However, the higher the amount of sample points, the higher the probability that any given subregion

of the sample space is evaluated. The fact that the Gaussian cumulative distribution function is closely related tot the error function is used for the generation of random samples in this research. The relation is formulated in Equation 2.2.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \mathrm{d}t = \frac{1}{2} + \frac{1}{2} \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right) \tag{2.2}$$

The generation of one random sample is a two-step process. First, a pseudorandom number x in the interval $[0, 1]$ is drawn for each of the separate input variables and transformed to a Gaussian distributed variable r with zero mean and unit variance via Equation 2.3, which is a reformulation of Equation 2.2.

$$r = \Phi^{-1}(x) = \sqrt{2}\mathrm{erf}^{-1}(2x - 1) \tag{2.3}$$

Second, this standard normal distributed variable is modified to a normal distributed variable X with a mean $\mu \in \mathbb{R}$ and a standard deviation $\sigma \in (0, \infty)$ by the relation defined in Equation 2.4. These steps are taken for each of the input parameters. The normal distributed input variables together form one random sample. The whole process is repeated N times to obtain the desired number of random samples.

$$X = \mu + \sigma r \tag{2.4}$$

Latin hypercube sampling

One of the drawbacks of random sampling is that there is no guarantee that the whole sample space is covered. Latin hypercube sampling is designed specifically to ensure that the full range of all variables is covered. The samples are still random, but relatively uniformly distributed over the sample space. This sampling method becomes the method of choice when the number of samples required for random sampling is impractical, since it requires only a fraction of that. As an additional advantage, this method utilizes the same unbiased estimators for the mean and variance as random sampling.

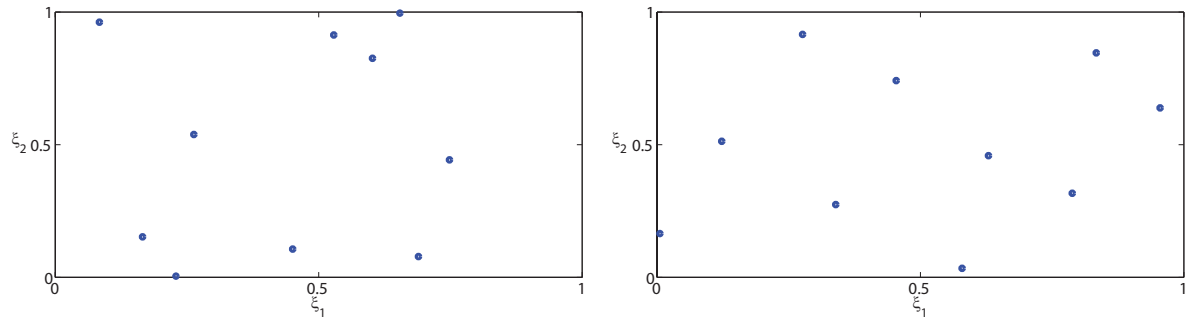
The sample space in each independent direction is divided into N intervals of equal probability. Within each interval a value is randomly selected. The random values for the first input parameter are randomly paired without replacement with the random values for the second input parameter. These pairs of values are then combined without replacement with the random values from the third parameter. This process continues until all input variables are included and N Latin hypercube samples are created. In this research the built-in Latin hypercube sampling algorithm in Matlab was used to generate Latin hypercube samples in the interval $[0, 1]$. The same transformation as for random sampling was applied to generate normally distributed samples, as demonstrated in Equation 2.3 and 2.4.

In Figure 2.2 a comparison between random sampling and Latin Hypercube sampling is shown for a low number of sample points. For both methods ten points were sampled from a uniform distribution in the interval $[0, 1]$. It can be clearly seen that in the case of random sampling large areas of the sample space are not covered, in the case of Latin Hypercube sampling a more uniform distribution of samples in space is visible.

Surface approximation

For medical purposes, in particular in radiotherapy, a sampling based method named the surface approximation is sometimes used in addition to the more established sampling based statistical methods described in the previous sections [12]. This method is also used for the optimization of robust treatment plans. The advantages of this method are that it only needs a very limited number of model evaluations $(1 + 2 \cdot \text{dimension})$ and that it is not necessary to have detailed knowledge of the probability density distribution of the modeled uncertainties.

The sample space consists of the nominal scenario and a limited number of error scenarios. In the case of three independent spatial variables, the sample space is a three-dimensional error-sphere. The radius equals an 85% confidence interval, or roughly $1.5 \cdot \sigma$. The model is evaluated at the origin and at the



(a) 10 random sample points

(b) 10 Latin hypercube sample points

Figure 2.2: Figure illustrating the difference between random sampling and Latin hypercube sampling. The latter is designed to ensure the sample space is covered more equally.

surface points along each of the three axes, which is depicted in Figure 2.3. In this approach it is assumed that any errors resulting from these surface scenarios will always be equal to or of larger magnitude than errors resulting from smaller input uncertainties. It is unsuitable to use this method quantitatively, as there is no mathematical proof that the expectation value or variance can be correctly calculated. As a consequence this method will solely be used for qualitative comparisons.

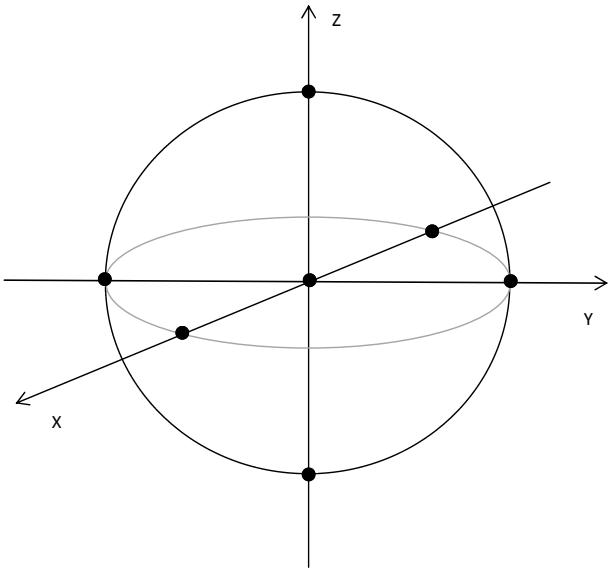


Figure 2.3: The surface approximation samples the origin (the nominal scenario) and the points marking the 85% confidence interval for all different input parameters. For a three-dimensional system, the sample space is a three-dimensional sphere.

2.2 Analysis of sampling based methods

The model is run using the generated samples as input parameters. The next step is the uncertainty analysis, in which the distributions of the model responses are analyzed in order to extract the effect of the uncertainty in the input parameters. For this, the concepts of central moments is used as defined in Equation 2.5, which can provide a systematic framework to distinguish between different probability distributions.

$$\mu_n = \mathrm{E}\left[(X - E\left[X\right])^n\right] = \int_{-\infty}^{\infty} (x-\mu)^n f(x)\mathrm{d}x \tag{2.5}$$

The second central moment is usually denoted by σ^2 and is more widely known as the aforementioned variance. The central moments, with n representing the order of the moment, combined with the expectation value of the distribution, gives a perfect description of the probability distribution in the case $n \to \infty$. For complex systems, it is often considered sufficient to accurately describe the mean and the variance of the underlying probability distribution. Random sampling and Latin hypercube sampling provide the same unbiased estimators for the mean and variance of a population. Equation 2.6 formulates the estimator for the mean of a sample population. Here N represents the number of samples and x_i the response of the i-th sample. For enough samples this estimator is expected to converge towards the true mean of the population (Equation 2.7).

$$\widehat{\mu_x} = \frac{\sum_i^N x_i}{N} \tag{2.6}$$

$$\mathrm{E}\left[\widehat{\mu_x}\right] = \mu_x \tag{2.7}$$

Because $\widehat{\mu_x}$ is a function of a random variable, the estimator itself is again a random variable with a probability distribution. In order to investigate the behavior of the estimator, it is possible to define a standard deviation for the mean σ_μ, defined in Equation 2.8, which gives information on the precision of the mean.

$$\sigma_\mu = \frac{\sigma_x}{\sqrt{N}} \tag{2.8}$$

This is not to be confused with the standard deviation of the population. With this tool it is possible to monitor the convergence of the mean for increasing population size N. This is graphically depicted in Figure 2.4a. As can be seen, a large number of samples is needed to obtain a sufficiently high accuracy of the estimator of the mean.

The variance of a random or Latin hypercube sampled population can be estimated without bias with Equation 2.9. For enough samples provided, this estimator is expected to converge towards the true variance (Equation 2.10).

$$\widehat{s^2} = \frac{\sum_i^N \left(x_i - \widehat{\mu_x}\right)^2}{N-1} \tag{2.9}$$

$$\mathrm{E}\left[s^2\right] = \sigma_x^2 \tag{2.10}$$

In the same way as for the mean, a standard deviation for the variance σ_{s^2} is defined in Equation 2.11. This quantity makes use of the fourth central moment μ_4 and the estimator for the variance $\widehat{s^2}$.

$$\sigma_{s^2} = \sqrt{\mathrm{Var}\left[s^2\right]} = \widehat{s^4}\left(\frac{2}{N-1} + \frac{\frac{\mu_4}{s^4} - 3}{N}\right) \tag{2.11}$$

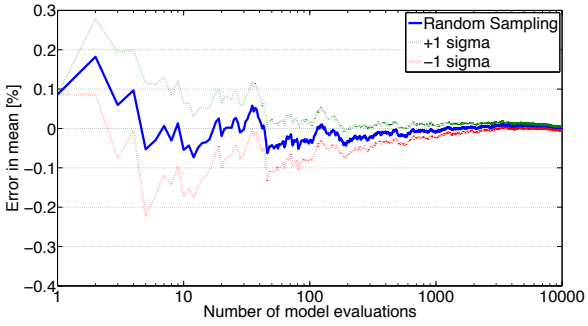
Generally, the mean is expected to converge much faster than the variance. In Figure 2.4 an example of the convergence behavior of the mean and variance is shown, where the dotted lines represent the $\pm 1\sigma$ of the mean and variance respectively. Values are shown relative to the mean and variance obtained after 10.000 model evaluations.

2.3 Spectral methods

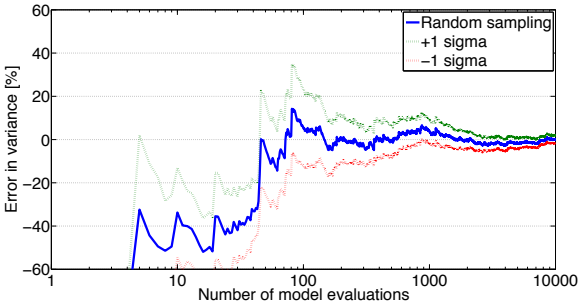
Spectral methods are based on the idea that a quantity of interest of a system can be constructed as a function of its input variables. The quantity of interest, or systems response, is reconstructed as a series expansion of basis vectors Ψ_k, each with a corresponding deterministic expansion coefficient r_k, formulated in Equation 2.12.

$$R = \sum_{k=0}^{\infty} r_k \Psi_k \tag{2.12}$$

The basis vectors Ψ_k are functionals that transform the given input vector into a scalar quantity. The systems response can thus be easily calculated for any given input vector, as the original (complex) model is essentially replaced by a series of functions and expansion coefficients. This property makes the method very suitable for S&U analysis.



(a) Convergence behavior of the mean



(b) Convergence behavior of the variance

Figure 2.4: Example illustrating the convergence behavior of the mean and the variance of the mean dose of the tumor. Lines are shown relative to the random sampling mean and variance after 10,000 random samples. Note the difference in scaling on the vertical axis between Figure 2.4a and Figure 2.4b.

Any process with a finite second order moment can be expressed in this manner, such that the expansion will eventually converge towards the process under consideration. Polynomial Chaos Expansion (PCE) is part of the family of spectral techniques and uses polynomials as basis functions. The idea behind polynomial chaos originates from Wiener [13] to describe Gaussian processes, but was later extended by Xiu and Karniadakis [14] to include processes with different underlying uncertainty distributions. In this section first the necessary mathematical definitions are explained. The next step is the choice of the basis vectors and the calculation of the expansion coefficients. Finally the numerical method needed for the calculation is discussed. The theory in this section is largely based on [15], [16] and [17].

The outcomes of all possible random events are contained in the sample space Θ. The collection Σ contains subsets of the sample space and is called the event space. To each event θ in the event space a probability is assigned by the probability measure $\mathbb{P}$. The entire probability space sums up to unity. This triple together makes up the probability space $(\Theta, \Sigma, \mathbb{P})$. Quantities of interest are real-valued scalar responses $R(\theta)$, mapping the sample space Θ to $\mathbb{R}$.

$$R(\theta) : \Theta \to \mathbb{R} \tag{2.13}$$

These real-valued responses of the system under consideration live in L_2-space, a vector space containing all square-integrable functions. The inner product of a response in L_2-space is finite and defined in Equation 2.14.

$$\langle R, R \rangle = \int_{\Theta} R(\theta) R(\theta) \, \mathrm{d}\mathbb{P}(\theta) < \infty \tag{2.14}$$

Uncertain input data is represented by a random event θ. Each random event $\theta \in \Theta$ can be written as a vector of N independent variables as formulated in Equation 2.15. Generally, each independent source of uncertainty is represented by a random variable, where the number of distinct sources N defines the dimension of the model. In this research the dimension is 3 when only setup uncertainties are investigated and 4 when range uncertainties are also included.

$$\underline{\xi}(\theta) = (\xi_1(\theta), \xi_2(\theta), \ldots, \xi_N(\theta))^T \tag{2.15}$$

2.3.1 Basis vectors

The basis vectors in polynomial chaos are sequences of polynomials orthogonal to one another with respect to the inner product in L_2. This means that the inner product of two different polynomials is always zero. This is formulated in Equation 2.16, in which α is the norm of the n-th polynomial and $\delta_{n,m}$ is the Kronecker delta. The weighting function $w(x)$ generally denotes the joint probability density functions of the orthogonal polynomials.

$$\langle P_n(x), P_m(x) \rangle = \int P_n(x) P_m(x) w(x) \, \mathrm{d}x = \alpha^2 \delta_{n,m} \tag{2.16}$$

The choice of the most appropriate basis vectors is dictated by the Wiener-Askey scheme, which is based on the fact that weighting functions for some type of polynomials are identical to the probability density function, denoted by $p_{\xi_j}(\xi_j)$, of certain types of random distributions. Thus, when an input has a certain probability density function and the corresponding family of polynomials is chosen, it can be shown that the model will converge the fastest. When a different expansion is used, the solution converges but at a clearly slower rate. [14]

Table 2.1: Probability distributions are coupled with corresponding orthogonal polynomials in the Wiener-Askey scheme.

	Random Distribution	Askey Polynomial	Support
Continuous	Gaussian	Normalised Hermite	[-∞,∞]
	Uniform	Legendre	[a,b]
	Gamma	Laguerre	[0,∞]
	Beta	Jacobi	[a,b]
Discrete	Poisson	Charlier	{0,1,2,...}
	Negative Binomial	Miexner	{0,1,2,...}
	Binomial	Krawtchouk	{0,1,...,N}
	Hypogeometric	Hahn	{0,1,...,N}

For this research, all input parameters were assumed to be Gaussian distributed, for which the output can best be modeled using Hermite polynomials. This type of polynomial is defined in Equation 2.17. For this type of polynomials it is possible to calculate the norm analytically, as demonstrated in Equation 2.18, which is later used in the calculation of the expansion coefficients (Section 2.3.3).

$$\mathrm{He}_n(x) = \frac{1}{(-1)^n \exp\left(x^2/2\right)} \frac{d^n}{dx^n} \left[\exp\left(-\frac{x^2}{2}\right)\right] \tag{2.17}$$

$$h_n^2 = \langle \mathrm{He}_n, \mathrm{He}_n \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[\mathrm{He}_n(x)\right]^2 \exp(-x^2/2) dx = n! \tag{2.18}$$

These polynomials are suited for a standard normal distribution Z with zero mean and unit variance. In order to modify this standard normal distribution to a normal distribution X with a mean $\mu \in \mathbb{R}$ and a standard deviation $\sigma \in (0, \infty)$, the transformation given in Equation 2.4 is applied. In Table 2.2 examples of one-dimensional low order probabilists' Hermite polynomials are given.

Table 2.2: Example of low order Hermite polynomials

Order	Notation	Polynomial
0	$\mathrm{He}_0(x)$	1
1	$\mathrm{He}_1(x)$	$2x$
2	$\mathrm{He}_2(x)$	$x^2 - 1$
3	$\mathrm{He}_3(x)$	$x^3 - 3x$
4	$\mathrm{He}_4(x)$	$x^4 - 6x^2 + 3$

2.3.2 Basis sets

The choice for the type basis vectors is now made. From the product of N one-dimensional Hermite polynomials, each depending on a different random input variable ξ_j, a basis vector is constructed as formulated in Equation 2.19. Here γ is used to distinguish between polynomials of different polynomial orders.

$$\Psi\left(\underline{\xi}\right) = \prod_{j=1}^{N} \mathrm{He}_\gamma\left(\xi_j\right) \tag{2.19}$$

The order of the multivariate polynomial $\lambda(o)$ is defined as the sum of the orders of the univariate polynomials in the basis vector. The basis set is then defined by all the possible basis vectors up to the chosen maximum allowed order O. To illustrate this more clearly, an example for the full PC basis set is given in Table 2.3 for two independent input variables (ξ_1, ξ_2) constructed from one-dimensional polynomials up to order $O = 2$.

Table 2.3: Example of the PC basis set up to second order

$\lambda(0)$	$\gamma = (0,0)$	$\Psi_0 = \mathrm{He}_0(\xi_1)\,\mathrm{He}_0(\xi_2)$
$\lambda(1)$	$\gamma = (1,0)$	$\Psi_1 = \mathrm{He}_1(\xi_1)\,\mathrm{He}_0(\xi_2)$
	$\gamma = (0,1)$	$\Psi_2 = \mathrm{He}_0(\xi_1)\,\mathrm{He}_1(\xi_2)$
$\lambda(2)$	$\gamma = (2,0)$	$\Psi_0 = \mathrm{He}_2(\xi_1)\,\mathrm{He}_0(\xi_2)$
	$\gamma = (1,1)$	$\Psi_0 = \mathrm{He}_1(\xi_1)\,\mathrm{He}_1(\xi_2)$
	$\gamma = (0,2)$	$\Psi_0 = \mathrm{He}_0(\xi_1)\,\mathrm{He}_2(\xi_2)$

The complete full PC basis set is formally defined by the multi-index set as defined in 2.20.

$$\mathcal{L}(O) = \bigcup_{o\in[0,1,\ldots,O]} \lambda(o) = \left\{ \underline{\gamma} : \sum_{j=1}^{N} \gamma \leq O \right\} \tag{2.20}$$

The basis vectors contained in the set are given by Equation 2.21.

$$\Gamma(O) = \bigcup_{o\in[0,1,\ldots,O]} \Gamma_o = \left\{ \cup_{\underline{\gamma}\in\mathcal{L}(O)} \prod_{j=1}^{N} \mathrm{He}_\gamma(\xi_j) \right\} \tag{2.21}$$

In order to use this technique for computational purposes, the infinite sum in Equation 2.12 must be truncated to a finite value P, formulated in Equation 2.22. The number of basis vectors is defined by the dimension of the problem and the chosen maximum polynomial order. The corresponding truncated expansion is expressed in Equation 2.23.

$$P + 1 = \frac{(N+O)!}{N!O!} \tag{2.22}$$

$$R(\underline{\xi}) \approx \sum_{k=0}^{P} r_k \Psi_k(\underline{\xi}) \tag{2.23}$$

2.3.3 Calculation of the expansion coefficients

With the construction of the basis sets completed, the next step is the calculation of the r_k expansion coefficients. Because the polynomial basis functions are orthogonal, it is possible to express the coefficient in the expansion by projecting the stochastic quantity of interest onto the respective polynomial

$$r_k = \frac{\langle R(\underline{\xi}), \Psi_k(\underline{\xi})\rangle}{\langle \Psi_k(\underline{\xi}), \Psi_k(\underline{\xi})\rangle} \tag{2.24}$$

Using the definition of the inner product in Equation 2.18, the expansion coefficient can be further expressed in Equation 2.25.

$$r_k = \frac{\int_\Theta R(\underline{\xi})\,\Psi_k(\underline{\xi})\,p_{\underline{\xi}}(\underline{\xi})\,\mathrm{d}\underline{\xi}}{\langle \Psi_k(\underline{\xi}), \Psi_k(\underline{\xi})\rangle} \tag{2.25}$$

Here $p_{\underline{\xi}}(\underline{\xi})$ represents the joint probability density function of the input variables, which in the case of independent variables is simply the product of the probability density functions of the individual variables. The expression above can be further evaluated when the integral is expressed as individual

integrals for each of the individual random variables over their respective domains. The inner product in the denominator is expressed in terms of Equation 2.16 and the k-th basis vector $\Psi_k\left(\underline{\xi}\right)$ is defined as $\underline{\gamma}_k = \left(\gamma_{k,1}, \ldots, \gamma_{k,N}\right)^T$. Equation 2.25 can now be further reformulated in Equation 2.25.

$$r_k = \frac{1}{h_k^2} \int \int \ldots \int R\left(\underline{\xi}\right) \prod_{j=1}^{N} He_{\gamma_k}\left(\xi_j\right) p_{\xi_j}\left(\xi_j\right) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \ldots \mathrm{d}\xi_N \tag{2.26}$$

R is the unknown response of interest that needs to be approximated, which means this projection integral needs to be solved. It is not possible to solve this integral analytically, instead an integration method is used. This means that the systems response R is evaluated at specific predefined points in the Θ space in order to acquire approximate knowledge on the behavior of the integrand. The original model is thus only used as a "black box".

For the evaluation of the multi-dimensional integral in Equation (2.26) a quadrature rule is used to evaluate the one-dimensional integrals. Quadrature rules are based on the idea that the integral of a general function f can be approximated as a weighted sum of the evaluation of that function at predefined points, as expressed in Equation 2.27.

$$I^{(1)} f = \int_a^b f\left(\xi_j\right) p_{\xi_j}\left(\xi_j\right) \mathrm{d}\xi_j \approx Q_{lev}^{(1)} f = \sum_{i=1}^{n_{lev}} f\left(\xi_{j,lev}^{(i)}\right) w_{lev}^{(i)} \tag{2.27}$$

The weights $w_{lev}^{(i)}$ and points $\xi_{j,lev}^{(i)}$ are determined by the chosen quadrature rule. In this research Gaussian quadrature rules are used, which can yield exact results for polynomials up to order $o = 2n_{lev} - 1$. The number n_{lev} is the number of points needed for the approximation and determines the accuracy of the quadrature formula which corresponds to a level index lev. The level index is also denoted by the grid order, which can be varied in such a way that a good balance between accuracy and computation time is obtained. In Figure 2.5 the relation between the grid order and the number and positions of the quadrature points is shown.

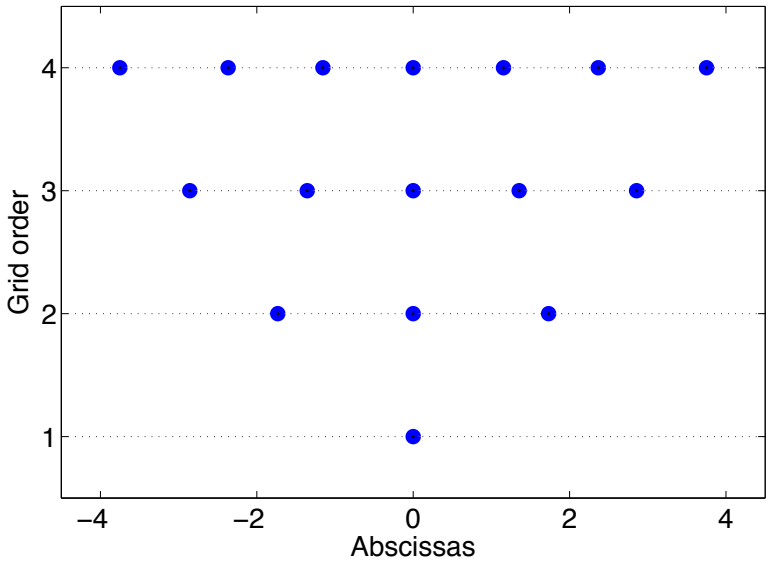


Figure 2.5: Quadrature points for consecutive levels for a one dimensional integral. Higher levels correspond to higher accuracy. Only the first level point is nested, which means that this point is contained in all higher level grids.

The one-dimensional quadrature formulae can be used to construct multidimensional cubature formulae via tensorization. Tensorization is basically the multiplication of the N one-dimensional quadrature

formulae (each depending only on a single variable), resulting in the cubature formula in Equation 2.28.

$$I^{(N)}f = \int_{a_1}^{b_1}\int_{a_2}^{b_2}\dots\int_{a_N}^{b_N} f\left(\underline{\xi}\right) p_{\underline{\xi}}\left(\underline{\xi}\right) \mathrm{d}\xi_1\mathrm{d}\xi_2\dots\mathrm{d}\xi_N \approx Q_{\underline{lev}}^{(N)}f = \left(Q_{lev_1}^{(i_1)} \otimes Q_{lev_2}^{(i_2)} \otimes \dots \otimes Q_{lev_N}^{(i_N)}\right) f = \sum_{i_1=1}^{n_{lev_1}}\sum_{i_2=1}^{n_{lev_2}}\dots\sum_{i_N=1}^{n_{lev_N}} f\left(\xi_{1,lev_1}^{(i_1)}, \xi_{2,lev_2}^{(i_2)}, \dots, \xi_{N,lev_N}^{(i_N)}\right) w_{lev_1}^{(i_1)} w_{lev_2}^{(i_2)} \dots w_{lev_N}^{(i_N)} \quad (2.28)$$

In this equation $\underline{lev} = (lev_1, lev_2, \dots, lev_N)^T$ is a vector containing the different quadrature levels in the different directions. This results in a summation over all possible combination of quadrature points. The number of grid points required to build the full quadrature grid is equal to n_{lev}^N if the grid is built to the same order in each direction, i.e. $n_{lev_1} = n_{lev_2} = \dots$. It is apparent that for higher dimension problems the number of needed system evaluations grows exponentially, which is often referred to as "the curse of dimensionality". Even for lower dimensional problems, but with a high required accuracy, the computational cost can increase rapidly. The cubature formulae are therefore substituted by sparse grids.

2.3.4 Sparse grids

Originally developed by the Russian mathematician Sergey Smolyak, sparse grid construction is a numerical technique to reduce the number of calculations needed for the integration or interpolation of high dimensional functions. Provided that certain smoothness requirements are fulfilled, this sparse grid will yield results nearly as good as conventional full grid methods, but with only a fraction of the needed system evaluations. Sparse grids are based on the difference between two consecutive quadrature levels as defined in Equation 2.29.

$$\Delta_{lev}^{(1)}f = Q_{lev}^{(1)}f - Q_{lev-1}^{(1)}f \quad (2.29)$$

Using Equation 2.29 and the definition $Q_0^{(1)}f = 0$, the quadrature rule can be written as a sum of difference formulae, formulated in Equation 2.30.

$$\begin{aligned}\sum_{l=1}^{lev}\Delta_l^{(1)}f &= \Delta_1^{(1)}f + \Delta_2^{(1)}f + \dots + \Delta_{lev}^{(1)}f \\ &= \left(Q_1^{(1)}f - Q_0^{(1)}f\right) + \left(Q_2^{(1)}f - Q_1^{(1)}f\right) + \dots + \left(Q_{lev}^{(1)}f - Q_{lev-1}^{(1)}f\right) \\ &= Q_{lev}^{(1)}f \end{aligned} \quad (2.30)$$

The cubature formula (with full tensorization) can then be written in the form of Equation 2.31.

$$Q_{\underline{lev}}^{(N)} = \sum_{l_1=1}^{lev_1}\sum_{l_2=1}^{lev_2}\dots\sum_{l_N=1}^{lev_N}\left(\Delta_{l_1}^{(1)} \otimes \Delta_{l_2}^{(1)} \otimes \dots \Delta_{l_N}^{(1)}\right) f \quad (2.31)$$

A more general formula for the cubature formula can be given when a multi-index $\underline{l} = (l_1, l_2, \dots, l_N)^T$ is used to distinguish between the different grids. If it is assumed that the different directions all have the same quadrature level ($lev_j = lev$), a full grid corresponding to n_{lev}^N grid points, formulated in Equation 2.32.

$$\mathcal{I}^{Full}(lev) = \{\underline{l} : l_j \leq lev \quad \forall\ j \in [1, \dots, N]\} \quad (2.32)$$

The idea behind sparse grid constructions is that not all directions contribute equally to the integral, and thus many terms in the sum can be discarded. The original Smolyak sparse grid construction is defined in Equation 2.33.

$$\mathcal{I}^{Smolyak}(lev) = \left\{\underline{l} : \sum_{j=1}^{N} l_j \leq lev + N - 1\right\} \quad (2.33)$$

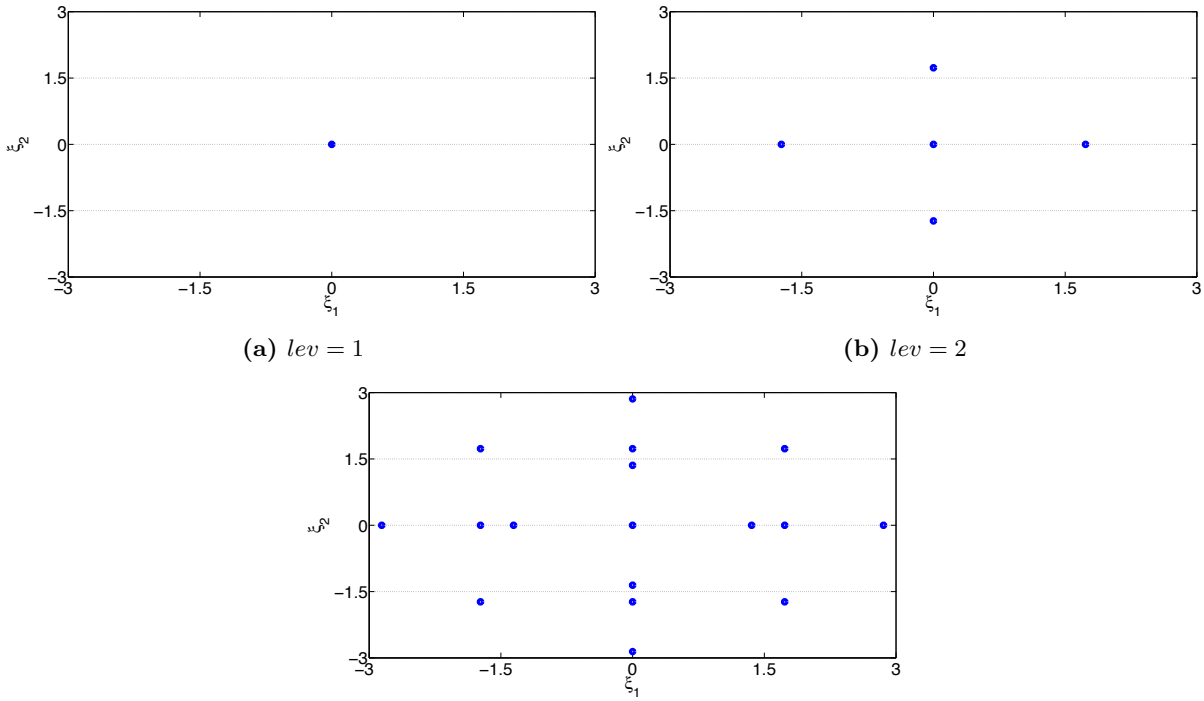
Each direction is thus at least integrated with one point ($lev = 1$) and an additional ($lev - 1$) extra integration levels can be distributed over the different dimensions. The integration level in any direction is at maximum lev. As a clarifying example, the subgrids and number of grid points for a three dimensional problem are given in Table 2.4. The subgrids described here are examples of 'prototype' subgrids. For

Table 2.4: Number of grid points in sparse grids for a three-dimensional system for different grid orders.

Grid order	Subgrids	Number of grid points
1	$(1,1,1)$	1
2	$(2,1,1)$	$(1+6)=7$
3	$(3,1,1),(2,2,1)$	$(7+12+12)=31$
4	$(4,1,1),(3,2,1),(2,2,2)$	$(31+18+48+8)=105$

example, the prototype subgrid $(2,1,1)$ has three possibilities, i.e. $(2,1,1)$, $(1,2,1)$ and $(1,1,2)$, which all consist of two extra points with respect to the grid with grid order 1.

The construction of the sparse grid can also be depicted graphically. In Figure 2.6 an example for a two dimensional grid is given for increasing grid orders.



(a) $lev = 1$ **(b)** $lev = 2$

(c) $lev = 3$

Figure 2.6: Sparse grid construction for increasing grid orders for a two-dimensional system.

2.4 Post-processing the polynomial chaos expansion

2.4.1 Analysis of mean and variance of polynomial chaos expansions

The mean and variance of the response of interest R are again assumed to sufficiently describe the underlying uncertainty distribution. Keeping in mind that the zeroth order polynomial $\Psi_0 = 1$, the mean, or expectation value, of R is given by the coefficient of the zeroth polynomial in the expansion, which gives the constant term. This is expressed in Equation 2.34. [15]

$$\mu_R = E\left[R\left(\theta\right)\right] = \left\langle R\left(\theta\right)\right\rangle = \left\langle \Psi_0 R\left(\theta\right)\right\rangle = \sum_{k=0}^{\infty} r_k \langle \Psi_0, \Psi_k\rangle = r_0 \tag{2.34}$$

The variance is given by a summation over all but the first coefficient in the expansion, which is expressed in Equation 2.35.

$$\sigma_R^2 = E\left[\left(R - E\left[R\right]\right)^2\right] = E\left[\left(\sum_{k=1}^{\infty} r_k \Psi_k\right)^2\right] = \sum_{k,l=1}^{\infty} r_k r_l \langle \Psi_k, \Psi_l \rangle = \sum_{k=1}^{\infty} r_k^2 h_k^2 \tag{2.35}$$

Other statistics of the distribution, such as higher moments or the probability density function, can be derived when the original model is replaced by the new metamodel made up by the basis functions and expansion coefficients. The expansion is resampled and statistics are derived as described in Section 2.2. This resampling is much less computationally intensive than the original system evaluations, as the new metamodel is much simpler than the original model.

2.4.2 Variance decomposition

In many applications it can be useful to determine which input contributes most to the uncertainty in the output. Sobol' indices are a powerful tool in sensitivity analysis for ranking the different input variables according to how much they contribute to the variance of the system responses of interest. For polynomial chaos, the Sobol' indices can be computed analytically from the coefficients in the expansion [18].

Again, assume the model $R = f\left(\underline{\xi}\right)$, where f represents the function that maps the sample space Θ to $\mathbb{R}$. $f\left(\underline{\xi}\right)$ can be decomposed into sums of increasing dimension according to the Sobol' decomposition formulated in Equation 2.36. The sums are referred to as summands.

$$f\left(\xi_1, \ldots, \xi_N\right) = f_0 + \sum_{i=1}^{N} f_i\left(\xi_i\right) + \sum_{1 \le i < j \le N} f_{ij}\left(\xi_i, \xi_j\right) + \cdots + f_{12\ldots N}\left(\xi_1, \ldots, \xi_N\right) \tag{2.36}$$

Here f_0 represents the mean value of the function in the same way as described before. The integral of each summand over any independent variable that is present in that summand equals zero. This has as consequence that the summands are orthogonal to each other, a property formulated in Equation 2.37.

$$\int f_{i_1 \ldots i_s}\left(\xi_{i_1}, \ldots \xi_{i_s}\right) \mathrm{d}\xi_{i_k} = 0 \qquad \text{for} \qquad 1 \le k \le s \tag{2.37}$$

$$\int f_{i_1 \ldots i_s}\left(\xi_{i_1}, \ldots \xi_{i_s}\right) f_{i_1 \ldots i_t}\left(\xi_{i_1}, \ldots \xi_{i_t}\right) \mathrm{d}\underline{\xi} = 0 \tag{2.38}$$

It can be shown that whenever the function $f\left(\underline{\xi}\right)$ is integrable over the domain, the decomposition in Equation (2.36) is unique and the terms in the decomposition can be derived analytically for polynomial chaos expansions [18]. For example, the one-index and two-index term can be expressed as

$$f_i\left(\xi_i\right) = \int f\left(\underline{\xi}\right) \mathrm{d}\underline{\xi}_{\sim i} - f_0$$

$$f_{ij}\left(\xi_i, \xi_j\right) = \int f\left(\underline{\xi}\right) \mathrm{d}\underline{\xi}_{\sim \{ij\}} - f_i\left(\xi_i\right) - f_j\left(\xi_j\right) - f_0$$

The notation $\sim \{ij\}$ here denotes the integration over all parameters except ξ_i and ξ_j. This means that any of the summands can be expressed in terms of a multidimensional integral and all lower order summands.

Recall that since the input parameters are random variables, the model response of interest R is also a random variable, whose (total) variance can be expressed as in Equation 2.39.

$$D = \mathrm{Var}\left[R\left(\underline{\xi}\right)\right] = \int f^2\left(\underline{\xi}\right) \mathrm{d}\underline{\xi} - f_0^2 \tag{2.39}$$

By integrating the square of Equation 2.36 and by making use of the orthogonality property of the summands, the variance can be decomposed in partial variances which together make up the decomposition

of the total variance, formulated in Equation 2.40.

$$D_{i_1 \dots i_s} = \int f^2_{i_1 \dots i_s} \left(\xi_{i_1}, \dots, \xi_{i_s}\right) \mathrm{d}\xi_{i_1} \dots \mathrm{d}\xi_{i_s} \tag{2.40}$$

$$D = \sum_{i=1}^{N} D_i + \sum_{1 \leq i < j \leq N} D_{ij} + \dots + D_{12 \dots N} \tag{2.41}$$

Finally, these partial variances can be used in the definition of the Sobol' indices (Equation 2.42), which are a sensitivity measure describing how much of the variance in the response can be contributed to any of the input variables.

$$S_{i_1 \dots i_s} = D_{i_1 \dots i_s} / D \tag{2.42}$$

The first order sensitivity indices measure the effect each individual parameter have on the output variance. Higher order sensitivity indices describe the interaction effects between combinations of the input parameters. For each input variable the total sensitivity index (TSI, defined in Equation 2.43) can be determined by including all the possible joint terms containing at least that particular input variable. For high-dimensional problems it is a reasonable assumption that only a few input parameters are of importance in the output of the model. This is sometimes known as the sparsity of effects principle.

$$\mathrm{TSI}(\xi_i) = S_i + \sum_{i \neq j} S_{ij} + \sum_{j,k \neq i, j<k} S_{ijk} + \dots + S_{ij,\dots,N-1} \tag{2.43}$$

The set of input variables can be ranked and classified on basis of their TSI [19]:

- *very important* if $0.8 <$ TSI
- *important* if $0.5 <$ TSI < 0.8
- *unimportant* if $0.3 <$ TSI < 0.5
- *irrelevant* if TSI < 0.3.

The computation of the sensitivity indices for polynomial chaos expansion is relatively straightforward. The polynomials in the expansion are grouped and summed according to the input parameters they depend on, which leads exactly to the Sobol' decomposition as defined earlier in Equation 2.36. In the same way as described in Equations 2.40 to 2.43, any of the Sobol' indices can be calculated. Once the expansion is computed, the sensitivity indices can thus be calculated with no additional cost.

The computation of sensitivity indices for Monte Carlo based sampling techniques is unfortunately not computationally cheap. Several algorithms are available for the calculation of (only) the first and total sensitivity indices. The computational costs depend both on the number of dimensions of the model and the original number of model evaluations. The cheapest method would require 30.000 model runs for a three dimensional problem and a 1% relative accuracy and would only give the first and total sensitivity indices [19]. Polynomial chaos therefore provides more complete information on all sensitivity indices without hardly any additional computational costs. Any sensitivity indices mentioned in the rest of this research are consequently solely calculated with polynomial chaos.

2.5 Implementation of Polynomial Chaos Expansion

For the application of polynomial chaos to the quantification of uncertainty in proton therapy treatment planning, the in house developed FANISP algorithm is used. FANISP stands for Fully Adaptive Non-Intrusive Spectral Projection and was developed by Perkó, originally intended for the quantification of uncertainty in generic problems [16]. In this section the algorithm itself is discussed in more detail.

Non-intrusive means that the original system does not have to be altered. The original code is simply used as a black box. A certain set of inputs, prescribed by the distributions, is given to the system. The corresponding system realizations are collected and subsequently used to approximate the behavior of

the integrals described in Section 2.3.3. In the end the complete original model can be replaced by a polynomial model. Computation time of the recalculations of the model can be decreased drastically, as the complete model is reduced to a sum of simple polynomial functions with corresponding expansion coefficients instead of, in this case, the original dose engine.

The user defines a number of parameters as input for the algorithm, the most important ones are:

- the maximum allowed polynomial order O and
- the maximum allowed grid order

During initialization the input of the algorithm is processed and the needed basis vectors up to the given maximum polynomial order are generated as the set $\mathcal{L}(O)$. The algorithm then determines the weights and abscissas according to the Gauss quadrature rules up to the desired grid order as described in Section 2.3.3 and creates the difference formulas for the numerical integration. For each grid $\underline{l}$ an error indicator is defined as in Equation 2.44, with $n_{\underline{l}}$ the number of quadrature points in the grid.

$$g_{\underline{l}} = \max\left(\alpha \frac{\Delta_{\underline{l}}^{(N)} f}{\Delta_{\underline{1}}^{(N)} f}, (1-\alpha)\,\frac{n_{\underline{1}}}{n_{\underline{l}}}\right) \tag{2.44}$$

The first term of the error indicator gives a measure of how much that particular grid contributes to the value of the integral (as defined in Equation 2.28) relative to the first grid (denoted by $\underline{1}$). The contribution of the grid to some parameter f is denoted by $\Delta_{\underline{l}}^{(N)} f$. The second term indicates the relative computational cost when compared to the first grid. The relative importance between the two terms is determined by $\alpha \in [0, 1]$, which can be set by the user. In this research α is set to 1, which means that we are not interested in the computational cost of the grid, but only in the relative contribution of that grid to some not yet defined parameter f.

The grids are divided into two sets, an old set $\mathcal{O}$ and an active set $\mathcal{A}$. In each step of the algorithm the grid with the highest error indicator in the active set is added to the old set. For all the grids the contribution to the basis vectors is calculated, which means that the expansion coefficients of the basis vectors are updated. New grids are added to the active set by identifying the forward neighbors of the grid $\underline{l}$ denoted by $\mathcal{F}_{\underline{l}}$.

The expansion coefficient for the constant term gives the mean of our parameter of interest. The most up to date estimate for the mean $\tilde{\mu}_R$ is given by Equation 2.45.

$$\tilde{\mu}_R = \tilde{r}_0 = \sum_{\underline{l} \in \mathcal{O} \cup \mathcal{A}} \Delta_{\underline{l}}^{(N)} r_0 \tag{2.45}$$

Here $\Delta_{\underline{l}}^{(N)} \mu_R$ stands for the contribution of grid $\underline{l}$ to the mean. This can be calculated according to Equation 2.46.

$$\Delta_{\underline{l}}^{(N)} \mu_k = \Delta_{\underline{l}}^{(N)} r_0 \tag{2.46}$$

The sum of the squares of the other expansion coefficients, multiplied by the norm of their corresponding polynomials, gives the variance of our parameter of interest. The most up to date estimate for the variance, $\tilde{\sigma}_R^2$, is given in Equation 2.47.

$$\tilde{\sigma}_R^2 = \sum_{k=1}^{P} \tilde{r}_k^2 h_k^2 = \sum_{k=1}^{P} \frac{1}{h_k^2} \left(\sum_{\underline{l} \in \mathcal{O} \cup \mathcal{A}} \Delta_{\underline{l}}^{(N)} r_k \right)^2 h_k^2 \tag{2.47}$$

The contribution of grid $\underline{l}$ to the variance, $\delta_{\underline{l}}^{(N)} \sigma_R^2$, can unfortunately not be calculated so easily and is approximated as in Equation 2.48.

$$\delta_{\underline{l}}^{(N)} \sigma_R^2 \approx \delta_{\underline{l}}^{(N),h} \sigma_R^2 = \sum_{k=1}^{P} \left[\tilde{r}_k^2 - \left(\tilde{r}_k - \frac{1}{h_k^2} \Delta_{\underline{l}}^{(N)} r_k \right)^2 \right] h_k^2 \tag{2.48}$$

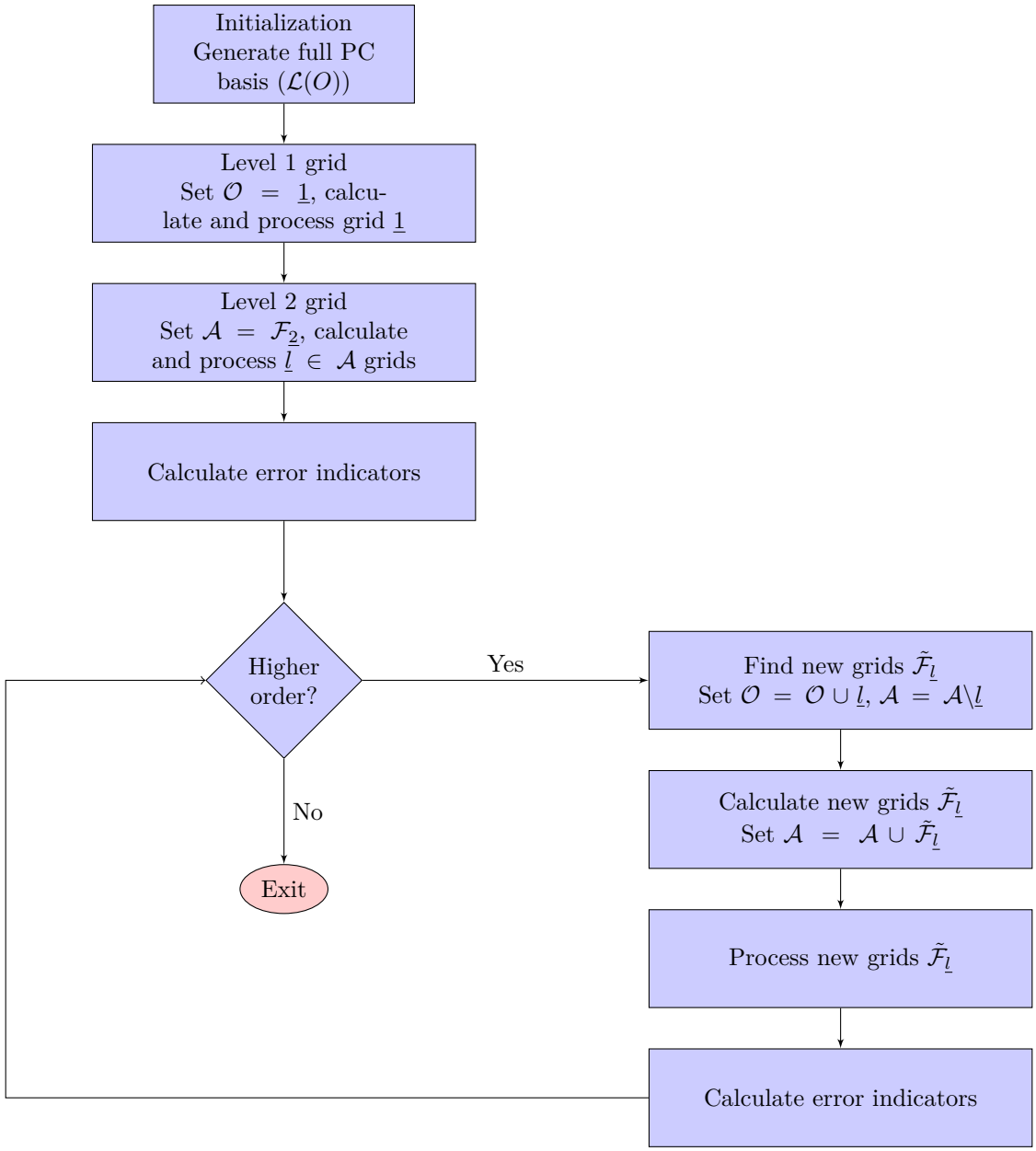


Figure 2.7: Flow chart of the FANISP algorithm. The non-adaptive part of the algorithm is shown. For the full flow chart, see [16].

This process can be best depicted in the flow diagram displayed in Figure 2.7, describing the most important steps.

As a first step the level 1 grid is calculated with all the input parameters at their mean value, i.e. unperturbed values for all the input parameters. This is the level 1 grid corresponding to the multi-index $\underline{1}$. The old set is then initialized to $\mathcal{O} = \underline{1}$ and the first estimates of the PC coefficients are calculated as $\tilde{r}_k = \frac{1}{h_k^2} \Delta_{\underline{1}}^{(N)} r_k$. The set of processed grids is initialized as $\mathcal{I} = \underline{1}$.

In the second step, the level 2 grids are processed. Processing a grid in this context means that the active grid set is initialized with $lev = 2$ grids and the model is run for the separate input realizations indicated in Figure 2.5 for all N directions. Therefore, this requires $2 \cdot N$ calculations in order to obtain the desired output realisations $R\left(\underline{\xi}\right)$. For example, in the case of only modeling setup uncertainties, the subgrids are $(1,1,2)$, $(1,2,1)$ and $(2,1,1)$. With the collected output realisations, the polynomial chaos basis coefficients are updated with the most up to date values.

For each of the subgrids the error indicator is calculated according to Equation 2.44. If allowed by the user, higher order level grids are added to the active set. These subgrids are consequently calculated and processed. This loop continues until the active set is empty and no more grids can be added to the grid.

Adaptivity

The algorithm can also be used in an adaptive way. Adaptivity is based on the idea that if a certain grids have a small contribution to the integral then the contribution of its forward neighbours will be even less significant. Only grids with a significant contribution will be calculated and neighbours of grids with a lower contribution will thus be left out of the calculation. In that way less grids are needed to obtain results with the same accuracy and calculation time can be decreased significantly.

In order to use FANISP adaptively, the user must specify additional input parameters for the program. These are:

- Tolerance level for the mean ϵ_μ
- Tolerance level for the variance ϵ_{σ^2}
- Parameter of interest that needs to be calculated adaptively
- Method for stopping the algorithm

For this purpose, the earlier defined error indicator in Equation 2.44 is used in such a way that the algorithm stops adding new subgrids to the active set when the error indicator falls below a certain user-defined tolerance limit. In this research the improved Gerstner method is used, which takes into account both the contribution of subgrids to the mean as well as to the variance of a parameter of interest [16]. The error indicator for the improved Gerstner method is defined in Equation 2.49.

$$\underline{g}_{\underline{l}}^{IG} = \left(\frac{\Delta_{\underline{l}}^{(N)} \mu_R}{\tilde{\mu}_R}, \frac{\delta_{\underline{l}}^{(N)} \sigma_R^2}{\tilde{\sigma}_R^2} \right) \in \mathbb{R}^2 \tag{2.49}$$

By separately registering the contribution of the grids to μ_R and σ_R^2, the algorithm ensures that both the mean and the variance are calculated up to the specified tolerance level. Generally, the mean converges much faster than the variance, thus the tolerance level for the variance may be somewhat relaxed by the user. The tolerance levels always have to be specified during the initialization of FANISP, thus when the algorithm is used in a non-adaptive way, the tolerances are both set to zero.

Other adaptive methods are also available within FANISP. However, the focus of this research is not on adaptivity, thus for a flow chart for the adaptive algorithm or other methods for stopping the algorithm, see [16].

Chapter 3

Application

The theoretical framework introduced in Chapter 2 is in this project applied to proton therapy treatment plans. All aspects regarding the implementation and application of polynomial chaos to this particular project is summarized in this chapter. Section 3.1 explains all relevant anatomy for the investigated patients. The codes responsible for the generation of treatment plans and the modeling of uncertainties are introduced in Section 3.2. The tools necessary to evaluate a treatment plan from a medical point of view are given in Section 3.3. The polynomial chaos method can be applied via two different routes, which are explained in Section 3.4. Finally, the methods used to validate the polynomial chaos results in terms of accuracy are described in Section 3.5.

3.1 Anatomy

Tumors can be classified by the primary site of origin in the body, by tissue type, size of the tumor and stage of the cancer. In this project head-and-neck cancer patients are studied. Head and neck cancer is the sixth most common type of cancer, representing about 6% of all cases [20]. This type of cancer is especially suited for proton therapy, as tumors are often located in the vicinity of important radiation sensitive organs. In many cases surgery can be too risky or even impossible, and the radiation dose given to the healthy tissue by conventional radiotherapy can be too high. In this section the medical terms describing the tumors are introduced and a short overview of the organs at risk is presented.

The tumor visible on the CT scan is called the gross target volume (GTV). Surrounding the GTV is a small area where, based on years of clinical experience, physicians suspect that the tumor is also present. This margin, together with the GTV forms the clinical target volume (CTV). It is this volume that is to be irradiated and has the highest priority of receiving the prescribed dose in the treatment planning process. The CTV is further divided into a high dose area denoted by *'CTV high'*, which contains the primary tumor, and a lower dose area, denoted by *'CTV low'*, which contains lymph glands that have a high risk of metastases. The CTV high and low are prescribed with different doses. Organs surrounding the tumor are denoted as organs at risk (OAR). Radiation damage to any of these organs can severely impact the patient during and after treatment as they are critical for functions in everyday life. The quality of life after treatment is significantly influenced by how well the patient is able to perform in areas such as eating and speaking [21], processes in which the salivary glands play a crucial role.

The major salivary glands are the *parotid glands* and the *submandibular glands*. Their main function is to produce saliva and the enzyme amylase. Among other things, saliva protects the oral cavity against bacteria and begins the digestion of starches. During radiotherapy treatment the salivary glands can get severely damaged and start malfunctioning. Quite often this leads to xerostomia, also known as dry mouth disease, a serious condition which symptoms include eating and speaking difficulties and severe dental problems. Xerostomia is the most common radiation-induced side-effect for head and neck patients and its occurrence depends on the dose delivered both to the parotid glands and to the submandibular

glands. When the mean dose remains below 39 Gy, the salivary glands tend to recover over time and prolonged xerostomia can be avoided [22].

The parotid glands are the biggest of the salivary glands and are located on the left and right of the head, right next to the ears. They produce approximately 25% of the saliva. The submandibular glands are located beneath the lower jaw (*mandible*) and even though they are much smaller than the parotid glands, they produce around 70% of the saliva. The third type of salivary glands, the sublingual glands, are not evaluated separately in this research, because they contribute only marginally to the saliva production (5%). In Figure 3.1 the location and relative size of the salivary glands are shown.

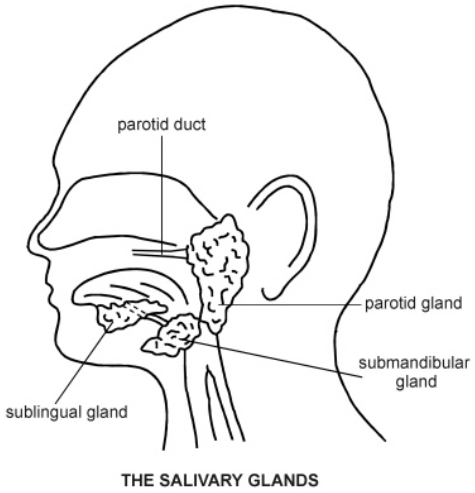


Figure 3.1: Location and relative size of the salivary glands in the head [23]. The parotid and submandibular glands are crucial for saliva production, a process affected by radiotherapy treatment.

The *spinal cord* together with the brain forms the central nervous system of the body. It can be seen as the most important communicator between the brain and the rest of the body. The brain sends movement commands to the body and receives back sensory information via neural signals running through the spinal cord. The spinal cord is also independently responsible for a number of body reflexes, such as jerking away from sources of heat or pain before the brain is even aware of it. The *brainstem* is located at the top of the spinal cord, connecting the cord with the motor and sensory systems in the main part of the brain. The brainstem has a wide variety of important functions, ranging from the regulation of the cardiac and respiratory functions to allowing the brain to detect important sensations such as touch, vibration, temperature and pain. Radiation induced complications can for example include acute paralysis or hemorrhages in the spinal cord.

Finally, organs involved in swallowing and speaking are researched in this thesis. Swallowing starts in the *oral cavity*, where the food is moistened with saliva and crushed until it is suited for digestion. The food is then transported by a number of muscles, which from top to bottom are the superior, middle and inferior constructor muscle (*SCM, MCM* and *MCI* respectively) and the cricopharyngeus muscle (*MCP*). Together with the top of the *esophagus* they form the swallowing muscles. The esophagus is the last step in the swallowing mechanism and transports the food and liquids to the stomach. The *larynx* is involved in breathing and any type of sound production, explaining its more commonly known name of the voice box.

3.2 Treatment planning

The goal of treatment planning is to make a plan that ensures that a specified desired dose is delivered to the tumor, whereby the probability that the tumor can be controlled is maximized. When making the plan, the delivered dose to the surrounding tissue and specifically the critical organs mentioned in Section

3.1 is also taken into account. The probability that any serious complications occur in this normal tissue should be minimized, or else the quality of life of the patient may be affected too severely. Treatment planning can thus be seen as a balancing act between curing the patient and sparing healthy tissue as much as possible. The prescription dose for the tumor and any maximum doses for surrounding organs are set by the radiation oncologist. A treatment plan consists of the selection of beam directions and subsequent beam weights to obtain a desirable dose distribution within the patient. The delivered dose is directly proportional to the number of protons in the beam, a value captured in the beam weight.

Treatment plans are generated in the iCycle treatment planning system, developed by the Erasmus Medical Center Cancer Institue. Erasmus iCycle is a multi-criteria prioritized optimization algorithm, originally developed for Intensity Modulated Radiation Therapy, which was extended with proton pencil beam scanning [24]. In this section all steps necessary to create a treatment plan are explained.

3.2.1 Generation of treatment plans

The process of generating a treatment plan starts with a planning CT of the patient. A radiation oncologist delineates the tumor and all relevant structures on this planning CT. A radiotherapy technologist then selects suitable beam angles manually based on experience. Often for head and neck patients the selected beam angles are $(-60°, 60°, 180°)$ or $(-50°, 50°, 180°)$ [22]. In 2013 a novel type of beam angle selection was implemented in Erasmus iCycle for proton therapy treatment planning [25]. In beam angle optimization the algorithm starts with the 'standard' beam angles and optimizes the beam angles. For each beam direction, iCycle determines all possible positions in the body that can be irradiated with these specific angles and with the available medical equipment.

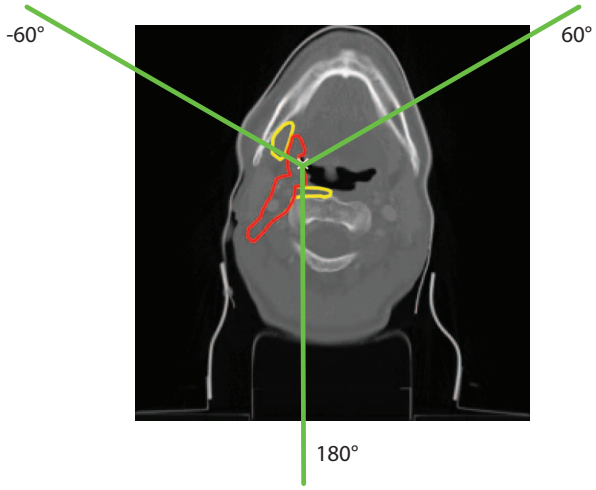


Figure 3.2: 'Standard' beam angles $(-60°, 60°, 180°)$ for patient 1. The beams coincide in the treatment isocenter.

A dose engine is used to determine where in the body the dose is deposited. The dose algorithm used for Erasmus iCycle to calculate the delivered dose was developed at the Massachusetts General Hospital and Harvard Medical School and is based on analytical calculations [26]. For the calculations the concept of pencil beams, introduced in Section 1.2, is used.

The dose calculation algorithm uses for each beam direction a so-called 'bixel grid' to account for density heterogeneities inside the patient. Firstly, a dose matrix is calculated for a very finely meshed (2x2mm) two-dimensional grid in the plane perpendicular to the beam axis (the bixel grid). This dose matrix describes in which voxels within the patient what amount of dose would be deposited if any of the bixels in the grid is 'activated'. The width of the pencil beam is considerably larger than a single bixel. The intensity of the beam is described by a Gaussian distribution with a standard deviation of 3-7mm, thus with a higher intensity in the (center) of the beam and lower towards the edges. In a second step, each pencil beam is convoluted with the two-dimensional bixel grid, thereby determining how much each part

of the bixel grid is 'activated' and thus where and how much that particular pencil beam deposits its energy. This gives the dose distribution in the patient for a unit weight pencil beam. This process is repeated for all beam directions.

The next step is to determine which of the possible pencil beams is to be used and what weight is to be assigned to the individual pencil beams in order to obtain the desired dose distribution in the patient. Pencil beams are selected according to the resampling method [27]. This method has as advantage that it can reduce the optimization time dramatically. In this method a relatively small subset of candidate pencil beams is randomly selected. The number of pencil beams in the subset is pre-defined by the user, in the treatment plans generated in this research subsets of 3000 or 5000 pencil beams were used. For this subset of candidate pencil beams the dose matrix is calculated and the spot weights are inversely optimized. Inverse optimization means that the desired dose is known and the pencil beam weights are optimized in such a way that it best resembles the desired dose. The pencil beams that contribute the least to the dose distribution are eliminated to reduce the number of pencil beams. In the next iteration a new subset of randomly selected pencil beams is added to the existing ones. For this new collection of pencil beam again the dose matrix and the pencil beam weights are calculated and, if needed, the number of pencil beams is again reduced. The algorithm stops when the addition of new pencil beams does not contribute any more to better plan quality. The construction of the full dose distribution in the patient is complete when the contribution of all individual pencil beams is added.

Optimization

The core idea behind iCycle is the wish-list, a list containing desired tumor and organ parameters. These parameters can either be constraints or objectives. The difference between them is that constraints must always be met, while the objectives are accomplished as much as possible. For physicians the tumor coverage will always have the highest priority in the treatment plan. This can be accomplished by setting the minimum dose delivered to the CTV as a constraint. This is the requirement that, in order for the plan to be accepted, has to be met. The dose values of the CTV are chosen in such a way that the V_{95}, the relative volume receiving 95% of the prescribed dose, is always above 98% and the V_{107}, the relative volume receiving 107% of the dose, is not higher than 2% of the CTV.

Objectives are much less stringent than the constraints. For serial organs, such as the spinal cord, the maximum dose delivered to the organs is kept as low as possible. This is because serial organs can stop functioning after one of its subvolumes is damaged. For parallel organs, such as the salivary glands, the mean dose to the organ is minimized as much as possible. They will only lose complete functionality if the whole organ is damaged. For the objectives the goal is not to meet the specified dose exactly, but if possible to further minimize the dose to the organ while still satisfying the dose prescription of the target. Objectives are optimized in the order in which they appear in the wish-list. The higher the priority, the more importance is given to this particular parameter in the optimization process. The objectives are optimized one by one. After each objective minimization, and based on the obtained result, a constraint is defined for the just minimized objective, which is used as an extra constraint in the following minimizations of lower prioritized objectives [24]. The minimization of low priority objectives will in this case never jeopardize the obtained results for the higher prioritized objectives. The optimization stops when none of the objectives can be optimized any further. In Table 3.1 the wish lists used in this research are shown.

Robust treatment plans

Dose delivery by proton beams can be very sensitive to uncertainties (Section 1.3). In order to reduce the sensitivity to uncertainties, especially in target coverage, robust treatment planning has been developed over the past few years. In robust treatment planning several error scenarios, i.e. displacements of the patient and deviation in the range of the beams, are taken into account in the optimization. The method used for Erasmus iCycle is called minimax optimization, which aims at minimizing the penalty of the worst case scenario [28].

Table 3.1: Simplified wish-list containing the constraints and objectives for the patients used in this research

Area of interest	Priority	Type	Goal non-robust plan	Goal robust plan
'CTV high'	Constraint	Minimum	0.99×66 Gy	0.98×66 Gy
'CTV low'	Constraint	Minimum	0.98×54 Gy	0.98×54 Gy
'CTV high'	1	Maximum	1.07×66 Gy	1.07×66 Gy
'CTV low'	1	Maximum	1.08×54 Gy	1.07×54 Gy
Parotid glands	3	Mean	0	0
Submandibular glands	4	Mean	0	0
Spinal cord	5	Maximum	20	20
Brainstem	5	Maximum	20	20
Larynx	6	Mean	0	0
Oral cavity	6	Mean	0	0
Swallowing muscles	7	Mean	0	0

For each of the error scenarios relevant dose parameters are calculated. In the case of robust optimization, the worst case of the relevant parameters is optimized. If for example the minimum dose of the tumor is to be optimized in a robust way, first the minimum doses are derived from the full dose distributions of all the error scenarios. The worst result is then optimized and limited to a more acceptable value. Robust optimization can come with the price that higher doses are delivered to the surrounding tissue [25].

For the researched robust plans, the error scenarios that were taken into account in the optimization were 6 setup error scenarios with displacements of ±3mm along the 3 main axes defined within the patient and 2 range error scenarios of an undershoot or overshoot with 3% relative range error and 1mm absolute range error.

3.2.2 Modeling errors

In order to accurately describe set-up uncertainties, three main axes in the body are defined. These three main axes are conveniently described as the x-, y- and z-axis and they form the usual orthogonal Cartesian coordinate system. The origin of the system is the treatment isocenter, which is defined by the radiation oncologist and is typically located in the center of the target volume. The x-axis is defined from the right of the body to the left, the y-axis runs from the front to the back and the z-axis from the bottom to the top. The plane defined by the x- and y-axis is called the transverse plane. The plane defined by the x- and z-axis is called the coronal plane and the plane defined the y- and z-axis is called the sagittal plane. In Figure 3.3 the different axes and planes are graphically depicted.

Setup errors are modeled as displacements of the treatment plan isocenter, corresponding to a misalignment of the patient relative to the gantry [28]. The uncertainty distribution of the setup errors is assumed to be Gaussian (Equation 2.1) and equal in all directions. In literature the uncertainty in head-and-neck cancer patients is given as 3mm [8] (85% confidence interval), which roughly corresponds to a standard deviation of 2mm. The uncertainty in any of the three spatial variables is then modeled with a Gaussian distribution with a mean of $\mu = 0$m, which corresponds to the nominal planned scenario, and the standard deviation is $\sigma = 2$mm. For the robust treatment plan optimization the error scenarios are all calculated with a deviation from the origin of 3mm.

Range errors are essentially uncertainties in the material parameter. This influences the stopping power of the protons and therefore the range of the proton beam. An alternative cheap way of calculating this type of uncertainties is to model them as a deviation from the nominal intended energy of the proton beam. For the quantification of the range errors, only a relative range error is considered. The distribution is again assumed to be Gaussian, with the mean $\mu = E_0$, corresponding to the energies in the nominal planned scenario and the standard deviation $\sigma = 2\% \cdot E_0$. This again translates to a 85% confidence interval of 3%, a value often found in literature ([8],[30]) and also used in the robust optimization of the

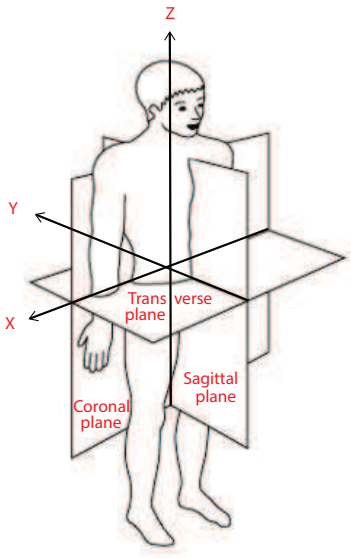


Figure 3.3: Figure depicting the major anatomical axes and planes defined in the body. Based on [29].

plans.

It is relatively straightforward and not computationally intensive to recalculate a new dose distribution with a perturbed input. To this end, the same very finely spaced two-dimensional grid as described in Section 3.2 is again calculated. This time however the pencil beams are laterally shifted over the bixel grid to model the displacement. Again this is multiplied with the beam weights and added, which gives the recalculated dose distribution. Theoretically, an infinite number of scenarios can be recalculated from this one very fine bixel grid, in practice Matlab can handle around 100 scenarios. Because the two-dimensional grid does not have to be recalculated for every scenario separately, the procedure to recalculate X number of scenarios does not scale linearly with X.

3.3 Medical evaluation of treatment plans

The generated treatment plan is evaluated by a doctor before treatment commences. Key objective is to find out if the goal of treatment planning, maximizing the curing of the patient balanced with the minimization of treatment side effects, is achieved with this treatment plan for this particular patient. Evaluation is typically done in the following three steps.

1. Visual inspection of the 3D dose distribution
2. Investigation of the dose-volume histograms of the tumor and relevant organs
3. Checking the values of a list of tumor and organ parameters

Each of these points is discussed in more detail in the coming subsections. Keep in mind that in this research the focus lies not only on the absolute values of the dose distribution and derived quantities, but primarily on how well polynomial chaos expansion is able to approximate these values.

3.3.1 Full dose distribution

The full 3D dose distribution gives the physician a clear view of how the different structures in the head and neck will be affected by the proton beams. One of the questions in this research is how well

polynomial chaos can approximate the (expected) full dose distribution. The first step in this research was to investigate how well the dose in individual random voxels can be approximated and what kind of grid order is needed to achieve a reasonable accuracy. The second step is to approximate the full dose distribution, which consists of over 10 million voxels.

3.3.2 Dose Volume Histogram

One of the most widely used ways of visualizing the dose distribution inside the patient is with the help of a Dose Volume Histogram (DHV). This visual tool converts the 3D dose distribution inside the body to a convenient 2D representation. On the vertical axis the relative volume is depicted and on the horizontal axis the received dose. Each different organ or tumor area of interest is depicted with a different line. Along the line the doctor can read what part of the volume of an organ receives what amount of the dose. It is common practice to display all of the areas of interest in one single graph, giving the evaluating physician a direct overview of the dose distribution, see also Figure 3.4. A downside of this type of representation is that it does not give any spatial information, so it is not possible for example to see where in an organ a radiation hot spot or cold spot would occur.

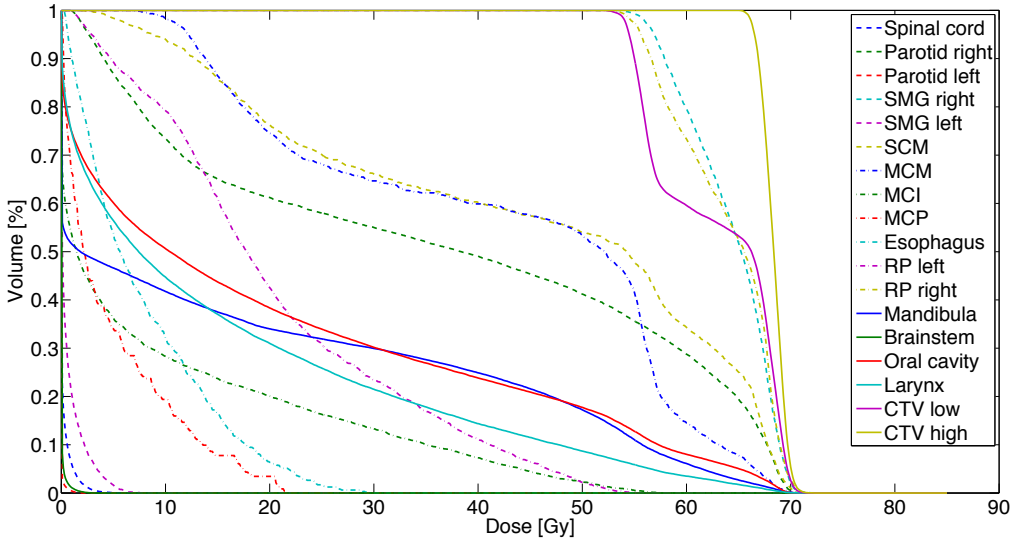
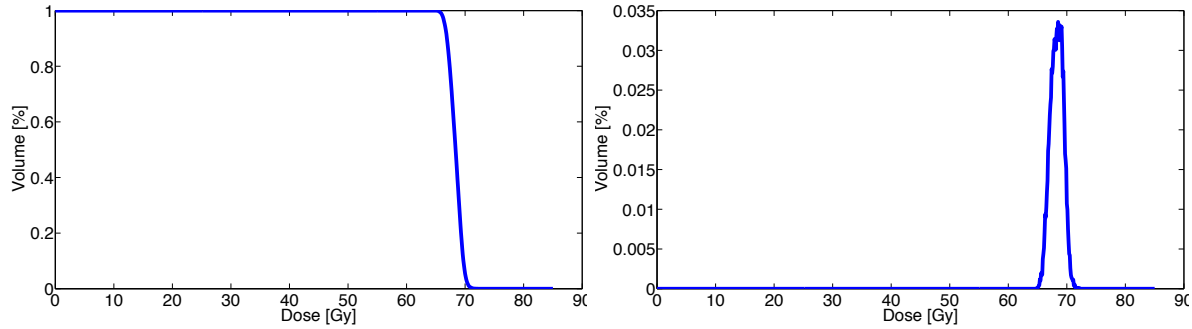


Figure 3.4: Cumulative dose-volume histogram of patient 1. All relevant organs and structures that are evaluated are shown in this single graph.

A DVH exists in two forms, the most commonly used is the cumulative dose volume histogram. An example of a cumulative DHV can be seen also in Figure 3.5a. The construction of a cumulative DVH is relatively straightforward. The histograms in this research are built up with a dose resolution of 0.1 Gy. To make sure all relevant information is included, the DVH is constructed from 0 to 85 Gy for all organs of interest. At each dose step the number of voxels is counted that have a dose higher than that particular dose. This number is divided by the total number of voxels in the organs to obtain a relative volume. This process is repeated for relevant organs and structures. For the irradiation of the tumor, it is preferred to have a homogeneous dose throughout the tumor, preferably resulting in a horizontal line at the top of the graph, ending in a vertical drop around the prescribed dose. On the other hand, critical organs should receive as low dose as possible, resulting in lines closer to the axes.

The second type of DVH is a differential dose volume histogram, in which for each dose interval the relative volume is depicted that receives a dose lying in that interval. This is mostly useful for the low and high dose CTV areas, because their cumulative DVH has a very steep gradient. A differential DVH might offer a more information in that case. In Figure 3.5b an example of a differential DVH is given. The construction of a differential DVH is similar to a cumulative DVH. The same dose resolution is used,



(a) Cumulative dose-volume histogram **(b)** Differential dose-volume histogram

Figure 3.5: Examples of the two types of dose-volume histograms. Structure of interest depicted here is the 'CTV high' of patient 1

but in this case for each dose interval the number of voxels receiving a dose that lies within the specific dose interval is counted. Again this value is normalized with the total number of voxels in the area of interest. This process is repeated for both tumor areas.

The depicted DVH in Figure 3.4 represent dose-volume histograms of the treatment plan. In practice, due to uncertainties in the dose delivery, the DVH will be slightly different with each perturbation. In order to give doctors an estimate of how accurate the planned DVH is when including uncertainty, a bandwidth can be constructed around the planned DVH. For each dose point the 5^{th} percentile is determined as a lower boundary and the 95^{th} percentile as the upper boundary. The bandwidth gives the doctor a clear view of the uncertainties that can occur during dose delivery. In Figure 3.6 an example for the bandwidth is shown.

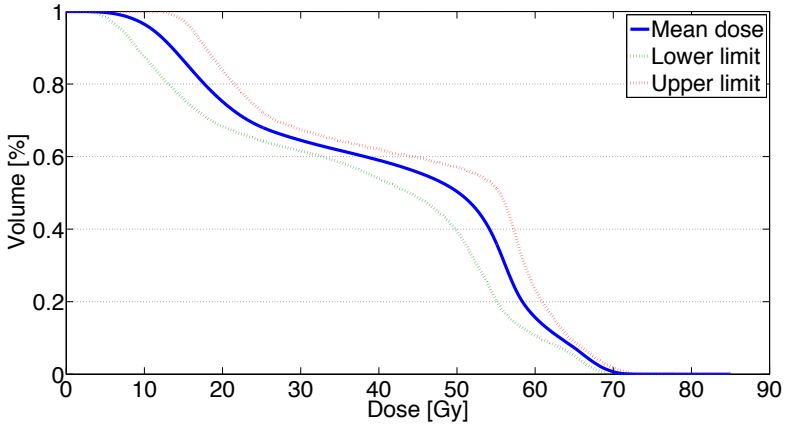


Figure 3.6: Dose-volume histogram of the MCM. The lower and upper limit of the bandwidth have been constructed from the 5^{th} and 95^{th} percentile, obtained after 10,000 random samples.

3.3.3 List of parameters

The first priority of the physician will always be on the irradiation of the tumor area. In order to maximize the probability that the tumor can be controlled, it is of great importance that the whole tumor area receives enough dose, even in the presence of uncertainties. For this reason, the distributions of the minimum dose, the mean dose and the maximum dose are evaluated separately. The distributions are described by their mean and variance.

Tumor coverage (V_{95} and V_{107}) For the description of tumor coverage, the well established definitions of V_{95} and V_{107} are used. V_{95} is the percentage of the target volume that receives 95% of the prescribed dose. In a similar way, V_{107} represents the percentage of the target volume that receives 107% of the prescribed dose. A plan is considered acceptable if V_{95} is above 98%, whereas the preferred value for V_{107} is below 2%.

Uniformity index (UI) The homogeneity of the dose in the tumor is evaluated by the uniformity and homogeneity index. The uniformity index (UI), defined in Equation 3.1, compares the minimum dose delivered to 5% of the CTV (D_5) to the minimum dose delivered to 95% of the CTV (D_{95}). It essentially thus compares the lowest dose to the highest dose in the tumor. The larger the difference between the lowest dose and the highest dose, the less uniform the distribution in the tumor will be and the smaller the uniformity index. A low uniformity index could potentially indicate the presence of hot spots in the tumor.

$$UI = \frac{D_5}{D_{95}} \tag{3.1}$$

Homogeneity index (HI) The radiation dose homogeneity index, abbreviated as HI, compares the difference between the minimum dose delivered to 99% of the CTV (D_{99}) and the minimum dose delivered to 1% of the CTV (D_1) with the prescribed ($\text{Dose}^{\text{prescribed}}$) dose in the CTV. In simpler terms, it subtracts the lowest dose from the highest dose in the tumor and compares this to the prescribed dose. The larger the difference between the highest and the lowest dose, the less homogeneous the distribution is and subsequently the homogeneity index will be larger. Equation 3.2 gives the definition of the homogeneity index.

$$HI = \frac{D_{99} - D_1}{\text{Dose}^{\text{prescribed}}} \tag{3.2}$$

All the mentioned target coverage parameters, i.e. V_{95}, V_{107}, UI and HI, are derived both for the 'CTV low' and the 'CTV high'. In this thesis they are grouped as the 'tumor parameters'.

Only when tumor coverage is ensured, the different organs in the vicinity of the tumor are evaluated. All organs and structures mentioned in Section 3.1 are the subject of evaluation in this research. For all the organs, the distributions of the minimum dose, the mean dose and the maximum dose are derived and described by their mean and their variance. In this thesis these three quantities (the minimum, the mean and the maximum dose), are denoted as dose parameters, this name holds both for the tumor areas and the organs at risk.

3.4 Direct route versus indirect route

In the case of proton therapy treatment planning any parameter of interest can be calculated in two ways, via the direct route or via the indirect route. A short description of both routes is given in this section, along with the main advantages of the routes and a discussion on why the different routes are relevant.

In the direct route parameters of interest and data points for the dose-volume histograms are directly chosen as responses of the system. This method can be used for any number of parameters. The responses of interest are simply collected in a vector and the corresponding expansions are computed simultaneously. The means and variances of the responses are simply given as explained in Section 2.4. In the case of the dose-volume histogram, even though the mean value can be read immediately, the construction of a bandwidth requires the polynomial chaos expansions of the different data points to be resampled.

In the indirect route the dose distribution in *all* voxels is calculated up to the specified grid and polynomial order. In this way the complete dose calculation algorithm is replaced by a reduced order system of polynomials and calculated coefficients for each voxel. The main advantage of this method is that it contains all possible information about the full dose distribution, so it can also be used to derive any dose parameter or dose-volume histogram of interest. In order to obtain these parameters or histograms, (parts of) the full distribution need to be resampled with enough samples to obtain a good relative accuracy

for the desired responses. The statistics of this resampled distribution can be obtained as described in Section 2.2. Computationally this comes at a price. All relevant organs together contain around 200,000 voxels that need to be resampled at least 10,000 times in order to obtain results with an acceptable accuracy for only approximately 50 parameters of interest and 15,000 points for all relevant dose-volume histogram data points. The difference between the two routes is depicted in the flow-chart in Figure 3.7.

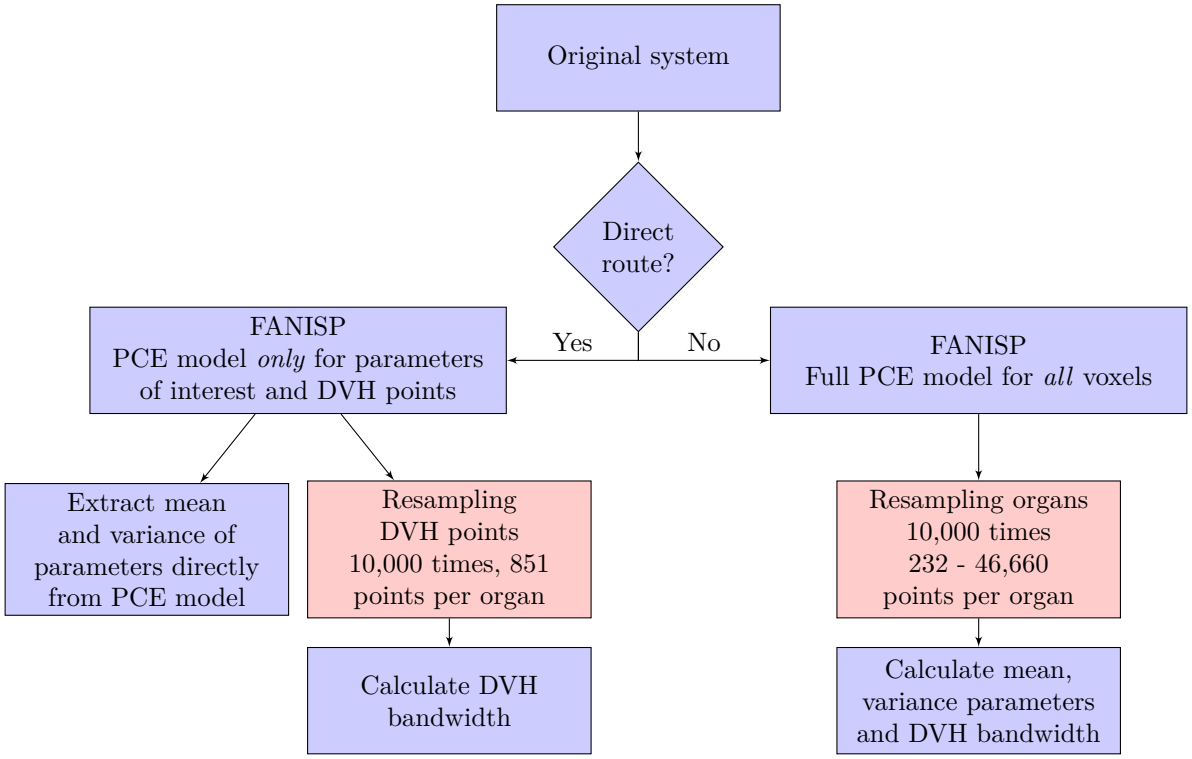


Figure 3.7: Flow chart showing the difference between the direct and the indirect route. Example for determining a dose parameter and the bandwidth of a dose-volume histogram. Blocks indicated in red take up most of the computation time. Dose-volume histograms are constructed in the range 0-85 Gy with a dose resolution of 0.1 Gy, resulting in 851 data points per DVH. The size of the organs ranges from 232 to 46,660 voxels.

The large scale of the full dose distribution, i.e. the large number of voxels poses two problems. First, the computational time needed for the processing of subgrids takes substantially more time than for smaller applications. Second, the matrices within FANISP become unpractically large, Matlab runs normally out of memory before higher order grids can be processed. This indirect route is thus only possible for a second order grid. A compromise solution in terms of computation time would be to derive the list of parameters named in Section 3.3 and the dose volume histograms of all organs of interest via the direct route and the full dose distribution (obviously) via the indirect route.

The two routes do not necessarily give the same values. The suggested compromise solution then could give problems in a consistent representation of the S&U analysis of the treatment plan. In this research the differences between the different routes are investigated in more detail in order to determine if the differences are clinically relevant.

3.5 Methods of comparison

Results for the polynomial chaos expansion are validated with a reference value, created by one of the statistical methods introduced in Section 2.1. For each of the treatment plans of patient 1 10,000 random sampling simulations of the full dose distribution were run, from which all relevant dose-volume histograms and parameters were derived. Furthermore, for all treatment plans also 1,000 Latin hypercube sampling simulations were run. The Latin hypercube sampling results proved to be similar enough to the random sampling results. It was therefore decided to only create 1,000 Latin hypercube samples for the treatment

plans of patient 2 and use the obtained values as a reference value for the polynomial chaos expansion results. All treatment plan simulations were run once with only setup uncertainties (3 dimensions) and once with both setup and range uncertainties (4 dimensions).

3.5.1 Full dose distribution

The approximation of the full dose distribution was first tested on a small scale with the approximation of dose distributions in ten individual voxels. The accuracy of the probability density function calculated by polynomial chaos expansions with increasing orders is determined by the mean squared percentage error (MSPE), defined in Equation 3.3. In each bin i of the histogram the reference value, denoted by D^{ref}, and the polynomial chaos value that is to be evaluated, D^{eval}, are compared.

$$MSPE = \frac{1}{M} \sum_{i=1}^{M} \left(\frac{D^{eval} - D^{ref}}{D^{ref}} \right)^2 \tag{3.3}$$

The accuracy of the full dose distribution is determined several ways. The polynomial chaos model is tested for a set of inputs and the resulting dose distributions are compared with the actual dose distribution for that same set of specific inputs. This gives an idea how well the polynomial chaos model is able to reproduce the calculations of the dose engine. For all treatment plans the mean squared percentage error of the full dose distribution is calculated. In this case i in Equation 3.3 represents the individual voxels. This statistics measure can only be calculated for voxels with a nonzero reference dose. Furthermore, for all plans the absolute error is calculated, which gives the possibility of comparing between the different treatment plans. Finally, the accuracy of the full dose distribution and the dose distributions in the various organs are depicted in error-volume histogram, which shows the absolute errors in relation to the percentage of the volume.

3.5.2 Dose-volume histograms

For each dose-volume histogram three parameters can be evaluated. The first parameter is the accuracy of the approximation of the mean dose-volume histogram. The second and third parameters are the upper and lower limit of the bandwidth around the mean. Visual comparisons are made between dose-volume histograms of increasing order, cumulative and differential histograms, the direct route and the indirect route, and the influence of the range uncertainty on the size of the bandwidth is investigated. Finally, a qualitative comparison with the surface approximation will be made for dose-volume histograms.

3.5.3 Parameters

Suitable comparison criteria for the parameters have been established for the evaluation of the performance of the polynomial chaos expansion method. Parameters can be classified as good, acceptable or bad. These criteria are partly based on the theory described in Chapter 2 and partly on intuition. Recall that the probability density distributions of the various parameters are assumed to be described sufficiently in terms of their means and variance. These quantities are evaluated both in absolute and relative terms. The absolute difference between the reference value and the polynomial chaos value is compared against the absolute criteria defined in Table 3.2. These criteria are based on intuition. The relative difference is calculated with Equation 3.4.

$$\epsilon_{rel} = \frac{D^{eval} - D^{ref}}{D^{ref}} \times 100\% \tag{3.4}$$

This relative value is compared against $\sigma_\mu / \mu_x \times 100\%$ as defined in Equations 2.6 and 2.8. For the variance, the obtained relative value is compared against $\sigma_{s^2} / s^2 \times 100\%$ as defined in Equations 2.9 and 2.11. The criteria are given in Table 3.2. A distinction is made between the dose parameters and the tumor parameters. Since the tumor parameters are fractional values between 0 and 1, it is not possible

to establish an objective criterium suited for absolute differences and only the relative errors are used in the evaluation.

Table 3.2: Performance criteria for the dose and tumor parameters based on the absolute or relative difference between the reference value and the polynomial chaos expansion method.

	Dose parameters		Tumor parameters	
Good	Absolute difference	< 0.1 Gy	Relative difference	$< \pm 1.5\sigma$
	or Relative difference	$< \pm 1.5\sigma$		
Acceptable	Absolute difference	< 1.0 Gy	Relative difference	$< \pm 3\sigma$
	or Relative difference	$< \pm 3\sigma$	Relative difference	$< \pm 1\%$
	or Relative difference	$< \pm 1\%$		
Bad	None of the above		None of the above	

Chapter 4

Results

In this chapter the results of the research are presented. First the investigated patients and plans are introduced in Section 4.1. The results of the proof of concept study are presented in three sections. First, the results of the approximation of the dose in individual voxels and the full dose distribution are discussed in Section 4.2. Second, polynomial chaos expansions for dose-volume histograms are discussed in detail in Section 4.3. The proof of concept study is concluded with the results for the approximation of various dose and tumor parameters in Section 4.4. A qualitative comparison with the surface approximation is made in Section 4.5. Section 4.6 shows the outcome of the discussion on consistency between the direct and the indirect route introduced in Section 3.4. The performance of polynomial chaos is evaluated both for dose-volume histograms and dose and tumor parameters. Next, the influence of robust treatment planning is analyzed in Section 4.7. Of interest are the sensitivities to uncertainties in the different plans and the performance of the polynomial chaos method. This chapter concludes with the results for variance decomposition in Section 4.8.

4.1 Investigated patients and plans

(a) Patient 1 has a unilateral tumor, located on one side of the head (CT scan is seen from below)

(b) Patient 2 has a bilateral tumor, located across the entire head

Figure 4.1: CT scans for both patients. The CTV high is delineated in red, the CTV low in yellow. In head-and-neck cases the standard prescriptions for the CTV low and the CTV high are 54 Gy and 66 Gy respectively. The white cross denotes the treatment isocenter.

For this research project two patient cases have been selected, which will be denoted by patient 1 and patient 2. Patient 1 has a unilateral oropharyngeal tumor. Unilateral indicates that the tumor is located on one side and oropharyngeal denotes the location of the tumor. This type of cancer patients fall in the category of head-and-neck patients. Patient 1 is depicted in Figure 4.1a. For this patient a non-robust

treatment plan was created to conduct a full proof of concept study on the performance of polynomial chaos expansion in treatment plans with setup uncertainties, including the approximation of individual voxels, dose-volume histogram and various dose and tumor parameters. Later also range uncertainty was added to the model. Next, a robust treatment plan was created for patient 1 to investigate the difference between robust and non-robust planning. Finally, a second patient was introduced to see if the general conclusions drawn from the first part of the research were also applicable to a different patient with a completely different tumor shape. Patient 2 has a bilateral tumor, located across the entire width of the patient's head. In Figure 4.1b a slice of the CT scan is shown with the tumor delineated.

Characteristics of the various treatment plans are summarized in Table 4.1. All treatment plans are analysed twice, once including only setup uncertainties and once including both setup and range uncertainties, making it a total of six researched models.

Table 4.1: Summary of patient treatment plan characteristics.

	Patient 1 Plan 1	Patient 1 Plan 2	Patient 2 Plan 1
Tumor classification	unilateral	unilateral	bilateral
Beam angles	$(0°, -60°, -120°)$	$(-60°, 60°, 180°)$	$(-60°, 60°, 180°)$
Beam optimization	Yes	No	No
Robust optimization	No	Yes	Yes
Number of voxels	12,743,136	12,743,136	14,469,840

4.2 Approximation of the full dose distribution

A visual representation of the full dose distribution allows the physician to have an exact overview of where the dose is deposited in the patient. This visual representation can be given for the nominal scenario, but also for the expected dose distribution. In that case, the dose in each of the individual voxels can be described with its own probability density function, and its own mean and variance.

4.2.1 Approximation of the dose in individual voxels

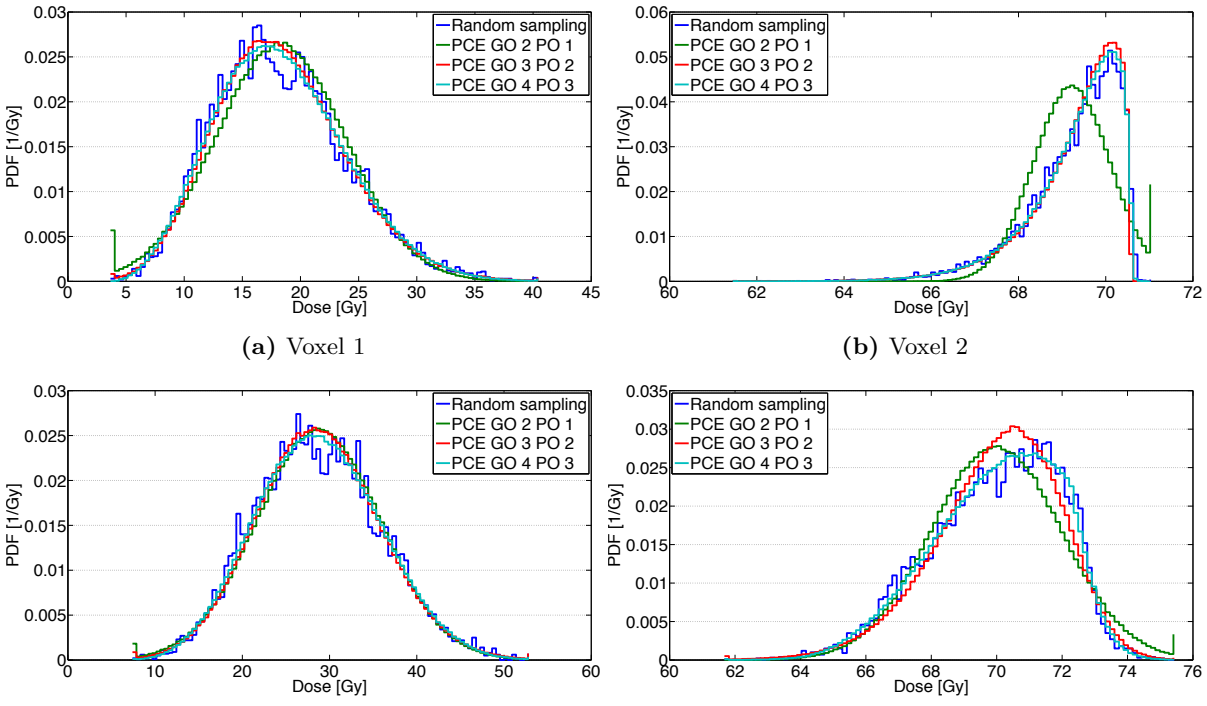
Naturally the first step in the approximation of the full dose within the patient is to see how well individual voxel doses can be approximated. For this purpose ten individual voxels have been randomly selected from patient 1 having various nominal dose levels. From these ten voxels, four voxels have been selected as interesting cases which will be studied further below. For each of the voxels polynomial chaos expansions with increasing grid orders have been calculated. The probability density functions produced by polynomial chaos are compared against the reference probability density functions in each of the voxels. The accuracy of the polynomial chaos PDFs is determined by the mean squared percentage error (MSPE), as defined in Equation 3.3. MSPE is a measure for the average relative error of the polynomial chaos expansion with respect to the reference distribution. The accuracy of expansions with increasing grid and polynomial orders are compared in Table 4.2. It can be seen that a higher grid order delivers a higher accuracy, a trend found in almost all voxels studied.

The impact of higher grid and polynomial order expansions however strongly depends on the shape of the probability density function in the particular voxels. This is also illustrated in Figure 4.2, where the reference probability density functions for the four selected voxels are depicted together with the probability density functions produced by the different polynomial chaos expansions. The PDF is constructed by making a histogram with 100 bins for the reference distribution. The same bins are used to make a histogram for the PDF of the polynomial chaos expansion. The presence of a high peak on either the left side (Figure 4.2a) or on the right side (Figure 4.2b) indicates that polynomial chaos predicts more values to the right or left respectively, but they are all collected in the outermost bin. Generally such a peak is only present for lower order approximations. In voxels with a rapid fall-of in the dose, an example depicted in Figure 4.2b, the lowest order expansion is not sufficient to approximate the distribution

Table 4.2: Mean squared percentage errors for several polynomial orders (PO) expansions with different grid orders (GO) in individual voxels. For all voxels increased computational time (increased grid order) delivers a higher accuracy (lower mean squared percentage error).

	Mean dose [Gy]	GO 2, PO 1 MSPE	GO 3, PO 2 MSPE	GO 4, PO 3 MSPE
Voxel 1	18.0	1.8416	0.8558	0.7059
Voxel 2	69.2	16.5728	1.1311	0.5405
Voxel 3	28.5	1.0481	0.9684	0.7936
Voxel 4	69.9	3.8487	2.1577	0.4842
Voxel 5	7.3	5.3891	0.4925	0.3966
Voxel 6	12.8	3.6560	0.5663	0.5011
Voxel 7	43.0	0.7238	0.8881	0.7056
Voxel 8	50.1	6.1788	0.3160	0.2826
Voxel 9	60.8	1.9127	0.4780	0.4889
Voxel 10	68.9	3.4178	0.8629	0.4239

correctly. In voxels where the distribution appears Gaussian (Figure 4.2c), the low order approximation is already quite accurate. Higher grid orders improve the approximation somewhat, but not drastically, something that could be seen in roughly half of the investigated voxels. For the other voxels however higher orders are preferred to resolve the skewness in the distribution. Figure 4.2d clearly shows that higher polynomial expansions sometimes are needed to correctly follow the shape of the distribution, even though a first order can give a fair approximation of the mean and variance of the distribution. In all graphs the coarse shape of the random sampling distribution suggests that 10,000 random samples might not be enough to create a smooth probability density function. This is less of a problem for the polynomial chaos expansion which has been resampled with 1,000,000 input realizations. The computational cost of this resampling is negligible compared to the computational time needed to produce random samples. It takes about one hour to produce 100 random samples, whereas it takes under one minute to produce the same amount of resamples from the polynomial chaos expansion.



(a) Voxel 1 **(b)** Voxel 2

(c) Voxel 3 **(d)** Voxel 4

Figure 4.2: Probability density function for four randomly selected voxels in patient 1. The impact of higher order expansions strongly depends on the shape of the probability density function of the dose in the voxel.

4.2.2 Approximation of individual dose distributions

The dose in the individual voxels can thus be approximated up to a certain accuracy. The next step is to determine how well a full dose distribution for a specific set of inputs can be approximated when those same inputs are given to the polynomial chaos expansion of the full dose distribution. First, a polynomial chaos expansion was created for all voxels in the patient. As mentioned in Section 3.4, full dose calculations are limited due to memory issues in Matlab due to the large number of voxels. The dose distribution can therefore only be approximated by a first order polynomial and a second order grid expansion (containing only seven quadrature points). From the results of the approximation of the dose in individual voxels it is already known that the low order will be insufficient for an accurate description of the probability density function in some of the voxels. A set of ten random inputs was generated. For each of these random inputs the reference dose distribution was calculated by the dose algorithm. The random inputs were then given to the full polynomial chaos expansion. The produced dose distribution was compared on a voxel-to-voxel basis with the reference dose distribution. The accuracy of the polynomial chaos was again determined by the mean squared percentage error.

A strong dependence was found between the accuracy of the resulting approximated dose distribution and the deviation from the isocenter of the treatment plan. This relationship is depicted in Figure 4.3 and it can be concluded that only plans with deviations close to the treatment isocenter can be recreated with sufficient accuracy. The high values for the MSPE is caused by two factors. First, the dose in all voxels in only approximated by a first order polynomial and a second order grid. A fall-off in accuracy with increasing distance can therefore be expected. Second, in a large number of voxels a near-zero dose is calculated, which can result in a high relative error.

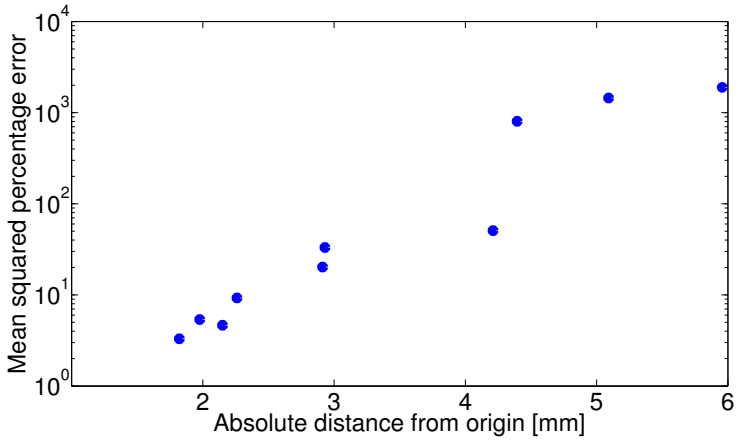
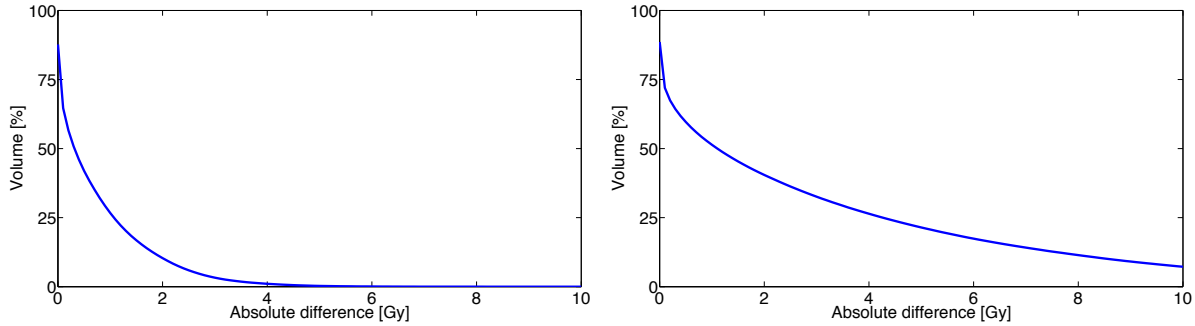


Figure 4.3: Dependence of the accuracy of the approximated dose for random error scenarios as a function of the distance from the origin. The accuracy strongly decreases when the distance from the origin increases.

In order to further determine the accuracy of the dose distribution produced by the polynomial chaos expansion, an error-volume histogram is used. In this type of graph only the absolute error is of interest. The absolute difference between the reference dose distribution and the polynomial chaos dose distribution is determined voxel-by-voxel. At different dose levels, the number of voxels with an absolute error larger than that dose level is calculated and normalized with the total number of voxels evaluated. This is only done for the volume receiving a nonzero dose, as the large amount of zero dose voxels gives an overestimation of how good the approximation is. Approximately 10% of the voxels receives a nonzero dose in the reference dose distribution. The line in Figure 4.4 then represent how accurately the dose is predicted with the polynomial chaos method in the voxels. For a plan with only a small deviation from the treatment isocenter (Figure 4.4a) the polynomial chaos expansion can approximate the full dose distribution with a maximum dose difference of approximately 4 Gy, while for a plan farther from the isocenter (Figure 4.4b) the maximum dose difference increases to well over 10 Gy.



(a) Displacement from origin is 0.1222 mm.

(b) Displacement from origin is 2.2805 mm.

Figure 4.4: Error volume histograms for full dose distributions with different distances from the origin. The smaller the displacement from the origin is, the more accurate the distribution becomes. Only voxels with a nonzero dose in the reference dose distribution are taken into account in the calculations.

4.2.3 Approximation of the expected dose distribution

Finally, the results for the approximation of the *expected* dose distribution can be discussed. The reference expected dose distribution is constructed by taking the mean value in each of the voxels after 10,000 random input realizations. The expected dose distribution via polynomial chaos is constructed by the same polynomial chaos expansion for all voxels which was also used in Section 4.2.2. The mean values in each of the voxels, the first coefficient of the expansions, are collected in an expected full dose distribution. The two distributions have been compared to determine the relative and the absolute error in each voxel. The relative errors were again averaged for all nonzero voxels (in the reference dose distribution) to produce the mean squared percentage error (MSPE). This analysis was repeated for all treatment plans and the outcomes are displayed in Table 4.3.

Table 4.3: Accuracy of the approximation of the polynomial chaos method for the expected dose distribution. Polynomial chaos is here calculated with first order polynomials and a second order grid.

		MSPE	Max absolute error [Gy]
Patient 1 Plan 1	Setup	0.1387	0.3159
	Setup and Range	0.1582	1.4330
Patient 1 Plan 2	Setup	0.1222	0.5080
	Setup and Range	0.1320	2.9947
Patient 2 Plan 1	Setup	0.0793	0.3021
	Setup and Range	0.0950	1.4360

For all treatment plans both the relative and the absolute error of the approximation increase when range uncertainties are added to the model. This is expected since it is more difficult for the polynomial chaos method to predict the correct value when more input parameters are present. The polynomial chaos method is able to reproduce the expected dose distribution for all treatment plans with an absolute maximum error of 0.5 Gy in the presence of only setup uncertainties with just seven quadrature points.

The accuracy of the full dose approximation for patient 1 plan 1 is displayed graphically in Figure 4.5 again with the help of an error-volume histogram. The line represents how accurately the mean dose in each of the nonzero dose voxels is approximated by the polynomial chaos expansion. The accuracy is around 0.2 Gy for the vast majority of the voxels in the distribution. Again, adding range uncertainties to the system makes it more difficult for polynomial chaos to correctly predict the correct mean. In that case the accuracy for predicting the mean correctly in virtually all nonzero dose voxels drops to around 0.4 Gy.

Such an error volume histogram can be constructed not only for the full dose distribution, but also for the

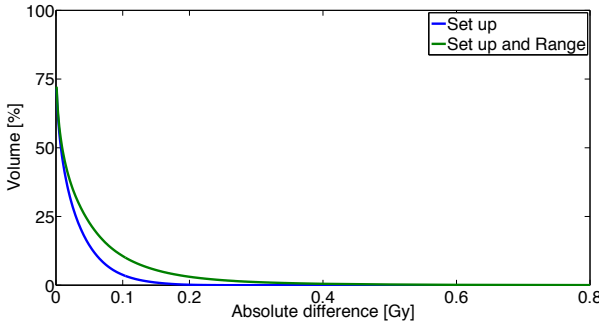
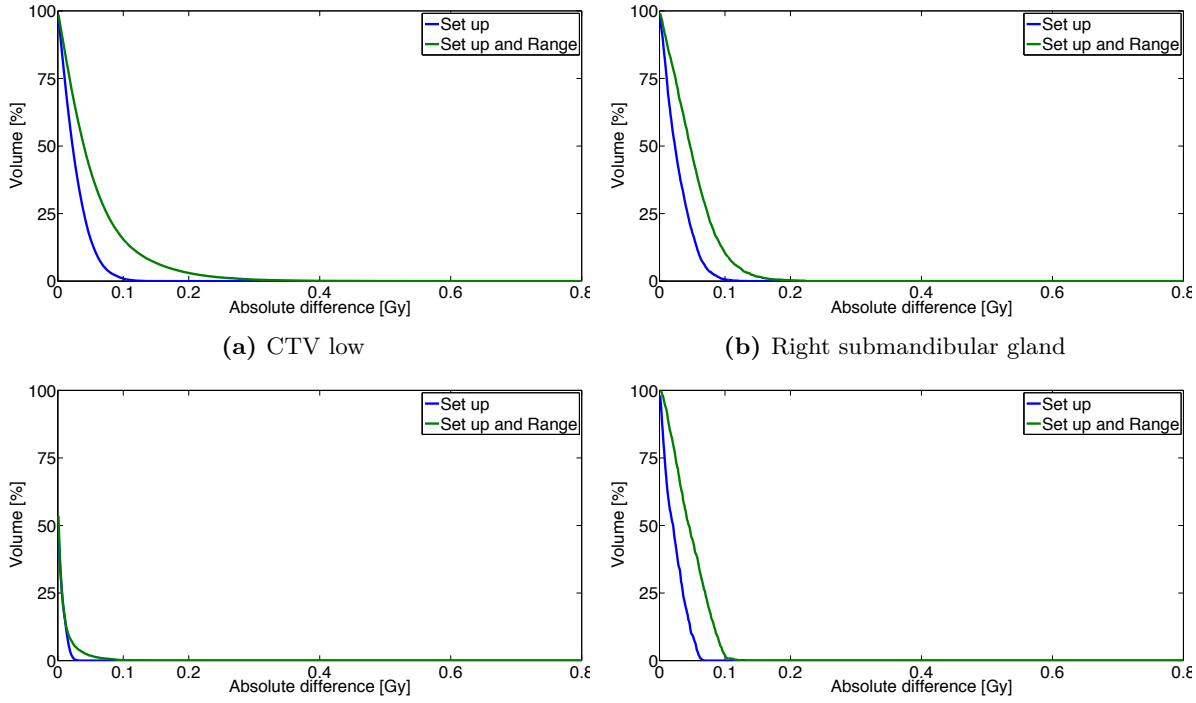


Figure 4.5: Error volume histogram as a function of the absolute difference between the reference expected full dose distribution and the polynomial chaos produced expected dose distribution. For setup errors (blue) the mean is predicted with an accuracy of 0.2 Gy for virtually all voxels. When range is included (green), this accuracy drops to 0.4 Gy. Only voxels with a nonzero dose are taken into account in the calculations.

organs at risk under consideration in this study. In Figure 4.6 a few representative cases are shown. The figures depicted here represent various cases in terms of organ size (number of voxels to be approximated) and received dose. The trend in the expected full dose distribution that the accuracy drops when range uncertainties are added to the model, is also present in all organs. Generally, the accuracy of the expected dose distribution in the separate organs is higher than that of the full distribution. A possible explanation for this is that more errors occur at the edges of the dose distribution, whereas most organs lie more in the center of the distribution.



(a) CTV low

(b) Right submandibular gland

(c) Spinal cord

(d) Esophagus

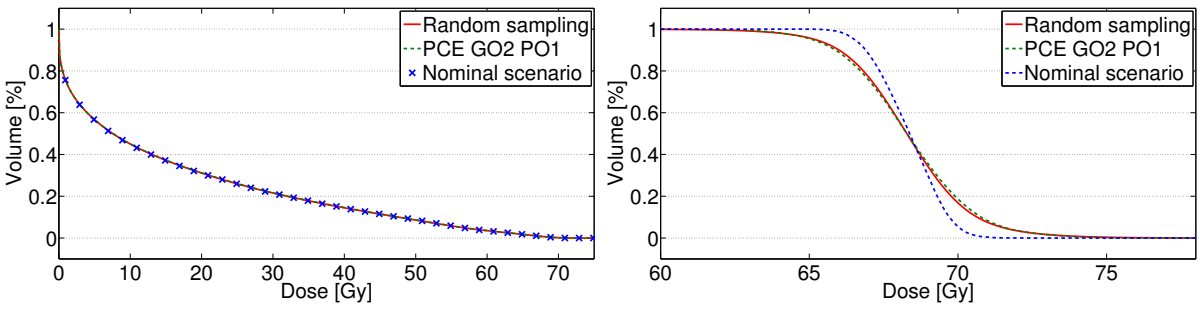
Figure 4.6: Error volume histogram for the expected dose in the separate organs as a function of the absolute difference between the reference and the polynomial chaos expansion. The graphs show similar results, regardless of the size of the organ (the esophagus is small and the CTV low is big) or the received dose (spinal cord receives a low dose and the submandibular gland receives a high dose). Only voxels with a nonzero dose are taken into account in the calculations.

4.3 Approximation of dose-volume histograms

Dose-volume histograms summarize the full dose distribution into convenient 2D graphs which give the physician a very quick overview of the delivered dose to the patient. During treatment planning a DVH for each organ is generated for the planned dose delivery scenario. In this section the approximation of the dose-volume histograms is investigated in more detail. Due to setup and range uncertainties the planned (nominal) dose distribution in general will not coincide with the expected dose distribution. This also means that the dose-volume histograms for the nominal scenario will not necessarily overlap with the dose-volume histograms of the expected dose. The expected dose distribution in the organs could already be approximated accurately with first order polynomials and second order grids as could be seen in Figure 4.6.

As explained in Section 3.3.2 a dose-volume histogram is constructed by counting the number of voxels that have a dose higher than the particular dose at each dose step, i.e. at every 0.1 Gy. The obtained value is divided by the total number of voxels in the organs to obtain a relative volume. For each input realization the values in all dose points are stored. With 10,000 (random sample) input realizations, each dose point has its own probability density function. For the construction of the reference dose-volume histogram of the expected dose the mean values in each dose are taken and combined in a single line. This reference expected dose-volume histogram is compared against the expected dose-volume histogram produced by direct route of the polynomial chaos method. For each dose point a separate expansion is calculated. The mean values at each of the points given by the first coefficient in the expansion are collected in a single line that forms the dose-volume histogram of the expected dose. They are compared on visual inspection of the produced figures. For the approximation of the dose-volume histogram for the expected dose again first order polynomials and second order grids proved sufficient and it is found unnecessary to use higher order polynomial chaos expansions. Including higher orders only marginally increases the accuracy of the approximation of the expected dose and are therefore not shown here.

In roughly half of the investigated organs the nominal DVH is very similar to the expected DVH, an example is illustrated in Figure 4.7a. This is especially true for low dose organs or for organs whose DVH is a smooth function. In the other half of the organs the nominal scenario does not correspond at all to the expected DVH, an example of this is depicted in Figure 4.7b. Often the nominal scenario incorrectly predicts a lower dose to parts of the organ, in some cases however an overestimation of the dose is predicted. This is the case in both the CTV low and the CTV high, leading to an unrealistic expectation of the dose delivered to the tumor.



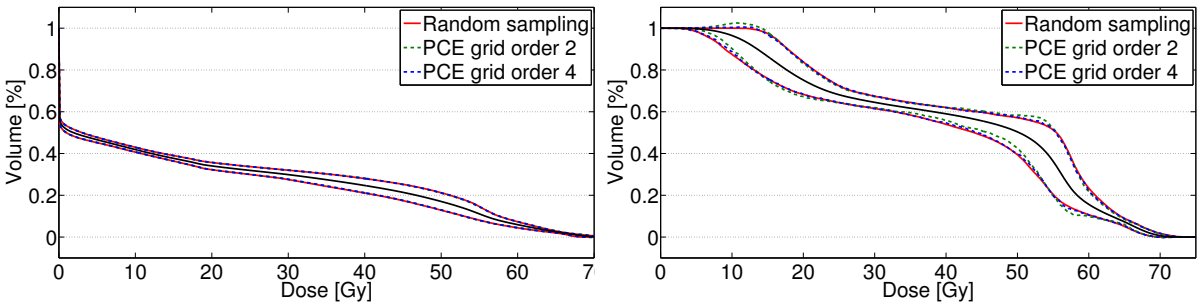
(a) Larynx **(b)** CTV high

Figure 4.7: Illustration of the difference between the nominal scenario and the expected dose distribution as predicted by random sampling and polynomial chaos expansion with grid order 2. For some organs, e.g. the larynx, the nominal scenario is an excellent approximation, where for other organs, e.g. the CTV high, the nominal scenario is not a good approximation of the expected dose distribution. Only setup uncertainties were taken into account in these figures.

For each of the organs a bandwidth is constructed around the expected DVH. This bandwidth allows for a more realistic view of possible scenarios that could happen during treatment. The reference bandwidth is constructed by taking the 5^{th} and 95^{th} percentile of the reference distribution at every data point (every 0.1 Gy). This reference bandwidth is compared against the bandwidth produced by the direct route of polynomial chaos (the indirect route will be treated in Section 4.6). In the direct route a separate

expansion is calculated at each data point. This expansion is resampled and from these resampled distributions (at every 0.1 Gy) the 5^{th} and 95^{th} percentile are taken and combined in the lower and upper limit of the bandwidth.

First order polynomials and second order grids were only for a part of the organs successful in accurately approximating the bandwidth, but for the majority of the organs higher order expansions were preferred to perfectly coincide with the reference bandwidth. The organs can roughly be divided into three equally sized groups based on the impact higher grid orders have on the accuracy of the predicted bandwidth. In Figure 4.8a an example of an organ is given for which higher order expansions only contribute marginally to the accuracy of the predicted bandwidth as it is already very precise. This group consists mainly of organs for which the nominal scenario was in very good agreement with the expected dose distribution. The second group consists of organs for which the majority of the bandwidth can be predicted quite accurately with a second order grid. Only small parts of the bandwidth deviate from the reference. For these parts higher order expansions are needed to improve the approximation. The last group consists of organs for which higher orders visibly improve the accuracy of the bandwidth approximation in multiple places, an example is depicted in Figure 4.8b. In Figure 4.8 a comparison between a first order polynomial, second order grid expansion and a third order polynomial, fourth order grid expansion is shown. The analysis was also done for a second order polynomial, third order grid, but this is not shown for visibility reasons.



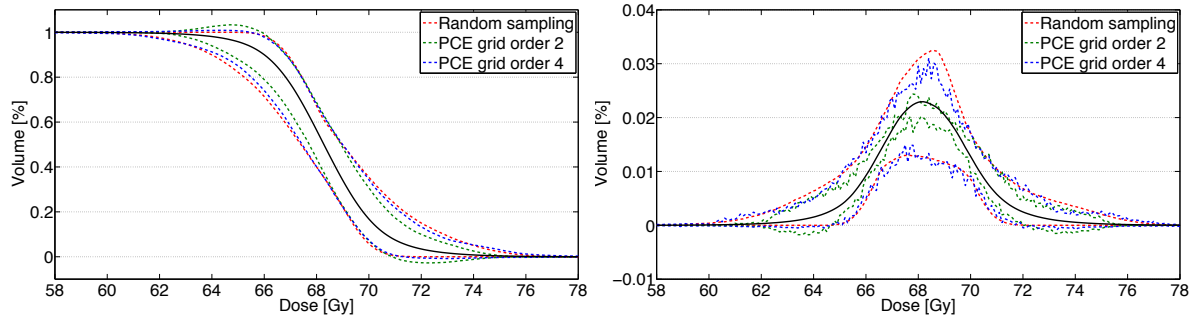
(a) Mandible **(b)** Left submandibular gland

Figure 4.8: Comparison between lower (GO 2, PO 1) and higher (GO 4, PO 3) grid orders in the approximation of the bandwidth of the DVH for different organs. For some of the organs low order expansions are enough to accurately predict the bandwidth of the DVH. For other organs higher order expansions are needed to perfectly overlap the reference bandwidth created by random sampling. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures.

The cumulative dose-volume histogram is very much preferred over the differential dose-volume histogram in the medical world. For the tumor, however it would also be of interest to take a look at the differential dose-volume histogram, in order to investigate if this could be approximated more accurately at lower order expansions. As depicted in Figure 4.9a, even higher order polynomial expansions are not capable of perfectly approximating the bandwidth for the CTV high. The construction of the differential DVH is done by determining for each dose interval of 0.1 Gy what percentage of the voxels has a dose corresponding to that interval. The reference bandwidths are constructed by looking at the distribution of the random input realizations and taking the 5^{th} and 95^{th} percentile. For polynomial chaos, each dose interval is approximated by a separate expansion, resampled and again the 5^{th} and 95^{th} percentile are taken.

The differential DVH for the CTV high is shown in Figure 4.9b. Here the impact of higher grid orders is even more present than in the cumulative DVH, as the lowest grid expansion only very roughly follows the intended shape, where the higher grid order is much more similar to the reference bandwidth. An explanation for this is found in the way the differential DVH is constructed. The distribution in each of the dose intervals are often very non-Gaussian, which cannot be predicted with first order polynomial chaos. The noisy character is because even adjacent points are found to have very dissimilar probability density functions.

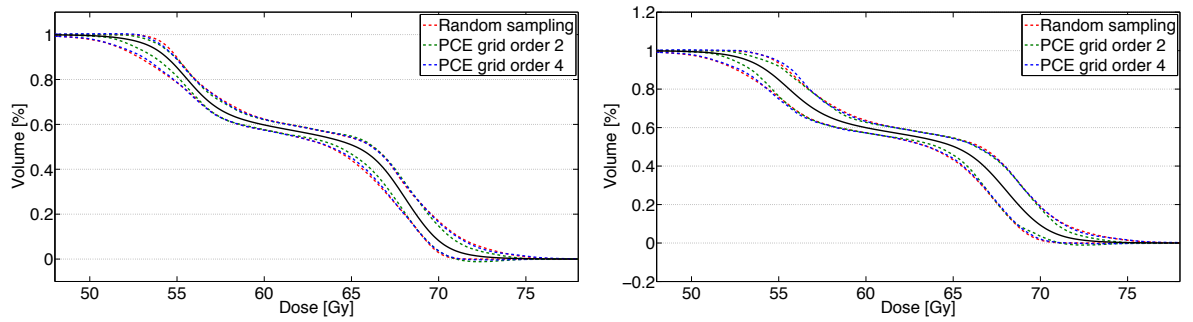
All cases above only take setup uncertainties into account. The inclusion of range in the models is found to have no significant effect on the DVH of the expected dose in the organs. For 7 out of the 18 researched



(a) Cumulative DVH **(b)** Differential DVH

Figure 4.9: Comparison between the cumulative and differential dose-volume histograms of the CTV high. In both figures the difference in grid order is clearly visible, as higher grid orders deliver a much higher accuracy. The differential DVH is not easier to approximate using polynomial chaos expansions, even though this was expected beforehand. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures.

organs adding range did not have an effect on the DVH at all. For the rest of the organs, the bandwidths broadened as expected. An example is given in Figure 4.10b. Low order expansions are not always accurate enough to describe the bandwidth, but higher order grids can again solve this problem.



(a) Only setup uncertainties **(b)** Both setup and range uncertainties

Figure 4.10: The inclusion of range uncertainty broadens the bandwidth of the dose-volume histogram for the majority of the organs. The performance of the polynomial chaos method is comparable to the model in which only setup uncertainties are included. The DVH of the expected dose is depicted as a solid black line.

4.4 Approximation of dose and tumor parameters

The list of dose and tumor parameters is the easiest way for a physician to determine whether a plan is acceptable or not. The evaluated parameters are the minimum, mean and maximum dose to 16 organs at risk and the tumor, divided in the CTV low and the CTV high. For the latter two, an additional four tumor parameters are evaluated, making it a total of 62 parameters. For some organs the minimum dose is always zero, these were taken out of the evaluation (even though polynomial chaos correctly predicts a zero mean value in these cases). The amount of parameters that were taken out of the evaluation in this way varies between 4 and 8 depending on the treatment plan and the inclusion of range uncertainties.

For each of the evaluated parameters a reference probability density function is determined from 10,000 random input realizations which is assumed to be described sufficiently by its mean and variance. The reference values are compared by results produced by the direct route of the polynomial chaos expansion (the indirect route will be treated in Section 4.6). Each of the parameters is directly chosen as a response of the system and individual polynomial chaos expansions are calculated for each of them. Unlike the dose-volume histogram, there is no need for resampling. The mean and variance are simply given by the

coefficients of the expansion as described in Section 2.4.1. The relative error in the mean and variance with respect to the reference values are calculated and compared to the criteria given in Section 3.5. For clarity reasons, the criteria are repeated in Table 4.4.

Table 4.4: Performance criteria for the dose and tumor parameters based on the absolute or relative difference between the reference value and the polynomial chaos expansion method.

	Dose parameters		Tumor parameters	
Good	Absolute difference	< 0.1 Gy	Relative difference	$< \pm 1.5\sigma$
	or Relative difference	$< \pm 1.5\sigma$		
Acceptable	Absolute difference	< 1.0 Gy	Relative difference	$< \pm 3\sigma$
	or Relative difference	$< \pm 3\sigma$	Relative difference	$< \pm 1\%$
	or Relative difference	$< \pm 1\%$		
Bad	None of the above		None of the above	

In Table 4.5 the performance of the polynomial chaos method is shown. As is expected, the accuracy with which parameters can be approximated by polynomial chaos expansion increases with increasing grid orders. Already 1/3 of the parameters is predicted correctly, both the mean and the variance, with only a first order polynomial and second order grid expansion. The performance goes up to over 60% with the inclusion of second order polynomials and third order grids (31 quadrature points). The inclusion of range uncertainty in the model slightly degrades the performance of the polynomial chaos method. This effect holds for all grid orders in both the mean and the variance. The trend that higher grid orders increase the performance of PCE obviously also holds when range errors are included.

Table 4.5: The performance of the polynomial chaos method in predicting the mean and variance for all dose and tumor parameters is compared for different grid orders for patient 1 plan 1 for only setup uncertainties.

	Grid order	Good [%]	Acceptable [%]	Bad [%]
Mean	2	72.2	22.2	5.6
	3	81.5	16.7	1.9
	4	88.9	9.3	1.9
Variance	2	33.3	46.3	38.9
	3	61.1	24.1	14.8
	4	70.4	22.2	7.4

The performance of the polynomial chaos method strongly depends on the set criteria. Although the criteria are partly based on theory, they are still arbitrary and might be too strict. This is especially the case for the tumor parameters, which are mostly classified as bad, as they are solely judged on relative criteria, where the absolute difference can sometimes be negligible. Unfortunately, no consensus on suitable criteria for accuracy is reached in radiotherapy.

One of the evaluated parameters, the minimum dose in the CTV low, was selected for further analysis. The choice for this particular parameter was chosen to illustrate the need for higher grid orders. In Figure 4.11a the reference probability density function is shown together with the distribution as calculated from the polynomial chaos expansion. The rapid fall-off in the dose requires higher grid orders for a more accurate approximation, an effect also seen in some of the dose distributions in the individual voxels in Section 4.2. The more Gaussian the distribution of the parameter is, the lower the grid order can be for a correct prediction of the pdf.

As can be seen in Table 4.5, the mean of a parameter can in a majority of the cases be approximated with only a few grid points. In Figures 4.11b it is clearly visible that the mean is within acceptable limits after as few as seven quadrature points. The value for the mean is predicted with an error of 0.7%. This can be compared to the random samples which require approximately 100 model evaluations before the mean can be described with that accuracy (i.e. before the confidence bandwidth indicated by $\pm\sigma$ reaches that value). Higher grid orders improve the mean even further, illustrated in Figure 4.11d and 4.11f.

Generally, the convergence of the variance is much slower, this holds both for random sampling and polynomial chaos expansions. Clearly, a second order grid expansion is by far not enough to give an accurate value for the variance for this particular parameter, depicted in Figure 4.11c. Including more quadrature points has a huge impact on the accuracy of the variance (Figure 4.11e). After the full third order polynomials, fourth order grid expansion the variance is predicted with an error of 6.1%, indicated by the arrow in Figure 4.11g. In comparison, the variance as calculated by the random sampling method with a similar number of model evaluations is over 20% off from the reference value for the variance.

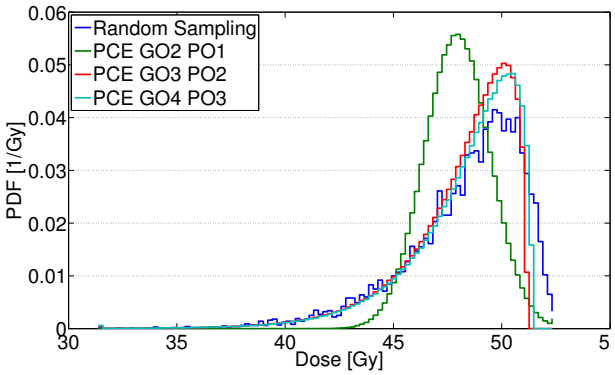
4.4.1 Dose parameters run adaptively

FANISP was originally intended to run adaptively to decrease computation time without decreasing the accuracy. The minimum, mean and maximum dose in every organ have all been calculated individually in an adaptive way. The algorithm was set with a tolerance level for the mean and variance of 0.001 and 0.01 respectively. The number of quadrature points needed for the algorithm to converge is then compared to the number of grid points needed for a full fourth order grid calculation (105 quadrature points). Results of this comparison are summarized in Table 4.6.

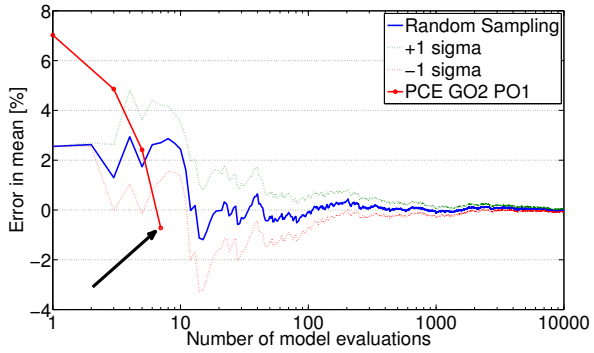
Table 4.6: Computational gain for the adaptive algorithm. The gain is expressed as the relative change in the number of needed grid points compared to a full fourth order grid consisting of 105 grid points. A dash indicates that that quantity was not calculated as the minimum dose was zero for all random sampling input realizations. Only setup uncertainties are taken into account.

		Minimum dose	Mean dose	Maximum dose
Tumor	'CTV high'	0%	-57%	0%
	'CTV low'	0%	-21%	0%
Parotid glands	Right	-21%	-70%	0%
	Left	-	0%	0%
Submandibular glands	Right	0%	-70%	0%
	Left	-	-57%	-21%
Spinal cord		-	-21%	0%
Brainstem		-	0%	0%
Larynx		-	-70%	-21%
Oral cavity		-	-70%	0%
Swallowing muscles	SCM	0%	-63%	0%
	MCM	0%	-70%	0%
	MCI	-	-70%	0%
	MCP	-21%	-70%	-63%
	Esophagus	-21%	-70%	-63%
Other	Mandible	-	-70%	-21%
	RP Right	0%	-63%	0%
	RP Left	0%	-70%	-21%

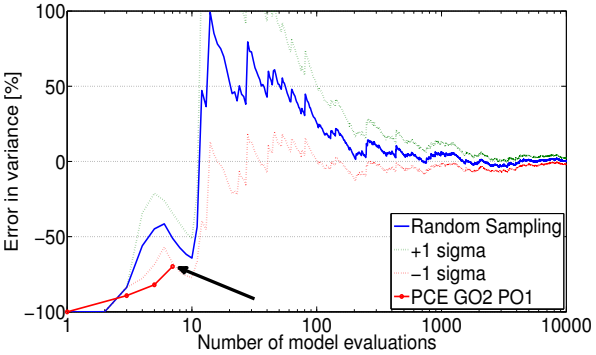
It can be concluded that the mean dose in any of the organs can benefit most from adaptivity in terms of decreased computation time. This is because the mean dose is the average of sometimes tens of thousands of voxels, whereas the minimum dose or maximum dose can depend on a few voxels with a very irregular dose probability density function. The computational benefit of the adaptivity strongly depends on the values given to the tolerance levels for the mean and the variance. A higher tolerance level can increase the computational gain even further, although this decreases the accuracy of the obtained results. The discussion on the optimum between computational gain and accuracy is not considered in this thesis.



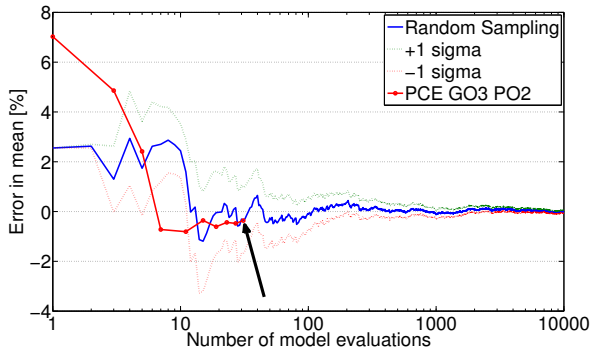
(a) PDF for different orders



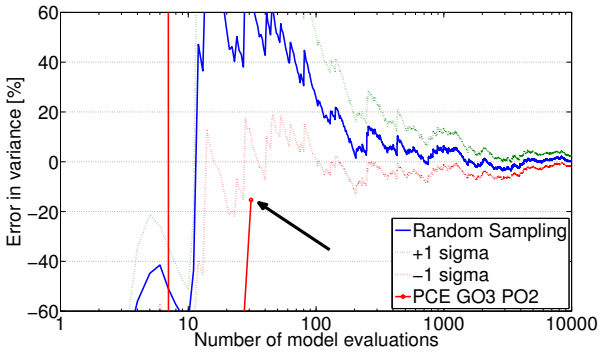
(b) Convergence mean, grid order 2



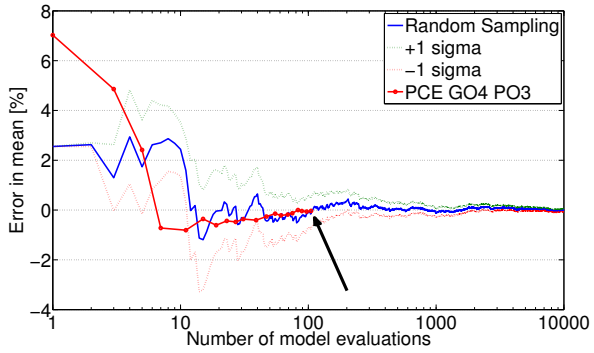
(c) Convergence variance, grid order 2



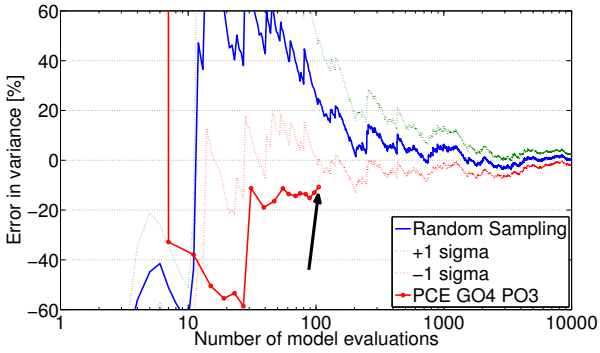
(d) Convergence mean, grid order 3



(e) Convergence variance, grid order 3



(f) Convergence mean, grid order 4



(g) Convergence variance, grid order 4

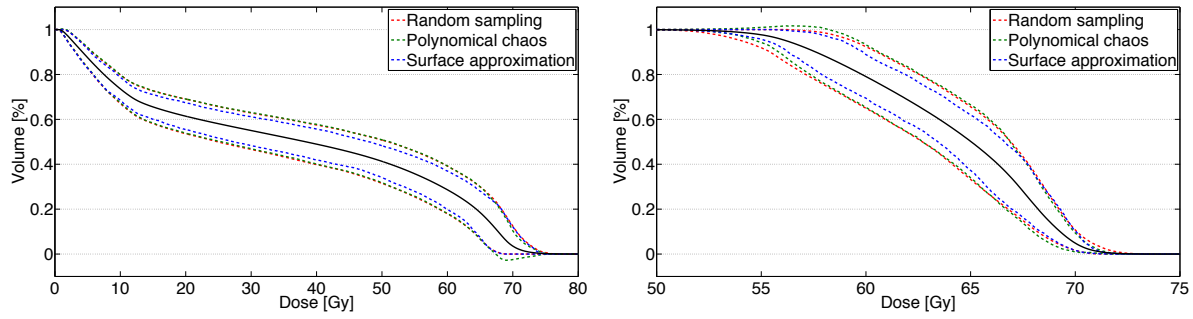
Figure 4.11: Illustration of the accuracy of different grid and polynomial orders for the minimum dose in the CTV low of patient 1 plan 1. Depicted are the the approximation of the probability density function, the convergence of the mean and the convergence of the variance. The variance especially can require higher grid orders to reach acceptable values. Second order grids contain 7 quadrature points, fourth order grids 105 quadrature points. Only setup uncertainties are taken into account.

4.5 Comparison with the surface approximation

The treatment plan is now fully evaluated. At this point it is also possible to make a comparison between polynomial chaos expansions and the surface approximation. In the case of setup uncertainties the surface approximation requires the calculation of seven error scenarios, which is the same amount of calculations required for a first order polynomial and second order grid expansion. When range uncertainties are also included, the number of calculations needed increases to nine. It was decided to only compare the surface approximation to first order polynomial expansions as they require the same number of calculations. Higher order expansions are therefore not included in this analysis. For the surface approximation it is unfortunately not possible to define an expected dose distribution or expected dose parameters. It is however possible to make a qualitative comparison for the dose-volume histograms.

For the construction of the lower and upper limit of the bandwidth the minimum and maximum values produced by the dose distributions of the surface approximation are taken at each dose point. Because the set of inputs used for the surface approximation mark an 85% confidence interval, the polynomial chaos bandwidth is constructed from the 7.5^{th} and 92.5^{th} percentiles, which also makes an 85% confidence bandwidth. Note that this is different from the standard 95% confidence bandwidth used earlier in this chapter.

In the majority of the organs the surface approximation underestimates the bandwidth, as illustrated in Figure 4.12. For several organs the bandwidths perfectly overlap, and in three organs the surface approximation provided a more accurate bandwidth than the polynomial approximation. In these cases higher order expansions would solve the accuracy issues for the polynomial chaos method. It is also possible to add more points in the surface approximation, but there is no recipe available for how many points would have to be added or where those points should lie on the error-sphere.



(a) Right parotid gland **(b)** Right submandibular gland

Figure 4.12: Qualitative comparison between a seven point polynomial chaos expansion (first order polynomial, second order grid) and a seven point surface approximation. The surface approximation is found to underestimate the width of the bandwidth in the majority of the researched organs. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures.

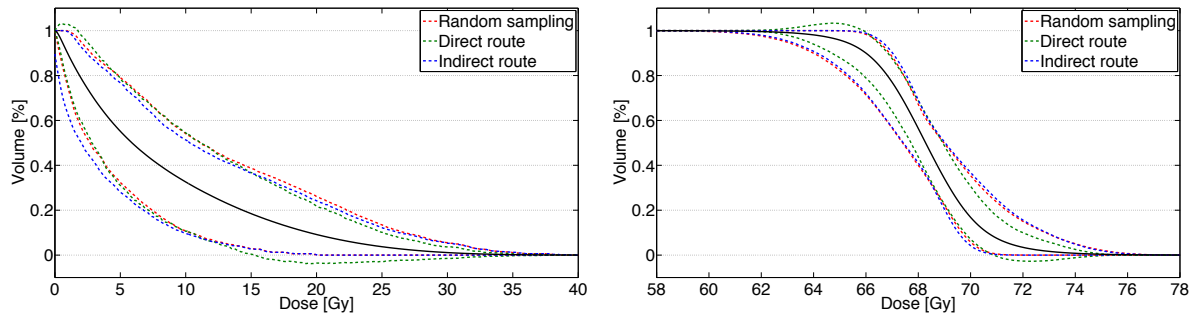
The surface approximation and a first order polynomial expansion both require the same number of calculations. With the same number of error scenarios, the polynomial chaos method delivers much more information as the summary below shows.

- PCE can produce any desired confidence interval for the bandwidth, the surface approximation produces only the 85% confidence interval
- PCE can produce an accurate expected cumulative dose-volume histogram for all organs
- PCE can produce the expected mean dose for the full dose distribution
- PCE can produce expected values for any desired dose or tumor parameter

4.6 The indirect route versus the direct route

For this particular application of the polynomial chaos method, it is possible to calculate some quantities both via a direct and an indirect route. In the indirect route a polynomial chaos expansion is made for every voxel inside the patient. From this expansion it is possible to derive any relevant dose and tumor parameter and all dose points needed for the construction of (the bandwidths of) the dose-volume histograms. Relevant polynomial expansions of the full dose distribution have to be resampled to obtain desired quantities. In the direct route the dose and tumor parameters and data points to be investigated have to be chosen beforehand. These parameters and dose points are taken as responses of the system and for each of them a separate polynomial chaos expansion is calculated simultaneously. The indirect route is much more computationally intensive for two reasons. First, the processing of subgrids takes substantially longer than in the direct route and second, the matrices within FANISP become so large, that Matlab runs out of memory before it is able to process higher order grids. The indirect route is therefore limited to first order polynomials and second order grids. The direct route does not have these computational limitations and can easily increase grid and polynomial order to obtain a higher accuracy. A computational compromise would therefore be to calculate only the expected dose of the full dose distribution via the indirect route and to calculate all dose and tumor parameters and data points for the dose-volume histograms via the direct route. However the problem with this compromise is that the values obtained from both routes do not necessarily coincide. Using them side by side could lead to an inconsistent representation in the sensitivity and uncertainty analysis of a treatment plan if the differences between them become clinically relevant. In this section the errors the two routes produce are compared for the dose-volume histograms and the parameters. In order to make a fair comparison, the direct route is in this section always calculated with first order polynomials and second order grids.

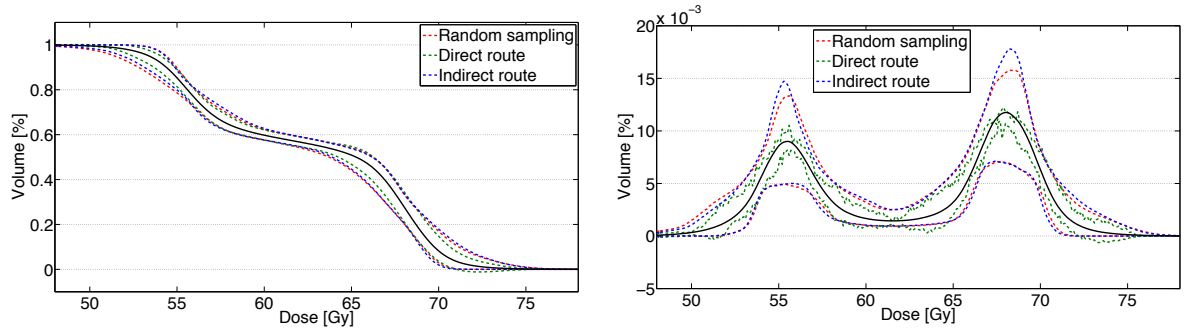
The dose-volume histogram for the expected dose is identical for both routes in all organs as they both overlap perfectly with the reference value. The bandwidths produced for half of the investigated organs are identical in both routes. This is not surprising since for these same organs first order polynomials and second order grid expansions proved sufficient in the approximation of the bandwidth. For the other half of the studied organs it was already known that these low order expansions not always give an accurate bandwidth and for these cases the bandwidth generated by the indirect route are slightly better. Often this slightly worse approximation by the direct route is because it is possible in the direct route to produce unphysical values for the volume. An example for this happens around 0 Gy in Figure 4.13a and 64 Gy in Figure 4.13b where the expansion predicts a volume larger than 100%. The opposite can also happen, when a volume less than 0% is predicted around 15 Gy and 71 Gy In Figure 4.13a and 4.13b respectively. The reason that this is possible is because the algorithm does not take underlying physics into account. For each of the data points the corresponding polynomials and coefficients are calculated, which are not restricted by a physical volume. In the indirect route this is never a problem, as the dose-volume histogram is derived from a sampled full dose distribution for that organ. For the indirect route it would however be possible that the DVH extends to the left, as some voxels may be approximated by negative values.



(a) Esophagus **(b)** CTV high

Figure 4.13: Illustrations of the difference in the DVH constructed via the direct or the indirect route. Around 1/3 of the DVHs are slightly better approximated via the indirect route, the others are very similar. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures. The expansions were both calculated with first order polynomials and second order grids.

In Section 4.3 a comparison was made between the cumulative and the differential dose-volume histogram, both obtained via the direct route. The approximation of the bandwidths in the differential DVH with first order polynomials was not sufficient. The indirect route however is much more able to approximate the bandwidths. The difference is depicted in Figure 4.14b. This was not expected from the cumulative DVH in Figure 4.14a, where the bandwidths produced by both routes are similar. A possible explanation for the large difference lies in the construction of the bandwidths. As mentioned in Section 4.3, the direct route is unable to reproduce the bandwidth accurately because the distributions of the data points are often non-Gaussian. The indirect route calculates for each dose distribution a differential dose-volume histograms, which is a smoother function. All the calculated differential DVHs are then combined and at each dose point the wanted percentile is calculated. This is bound to give a smoother bandwidth and a more accurate approximation.



(a) Cumulative DVH

(b) Differential DVH

Figure 4.14: Comparison between the indirect and direct route for the two types of dose-volume histograms for the CTV low. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures. The expansions were both calculated with first order polynomials and second order grids.

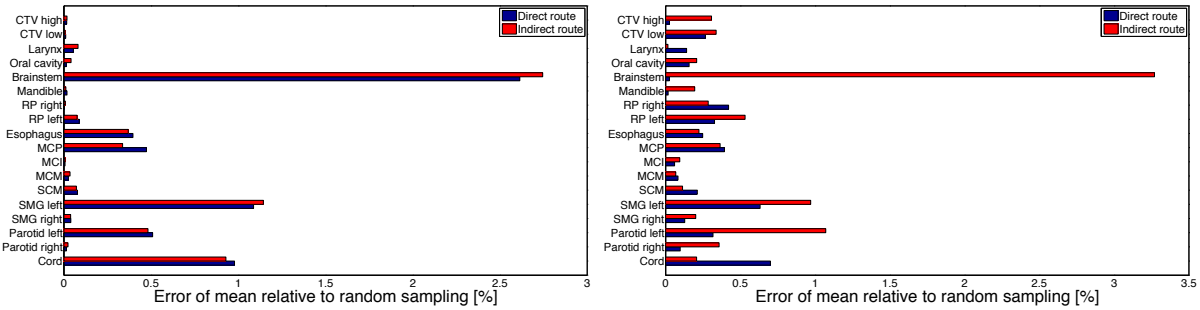
Aside from the dose-volume histograms, a comparison is made between the values obtained via the two routes for all dose and tumor parameters. This is done in terms of the performance of the polynomial chaos method evaluated by the set criteria as listed in Table 4.4. The dose and tumor parameters in all three treatment plans are scored according to these criteria and compared in Table 4.7. In all treatment plans the direct route scores the highest percentage of parameters classified as 'good' in both the mean and the variance. For the indirect route, the problems for correctly predicting the mean of a parameters arise either in the minimum dose or maximum dose, the mean dose is always predicted correctly. Advantage for the direct method is that these scores can easily go up when higher grid orders are included in the expansion, something that is not possible for the indirect route because of computational limitations in Matlab.

In Figure 4.15 the two routes are further compared via the absolute relative errors the expansions produce in the mean and maximum doses in the various organs. The differences in the mean are very small. This makes sense as this is the easiest parameter of an organ to predict. The difference in the maximum dose is more varied. Generally, the direct route approximates the maximum dose more accurately, with only a few exceptions. In principle, Figure 4.15 could also be constructed for the minimum doses. However, for 1/3 of the organs, the direct route correctly predicts a zero mean for the minimum dose and it is not possible to calculate a relative difference with the reference. For these cases the indirect route gives a negative value for the minimum dose. Negative dose values are of course unphysical, but again the algorithm never takes underlying physics into account.

For the tumor parameters the indirect route performs generally much better than the direct route in terms of the relative error, depicted in Figure 4.16. In the majority of the tumor parameters, the indirect route has a much smaller relative error compared with the direct route. The absolute values are often very similar to the reference results, but because for the variance typical values are in the order of 10^{-4}, the relative error is large. Seeing that the uncertainty in these parameters is thus already very small, the 'poorly performing' tumor parameters are no cause for concern.

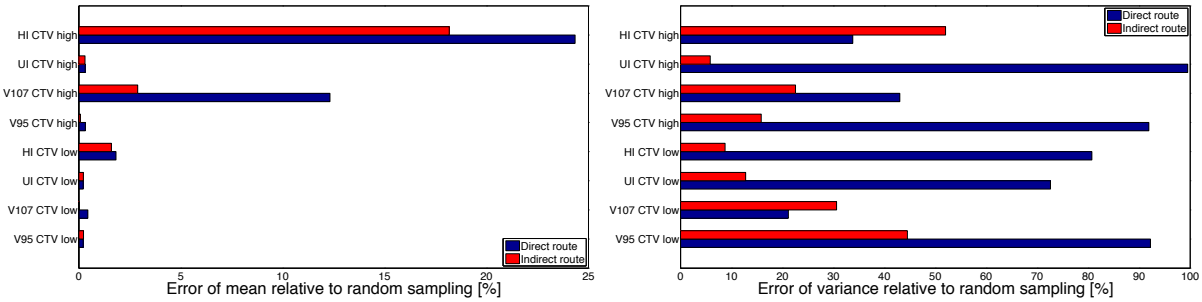
Table 4.7: Difference in performance of the polynomial chaos method for parameters calculated via the direct or indirect route. In all plans directly calculated parameters score a higher percentage of correctly predicted parameters. Only setup uncertainties were considered in this analysis. The expansions were both calculated with first order polynomials and second order grids.

			Good [%]	Acceptable [%]	Bad [%]
Patient 1	Mean	Direct	72.2	22.2	5.6
Plan 1		Indirect	61.1	33.3	5.6
	Variance	Direct	33.3	27.8	38.9
		Indirect	29.6	40.7	29.6
Patient 1	Mean	Direct	73.6	22.6	3.8
Plan 2		Indirect	62.3	34.0	3.8
	Variance	Direct	45.3	26.4	28.3
		Indirect	22.6	43.4	34.0
Patient 1	Mean	Direct	75.9	19.0	5.2
Plan 2		Indirect	53.4	41.4	5.2
	Variance	Direct	41.4	31.0	27.6
		Indirect	32.8	27.6	39.7



(a) Mean doses **(b)** Maximum doses

Figure 4.15: Comparison in absolute relative errors between the direct and indirect route for the mean and maximum doses in the organs. Reference is the value obtained from random sampling. Only setup errors were considered in this analysis. Both expansions were calculated with first order polynomials and second order grids.



(a) Mean **(b)** Variance

Figure 4.16: Comparison in the absolute relative errors made between the direct and the indirect route in the tumor parameters. Reference is the value for random sampling. Only setup errors were considered in this analysis.

A few concluding remarks on the differences between the direct and indirect route must be given. The differences in performance for the parameters are generally quite small. Exception to this is the approximation of the differential dose-volume histogram, which is superior via the indirect route. The differential DVH however is a tool that is hardly used, thus the question remains what this superior approximation is worth. Both routes have the drawback that the system is not able to include information on the

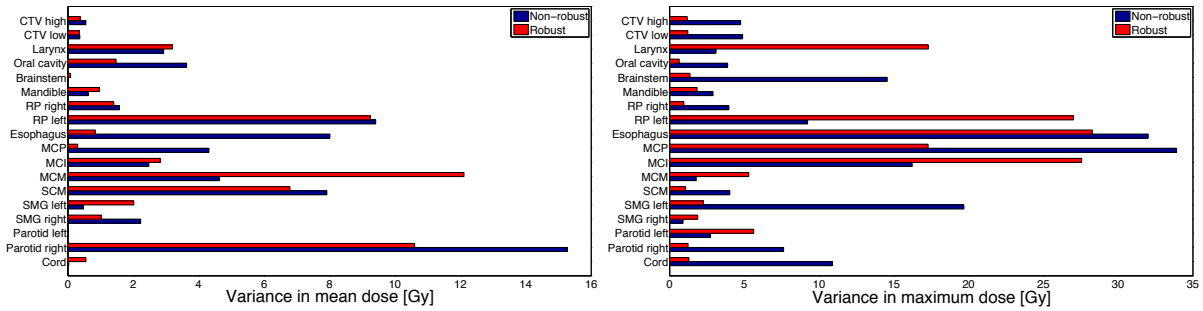
underlying physics in the expansion, leading to negative doses (minimum doses via the indirect route) or unrealistic volumes (DVH bandwidth via the direct route). This can however be fixed during the post-processing of the results and should be no reason for concern. Finally, the direct route has as advantage that it is much easier to include higher order grids in the expansion, which makes the parameters much more accurate. This is not possible for the indirect route because of computational limitations in Matlab.

4.7 Robust versus non-robust treatment planning

For patient 1 two plans were created, a non-robust plan (plan 1) and a robust plan (plan 2). This gives the opportunity to compare the influence robust treatment planning has on the absolute values of the variance in both plans and the effect on the performance of the polynomial chaos expansion. This analysis focuses solely on the dose and tumor parameters in the plan. Unfortunately the treatment plans were not created with the same beam directions, so a true comparison is not possible and this also becomes apparent in this analysis. In the robust plan the beam directions are evenly spaced, whereas for the non-robust plan the beam directions are all on the left side of the patient.

Beforehand it was expected that the parameters in the robust treatment plan would have a lower variance than the same parameters in the non-robust treatment plan. This would then indicate a lower sensitivity to uncertainties. In Figure 4.17a the absolute values of the variance in the mean doses for each organ for both the robust and the non-robust plan are compared. The salivary glands located near the tumor and several organs involved in the swallowing mechanism have a reduced variance in the mean in the robust plan when compared to the non-robust plan. The more evenly spread beam directions in the robust plan significantly lower the mean dose in these organs. The opposite is true for the mean dose in the spinal cord and the left submandibular gland which receive a much higher mean dose in the robust plan compared to the non-robust plan. The MCM and oral cavity however receive similar doses in both plans and the deviating values in the variance of the mean remain unexplained. The variance in the mean dose delivered to the tumor is the same in both plans.

The same analysis was repeated for the variances of the maximum dose delivered to the organs and depicted in Figure 4.17b. The first thing to notice is that both the CTV high and the CTV low have a strong reduction in the variance for the robust plan. This indicates that the maximum dose delivered to the tumor is less sensitive to uncertainties. In the majority of the organs the variance in the maximum dose is reduced in the robust plan when compared to the non-robust plan. There is no relationship found between higher maximum dose and subsequent higher variances or vice versa. It is possible that because of the changed beam directions parts of the organs have much higher variability in the dose they receive.



(a) Mean **(b)** Max

Figure 4.17: Comparison between a robust and a non-robust plan in terms of the absolute values of the variances in the mean and maximum doses for all organs for patient 1. Only setup errors were considered in this analysis.

For both plans all parameters are judged on how well they approximate the reference value, based on the same criteria explained in Section 4.4. The performance of the polynomial chaos expansions are compared in Table 4.8. For all grid orders the polynomial chaos performs better for the robust treatment plan. This holds both for the mean and the variance for the different parameters. This trend also holds when range errors are added to the model.

Table 4.8: Comparison between the performance of the polynomial chaos method of a robust and non-robust treatment plan. For all grid orders the robust plan score slightly better. Only setup uncertainties are taken into account in this analysis.

			Good [%]	Acceptable [%]	Bad [%]
Grid order 2	Mean	Non-Robust	72.2	22.2	5.6
Polynomial order 1		Robust	75.4	21.1	3.5
	Variance	Non-robust	33.3	46.3	38.9
		Robust	49.1	24.6	26.3
Grid order 3	Mean	Non-robust	81.5	16.7	1.9
Polynomial order 2		Robust	82.5	15.8	1.8
	Variance	Non-robust	61.1	24.1	14.8
		Robust	64.9	28.1	7.0
Grid order 4	Mean	Non-robust	88.9	9.3	1.9
Polynomial order 3		Robust	93.0	5.3	1.8
	Variance	Non-robust	70.4	22.2	47.4
		Robust	82.5	10.5	7.0

4.8 Variance decomposition

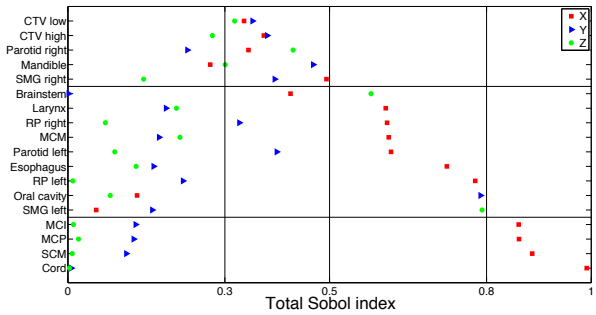
Variance decomposition is a feature of polynomial chaos which can not be easily executed with Monte Carlo methods. In variance decomposition the variance in a parameter is attributed to uncertainties in the individual input directions. In our case the individual inputs are the x-, y- and z-direction and in the case of range uncertainties the range is the fourth input. If it is known beforehand what the major cause of the uncertainty is, it would be possible to put extra focus on outlining the patient correctly in that particular direction. For each of the dose and tumor parameters the total Sobol' indices are calculated from their polynomial chaos expansions. These expansions have been calculated with third order polynomials and fourth order grids. The total Sobol' indices can be classified in one of four categories, as introduced in Section 2.4.2, and are repeated here.

- *very important* if $0.8 <$ TSI
- *important* if $0.5 <$ TSI < 0.8
- *unimportant* if $0.3 <$ TSI < 0.5
- *irrelevant* if TSI < 0.3.

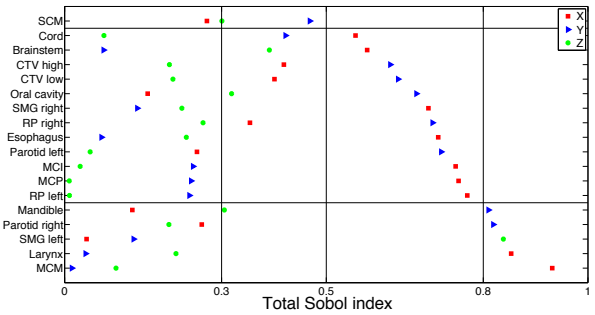
The boundaries of these classifications are marked by a vertical solid line in Figure 4.18. The organs are then subsequently grouped depending on the classification of the input variable with the highest total Sobol' index, i.e. all organs with one very important input parameter are grouped, all organs with one important parameter are grouped, etc.

A few interesting cases have been selected to present in this section. The reference case in this analysis of variance decomposition are the total Sobol' indices for the mean doses for patient 1 plan 2, depicted in Figure 4.18a. In the majority of the organs one of the input variables is considerably more important than the other input parameters, which for this patient in most cases is the x-coordinate. It can be seen that in the tumor all of the input parameters are roughly equally important. The sum of the total Sobol' indices is quite close to unity for all of the organs. This suggests that the first order sensitivity indices constitute the major part of the total Sobol' index, and that mixed influences are negligible for most organs.

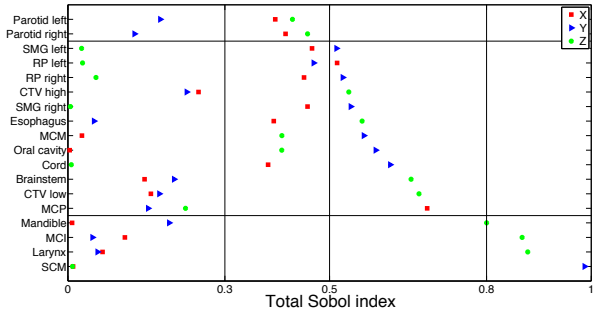
The first case is depicted in Figure 4.18b. For the same patient, the same analysis is done but now for the maximum doses in each of the organs. Figure 4.18b might seem similar to Figure 4.18a, but a closer look reveals that the order in which the organs appear on the right side of the figure has changed. For example, the variance in the mean dose in the spinal cord comes solely from the x-coordinate, but the influence of the x-coordinate in the variance of the maximum dose in the spinal cord is decreased to roughly half that



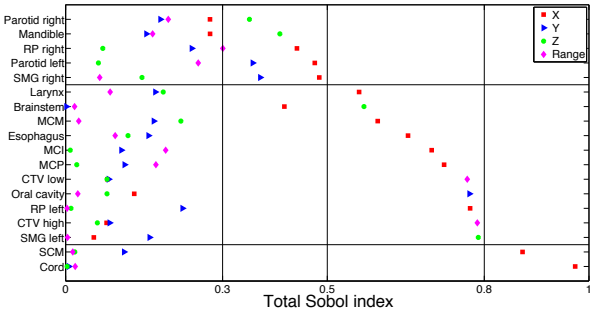
(a) Total Sobol' indices for the mean doses in patient 1 plan 2, used as a reference case.



(b) Total Sobol' indices for the maximum doses in patient 1 plan 2.



(c) Total Sobol' indices for the mean doses in patient 2.



(d) Total Sobol' indices for the mean doses in patient 1 plan 2, with range errors included.

Figure 4.18: Analysis of the total Sobol' indices for various cases. The reference case used are the total Sobol' indices for patient 1 plan 2.

value. The variance in the mean of the right parotid gland is more or less equally distributed over the three input parameters, with the y-coordinate being the least important. In the maximum dose of the same organ the y-coordinate determines over 80% of the variance. Generally it can be concluded that the Sobol' indices for a dose parameter in one specific organ are not representative for other dose parameters in that organ.

The second case is the analysis of the Sobol' indices for patient 2 in Figure 4.18c. The comparison is made between the mean dose for all organs with only setup errors included. What can be noticed immediately is that the importance of the x-coordinate in the variance of patient 1 is not present in the variance of the organs in patient 2. The most logical explanation for this can be found in the tumor shape. Patient 2 has a bilateral tumor, which is more or less symmetric in the x-direction and thus displacements in this direction are bound to have less effect. Patient 1 has a unilateral tumor, which can be expected to be much more sensitive to movements in the x-direction.

The final case is highlighted in Figure 4.18d. For this analysis the reference case is compared with the total Sobol' indices when range errors are also included in the model. The general order in which organs appear on the left side seems more or less unchanged. For all investigated organs, the influence of the range error on the variance of the mean doses is irrelevant. An important exception is the tumor. Both for the CTV low and the CTV high the range input parameter is the most important parameter in terms of the variance, where the setup uncertainties are considered irrelevant. Analysis of the tumor parameters shows that this influence of the range parameter is only apparent in V_{107}, both for the CTV low and the CTV high. The same trend is visible in the tumor parameters of patient 2, but is surprisingly much less apparent in plan 1 for patient 1. It is possible that robust treatment planning has an influence on this.

The total Sobol' indices in this section are all calculated with third order polynomials and a fourth order grid. Obviously, the higher the order of the expansion, the more accurate the parameter is approximated, and thus the more accurate the Sobol' indices are assumed to be. During the analysis of the Sobol' indices it was found that the total Sobol' index for any of the inputs is mostly built up from the first

order sensitivity Sobol' index. Only in a few dose parameters mixed influences of multiple inputs have any effect. This means that for a quick approximation of the total Sobol 'indices it would possibly be enough to only calculate the first order Sobol' indices. These could also be calculated when the parameters are approximated by first order polynomials and a second order grid, which would save computation time.

For each of the parameters the first order Sobol' indices were calculated with a first order polynomial expansion. For three different treatment plans it is evaluated how much the classification of the input parameters changes compared to the higher grid order Sobol' indices. For approximately 50% of the parameters the classification of all input parameters stays the same, for 25% one of the input parameters is classified differently and for the other 25% two or more input parameters are classified differently. These changes in classification are often caused by the fact that Sobol' indices are close to the boundary. Generally, the first order Sobol' indices derived from first order polynomial expansions give a reasonable approximation for the classification of more accurately calculated total Sobol' indices.

Chapter 5

Conclusions and Recommendations

The objective of this research is to investigate the applicability of polynomial chaos methods to sensitivity and uncertainty analysis for proton therapy treatment plans. In this chapter the conclusions of the research are summarized and some recommendations for further research are presented.

5.1 Conclusions

The polynomial chaos method was proven to be able to accurately produce a full evaluation of the sensitivity and uncertainty analysis of proton therapy treatment plans. The expected dose distribution can be approximated with an accuracy of 0.2 Gy in the presence of setup uncertainties. Dose-volume histograms for 18 investigated structures, including the CTV high and low and all important organs at risk, can be perfectly reproduced and it is possible to construct any wanted confidence intervals around the expected dose with high accuracy. Higher order expansions or post-processing of these bandwidth are necessary in the case of half of the organs to avoid unphysical results. The uncertainty distribution of over 50 researched dose and tumor parameters were found to be predicted within acceptable limits for a second order grid expansion. The performance is judged on criteria set by the author, which for medical purposes might be considered too strict or not strict enough. The discussion on what criteria are considered the most appropriate is outside the scope of this thesis project. Finally, with the polynomial chaos the variance in the parameters can be attributed to the uncertainty in the individual input parameters, a process that is too computationally intensive for sampling based methods.

The inclusion of range uncertainties slightly decreases the performance of the polynomial chaos method. The range uncertainty only has a significant influence on the variance of the tumor, for the organs this influence is negligible. The conclusions drawn from the proof of concept study were found to hold both for robustly and non-robustly generated treatment plans as well as for a second patient case with a completely different shaped tumor.

Tumor and dose parameters and dose-volume histograms can be calculated directly or indirectly for this particular application of the polynomial chaos method. There are no visible differences for the expected dose in the dose-volume histogram produced by the two routes. The indirect route reproduces the bandwidths for the dose-volume histograms generally better than the direct route, as the latter can produce unphysical results. The prediction of the bandwidths for the differential DVH is dramatically better for the indirect route. The direct route performs better on the parameters, where the indirect route sometimes produces unphysical results. Any unphysical result for both routes can easily be fixed during post-processing of the results and should be no reason for concern. The direct route has as largest advantage that it is very easy to include higher order expansions for an even higher accuracy, whereas the indirect route is computationally limited to first order polynomials. It can be concluded from this research that the suggested compromise, i.e. approximate the full dose via the indirect route and approximate the parameters and dose-volume histograms via the direct, delivers no clinically relevant inconsistencies

in the representation of the sensitivity and uncertainty analysis of the treatment plan.

5.2 Recommendations for further research

Further research for the application of the polynomial chaos method can continue in two directions. The first direction is to use the method as applied in this research but to extend the model with one of the following options:

- The errors modeled in this research only take into account the position (x-, y- and z-component) and range uncertainties. A third type of uncertainty that could be considered are rotational errors. For this type of research the dose recalculation algorithm would need to be expanded to make calculations possible.
- In this research the planning CTs are taken as a fixed reference for the patient and the modeled uncertainties in this research then represent only the random fluctuation of a patient's position in time. As the CT is only a snap-shot in time, a systematic error occurs as the result of uncertainties in the planning CT. This systematic error is fixed for a specific patient, but for a population of people receiving this type of treatment, this systematic error can be assumed to also have an uncertainty distribution. The impact of the systematic error for a specific patient or for a population of patients, could be investigated by means of the polynomial chaos method.
- Patients receive their treatment in fractions. The dose delivered in each fraction is assumed constant, but can in fact also be subjected to uncertainties. The fraction dose can be modeled as an additional uncertain input variable and the effect of fractionation on the total dose can be researched.
- Sensitivity and uncertainty analysis for proton therapy treatment plans could potentially benefit from adaptivity. Mean doses in the researched organs can often be correctly predicted by a first order polynomial, the minimum and maximum dose can require higher order polynomials to be predicted more accurately. For the use of adaptivity, a specific response of the system needs to be chosen to steer the algorithm in adaptively building up the grids. With the 50+ researched parameters in the treatment plans, it is not possible to prefer one specific parameter to focus on. Further research could be done how the combination of all these parameters could be calculated adaptively in a clever way.

The second direction of research involves the investigation of the implementation of the polynomial chaos method into the optimization algorithm. Currently robust treatment planning is done by means of the surface approximation, which involves the calculation of several error scenarios. The same number of error scenarios is needed for the calculation of a full first order polynomial chaos expansion, which proved to give overwhelmingly more information than the surface approximation. In this way information on the sensitivity to uncertainties of the plan could already be taken into account in the optimization, possibly leading to different choices in the optimization scheme.

Bibliography

[1] Integraal Kankercentrum Nederland. Cijfers over kanker.

[2] Centraal Bureau voor de Statistiek. Cbs doodsoorzakenstatistiek.

[3] F. Bray, J.S. Ren, E. Masuyer, and J. Ferlay. Estimates of global cancer prevalence for 27 sites in the adult population in 2008. *International Journal of Cancer*, 132:1133–1145, 2013.

[4] R. Misra and M. Waalkes. Genetic toxicology. *Encyclopedia of Occupational Health and Safety*, 2011.

[5] The Children's Hospital of Philadelphia.

[6] W.P. Levin, H. Kooy, J.S. Loeffler, and T.F. DeLaney. Proton beam therapy. *British Journal of Cancer*, 93:849–854, 2005.

[7] Health Council of the Netherlands. Proton radiotherapy Horizon scanning report. 2009.

[8] F. Albertini, E.B. Hug, and A.J. Lomax. Is it necessary to plan with safety margins for actively scanned proton therapy? *Physics in Medicine and Biology*, 56:4399–4413, 2011.

[9] H. Jiang, J. Seco, and H. Paganetti. Effects of Hounsfield number conversion on CT based proton Monte Carlo dose calculations. *Medical Physics*, 34:1439–1449, 2007.

[10] D.G. Cacuci and M. Ionescu-Bujor. A Comparative Review of Sensitivty and Uncertainty Analysis of Large-Scale Systems - II: Statistical Methods. *Nuclear Science and Engineering*, 147:204–217, 2004.

[11] A. Bolsi, A.J. Lomax, E. Pedroni, G. Goitein, and E.B. Hug. Experiences at the Paul Scherrer Institute with a remote patient positioning procedure for high-throughput proton radiation therapy. *International Journal of Radiation Oncology, Biology and Physics*, 71:1581–1590, 2008.

[12] M. Casiraghi, F. Albertini, and A.J. Lomax. Advantages and limitations of the 'worst case scenario' approach in IMPT treatment planning. *Physics in Medicine and Biology*, 58:1323–1339, 2013.

[13] N. Wiener. The Homogeneous Chaos. *American Journal of Mathematics*, 60:897–936, 1938.

[14] D. Xiu and G.E. Karniadakis. The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations. *SIAM Journal on Scientific Computing*, 24:619–644, 2002.

[15] O.P. Le Maître and O.M. Knio. *Spectral Methods for Uncertainty Quantification*. Springer, 2010.

[16] Z. Perkó, L. Gilli, D. Lathouwers, and J.L. Kloosterman. Grid and Basis Adaptive Polynomial Chaos Techniques for Sensitivity and Uncertainty Analysis. *Journal of Computational Physics*, 260:54–84, 2014.

[17] L. Gilli, D. Lathouwers, J.L. Kloosterman, T.H.J.J. van der Hagen, A.J. Koning, and D. Rochman. Uncertainty quantification for criticality problems using non-intrusive and adaptive Polynomial Chaos Techniques. *Annals of Nuclear Energy*, 56:71–80, 2013.

[18] B. Sudret. Global sensitivity analysis using polynomial chaos expansion. *Reliability Engineering and System Safety*, 93:964–979, 2008.

[19] I. Dimov and R. Georgieva. Monte Carlo algorithms for evaluating Sobol' sensitivty indices. *Mathematics and Computers in Simulation*, 81:506–514, 2010.

[20] A. Argiris, M.V. Karamouzis, D. Raben, and R.L. Ferris. Head and neck cancer. *The Lancet*, 371:1695–1709, 2008.

[21] B.A. Murphy, S. Ridner, N. Wells, and M. Dietrich. Quality of life research in head and neck cancer: A review of the current state of the science. *Critical Reviews in Oncology and Hematology*, 62:251–267, 2007.

[22] T.A. van de Water, A.J. Lomax, H.P. Bijl, C. Schilstra, E.B. Hug, and J.A. Langendijk. Using a reduced spot size for Intesity-Modulated Proton Therapy potentially improves salivary gland-sparing in oropharyngeal cancer. *International Journal of Radiation Oncology, Biology and Physics*, 82:e313–319, 2012.

[23] B. Kenny. The salivary glands. http://www.patient.co.uk/diagram/salivary-glands-diagram, June 2010.

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

[25] I. van Dam. The price of robustness: impact of wost-case optimization and beam-angle optimization in intensity-modulated proton therapy for oropharyngeal cancer patients. Master's thesis, Delft University of Technology, November 2013.

[26] H.M. Kooy, B.M. Clasie, H.M. Lu, T.M. Madden, H. Bentefour, N. Depauw, J.A. Adams, A.V. Trofimov, D. Demaret, T.F. Delaney, and J.B. Flanz. A case study in proton pencil-beam scanning delivery. *International Journal of Radiation Oncology, Biology and Physics*, 76:624–630, 201.

[27] S. van de Water, A.C. Kraan, S. Breedveld, W. Schillemans, D.N. Teguh, H.M. Kooy, T.M. Madden, B.J.M. Heijmen, and M.S. Hoogeman. Improved efficiency of multi-criteria IMPT treatment planning using iterative resampling of randomly placed pencil beams. *Physics in Medicine and Biology*, 58:6969–6983, 2013.

[28] A. Fredriksson, A. Forsgren, and B. Hårdemark. Minimax optimization for handling range and setup uncertainties in proton therapy. *Medical Physics*, 38:1672–1684, 2011.

[29] M.J. Geck. Osteotomy surgery for severe scoliosis. http://scoliosistexas.com/osteotomy-surgery.php.

[30] A.C. Kraan, S. van de Water, D.N. Teguh, A. Al-Mamgani, T. Madden, H.M. Kooy, B.J.M. Heijmen, and M.S. Hoogeman. Dose Uncertainties in IMPT for Oropharyngeal Cancer in the Presence of Anatomical, Range and Setup Errors. *International Journal of Radiation Oncology, Biology and Physics*, 87:888–896, 2013.

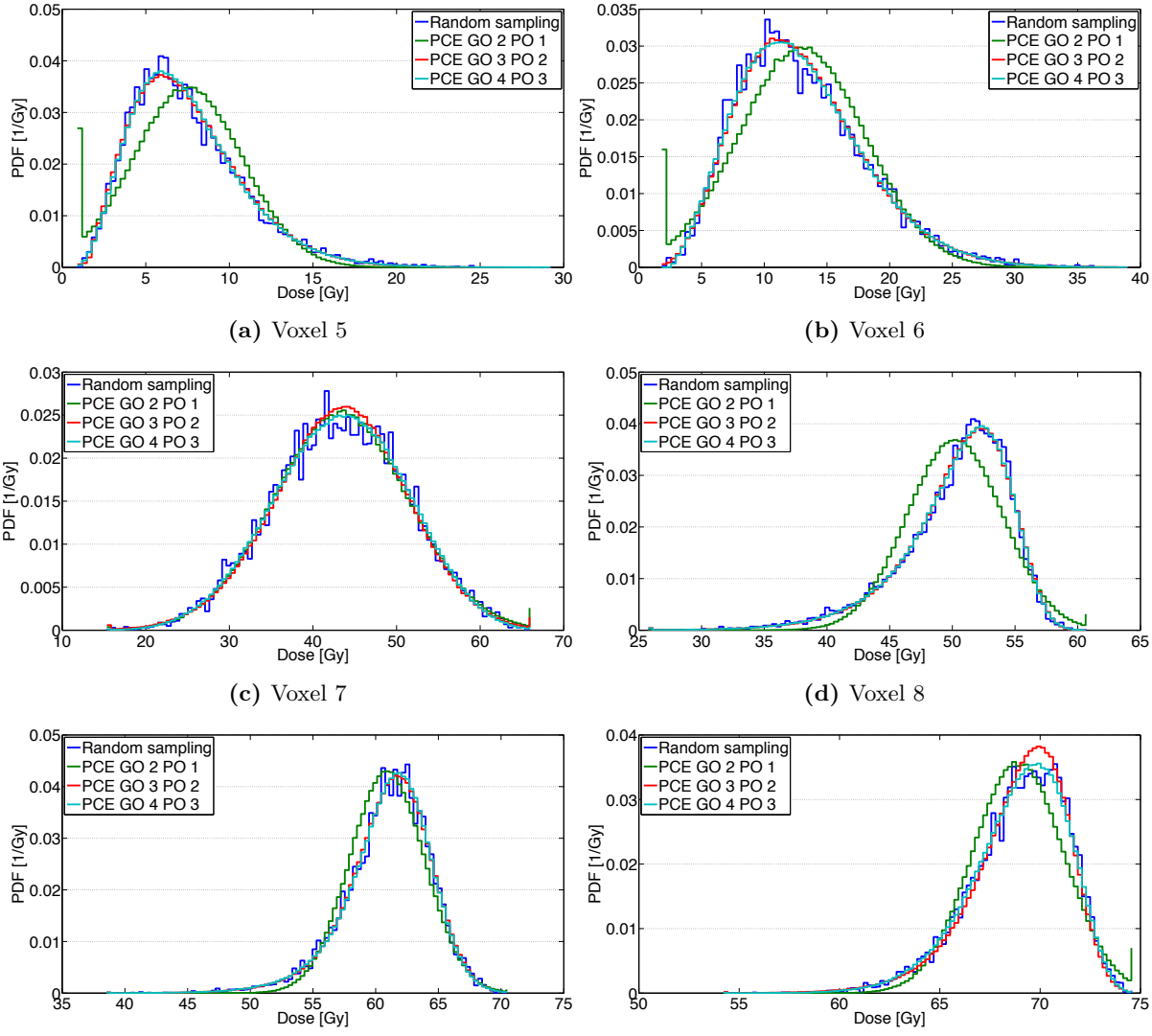
Appendix A

Appendix

In this appendix a broader selection of figures is presented. Due to lack of space not all possible figures are shown. Instead it is chosen to show a representative selection. The figures are presented in the same order as they appear in Chapter 4.

Full dose distribution

This section presents figures complementary to Section 4.2. In Figure A.1 the probability density functions for the six voxels that were not selected are shown. Figure A.2 shows the error-volume histograms for the full dose distributions at different distances from the origin. The error-volume histograms for the organs not displayed in Chapter 4 are depicted in Figure A.3.

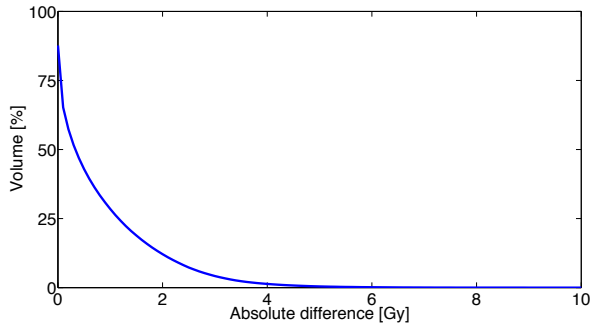


(a) Voxel 5 **(b)** Voxel 6

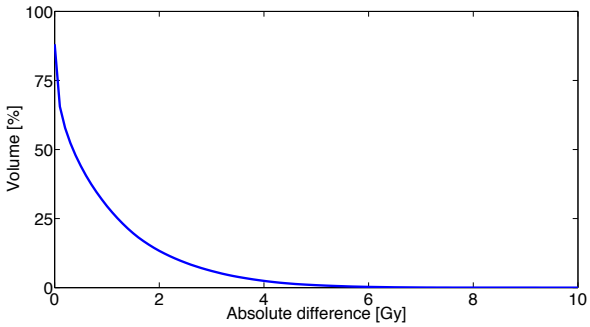
(c) Voxel 7 **(d)** Voxel 8

(e) Voxel 9 **(f)** Voxel 10

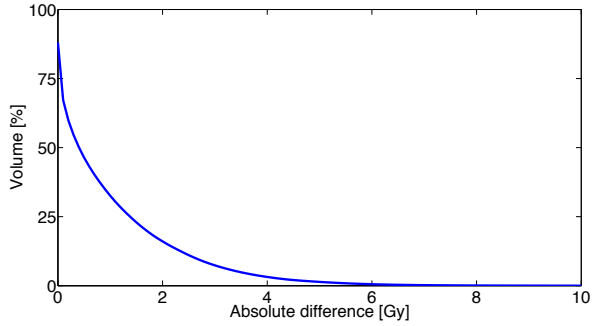
Figure A.1: Probability density function for six randomly selected voxels in patient 1. The impact of higher order expansions strongly depends on the shape of the probability density function of the dose in the voxel.



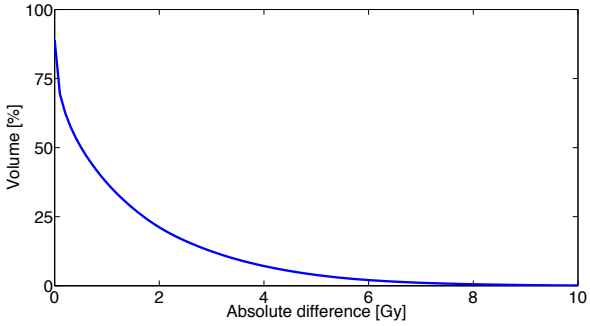
(a) Displacement from origin is 0.1413 mm.



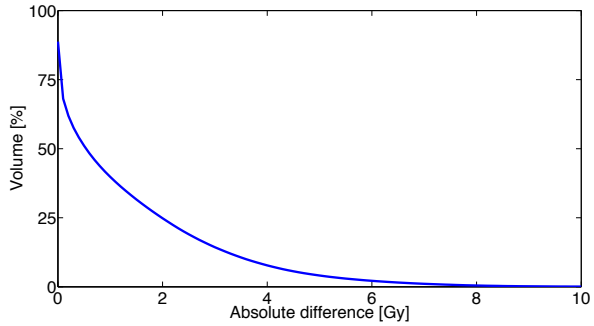
(b) Displacement from origin is 0.1705 mm.



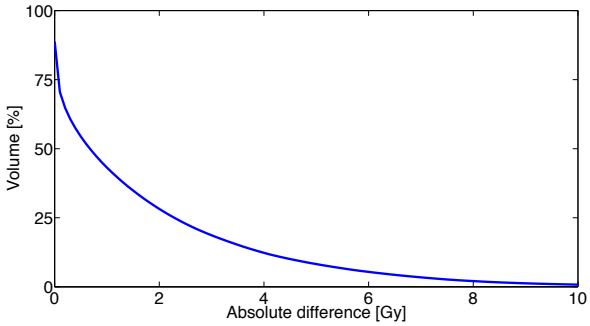
(c) Displacement from origin is 0.2099 mm.



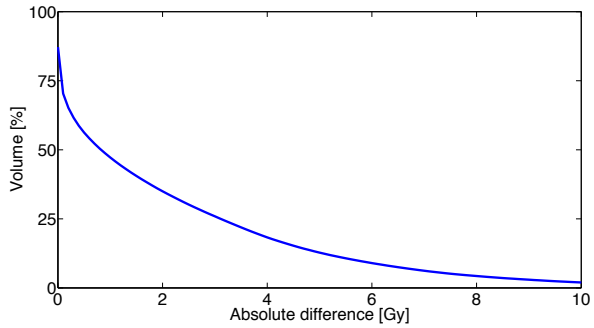
(d) Displacement from origin is 0.3522 mm.



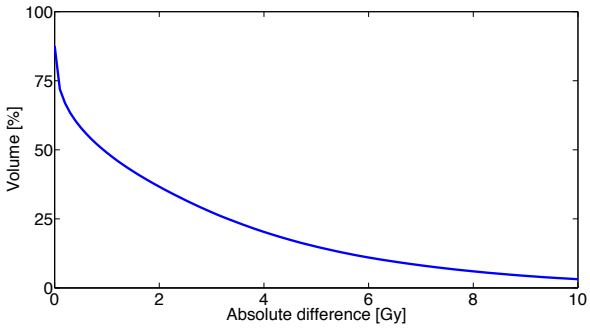
(e) Displacement from origin is 0.3865 mm.



(f) Displacement from origin is 0.6365 mm.

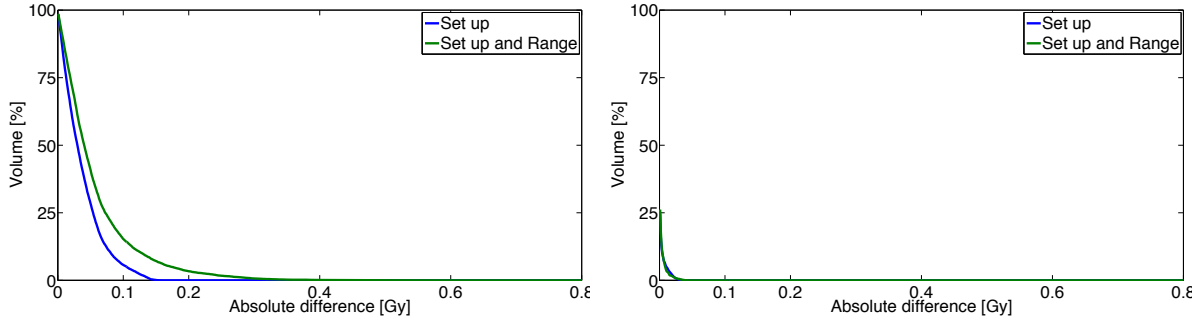


(g) Displacement from origin is 0.9788 mm.



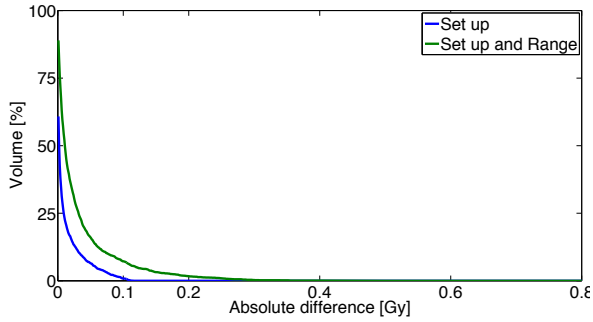
(h) Displacement from origin is 1.2479 mm.

Figure A.2: Error volume histograms for full dose distributions with different distances from the origin. The smaller the displacement from the origin is, the more accurate the prediction of the distribution becomes. Only voxels with a nonzero dose in the reference dose distribution are taken into account in the calculations.

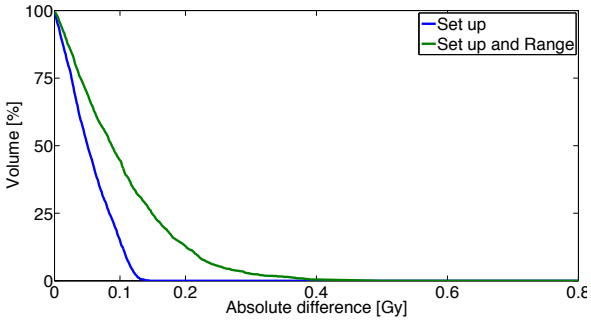


(a) Right parotid gland

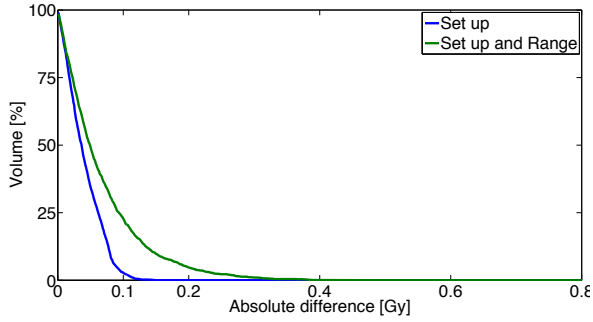
(b) Left parotid gland



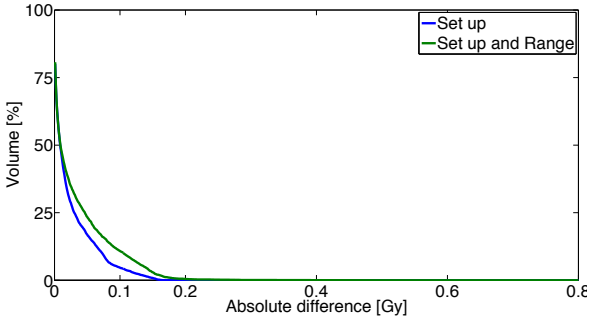
(c) Left submandibular gland



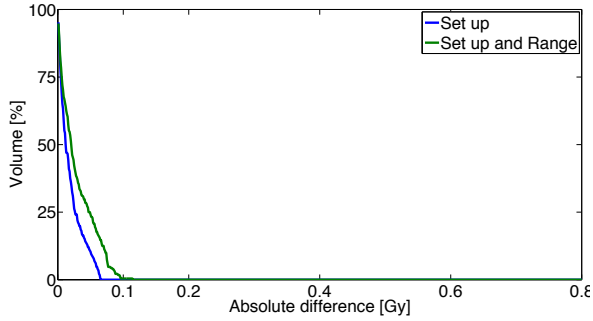
(d) SCM



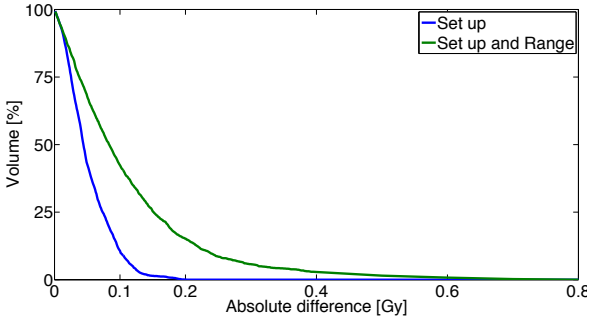
(e) MCM



(f) MCI

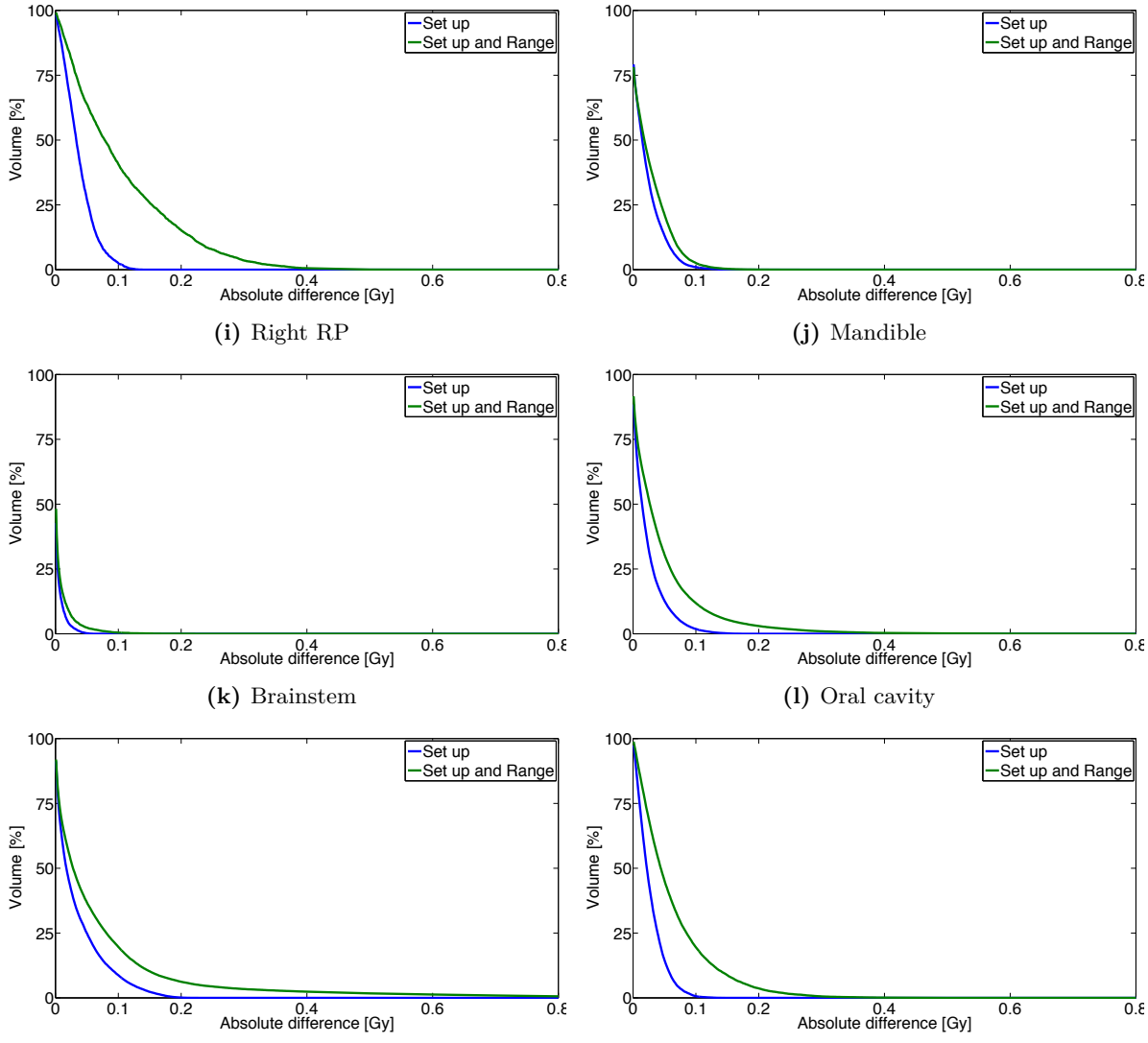


(g) MCP



(h) Left RP

Figure A.3: Error volume histogram for the expected dose in the separate organs as a function of the absolute difference between the reference and the polynomial chaos expansion. Only voxels with a nonzero dose are taken into account in the calculations.



(i) Right RP

(j) Mandible

(k) Brainstem

(l) Oral cavity

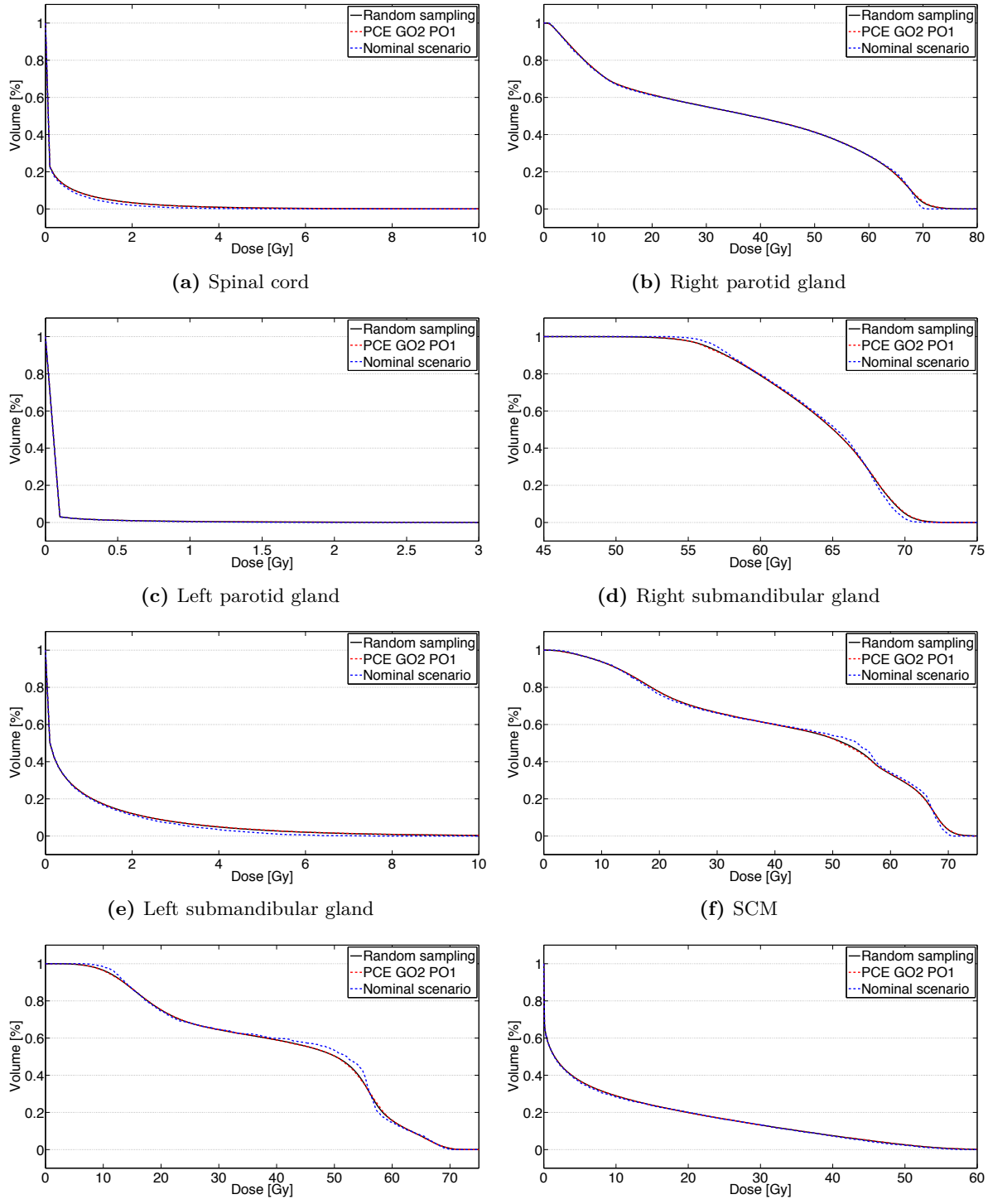
(m) Larynx

(n) CTV high

Figure A.3: *(Continued)* Error volume histogram for the expected dose in the separate organs as a function of the absolute difference between the reference and the polynomial chaos expansion. Only voxels with a nonzero dose are taken into account in the calculations.

Dose volume histogram

This section displays figures complementary to Section 4.3. First the difference between the nominal and the expected dose-volume histograms is shown for all organs in Figure A.4. The impact of higher order expansions on the accuracy of the produced bandwidths are shown in Figure A.5. This section concludes with a comparison between the cumulative and differential dose-volume histograms for the CTV low is displayed in Figure A.6.



(a) Spinal cord

(b) Right parotid gland

(c) Left parotid gland

(d) Right submandibular gland

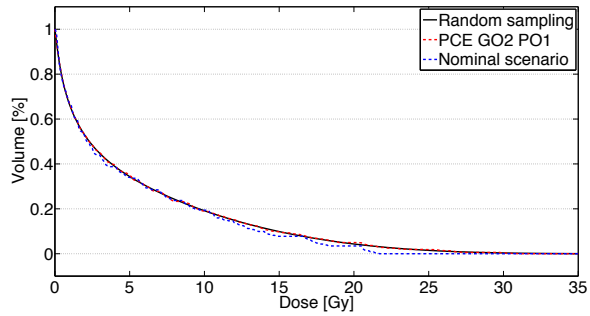
(e) Left submandibular gland

(f) SCM

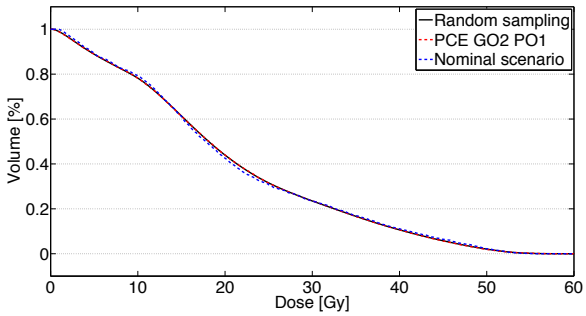
(g) MCM

(h) MCI

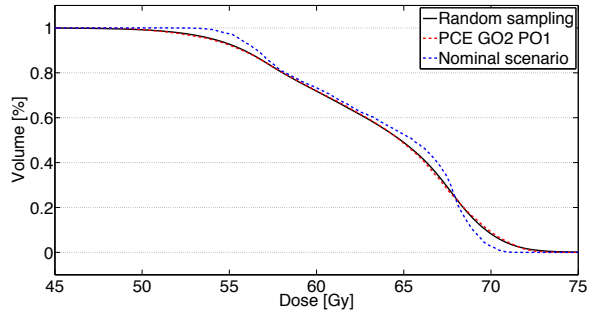
Figure A.4: Illustration of the difference between the nominal scenario and the expected dose distribution as predicted by random sampling and polynomial chaos expansion with first order polynomials and grid order 2. For some organs the nominal scenario is an excellent approximation, but for other organs the nominal scenario is not a good approximation of the expected dose distribution. Only setup uncertainties were taken into account in these figures.



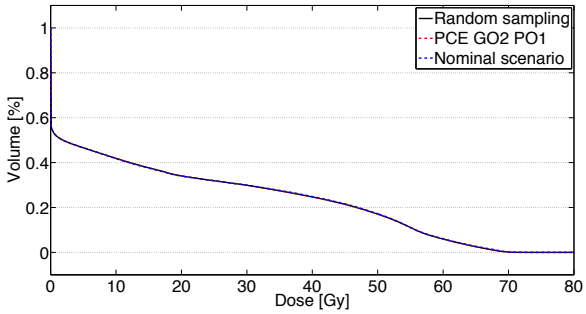
(i) MCP



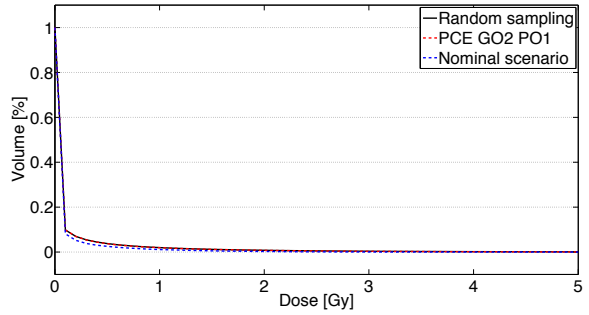
(j) Left RP



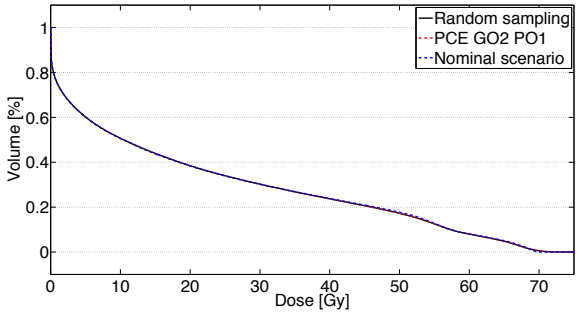
(k) Right RP



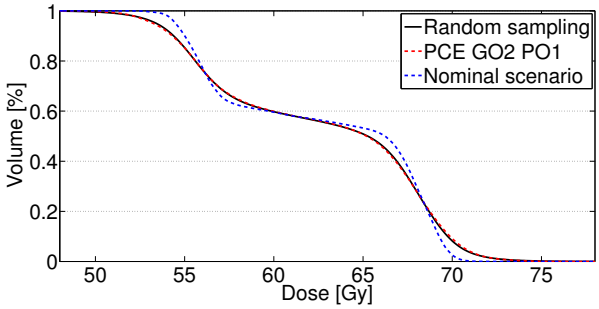
(l) Mandible



(m) Brainstem

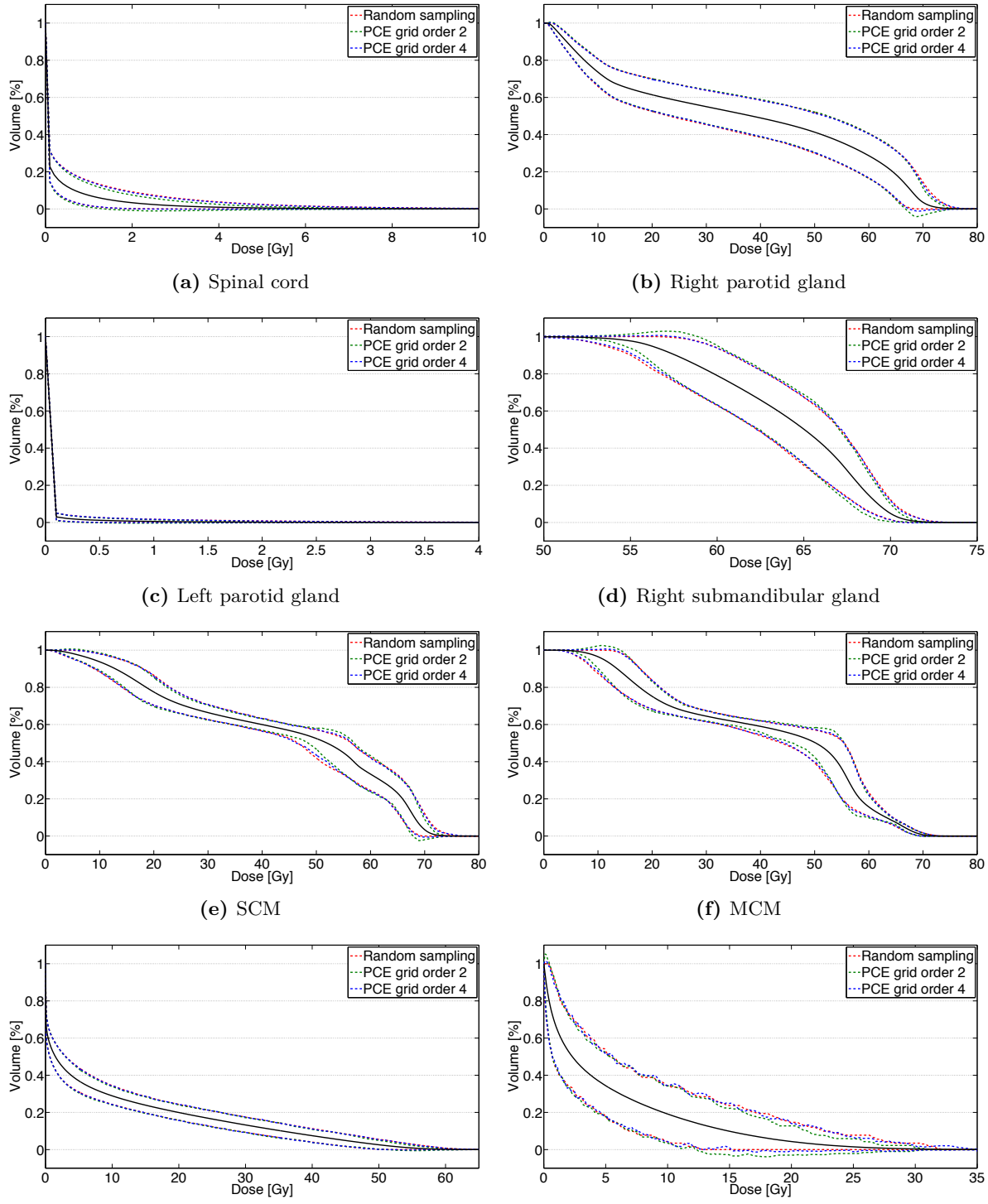


(n) Oral cavity



(o) CTV low

Figure A.4: *(Continued)* Illustration of the difference between the nominal scenario and the expected dose distribution as predicted by random sampling and polynomial chaos expansion with first order polynomials and grid order 2. For some organs the nominal scenario is an excellent approximation, but for other organs the nominal scenario is not a good approximation of the expected dose distribution. Only setup uncertainties were taken into account in these figures.



(a) Spinal cord

(b) Right parotid gland

(c) Left parotid gland

(d) Right submandibular gland

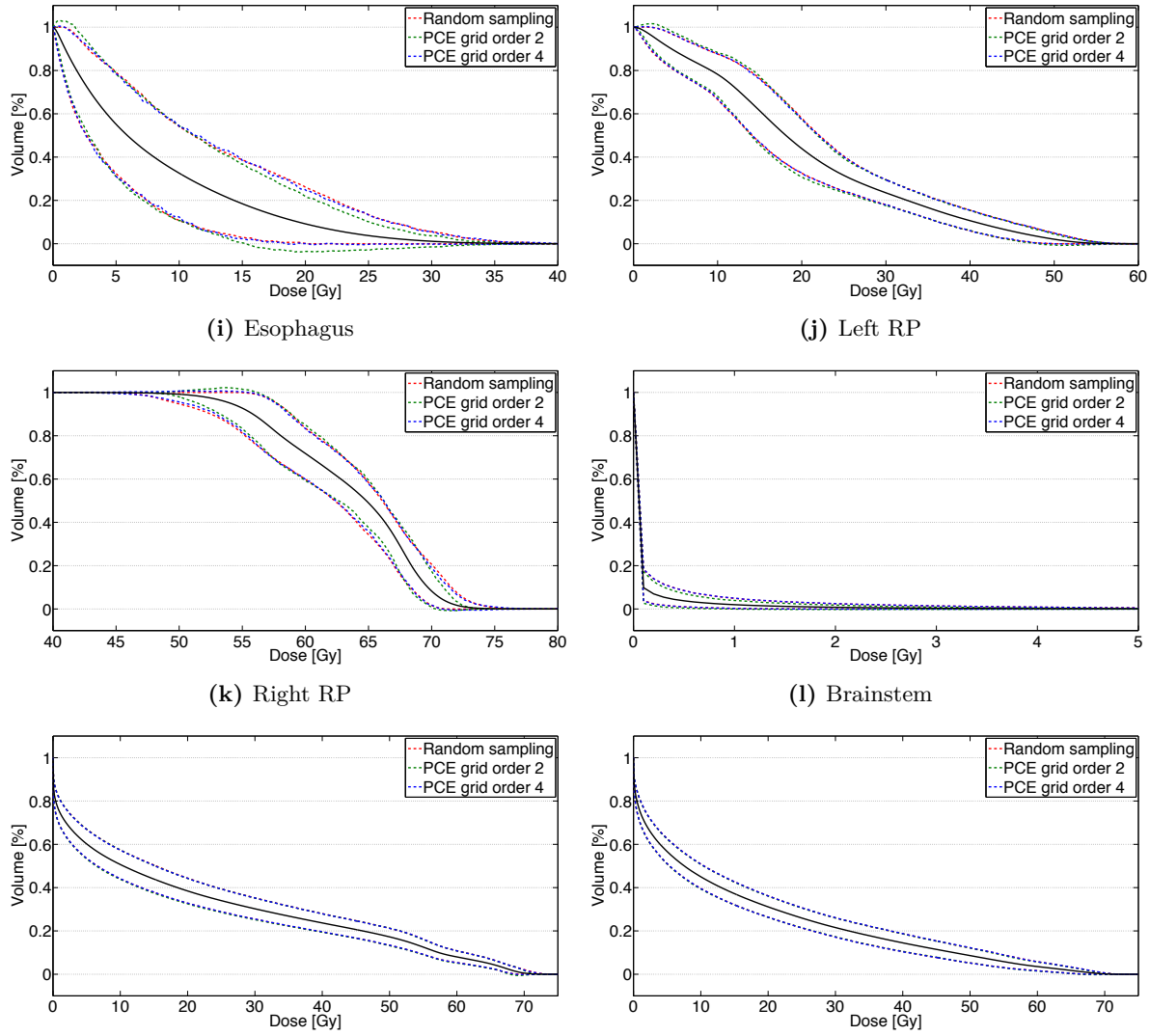
(e) SCM

(f) MCM

(g) MCI

(h) MCP

Figure A.5: Comparison between lower (GO 2, PO 1) and higher (GO 4, PO 3) grid orders in the approximation of the bandwidth of the DVH for different organs. Higher orders have a variable impact on the accuracy with which the bandwidth of the DVH can be predicted. Higher order expansions perfectly overlap the reference bandwidth created by random sampling. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures.



(i) Esophagus

(j) Left RP

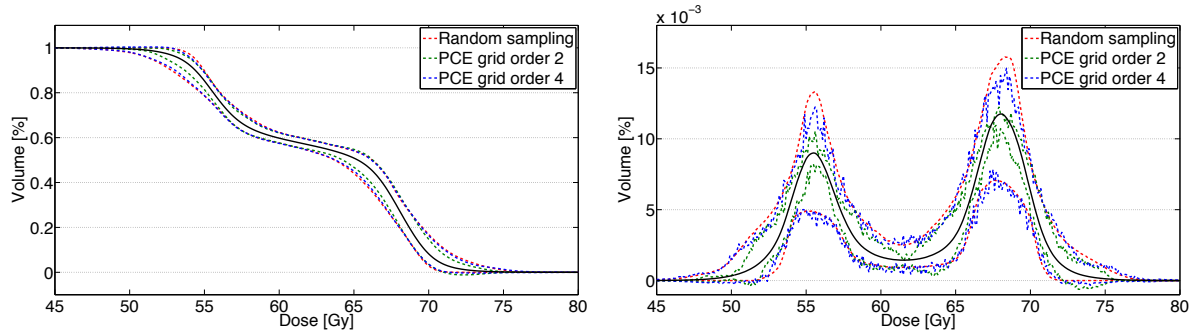
(k) Right RP

(l) Brainstem

(m) Oral cavity

(n) Larynx

Figure A.5: *(Continued)* Comparison between lower (GO 2, PO 1) and higher (GO 4, PO 3) grid orders in the approximation of the bandwidth of the DVH for different organs. Higher orders have a variable impact on the accuracy with which the bandwidth of the DVH can be predicted. Higher order expansions perfectly overlap the reference bandwidth created by random sampling. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures.

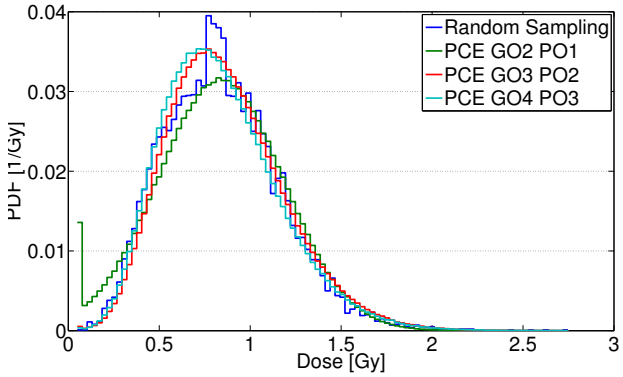


(a) Cumulative DVH **(b)** Differential DVH

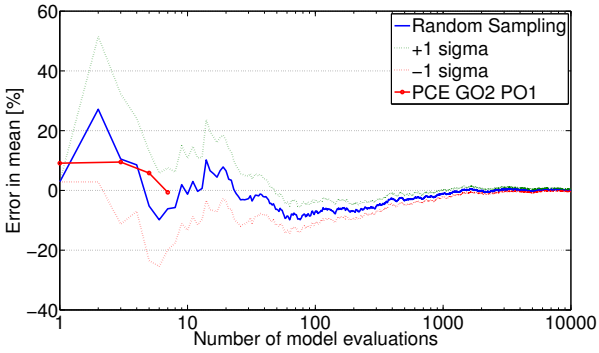
Figure A.6: Comparison between the cumulative and differential dose-volume histograms of the CTV low. In both figures the difference in grid order is clearly visible, as higher grid orders deliver a much higher accuracy. The differential DVH is not easier to approximate using polynomial chaos expansions, even though this was expected beforehand. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures.

Dose and tumor parameters

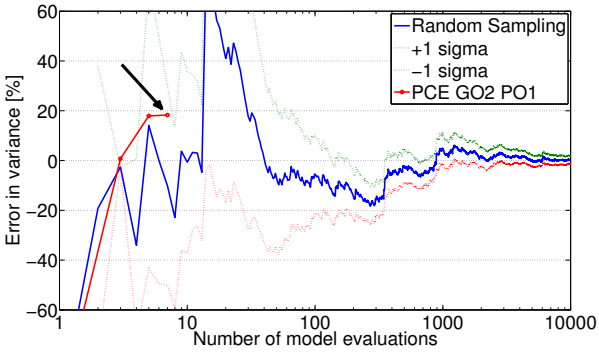
Over 60 parameters were evaluated in this research from 18 different organs. For each of these parameters, convergence graphs for both the mean and the variance for several polynomial and grid orders can be constructed. Instead it is chosen to select six parameters that vary in organ and type of parameter. A full analysis of the minimum dose of the right parotid gland is depicted in Figure A.7. The mean dose has been selected twice, first the mean dose in the brainstem is depicted in Figure A.8 and second the mean dose in Figure A.9. For the maximum dose the right submandibular gland is chosen and displayed in Figure A.10. Two tumor parameters have been selected to complete this section. The homogeneity index for the CTV low is shown in Figure A.11 and the V95 parameter for the CTV high is given in Figure A.12.



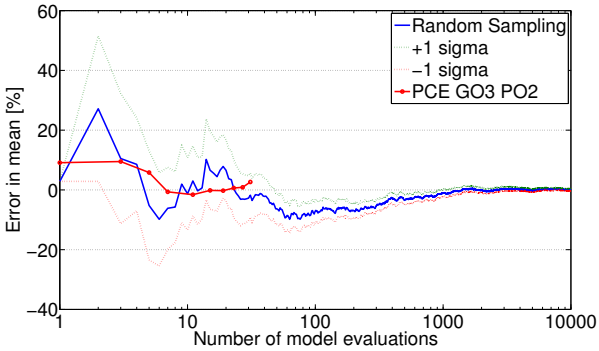
(a) Probability density function



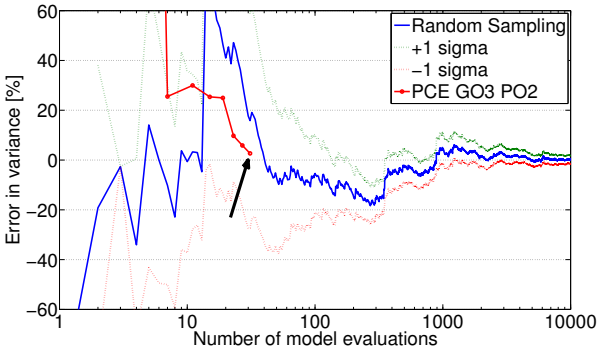
(b) Convergence mean GO2 PO1



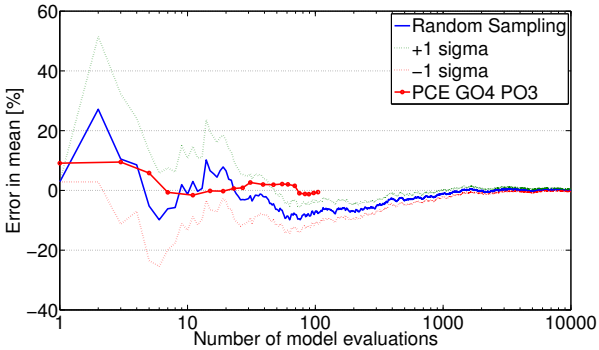
(c) Convergence variance GO2 PO1



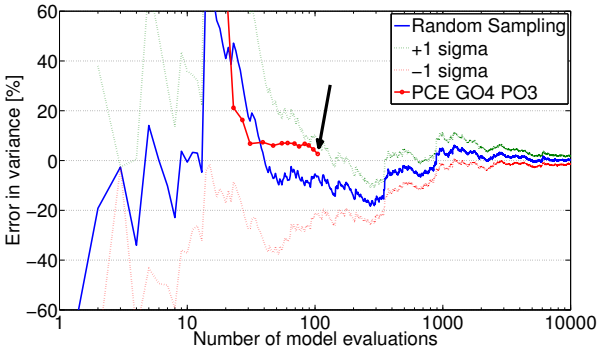
(d) Convergence mean GO3 PO2



(e) Convergence variance GO3 PO2

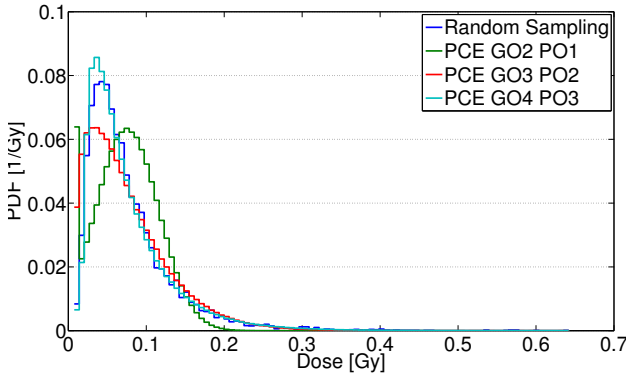


(f) Convergence mean GO4 PO3

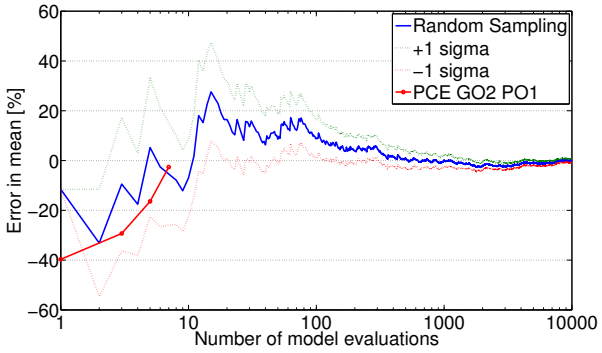


(g) Convergence variance GO4 PO3

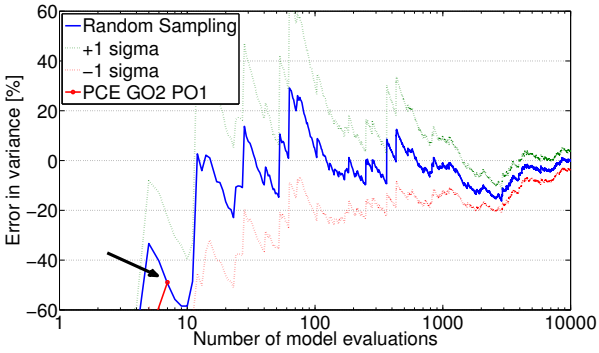
Figure A.7: Illustration of the accuracy of different grid and polynomial orders for the minimum dose in the right parotid gland of patient 1 plan 1. Only setup uncertainties are taken into account.



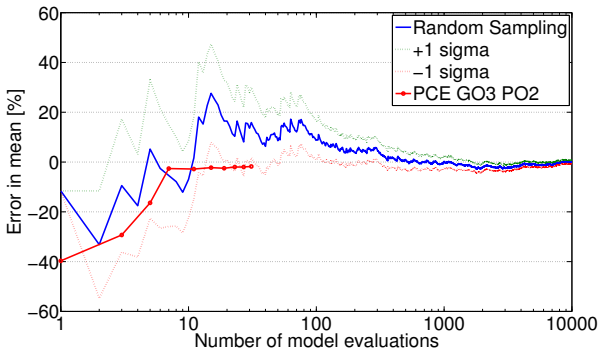
(a) Probability density function



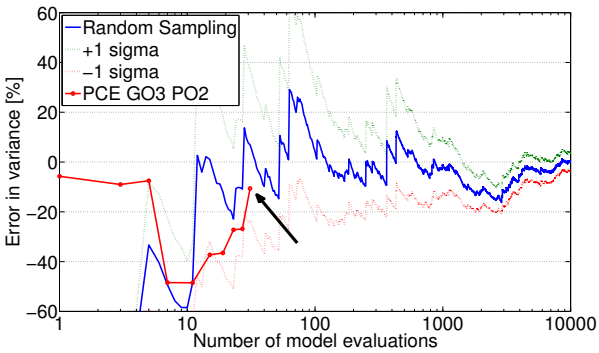
(b) Convergence mean GO2 PO1



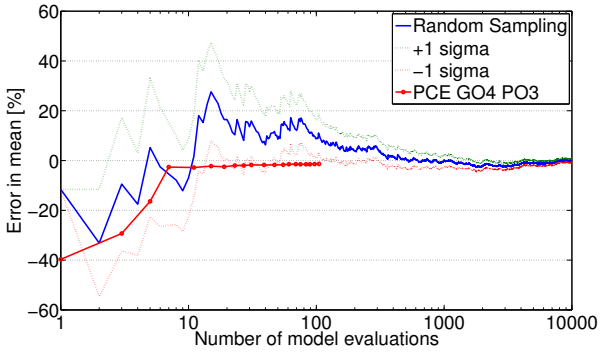
(c) Convergence variance GO2 PO1



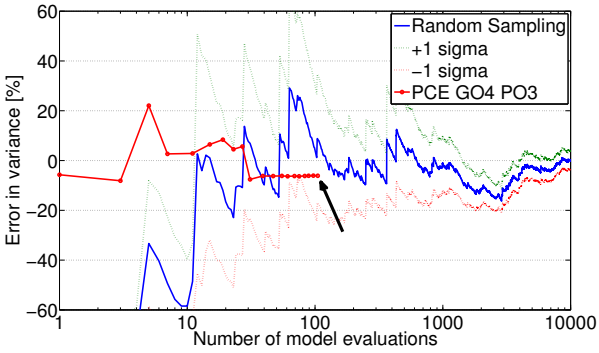
(d) Convergence mean GO3 PO2



(e) Convergence variance GO3 PO2

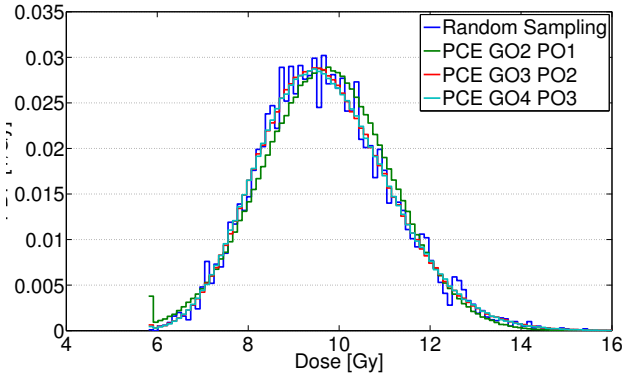


(f) Convergence mean GO4 PO3

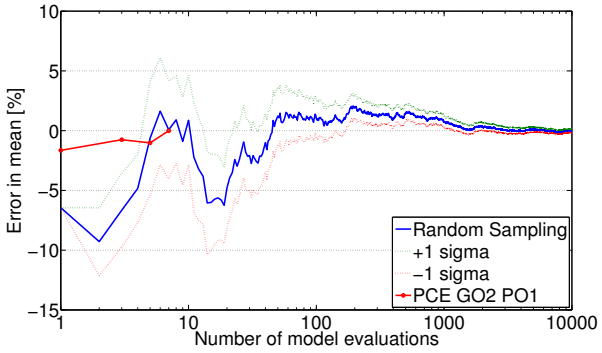


(g) Convergence variance GO4 PO3

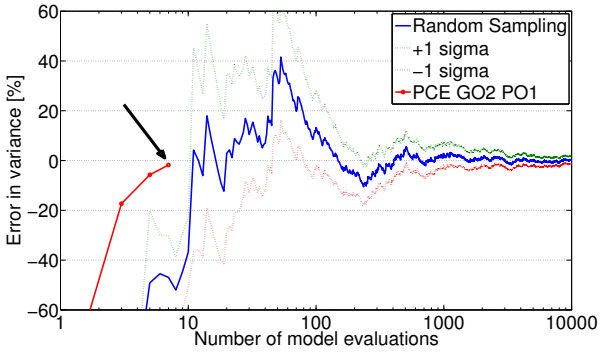
Figure A.8: Illustration of the accuracy of different grid and polynomial orders for the mean dose in the brainstem of patient 1 plan 1. Only setup uncertainties are taken into account.



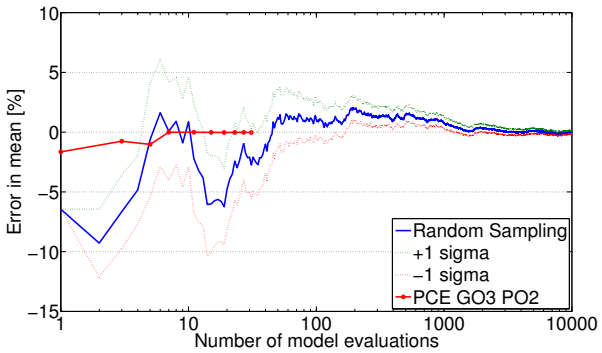
(a) Probability density function



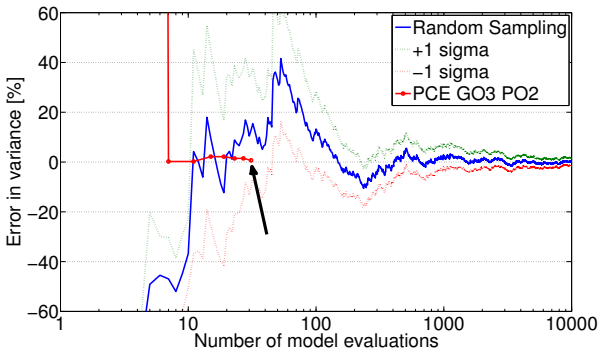
(b) Convergence mean GO2 PO1



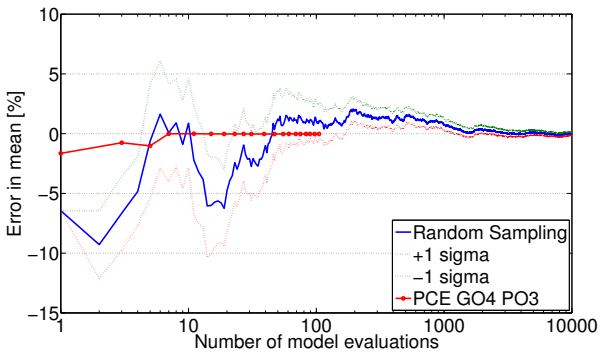
(c) Convergence variance GO2 PO1



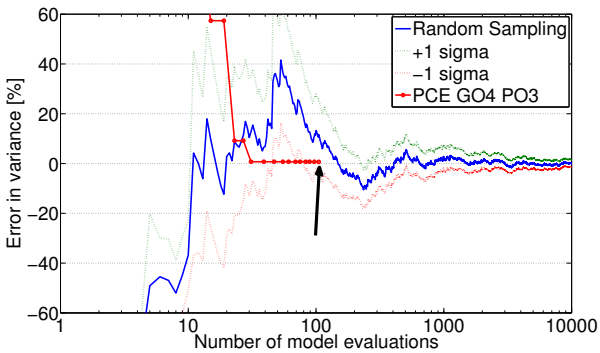
(d) Convergence mean GO3 PO2



(e) Convergence variance GO3 PO2

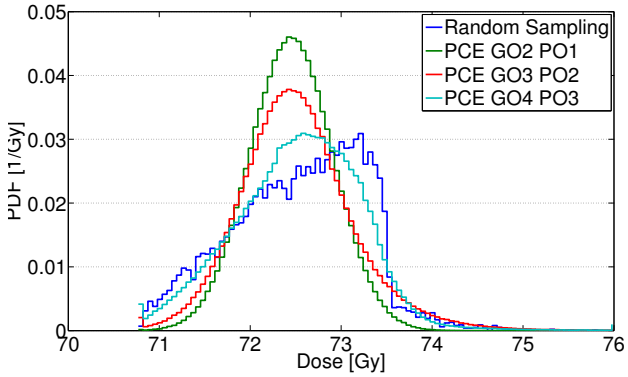


(f) Convergence mean GO4 PO3

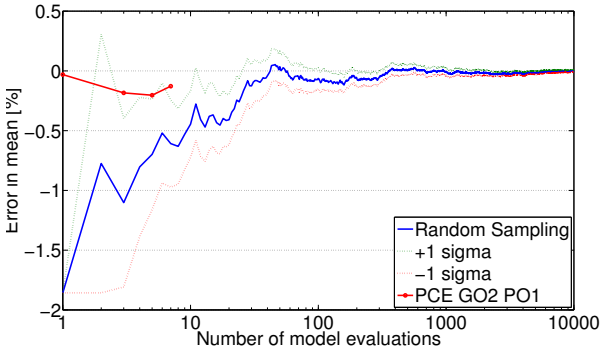


(g) Convergence variance GO4 PO3

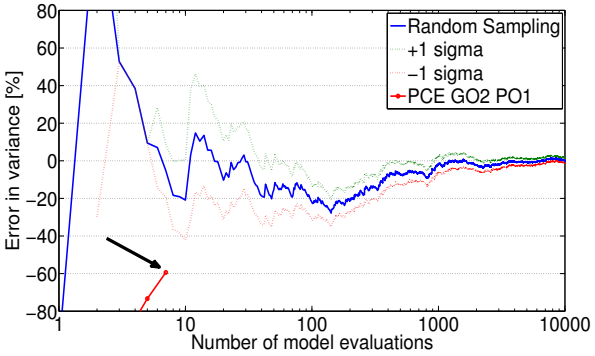
Figure A.9: Illustration of the accuracy of different grid and polynomial orders for the mean dose in the MCI of patient 1 plan 1. Only setup uncertainties are taken into account.



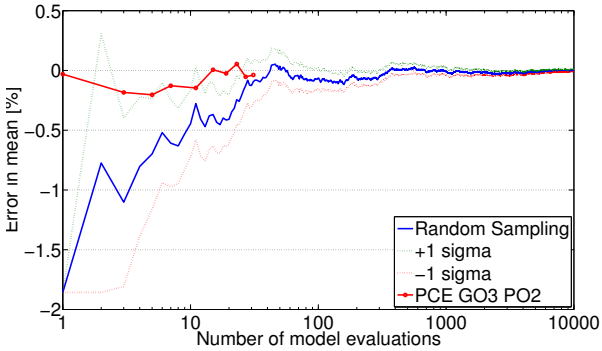
(a) Probability density function



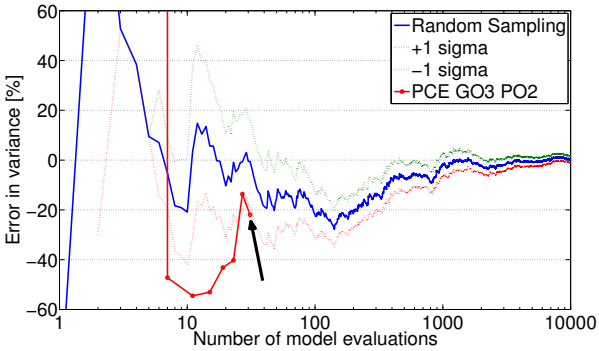
(b) Convergence mean GO2 PO1



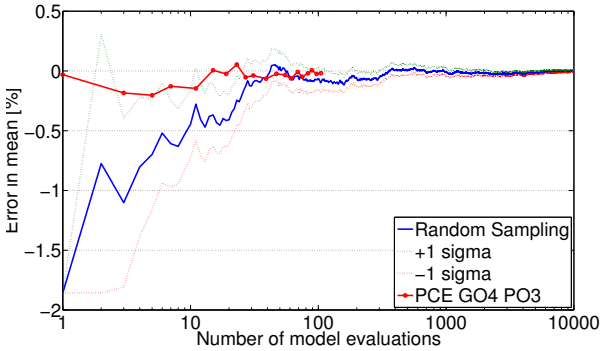
(c) Convergence variance GO2 PO1



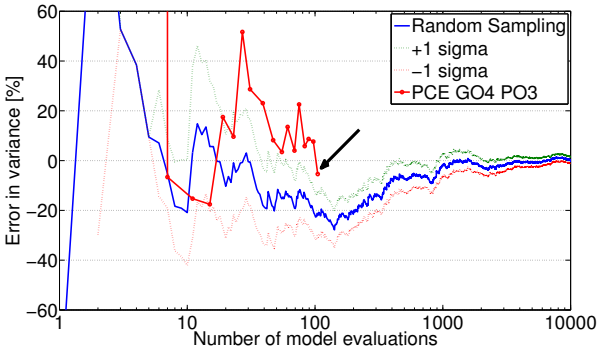
(d) Convergence mean GO3 PO2



(e) Convergence variance GO3 PO2

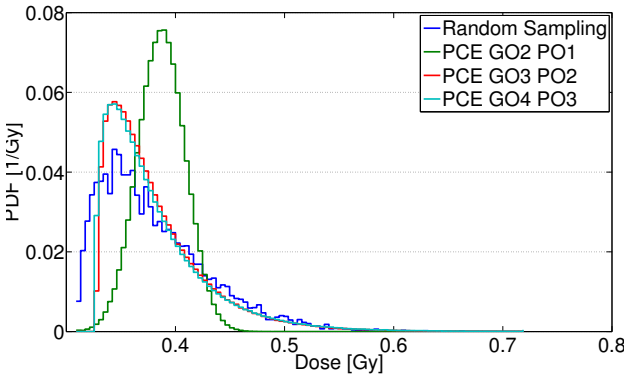


(f) Convergence mean GO4 PO3

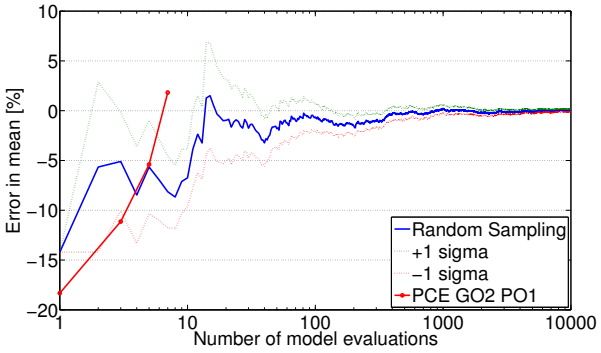


(g) Convergence variance GO4 PO3

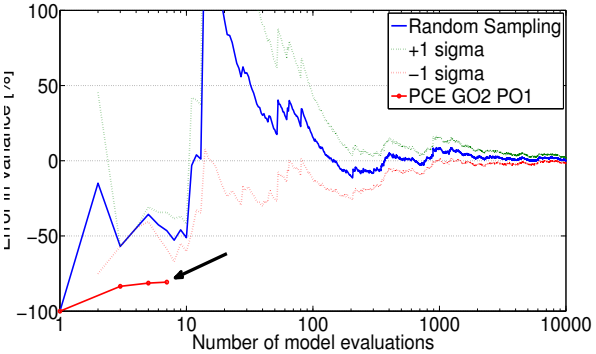
Figure A.10: Illustration of the accuracy of different grid and polynomial orders for the maximum dose in the right submandibular gland of patient 1 plan 1. Only setup uncertainties are taken into account.



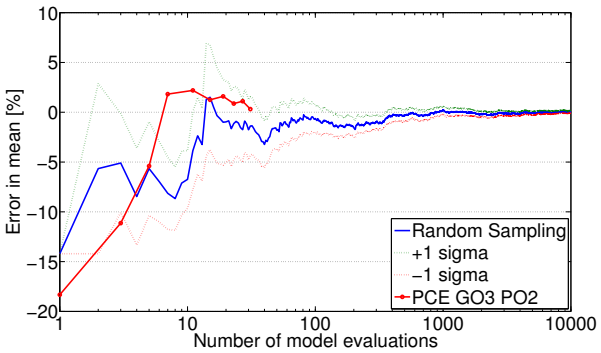
(a) Probability density function



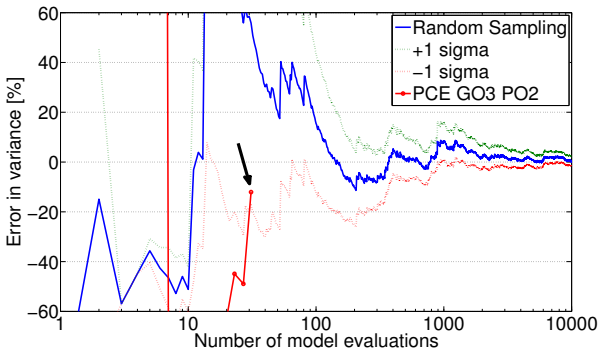
(b) Convergence mean GO2 PO1



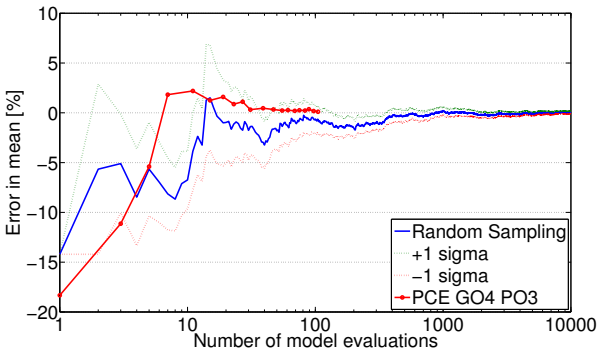
(c) Convergence variance GO2 PO1



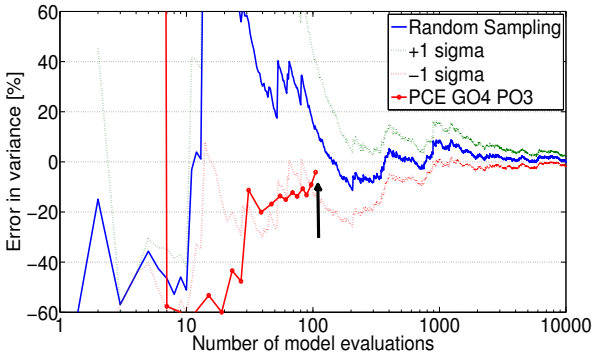
(d) Convergence mean GO3 PO2



(e) Convergence variance GO3 PO2

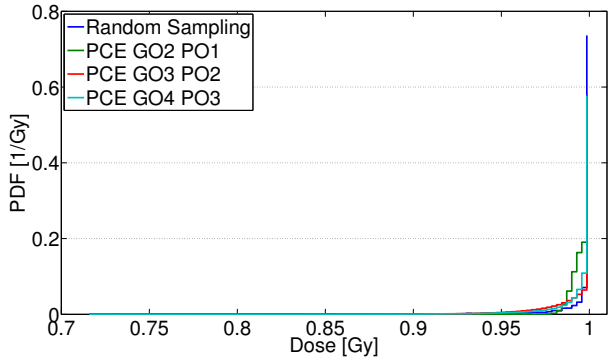


(f) Convergence mean GO4 PO3

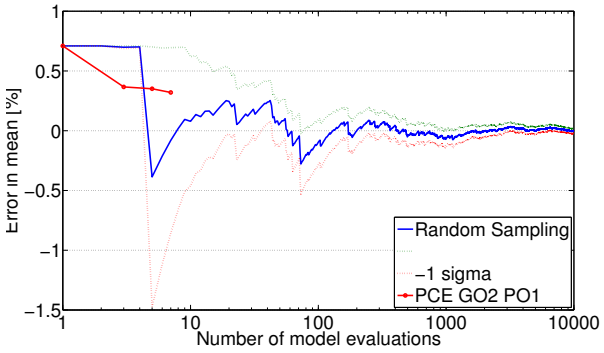


(g) Convergence variance GO4 PO3

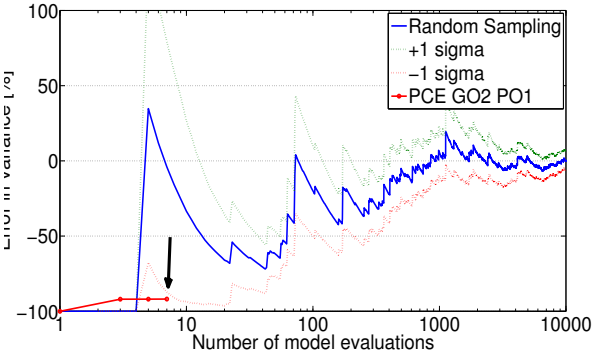
Figure A.11: Illustration of the accuracy of different grid and polynomial orders for the homogeneity index of the CTV low of patient 1 plan 1. Only setup uncertainties are taken into account.



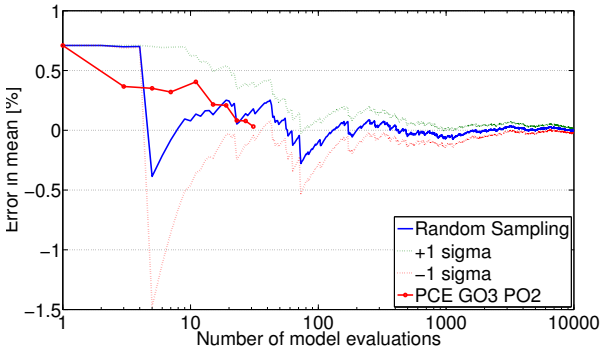
(a) Probability density function



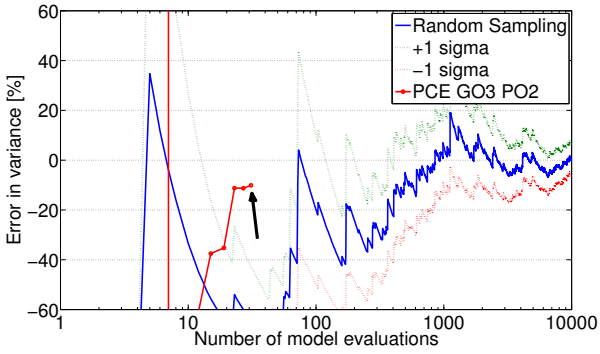
(b) Convergence mean GO2 PO1



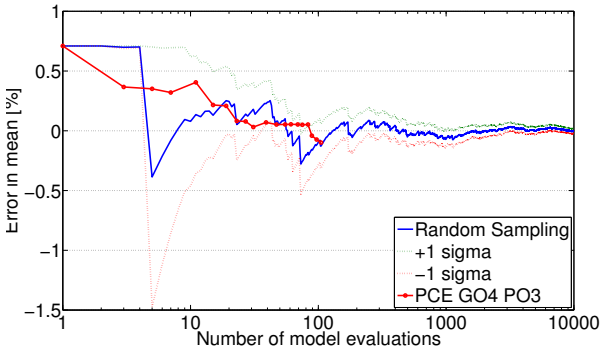
(c) Convergence variance GO2 PO1



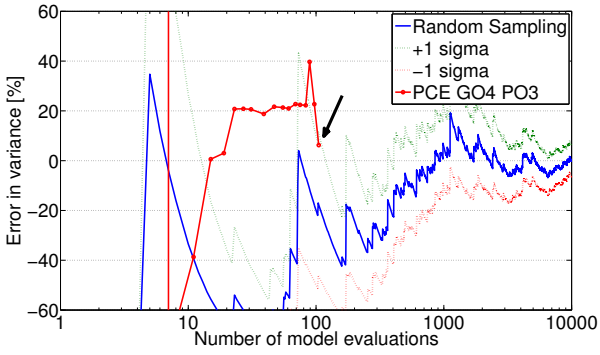
(d) Convergence mean GO3 PO2



(e) Convergence variance GO3 PO2



(f) Convergence mean GO4 PO3

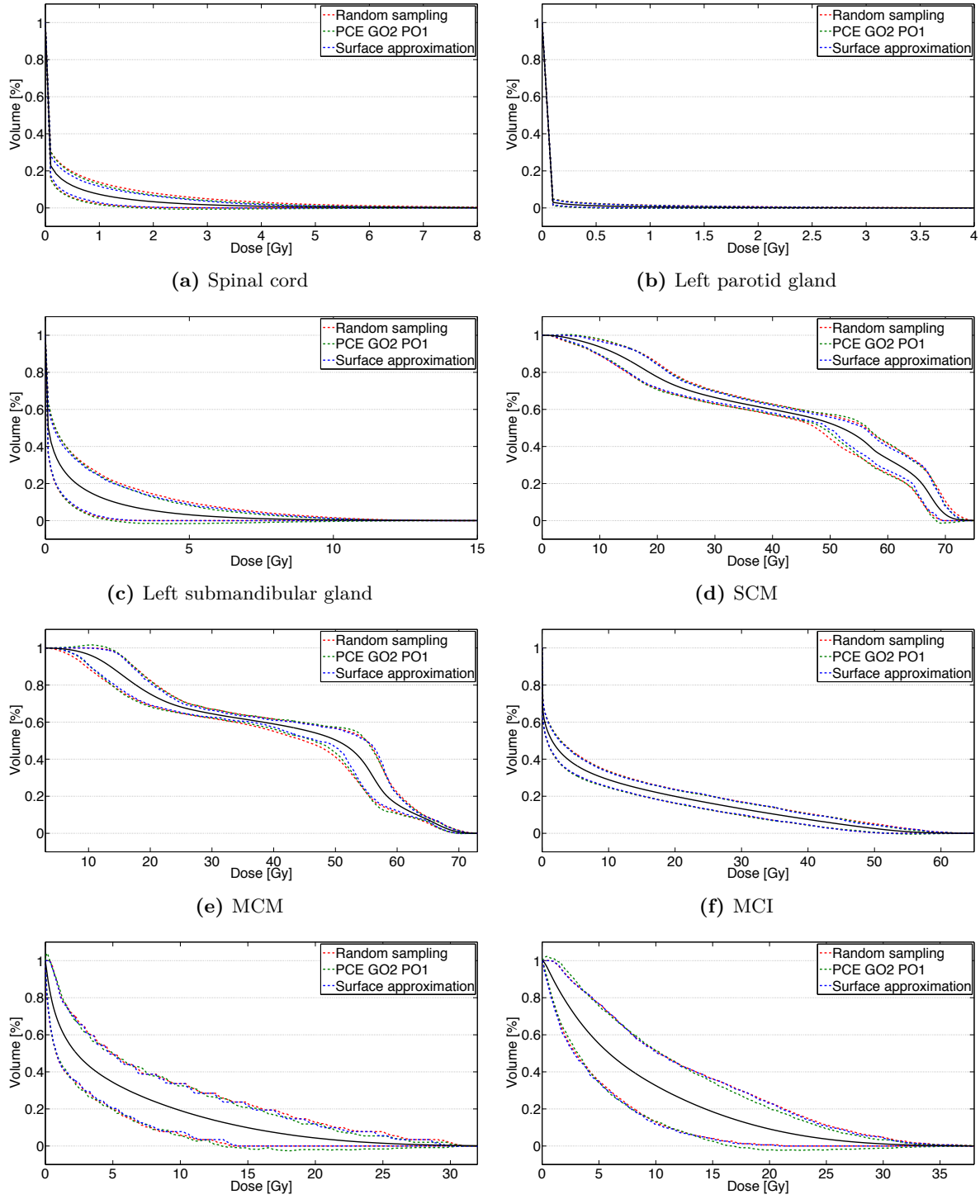


(g) Convergence variance GO4 PO3

Figure A.12: Illustration of the accuracy of different grid and polynomial orders for the V95 of the CTV high of patient 1 plan 1. Only setup uncertainties are taken into account.

Comparison with the surface approximation

In Section 4.5 a qualitative comparison was made between the bandwidths created by the surface approximation and the polynomial chaos method. In this section the remaining organs are depicted in Figure A.13



(a) Spinal cord

(b) Left parotid gland

(c) Left submandibular gland

(d) SCM

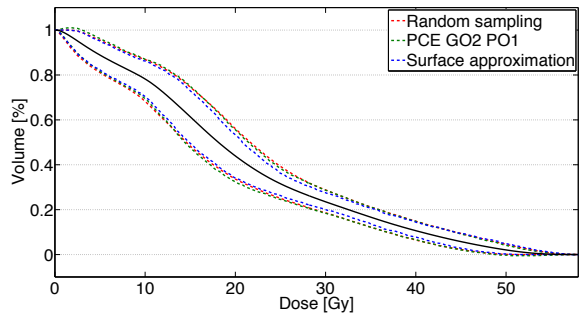
(e) MCM

(f) MCI

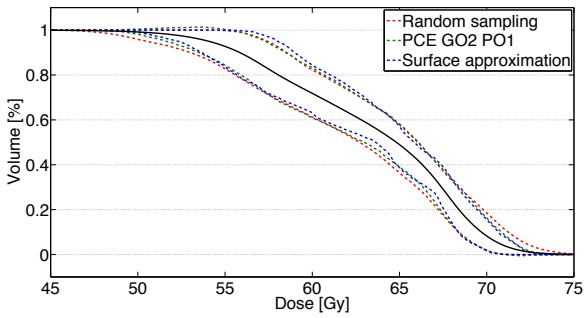
(g) MCP

(h) Esophagus

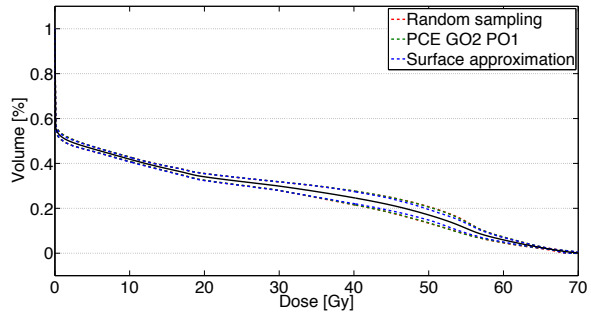
Figure A.13: Qualitative comparison between a seven point polynomial chaos expansion (first order polynomial, second order grid) and a seven point surface approximation for different organs. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures.



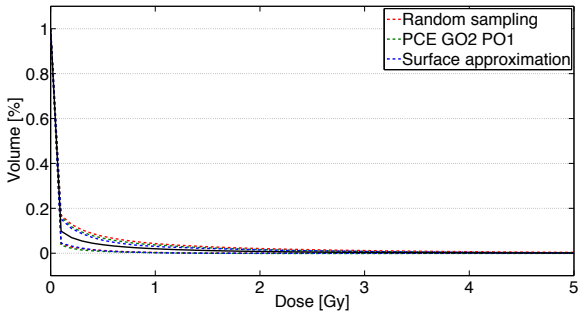
(i) Left RP



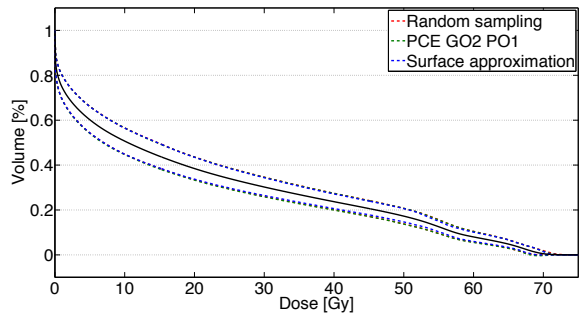
(j) Right RP



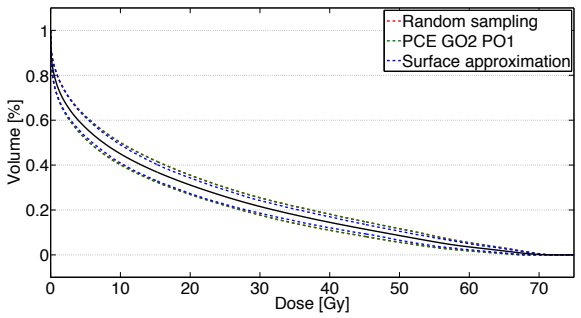
(k) Mandible



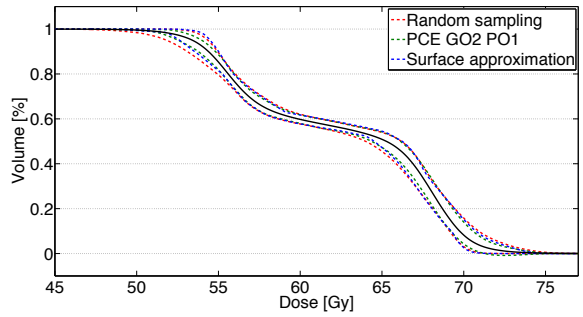
(l) Brainstem



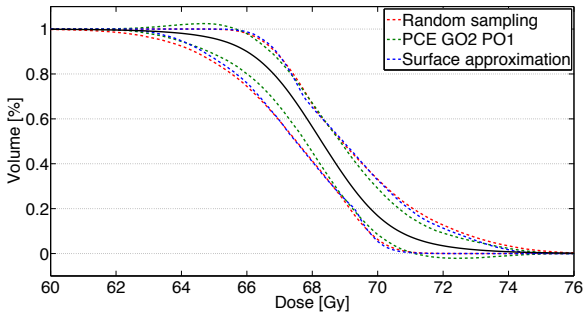
(m) Oral cavity



(n) Larynx



(o) CTV low

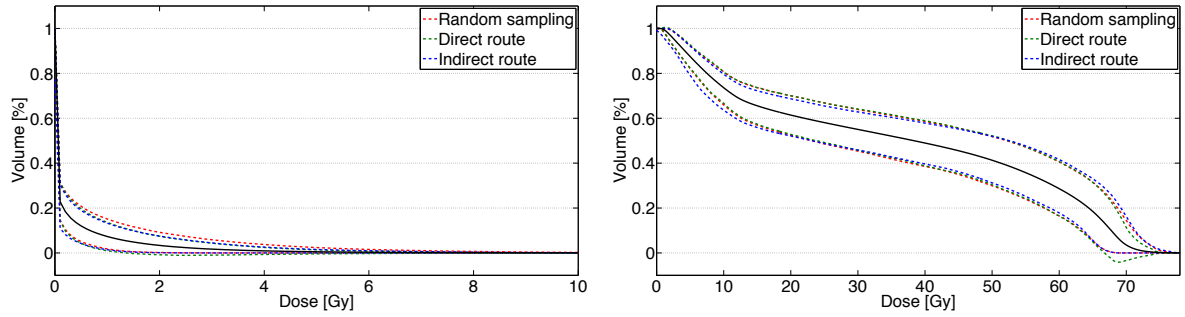


(p) CTV high

Figure A.13: *(Continued)* Qualitative comparison between a seven point polynomial chaos expansion (first order polynomial, second order grid) and a seven point surface approximation for different organs. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures.

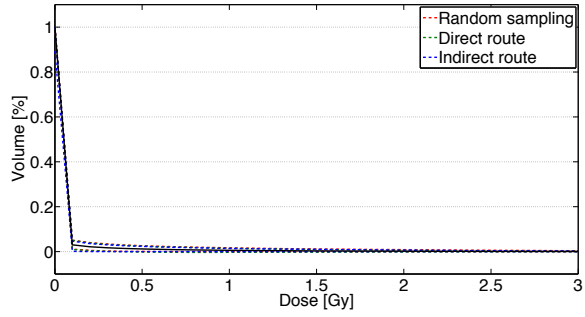
The indirect route versus the direct route

In this section the rest of the figures belonging to Section 4.6 are given. In Figure A.14 the construction of the bandwidths via the direct and the indirect route are compared for organs not displayed in Section 4.6. In Figure A.15 the two types of dose-volume histograms are compared for the CTV high.

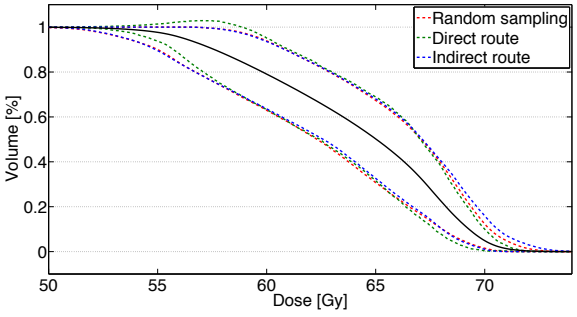


(a) Spinal cord

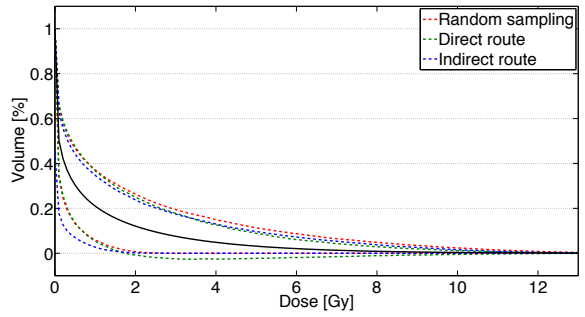
(b) Right parotid gland



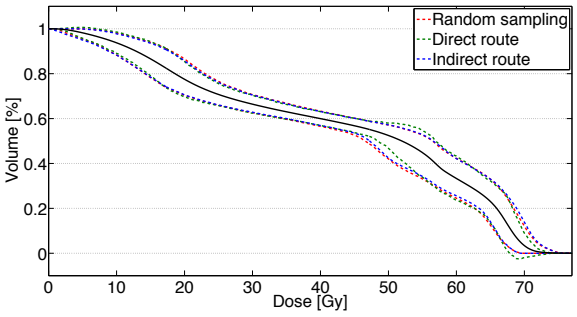
(c) Left parotid gland



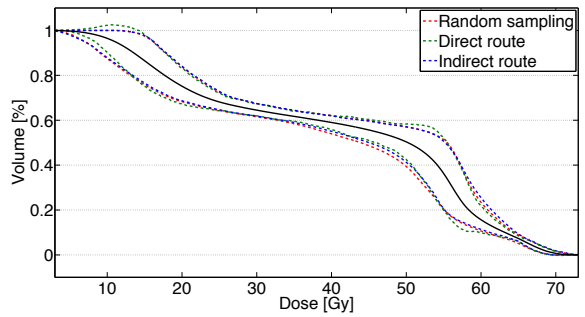
(d) Right submandibular gland



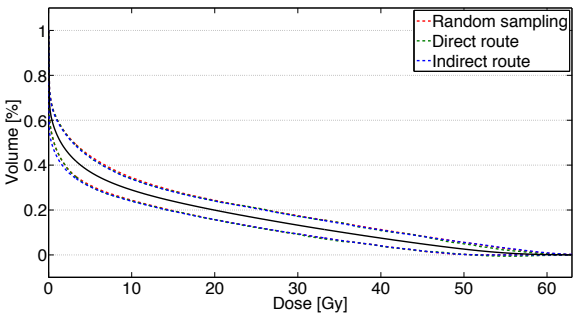
(e) Left submandibular gland



(f) SCM

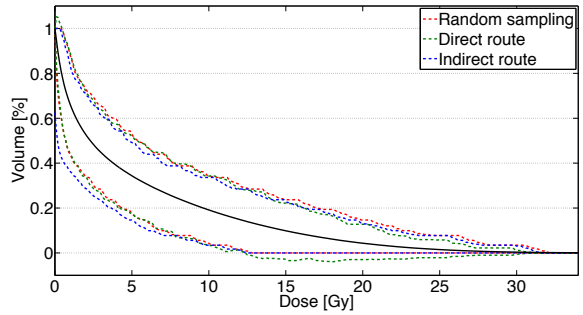


(g) MCM

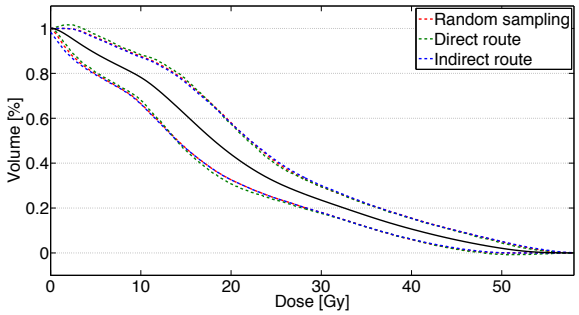


(h) MCI

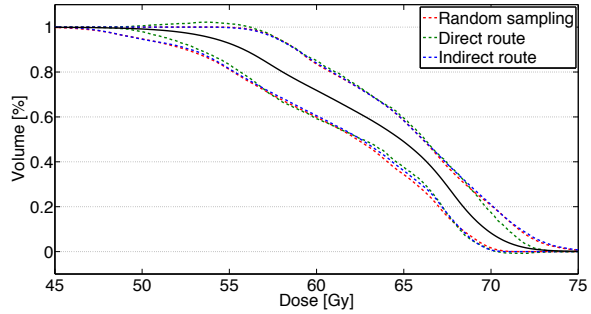
Figure A.14: Comparison between the indirect and direct route in the construction of the bandwidths around the expected dose-volume histograms for various organs. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures. Both expansions were calculated with first order polynomials and second order grids.



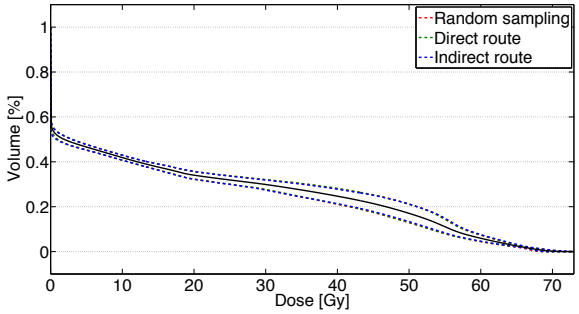
(i) MCP



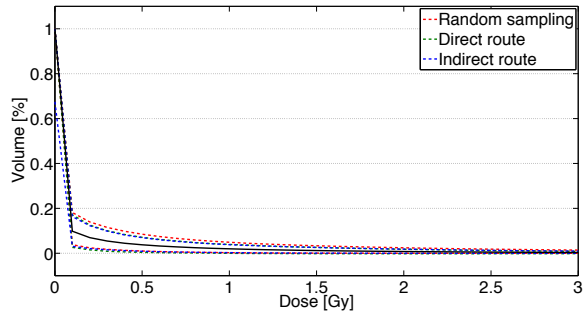
(j) Left RP



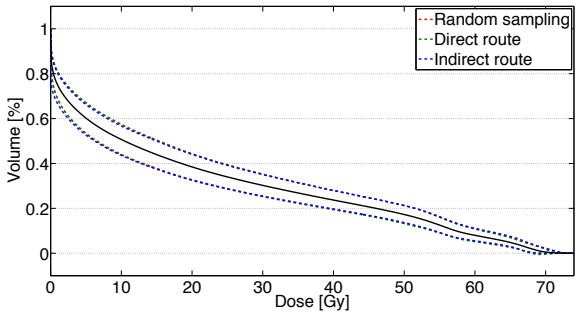
(k) Right RP



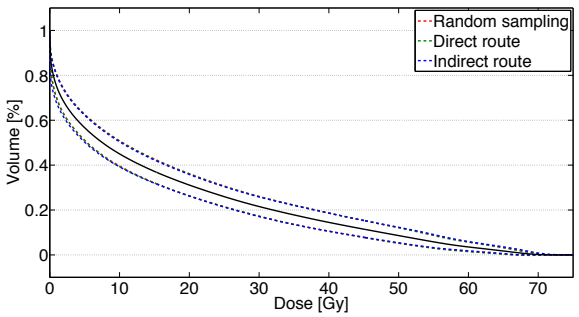
(l) Mandible



(m) Brainstem

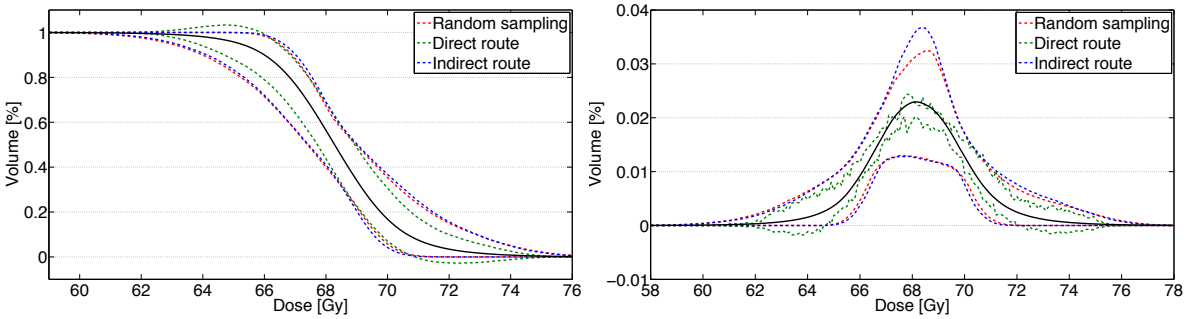


(n) Oral cavity



(o) Larynx

Figure A.14: *(Continued)* Comparison between the indirect and direct route in the construction of the bandwidths around the expected dose-volume histograms for various organs. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures. The expansions were both calculated with first order polynomials and second order grids.

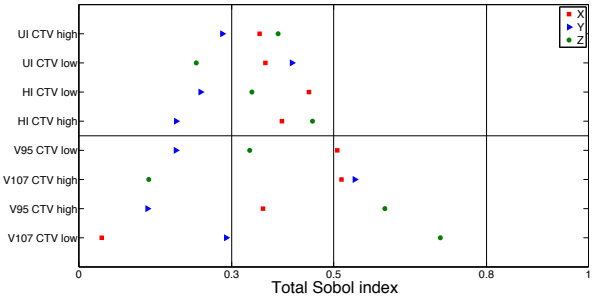


(a) Cumulative DVH **(b)** Differential DVH

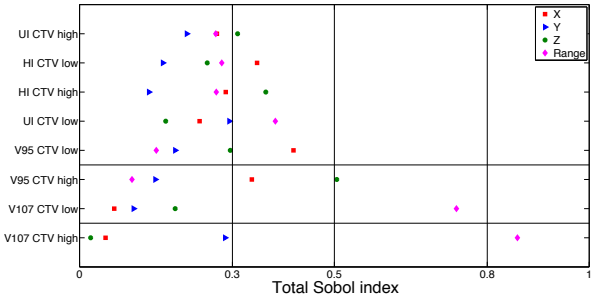
Figure A.15: Comparison between the indirect and direct route for the two types of dose-volume histograms for the CTV high. The DVH of the expected dose is depicted as a solid black line. Only setup uncertainties were taken into account in these figures. The expansions were both calculated with first order polynomials and second order grids.

Variance decomposition

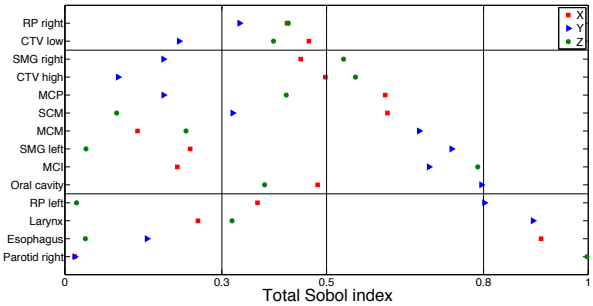
This Appendix concludes with the remaining results for the variance decomposition. First the treatment plan that was used as a reference case in Section 4.8, the second plan for patient 1 is depicted in Figure A.16. Next, the variance decomposition for the first plan for patient 1 is displayed in Figure A.17. Finally, the Sobol' analysis for the second patient is shown in Figure A.18.



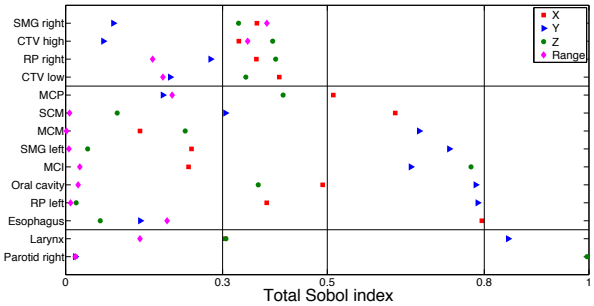
(a) Setup, tumor parameters



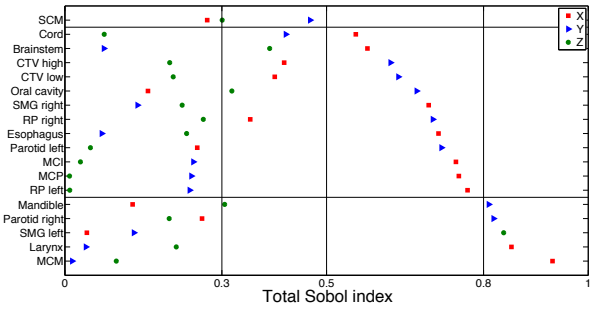
(b) Setup and range, tumor parameters



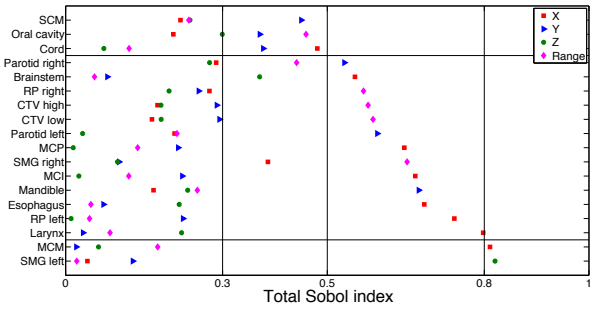
(c) Setup, minimum doses



(d) Setup and range, minimum doses

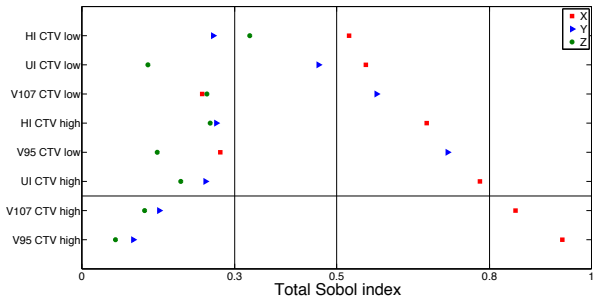


(e) Setup, maximum doses

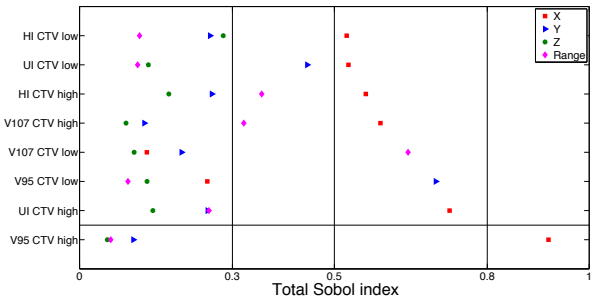


(f) Setup and range, maximum doses

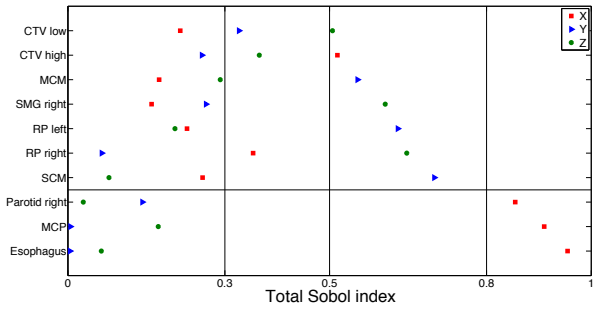
Figure A.16: Sobol' analysis for patient 1 plan 2 for both setup uncertainties and setup and range uncertainties.



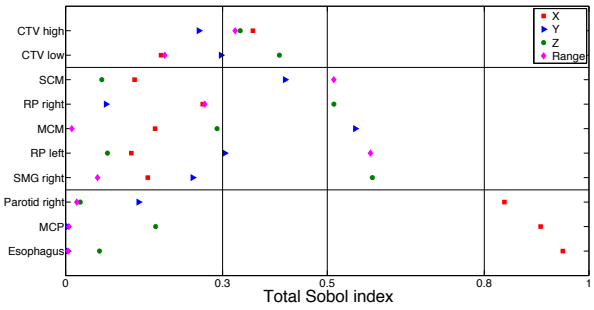
(a) Setup, tumor parameters



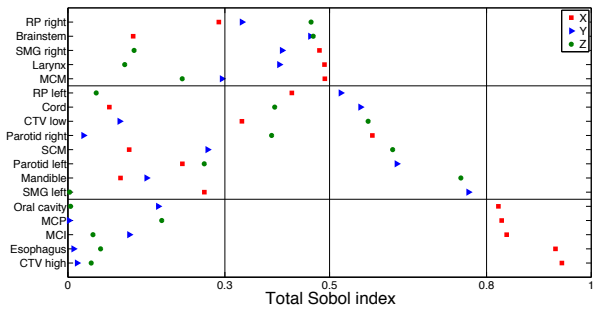
(b) Setup and range, tumor parameters



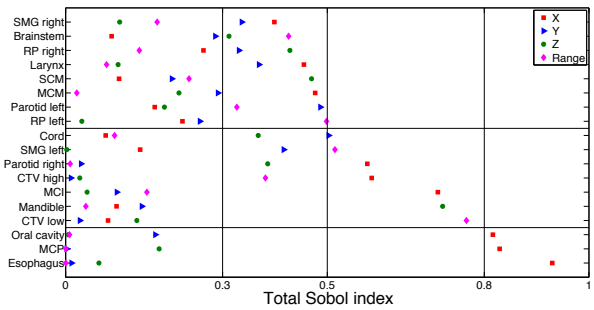
(c) Setup, minimum doses



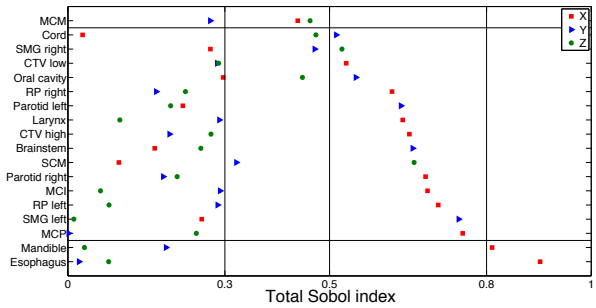
(d) Setup and range, minimum doses



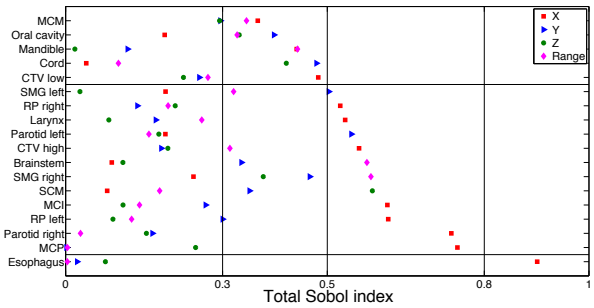
(e) Setup, mean doses



(f) Setup and range, mean doses

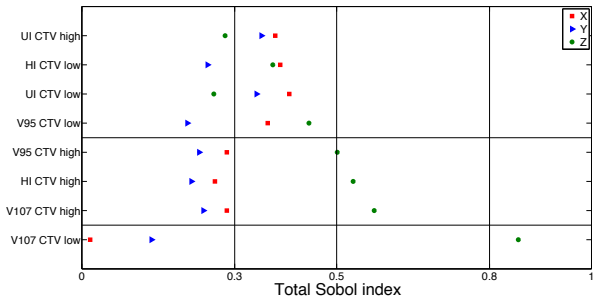


(g) Setup, maximum doses

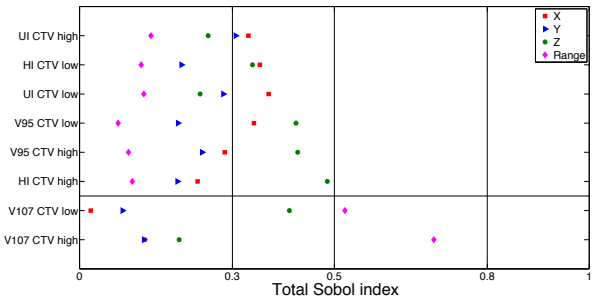


(h) Setup and range, maximum doses

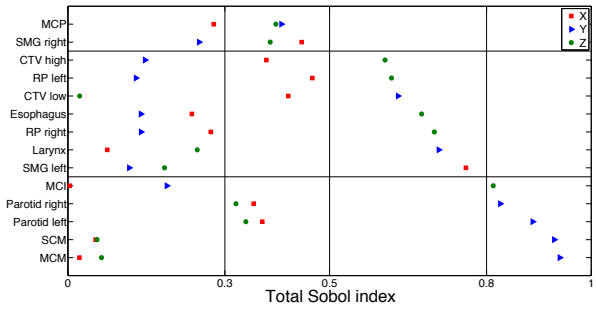
Figure A.17: Sobol' analysis for patient 1 plan 1 for both setup uncertainties and setup and range uncertainties.



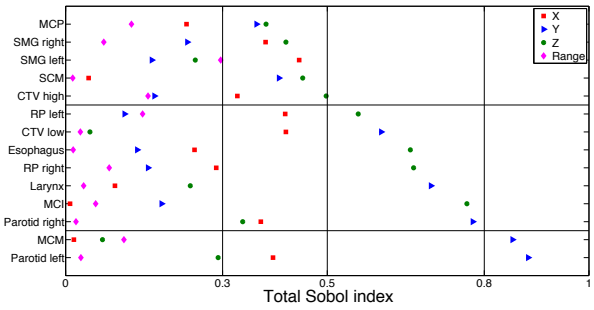
(a) Setup, tumor parameters



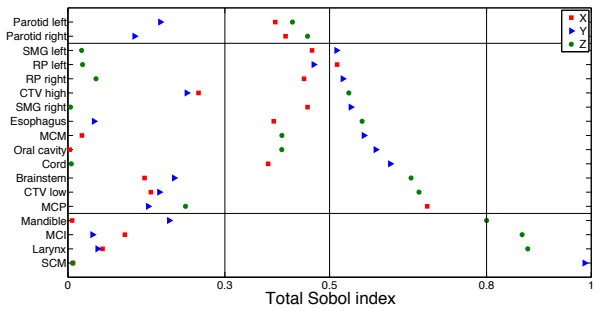
(b) Setup and range, tumor parameters



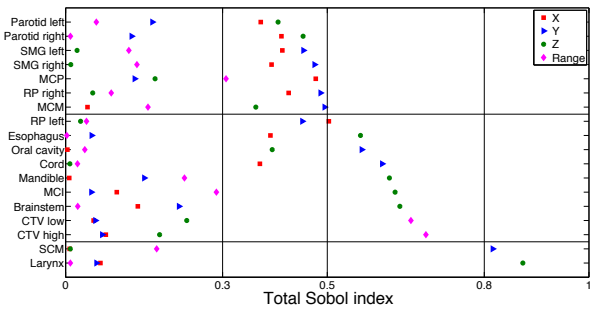
(c) Setup, minimum doses



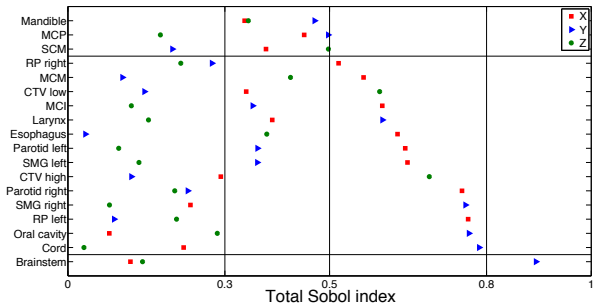
(d) Setup and range, minimum doses



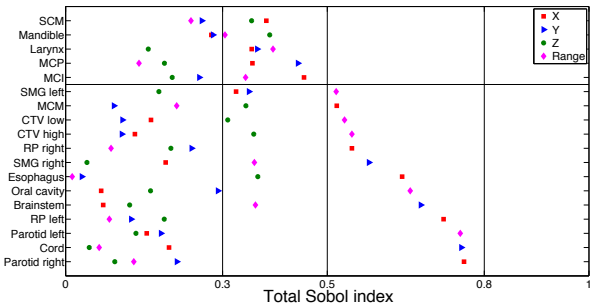
(e) Setup, mean doses



(f) Setup and range, mean doses



(g) Setup, maximum doses



(h) Setup and range, maximum doses

Figure A.18: Sobol' analysis for patient 2 for both setup uncertainties and setup and range uncertainties.