Impact of Gaussian uncertainty assumptions on probabilistic optimization in particle therapy

H.P. Wieser[1,2], **C.P. Karger**[1,3], **N. Wahl**[1,4], **M. Bangert**[1,4]

[1]Department of Medical Physics in Radiation Oncology, German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany
[2] Faculty of Experimental and Medical Physics, Ludwig-Maximilians-Universität München, Am Coulombwall 1, D-85748 Garching, Germany
[3] Heidelberg Institute for Radiation Oncology (HIRO) and National Center for Radiation Research in Oncology (NCRO), Heidelberg, Germany
[4] Heidelberg Institute for Radiation Oncology (HIRO), Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

E-mail: `h.wieser@lmu.de`

August 2019

Abstract. Purpose: Range and setup uncertainties in charged particle therapy may induce a discrepancy between planned and delivered dose. Countermeasures based on probabilistic (stochastic) optimization usually assume a Gaussian probability density to model the underlying range and setup error. While this standard assumption is generally taken for granted, this work explicitly investigates the dosimetric consequences if the actual range and setup errors obey a different probability density function (PDF) over the course of treatment than the one used during probabilistic treatment plan optimization.
Methods: Discrete random sampling was performed for conventionally and probabilistically optimized proton and carbon ion treatment plans utilizing various probability density functions modeling the setup and range error. This method allowed to assess the treatment plan robustness against different probability density functions of conventional and probabilistic plans, which both explicitly assume Gaussian uncertainties. The induced uncertainty in dose was quantified by estimating the expectation value and standard deviation of the RBE-weighted dose for each probability density function on the basis of 2500-5000 random dose samples. Probabilistic dose metrics and standard deviation volume histograms were computed to quantify treatment plan robustness of both optimization approaches.
Results: It was shown that the classical PTV-margin extension concept did not compensate the influence of range and setup errors and consequently resulted in a non-negligible average standard deviation in dose of 7.3 % throughout the CTV. In contrast, probabilistic optimization on normally distributed errors yielded treatment plans that not only entailed a lower standard deviation against normally distributed errors accounted for during optimization but also lower standard deviations for other symmetric PDFs. It was shown that the impact of an incorrect probability distribution assumption is of lower importance after probabilistic optimization as the average uncertainty in the CTV drops to 3.9 %.
Conclusions: Probabilistic optimization is an effective tool to create robust particle treatment plans. Normally distributed range and setup error assumptions for probabilistic optimization are a reasonable first approximation and yield treatment plans that are also robust against other PDFs.

Keywords: intensity-modulated proton therapy, range uncertainties, probability distributions, robust optimization, uncertainty quantification, probabilistic optimization

1. Introduction

Physical uncertainties in charged particle therapy, e.g., range and setup errors may induce a discrepancy between planned and delivered dose and hinder the full exploitation of intensity-modulated particle therapy (Knopf & Lomax 2013). The primary concern in this regard is to avoid underdosage of the target volume and secondly to spare dose to adjacent healthy tissue.

Physical uncertainties induce an uncertainty in deposited dose and originate from the imperfect knowledge of the exact integral stopping power of the charged particle's trajectory. The uncertainty in dose can be attributed to multiple factors, for instance, dose-calculation errors, inaccuracies in the patient's relative stopping powers derived from photon attenuation, CT artifacts, patient positioning errors, and anatomical changes (Lomax 2008a, Lomax 2008b, Paganetti 2012, Graeff 2014). Moreover, the advancement of magnetic spot-scanning systems (Haberer et al. 1993, Krämer & Scholz 2000) in combination with intensity-modulated multi-field optimization (MFO) allows to create highly modulated dose distributions for individual treatment fields. Consequently, the treatment plan becomes more susceptible to setup and range uncertainties. In this context, it is of particular concern that pencil beams with the largest ranges, located at the distal part of the target, apply the highest dose.

To mitigate physical- and beam application-uncertainties in the treatment planning system (TPS), robust beam angles (Cao et al. 2015) and either an isotropic or range error specific expansion of the clinical target volume (CTV) is defined to obtain the planning target volume (PTV) (Van Herk 2004, Park et al. 2012, Knopf et al. 2013, Schuemann et al. 2014a). Two important aspects need to be considered in this regard. First, the margin expansion was originally developed for photons and second it is standard to assume normally distributed errors for the margin expansion. While these measures are generally sufficient to guarantee adequate target coverage for photon irradiation, it may still be problematic for particle irradiation (Unkelbach & Paganetti 2018). In strong contrast to photons, a homogenous dose distribution might not be maintained for the PTV with charged particles since the ion dose deposition is heavily influenced by alterations in the integral stopping power along the beam path - especially within heterogeneous patient anatomies (e.g. lung or cranial tumors).

For moderate to complex patient cases, it is possible to explicitly account for setup and range uncertainties at the state of inverse planning via single-field uniform dose (SFUD), worst-case or probabilistic (stochastic) optimization (Unkelbach et al. 2007, Pflugfelder et al. 2008, Albertini et al. 2010, Fredriksson et al. 2011, Liu et al. 2012, Fredriksson 2012, Bangert et al. 2013, Steitz et al. 2016, Lomax 2016, Lowe et al. 2016, Fredriksson & Bokrantz 2016). An overview is given in (Unkelbach et al. 2018).

Focusing on probabilistic optimization, it is a prerequisite to associate individual error scenarios with an occurrence probability which is different from worst-case optimization approaches where each scenario has equal probability (Unkelbach et al. 2018). In order to assign probabilities to individual error scenarios it is necessary to describe range and setup uncertainties with a probability distribution function (PDF). Therefore, the choice of PDF directly influences probabilistic optimization. Moreover, accurate uncertainty quantification methods either in the form of meta-model approaches (Bangert et al. 2013, Perkó et al. 2016, Wahl et al. 2018, Wieser et al. 2017b) or discrete sampling strategies (Park et al. 2013, Sakama et al. 2016, Wahl et al. 2017) have to make assumptions about the probability distribution of input uncertainties. This clearly shows a dependency of uncertainty quantification and probabilistic optimization on the choice of the PDF representing setup and range uncertainties. In addition, there are also other uncertainty quantification approaches that do not rely on the exact knowledge of the PDF such as dose blurring (van Herk et al. 2002, Baum et al. 2004), min-max estimates (Goitein 1985, Pflugfelder et al. 2008, Fredriksson et al. 2011, Liu 2016) or dose-error-bar distributions which aim to derive

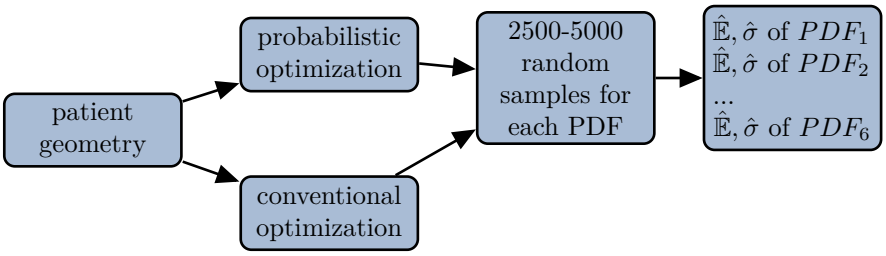

Figure 1: Uncertainty quantification pipeline. For each patient, conventional and probabilistic optimization was carried out assuming a normally distributed PDF. The resulting pencil beam intensities were then used as input for random sampling for six different PDFs. Each PDF yielded a sampling set that allowed in a final step to estimate the expectation value and standard deviation in dose.

the maximum dose error by employing a limited number of dose samples taken from the surface of a spherical error-space (Albertini et al. 2011, Lowe et al. 2016). However, the former two approaches might not reveal the true uncertainty in dose (Ploquin et al. 2006) and the latter - error bar dose distributions - exhibit limitations on a voxel by voxel comparison (Casiraghi et al. 2013).

Only limited studies are available investigating and quantifying not only the tumor site and patient-specific but also machine-specific probability distributions for setup and range uncertainties (Bolsi et al. 2008, Liebl et al. 2014, Holloway et al. 2017). Due to the high effort to quantify PDFs of physical uncertainties based on measured data, all aforementioned approaches usually assume setup and range errors to follow a Normal distribution (Unkelbach et al. 2007, Bangert et al. 2013, Perkó et al. 2016, Wahl et al. 2018, Wieser et al. 2017b). This means that particle centers have to make uncertainty assumptions of quantities with unknown PDF, which can become problematic, if the selected PDF does not reflect reality.

While normally distributed errors are generally taken for granted for the margin expansion as well as for probabilistic optimization due to lack of better knowledge, this study investigates the dosimetric impact when the actual range and setup uncertainties obey a different probability distribution (PDF) over the course of treatment. For this, we quantify the treatment plan robustness for PDFs that differ from the Normal distribution that was originally used for conventional (margin based) and probabilistic optimization.

2. Methods

2.1. Uncertainty Quantification Pipeline

First, a conventional treatment plan was optimized mimicking the clinical standard by considering the PTV as target structure. Here, *conventional* indicates that uncertainties were accounted for by generic margin expansion and not explicitly incorporated into the optimization process. Then, the conventionally optimized pencil beam intensities were used as input for discrete random sampling assuming different probability density functions (PDFs) for range and setup uncertainties. 5000 random samples were computed for one-dimensional examples and 2500 random samples for three-dimensional patient cases to ensure accurate uncertainty estimates in deposited dose for each of the six $PDFs_{1...6}$. The term random sampling refers to the procedure of independently sampling individual treatment planning scenarios from a predefined probability distribution (population) which is characterized by a multivariate PDF.

Next, treatment plan uncertainty quantification was performed by estimating the expectation

value (first non-central moment) and the standard deviation (second central moment) of the RBE-weighted dose from the respective random sampling set. The sequential uncertainty quantification pipeline to determine the treatment plan robustness is presented in Figure 1.

The exact same procedure was now performed for probabilistic optimization on the basis of normally distributed setup and range uncertainties. Thus, Normal distributions were continually assumed for probabilistic optimization. The resulting probabilistically optimized pencil beam intensities were then again considered as input for random sampling assuming each time one of the six different PDFs. This methodology allowed to analyze the robustness of a treatment plan to different realizations of the real uncertainty model upon delivery, which can also be understood as testing the sensitivity of treatment plan against different uncertainty realizations. With the pipeline shown in Figure 1 uncertainties in RBE-weighted dose were quantified for six distinct PDFs on the basis of two different treatment plans (two sets of pencil beam weights), a conventionally and a probabilistically optimized plan.

2.2. Uncertainty model

Setup uncertainties were not modeled in the patient geometry (voxel domain) but in beam's eye view coordinates. Hence instead of shifting the patient, pencil beam positions were shifted in opposite direction to model the setup error and resulting range errors, which is in line with existing worst case optimization approaches. In contrast, range uncertainties were realized by scaling the radiological depths for each pencil beam dose calculation.

Throughout this manuscript random uncertainties were neglected and only systematic uncertainties were modeled to ensure an accurate uncertainty estimation. Accounting only for systematic uncertainties reduced sampling requirements significantly as no interplay with random uncertainties needs to be considered. Although the contribution of random uncertainties in fractionation can be approximated for certain conditions without explicilty sampling each fraction (Unkelbach & Oelfke 2004, Fredriksson 2012, Lowe et al. 2016, Wahl et al. 2018), accurately capturing their interplay with systematic errors in the context of benchmarking requires fractionated sampling, which increases runtimes linearly with the number of fractions.

Figure 2 presents in total six different one-dimensional $PDFs_{1...6}$ that were used to model range uncertainties for uncertainty quantification of conventional and probabilistic optimized treatment plans. As the underlying PDFs of each error source are practically unknown, $PDFs$ with a wide spectrum of different and partly unrealistic shapes were considered for the robustness analysis. In detail, Figure 2 shows the relative range uncertainty with respect to the corresponding range of one particular pencil beam j. Throughout this manuscript the range was defined as the spatial location behind the Bragg peak where the dose deposition reaches 80 % of its maximal value (R80) (Schuemann et al. 2014b). Specifically, Figure 2(a) shows an unbounded Normally distributed range error $p(\Delta^z)$ with mean zero and $\sigma(\Delta^z) = 3.5\,\%$ which we label *standard assumption*. Here, we intend to mimic current clinical range error assumptions. The remaining Figures 2(b)-(f) present other PDFs that were additionally used for uncertainty quantification whereby the uniform, the triangular, and the bimodal bitriangular distribution range from $-7\,\%$ to $7\,\%$. Different σ-values were chosen to cover not only different shapes but also different uncertainty magnitudes. In detail, the uniform distribution represents knowledge of upper and lower error bounds but ignorance about the shape. Thus, the range error can take any value within the bounds with equal probability. The triangular distribution also represents a bounded distribution with an additional assumption about the shape, e.g. the mean value being most likely. Further, the triangular distribution exhibits a pronounced statistical interference (overlap of probability mass) with the Gaussian distribution. Therefore this distribution only slightly deviates from the Gaussian case and mimics consequently clinical realizations close to the standard assumptions,

however, with defined limits.

The most extreme non-Gaussian distribution was given by the bounded bitriangular distribution shown in Figure 2(d) for which a range over and underestimation of 7 % is most likely. This assumption translates to severe systematic density changes along the beam path. Such rather extreme over and undershoot scenarios are clinically employed in scenario-based worst case optimization approaches and can in first approximation be modeled with a bimodal distribution. In order to test the treatment plan robustness against such extreme deviations, also PDFs (e.g. bitriangular) that may be considered rather unrealistic from a clinical standpoint have been included. A range error of 7 % directly corresponds to an occurrence probability of $\leq 4.6\,\%$ $\widehat{=} 1 - P(2\sigma)$ of the Normally distributed range error of Figure 2(a).

Figure 2(e) depicts a skewed Normal distribution, which is the only asymmetric distribution with a non-zero mean value used for uncertainty quantification. Such a distribution might occur if, for instance, the Hounsfield unit to RSP conversion favours a certain direction or density changes in one direction are more likely.

Lastly, Figure 2(f) shows again a Normal distribution, however, this time with $\sigma(\Delta^z) = 5.25\,\%$ and is referred in the manuscript to as broad Normal distribution. The latter case is intended to represent scenarios with large standard deviation and is of particular interest for uncertainty quantification of probabilistic treatment plans on the basis of the standard assumption because it highlights the dosimetric impact for underestimating the magnitude of uncertainties during probabilistic optimization.

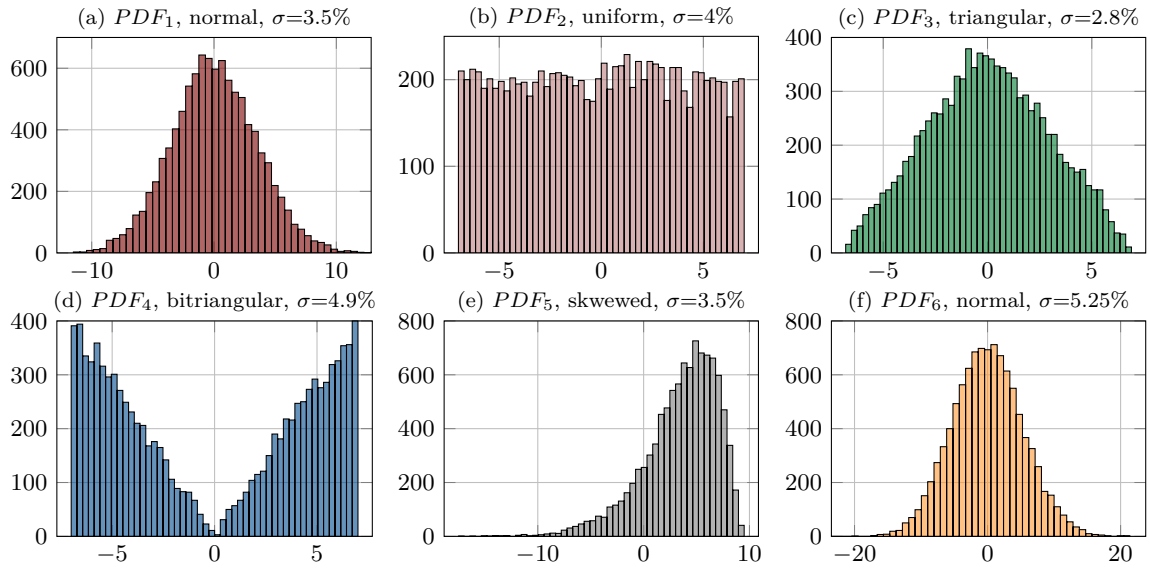

Figure 2: Sampled histograms from various probability density functions (PDFs) modeling the relative range errors $p(\Delta^z)$ of charged particles. The PDFs were used for uncertainty quantification via discrete random sampling.

As treatment plans were comprised of multiple thousand pencil beams J, the corresponding J-dimensional multivariate PDFs were actually employed for discrete random sampling. Assuming multivariate PDFs required the definition of co-variance information for each pencil beam combination($\in \mathbb{R}^{J\times J}$). The co-variance determines in that regard the error-dependence of pencil

beam pairs for a specific uncertainty source. In this study we used a ray-wise correlation model for range uncertainties assuming pencil beams impinging from the same beam orientation at the same lateral x,y-position (= same ray) are perfectly correlated as the same tissue is penetrated. Thus, the relative range error for pencil beams on the same ray behave consistently in the same direction. On the other hand, pencil beam pairs at different x,y-positions (= different rays) were assumed to be completely uncorrelated (Unkelbach et al. 2007, Bangert et al. 2013, Wieser et al. 2017b). This modeling ansatz is in contrast to worst case optimization approaches, which assume fully correlated under and overshoot dose scenarios.

For setup errors, modeled in the lateral x and y dimension, we considered the same distributions from Figure 2, but with different parameters. The standard assumption for the normally distributed setup error featured an absolute standard deviation of $\sigma(\Delta^{x/y}) = 2\,\text{mm}$ whereas the uniform, triangular, and bitriangular distribution ranged from $-4\,\text{mm}$ to $4\,\text{mm}$. Different to range errors modeled via the ray-correlation model, the setup error was assumed to be perfectly correlated for pencil beam combinations belonging to the same beam direction and uncorrelated otherwise (Unkelbach et al. 2007, Bangert et al. 2013, Wieser et al. 2017b). Hence, each beam direction models the setup error independently which is different to single iso-center shifts.

Note that the expression *standard assumption* refers in this manuscript to normally distributed range $\sigma(\Delta^{z}) = 3.5\,\%$ (see Figure 2(a)) and setup errors $\sigma(\Delta^{x/y}) = 2\,\text{mm}$ and was consistently used for probabilistic optimization. Furthermore, range and setup uncertainties caused by anatomical changes such as intra-fractional tumor motion are not explicitly modeled in this study.

2.3. Probabilistic Optimization

To save computation time for sampling, an analytical uncertainty quantification method described in (Bangert et al. 2013, Wahl et al. 2017, Wieser et al. 2017b) was employed to calculate the expectation value ($\mathbb{E}$) and standard deviation (σ) of the RBE weighted dose for protons and carbon ions in closed-form. This approach is called analytical probabilistic modeling (APM) for which we assumed normally distributed physical uncertainties according to the *standard assumption* and the correlation model described in the previous section. In detail, APM is a probabilistic dose calculation that allows to calculate the expected dose influence matrix and a co-variance influence tensor in closed-form which can in a subsequent step be used for probabilistic optimization on the expected objective function value. Probabilistic optimization generally optimizes the expectation of objective functions translating to lower variability in the objective function and therefore dose. It has been shown elsewhere that this methodology can be used similar to discrete sampling based optimization methods to provide uncertainty estimates for the generation of probabilistic treatment plans (Bangert et al. 2013, Wahl et al. 2017, Wieser et al. 2017b, Wahl et al. 2018).

For protons, the physical dose $\boldsymbol{d}$ is used for APM and probabilistic optimization and then ultimately scaled up by $10\,\%$ to consider a RBE of 1.1. Whereas for carbon ions, not the dose $\boldsymbol{d}$ but the biological effect ε was considered for optimization and then converted back to the RBE weighted dose domain (Wieser et al. 2017b). Throughout this manuscript, conventional optimization was carried out considering the PTV as target structure. In contrast, probabilistic optimization assumed the CTV to be the target structure. To compute the RBE-weighted dose for carbon ions we employed a variable RBE on the basis of LEM IV (Scholz et al. 1997, Friedrich et al. 2012) modeling a tissue characterized by $\alpha_x = 0.1\,\text{Gy}^{-1}$ and $0.05\,\text{Gy}^{-2}$ and assumed according to current clinical practice a single biological system with constant linear quadratic model (LQM) parameters for all voxels.

We adapted the basic dose calculation and optimization functionalities for charged particles of the open-source treatment planning system matRad (Wieser et al. 2017a). On the basis of APM

implementations used for (Wieser et al. 2017b), the code was further optimized using MEX-files and parallelization. Moreover, the existing random sampling framework on the basis of multivariate Normal distributions was extended to use the six multivariate PDFs shown for the one-dimensional range error in Figure 2. Treatment plan optimization was performed based on objective and gradient function interfaces to the IPOPT optimization package (Waechter & Biegler 2006).

2.4. Patient cases

The proposed methodology was first tested for protons and carbon ions on a one-dimensional artificial water phantom using an opposing beam configuration to investigate the range error. Figure 3(a) shows the conventionally optimized proton treatment plan for this configuration assuming a RBE of 1.1. The white area represents normal tissue (NT), the light-gray area from 100 mm to 150 mm represents the CTV used for probabilistic optimization and the dark-grey area the margin expansion to obtain the PTV structure which is considered for conventional treatment plan optimization. The red area from 72 mm to 80 mm in 3(b) indicates an OAR. Each proton (carbon ion) beam direction is composed of 33 (45) pencil beams.

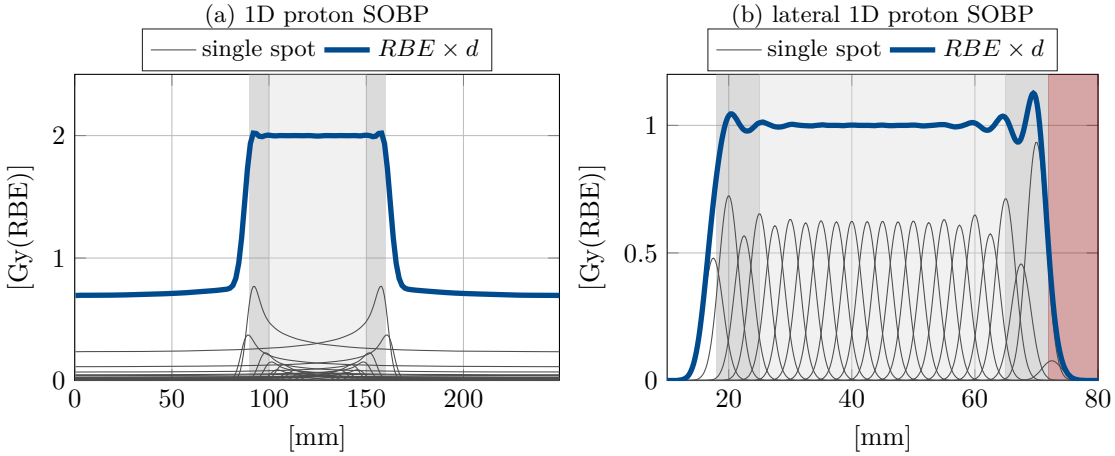

Figure 3: Figure 3(a) presents a one-dimensional proton SOBP assuming two opposing beam directions. The solid blue line denotes the RBE weighted dose assuming a RBE of 1.1. Thin gray solid lines show the physical dose deposition of individual pencil beams. Figure 3(b) presents a separate independent one-dimensional lateral dose profile imitating a lateral proton SOBP. This treatment plan is comprised of 23 pencil beams which are represented by single Gaussian components of varying width. White regions depict normal tissue (NT), light grey regions the CTV, dark gray areas the PTV margin and the red area highlights an OAR.

Figure 3(a) emphasizes the concern of MFO with respect to range uncertainties mentioned in the introduction. Each beam direction primarily covers the distal part of the PTV. Notably the highest modulation can be found in the margin itself thereby introducing a high linear energy transfer (LET) in these regions.

Complementary, the setup error is also investigated in a separate artifical lateral dose profile (see Figure 3(b)). Each lateral pencil beam dose is given by a single Gaussian component. The linearly spaced mean positions and widths of the single Gaussian components were fixed and an inverse optimization was carried out to find weights/intensities that sufficiently cover the PTV with 1 Gy(RBE). The OAR causes an asymmetry of the optimized lateral dose profile due to

different trade-offs during optimization. In particular, a higher dose gradient on the PTV/OAR interface can be observed because of the higher penalty on the OAR compared to NT. In this configuration, setup uncertainties were modeled by being uncertain in the lateral position of each Gaussian component. Further, all pencil beams originated in this example from one single beam direction and are assumed to be perfectly correlated. Figure 3(b) was simulated separately from the depth dose SOBP example in Figure 3(a).

Since the one-dimensional results of protons and carbon ions were comparable, the second part of the analysis was restricted to three-dimensional proton treatment plans only. Figure 4 presents the nominal RBE weighted dose distributions on transversal slices next to the iso-center of three proton treatment plans. These treatment plans are based on conventional optimization (conv.opt.) considering the PTV as target structure. In Figure 4 the solid black contour represents the CTV and the solid pink contour the corresponding critical organ at risk. The corresponding patient and treatment plan parameters are summarized in the Appendix in Table 5.

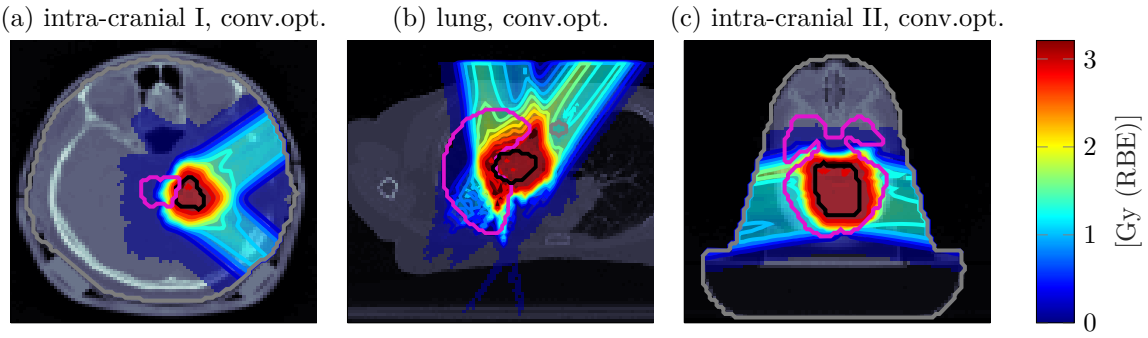

Figure 4: Three-dimensional intensity-modulated proton treatment plans with two angular incident beams in (a) and (b) and an almost opposing beam configuration in (c). conv.opt indicates conventional optimization and hence does not explicitly account for setup and range uncertainties during optimization. Each panel shows the transversal iso-center slice of the CT scan overlaid by transparent dose colorwash and solid isodose lines. The solid black contour represents the CTV and the solid pink contour the corresponding critical organ at risk.

2.5. Robustness Analysis

To quantify the dosimetric impact of different PDFs on conventional and probabilistic treatment plans, six different probabilistic dose metrics were calculated for the CTV. The six probabilistic dose metrics are shown in Table 1 and have been calculated for each $PDF_{1...6}$ based on 5000 samples for one-dimensional and 2500 samples for three-dimensional patient cases. Computing these quantities for conventional and probabilistic treatment plans allowed then to compare treatment plan robustness. Exemplary, $\mu[\sigma]$ denotes the average value of the standard deviation of the RBE weighted dose and consequently represents the mean uncertainty level for a specific structure (CTV or OAR).

Moreover, we also evaluated for the patient cases the probabilistic dose metrics for the organ at risk represented by the pink contours in Figure 4.

In addition, standard deviation volume histograms (SDVH) were calculated for three-dimensional patient cases to consider a further metric for analyzing the treatment plan robustness.

Table 1: Probabilistic dose metrics used to quantify treatment plan robustness. These quantities always refer to a specific volume of interest. For the last two rows, note that 95 % of the prescribed dose corresponds to a dose level of 1.9 Gy(RBE) for the one-dimensional SOBP and 2.85 Gy(RBE) for the three-dimensional patient cases.

metric	unit	meaning
$\mu[\mathbb{E}]$	[Gy(RBE)]	average value of the expected $RBExD$
$\mu[\sigma]$	[Gy(RBE)]	average value of the standard deviation of the $RBExD$
$\mu[D95]$	[Gy(RBE)]	average dose given to 95 % of the volume
$\sigma[D95]$	[Gy(RBE)]	standard deviation of dose given to 95 % of the volume
$\mu[V95]$	[%]	average volume receiving 95 % of dose
$\sigma[V95]$	[%]	standard deviation of the volume receiving 95 % of dose

The SDVH allowed an in-depth comparison of standard deviations originating from different uncertainty assumptions and optimization techniques.

3. Results

3.1. One-dimensional Spread Out Bragg Peak

Figure 5 illustrates the nominal dose together with the corresponding standard deviation for the PTV based conventional optimization (conv.opt) by solid solid lines and CTV-based probabilistic optimization (prop.opt) outcome by dotted lines assuming identical penalties for both optimizations. The light grey shaded area depicts the CTV and the dark grey area the PTV margin. Conventional optimization aims for steep dose gradients at the PTV-tissue interfaces whereas probabilistic optimization, which explicitly models range uncertainties, aims for shallower dose gradients to reduce the variance in the CTV. As a consequence, a lower standard deviation highlighted by the red dotted line can be observed in the case of probabilistic optimization for the price of an increased entrance dose. The steepness of the dose gradient is determined by both the uncertainty assumption and the relative importance relation of target and normal tissue. This becomes clear, when defining a higher penalty on the NT for probabilistic optimization (prop.opt2), which results in dash dotted profiles in Figure 5. The entrance dose equals the one of a conventional PTV based optimization, however, the uncertainty in dose can only be mitigated to a minor extend. This optimization result (prop.opt2) is not considered for the remaining analysis but highlights the influence of the chosen penalty on the optimization outcome.

Figure 6 presents the comparison of statistical moments for different proton range uncertainty assumptions for the homogenous one-dimensional phantom case. In detail, Figure 6(a) shows the expectation value ($\mathbb{E}[RBE \times d]$) of the one-dimensional intensity-modulated proton Spread Out Bragg peak (SOBP) on the basis of conventional optimization (conv.opt.) assuming different PDFs. The nominal dose profile covering the whole PTV is not shown here but in Figure 3. Complementary, Figure 6(c) illustrates the corresponding standard deviations ($\sigma[RBE \times d]$). Because of using an opposing beam configuration for the homogenous phantom, only minimal discrepancies can be observed for the $\mathbb{E}[RBE \times d]$ on the basis of different PDFs.

Moreover, Figures 6(b) and 6(d) depict the results obtained from random sampling on the basis of pencil beam intensities derived from probabilistic optimization (prob.opt). Thin solid gray lines in Figures 6(a) and 6(c) denote the dose distributions of individual proton pencil beams. Due to probabilistic optimization, the following observations can me made from Figure 6: (i) the modulation of individual pencil beams in Figure 6(b) is less pronounced compared to Figure 6(a),

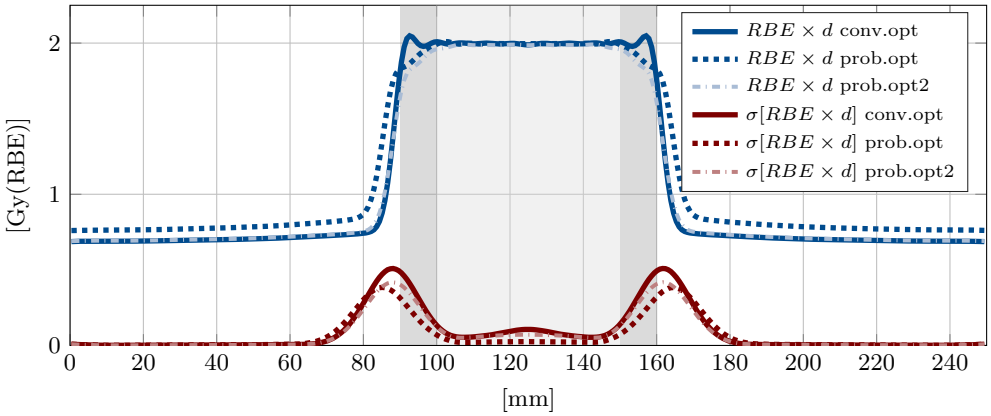

Figure 5: Nominal dose ($RBE \times d$) and standard deviation ($\sigma[RBE \times d]$) of conventional PTV based optimization (solid lines) and probabilistic optimization (dotted lines) using the same penalties for the one-dimensional proton SOBP in water introduced in 3(a). Dashed dotted lines are based on probabilistic optimization with an increased penalty on the normal tissue (white area). The light grey area depicts the CTV and the dark grey area the PTV margin expansion.

(ii) the dose gradients of $\mathbb{E}[RBE \times d]$ in the fall off region are lower in 6(b) than in 6(a) and (iii) the dose level at the entrance regions is higher in 6(b) than in 6(a). The latter can be explained by the fact of using a relatively low penalty for the surrounding tissue. Therefore, the optimizer favours increased dose to surrounding tissue for a lower standard deviation in the target in return.

Table 2: The probabilistic dose metrics are described in Table 1 and are given here for the CTV. The upper half of the table (range error) shows the robustness analysis of the one-dimensional proton SOBP and the lower half of the table refers to the robustness analysis of the one-dimensional lateral dose profile modeling a setup error. The corresponding dose profile can be found in Figure 7. Columns 2-7 show the probabilistic dose metrics for conventional optimization and columns 8-13 are based on probabilistic optimization. The columns $\mu[\mathbb{E}]$, $\mu[\sigma]$, $\mu[D95]$ and $\sigma[D95]$ are given in [Gy(RBE)] whereas $\mu[V95]$ and $\sigma[V95]$ are given in [%].

range error	**conv.opt.**						**prob.opt.**					
PDF	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[D95]$	$\sigma[D95]$	$\mu[V95]$	$\sigma[V95]$	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[D95]$	$\sigma[D95]$	$\mu[V95]$	$\sigma[V95]$
normal	1.99	0.089	1.91	0.17	92.8	14.4	1.99	0.036	1.93	0.091	98.1	3.5
uniform	1.99	0.087	1.91	0.10	89.2	18.7	1.99	0.030	1.93	0.036	97.2	3.2
triangular	1.99	0.057	1.96	0.06	95.8	9.9	1.99	0.021	1.95	0.026	99.2	1.6
bitriangular	1.99	0.108	1.87	0.12	83.2	24.7	1.99	0.035	1.91	0.039	95.7	3.1
skewed	1.91	0.081	1.82	0.13	61.3	32.0	1.96	0.034	1.90	0.081	93.5	6.7
broad normal	1.97	0.016	1.77	0.32	83.2	24.2	1.98	0.081	1.84	0.233	94.3	7.5
setup error	**conv.opt.**						**prob.opt.**					
normal	0.99	0.033	0.90	0.14	97.19	2.4	0.98	0.024	0.89	0.13	96.1	2.4
uniform	0.98	0.034	0.88	0.13	96.41	2.3	0.98	0.024	0.86	0.14	95.43	2.3
triangular	0.99	0.026	0.94	0.07	97.67	1.8	0.98	0.017	0.91	0.09	96.52	1.9
bitriangular	0.97	0.038	0.82	0.16	95.12	1.9	0.98	0.029	0.80	0.17	94.22	2.2
skewed	0.97	0.04	0.76	0.25	95.17	2.7	0.97	0.028	0.72	0.22	93.56	2.8
broad normal	0.97	0.056	0.79	0.27	95.64	3.5	0.98	0.044	0.80	0.23	94.63	3.5

It can be seen from Figures 6(a) and (c) that different PDFs have a distinct impact on the resulting standard deviation σ. The main observations made from Figure 6 are that probabilistic

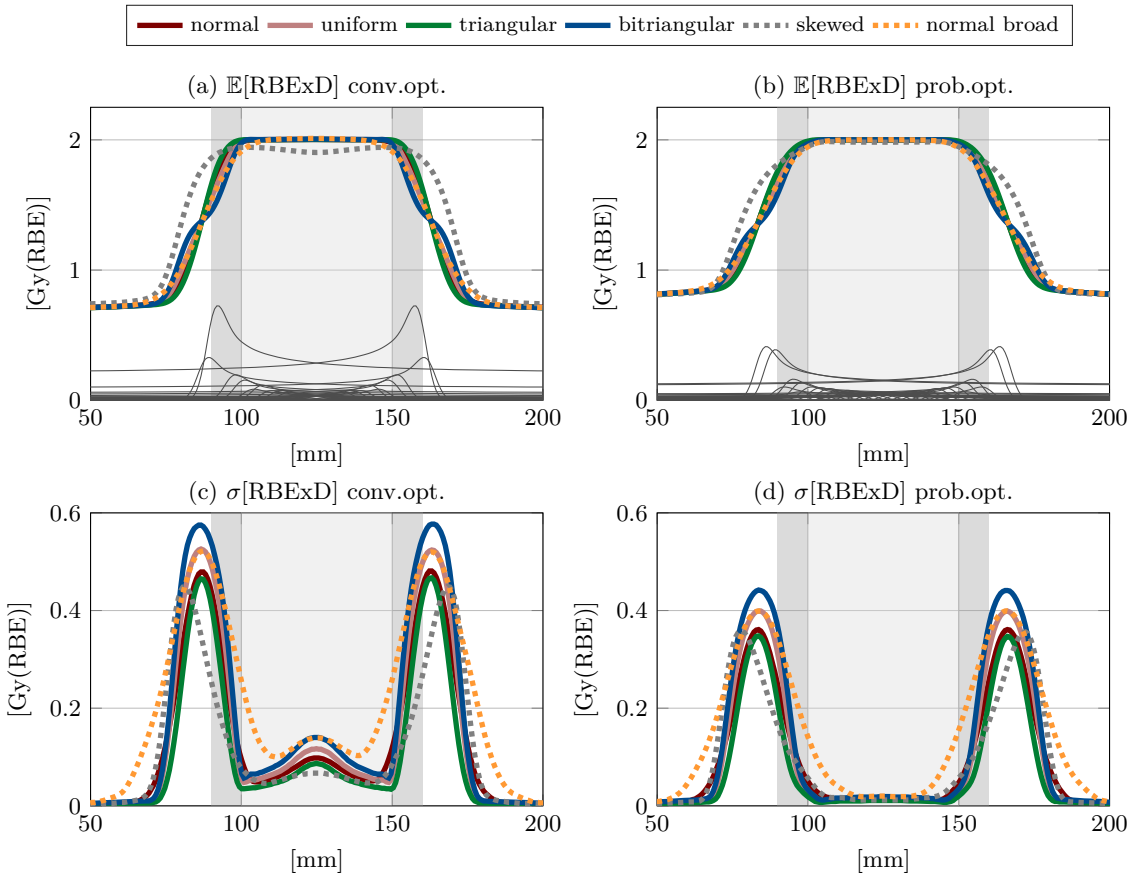

Figure 6: Statistical moments ($\mathbb{E}[RBE \times d]$, $\sigma[RBE \times d]$) of a one-dimensional proton SOBP in water. Figures (a) and (b) show the expectation value and Figures (c) and (d) the standard deviation of the RBE weighted dose. Different probability density functions were employed for random sampling to model the relative range errors of protons. Figures (a) and (c) are grounded on pencil beam intensities obtained from conventional optimization on the PTV and Figures (b) and (c) are based on probabilistically optimized pencil beam intensities on the CTV.

optimization yields lower absolute standard deviations bumps in high dose gradient regions and creates a lower uncertainty level throughout the CTV. The most significant influence can be observed in Figure 6 when range errors were modeled with a skewed, bitriangular and a broad Normal distribution. Interestingly, in this one-dimensional opposing proton beam configuration, the underdosage of the CTV assuming a skewed PDF is partly compensated by probabilistic optimization. Focusing on the standard deviations ($\sigma[RBE \times d]$) in Figures 6(c) and 6(d) then two aspects can be observed. First, the standard deviation at high dose gradient regions at ∼90 mm and ∼160 mm was lower for all PDFs after probabilistic optimization but the spread of $\sigma[RBE \times d]$ profiles in these regions remained similar. Second, when analyzing the standard deviation profiles in the light grey CTV region from 100 mm to 150 mm in Figures 6(c) and 6(d), a lower variability between different profiles can be observed due to probabilistic optimization.

In line with Figure 6, these observations also manifest in Table 2 which presents probabilistic

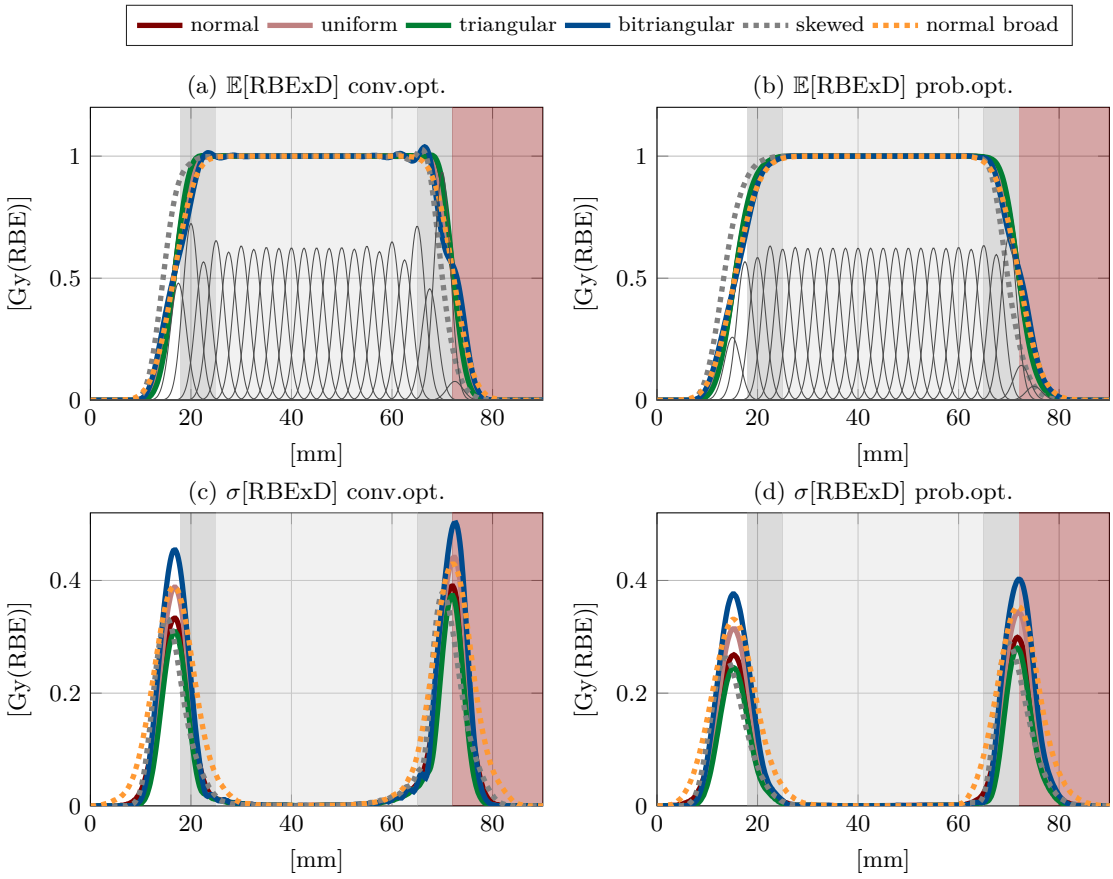

Figure 7: Lateral SOBP profile of a single beam direction. Different probability density functions (PDFs) were used to model the setup error of charged particles. Figures 7(a) and (c) show statistical moments obtained from conventional optimization and Figures 7(b) and (d) relate to probabilistic optimization.

dose metrics for this setup. Focusing on the average value of the expected RBE weighted dose $\mu[\mathbb{E}]$ shows that both conventional and probabilistic optimization sufficiently cover the target for the homogenous one-dimensional phantom even for different PDFs. Only for the skewed Normal distribution a more pronounced under-dosage can be observed. However, the upper half of Table 2 indicates the skewed, bitriangular and the broad Normal distribution disturb the optimized one-dimensional proton treatment plans the most. Notwithstanding, their impact was minimized by probabilistic optimization as for instance indicated by $\sigma[V1.9]$ which dropped from 24.7, 32 and 24.2 % to 3.1, 6.7 and 7.5 %, respectively. Furthermore, $\mu[\sigma]$ in Table 2 reveals that the uncertainty level throughout the CTV is reduced after probabilistic optimization by factor $\sim$2. In general, Table 2 demonstrates that absolute numbers of average quantities $\mu[\mathbb{E}]$, $\mu[D95]$ and $\mu[V1.9]$ tend to be higher after probabilistic optimization and standard deviation quantities $\sigma[D95]$ and $\sigma[V1.9]$ tend to be smaller.

Probabilistic dose metrics for the one-dimensional lateral proton SOBP are shown in the lower half of Table 2. Here, the results are analogous but are compared to the range error less pronounced indicating the PTV margin expansion covers the CTV in this single beam direction simulation

sufficiently against different error realizations. The underlying conventional and probabilistic dose profiles can be seen in Figure 7. As all pencil beams are considered to be perfectly correlated expressing coherent beam shifts in water, the standard deviation profiles drop down to zero inside the CTV. In contrast, the maximum standard deviation can be observed in the lateral dose fall off region around 20 mm and 70 mm (see Figure 7(c) and (d)). Due to a different trade-off during optimization, the OAR on the right side of the PTV causes the standard deviation maxima at 70 mm to be higher than at 20 mm. The integral dose to the OAR after conventional optimization results in 1.02 Gy mm and to 1.17 Gy mm for probabilistic optimization. The analysis of the NT results in 1.77 Gy mm and 2.7 Gy mm, respectively. Large penalty differences of adjacent volumes (e.g., normal tissue and CTV) yields after probabilistic optimization to a decrease of dose uncertainties by reducing dose gradients which itself entails an increase of integral dose to the NT. Contrary, similar penalties (e.g., OAR and CTV) limits the reduction of dose uncertainties.

For case presented in Figure 7 probabilistic optimization reduces the weights of the outer Gaussian components thereby creating lower dose gradients. This leads to a reduction of standard deviation of up to 0.1 Gy which corresponds to 10 %.

Due to similar Bragg peak characteristics, similar observations were made for the one-dimensional carbon ion SOBP which can be found in the Appendix in Figure 12 and Table 6.

3.2. Three-dimensional proton patient cases

This result section focuses first on presenting the uncertainty in dose by means of standard deviations plots to illustrate the impact of different PDFs on the uncertainty in dose. Hence, Figure 8 shows a transversal CT slice next to the iso-center overlaid by a transparent dose colorwash representing the standard deviation ($\sigma[RBE \times d]$) obtained from discrete random sampling on the basis of the PDFs illustrated in Figure 2. Due to range uncertainties the standard deviation is peaked at the boarder between the CTV and brainstem. All the underlying pencil beam intensities for Figure 8 were obtained from the same conventional optimization on the PTV. It can be observed from Figure 8 that the standard deviation reaches ∼20 % of the prescribed dose of 3 Gy(RBE). Analog to the one-dimensional example, each PDF caused a distinct standard deviation distribution. The bitriangular PDF highlighted in Figure 8(d) and the broad Normal PDF illustrated in Figure 8(f) induced, compared to normally distributed errors in Figure 8(a), the highest uncertainty in RBE weighted dose. Summarized, all uncertainty quantifications on the basis of conventional optimization revealed a non-negligible uncertainty in dose throughout the CTV.

Next, probabilistic treatment plan optimization on the basis of normally distributed uncertainties according to the *standard assumption* was carried out. The resulting pencil beam intensities were then considered for discrete random sampling. The resulting standard deviations are presented in Figure 9. Note that Figure 9(a) shows the standard deviation ($\sigma[RBE \times d]$) of the *standard assumption* which was used for probabilistic optimization. Comparing the uncertainty in dose from conventional optimization in Figure 8(a) to the result obtained from probabilistic optimization in Figure 9(a), it becomes clear that the latter treatment plan exhibits a reduced uncertainty in dose throughout the CTV. Analyzing the standard deviations of the remaining PDFs in Figures 9(b)-(f) showed that each PDF caused mainly different standard deviations around the CTV and reached the highest absolute values for the bitriangular and the broad Normal distribution. However, the uncertainty level inside the CTV in Figure 9 remained almost unaffected for all PDFs. Hence probabilistic optimization on normally distributed physical errors created intensity-modulated treatment plans that revealed a lower standard deviation in dose throughout the CTV against other PDFs. Comparable observations could be made for the standard deviations of the lung and intra-cranial II patient case which are presented separately in the Appendix in

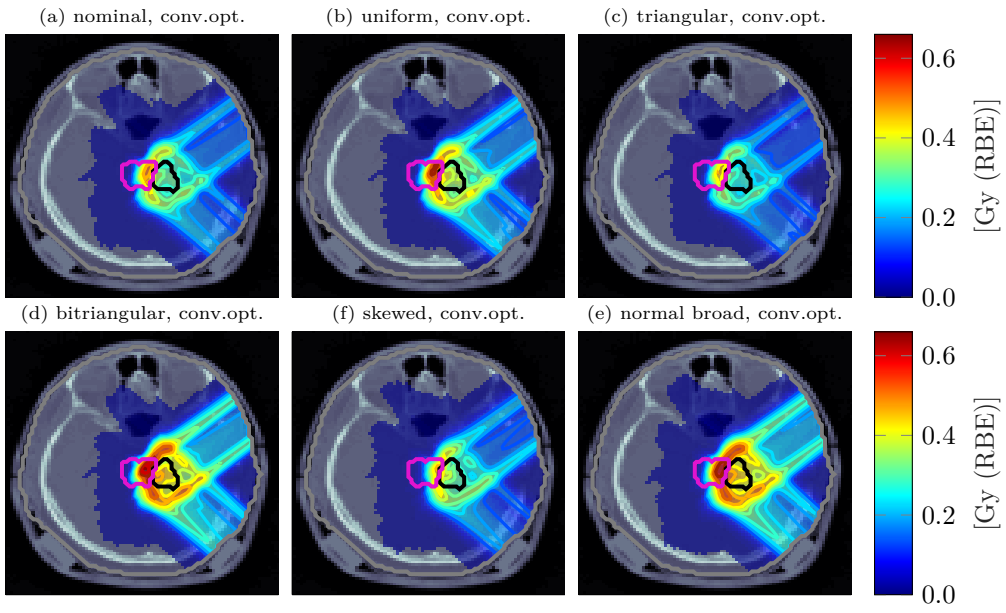

Figure 8: Transversal CT next to the iso-center of the intra-cranial case superimposed by transparent dose colorwash representing the standard deviation assuming different PDFs. Solid lines indicate iso-dose lines whereas the solid black contour denotes the CTV and the pink contour the brainstem. Pencil beam intensities are based on conventional optimization on the PTV (not shown). The two proton beam direction with gantry angles 50° and 135° were utilized.

Figures 13 and 14. Although in Figure 9 the bitriangular PDF caused the highest uncertainty in dose, the almost opposing beam configuration for the intra-cranial II patient case in Figure 14(d) resulted in relatively low range uncertainties also for this extreme distribution.

The corresponding probabilistic dose metrics for the CTV of all three patient cases are presented in Table 3. Analyzing $\mu[\mathbb{E}]$, the average value of the expected $RBE \times d$, indicates for conventional optimization a slight to moderate underdosage of the CTV. Performing probabilistic optimization partly compensates this effect as $\mu[\mathbb{E}]$ is shifted back closer to the prescribed dose of 3 Gy(RBE). Focusing on the average standard deviation in the CTV highlights the advantageous influence of probabilistic optimization on the treatment plan robustness. Similar to the one dimensional scenario, the average $\mu[\sigma]$ of all three patient cases yields 7.3 % (average of the third column in Table 3). This number drops to 3.9 % after probabilistic optimization. Moreover, the combined consideration of $\mu[V95]$ and $\sigma[V95]$ provides detailed insight into treatment plan robustness. Not only is $\mu[V95]$ higher after probabilistic optimization but also the spread of $V95$ is reduced.

In line with the one-dimensional examples, a greater influence on the treatment plan robustness was identifiable in Table 3 for the bitriangular, skewed and the broader Normal PDF. Particularly the asymmetric skewed Normal distribution decreased $\mu[\mathbb{E}]$ down to 2.80 Gy(RBE) as a range undershoot was much more likely due to asymmetry and caused consequently an underdosage of the CTV.

Probabilistic dose metrics for the OAR of each respective case are shown in Table 4. Comparing the average expectation value $\mu[\mathbb{E}]$ between the two optimization approaches exhibits a decrease in dose for probabilistic optimization. In addition, also a moderate reduction of the standard

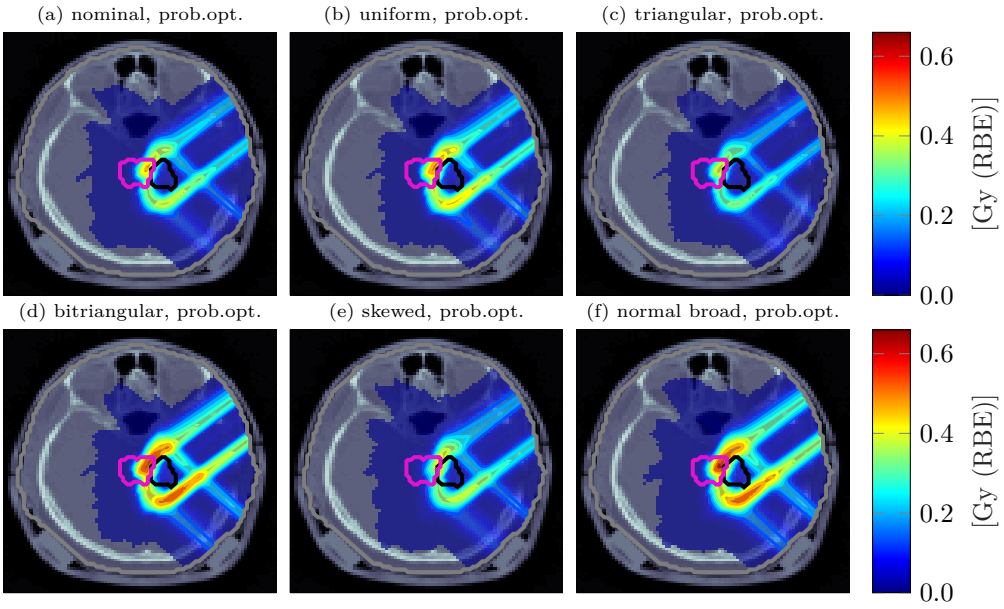

Figure 9: Uncertainty quantification result on the basis of probabilistic optimization. Transversal CT image close to the iso-center of the intra-cranial case superimposed by transparent dose colorwash indicates the standard deviation of the RBE-weighted dose on the basis of various PDFs.

deviation in dose can be observed indicating the dose received by the OAR is not only smaller in absolute numbers but also varies to a lesser extend. These effects can also be observed in Figures 8 and 9 when focusing on the brainstem, which is represented by the pink contour.

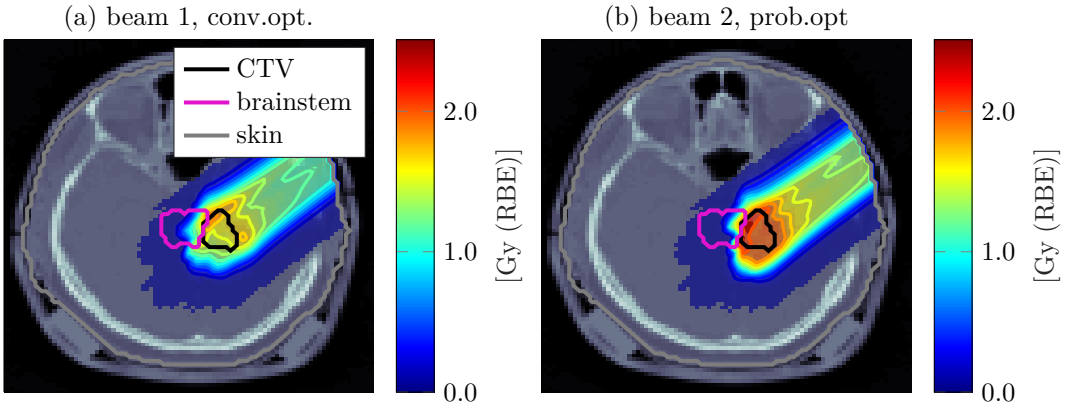

Figure 10: RBE-weighted dose distribution of the first beam based on conventional optimization on the PTV in Figure 10(a) and probabilistic optimization on the CTV in 10(b).

Figure 10 presents the beam dose distributions of the first beam from the conventional PTV based optimization in 10(a) and from probabilistic optimization on the CTV in 10(b) employing the *standard assumption*. This result shows that more dose is given by the first beam direction of

Table 3: Robustness analysis of the CTV by means of probabilistic dose metrics for the head, lung and cranial patient case from Figure 4 assuming different PDFs. Statistical moments were estimated using 2500 discrete random samples. The abbreviation conv.opt denotes conventional optimization and prob.opt. depicts probabilistic optimization. The columns $\mu[\mathbb{E}]$, $\mu[\sigma]$, $\mu[D95]$ and $\sigma[D95]$ are given in [Gy(RBE)] whereas $\mu[V95]$ and $\sigma[V95]$ are given in [%].

intra-cranial I	**conv.opt.**						**prob.opt.**					
PDF	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[D95]$	$\sigma[D95]$	$\mu[V95]$	$\sigma[V95]$	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[D95]$	$\sigma[D95]$	$\mu[V95]$	$\sigma[V95]$
normal	2.97	0.270	2.64	0.21	68.1	29.2	2.99	0.105	2.68	0.2	88.8	6.8
uniform	2.93	0.311	2.52	0.22	60.9	30.1	2.95	0.121	2.52	0.21	81.9	7
triangular	2.98	0.229	2.7	0.17	72.8	26.5	3.00	0.085	2.76	0.15	92	4.7
bitriangular	2.88	0.373	2.37	0.24	54.6	30.6	2.93	0.151	2.3	0.25	71.5	7.9
skewed	2.85	0.296	2.21	0.24	55.8	24.8	2.89	0.163	1.83	0.28	69	6
broad normal	2.88	0.389	2.37	0.3	55.5	30.9	2.89	0.181	2.35	0.31	72.5	12.2
lung	**conv.opt.**						**prob.opt.**					
normal	2.97	0.165	2.71	0.19	82.3	16.7	3.00	0.088	2.85	0.1	94.9	5.6
uniform	2.96	0.190	2.64	0.22	77.4	18.5	2.99	0.108	2.78	0.13	91.7	6.1
triangular	2.99	0.135	2.76	0.14	86.4	13.3	3.01	0.079	2.87	0.07	96.1	3.6
bitriangular	2.94	0.228	2.54	0.26	71.8	20.1	2.97	0.131	2.67	0.18	86.7	7.4
skewed	2.80	0.203	2.24	0.27	56.6	17.9	2.95	0.134	2.45	0.25	81.8	8.1
broad normal	2.94	0.243	2.55	0.29	73	20.9	2.98	0.136	2.7	0.22	88	9.7
intra-cranial II	**conv.opt.**						**prob.opt.**					
normal	2.98	0.127	2.74	0.09	87.2	6	2.99	0.086	2.71	0.2	92.2	3.7
uniform	2.97	0.151	2.69	0.08	81.9	5.9	2.97	0.109	2.63	0.2	89.4	3.7
triangular	2.99	0.113	2.78	0.06	89.7	4.6	3.00	0.078	2.78	0.14	93.9	2.5
bitriangular	2.95	0.177	2.59	0.09	75.6	6.2	2.96	0.134	2.49	0.25	84.4	4.3
skewed	2.91	0.153	2.54	0.11	70.5	7.1	2.96	0.118	2.57	0.15	83.3	5.7
broad normal	2.94	0.191	2.56	0.18	76.3	8.9	2.95	0.137	2.5	0.3	84.6	7.5

the probabilistic plan than from the conventional plan. Further, the beam dose of the probabilistic plan is more homogenous except for the high-dose region close to the brainstem, which is preferably covered by the first beam direction rather than the second. To avoid giving a high dose to the tumor area close to the brainstem from the second beam direction, which is associated with higher range uncertainties, probabilistic optimization favours the solution of putting higher pencil beam weights on the first beam direction.

Table 4: Robustness analysis of organs at risk. For each patient case, a single organ at risk was used to compute the average expectation value $\mu[\mathbb{E}]$ and the average standard deviation in dose $\mu[\sigma]$. All columns are given in [Gy(RBE)].

	brainstem				**lung**				**brain**			
	conv.opt.		**prob.opt.**		**conv.opt.**		**prob.opt.**		**conv.opt.**		**prob.opt.**	
PDF	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[\mathbb{E}]$	$\mu[\sigma]$
normal	0.39	0.138	0.27	0.101	0.44	0.091	0.3	0.06	0.22	0.101	0.08	0.054
uniform	0.40	0.161	0.29	0.117	0.44	0.104	0.3	0.069	0.22	0.115	0.09	0.059
triangular	0.39	0.120	0.27	0.085	0.44	0.078	0.3	0.05	0.21	0.082	0.08	0.039
bitriangular	0.42	0.189	0.30	0.138	0.43	0.122	0.3	0.082	0.24	0.144	0.10	0.075
skewed	0.24	0.11	0.16	0.076	0.36	0.087	0.26	0.06	0.27	0.117	0.13	0.07
broad normal	0.4	0.194	0.29	0.143	0.44	0.126	0.3	0.087	0.24	0.157	0.10	0.089

The standard deviation volume histogram (SDVH) was used as a second robustness evaluation method. The following analysis only focused on the bitriangular, skewed and broad Normal distribution as the corresponding probabilistic dose metrics deviate in Table 3 the most from dose

metrics assuming a Normal distribution and are consequently the most interesting configurations.

Figure 11(a) shows the respective SDVHs for the CTV of the intra-cranial I case. The solid blue line represents the reference result obtained by conventional optimization on the PTV and the dashed blue line denotes the robust reference based on probabilistic optimization utilizing the *standard assumption*. The red, green and orange line denote the SDVH from a skewed, bitriangular and a broader normal PDF. All three PDFs induce for pencil beam intensities derived from conventional optimization (solid lines) an increased standard deviation in dose compared to the solid blue reference line.

A clear separation of SDVH profiles can be observable in Figure 11(a). Interestingly, the standard deviation profiles in Figure 11 are notably lower after probabilistic optimization (dashed lines). The dashed blue line stems from a probabilistic optimization using the standard assumption, for which reason the same uncertainty assumption has been used for optimization and sampling. Hence, this configuration reveals the lowest SDVH profile of all. Furthermore, dashed lines which are based on probabilistic optimization show a lower intra-variability than the counterpart obtained from conventional optimization (solid lines). Moreover, Figure 11(a) illustrates that the probabilistically optimized treatment plan entails mostly lower SDVHs, even for different PDFs used for re-sampling, than the solid blue reference originating form conventional optimization.

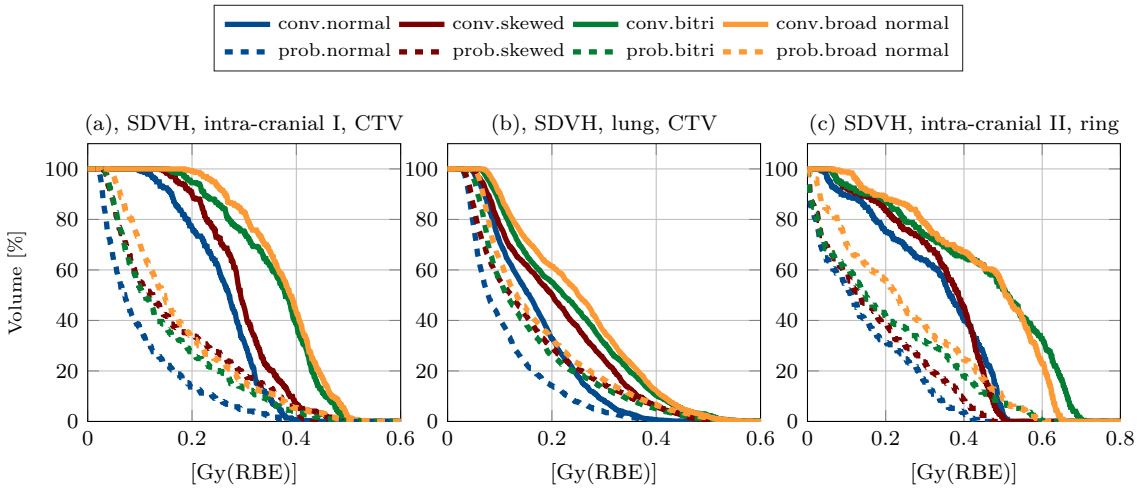

Figure 11: Standard deviation volume histogram (SDVH) of the intra-cranial I, lung and intra-cranial II patient case. Solid lines are based on conventionally optimized pencil beam intensities whereas dashed lines refer to probabilistically optimized pencil beam intensities.

Figure 11(b) highlights the SDVH for the CTV of the lung patient case. The SDVH profiles were not as clearly separated as before in the intra-cranial I case from Figure 11(a) but the main tendencies can still be observed. Despite not specifically optimizing for a skewed, bitriangular or broader normal distribution the resulting standard deviation from probabilistic optimization are substantially lower to that of conventional optimization. Further, all dashed lines indicate for the most part of the CTV Volume from 30 % to 100 % a lower uncertainty in dose than conventional optimization assuming the *standard assumption* (solid blue line).

Lastly, Figure 11(c) shows the SDVH for the ring structure of the intra-cranial II patient case. The ring structure was not explicitly plotted in Figure 4 as it was derived from a simple isotropic expansion of the CTV by 15 mm. Figure 11(c) demonstrates the robustness of probabilistic

optimized treatment plans against other PDFs assumptions. Again two groups of profiles were distinguishable in Figure 11(c). First, SDVH profiles originating from probabilistically optimized pencil beam intensities represented by dashed lines are significantly lower to the uncertainty in dose stemming from conventional optimization as indicated by solid lines.

4. Discussion

In this manuscript, we investigated the robustness of conventional and probabilistic charged particle treatment plans assuming various probability density functions (PDFs) to model range and setup uncertainties.

For each of the six PDFs discrete random samples were calculated, followed by an estimation of the expectation value and standard deviation of the RBE-weighted dose. In a further step, probabilistic optimization was carried out considering the *standard assumption* for setup and range uncertainties because the Normal distribution is acknowledged to be the standard probability distribution in radiotherapy. With APM it is in principle possible to consider other non-Gaussian PDFs either via a direct analytical description or by representing a specific PDF with a superposition of Gaussian components. The resulting pencil beam intensities were then as the robust reference plan and were in turn used for random sampling to estimate moments from other PDFs. With this pipeline, we could investigate the dosimetric impact if the PDF during irradiation is different to the one used for treatment planning.

Random sampling for one case took up to six hours per PDF within matRad using a regular desktop PC and Matlab's parallelization capabilities. The setup error was modeled with a beam-wise correlation model. If more than two beam directions are used, setup error do not compensate due to the high modulated field doses of different beam directions. These setup assumptions are different to a perfect correlation model which corresponds to single iso center shifts. If range errors were modeled as perfectly correlated, which corresponds to consistent under- and overshoot scenarios, the resulting uncertainty in dose would be higher than for the utilized ray-wise correlation model. Therefore it has to be emphasized that the utilized correlation model for setup and range errors impacts the uncertainty in dose for which reason the results cannot be transferred one-to-one to other correlation assumptions.

The uncertainty in dose was analyzed first for one-dimensional artificial examples mimicking a lateral and a depth SOBP dose profile. Each PDF caused a distinct expectation value and standard deviation profile and similar characteristics between protons and carbon ions could be observed.

In these simple geometries we saw for both examples after probabilistic optimization an increased coverage of the CTV and a lower uncertainty in dose. This statement is especially true for the CTV structure and was more pronounced for the depth SOBP dose profile than for the lateral dose profile. Since protons and carbon ions have similar Bragg peak characteristics similar conclusions could be drawn from the one-dimensional proton and carbon ion treatment plan. The latter is shown in the Appendix in Figure 12.

In a further step, multiple three-dimensional proton treatment plans were investigated. For conventional optimization utilizing the PTV-margin concept, different PDFs induced different expectation value distributions and standard deviation distributions. A non-negligible level of uncertainty could be observed for all PDFs indicating that setup and range errors are not sufficiently compensated by the PTV margin. Performing probabilistic optimization on normally distributed range and setup errors (*standard assumption*) not only minimized the uncertainty in dose but also reduced the spread of statistical moments originating from different PDFs.

The simplified 1D treatment plan from Figures 5 and 6 yielded shallower dose gradients and a higher dose to the surrounding tissue after probabilistic optimization. The low penalty on the objective function limiting dose to the surrounding tissue causes dose to be redistributed to

less critical normal tissue. The penalty ratio defines the trade-off between more dose to normal tissue/OAR and lower uncertainty in dose. The situation in 3D patient cases is different, as a higher relative penalty is used for the OAR forcing the optimizer to project the standard deviation to less important normal tissue regions around the CTV. As a result the expected dose and standard deviation is reduced in the OAR, which is in line with previous findings (Fredriksson et al. 2011, Liu et al. 2012, Unkelbach et al. 2018). The probabilistic treatment plan reduces the impact of range under and overshoot scenarios which consequently yields on average to a lower OAR dose. Therefore, the normal tissue complication probability (NTCP) becomes lower which might allow for a more advantageous balancing of tumor control probability (TCP) and NTCP due to a broader therapeutic window. But also for probabilistic and worst case optimization there is a trade-off between target coverage and OAR sparing. It depends on the very choice of objective function (incl. penalty) and the uncertainty assumption how this trade-off is used during optimization (Mescher et al. 2017, Arts et al. 2017, van der Voort et al. 2016).

In general, an optimal treatment occurs per design, if the same error distribution is present for both treatment planning and the actual patient irradiation. In contrast, if differences between the error distributions occur, then the resulting treatment plan might either be overly conservative or too liberal.

Although identical error distributions for treatment planning and patient irradiation cannot be guaranteed, we showed that probabilistic optimization on normally distributed errors increased firstly the average treatment plan quality and secondly revealed a lower uncertainty in dose for other PDFs. This can be understood by two facts. First, other PDFs are to a certain extent covered by the Normal distribution and only differences to the Normal distribution cause the plan to be less robust. The greater the dissimilarly between planned and actual unknown PDF, the less robust the plan becomes compared to the utilized PDF for optimization. Second, the probabilistic objective function aims to minimize the variance and the difference to the expected value. Therefore, distributions which induce a similar variance and expectation value in dose compared to the standard assumption, will also be mitigated to a comparable degree. The broader Normal and the bitriangular distribution are out of line in this consideration due to the increased variance and the fact that improbable scenarios become most probable. Notwithstanding, even for these rather extreme distributions the resulting variance of the probabilistic treatment plan is lower compared to the conventionally optimized plan. Probabilistic dose metrics and standard deviation volume histograms confirmed the observations made for the one and three dimensional examples. This demonstrated that the probabilistically optimized treatment plans revealed a lower variance against other PDFs.

The highest dosimetric impact on the expectation value was observed for the asymmetric skewed Normal distribution. Although probabilistic optimization on normally distributed errors could compensate dosimetric differences for the one-dimensional example, the situation was different for three-dimensional examples as $\mu[\mathbb{E}]$ was still below 2.96 Gy(RBE) for all patient cases. Nevertheless, the corresponding SDVH still agreed with the main trend of other PDFs as pointed out in Figure 11.

The shape of the PDF modeling range and setup errors is of lower importance after probabilistic optimization as the corresponding treatment plans revealed a lower variance for other PDFs. Even when the PDFs differ considerably (e.g. bitriangular) from the one used for probabilistic optimization, the resulting treatment plans exhibited, compared to conventional treatment plan based on the PTV-margin concept, a smaller uncertainty in dose for the CTV as well as for the considered OAR. Moreover, the average expected dose $\mu[\mathbb{E}]$ was for the CTV tendentiously higher after probabilistic optimization. Complementary, an opposing trend could be observed for the OAR. This suggests that a vague knowledge about the underlying exact PDF may be sufficient to reduce uncertainties during probabilistic optimization. In further consequence,

the approximation of modeling the unknown setup and range uncertainties with a multivariate symmetric Normal distribution is considered as a reasonable first approach. However, a critical eye must be kept on asymmetric distributions such as the skewed Normal distribution with a mean value different than zero. For such cases, probabilistic optimization on the basis of symmetric Normal distribution did improve the uncertainty in dose but did not avoid potential under/overdosage of the CTV.

We acknowledge that the facilitated analytical pencil beam dose calculation model has limitations in heterogeneous tissues such as the lung. Although this work facilitates an improved version of the pencil beam algorithm which takes the actual radiological depth of voxel i into account and not its projection onto the pencil beam central axis, the dose distribution might still be inaccurate for certain tumor sites (e. g. lung). However, it can be assumed that this does only moderately compromise the observations regarding the relative improvement in treatment plan robustness. The two optimization approaches output different set of pencil beam intensities in which probabilistic optimization aims to produce shallower in-field dose gradients by reducing the intensities of distal pencil beams. This relative improvement depends only minor on the underlying dose engine (analytical or Monte Carlo).

On the basis of this knowledge, a more accurate uncertainty estimate could be derived via a Monte Carlo dose calculation algorithm for the cost of a substantially increased computational effort. Another limitation of the presented method is to model setup and range uncertainties in the pencil beam space at the same time leaving out a direct consideration of uncertainties caused by intra-fractional anatomical changes (e. g. lung motion). Although anatomical changes such as organ movements affect setup and range uncertainties, they are not incorporated into the presented uncertainty quantification and probabilistic optimization. Nevertheless, we argue that probabilistically optimized treatment plans considering setup and range uncertainties would still be more robust against motion than the conventionally optimized counterpart due to reduced in-field dose gradients, especially in the distal dose fall off. Notwithstanding, future studies need to elaborate on this aspect to hold. One solution for APM might be given by incorporating sophisticated tumor-site specific correlation models in the pencil beam space or voxel domain to consider intra-fractional motion uncertainties.

This study is restricted to systematic errors in order to reduce sampling requirements. Although, it was shown in previous studies (Bangert et al. 2013, Wahl et al. 2017) that APM allows to efficiently define random and systematic contributions separately within the correlation model in the pencil beam space, we neglect fractionation effects for the reduced computing time. Sampling from systematic and random error sources would result in an increased sampling effort but allows to obtain a more accurate uncertainty quantification in dose for a specific fractionation scheme.

Our study was restricted on the question how crucial exact knowledge about the true PDF of range and setup uncertainties is for probabilistic treatment plan optimization. In this context, we have shown that the general benefits of probabilistic optimization outweigh potential modeling deficiencies compared to conventional multi-field IMPT optimization. The results might generalize to other uncertainty mitigation approaches as they also tend to flatten high dose gradients by reducing the intensity of highly modulated pencil beam contributions, however, a final concluding statement cannot be made at this point. In the next step these results may be tested for individual indications against other methods for uncertainty mitigation in particle therapy such as singe field uniform dose (SFUD) optimization and worst case optimization.

5. Conclusion

Precise uncertainty quantification and stringent probabilistic optimization requires exact knowledge about the underlying probability distribution. However, in reality the underlying probability distribution is mostly unknown. We investigated the dependence of dosimetric uncertainty on the choice of the underlying uncertainty model in intensity-modulated treatment planning for proton and carbon ion therapy. We analyzed the treatment plan robustness of conventionally and probabilistically treatment plans assuming that range and setup error follow six different probability distributions. As expected, different uncertainty assumptions yielded distinct expectation values, standard deviations and probabilistic dose metrics. For conventional optimization utilizing the PTV as target, a non-negligible average standard deviation in dose of 7.3 % was observed throughout the CTV. Probabilistic optimization assuming normally distributed range and setup errors provided dose distributions with lower standard deviation in dose than the conventional reference plan, also considering different probability distributions for uncertainty quantification than for optimization. This statement is in particular true for the CTV structure for which the average standard deviation in dose dropped for different probability distributions to 3.9 %. Our study provides evidence that probabilistic optimization provides substantial benefits in treatment plan robustness also if an incorrect probability density occurs during patient irradiation.

6. Disclosure of Conflicts of Interest

The authors have no relevant conflicts of interest to disclose.

7. Acknowledgments

We thank Dr. Malte Ellerbrock from the Heidelberg Ion-Beam Therapy Center (HIT) for the provision of an anonymized lung cancer case and Dr. Wenuha Cao from the MD Anderson Cancer Center in Houston for the supply of an anonymized cranial head case. Financial support from the German Research Foundation, Grant No. BA 2279/3-1, is gratefully acknowledged. Further the authors would like to thank *Mathworks Inc.* for providing MATLAB software licenses.

8. References

Albertini, F., Gaignat, S., Bosshardt, M. & Lomax, A. J. (2010). Planning and Optimizing Treatment Plans for Actively Scanned Proton Therapy, *Biomedical mathematics : promising directions in imaging, therapy planning, and inverse problems*, Medical Physics Publishing, pp. 1–18.

Albertini, F., Hug, E. B. & Lomax, A. J. (2011). Is it necessary to plan with safety margins for actively scanned proton therapy?, *Physics in Medicine and Biology* **56**(14): 4399–4413.

Arts, T., Breedveld, S., de Jong, M. A., Astreinidou, E., Tans, L., Keskin-Cambay, F., Krol, A. D., van de Water, S., Bijman, R. G. & Hoogeman, M. S. (2017). The impact of treatment accuracy on proton therapy patient selection for oropharyngeal cancer patients, *Radiotherapy and Oncology* **125**(3): 520–525.

Bangert, M., Hennig, P. & Oelfke, U. (2013). Analytical probabilistic modeling for radiation therapy treatment planning., *Physics in medicine and biology* **58**: 5401–19.

Baum, C., Alber, M., Birkner, M. & Nüsslin, F. (2004). Treatment simulation approaches for the estimation of the distributions of treatment quality parameters generated by geometrical uncertainties., *Physics in medicine and biology* **49**(24): 5475–5488.

Bolsi, A., Lomax, A. J., Pedroni, E., Goitein, G. & Hug, E. (2008). Experiences at the Paul Scherrer Institute With a Remote Patient Positioning Procedure for High-Throughput Proton Radiation Therapy, *International Journal of Radiation Oncology*Biology*Physics* **71**(5): 1581–1590.

Cao, W., Lim, G. J., Li, Y., Zhu, X. R. & Zhang, X. (2015). Improved beam angle arrangement in intensity modulated proton therapy treatment planning for localized prostate cancer, *Cancers* **7**(2): 574–584.

Casiraghi, M., Albertini, F. & Lomax, a. J. (2013). Advantages and limitations of the 'worst case scenario' approach in IMPT treatment planning., *Physics in medicine and biology* **58**(5): 1323–39.

Fredriksson, A. (2012). A characterization of robust radiation therapy treatment planning methods-from expected value to worst case optimization, *Medical Physics* **39**(8): 5169–5181.

Fredriksson, A. & Bokrantz, R. (2016). The scenario-based generalization of radiation therapy margins., *Physics in medicine and biology* **61**(5): 2067–82.

Fredriksson, A., Forsgren, A. & Hardemark, B. (2011). Minimax optimization for handling range and setup uncertainties in proton therapy., *Medical physics* **38**(3): 1672–84.

Friedrich, T., Scholz, U., Elsässer, T., Durante, M. & Scholz, M. (2012). Calculation of the biological effects of ion beams based on the microscopic spatial damage distribution pattern, *International Journal of Radiation Biology* **88**(1-2): 103–107.

Goitein, M. (1985). Calculation of the uncertainty in the dose delivered during radiation therapy, *Medical Physics* **12**(5): 608–612.

Graeff, C. (2014). Motion mitigation in scanned ion beam therapy through 4D-optimization, *Physica Medica* **30**(5): 570–577.

Haberer, T., Becher, W., Schardt, D. & Kraft, G. (1993). Magnetic scanning system for heavy ion therapy, *Nuclear Inst. and Methods in Physics Research, A* **330**(1-2): 296–305.

Holloway, S. M., Holloway, M. D. & Thomas, S. J. (2017). A method for acquiring random range uncertainty probability distributions in proton therapy, *Physics in Medicine & Biology* **63**(1): 01NT02.

Knopf, A.-C., Boye, D., Lomax, A. & Mori, S. (2013). Adequate margin definition for scanned particle therapy in the incidence of intrafractional motion, *Physics in Medicine and Biology* **58**(17): 6079–6094.

Knopf, A. C. & Lomax, A. (2013). In vivo proton range verification: A review, *Physics in Medicine and Biology* **58**(15): 131–160.

Krämer, M. & Scholz, M. (2000). Treatment planning for heavy-ion radiotherapy: calculation and optimization of biologically effective dose, *Physics in Medicine and Biology* **45**(11): 3319–3330.

Liebl, J., Paganetti, H., Zhu, M. & Winey, B. A. (2014). The influence of patient positioning uncertainties in proton radiotherapy on proton range and dose distributions, *Medical Physics* **41**(9): 091711.

Liu, W., Zhang, X., Li, Y. & Mohan, R. (2012). Robust optimization of intensity modulated proton therapy, *Medical Physics* .

Liu, W. (2016). Robustness Quantification and Worst-Case Robust Optimization in Intensity-Modulated Proton Therapy, *Particle Radiotherapy*, Springer India, Liu2016, pp. 139–155.

Lomax, A. (2016). SFUD, IMPT, and Plan Robustness, *Particle Radiotherapy*, Springer India, New Delhi, pp. 169–194.

Lomax, A. J. (2008a). Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties., *Physics in medicine and biology* **53**(4): 1027–1042.

Lomax, A. J. (2008b). Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions, *Physics in Medicine and Biology* **53**(4): 1043–1056.

Lowe, M., Albertini, F., Aitkenhead, A., Lomax, A. J. & MacKay, R. I. (2016). Incorporating the effect of fractionation in the evaluation of proton plan robustness to setup errors, *Physics in Medicine and Biology* **61**(1): 413–429.

Mescher, H., Ulrich, S. & Bangert, M. (2017). Coverage-based constraints for IMRT optimization, *Physics in Medicine & Biology* **62**(18): N460–N473.

Paganetti, H. (2012). Range uncertainties in proton therapy and the role of Monte Carlo simulations, *Physics in Medicine and Biology* **57**: R99–R117.

Park, P. C., Cheung, J. P., Zhu, X. R., Lee, A. K., Sahoo, N., Tucker, S. L., Liu, W., Li, H., Mohan, R., Court, L. E. & Dong, L. (2013). Statistical assessment of proton treatment plans under setup and range uncertainties, *International Journal of Radiation Oncology Biology Physics* **86**(5): 1007–1013.

Park, P. C., Zhu, X. R., Lee, A. K., Sahoo, N., Melancon, A. D., Zhang, L. & Dong, L. (2012). A Beam-Specific Planning Target Volume (PTV) Design for Proton Therapy to Account for Setup and Range Uncertainties, *International Journal of Radiation Oncology*Biology*Physics* **82**(2): e329–e336.

Perkó, Z., van der Voort, S. R., van de Water, S., Hartman, C. M. H., Hoogeman, M. & Lathouwers, D. (2016). Fast and accurate sensitivity analysis of IMPT treatment plans using Polynomial Chaos Expansion., *Physics in medicine and biology* **61**(12): 4646–64.

Pflugfelder, D., Wilkens, J. J. & Oelfke, U. (2008). Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy, *Phys. Med. Biol. Phys. Med. Biol* **53**(53): 1689–1700.

Ploquin, N., Kay, I., Rangel-Baltazar, A., Lau, H. & Dunscombe, P. (2006). A comparison of techniques for simulating set-up error and uncertainty in head and neck IMRT, *Medical Physics* **33**(9): 3213–3219.

Sakama, M., Kanematsu, N. & Inaniwa, T. (2016). A robustness analysis method with fast estimation of dose uncertainty distributions for carbon-ion therapy treatment planning, *Physics in Medicine and Biology* **61**(15): 5818–5836.

Scholz, M., Kellerer, a. M., Kraft-Weyrather, W. & Kraft, G. (1997). Computation of cell survival in heavy ion beams for therapy, *Radiation and Environmental Biophysics* **36**(1): 59–66.

Schuemann, J., Dowdell, S., Grassberger, C., Min, C. H. & Paganetti, H. (2014a). Site-specific range

uncertainties caused by dose calculation algorithms for proton therapy, *Physics in Medicine and Biology* **59**(15): 4007–4031.

Schuemann, J., Dowdell, S., Grassberger, C., Min, C. H. & Paganetti, H. (2014b). Site-specific range uncertainties caused by dose calculation algorithms for proton therapy., *Physics in medicine and biology* **59**(15): 4007–31.

Steitz, J., Naumann, P., Ulrich, S., Haefner, M. F., Sterzing, F., Oelfke, U. & Bangert, M. (2016). Worst case optimization for interfractional motion mitigation in carbon ion therapy of pancreatic cancer, *Radiation Oncology* **11**(1): 134.

Unkelbach, J. & Oelfke, U. (2004). Inclusion of organ movements in IMRT treatment planning via inverse planning based on probability distributions., *Physics in medicine and biology* **49**(17): 4005–4029.

Unkelbach, J., Alber, M., Bangert, M., Bokrantz, R., Chan, T. C. Y., Deasy, J. O., Fredriksson, A., Gorissen, B. L., van Herk, M., Liu, W., Mahmoudzadeh, H., Nohadani, O., Siebers, J. V., Witte, M. & Xu, H. (2018). Robust radiotherapy planning, *Physics in Medicine & Biology* **63**(22): 22TR02.

Unkelbach, J., Chan, T. C. Y. & Bortfeld, T. (2007). Accounting for range uncertainties in the optimization of intensity modulated proton therapy, *Physics in Medicine and Biology* **52**(10): 2755–2773.

Unkelbach, J. & Paganetti, H. (2018). Robust Proton Treatment Planning: Physical and Biological Optimization, *Seminars in Radiation Oncology* **28**(2): 88–96.

van der Voort, S., van de Water, S., Perkó, Z., Heijmen, B., Lathouwers, D. & Hoogeman, M. (2016). Robustness Recipes for Minimax Robust Optimization in Intensity Modulated Proton Therapy for Oropharyngeal Cancer Patients, *International Journal of Radiation Oncology*Biology*Physics* **95**(1): 163–170.

Van Herk, M. (2004). Errors and Margins in Radiotherapy, *Seminars in Radiation Oncology* **14**(1): 52–64.

van Herk, M., Remeijer, P. & Lebesque, J. V. (2002). Inclusion of geometric uncertainties in treatment plan evaluation., *International journal of radiation oncology, biology, physics* **52**(5): 1407–22.

Waechter, A. & Biegler, L. T. (2006). *On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming*, Vol. 106.

Wahl, N., Hennig, P., Wieser, H. P. & Bangert, M. (2017). Efficiency of analytical and sampling-based uncertainty propagation in intensity-modulated proton therapy., *Physics in medicine and biology* **62**(14): 5790–5807.

Wahl, N., Hennig, P., Wieser, H.-P. & Bangert, M. (2018). Analytical incorporation of fractionation effects in probabilistic treatment planning for intensity-modulated proton therapy, *Medical Physics* **45**(4): 1317–1328. **URL:** *http://doi.wiley.com/10.1002/mp.12775*

Wieser, H.-P., Cisternas, E., Wahl, N., Ulrich, S., Stadler, A., Mescher, H., Muller, L.-R., Klinge, T., Gabrys, H., Burigo, L., Mairani, A., Ecker, S., Ackermann, B., Ellerbrock, M., Parodi, K., Jakel, O. & Bangert, M. (2017). Development of the open-source dose calculation and optimization toolkit matRad, *Medical Physics* **44**(6).

Wieser, H.-P, Hennig, P., Wahl, N. & Bangert, M. (2017b). Analytical probabilistic modeling of RBE-weighted dose for ion therapy, *Physics in Medicine and Biology* **62**(23).

9. Appendix

Table 5: Treatment plan parameters of three-dimensional intensity-modulated proton treatment plans. The lateral spot distance determines the pencil beam spacing at iso-center and therefore the lateral density of pencil beam spots. '# voxels of σ' denotes the number of voxels for which standard deviation was computed.

	unit	intra-cranial I	lung	intra-cranial II
gantry angles	[°]	50°, 135°	35°, 350°	100°, 260°
couch angles	[°]	0°, 0°	0°,0°	0°,0°
lat. spot distance	[mm]	3	5	5
resolution	[mm]	$(1.2 \times 1.2 \times 3)$	$(3 \times 3 \times 2.5)$	$(3 \times 3 \times 2.5)$
# pencil beam	-	2022	5182	7275
# voxels of σ	-	0.95×10^5	0.65×10^5	0.4×10^5
# fractions f	-	20	20	20
pres. fraction dose	[Gy(RBE)]	3	3	3

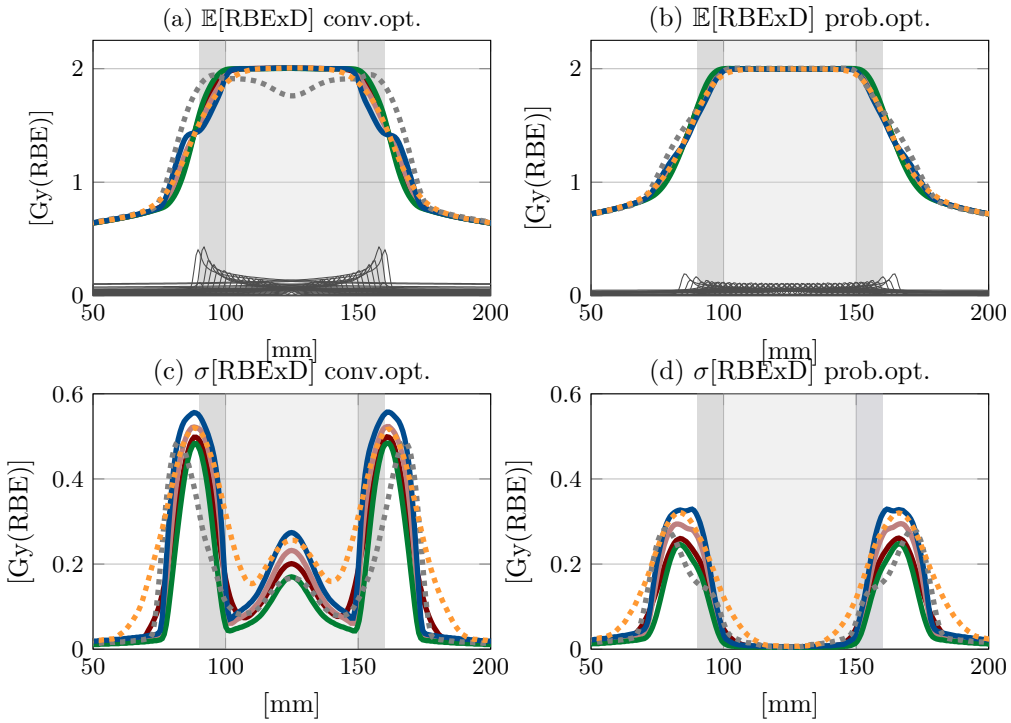

Figure 12: Statistical moments of a one-dimensional carbon ion SOBP. Different probability density functions were employed for random sampling to model the relative range errors. The statistical moments obtained from probabilistic optimization (prob.opt.) show first a lower uncertainty in dose throughout the CTV and a decreased variability between different $\sigma[RBExD]$.

Table 6: Robustness analysis of the one-dimensional carbon ion SOBP employing different PDFs. The corresponding dose profiles can be found in the Appendix in Figure 12. The shown probabilistic dose metrics refer to the CTV structure. Columns 2-7 show the probabilistic dose metrics for conventional optimization and columns 8-13 are based on probabilistic optimization. The columns $\mu[\mathbb{E}]$, $\mu[\sigma]$, $\mu[D95]$ and $\sigma[D95]$ are given in [Gy(RBE)] whereas $\mu[V95]$ and $\sigma[V95]$ are given in [%].

range error	**conv.opt.**						**prob.opt.**					
PDF	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[D95]$	$\sigma[D95]$	$\mu[V95]$	$\sigma[V95]$	$\mu[\mathbb{E}]$	$\mu[\sigma]$	$\mu[D95]$	$\sigma[D95]$	$\mu[V95]$	$\sigma[V95]$
normal	1.99	0.138	1.84	0.24	84.2	22.9	1.99	0.017	1.96	0.063	98.6	2.8
uniform	1.98	0.144	1.82	0.19	79.9	26.9	1.99	0.014	1.96	0.034	98,1	2.1
triangular	1.99	0.09	1.92	0.12	88.5	19.4	1.99	0.008	1.98	0.015	99.6	1.0
bitriangular	1.97	0.173	1.73	0.21	74.3	32.0	1.99	0.017	1.94	0.036	96.8	2.2
skewed	1.86	0.117	1.71	0.17	43.4	34.7	1.99	0.024	1.93	0.080	95.9	3.7
broad normal	1.96	0.227	1.66	0.38	74.9	28.4	1.98	0.046	1.88	0.159	94.9	6.7

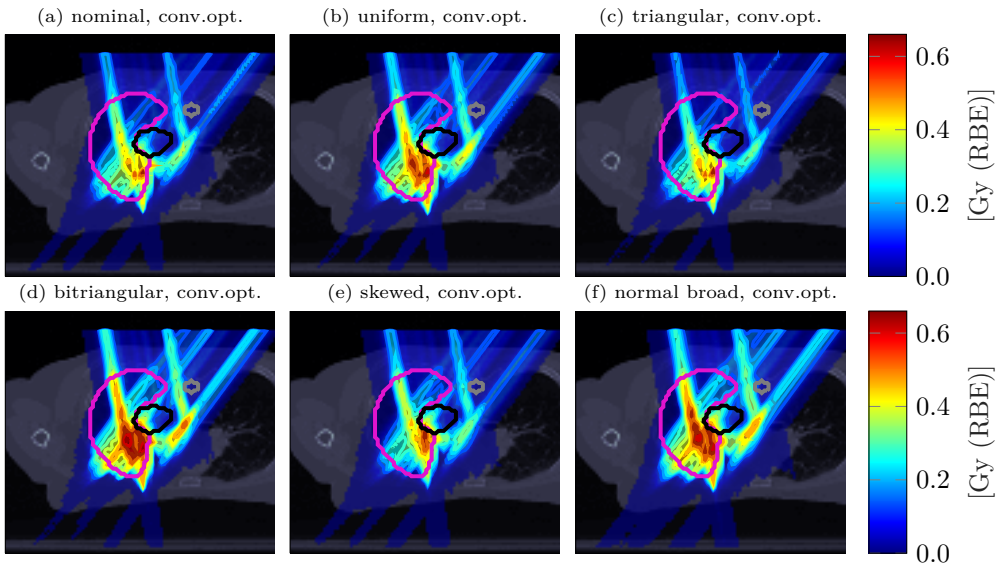

Figure 13: illustrates the transversal CT scan of the lung patient case close to the iso-center. Transparent dose colorwash indicates the standard deviation of the RBE weighted dose assuming different PDFs. The proton treatment plan is based on conventional optimization on the PTV (not shown). Two proton beams are impinging from gantry angles of 35° and 350°. The black contour represents the CTV boundary and the pink contour depicts the left lung.

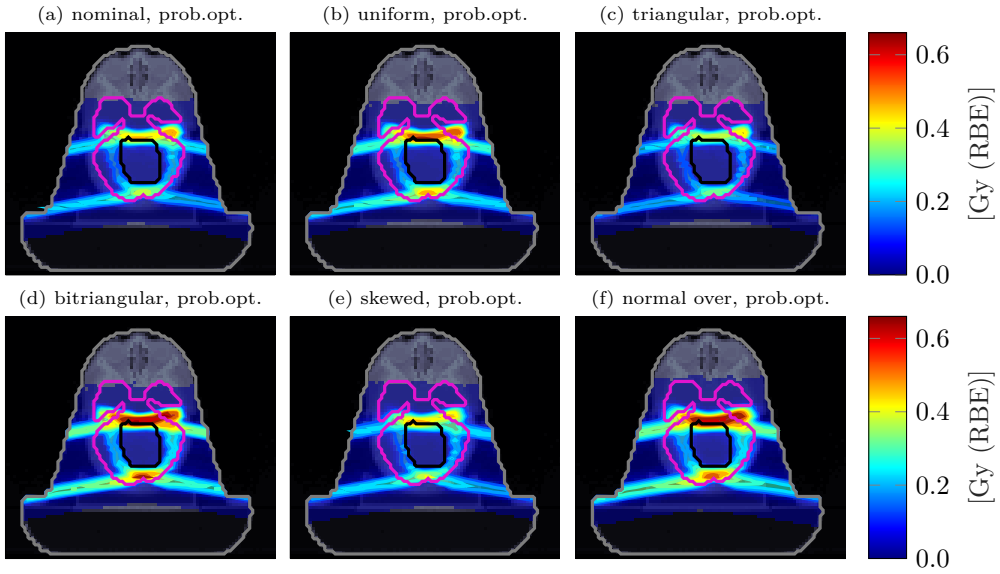

Figure 14: Uncertainty quantification results for the intra-cranial II case. Transversal CT image is overlaid by transparent dose colorwash representing the standard deviation of the RBE weighted dose assuming different PDFs. Pencil beam intensities are based on probabilistic optimization on the CTV. Two proton beam directions from 100° and 260° were defined. The black contour denotes the CTV whereas the pink contour depicts the brain structure.