

Dynamic evaluation of 4D robust optimisation
for motion management in scanned proton therapy
of hepatocellular carcinoma

DISSERTATION

Zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)

Fakultät Physik
Technische Universität Dortmund

vorgelegt von

M.Sc. Tina Pfeiler
geboren am 27.10.1989 in Oberhausen

Essen im Dezember 2018

Dissertation to obtain the degree Doctor Rerum Naturalium (Dr. rer. nat.) at the Faculty of Physics, Technical University of Dortmund.

1st Referee: Prof. Dr. Bernhard Spaan
Technical University of Dortmund

2nd Referee: Prof. Dr. Kevin Kröniger
Technical University of Dortmund

This thesis has been conducted at the West German Proton Therapy Centre Essen (WPE) under supervision of Dr. Christian Bäumer.

Abstract

Pencil beam scanning (PBS) proton therapy is due to its steep dose gradients a promising treatment option for many cancer patients, particularly if the tumour is surrounded by critical organs. Currently, organ motion often poses an obstacle for the irradiation of tumours in the thorax or abdomen since the interference of pencil beam and tumour motion can lead to distortions of the intended dose distribution and requires dedicated motion mitigation techniques. A technique which has been recommended by recently published guidelines to reduce motion effects in PBS therapy of thoracic malignancies is 4D robust optimisation. 4D robust optimisation uses time-resolved CT images from different respiratory phases to incorporate uncertainties due to physiologic motion in treatment plan optimisation. Its effect has not yet been sufficiently examined and further research is needed to establish it in clinical routine.

Within this thesis, the utility of 4D robust optimisation to mitigate motion effects was investigated for hepatocellular carcinoma, representative for mobile abdominal targets. For this purpose, a 4D dynamic dose calculation routine including an empirical beam delivery time model was developed based on a non-clinical, preliminary script from a global software provider for radiation therapy. The customised 4D tool enables prospective estimations before treatment, 4D dose reconstruction after treatment and comprehensive *in silico* examinations for research purposes. The validity of the routine and the time model was confirmed for typical beam settings in clinical end-to-end tests. In the treatment planning study on 4D robust optimisation, the routine revealed no significant improvement of the target coverage under motion for the investigated patients. A great advantage over conventional PBS plans was, however, that the same level of robustness against motion was reached for lower organ at risk doses. Since dose inhomogeneities averaged out for the applied fractionation scheme of 15 fractions, 4D robust optimisation was found to generate clinically acceptable sum plans. Due to the unknown biological effect of strong over- and underdosage in a single fraction, the method should be supplemented by further motion mitigation techniques in case of large tumour motion.

The use of computer generated CTs for more comprehensive 4D evaluations was successfully tested in a proof of concept study and will allow to model a broad range of clinical cases in upcoming research projects. Thus, the benefit of 4D robust optimisation can be systematically investigated on a large scale for different subgroups. Although the elaboration of operating procedures for the irradiation of moving targets at the West German Proton Therapy Centre Essen (WPE) is still in process, the 4D dynamic dose calculation routine is already in clinical use to assess potential motion effects for current patients with minor tumour motion.

Zusammenfassung

Die Protonentherapie mit magnetisch gelenkten Nadelstrahlen, englisch pencil beam scanning (PBS), stellt aufgrund ihrer steilen Dosisgradienten eine erfolgversprechende Behandlungsoption für viele Tumorpatienten dar. Dies gilt insbesondere, wenn der Tumor von lebenswichtigen Organen umgeben ist. Aktuell sind atmungsbedingte Organbewegungen oft ein Hindernis für die Bestrahlung von Tumoren im Thorax oder Abdomen, weil das Zusammenspiel von Tumor- und Strahlbewegung zu Verzerrungen der geplanten Dosisverteilung führen kann. Die Anwendung spezieller Techniken zur Reduzierung von Bewegungseffekten ist daher erforderlich. Kürzlich veröffentlichte Richtlinien für die Bestrahlung von thorakalen Tumoren mit PBS empfehlen die Anwendung von 4D-robuster Optimierung. Diese berücksichtigt physiologische Organbewegungen bei der Optimierung von Bestrahlungsplanungsparametern anhand von zeitaufgelösten CT-Aufnahmen zu verschiedenen Atemphasen. Ihre Wirkung ist noch nicht ausreichend untersucht und es bedarf weiterer Forschungsarbeit, um die Technik in der Klinik zu etablieren.

Diese Arbeit hatte zum Ziel, den Nutzen von 4D-robuster Optimierung zur Abschwächung von Bewegungseffekten für hepatozelluläre Karzinome zu untersuchen, repräsentativ für bewegte Ziele im Abdomen. Zu diesem Zweck wurde eine Routine entwickelt, die retrospektive, dynamische 4D-Dosisberechnungen ermöglicht. Sie basiert auf einer nicht klinischen Vorläuferversion eines führenden Anbieters für Bestrahlungsplanungs-Software. Ein empirisches Zeitmodell der Feldapplikation vervollständigt die Routine und erlaubt umfangreiche in silico Studien zu Forschungszwecken sowie prospektive Untersuchungen zur Abschätzung von Risiken vor der Behandlung. Die Gültigkeit der Routine und des integrierten Zeitmodells wurde in klinischen End-to-End-Tests für typische Strahlparameter nachgewiesen. In der Bestrahlungsplanungsstudie zu 4D-robuster Optimierung zeigte die Routine keine signifikante Verbesserung in der Zielvolumenabdeckung bei Atembewegungen für die untersuchten Patienten. Ein großer Vorteil war jedoch, dass, im Vergleich zu konventionellen PBS-Plänen, der gleiche Robustheitsgrad gegenüber Bewegungen bei geringerer Belastung des gesunden umliegenden Gewebes erzielt werden konnte. Da sich Dosisinhomogenitäten für das angewandte Fraktionierungsschema von 15 Fraktionen herausmittelten, generierte 4D-robuste Optimierung klinisch akzeptable 4D-Summenpläne. Aufgrund der nicht ausreichend bekannten biologischen Wirkung von starken Unter- und Überdosierungen in einzelnen Fraktionen, ist es jedoch erforderlich, im Fall von großen Bewegungsamplituden, begleitend zusätzliche Maßnahmen zur Einschränkung von Bewegungseffekten zu ergreifen.

Die Anwendung von computergenerierten CTs für umfangreichere 4D-Evaluierungen wurde erfolgreich in einer Machbarkeitsstudie getestet. Sie macht es möglich, zukünftig eine große Bandbreite an klinischen Fällen in Forschungsprojekten zu modellieren. Somit kann der Nutzen von 4D-robuster Optimierung systematisch im großen Rahmen für verschiedene Untergruppen untersucht werden. Obwohl die Ausarbeitung von Arbeitskonzepten für die Behandlung von bewegten Tumoren am Westdeutschen Protonentherapiezentrum Essen (WPE) noch nicht abgeschlossen ist, wird die 4D-Dosisberechnungsroutine bereits für derzeitige Patienten mit geringer Tumorbewegung klinisch genutzt, um potenzielle Bewegungseffekte zu erfassen und zu bewerten.

Publications related to this work

- **Pfeiler T**, Bäumer C, Engwall E, Geismar D, Spaan B, Timmermann B. Experimental validation of a 4D dose calculation routine for pencil beam scanning proton therapy. *Z Med Phys* 2018;28:121–33.
- **Pfeiler T**, Ahmad Khalil D, Bäumer C, Blanck O, Chan M, Engwall E, Geismar D, Peters S, Plaude S, Spaan B, Wulff J and Timmermann B. 4D robust optimization in pencil beam scanning proton therapy for hepatocellular carcinoma. *J Phys Conf Ser*. Submitted 2018.
- **Pfeiler T**, Ahmad Khalil D, Ayadi M, Bäumer C, Blanck O, Chan M, Engwall E, Geismar D, Peters S, Plaude S, Spaan B, Timmermann B, Wulff J. Motion effects in proton treatments of hepatocellular carcinoma - 4D robustly optimised pencil beam scanning plans vs double scattering plans. *Phys Med Biol* 2018;63 235006.

Conference contributions

- **Pfeiler T**, Bäumer C, Engwall E, Spaan B, Timmermann B. A routine for the evaluation of interplay effects in Intensity-Modulated Proton Therapy (IMPT). 47th Annual Meeting of the DGMP, Würzburg, Germany, 2016. Oral presentation.
- **Pfeiler T**, Bäumer C, Engwall E, Spaan B, Timmermann B. A routine for the evaluation of interplay effects in pencil-beam scanning proton therapy. 4D Treatment Planning Workshop, Groningen, Netherlands, 2016. Poster.
- **Pfeiler T**, Bäumer C, Engwall E, Spaan B, Timmermann B. 4D dose computation in pencil beam scanning proton therapy has been clinically implemented and experimentally verified. PTCOG 56, Yokohama, Japan, 2017. Oral presentation.
- **Pfeiler T**, Ahmad Khalil D, Bäumer C, Blanck O, Chan M, Engwall E, Geismar D, Peters S, Plaude S, Spaan B, Wulff J, Timmermann B. 4D robust optimization in pencil beam scanning proton therapy for hepatocellular carcinoma. MMND, Mooloolaba, Australia, 2018. Oral presentation.
- **Pfeiler T**, Ahmad Khalil D, Bäumer C, Blanck O, Chan M, Engwall E, Geismar D, Peters S, Plaude S, Spaan B, Wulff J, Timmermann B. 4D evaluation of proton pencil beam scanning and double scattering for hepatocellular carcinoma. ESTRO 37, Barcelona, Spain, 2018. E-Poster.
- Moustakis C, Blanck O, Andratschke N, Boda-Heggemann J, Wilke L, Tanadini-Lang S, Gauer T, Alraun M, Bäumer C, Beckers E, Blaschek M, Brunner T, Chan MK, Dierl M, Droege S, Duma MN, Ebrahimi F, Eich HT, Fleckenstein J, Ganswindt U, Garbe S, Gkika E, Heinz C, Henkenberens C, Hennig A, Henning S, Köhn J, Kornhuber C, Kretschmer M, Krieger T, Loufti-Kraus B, Mayr M, Mensing T, **Pfeiler T**, Pollul G, Ramm U, Rieder F, Rieken S, Schmidhalter D, Schmitt D, Schöffler J, Ulm C, Walke M, Weigel R, Wiehle R, Wiezorek T, Wolf U, Zimmer J, Baus W, Guckenberger M. Referenz Planvergleichsstudie für Leber-SBRT – Ergebnisse DEGRO AG/DGMP AK Stereotaxie. *Strahlenther Onkol* 2018;194:1.

Nomenclature

| | |
|-----------------|--|
| 3D PBS plan | non-robustly optimised PBS plan |
| 4D CT | respiratory correlated computed tomography |
| 4D PBS plan | 4D robustly optimised PBS plan |
| 4D XCAT phantom | 4D extended cardiac-torso phantom |
| 4DD | 4D accumulated dose |
| 4DDD | 4D dynamically accumulated dose |
| AP | anterior-posterior |
| CT | computed tomography |
| CTV | clinical target volume |
| DIR | deformable image registration |
| DS | double scattering |
| DSC | Dice similarity coefficient |
| gEUD | generalised EUD |
| GTV | gross tumour volume |
| HCC | hepatocellular carcinoma |
| HI | homogeneity index |
| HU | Hounsfield unit |
| IMPT | intensity-modulated proton therapy |
| ITV | internal target volume |
| LET | linear energy transfer |
| LKB model | Lyman-Kutcher-Burman model |
| LQ model | linear-quadratic model |
| LR | left-right |
| MC | Monte Carlo |
| MCS | multiple coulomb scattering |
| MU | monitor unit |
| NLP | non-linear optimisation problem |
| NTCP | normal tissue complication probability |
| OAR | organ at risk |
| PBS | pencil beam scanning |
| PLD file | PBS layer definition file |
| PMMA | polymethyl methacrylate |

| | |
|------|---|
| POI | point of interest |
| PTV | planning target volume |
| QP | quadratic programming |
| RBE | relative biological effectiveness |
| RF | radio frequency |
| RFA | radiofrequency ablation |
| RILD | radiation-induced liver disease |
| ROI | region of interest |
| SBRT | stereotactic body radiotherapy |
| SD | standard deviation |
| SEM | standard error of the mean |
| SI | superior-inferior |
| SOBP | spread-out Bragg peak |
| SQP | sequential quadratic programming |
| TACE | transarterial chemoembolisation |
| TPS | treatment planning system |
| WPE | West German Proton Therapy Centre Essen |

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 2 | Research principles and background of proton beam therapy | 3 |
| 2.1 | Physical fundamentals | 3 |
| 2.1.1 | Proton interactions in matter | 3 |
| 2.1.2 | Dosimetric quantities | 3 |
| 2.1.3 | Depth dose profile | 4 |
| 2.2 | Cell survival and biological effectiveness | 6 |
| 2.2.1 | Cell survival | 6 |
| 2.2.2 | Relative biological effectiveness | 7 |
| 2.2.3 | Biological dose | 8 |
| 2.2.4 | Equivalent uniform dose | 8 |
| 2.3 | Proton beam therapy system | 8 |
| 2.3.1 | Components | 8 |
| 2.3.2 | Beam delivery: Pencil beam scanning (PBS) and double scattering (DS) | 10 |
| 2.4 | Treatment planning | 14 |
| 2.4.1 | Treatment planning system (TPS) | 14 |
| 2.4.2 | Computed tomography (CT) | 14 |
| 2.4.3 | Contouring | 16 |
| 2.4.4 | Deformable image registration | 17 |
| 2.4.5 | Proton dose calculation | 18 |
| 2.4.6 | Treatment planning in pencil beam scanning and 4D robust optimisation | 21 |
| 2.4.7 | Treatment planning in double scattering | 23 |
| 2.5 | Treatment of mobile targets | 23 |
| 2.5.1 | Intra- and inter-fraction motion | 24 |
| 2.5.2 | Impact of organ motion | 24 |
| 2.5.3 | Interplay effects | 24 |
| 2.5.4 | Respiratory correlated computed tomography | 25 |
| 2.5.5 | Motion mitigation techniques | 25 |
| 3 | Implementation of 4D dynamic dose accumulation in PBS treatment planning | 27 |
| 3.1 | Modelling of interplay effects | 27 |
| 3.2 | Experimental validation of the 4D dose computation | 28 |
| 3.2.1 | Materials and methods | 28 |
| 3.2.2 | Results | 35 |
| 3.2.3 | Discussion | 37 |
| 3.3 | Development and validation of a beam delivery time model | 42 |
| 3.3.1 | Materials and methods | 42 |
| 3.3.2 | Results | 43 |
| 3.3.3 | Discussion | 44 |
| 3.4 | Clinical aspects | 47 |
| 4 | Evaluation of 4D robust optimisation for hepatocellular carcinoma patients | 49 |
| 4.1 | Hepatocellular carcinoma (HCC) | 49 |
| 4.1.1 | Clinical background | 49 |
| 4.1.2 | Treatment options and the role of proton therapy | 50 |
| 4.2 | Design of the treatment planning study | 51 |
| 4.2.1 | Patient collective | 51 |
| 4.2.2 | CT scanner calibration, contouring and deformable image registration | 53 |
| 4.2.3 | Treatment plan optimisation in PBS and DS | 53 |
| 4.2.4 | CTV coverage and robustness evaluation | 54 |
| 4.2.5 | 4D dose accumulation in pencil beam scanning and double scattering | 55 |
| 4.2.6 | Evaluation methods and criteria | 55 |

| | | |
|----------|---|------------|
| 4.2.7 | Supplementary analyses | 57 |
| 4.3 | Results | 58 |
| 4.3.1 | Evaluation of the deformable image registration | 58 |
| 4.3.2 | Perturbation analysis of the nominal treatment plans | 58 |
| 4.3.3 | Analysis of the 4D dynamically accumulated dose distribution | 58 |
| 4.3.4 | Supplementary analyses | 62 |
| 4.4 | Discussion | 71 |
| 4.4.1 | Evaluation of the deformable image registration | 71 |
| 4.4.2 | Perturbation analysis of the nominal treatment plans | 71 |
| 4.4.3 | Analysis of the 4D dynamically accumulated dose distribution | 72 |
| 4.4.4 | Supplementary analyses | 77 |
| 4.4.5 | Robust optimisation for motion management of hepatocellular carcinomas | 78 |
| 5 | Usage of computer generated 4D CTs in interplay studies - a proof of concept | 81 |
| 5.1 | 4D XCAT phantom and CT simulation | 81 |
| 5.2 | Study design | 82 |
| 5.2.1 | Simulated 4D CT data and deformable image registration | 82 |
| 5.2.2 | Treatment planning | 83 |
| 5.2.3 | Evaluation methods | 83 |
| 5.3 | Results | 84 |
| 5.3.1 | Evaluation of the deformable image registration | 84 |
| 5.3.2 | CTV coverage and perturbation analysis of the nominal treatment plans | 84 |
| 5.3.3 | Analysis of the 4D dynamically accumulated dose distribution | 84 |
| 5.3.4 | Feasibility of interplay studies using simulated 4D CT data | 85 |
| 5.4 | Discussion | 85 |
| 5.4.1 | Evaluation of the deformable image registration | 85 |
| 5.4.2 | Analysis of the 4D dynamically accumulated dose distribution | 86 |
| 5.4.3 | Feasibility of interplay studies using simulated 4D CT data | 86 |
| 6 | Summary and outlook on future research | 87 |
| 6.1 | Summary - Implementation of 4D dynamic dose accumulation | 87 |
| 6.2 | Summary - Evaluation of 4D robust optimisation for HCC patients | 88 |
| 6.3 | Summary - Usage of computer generated 4D CTs in interplay studies | 89 |
| 6.4 | Outlook on future research | 89 |
| 7 | Conclusion | 91 |
| A | Appendix | 93 |
| A.1 | Scripting in RayStation | 93 |
| A.1.1 | Python file | 94 |
| A.1.2 | XAML file | 98 |
| | References | 101 |

1 Introduction

Cancer represents a major and growing public health problem. According to the World Health Organisation (WHO), in 2016, malignant neoplasms caused worldwide about 9.0 million deaths and were thus the second leading cause of mortality after cardiovascular diseases [1]. This fact underlines the need for more effective cancer treatment. Proton therapy enables high-precision irradiation of malignant tumours owing to the characteristic absorption profile of swift protons in matter. Unlike conventional photon beams, which lead to an exponential dose decay after a short buildup region, therapeutic proton beams exhibit a distinct dose maximum (Bragg peak) followed by a steep dose fall-off at the end of the particle range. This characteristic allows for normal-tissue sparing while escalating the dose to the tumour to improve local control and overall survival (see for example Ref. [2, 3, 4, 5]). An advanced form of proton beam delivery is pencil beam scanning (PBS) which uses two perpendicular magnetic fields to deflect the proton beam laterally to predefined spot positions for a series of iso-energy layers. PBS offers a high degree of dose conformity even in anatomically challenging cases. However, the susceptibility of proton therapy to range uncertainties, setup errors and organ motion may affect the dose distribution. Due to the finite range of protons, not only the lateral but also the longitudinal conformity is compromised in the presence of organ motion. Active delivery methods such as PBS are additionally affected by the interference of organ motion and beam scanning which can lead to over- and underdosage of the target and exposure of organs at risk (OARs), an effect known as interplay effect [6, 7, 8, 9]. Nonetheless, many proton therapy centres plan to apply PBS for moving targets and first centres have even started treating tumours in liver and lung with scanned pencil beams [10, 11]. The reasons behind this are (1) the greater tumour conformity enabled by PBS technique and (2) the fact that the majority of recently built particle therapy centres is equipped with active scanning systems only [10].

The trend towards PBS for the treatment of moving targets requires, even more than passive delivery methods, adequate motion mitigation techniques. Various approaches have been made to mitigate motion induced effects, such as abdominal compression, rescanning (the beam visits a spot position several times per treatment until the prescribed dose is delivered), gating (the beam is turned off for certain phases of the respiratory cycle) or four-dimensional treatment planning (the fourth dimension refers to the time) [12, 13, 14]. However, for some of the above supplementary techniques, interplay effects – although reduced – persist on a low level. Four-dimensional dynamically accumulated dose (4DDD) calculations therefore represent a key ingredient of motion management in PBS to evaluate the available treatment options and determine residual deviations from the intended dose distribution.

Current challenges to implementing PBS for moving targets arise from the facts that, firstly, not all proton therapy centres have the expertise and tools (hardware and/or software) to appropriately reduce motion effects and, secondly, commercial treatment planning systems (TPSs) do not provide a clinical 4DDD computation tool [15].

A motion mitigation technique which does not require any additional hardware and is already implemented in some commercial TPSs is 4D robust optimisation [16, 17, 18]. Clinical 4D robust optimisation tools do not explicitly account for interplay effects but consider anatomical changes in different respiratory phases. Recently published guidelines of the Particle Therapy Co-Operative Group (PTCOG) recommend the application of 4D robust optimisation for the treatment of thoracic malignancies in PBS [15]. The methodology of 4D robust optimisation has not yet been established. The recommendation on the use is based on two exploratory studies only, which used an in-house developed TPS and investigated two different tumour sites, lung and oesophagus [17, 18]. Dosimetric investigations of 4D robust optimisation for a wide range of tumour sites and optimisation algorithms are lacking not least because an appropriate, site-specific 4DDD computation tool is required. Such a tool depends on the knowledge of the time structure of the beam delivery. As there is no standard time structure and vendors of TPSs and delivery machines do not yet provide the option to access relevant scanning control system information from the TPS, prospective interplay effect evaluations currently require in-house development.

This dissertation focuses on the question whether treatment planning based on 4D robust optimisation can help to reduce motion effects and the dose to normal tissue in PBS proton therapy for hepatocellular carcinomas (HCCs). HCC, which is the most common type of primary liver cancer and the third leading cause of cancer related mortality worldwide [19, 20], was selected as a model system for moving targets because it is less affected by tissue heterogeneities and related range uncertainties than thoracic targets [21, 22, 23, 24, 25]. The work was divided into two main parts:

1. A customised 4DDD routine simulating interplay effects in PBS had to be implemented in treatment planning. The routine was experimentally validated in an end-to-end test using a dynamic phantom. A time model of the beam dynamics was established based on empirical measurements to allow for prospective evaluations and the simulation of possible variations in the beam delivery time structure. Although interplay effects have been subject of numerous investigations both experimental and *in silico* (see for example Ref. [26, 27, 28, 29, 30]), there are only a few approaches for the experimental validation of the 4DDD calculation itself [31, 32, 33]. This part of the dissertation extends previous works with its beam time model and a more realistic scenario by utilising different tissue equivalent phantom materials, creating treatment plans with multiple energy layers and including measurement points close to the target boundary.
2. The developed 4DDD routine was applied to evaluate 4D robust optimisation regarding target coverage and OAR sparing in comparison to conventional PBS planning, in which uncertainties are considered by uniform safety margins, and to the former standard for moving targets, double scattering (DS). DS, a passive beam delivery technique, is less susceptible to motion than PBS since the entire tumour is continuously repainted by the usage of scatterers and a fast spinning range modulator wheel [34]. In contrast to PBS, DS requires patient-specific beam shaping devices, deposits unnecessary dose proximal to the target and does not allow for intensity modulation. The novelty of the *in silico* study on 4D robust optimisation is that a different methodology to achieve plan robustness against motion was investigated than in previous publications. Furthermore, the performance was evaluated for another target site and compared to DS. Since the evaluated optimisation tool is part of a commercial TPS, the outcome is of relevance for many institutes.

Within the scope of the treatment planning study, the idea for a follow-up project arose, for which a first proof of concept will be provided in this work. Due to the poor availability of suitable patient data, 4D treatment planning studies often have two major deficits: (1) The patient cohort cannot be divided into subgroups of sufficient size to systematically investigate different influencing factors and (2) the study results are not representative for the entire spectrum of patients. Thus, the validity is restricted to the underlying conditions and it is not possible to draw universal conclusions. In a feasibility study, the use of simulated computed tomography (CT) images was tested to overcome these issues. To this end, the computer generated image data were imported in the TPS and had to go through the whole workflow of 4D treatment planning and evaluation. The study examined exemplarily the dependence of interplay effects on the target size.

The structure of this dissertation is as follows. Chapter 2 provides an overview on the physical and radiobiological fundamentals of proton beam therapy and introduces treatment planning concepts and the topic of moving targets. Chapter 3 deals with the implementation and experimental validation of a four-dimensional dynamic dose calculation routine to assess the impact of organ motion on the dose delivery for scanned pencil beams. In Chapter 4, the capability of 4D robust optimisation to maintain target dose coverage and spare normal tissue in the presence of organ motion is studied for hepatocellular carcinomas based on the developed routine. Chapter 5 presents a proof of concept study which examines the application of computer generated CTs to systematically analyse motion interplay effects in scanned proton beam treatments. A summary of the results and future directions are given in Chapter 6. Chapter 7 concludes the dissertation.

2 Research principles and background of proton beam therapy

2.1 Physical fundamentals

2.1.1 Proton interactions in matter

There are three predominant types of interactions of swift protons with matter: (1) inelastic Coulomb interactions with atomic electrons, (2) elastic Coulomb interactions with atomic nuclei and (3) nuclear reactions [35].

Inelastic Coulomb interactions with atomic electrons dominate the slowing down process of a proton in an absorber. The energy loss per interaction is small and multiple ionisation events lead to a quasi-continuous energy loss. The range of the liberated electrons can reach a few millimetres but is usually less than 1 mm in tissue so that the energy is deposited locally. Even for a high proton energy of 200 MeV, the maximum energy of secondary electrons is only around 400 keV [36]. Due to the mass ratio between protons and atomic electrons, most of the protons do not experience considerable changes in their direction of motion and travel almost along a straight line. Only protons passing close to a nucleus will be elastically scattered or deflected by the repulsive Coulomb force. The energy loss of a proton is negligible for this kind of interaction but even small deviations of the initial trajectory due to scattering will have an impact on the spatial distribution of the energy deposition. Nuclear reactions in which the proton retains approximately its initial energy and angle are of minor importance compared to occasional (non-elastic) nuclear reactions in which the absorption of a swift proton by a nucleus lead to the liberation of neutrons, secondary protons, deuterons, tritons or heavier ions [35]. These reactions are less frequent than Coulomb interactions. Nonetheless, they have a noticeable effect on the dose profile since they remove primary protons from the incident beam (see Sec. 2.1.3).

2.1.2 Dosimetric quantities

The mass stopping power $\frac{1}{\rho}S$ describes the average energy loss dE per path length dl of an ion traversing an absorber with density ρ . It is given by the sum of

- the mass electronic (or collision) stopping power $\frac{1}{\rho} \left(\frac{dE}{dl} \right)_{el}$ caused by inelastic collisions with the atomic shell as well as binary inelastic collisions with single shell electrons,
- the mass radiative stopping power $\frac{1}{\rho} \left(\frac{dE}{dl} \right)_{rad}$ arising from the emission of bremsstrahlung and
- the mass nuclear stopping power $\frac{1}{\rho} \left(\frac{dE}{dl} \right)_{nuc}$ due to elastic Coulomb interactions (the energy loss due to non-elastic nuclear reactions is usually not described by the nuclear stopping power) [37]:

$$\frac{S}{\rho} = -\frac{1}{\rho} \left(\frac{dE}{dl} \right)_{el} - \frac{1}{\rho} \left(\frac{dE}{dl} \right)_{rad} - \frac{1}{\rho} \left(\frac{dE}{dl} \right)_{nuc}, \quad \left[\frac{S}{\rho} \right] = \frac{J m^2}{kg}. \quad (2.1)$$

At therapeutic proton beam energies from about 70 MeV to 250 MeV, the slowing-down process of protons is, as already mentioned above, dominated by the electronic stopping power and can be expressed analytically by the Bethe-Bloch formula [38, 39]. Here, the relativistic version derived by Fano [40] is stated including a shell correction term $\frac{C}{Z_t}$ and a density effect correction term $\frac{\delta}{2}$:

$$\frac{S}{\rho} \approx -\frac{1}{\rho} \left(\frac{dE}{dl} \right)_{el} = NZ_t \frac{4\pi Z_p^2 e^4}{m_e v^2} \left[\ln \frac{2m_e v^2}{I} - \ln(1 - \beta^2) - \beta^2 - \frac{C}{Z_t} - \frac{\delta}{2} \right], \quad (2.2)$$

where N denotes the number of target atoms (or molecules) per unit volume, Z_t the atomic number of the target atom, Z_p the nuclear charge of the projectile, e the charge and m_e

the mass of an electron, v the velocity of the projectile, I the mean ionisation energy of the absorbing material and $\beta = v/c$ the ratio of the projectile velocity and the speed of light. The shell correction term is significant at proton energies below approximately 10 MeV only [41]. It corrects for the fact that the approximation of stationary electrons is no longer valid at small particle velocities. The density correction term takes into account that incident protons polarise the target medium leading to an effective shielding of the electric field far from the particle track. This effect reduces the stopping power but is only relevant for proton energies above several hundred MeV [41].

The mass stopping power is proportional to the inverse square of the velocity of the projectile, the square of the particle charge and the electron density of the absorber (NZ_t).

The linear energy transfer (LET) L_Δ is the differential quotient of deposited energy per unit length of a charged particle within a tube-shaped volume around the particle track. The volume is determined by the range of secondary electrons whose kinetic energy equals the cut-off value Δ . The LET can thus be written as the difference of the electronic stopping power and the kinetic energy of electrons leaving the regarded volume (E_Δ^δ) on the line segment dl :

$$L_\Delta = \left(\frac{dE}{dl} \right)_{el} - \frac{dE_\Delta^\delta}{dl}, [L_\Delta] = \frac{keV}{\mu m}, \quad (2.3)$$

where electrons with kinetic energies smaller than Δ are considered to be part of the track core and L_∞ equals the electronic stopping power [37]. Heavy charged particles, such as carbon ions, are rated as high-LET radiation, i.e. densely ionising, whereas photons or electrons are rated as low-LET radiation, i.e. sparsely ionising. Therapeutic proton beams are usually considered as low-LET radiation.

The absorbed dose D is given by the mean local energy $d\bar{\epsilon}$ delivered by ionising radiation to a mass element dm with density ρ and volume dV :

$$D = \frac{d\bar{\epsilon}}{dm} = \frac{1}{\rho} \cdot \frac{d\bar{\epsilon}}{dV}, [D] = Gy, \quad (2.4)$$

where one gray (Gy) equals one Joule per kilogram [37]. It can be calculated based on the stopping power and particle fluence Φ (number of particles per unit area) using the following formula [42]:

$$D = \frac{S}{\rho} \cdot \Phi. \quad (2.5)$$

2.1.3 Depth dose profile

The depth dose profile of monoenergetic protons is depicted in Figure 2.1a. With increasing depth different effects contribute to the Bragg peak curve:

- **Electronic buildup region:** Incident high-energy protons liberate secondary electrons which have sufficient energy to produce further ionisation in the tissue several millimetres away from the primary proton beam (delta rays). The electronic buildup region extends from the surface to a depth which corresponds to the highest range of a recoil electron ($E_{electron}^{max} \approx E_{proton}/500$ [36]). Within that region, the dose increases approaching asymptotically the dose level of the sub-peak region. The electronic buildup might not be observable if material upstream of the surface, e.g. beam shaping devices or patient immobilisation devices, create a charged particle equilibrium or if it is superimposed by the protonic buildup.
- **Protonic buildup region:** Secondary protons arising from non-elastic, proton-induced nuclear interactions cause an increase of the absorbed dose with depth near the surface of the absorber.

- Sub-peak region: The sub-peak region extends from the surface to the beginning of the peak. It results from the $1/v^2$ dependence of the energy-loss rate and nuclear interactions which remove protons from the incident beam and liberate secondary particles. Secondary particles exhibit shorter ranges and larger angles with respect to the incident beam. Therefore, they deposit their energy further upstream than primary protons. For small fields, the proton fluence on the central axis is reduced due to the lack of lateral multiple Coulomb scattering equilibrium.
- Bragg peak: As protons slow down, the energy-loss rate increases (Eq. 2.2) leading to a Bragg peak near the end of range, where protons are absorbed by the medium. Statistical fluctuations of the energy loss broaden the Bragg peak and lower the maximum deposited energy (energy-/ range-straggling) [35]. This effect is more pronounced for high energy particles which travel longer distances. The range of a monoenergetic proton beam is consequently an average quantity and defined as the depth in which 50% of the incident protons have stopped. This depth coincides with the position in the Bragg peak curve where the absorbed dose falls down to 80% of the maximum dose [43]. A good approximation of range for clinical applications, assuming a continuous energy loss and neglecting lateral scattering, is the projected range. It is given by the integral of the inverse stopping power over the kinetic energy in the interval $[E_{max}, E_{min}]$:

$$R(E) = \int_{E_{max}}^{E_{min}} \frac{1}{S(E')} dE'. \quad (2.6)$$

The introduction of a lower energy limit E_{min} serves to exclude the case $S(0) = 0$. It does not affect the result since very low particle energies only have minor contribution to the total path length. The correlation between range and energy determines the clinical energy spectrum for a therapeutic application [44].

- Distal fall-off region: The region downstream of the Bragg peak is called distal fall-off region. It is caused by range straggling. In addition a momentum spread, attributable to the energy selection system (Sec. 2.3.1), contributes to this region [45].

The dosimetric manifestation of the predominant interaction types can be summarised as follows: inelastic Coulomb interactions with atomic electrons determine the proton range in an absorber, elastic Coulomb interactions with atomic nuclei the lateral penumbra and nuclear reactions the peak-to-plateau ratio.

In proton therapy, the initial beam energy is varied to reach different depths. A homogeneous dose in the target volume can be created by combining multiple Bragg peaks with appropriate weighting factors to a so-called spread-out Bragg peak (SOBP) (see Fig. 2.1b).

Comparison with photon beams

Conventional radiotherapy usually utilises megavoltage x-rays with energies ranging from about 1 MeV to 20 MeV [46]. At this energy range, the Compton effect is the dominant interaction type in tissue. The photoelectric effect dominates in water only for energies below 30 keV and can be neglected above 100 keV. Photons with energies higher than 1.022 MeV can produce electron-positron pairs in interactions with the nuclear Coulomb field. But pair production only becomes relevant for energies above approximately 10 MeV in tissue.

When passing through matter, the fluence of a monoenergetic photon beam decreases exponentially. However, the depth dose curve does not reach its maximum at the surface as can be seen in Figure 2.1b. This is due to the buildup effect arising from Compton interactions. Secondary electrons produced close to the surface spread their kinetic energy over their range. Thus, they contribute to the dose deposition in deeper layers. The full buildup is reached at a depth which corresponds to the average range of secondary electrons. The range depends on the initial beam energy and is typically in-between 5 mm and 5 cm [44]. This allows better skin sparing compared to proton therapy. On the downside, photon beams deposit a high amount

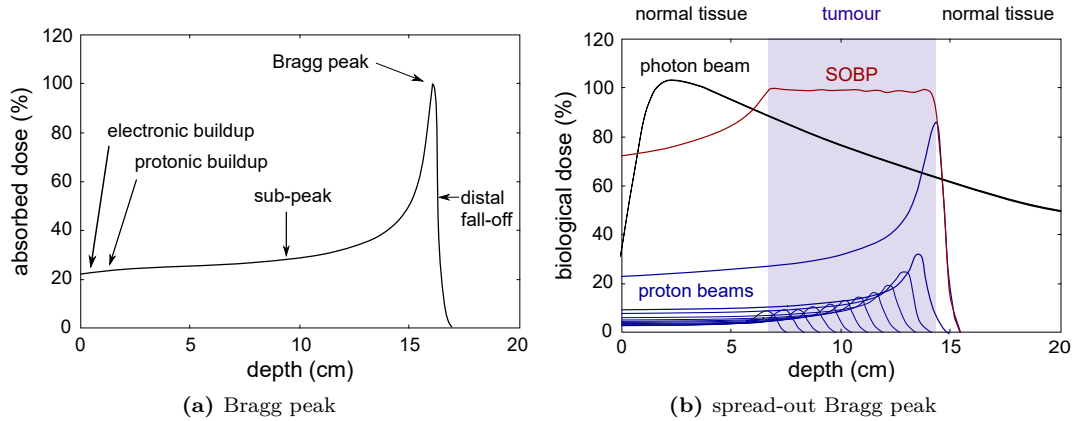


Figure 2.1: (a) Depth dose profile of a broad, monoenergetic proton beam (reproduction according to Ref. [35]): Due to the low entrance dose in the sub-peak region and the steep dose fall-off behind the Bragg peak, proton beams are particularly suitable for clinical use. The different regions of the Bragg peak curve and the underlying interactions are described in Section 2.1.3. (b) Depth dose profile of a spread-out proton Bragg peak (SOBP): Multiple pristine Bragg peak curves (blue lines) can be superimposed to an SOBP (red line) when using appropriate beam energies and weighting factors. Thus, it is possible to deliver a homogeneous dose to the tumour. In contrast to conventional photon beams (black line), proton therapy enables normal tissue sparing proximal and distal to the tumour. The biological dose states the photon dose which would yield the same biological effect as the absorbed proton dose (see Sec. 2.2.3).

of dose in the normal tissue proximal and distal to the tumour, while protons exhibit no exit dose and a lower entrance dose.

2.2 Cell survival and biological effectiveness

The availability of data on treatment outcome is limited for proton therapy which is why treatment concepts are usually based on experiences in the field of photon therapy. Although protons are in general considered as low-LET irradiation, their microscopic energy deposition pattern differs from that of photons resulting in a different biological effect [47]. This must be taken into account in treatment planning. The most studied endpoint to determine the relative biological effectiveness of protons is clonogenic cell survival which will be explained in the following. Molecular mechanisms of DNA and chromosome damage are not discussed in the scope of this thesis. A detailed description can be found in Ref. [48].

2.2.1 Cell survival

There are different definitions for the term cell survival and its antonym cell death. In radiobiology, cell death means the loss of ability to proliferate indefinitely. This is due to the fact that the prevention of tumour growth and dissemination of malignant tumour cells in the body is of major relevance. A cell survived by definition if it is able to produce a colony of at least fifty daughter cells within a certain period of time [49].

The biological effect of a test irradiation on a tissue can be studied based on cell survival curves. To this end, a known number of cells is irradiated in a Petri dish and after one to two weeks of incubation cell colonies are counted. The fraction of surviving cells is plotted semi-logarithmically against the dose. Characteristic curves for low- and high-LET radiation are depicted in Figure 2.2a. Low-LET radiation exhibits a finite initial slope at low doses, followed by a bending region (shoulder) and a linear decrease at higher doses (note the semi-logarithmic scale). In contrast, the curve for high-LET radiation does not feature a shoulder and runs almost straight over the whole dose range, i.e. the surviving fraction decreases exponentially with dose. As a consequence, a lower amount of dose is needed to reach the same level of

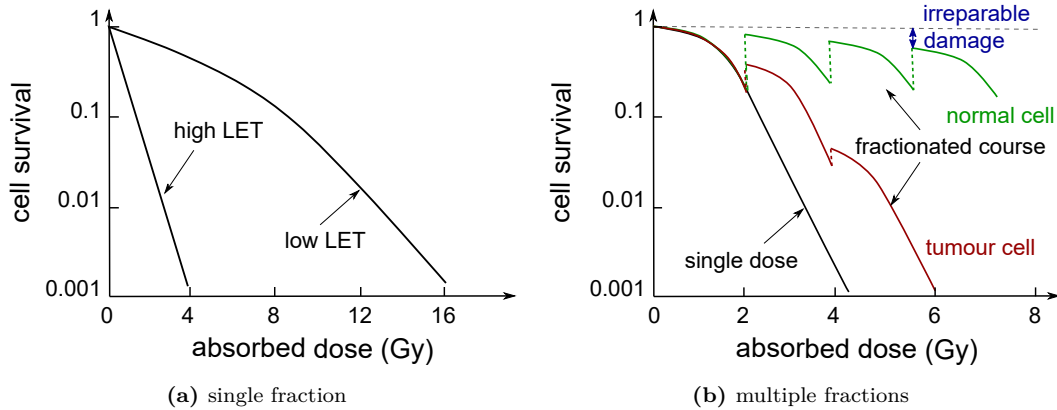


Figure 2.2: Schematic illustration of cell survival (reproduction according to Ref. [50]): (a) comparison of a densely ionising radiation (high LET) and a sparsely ionising radiation (low LET) for a single fraction, (b) comparison of a single fraction and a fractionated radiation course for normal tissue and tumour cells.

survival as with low-LET radiation. The shoulder in the low-LET curve indicates the ability of the cell to repair sub-lethal damage. A simple interpretation is given by the biologic concept of dual-radiation action below. The time needed for repair depends on the tissue type and takes several hours. Conventional radiotherapy exploits the fact that tumour cells have a lower ability to repair from sub-lethal damage than normal cells. The prescribed dose is given in a fractionated course with small daily doses. Thus, the normal tissue can recover in-between the treatment sessions (see Fig. 2.2b).

Linear-quadratic model

Mathematically, cellular survival can be described using the linear-quadratic (LQ) model. The model is based on the assumption that two components, a linear (proportional to dose) and a quadratic (proportional to the square of dose), contribute to the cell damage by ionising radiation. The fraction of cells surviving a dose D is given by

$$S = \exp(-\alpha D - \beta D^2) \quad (2.7)$$

and is usually fitted by a second-order polynomial in a semi-logarithmic plot: $-\ln(S) = \alpha D + \beta D^2$ (see Fig. 2.2). The constants α (Gy^{-1}) and β (Gy^{-2}) are chosen to obtain the best agreement with the experimental data and depend therefore on the irradiation and tissue type. The mathematic formulation as a two-component model might be justifiable by the biologic concept of dual-radiation action. According to the concept, the probability that two chromosome breaks are caused by a single particle track and form a lethal exchange-type aberration¹ increases linearly with dose. If the two breaks result from two independent tracks, however, the probability is proportional to the square of the dose. For $D = \alpha/\beta$, both components contribute equally. The ratio of α to β determines the shape of the curve and therefore the radiation sensitivity of a tissue. Late-responding tissues exhibit small α to β ratios, whereas early-responding tissues are characterised by a large α to β ratio [48].

2.2.2 Relative biological effectiveness

Cell experiments have demonstrated the higher biological effectiveness of protons compared to photons. The relative biological effectiveness (RBE) is defined by the ratio of the reference

¹Chromosomal aberrations: changes in the normal structure or number of chromosomes.

dose D_{ref} , e.g. the dose of 6 MV photons, and the proton dose D_{test} which lead to the same biological endpoint (e.g. cell survival):

$$RBE = \left[\frac{D_{ref}}{D_{test}} \right]_{isoeffect}. \quad (2.8)$$

The RBE depends on various parameters, such as dose, biological endpoint, irradiated tissue, number of dose fractions, dose rate and LET. In proton therapy, a generic RBE value of 1.1 is used. This value is a conservative estimate to ensure that the tumour receives the prescribed dose despite considerable uncertainties in the experimental RBE determination [47].

2.2.3 Biological dose

The biological dose D_{RBE} is defined as the product of the absorbed dose and the RBE:

$$D_{RBE} = D \cdot RBE. \quad (2.9)$$

It equals the photon dose which would achieve the same biological effect as the radiation under consideration. D_{RBE} is an important quantity in particle therapy because it incorporates the higher biological efficacy of charged particles in the optimisation process.

2.2.4 Equivalent uniform dose

The equivalent uniform dose is the absorbed dose which yields the same radiobiological effect (cell survival), when given homogeneously to the whole tumour, as a studied non-uniform dose distribution. A general formula, valid for tumours and normal tissue, has been introduced by Niemierko in 1997 [51] and is hereinafter referred to as generalised equivalent uniform dose (gEUD):

$$gEUD = \left(\sum_i v_i D_i^{1/n} \right)^n, \quad (2.10)$$

where v_i denotes a fractional organ volume and D_i the corresponding absorbed dose. The tissue-specific parameter n indicates the dependence on the irradiated partial volume and is the only parameter which has to be fitted for a given radiation and tissue type. The formula is based on the power law behaviour of cell survival with dose.

2.3 Proton beam therapy system

The West German Proton Therapy Centre Essen (WPE) is equipped with a cyclotron-based IBA Proteus®Plus proton therapy system (IBA, Louvain-La-Neuve, Belgium). It has four different treatment rooms: three 360° beam rotation isocentric gantry rooms and one 90° fixed beam line room including an eyeline, which is not yet clinically used. A floor plan of the proton therapy facility can be seen in Figure 2.3. The following sections describe the generation of a therapeutic proton beam at WPE.

2.3.1 Components

Proton source

Protons are produced by a hot filament Penning Ion Gauge (PIG) source in the centre of the cyclotron. Free electrons originating from the hot filament are accelerated in a strong electric field and ionise inflowing molecular hydrogen (H_2) gas in a hollow cylinder (chimney). The ions and electrons form a plasma which shields external electric fields. Only ions diffusing through a slit in the wall of the gas chimney experience the electric field of the closest cyclotron electrode. The beam intensity is regulated by the filament heating through a feedback loop [53, 54, 55].

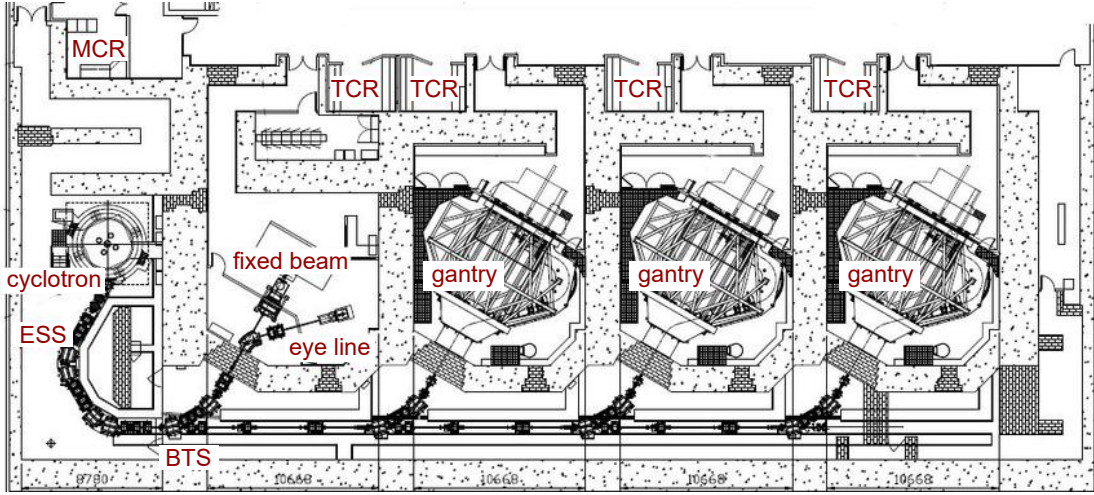


Figure 2.3: Layout of the IBA proton therapy system at WPE showing the main control room (MCR), treatment control rooms (TCRs), the cyclotron, the energy selection system (ESS), the beam transport and switching system (BTS), the fixed beam treatment room including an eye line and three gantry treatment rooms [52].

Isochronous cyclotron

An isochronous cyclotron accelerates the protons to a fixed energy of about 227 MeV. In contrast to a classical cyclotron, the magnetic field increases with radius to compensate for the relativistic mass increase at high particle energies [53]:

$$B(r) = m(r) \frac{B_0}{m_0} = \gamma B_0, \quad (2.11)$$

where $\gamma = 1/\sqrt{1-(v/c)^2}$ denotes the Lorentz factor. Thus, the azimuthal frequency of the protons becomes independent of the radius and remains synchronised with the radio frequency (RF) voltage between the cyclotron electrodes:

$$\omega = \frac{v}{r} = \frac{qB(r)}{m(r)} = \text{constant}. \quad (2.12)$$

The main advantage of an isochronous cyclotron is the quasi-continuous character of the beam. Prior to the invention of isochronous cyclotrons, the RF frequency had to be adjusted synchronously to the proton mass (synchrocyclotron) so that only particle bunches with phase parameters matching the RF frequency could be accelerated [54]. In IBA cyclotrons, the increase of the magnetic field with radius is realised by a decreasing gap between the magnetic poles. The increasing magnet field, however, compromises the vertical beam stability. Without corrections vertical deflected protons would be pushed towards the poles and be lost in collisions [53]. To reduce the particle loss, hills with stronger magnetic fields (small gaps) and valleys with weaker magnetic fields (large gaps) are added to the magnetic pole. The azimuthally varying magnetic field provides a focusing force keeping the protons within a horizontal plane.

Energy selection system and beam transport and switching system

At the exit of the cyclotron, the beam energy is reduced to an appropriate therapeutic value by the energy selection system. The energy selection system consists of a wheel-mounted energy degrader wedge of variable thickness and material, a magnetic analyser and slits defining the energy width. Beam energies from 100 MeV to 227 MeV are clinically available. From the energy selection system, the proton beam is transported within a vacuum tube to the treatment rooms, whereby dipole and quadrupole magnets serve to bend and focus the beam. For the

field shaping, it is important that the beam transport and switching system provides a centred beam with appropriate optical characteristics such as spot size, beam intensity, average energy, energy spread and angular spread (emittance).

Nozzle

The delivery technique and thus the beam characteristics are determined by the beam delivery system ('nozzle') at the exit of the beam line. The nozzle contains all technical devices which are required to shape the irradiation field and to monitor the beam properties. It is either mounted horizontally (fixed beam treatment room) or on a 360° beam rotation isocentric gantry. At WPE, dedicated pencil beam scanning nozzles and universal nozzles supporting different delivery techniques are available for treatment (see below).

2.3.2 Beam delivery: Pencil beam scanning (PBS) and double scattering (DS)

In order to distribute a prescribed dose over the whole target volume the irradiation field has to be shaped in lateral and longitudinal direction. The longitudinal modulation of the beam is done by adjusting the beam energy according to the depth and thickness of the tumour in beam direction. As already mentioned, a homogeneous depth dose distribution (SOBP) can be created by combining different beam energies with appropriate weighting.

The lateral shaping of the field can be active or passive. In an active system, a narrow beam is scanned across the target energy layer by energy layer. Whereas in a passive system, the beam is scattered to increase the beam cross section and to spread the dose over the target. In the following, the two delivery techniques which are referred to in this thesis will be described.

Pencil beam scanning is an active delivery technique in which the narrow beam scans the target three-dimensionally. A dedicated pencil beam scanning nozzle is depicted in Figure 2.4a. It includes

- an ionisation chamber (IC1) to monitor the centring of the beam at the nozzle entrance,
- two quadrupoles to focus the beam,
- a pair of dipole magnets to scan the beam laterally,
- an x-ray tube which can be inserted and retracted to enable beam's eye view imaging,
- two ionisation chambers (IC2 and IC3) to measure the dose and to verify the position, size, and shape of the pencil beam
- and a removable range shifter which enables range degradation up to the patient's skin to treat superficial tumours.

At WPE, scanning is performed in discrete steps with a fast and a slow magnet [56] using a 'step-and-shoot' method (spot scanning). In contrast to passive systems, this technique allows to modulate the intensity and enables thus multi-field optimisation. In spot scanning, the intensity is modulated by varying the dwell time of the pencil beam at each spot position. Therefore, no modulation in scanning speed or beam intensity is required. The irradiation is delivered energy layer by energy layer and starts with the most distal range (highest energy) (see Fig. 2.5). Spots optimised by the treatment planning system are reordered by the IBA scanning control system aiming for a short delivery time. The spots are placed consecutively on a line along the fast scanning direction. After changing the position in the slow direction the order within the next line of spots is optimised with regard to the shortest scanning path in-between the lines. The difference in the scanning speeds results from the fact that the dedicated PBS nozzle is equipped with the same scanning magnets which were originally developed for optimum beam delivery in uniform scanning mode (diamond pattern). Besides, the magnets have a different distance to the isocentre (see Fig. 2.4a). The fast scanning magnet is ten times faster than the slow scanning magnet.

In the following, typical clinical values for spot size, spot distance, spot weight and energy layer spacing are presented. The spot size at the isocentre ranges from about 6.6 mm for 227 MeV to 13.4 mm for 100 MeV in air (full-width at half-maximum) [57]. The spot spacing can be scaled automatically depending on the radial spread in the Bragg peak for an energy layer or is chosen to be constant. Usually, a spot spacing of 6 mm to 8 mm is applied. Spot weights used in clinical practice range from 0.04 monitor units (MUs) to a few MUs, where 1000 MUs correspond to about 108 cGy within the entrance plateau of the Bragg peak curve for a 10 cm x 10 cm field, 160 MeV and a spot spacing of 2.5 mm [57]. The energy layer spacing is scaled depending on the width of the Bragg peak. The distance increases with larger scaling factor. In clinical scenarios, the energy of two adjacent layers differs by approximately 2 MeV to 4 MeV.

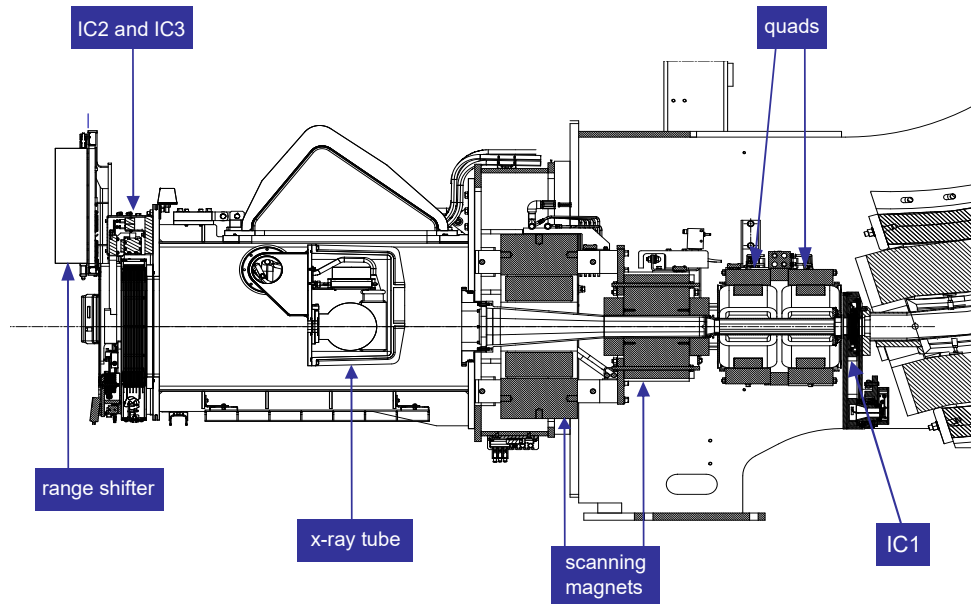
The beam delivery is documented in irradiation log files which record parameters such as the spot position and the amount of charge collected by the ionisation chambers on a 0.25 ms time grid. Variables which determine the time structure are the spot delivery time, the scanning speed and the energy layer switching time. The spot delivery time varies with the beam current. Depending on the material and the thickness of the energy degrader for a certain range, a higher beam current is requested by the IBA system to compensate for scattered protons and a reduced transmission efficiency of the energy selection system. The lowest energy selection system efficiency determines the maximum beam current requested from the cyclotron at the beginning of each irradiation. During the irradiation, the 'beam current regulation unit' adjusts the current for each energy layer by tuning the arc voltage of the ion source. A feed-back is received by an ionisation chamber at the exit of the cyclotron.

The scanning magnets are driven by a digitally controlled scanning magnet power supply, which consists of two pulse width modulation amplifiers. Due to the 'slew rate' of the magnet power supply, the time to move from one spot position to the next will never be zero [53]. The settling time for the magnetic field takes several hundred microseconds. Since the magnets are prone to hysteresis effects, longer settling times are required for larger distances.

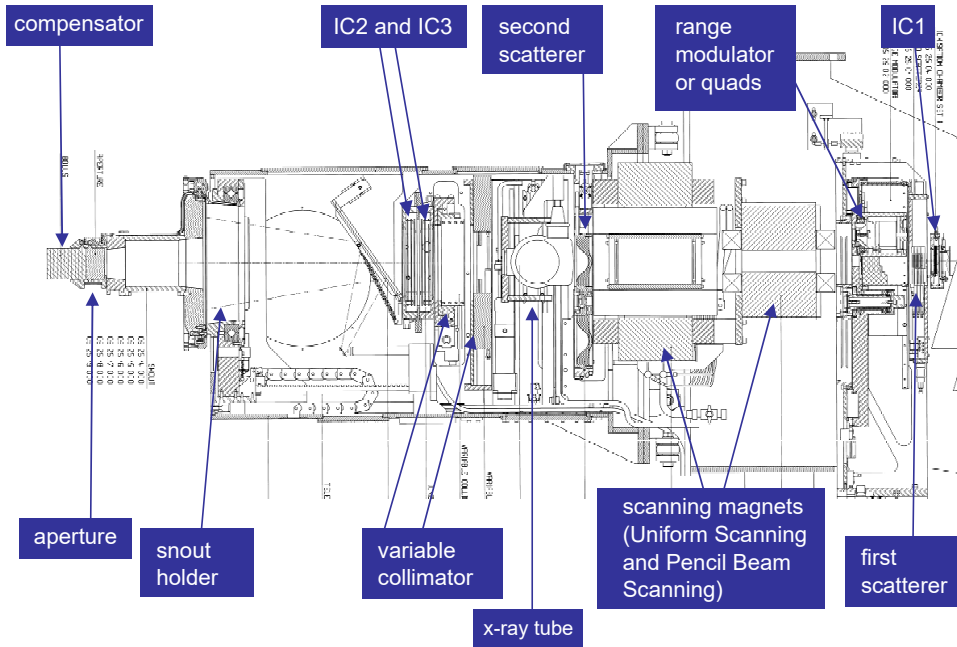
A change of the beam energy involves adaptations of all magnets in the beam line. The total energy layer switching time takes less than 2 s according to the Proteus[®]Plus product description [55].

Double scattering is a passive beam delivery technique in which the narrow beam is spread laterally by a pair of scattering devices on the central beam axis (see Fig. 2.6). The first spreading device, a lead foil, increases the width of the beam based on multiple Coulomb scattering in the high-Z material. After passing the first scatterer, the lateral beam profile has an approximate Gaussian shape. A flat and symmetric beam profile is however required to generate a uniform dose distribution. Without a second scatterer, only the beam fraction around the centre exhibiting a rather flat intensity profile is clinically usable (single scattering technique). Since most of the beam has to be blocked by a collimator in single scattering, the system efficiency is very low and only small target volumes can be treated. DS introduces a second scatterer composed of high- and low-Z material to flatten the beam on a large field area (maximum field radius: 12 cm). The thickness of the high-Z material (lead) decreases with increasing radius. Thus, protons passing the thicker, central part are scattered more than those hitting the outer regions of the foil. The dome shape of the lead component is designed such that the amount of in and out scattered protons yields a flat lateral intensity distribution at the height of the isocentre. To compensate for resulting differences in energy loss, a low-Z component (lexan) is added to the foil. The thickness of the lexan component is larger the thinner the lead foil is so that the range is the same for all beam rays. Due to the contoured scatter foil, an accurate centring of the beam is particularly important in DS [55].

The modulation of the beam energy in DS is done by a range modulator wheel which is located in-between the first and the second scatterer. The range modulator is divided into steps of different material thicknesses and spins at 600 rounds per minute [55]. It is composed of low-Z material (carbon/lexan) and high Z-material (lead), whereby the low-Z material serves to lower the beam energy. The high Z-material is used to adjust the amount of scattering such that each step has the same scattering effect on the beam independent of the range. The Bragg peaks



(a) IBA's pencil beam scanning nozzle



(b) IBA's universal nozzle

Figure 2.4: Technical drawing of a dedicated pencil beam scanning nozzle from IBA (a) and a universal nozzle supporting four different delivery types (b) [55]. A description of the elements is given in Section 2.3.2.

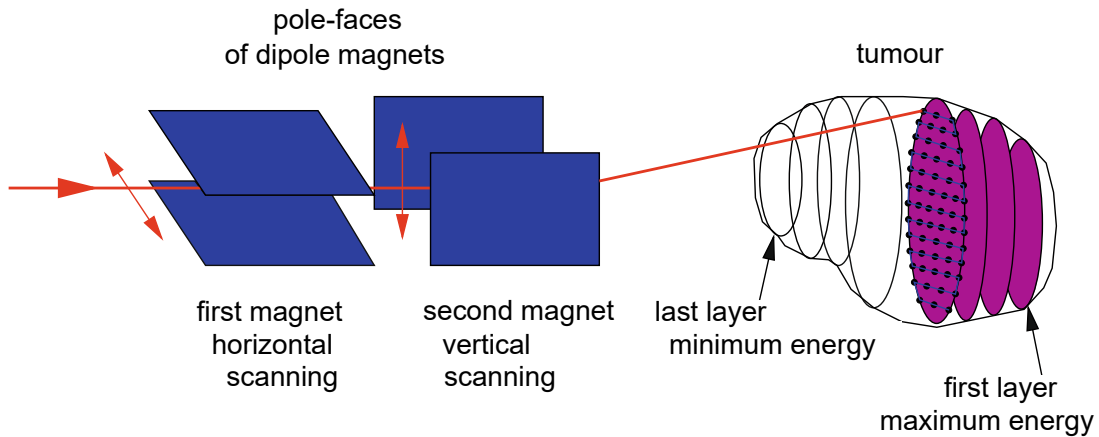


Figure 2.5: Schematic drawing of the pencil beam scanning (PBS) delivery technique described in Section 2.3.2 (reproduction according to Ref. [58]).

are weighted by the angular length of the corresponding wheel step in order to create a flat SOBP. The optimum Bragg peak weights depend on the beam energy at the nozzle entrance which is set by the energy selection system dependent on the most distal point of the target. To use the same wheel for different energy spans, the weight of each step has to be fine-tuned by modulating the beam current with angular rotation. The respective values are stored in a beam current modulation file. IBA developed three wheels each containing three concentric tracks dedicated for different clinical demands. Eight out of the nine tracks are available in DS to create SOBPs of variable modulation and range according to the maximum depth and thickness of the tumour.

Patient-specific beam shaping devices are required to adjust the field laterally and longitudinally to the shape of the target. In lateral direction, field shaping is done by the use of an aperture made of brass (see Fig. 2.7a). It limits the field to the size of the target from beam's eye view so that different apertures have to be fabricated for each beam angle. To minimise the penumbra the aperture is mounted close to the patient on the so-called 'snout'. The snout holds all patient-specific devices. Unfortunately, the majority of unwanted neutron dose received by the patient is produced by the aperture. A variable collimator behind the second scatterer serves to minimise the neutron amount by reducing the field size to the smallest rectangle that still encloses the patient aperture.

For longitudinal beam shaping, an individually produced range compensator is attached to the snout right behind the aperture. The range compensator represents the last element in front

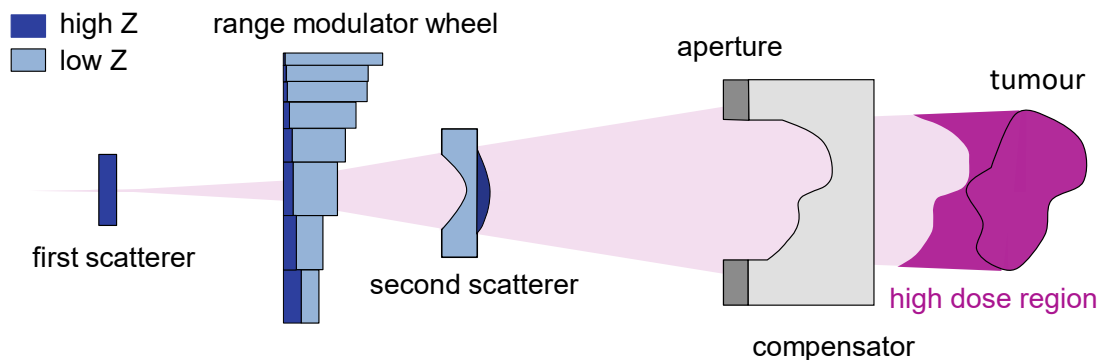


Figure 2.6: Schematic drawing of the double scattering (DS) delivery technique described in Section 2.3.2.

of the patient. It is made of polymethyl methacrylate (PMMA) and conforms the distal range of the SOBP to the distal shape of the target including a safety margin. The thickness of a compensator pixel is given by the treatment planning system. Due to multiple Coulomb scattering, setup uncertainties and anatomical changes, range gradients between neighbouring pixels cannot be infinitely large. In a smearing process, the thickness of each pixel is reduced to the smallest thickness of all neighbouring pixels whose centre is located within a specified smearing radius from the centre of the current pixel to add robustness against position uncertainties and to prevent from undershooting [59]. The arrangement of all technical devices in the universal nozzle can be seen in Figure 2.4b.

An advantage of DS is the continuous repainting of the whole tumour. There is negligible energy layer switching time. Thus, the treatment is faster and less sensitive to organ motion. Disadvantages of DS in comparison to PBS include a lower dose conformality proximal to the target, a higher neutron production, a higher sensitivity to beam misalignment compromising the lateral uniformity and additional costs for the production of individual beam shaping devices.

2.4 Treatment planning

Treatment planning aims at finding the optimal dose distribution for each individual patient to increase the probability of a curative treatment without side effects. A decisive factor for the treatment quality is the choice of the treatment plan parameters such as spot positions and weights in PBS or the shape of blocks and compensators in DS. Due to the large number of variables, forward planning, a manual trial-and-error process, is not a convenient strategy to find the best suited set of parameters. For that reason, most of the beam parameters are determined by automated computation methods based on specified plan criteria (inverse planning). Only the number of fields and the couch and gantry angles are chosen manually by the treatment planner. Finding the optimal dose distribution still remains an iterative process. After each dose calculation the optimised plan has to be analysed and specified criteria for targets and OARs may have to be adjusted in order to improve the plan quality. In general, a good treatment plan is characterised by a high target coverage and low doses to surrounding healthy tissues. However, in some cases OARs overlap with the target volume or are very close to it so that trade-offs must be made.

2.4.1 Treatment planning system (TPS)

The treatment planning system RayStation from RaySearch (RaySearch Laboratories AB, Stockholm, Sweden) is clinically used for treatment planning and evaluation at WPE since 2015. In Europe, WPE was the first proton therapy centre treating a patient with RayStation. The planning software provides dose engines for photon, electron, proton and carbon ion therapy and features advanced treatment planning techniques such as Multi-Criteria Optimisation or 4D treatment planning.

A scripting interface, in form of an embedded IronPython console, offers maximum flexibility. IronPython is an open-source implementation of the scripting language Python with access to the .NET Framework [60]. It allows for automation of treatment planning processes, execution of actions which are not available in the standard interface and systematic, retrospective data analysis of patient cohorts. Thus, scripting in RayStation does not only save valuable time in clinical routine, but also represents an important research tool. An example of a script, which was developed within the scope of the dissertation, is provided in section A.1 in the appendix. In this thesis, the research versions of RayStation 5, 6 and 7 operated within the clinical installation were used to compute dynamic dose distributions in the presence of motion and to evaluate and compare different treatment planning techniques for moving targets.

2.4.2 Computed tomography (CT)

Computed tomography is a diagnostic imaging technique in which a series of x-ray images is taken from different angles and is processed into cross-sectional images of inner physical structures to get a three-dimensional image of the patient. The attenuation of x-rays in the

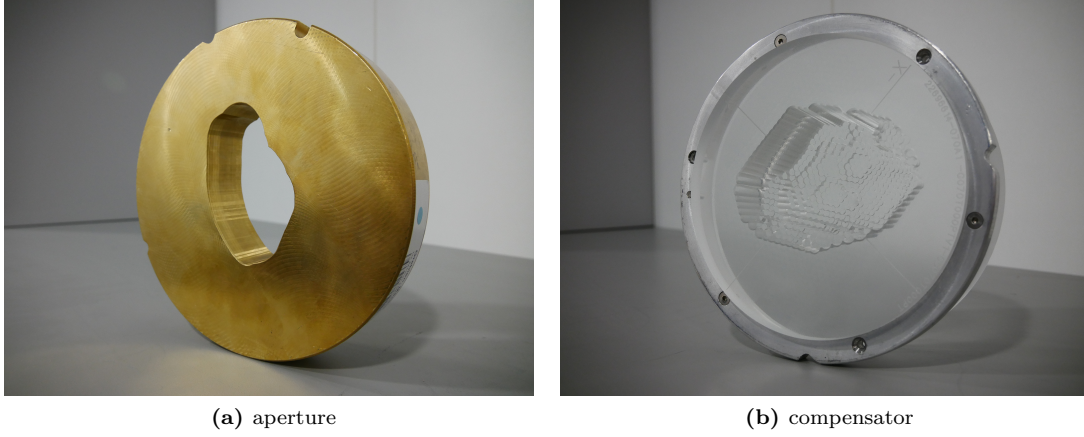


Figure 2.7: Patient and field specific beam shaping devices in passively scattered proton therapy.

patient's body yields information on tissue characteristics. In proton therapy planning, CT data sets are needed to determine the water equivalent path length of protons in tissue. The current generation of CT scanners produces narrow, cone-shaped x-ray beams. During a CT scan, the x-ray tube and an opposing detector array rotate around the patient while the patient table is moving through the scanner (see Fig. 2.8). Thus, it is possible to acquire intensity profiles from different angles. The attenuation of the photon intensity after passing objects can be described by

$$I = \int_0^{E_{max}} I_0(E) \cdot e^{-\int_0^L \mu(E) ds} dE, \quad (2.13)$$

where E_{max} is the maximum photon energy, I_0 the incident photon spectrum, L the total path length and μ the photon attenuation coefficient. The attenuation coefficient μ states the probability of photon interactions for a tissue/material and is given by the product of the atomic cross section σ_a and the number of atoms in the absorber volume n :

$$\mu = n \cdot \sigma_a = N_A \cdot \rho \cdot \frac{Z}{A} \cdot \sigma_e \quad (2.14)$$

with the density ρ , the atomic number Z , the mass number A , Avogadro's number N_A and the electron cross section σ_e [61]. μ depends on the electron density, the atomic number and the photon energy. The latter is expressed by the electron cross section in Equation 2.14. For diagnostic x-ray energies between 30 peak kilo voltage (kVp) and 140 kVp, photoelectric absorption processes as well as incoherent and coherent scattering contribute to the atomic cross section:

$$\sigma_a = \sigma_a^{photo} + \sigma_a^{coh} + \sigma_a^{incoh}. \quad (2.15)$$

Profile data from many directions are taken to calculate μ for individual volume elements (voxel) and reconstruct the image. Comparability between different CT scanners is established by defining a CT number which states the photon attenuation with reference to water [62]:

$$CT\ number = \frac{\mu_{tissue} - \mu_{water}}{\mu_{water}} \cdot 1000, [CT\ number] = HU. \quad (2.16)$$

The CT number is measured in Hounsfield units (HU). By definition, the CT number of water is 0 HU and the CT number of air -1000 HU. Bones can exhibit CT numbers of up to 3000 HU, whereas the CT numbers of soft tissues and fluids typically range from about -100 HU to +100 HU.

In proton therapy, a CT calibration relating CT numbers to stopping powers is required for dose computation. RayStation's dose engine calculates the stopping power based on the Bethe-Bloch formula (Sec. 2.1.2). As a consequence, mass density, mean ionisation energy and material composition of a voxel must be deduced from the CT.

The elemental composition and mean ionisation energy are defined based on a list of reference materials stored in RayStation. For each reference material, the mass fraction of atomic elements $\omega(Z)$ has been interpolated from ten well-established core materials which are found in patients, such as lung, muscle or aluminium [59]. The reference material with the most similar mass density is assigned to a voxel. The mass density itself is not taken from the reference material. Instead, CT numbers are converted to mass densities by linear interpolation from CT-to-density calibration curves. Since the CT number for a given mass density can vary with the CT scanner, the scan protocol or the patient geometry, it might be necessary to create different calibration curves [59]. At WPE, for example, calibration curves exist for CT scans of the patient's head at 120 kV, the patient's trunk at 120 kV and the patient's trunk at 140 kV, the latter being used for corpulent patients.

Common calibration approaches are the tissue substitute and the stoichiometric method. The tissue substitute calibration is straightforward. A CT scan of a phantom with different tissue surrogates of known densities is acquired. Subsequently, the mass densities are assigned to the measured CT numbers. The stoichiometric calibration requires more effort but yields higher accuracy for biological tissues [63]. It considers also the chemical composition of the materials which differs for tissue substitutes and real tissues and affects the attenuation of x-rays. For photon energies of up to 1 MeV, the attenuation of x-rays can be described by the photoelectric effect, incoherent scattering and coherent scattering. In the stoichiometric approach, the contribution of each of those effects to the linear attenuation coefficient is calculated as a function of the atomic number based on CT scans of tissue surrogates of known chemical composition. The contributions are specified by the so-called K-coefficients and depend on the CT scanner and the used scan protocol. Therefore, they have to be determined for each calibration individually. Once the K-coefficients are known, theoretical CT numbers can be calculated for a wide variety of biological tissues, whose chemical composition is taken from literature.

The tissue substitute method was applied within the scope of this thesis to create an individual calibration curve for the used, artificial phantom materials (Sec. 3.2.1), whereas a stoichiometric calibration curve was applied to convert CT numbers to stopping powers for CT data of real patients (Sec. 4.2.2).

If a material does not fit to a CT-to-density calibration curve or CT artefacts distort the measurement, it is possible to override the affected volume by a specified mass density and elemental composition.

2.4.3 Contouring

Contoured structures in form of regions of interest (ROIs) are required to guide the treatment plan optimisation. They are used for the specification of site-specific dose limits in PBS and for the definition of margins in the treat and protect settings in DS. Target volumes and nearby OARs are delineated manually on 2D slices of tomographic image data sets or with the aid of segmentation tools.

Target volume definitions in radiotherapy are described in ICRU Report 78 [65] and are illustrated in Figure 2.9. The gross tumour volume (GTV) refers to the macroscopic malignancy, which is clinically demonstrable. In clinic, an extra margin is added to the GTV to account for a sub-clinical, microscopic spread of tumour cells. The expanded volume is called clinical target volume (CTV). Expected changes in shape, size and position of the CTV due to physiological processes are considered by an internal margin. The corresponding volume is termed internal target volume (ITV). It can be generated for example by encompassing the CTV in all respiratory phases of a respiratory correlated CT (Sec. 2.5.4). Motion mitigation techniques as respiratory gating or abdominal compression reduce the size of the ITV (Sec. 2.5.5). The planning target volume (PTV) takes additionally uncertainties in patient positioning and proton range calculation into account.

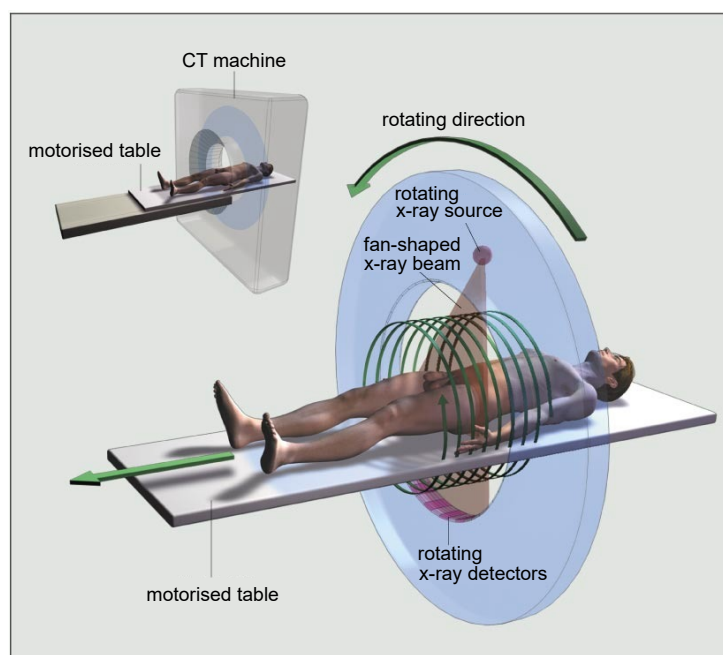


Figure 2.8: Schematic illustration of a CT scan [64]: The x-ray source and an opposing detector array rotate synchronously around the patient. At the same time, the motorised patient table moves continuously through the scanner resulting in a helical scan.

In RayStation, a ROI covering the patient’s body as well as all support and fixation devices is required for each plan. The so-called External defines the volume for the dose computation problem. Every structures which lies outside of the External or the dose grid is considered as vacuum [59]. In addition, a ROI which presents the patient’s body in the CT image is created in treatment planning. The body ROI facilitates contouring of OARs close to the surface with segmentation tools, can be applied as controlling structure or focusing region in deformable image registrations (see Sec. 2.4.4) and is used to determine the dose maximum within the patient.

2.4.4 Deformable image registration

Deformable image registrations (DIRs) were performed using the ANAtomically CONstrained Deformation Algorithm (ANACONDA) [66]. ANACONDA pursues a hybrid approach which combines intensity and geometric based algorithms. The registration involves the computation of a deformation vector field whose parameters are determined by solving a non-linear optimisation problem. The objective function, which is minimised during the optimisation process, is formed by the weighted sum of four non-linear terms:

1. an image similarity term to optimise the similarity between the deformed image and the target image by means of the Pearson correlation coefficient,
2. a grid regularisation term to ensure local smoothness and invertibility of the deformation vector field,
3. a shape based grid regularisation term to inhibit large shape deviations of selected contoured structures (controlling structures) and
4. a penalty term to guide the algorithm such that controlling structures in the reference image are deformed to the corresponding structures in the target image.

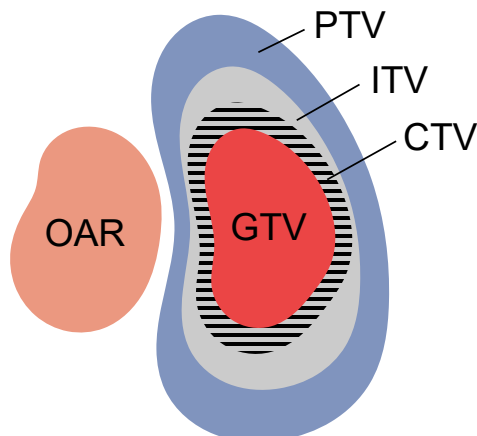


Figure 2.9: Schematic illustration of target volume definitions in radiotherapy: gross tumour volume (GTV), clinical target volume (CTV), internal target volume (ITV) and planning target volume (PTV). An organ at risk (OAR) is depicted next to the tumour.

A detailed mathematical description can be found in reference [66]. The image similarity term is invariant to linear intensity transformations. Thus, ANACONDA can be used when images describe the same data but were acquired with different imaging modalities, e.g. CT and cone beam CT. By including ROIs or points of interest (POIs) as controlling structures or focusing regions a higher accuracy of the DIR can be reached. A focusing ROI is applied to reduce the size of the region considered by the algorithm for large scans or to exclude regions containing image artefacts. ANACONDA has been validated in terms of landmark tracking accuracy, contour propagation accuracy and image similarity [66, 67, 68, 69].

2.4.5 Proton dose calculation

RayStation provides two different dose engines for PBS, one that is based on a Pencil Beam algorithm and one that is based on Monte Carlo (MC) simulations. The Pencil Beam algorithm is usually faster but less accurate than MC. Therefore, it is recommended to use the MC engine in cases which exhibit large inhomogeneities (e.g. lung tissue) or include a range shifter [59, 21, 23, 25, 24].

At the time when the PhD project started, the MC engine was not yet implemented in RayStation (version 5) and the latest TPS version (RayStation version 7) supports MC dose calculations only for PBS and not for DS. For that reason, the main focus of this section is on the Pencil Beam algorithm. Nonetheless, a short description of the MC dose engine is provided since the project was complemented by a review of the calculated PBS dose distributions with MC on a sample basis.

Pencil Beam dose engine

The Pencil Beam algorithm assumes that any incident proton beam can be regarded as an accumulation of smaller, narrow beams of pencil like shape, giving rise to the name of the algorithm. These pencil beams should not be confused with the physical pencil beams which are delivered in the PBS technique ('spots'). The algorithm uses a infinite slab approximation, i.e. the patient is treated as a stack of semi-infinite layers and lateral inhomogeneities are disregarded. The decomposition of spots into pencil beams is made to better account for inhomogeneities in the patient within the volume of a spot and thus reduces errors due to the infinite slab approximation. Given the differences in mean direction of protons at a certain point in the fluence plane for different spots, every spot has to be handled separately and it is not appropriate to use a common fluence grid in PBS. Instead, RayStation arranges 18 pencil beams per spot in two concentric circles around a high weighted centre pencil beam. The

number of 19 sub-spots is a compromise between computation time and sufficient accuracy in the sampling of the patient geometry and the reproduction of the initial spot fluence. The distance between sub-spots depends on the spot size.

In contrast to PBS, DS exhibits a uniform intensity distribution within an energy layer and the fluence plane can be discretised into an orthogonal grid, each pixel corresponding to one pencil beam.

For both, PBS and DS, the dose contribution of a pencil beam along its central axis z is determined by the product of the lateral proton fluence Φ and the integrated depth dose IDD :

$$dose(z, r) = \Phi(z, r) IDD(z), \quad (2.17)$$

where r is the radial coordinate in the lateral plane [59]. The lateral proton fluence can be deduced from the phase space parameters which are computed along the central axis of each pencil beam with a sub-spot multi tracing algorithm considering stopping power, multiple Coulomb scattering (MCS) and non-elastic nuclear interactions. The effect of energy loss straggling is not explicitly accounted for during ray tracing but it is included in the IDD curves.

In PBS, the initial phase space parameters of a spot are the mean direction of the protons, the energy spectrum dN/dE , spatial-angular distribution moments (angular variance Θ^2 , spatial-angular covariance $r\theta$ and spatial variance r^2 of the Gaussian spot fluence) and the absolute spot weight ω^a , whereas in DS the parameters for an energy layer e include the mean direction of the protons, the mean beam energy at the patient surface $E_{m,e}$, the integrated depth dose curve in water IDD_e^w (without compensator), the local angular variance at the compensator entrance plane $\Theta_{C,e}^2$ and the relative weight of the energy layer ω_e . The ray tracing grid equals the dose grid and the tracing process stops either when the ray leaves the grid or when the energy falls below 1 MeV. The average HU value of a voxel is calculated based on the planning CT, which often has a higher resolution.

The integrated depth dose for energy layer e and pencil beam j at depth z_{ej} is given by the reference integrated depth dose curve in water IDD_e^w at the water equivalent depth z_{ej}^w :

$$IDD_{ej} = IDD_e^w(z_{ej}^w). \quad (2.18)$$

z_{ej}^w equals the integral of the medium to water stopping power ratio over the path length z_{ej} :

$$z_{ej}^w(z_{ej}) = \int_0^{z_{ej}} \frac{S(z')}{S_w(z')} dz'. \quad (2.19)$$

The water reference IDD with the unit cGy cm²/proton is the weighted superposition of Monte Carlo pre-calculated, mono-energetic IDs and provided by the beam model.

The stopping power model used in RayStation is based on the Bethe-Bloch formula omitting the clinically non-relevant shell and density correction terms (Sec. 2.1.2). It serves to compute water equivalent depths (Eq. 2.19) and absolute energy loss.

The lateral spread of the proton beam is caused by MCS and nuclear interactions. It can be divided in a central Gaussian shaped distribution, the core, and a low dose region surrounding the core, the nuclear halo. The halo is largely attributable to charged secondary particles from nuclear interactions and has a radius of approximately one third of the beam range [70]. MCS is mainly associated with the Gaussian core but also contributes to the non-Gaussian tail of the distribution. Even though the nuclear halo exhibits a low dose which is spread over a large volume, it must be modelled with sufficient accuracy since it removes a significant fraction of the integrated dose of the core.

MCS is handled by the Fermi–Eyges theory [71]. The theory describes the phase space distribution of a Gaussian beam when it propagates through a stack of slabs of different materials. The Gaussian distribution remains Gaussian at any slab of the stack. Drawbacks of the theory

are that it neglects the non-Gaussian tail created by MCS and that it does not consider lateral inhomogeneities.

Nuclear interactions are modelled in RayStation using look-up formulas from Soukup et al. [59, 72]. A correction term is included to account for differences in the output for certain field sizes and energies which were detected in a field size study. Details on the calculation methods can be found in the RayStation reference manual [59].

In PBS, the total dose in point (x, y, z) is given by the sum of all spot contributions:

$$dose(x, y, z) = \sum_e \sum_j S_e \omega_{ej}^{MU} d_{ej}(x, y, z), \quad (2.20)$$

where d_{ej} is the dose contribution from spot j in energy layer e , S_e (protons/MU) the energy dependent dose monitor sensitivity and ω_{ej}^{MU} the meterset weight of the spot. d_{ej} is determined by the MCS dose contribution of each sub-spot k and the nuclear interaction (NI) dose contribution of spot j :

$$d_{ej}(x, y, z) = \sum_{k=1}^{19} (\omega_k \Phi_{ejk}^{MCS} IDD_{ejk}) + \Phi_{ej}^{NI} IDD_{ej}. \quad (2.21)$$

with ω_k being the relative weight of sub-spot k . The nuclear interaction dose can be approximated by a single pencil beam trace since its contribution is small compared to MCS.

In DS, the total dose in point (x, y, z) is calculated based on the dose contribution of each fluence pixel j , the meterset weight of the beam MU , the collimator transparency a_j and the number of protons per MU I_e :

$$dose(x, y, z) = MU \sum_j a_j \sum_e I_e (\Phi_{ej}^{MCS} + \Phi_{ej}^{NI}) IDD_{ej}. \quad (2.22)$$

I_e is the product of the relative weight of the energy layer ω_e , the dose monitor sensitivity S_N (protons/(MU cm²)) and the pixel area A : $I_e = \omega_e S_N A$.

Monte Carlo dose engine

This section gives an overview of the RayStation MC dose engine without going into mathematical details. Formulas can be found in the RayStation reference manual [59] and the references therein.

The current MC dose engine supports only PBS beams and was developed with the goal to achieve a good trade-off between computation time and sufficient accuracy of the applied physical models. Some of the concepts are identical to those used in the Pencil Beam dose engine. For example, both dose engines share the same initial phase space parameters for a spot. The protons are transported through the patient geometry through a sequence of so-called random hinges. Each step of the sequence is represented by two straight legs and a connecting hinge point. The angle between the legs simulates the deflection of the proton due to MCS and is randomly selected from a probability distribution defined by the Goudsmit-Saunderson theory [59, 73, 74]. One hinge can extend over several voxels of the transport grid. MCS and energy loss straggling are calculated once per random hinge to speed up the computation time. They are taken into account until the energy falls below an energy cut-off value of 30.7 MeV. Above this value, energy loss straggling is handled by the Bohr approximation which assumes that the spread can be described by a Gaussian distribution [59]. In contrast to MCS and energy straggling, energy loss and nuclear absorption have to be considered for each voxel which is traversed due to their impact on the kinetic energy of the proton and the elemental composition. As in the Pencil Beam dose engine, energy loss is calculated based on the Bethe-Bloch formula neglecting the shell and density correction terms.

Nuclear interactions which are considered by the MC engine are non-elastic nuclear reactions,

elastic proton-proton scattering and elastic proton-nucleus scattering. Non-elastic nuclear reactions are difficult to model which is why RayStation uses a cross section data library based on the ICRU report 63 [75]. A continuous slowing down approximation is applied for heavy secondary particles such as deuterons and alphas, i.e. MCS, nuclear absorption and energy straggling are disregarded. Elastic proton-proton scattering is modelled by replacing the primary proton by two secondary protons in the transport process. The differential cross sections are calculated from parametrised models described in Ref. [59, 76]. Elastic proton-nucleus scattering events at large angles are rare and the angular distribution is approximated by ICRU63 data on the elastic scattering of protons on hydrogen [75, 76]. Neutrons and gammas are not included in the transport process but their energy is considered in the energy balance. Secondary electrons are also disregarded by the transport mechanics since they deposit their energy very close to the track within a volume which is usually smaller than the voxel size.

2.4.6 Treatment planning in pencil beam scanning and 4D robust optimisation

In PBS, objective functions and constraints are used to penalise deviations from specified dose limits to reach the clinical goals. The chosen dose limits often coincide with the values stated in the clinical goals. But in some cases it might be necessary to choose higher or lower limits to avoid conflicts with competing optimisation functions with regard to the overall treatment quality. There are different optimisation functions which might lead to a certain goal. For instance, a homogeneous dose distribution in the target volume can be realised by penalising deviations from a desired dose level (Uniform Dose function) or by penalising doses under and above specified threshold values (a combination of Min and Max Dose functions). If a Uniform Dose function is applied, the optimizer will try to ensure that exactly the stated dose is delivered to each voxel of the ROI. Using Min and Max Dose functions provides more scope for the optimizer, especially in view of the fact that it is more important to eliminate cold spots than hot spots within the target. One example for a mathematical formulation of an objective function i is:

$$f_i(d) = \int_V \rho \left(d(p) - \hat{d} \right)^2 dp. \quad (2.23)$$

$d(p)$ denotes the dose at point p within the volume V of the considered ROI, \hat{d} indicates the reference dose level which shall not be exceeded or fallen below and ρ is a ramp function. Depending on whether a Min or Max function has to be applied, ρ is defined as $\rho(y) = \min\{y, 0\}$ or $\rho(y) = \max\{y, 0\}$. Further functions in RayStation relate to the dose volume histogram, the equivalent uniform dose or the dose fall-off (see Ref. [77, 78] for details). The composite function f of all objectives f_1, \dots, f_n and, where appropriate, constraint functions form together the optimisation problem. Since the functions f_1, \dots, f_n penalise deviations from desired dose criteria, a minimum of f has to be found. However, it is most likely that a common minimiser for all functions does not exist due to conflicting criteria involving high target doses and low doses to normal tissues. A trade-off between the desired qualities is made based on the relative weighting of the functions:

$$f = \sum_{i=1}^n \omega_i f_i, \quad (2.24)$$

where $\omega_1, \dots, \omega_n$ are non-negative weights.

The constraints are divided into physical constraints x and planning constraints c_1, \dots, c_m . The latter are optimisation functions chosen by the planner which represent important planning criteria that have to be fulfilled without making compromises. Physical constraints x arise from hardware limitations or laws of nature and are part of a given set of feasible parameters χ . Altogether, the non-linear optimisation problem (NLP) can be written as follows:

$$\begin{aligned} & \underset{\text{over } x}{\text{minimise}} && f(d(x)) \\ & \text{subject to} && c_j(d(x)) \leq 0, \quad j = 1, \dots, m \\ & && x \in \chi, \end{aligned} \quad (2.25)$$

where x stands for the optimisation variables and $d(x)$ for the corresponding dose distributions [79, 80]. In PBS, the optimisation variables are the spot weights.

A sequential quadratic programming (SQP) algorithm is used to solve the inverse planning problem in RayStation. SQP is an iterative procedure in which an approximate solution of a NLP is found by solving a sequence of quadratic programming (QP) subproblems. The objective function of each QP subproblem is given by the local quadratic Taylor series approximation of the Lagrangian function $\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j^T c_j(x)$, where λ is the Lagrangian multiplier vector associated with the constraints c_1, \dots, c_m :

$$\mathcal{L}(x, \lambda) \approx \mathcal{L}(x_k, \lambda_k) + \nabla \mathcal{L}(x_k, \lambda_k)^T (x - x_k) + \frac{1}{2} (x - x_k)^T \nabla_{xx}^2 \mathcal{L}(x_k, \lambda_k) (x - x_k). \quad (2.26)$$

With the linearisation of the constraints in the original problem, the QP subproblem in iteration k takes the form:

$$\begin{aligned} & \underset{\text{over } x}{\text{minimise}} && f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2} (x - x_k)^T \nabla_{xx}^2 \mathcal{L}(x_k, \lambda_k) (x - x_k) \\ & \text{subject to} && c_j(x_k) + \nabla c_j(x_k)^T (x - x_k) \leq 0, \quad j = 1, \dots, m \\ & && x \in \chi, \end{aligned} \quad (2.27)$$

where x_k is the initial estimate and the function and gradient values of the Lagrangian equal those of the objective function at $x = x_k$: $\mathcal{L}(x_k, \lambda_k) = f(x_k)$, $\nabla \mathcal{L}(x_k, \lambda_k) = \nabla f(x_k)$ [81]. The search direction $(x - x_k)$ for the next iteration is determined by the solution of the QP subproblem. Based on this, new variables (x_{k+1}, λ_{k+1}) are constructed such that the sequence converges to a local minimum of the NLP for $k \rightarrow \infty$. The optimisation of a subproblem is itself an iterative procedure including numerical approximations. A review on different SQP methods with detailed mathematical descriptions can be found in Ref. [82].

Robust optimisation

PBS enables a conformal dose delivery to the target in three dimensions, but at the same time it is susceptible to setup uncertainties, range uncertainties and especially organ motion due to its steep dose gradients [83, 84]. Setup uncertainties include positioning errors of the patient and mechanical inaccuracies of the equipment [65]. Range uncertainties are attributable to errors in the CT calibration and the conversion of Hounsfield units to stopping power [85, 86]. In conventional treatment planning, these uncertainties are taken into account by adding margins to the target volume as for example in the PTV concept introduced in Section 2.4.3. This concept is not necessarily suited for proton therapy, where the assumption of a rigid displacement of the dose cloud is not valid in the case of beam misalignment or interplay effects [87]. A possibility to increase the robustness of a treatment plan is to incorporate the uncertainties already in the optimisation process. This is done in RayStation by the minimax optimisation. The robust optimisation minimises the objective functions considering the worst case scenario for user defined uncertainties which are discretised into a set of scenarios S [88, 59]:

$$\begin{aligned} & \underset{\text{over } x}{\text{minimise}} && \max_{s \in S} \{f(d(x, s))\} \\ & \text{subject to} && x \in \chi. \end{aligned} \quad (2.28)$$

Each scenario s represents a certain motion phase, setup error and density error. Setup errors, in form of a misalignment of the patient relative to the gantry, are simulated by shifts of the isocentre. The resulting displacement of spot positions depends on the traversed material. Consequently, the dose deformation is non-rigid and gets large for steep density gradients. Density errors are modelled by a uniform scaling of the mass density which leads to a material dependent stretching or contraction of the depth dose curve. By computing the dose for every setup and density scenario on selected patient image data sets, e.g. on different motion phases of a 4D CT, uncertainties due to organ motion are taken into account. The number of scenarios which is

considered by the robust optimisation is given by the product of the selected patient image scenarios, setup error scenarios and density error scenarios. The nominal scenario assuming no errors is included in each category [59]. By determining the worst case based on scenarios and not independently for each voxel, only physically realisable scenarios are considered. If more than one function is selected to be robust, the worst case scenario is the scenario in which the weighted sum of these functions becomes maximal. It is possible to combine nominal and robust objectives and constraints in the formulation of the robust optimisation problem in Equation 2.28.

2.4.7 Treatment planning in double scattering

The 'Treat and Protect' concept forms the basis for treatment planning in DS. The planner has to specify for each beam which target ROIs shall be treated (hereinafter referred to as 'treat ROIs') and which OAR ROIs have to be protected by the shape of the aperture or compensator (hereinafter referred to as 'aperture protect ROIs' and 'compensator protect ROIs'). To this end, lateral, distal and proximal margins, the density uncertainty and the air gap must be defined. Optional settings include the smearing radius (see Sec. 2.3.2) and the maximum allowed gradient for the compensator. Based on these specifications, the TPS computes automatically the shape of the aperture, the compensator matrix, the range and modulation of the SOBP and the snout position.

The contour of the aperture is given by the projected area of all treat ROIs from beam's eye view plus the specified lateral margins. If the beam's eye view projection of any aperture protect ROI plus margins intersects this area, it is subtracted from the first contour to protect the OAR. In a last step, the contour might have to be expanded to make the aperture millable.

The compensator matrix is characterised by the milling tool diameter and the thickness of each pixel. It is calculated for an area which covers the aperture contour plus a margin of about 1 cm. The physical thickness of a pixel is either determined based on the maximum radiological depth of the treat ROIs at that point considering the distal margin and the density uncertainty or by the minimum radiological depth of the compensator protect ROIs at that point minus the proximal margin and the density uncertainty factor, dependent on which radiological depth is shorter. Furthermore, the thickness depends on the maximum allowed gradient and the smearing radius between neighbouring pixels (see Sec. 2.3.2). For both aperture and compensator design, the protection of OARs has a higher priority than the treatment of targets.

The range of the SOBP is defined by the maximum radiological depth of the target considering the distal margin and the density uncertainty. The modulation of the SOBP equals the SOBP range minus the minimum radiological depth of the target decreased by the proximal margin and the density uncertainty factor. Range and modulation are determined by the TPS such that the the whole target including margins and density uncertainty is covered by at least 95% of the prescribed dose, provided that there is no contradiction with the protection of OARs.

The snout is set to a position where the distance between the most downstream beam shaping device and the patient or patient fixation matches the specified air gap [59].

2.5 Treatment of mobile targets

Proton therapy aims at creating steep dose gradients and minimising the irradiated volume to escalate the dose to the tumour without harming nearby OARs. Target motion, however, requires an increase of the planning target volume/ larger uncertainties in the robustness settings to cover the tumour with sufficient dose at any point of the motion trajectory. As a consequence, the favourable conformity of proton plans is limited and the dose to the tumour might be restricted by dose constraints for nearby OARs. In some cases, this may result in a deterioration of the local control rate.

This section will explain the difference between intra- and inter-fraction motion, the impact of organ motion on the delivered dose distribution, interplay effects in PBS and techniques to mitigate tumour motion.

2.5.1 Intra- and inter-fraction motion

Motion can occur on different time scales. A distinction is made between intra- and inter-fraction motion. Intra-fraction motion takes place during the irradiation on a time scale of seconds. It is usually related to respiration motion or the heart beat. In the abdomen, respiratory-induced organ motion is mainly in superior-inferior (SI) direction. The displacements in anterior-posterior (AP) and left-right (LR) direction usually do not exceed a few millimetres [89]. The human respiratory cycle length is typically three to five seconds and thus much shorter than the treatment time which takes several minutes depending on the tumour volume, the prescribed dose and the number of irradiation fields.

The term inter-fraction motion is used for changes which occur between different treatment fractions. The time scale range from minutes to days. Sources of inter-fraction motion are for instance tumour shrinkage, loss in weight, changes in the baseline of the motion pattern, a different filling of the bladder or patient positioning errors [89, 90, 91, 92]. A treatment course extends over several weeks depending on the number of fractions.

Patient motion can cause both intra- and inter-fractional changes. In radiotherapy, it is common to suppress it by patient immobilisation masks.

This work focuses on intra-fraction motion of liver tumours. The displacement of a liver tumour due to regular breathing can be larger than 2 cm [93]. Hu et al. reported mean motion amplitudes of 3 mm (LR), 10 mm (SI) and 3 mm (AP) for 4D CT data sets of 46 liver cancer patients treated without abdominal compression and mean amplitudes of 3 mm (LR), 5 mm (SI) and 2 mm (AP) for 53 liver cancer patients treated with abdominal compression [94]. Mean motion amplitudes of 2 mm (LR), 8 mm (SI) and 4 mm (AP) were observed by Case et al. for 29 liver cancer patients which were partly treated with abdominal compression [95]. Kitamura et al. measured mean motion amplitudes of 4 mm (LR), 9 mm (SI) and 5 mm (AP) for 20 liver cancer patients during free-breathing (without abdominal compression) [96].

2.5.2 Impact of organ motion

Mobile targets such as lung or liver tumours, which continuously change their position during the respiratory cycle, pose a challenge for proton therapy. Since tissue is moving in and out of the irradiation field, steep dose gradients at the edge of the target volume are blurred. In addition, motion-induced range changes might lead to a deformation of the dose distribution. This is for example the case if ribs are located within the beam path and change their position with respect to the planning CT. Even small changes in the density of the traversed tissue have a great impact on the dose deposition due to the shift of the Bragg peak. By comparison, photon plans exhibit similar dose blurring as proton plans but are only marginally affected by density changes (except for the short buildup region) due to the exponential decay of the depth dose curve (see Fig. 2.1). Both, proton and photon beams, are sensitive to inter-fraction motion such as patient positioning errors or baseline drifts which affect the relative position of the tumour to the beam.

Active proton beam delivery techniques such as PBS are additionally affected by interplay effects between organ motion and beam scanning. The effect is explained in detail in the next section (2.5.3). In comparison to PBS, DS plans are more robust against motion (Sec. 2.3.2). The beam energy changes quasi-continuously because of the fast rotation speed of the range modulator wheel (600 revolutions per minute) in relation to the respiratory cycle length (about 3 s to 5 s).

2.5.3 Interplay effects

In radiotherapy, the term 'interplay effect' refers to the deformation of a planned dose distribution due to the temporal interference between organ motion and dynamic beam delivery. Interplay effects impact for example the dose deposition in lung and liver tumours in PBS therapy, where intra-fractional tumour motion occurs on a similar time scale as beam scanning. The simultaneous motion of the tumour and the proton beam results in a displacement of the

individual proton spots with respect to the planned positions, which may cause under- and over-dosage of the target and unintended exposure of OARs [6, 7, 8, 9]. The formation of interplay effects in PBS is schematically illustrated in Figure 2.10. The extent of the dose degradation depends on several factors such as the tumour motion amplitude, the beam delivery sequence, the relative direction of tumour and beam motion, the length of the respiratory cycle and the initial respiratory phase at the beginning of the irradiation. Modelling of interplay effects is a key ingredient of motion management in PBS and is discussed in Chapter 3.

2.5.4 Respiratory correlated computed tomography

Respiratory correlated computed tomography (4D CT) is utilised to assess tumour and organ motion for treatment planning. It requires respiratory monitoring devices to assign CT projection data to the respective motion phases. A common acquisition method, which is also used at WPE, is a helical CT scan with a very low couch speed. The couch speed must be low enough to ensure that each slice is illuminated by the x-rays at least for the duration of a respiratory cycle. It cannot be set arbitrarily low due to a maximum 'x-ray on time' of the CT scanner. Retrospectively, the acquired data is binned into a series of three-dimensional CT data sets, usually eight or ten, representing different motion phases. The binning is either based on the phase or the amplitude of a surrogate motion signal. Commercial available motion monitoring devices are for example air bellows that measures the air pressure in the bellow when the abdomen contracts or stretches or plastic blocks with infrared reflectors which are positioned on the abdomen of the patient and are recorded by a digital camera [97].

2.5.5 Motion mitigation techniques

In the following, existing techniques to mitigate motion effects in proton therapy are briefly summarised. The tumour motion amplitude can be reduced by abdominal compression using pressure plates or belts. Another strategy is to freeze the motion, e.g. by breath hold techniques or anaesthesia. In treatments under free breathing, it is possible to decrease the ITV by beam gating. Gating controls the field application depending on the respiratory cycle and requires motion monitoring systems. The beam is only turned on for a certain fraction of the respiratory cycle (gating window). As a consequence, the treatment time is prolonged (on average by a factor of two to five compared to a treatment without gating) [14]. The gating window determines the amount of residual motion. Usually, the residual motion is comparable to tumour motion under abdominal compression and can still cause under- and overdosage due to interplay effects for dynamic beam delivery techniques [12].

In PBS, a frequently studied motion mitigation technique is rescanning, also referred to as repainting. Rescanning smoothes the interplay pattern by irradiating the target volume multiple times per fraction with a reduced dose until the prescribed fraction dose is reached (averaging effect) [6, 12]. Different rescanning strategies exist like for example layered rescanning, in which all rescans of a layer are completed before the energy is changed, or volumetric rescanning, in which all energy layers are irradiated in each rescan. In contrast to gating, rescanning reduces the amount of over- and underdosage by dose smearing but requires a larger treatment volume to account for the internal motion. Thus, more normal tissue is exposed to high doses and rescanning used on its own is rather suited for moderate than large motion amplitudes. Effects similar to rescanning are achieved when increasing the number of irradiation fields [12]. A high number of treatment fractions and a large spot size in PBS are also associated with a reduction of motion effects [98]. A particularly robust approach to mitigate motion effects in PBS is to combine gating and rescanning and was investigated by Zhang et al. for different scanning dynamics [99]. However, prolonged treatment times of up to 20 minutes limit the applicability. There are also approaches in treatment planning to increase the robustness against tumour motion, such as 4D robust optimisation (Sec. 2.4.6). In this thesis, 4D robust optimisation was evaluated for liver cancer patients. The results are presented in Chapter 4.

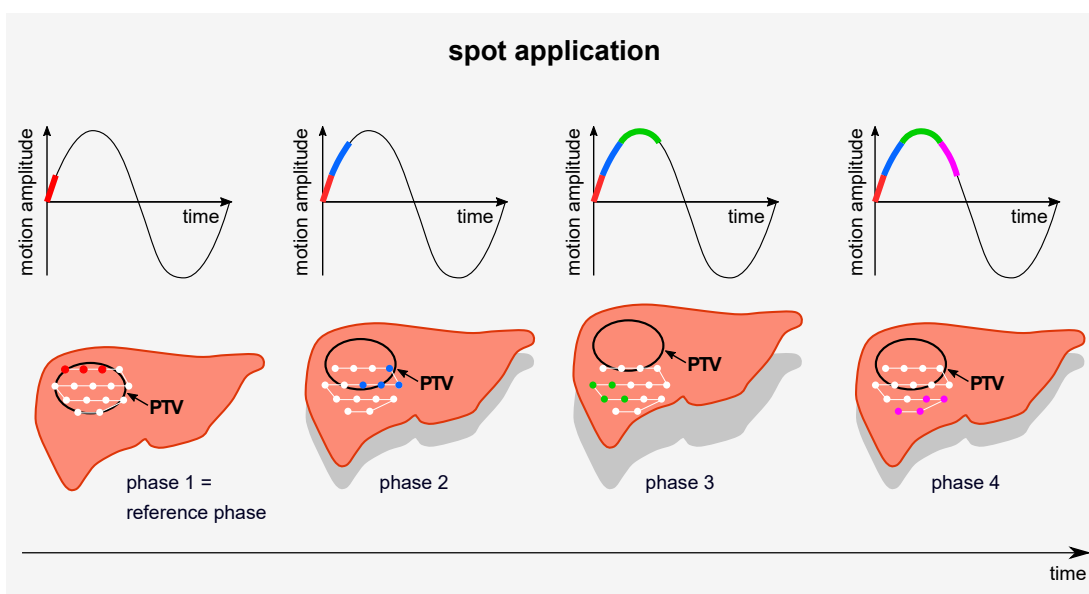
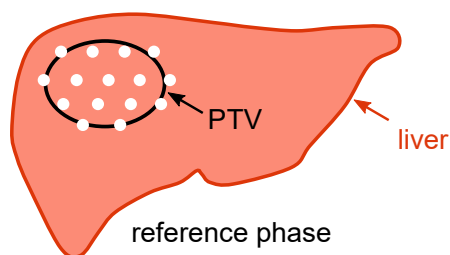
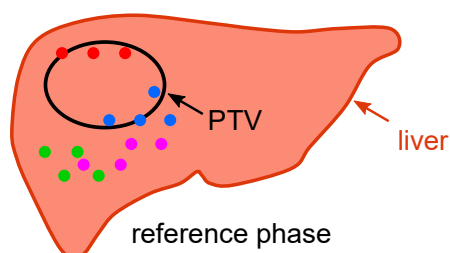
planned spot pattern:**actually delivered:**

Figure 2.10: Schematic representation of motion interplay effects in pencil beam scanning proton therapy: The planned spot positions for the irradiation of the planning target volume (PTV) are marked by white spots. During spot application, the pencil beam is deflected according to the predefined scanning path (white line). At the same time, the liver is moving due to respiratory motion. The reference position of the liver is indicated by a grey contour for each time interval. The spot delivery over time and the respective motion phases are displayed in red (phase 1), blue (phase 2), green (phase 3) and purple (phase 4) (simplified time model for illustration purposes). Since the liver continuously changes its position, some spots are delivered at different positions than planned. The image on the bottom right illustrates the resulting spot pattern on the reference phase. Local underdosage of the PTV occurs in regions where spots are missing, whereas overdosage of the PTV and exposure of the normal liver tissue occur for regions with too high spot densities. The delivered spot pattern in the presence of organ motion depends among others on the respiratory phase at the beginning of the irradiation, the respiratory cycle length, the tumour motion amplitude and the beam delivery sequence. In this example, the irradiation started within the reference phase.

3 Implementation of 4D dynamic dose accumulation in PBS treatment planning

A dedicated 4D dynamic dose computation routine, which enables the simulation of interplay effects, was created in the scope of this thesis and is described in Section 3.1. Interplay effect scripts, which easily encompass several hundred lines of code, require thorough validation due to their complexity. The experimental validation of the developed routine is presented in Section 3.2. End-to-end tests were performed using a dynamic phantom and included 4D imaging, treatment planning, irradiation and quality assurance according to clinical standards. An empirical beam time model was built to enable *in silico* studies. It was applied for the treatment planning study on hepatocellular carcinomas addressed in the next chapter. The creation of the beam time model is subject of Section 3.3. Partial results of the presented work in this chapter have been published in Ref. [100].

3.1 Modelling of interplay effects

A standard procedure for the evaluation of motion effects in radiotherapy is the averaging of the dose over all respiratory phases of a 4D CT set. However, this method neglects any dynamics of the beam delivery. In order to visualise interplay effects, information on the time structure of the field as well as on the temporally resolved organ motion is needed. A routine tailored to 4DDD computations of irradiations with an IBA Proteus[®]Plus proton therapy system has been developed based on a rudimentary script provided by RaySearch (E. Engwall, June 2015) using a similar approach as Zeng et al. [29]. The individual steps of the procedure are illustrated in Figure 3.1. Patient motion is assessed based on a 4D CT with respiratory phases CT1, CT2, ..., CT r . The phases, together with the respiratory cycle length and the initial phase at the beginning of the irradiation, determine the organ motion sequence. The respiratory cycle length is assumed to be constant and defined as the average breathing rate measured during the 4D CT scan. A treatment plan is created on the reference CT. Based on the optimised spot positions, weights and energies, the beam delivery sequence is generated either from irradiation log files (retrospective analysis) or from a dedicated beam time model. The advantage of a model-based computation is that possible fluctuations of the beam delivery parameters can be simulated and considered for planning studies or patient-specific quality assurance. Furthermore, it simplifies the clinical workflow. A description of the IBA implementation of PBS including the scanning control system is given in Section 2.3.2. The information on organ motion and beam delivery is used to determine in which respiratory phase each spot is delivered. The resulting fraction dose d_{fx,CT_i} for respiratory phase i is calculated and subsequently mapped to the reference phase using deformable image registrations (see Sec. 2.4.4 for details). This procedure is repeated for all respiratory phases. In the end, the deformed doses d_{fx,CT_i}^{Ref} and the dose delivered within the reference phase are added up to the fractional 'interplay dose distribution':

$$d_{fx} = d_{fx,CT1}^{Ref} + d_{fx,CT2}^{Ref} + \dots + d_{fx,CTr}^{Ref}, \quad (3.1)$$

where fx denotes the fraction number and r the number of respiratory phases. The 4DDD delivered over the whole treatment course is given by: $D = \sum_{fx=1}^N d_{fx}$.

Implementation in RayStation

Interplay effect evaluations were conducted using the scripting interface of RayStation (see Sec. 2.4.1) [60]. A graphical user interface has been created to facilitate the input of patient- and machine-specific parameters as well as different evaluation options (Fig. 3.2). Evaluation options include for example the evaluation type (single or multiple fractions) and the number of evaluation scenarios in prospective analyses (with varying energy layer switching times and starting phases).

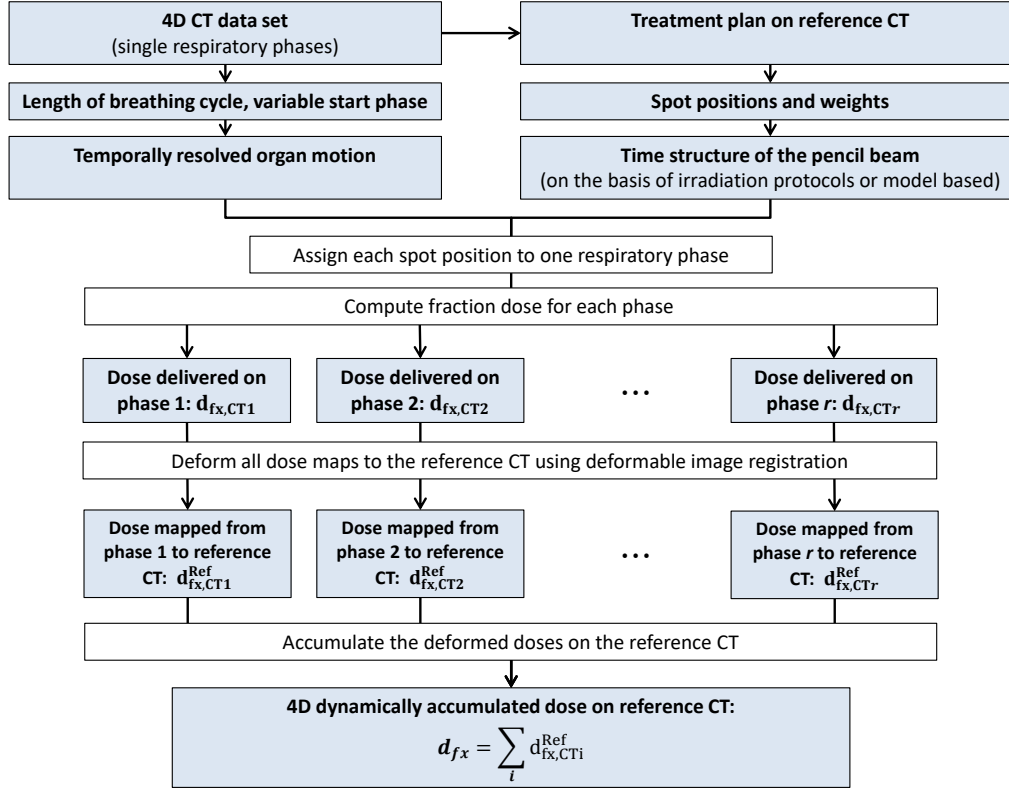


Figure 3.1: Flow chart illustrating the workflow of interplay effect simulations automated by a Python script in RayStation: The 4D dynamically accumulated fraction dose d_{fx} is calculated based on a time-resolved CT (4D CT), a treatment plan and the corresponding beam delivery sequence to assess interplay effects in pencil beam scanning proton therapy. White boxes indicate active processes and light blue boxes outputs/ inputs. In prospective analyses, the computation is repeated for different pencil beam time structures to consider possible fluctuations of the beam delivery parameters. Adopted from Ref. [100].

3.2 Experimental validation of the 4D dose computation

The accuracy of the interplay effect routine was examined by comparing calculated 4DDD distributions with two-dimensional, dynamic dose measurements. In a first experiment, the dose was measured by a static detector downstream of a dynamic buildup (setup A). Setup A served as a simplified technical model for dose validations in quasi-static anatomical regions behind moving structures, such as ribs. A second series of experiments was performed with a dynamic detector (setup B) simulating the dose validation within a moving tumour. Details on the experimental setups as well as on 4D CT acquisition, treatment planning and evaluation methods are given in the following section. Subsequently, the results are presented and discussed.

3.2.1 Materials and methods

MatriXX PT detector

The MatriXX PT detector (IBA Dosimetry, Schwarzenbruck, Germany), a two-dimensional ionisation chamber array, has been designed for machine quality assurance and pretreatment plan verifications in proton therapy. It consists of 1020 vented, plane parallel cylindrical chambers, each with a diameter of 4.2 mm and a centre-to-centre spacing of 7.6 mm [101]. The chambers are arranged in a 32 x 32 grid yielding an active measurement area of 24.4 cm x 24.4 cm (Fig. 3.3a). A 6 mm thick acrylonitrile butadiene styrene buildup covers the array [102]. Measurements are

Interplay effect evaluation

Beam delivery settings

Settling time of the fast scanning magnet (ms):

Settling time of the slow scanning magnet (ms):

Energy layer switching time

Average value (s):

Standard deviation (s):

Respiratory motion

Respiratory cycle length (s):

Evaluation settings

Evaluation type:

Single fraction Multiple fractions

Number of fractions:

Number of scenarios per fraction:

Supplementary motion mitigation technique:

None Gating

Plot DVH:

No Yes

Select respiratory start phase(s):

CT1 0% CT2 10%

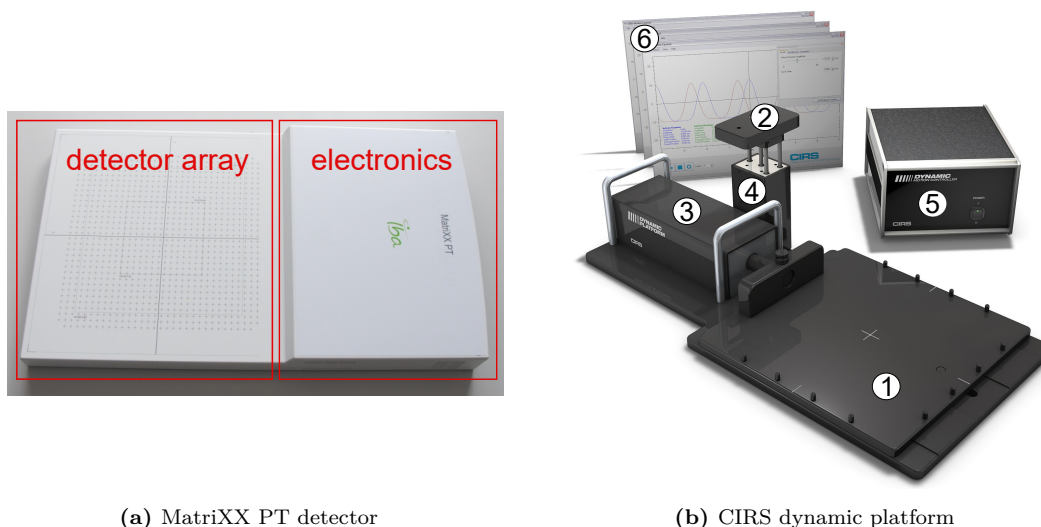
CT3 20% CT4 30%

CT5 40% CT6 50%

CT7 60% CT8 70%

CT9 80% CT10 90%

Figure 3.2: Graphical user interface of the interplay effect routine for prospective evaluations: The interface allows for quick adoptions of the beam delivery parameters, e.g. in the case of a software update of the scanning control system or to model different proton therapy facilities. To determine the time structure of the organ motion, the average respiratory cycle length of the patient must be entered. The evaluation settings include among others the evaluation type, the number of evaluation scenarios and the start phase(s) at which the irradiation commences. One of the selected respiratory start phases is randomly chosen by the routine for each scenario. For retrospective evaluations, another version of the interplay effect routine enables the input of the beam delivery sequence gained from log files. Only the respiratory start phase which was recorded at the time of the treatment is selected.



(a) MatriXX PT detector

(b) CIRS dynamic platform

Figure 3.3: (a) Top view of the MatriXX PT detector, a two-dimensional ionisation chamber array [101], (b) components of the CIRS dynamic platform [103]: 1. base platform, 2. surrogate platform, 3. phantom motion actuator box, 4. surrogate motion actuator box, 5. motion control unit, 6. CIRS motion control software.

controlled by the designated software OmniPro-I^mRT (IBA Dosimetry, Schwarzenbruck, Germany). Within the scope of this thesis, two measurement modes were used: single shot mode and movie mode. In single shot mode, one measurement is accumulated over a pre-selected integration time. In movie mode, a continuous set of single measurements is acquired during a pre-defined time. The acquisition time for a single measurement is specified by the user. The sum over all measurements of the set is created and saved automatically by the software.

CIRS dynamic platform

The CIRS dynamic platform (CIRS, Norfolk, USA) is designed for the study of tumour motion in radiotherapy and can be used in end-to-end tests including four-dimensional imaging, treatment planning and dose delivery. It is composed of a base platform, which supports almost any phantom up to 32 kg, a small surrogate platform, which imitates the motion of the chest wall, motion actuators and a multi-axis motion controller (Fig. 3.3b). Respiratory monitoring devices can be attached to the surrogate platform in order to perform 4D CT imaging or respiratory gated irradiations. The CIRS dynamic platform provides precise and repeatable inferior-superior motion of the base platform and posterior-anterior motion of the surrogate platform through bipolar stepper motors. Optical sensors control the accurate mechanical positioning. For both platforms, the motion is limited by a total range of 50 mm. The spatial accuracy is 0.05 mm and the accuracy of the motion period approximately 5 ms [103]. The CIRS dynamic platform is operated through the CIRS motion control software. Motion patterns can be selected from a menu or imported from tab delimited or comma separated files. The menu offers five different wave functions with variable amplitude, motion period and phase shift.

Setup

The two setups used for the validation of the interplay effect routine are depicted in Figure 3.4a and 3.4c. For setup A, the MatriXX PT detector was placed on the patient couch. The dose was measured in a horizontal plane perpendicular to the beam (gantry angle: 0°). 2.1 cm of RW3 plates (SP34 solid plate phantom, IBA Dosimetry, Schwarzenbruck, Germany) served as buildup material in front of the detector. The phantom consisted of a PMMA cylinder with a

Table 3.1: 4D CT scanning parameters used for the experimental validation of the 4D dynamic dose computation (Sec. 3.2)

| | |
|-----------------------|------------|
| slice thickness (mm) | 2 |
| increment (mm) | 1 |
| field of view (mm) | 600 |
| voltage (kV) | 120 |
| mAs/slice | 360 |
| reconstruction filter | standard B |
| rotation time (s) | 0.5 |
| pitch | 0.042 |

diameter of 16 cm, a height of 8.5 cm and five holes for tissue equivalent Gammex rods (GAMMEX INC., Middleton, USA). Surrogates for liver, cortical bone, adipose tissue, inner bone and lung were chosen as inserts. The first three of them were located within the irradiation field to simulate liver tumour irradiations (see Fig. 3.4b). The dimensions of the technical phantom, e.g. the thickness of the cortical bone surrogate, do not correspond to human anatomy and represent an extreme case. Phantom motion was realised by placing the cylinder onto the dynamic CIRS platform. The platform with the phantom had to be set up on a PMMA table in order to position it upstream of the detector in the beam path. During the 4D CT acquisition and irradiation, the phantom moved back- and forwards along the y axis (fixed reference system according to IEC 61217 [104]), which equals the main direction of breathing induced motion for a patient at a couch angle of 0° . The motion curve was set to a sinus function with an amplitude of 5 mm. A comparatively long motion period of 10s was chosen for the first test series due to the complexity of the dynamic dose assessment. After successful trials, the time period was reduced to 5s in a second test series, which is a realistic value for the human respiratory cycle length.

According to the manufacturer, the CIRS dynamic platform moves any phantom with sub-millimetre accuracy (± 0.1 mm) and reproducibility [103]. Measurements with a laser tracker (model Absolute Tracker 901B, LeicaGeosystems AG, Unterentfelden, Switzerland) could confirm the high accuracy in positioning with an average value of 0.02 mm and a standard deviation (SD) of 0.02 mm for the positional deviation.

A second series of experiments was performed with a dynamic detector (setup B). In this setup, the MatriXX PT was positioned on the moving CIRS platform (see Fig. 3.4c). There was no need for supplementary RW3 plates or a table since the cylinder could be placed directly onto the MatriXX PT. The same settings were used as for setup A.

4D CT acquisition

4D CTs were acquired with a Philips Brilliance CT BigBore (Philips GmbH, Hamburg, Germany). The settings for the CT scan are listed in Table 3.1. Retrospective spiral correlated imaging was chosen as scanning mode. The simulated breathing pattern was monitored by a pneumatic belt ('Philips bellows') so that the 4D projection data could be retrospectively sorted. Ten 3D CT sets were reconstructed from one respiratory cycle equally spaced at 10% time intervals (phase binning), where 0% corresponds to end-inspiration and 50% to end-expiration.

Treatment planning

Two separate treatment plans corresponding to setup A and setup B were created in a research version of RayStation 5 using the Pencil Beam dose algorithm and a 2 mm dose grid. For both plans, the PTV was defined by a rectangular shaped ROI with the following dimensions: 4 cm (LR), 9 cm (SI) and 1 cm (AP), where the SI direction corresponds to the direction of motion.

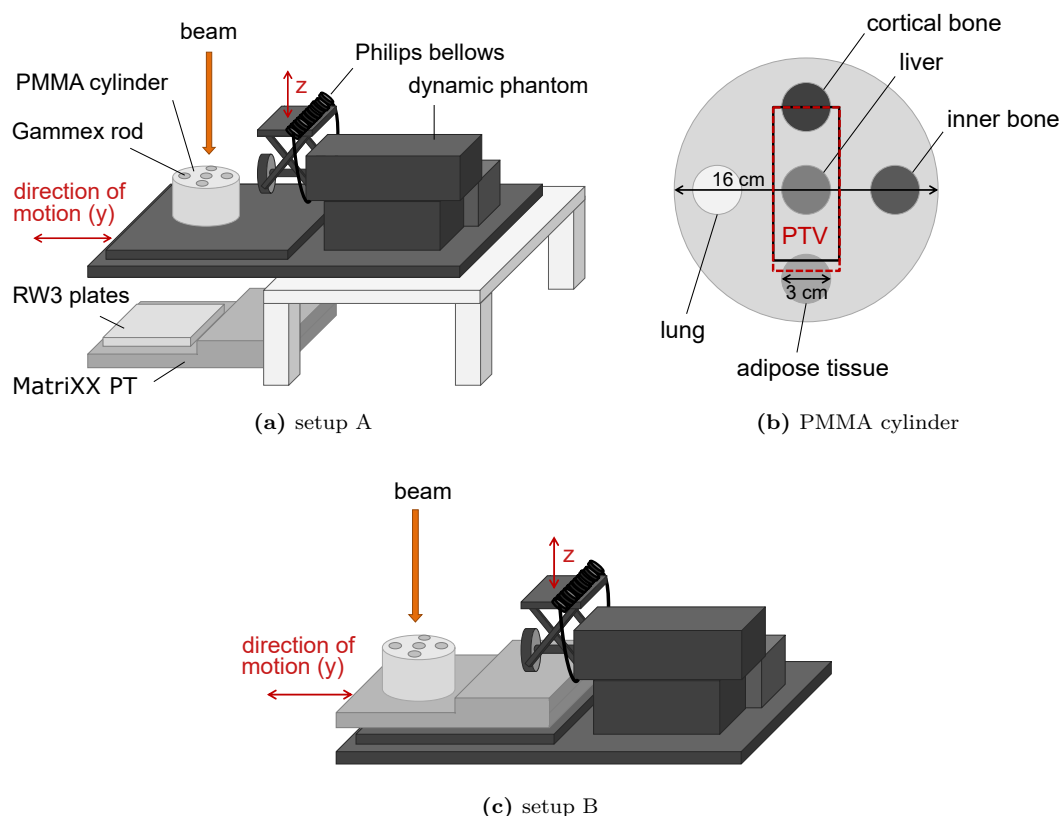


Figure 3.4: (a) Sketch of setup A: The dose is measured by a static, two-dimensional ionisation chamber array downstream of the dynamic phantom. The pneumatic belt (‘Philips bellows’) used for the 4D CT acquisition was removed prior to the irradiation, (b) sketch of the PMMA cylinder: The top view shows the arrangement of the tissue equivalent Gammex rods. The relative position of the planning target volume (PTV) to the cylinder geometry is indicated by a dashed red line for setup A (considering all motion phases) and by a solid black line for setup B (constant relative position), (c) sketch of setup B: The two-dimensional ionisation chamber array moves perpendicular to the incoming beam.

The small modulation width (AP direction) was chosen to be sensitive to changes in the proton range. In view of the HCC treatment planning study (Chap. 4), the field size of the PTV was adjusted to the geometrical arrangement of the Gammex rods, so that the surrogates for liver, cortical bone and adipose tissue were located within the irradiation field (see Fig. 3.5). The effective plane of measurement of the MatriXX PT, as defined in Ref. [102], was at the same height as the centre of the PTV. A total dose of $63 \text{ Gy}_{\text{RBE}}$ was prescribed to the PTV in 15 fractions referring to the fractionation scheme proposed by Bush et al. for HCC patients [2]. Neither robust planning nor repainting was applied in order to render interplay effects visible to the full extent. Each treatment plan consisted of a single field at 0° . The isocentre corresponded to the centre of the effective plane of measurement. None of the plans required the use of a range shifter. The air gap between the snout and the phantom was 20 cm for setup A and 33.8 cm for setup B². In addition, there was another air gap of 8.8 cm in-between the CIRS motion platform and the buildup material in front of the detector in setup A. The resolution of the dose grid was set to 2 mm per voxel in all directions and spots were placed on a regular grid with a spacing of 6 mm. The objective functions in the optimisation included maximum and minimum dose settings for the PTV, a dose fall-off objective for the External ROI (which

²At the time of the experiment, a movable snout minimising the air gap between a range shifter and the patient was not yet installed at WPE. Since no range shifter had to be applied, this did not represent a disadvantage for the study.

comprised the whole phantom and the patient couch) and maximum dose settings for technical planning structures proximal and distal to the PTV in order to improve the dose conformity.

CT data of synthetic tissue surrogates can deviate from clinical CT number to density calibration curves tailored to human tissue data. Therefore, a dedicated CT number to density calibration curve was created based on the known mass densities of the Gammex rods to ensure an accurate estimation of proton stopping power (see Fig. 3.6). Because PMMA has a distinct relation between mass density and CT number, a material override was used for PMMA in planning (density: 1.18 g/cm^3). An example of how the usage of the clinical calibration curve (Body, 120 kV, PBS) can impact the dose comparison is provided in Section 3.2.2 for setup A.

Image registration

Using the dynamic CIRS platform, the phantom was translated but not deformed. Consequently, a rigid registration algorithm would have been the method of choice for 4D image registration. However, for a 4D CT group, RayStation enables only deformable image registrations since the functionality has been developed for planning and evaluation in real patient geometries. Image registrations therefore had to be performed with ANACONDA (Sec. 2.4.4). Another issue was the symmetry of the ionisation chamber array, which made it impossible to rely solely on a CT intensity based registration approach. Multiple positions within the ionisation chamber array yielded the same image similarity measure leading to an irregular deformation vector field. To overcome these problems, a grid of 225 POIs was created and shifted for each phase according to the displacement of the phantom via scripting. The grid covered an area of $15 \text{ cm} \times 15 \text{ cm}$ at the height of the isocentre and was used to guide the DIR. Furthermore, the accuracy of the deformable registrations was investigated for the presented rigid geometry based on five controlling POIs in the measuring plane. The POIs were manually marked in each motion phase and mapped to the reference CT. The error introduced by the DIR in the direction of motion was on average 0.1 mm with a standard deviation of 0.4 mm.

Irradiation

A fraction dose of 4.2 Gy was delivered to the PTV in each experimental run. The delivery was neither synchronised with the phantom motion nor triggered manually based on the displayed motion pattern. In order to determine the start phase at the beginning of the irradiation for the interplay effect evaluation, a camera recorded the screen of the laptop showing the motion curve and a time display (synchronised with a time server). The time structure of the field delivery was gained from the corresponding irradiation log files which also contain time information referred to a time server. Each experiment was performed five times (two times with a motion period of 10 s and 3 times with 5 s). Two stationary irradiations served as reference. Measurements were performed in single shot mode and movie mode. The latter served for time resolved evaluations (described below).

Evaluation methods

The interplay effect routine was validated based on a relative comparison of the measured and the calculated dose distribution since an absolute dose comparison requires thorough calibration processes which are, contrary to measurements in a water phantom, not straightforward for the given setup. Discrepancies in the absolute dose were compensated by a normalisation of the dose distributions.

4DDD distributions were reconstructed in RayStation based on the acquired 4D CT data set (setup A or B), the corresponding irradiation log file and the determined start phase using the interplay effect routine. The dose plane at the depth of the isocentre was exported and compared to the measured dose distribution with the software OmniPro-ImRT. First, the grid resolution of the measured and the calculated dose plane was interpolated to 1 mm by means of cubic spline. Next, the distributions were normalised to a common point in the dose plane.

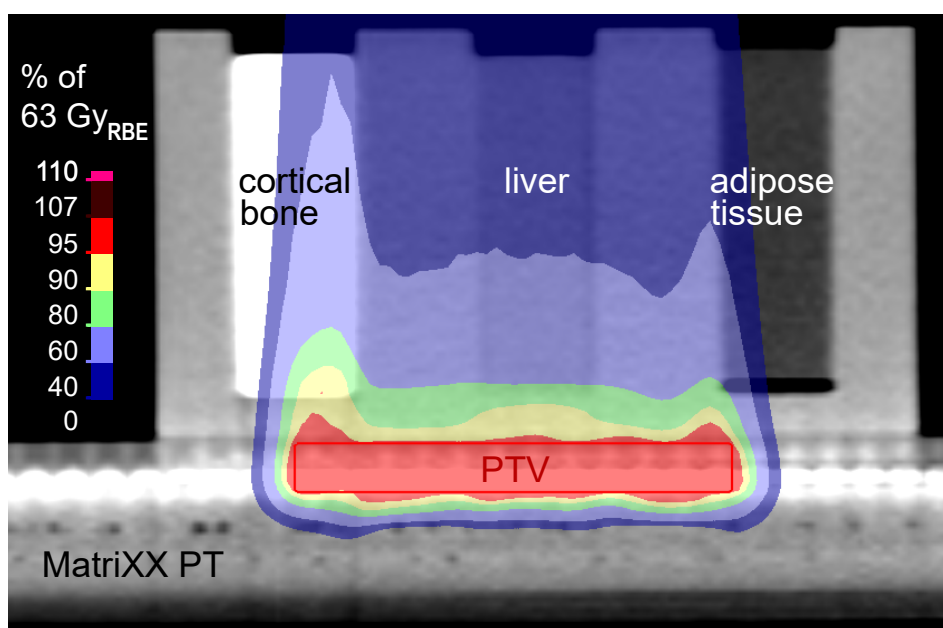


Figure 3.5: Sagittal view of the static dose distribution (overlaid as colour wash) for setup B in RayStation: The centre of the rectangular shaped PTV (red contour) is at the same height as the effective plane of measurement of the MatriXX PT. The PTV is located downstream of the surrogates for cortical bone, liver and adipose tissue. Depending on the motion phase, the pencil beam might pass a different material than in the nominal treatment plan. The direction of motion corresponds to the left-right direction in the image.

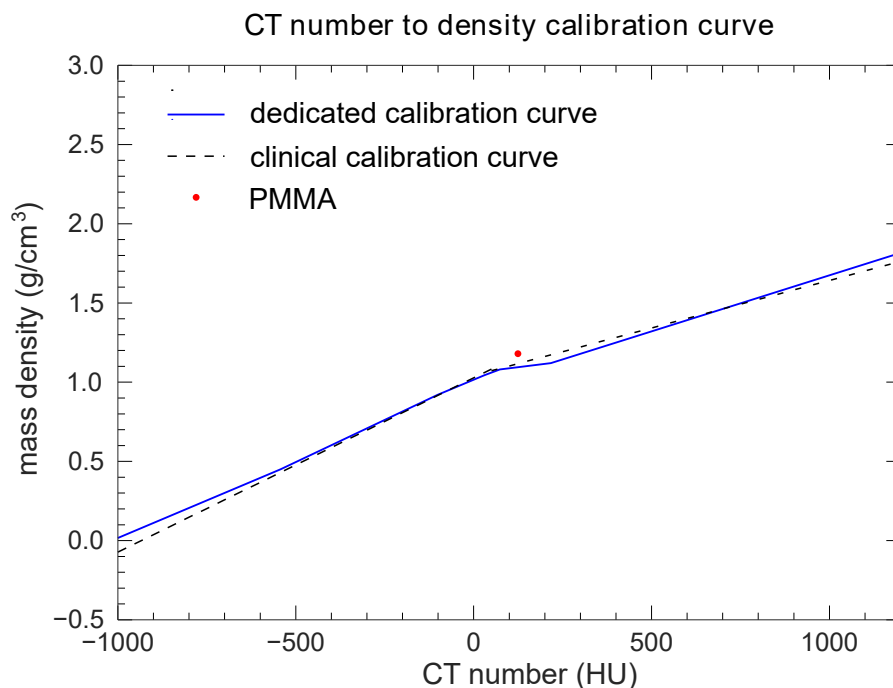


Figure 3.6: CT number to density calibration curve for the experimental validation of the 4D dose computation: To improve the accuracy of proton stopping power calculations in synthetic tissue surrogates a dedicated calibration curve (blue line) was created based on the mass densities taken from data sheets. Since the PMMA data (red point) did not fit to the composite curve, a material override was used for PMMA in planning. The clinical calibration curve (Body, 120 kV, PBS) is plotted for comparison purposes (black dashed line).

Analyses were performed using the gamma evaluation method by Low et al. [105, 106]. The gamma value for a measurement point r_m is given by

$$\gamma(r_m) = \min\{\Gamma(r_m, r_c)\} \forall \{r_c\} \quad (3.2)$$

with

$$\Gamma(r_m, r_c) = \sqrt{\frac{r^2(r_m, r_c)}{\Delta d_m^2} + \frac{\delta^2(r_m, r_c)}{\Delta D_m^2}}, \quad (3.3)$$

where Δd_m denotes the distance-to-agreement criterion and ΔD_m the dose-difference criterion. $r(r_m, r_c)$ and $\delta(r_m, r_c)$ are the differences in distance and dose between measurement and calculation:

$$r(r_m, r_c) = |r_c - r_m|, \quad (3.4)$$

$$\delta(r_m, r_c) = D_c(r_c) - D_m(r_m), \quad (3.5)$$

where D_c is the calculated dose in point r_c and D_m the measured dose in point r_m . The passing criteria Δd_m and ΔD_m were set to 3 mm and 3%, respectively, as used for the patient-specific quality assurance at WPE [107, 108]. A point r_m passes the gamma criterion if $\gamma(r_m) \leq 1$. The threshold dose value for relevant signals was set to 20%.

Even though the measurements could not be calibrated using a water phantom and in-house reference irradiation fields, the measured absolute dose was exemplarily compared with the 4D dose calculation for setup A.

The sensitivity of the routine to its input parameters was tested by comparing the original dose distribution to manipulated computations. To this end, the 4DDD distribution was calculated for different start phases and motion periods. Motion periods from 4.4 s to 5.6 s were investigated in increments of 0.2 s.

In addition to the evaluation of the whole fraction dose, time-resolved computations were compared with measurements in movie mode. For this purpose, the interplay effect routine had to be adapted so that partial dose distributions could be computed for any time interval during the irradiation. The functionality of the time-resolved interplay effect routine was tested by comparing exported dose planes from the original routine and the time-resolved version. First, it was demonstrated that both routines yielded the same dose distribution when choosing the total irradiation time as time interval. Next, the dose delivery was split into five time intervals. The sum of the resulting partial dose distributions equalled the dose distribution obtained by the original routine.

In time-resolved 4DDD experiments, the sampling time of the MatriXX PT was set to 2 s resulting in 14 partial dose distributions. The sampling time for the experimental test was more or less arbitrarily chosen. It should be low enough to visualise the dose delivery with time over different respiratory phases on the one hand and, on the other, it should be large enough to ensure a certain amount of dose per measurement to reduce the impact of inaccuracies induced by the design of the experiment. An offset had to be considered in the computation to account for the time difference between the start of the measurement and the beginning of the irradiation. The offset could not be recorded. It was determined using a try and error approach with the aim to maximise the gamma pass rate of a measurement sample which started and ended within an energy layer and was not separated by irradiation pauses from the previous or subsequent sample of the measurement set. Two time-resolved measurements were performed for setup A.

3.2.2 Results

Comparison with two-dimensional dose measurements

Figure 3.7 shows by the way of example the relative dose comparison of one experimental measurement and the corresponding 4D simulation for setup A. Both the measured dose and the 4D

reconstructed dose reveal a distortion of the planned dose distribution due to motion, e.g. a slight underdosage in the lower half of the PTV (colour wash: light red). The gamma analysis demonstrates a good agreement between the simulated and experimental data with a gamma pass rate of 99.7% (see Fig. 3.7f). The corresponding dose profiles differ only slightly with a maximum deviation of about 3% (see Fig. 3.7d and e).

For both test series (5 s and 10 s motion period), the gamma pass rate remained in-between 97.8% and 99.7% (average 98.7%) meeting the acceptance requirement of 95% used for the patient-specific quality assurance at WPE. The reference gamma pass rates of the two stationary irradiations were 98.6% and 98.3%.

Using the clinical CT number to density calibration curve tailored to human tissue data instead of the specially created calibration, the gamma pass rate decreased from 99.7% to 92.9% for the investigated test case.

For setup B, the dynamic detector, the distortion of the planned dose distribution was particularly pronounced as can be seen in Figure 3.8b. Interplay effects could be reproduced by the routine (Fig. 3.8c). A few locally restricted deviations of up to about 4% were detected (see Fig. 3.8e). Nevertheless, the gamma pass rate was 98.0%. Unlike setup A, setup B exhibited differences in the gamma pass rate depending on the motion period. For 10 s, the pass rate was better than 96.1%, whereas for 5 s it ranged from 88.8% to 90.2%.

The comparison of the absolute dose values revealed an overestimation of the dose of 7.6% by the routine for setup A. The two stationary reference irradiations deviated by 7.8% and 7.5%, respectively, from the planned dose distribution. These discrepancies were equalised by normalisations in the relative comparisons.

Sensitivity towards start phase and motion period

In the following, the results of the sensitivity analysis are presented for manipulated motion periods and start phases. A change of the motion period in the computation by ± 0.2 s resulted in a decrease of the gamma pass rate by about 1–2.5% compared to the original measurement. For a shift of 0.4 s and more, the percentage decrease of the gamma pass rate was already within the double-digit range.

Regarding the start phase, the deterioration of the gamma pass rate was less for a shift towards earlier phases. For instance, a change to the previous motion phase led in average to a decrease of 1.7% and a shift to the subsequent phase to 4.1%. A phase shift of π , representative for the highest deviation, resulted almost in a reduction of the gamma pass rate of 20%.

Time-resolved evaluation

The dose delivery over the course of time is depicted in Figure 3.9 and 3.10 for one measurement and the corresponding time-resolved 4DDD reconstruction. The 4DDD computation reproduces the measured dose distributions well on a qualitative level, in particular the shape of the irradiation field over time. Local differences in the intensity become visible in the second half of the irradiation for image number 8, 9, 10 and 13 (see Fig. 3.10). The gamma pass rate is on average 98.3% for the first and 73.1% for the second half of the irradiation.

The time-resolved dose distribution looked different for the repeated measurement since the irradiation started at another respiratory phase, the energy layer switching times changed and the offset between measurement and irradiation was not the same. The time-resolved interplay routine was able to reproduce these changes. Again, higher gamma pass rates were obtained for the first than for the second half of the irradiation (average 98.7% and 79.3%, respectively).

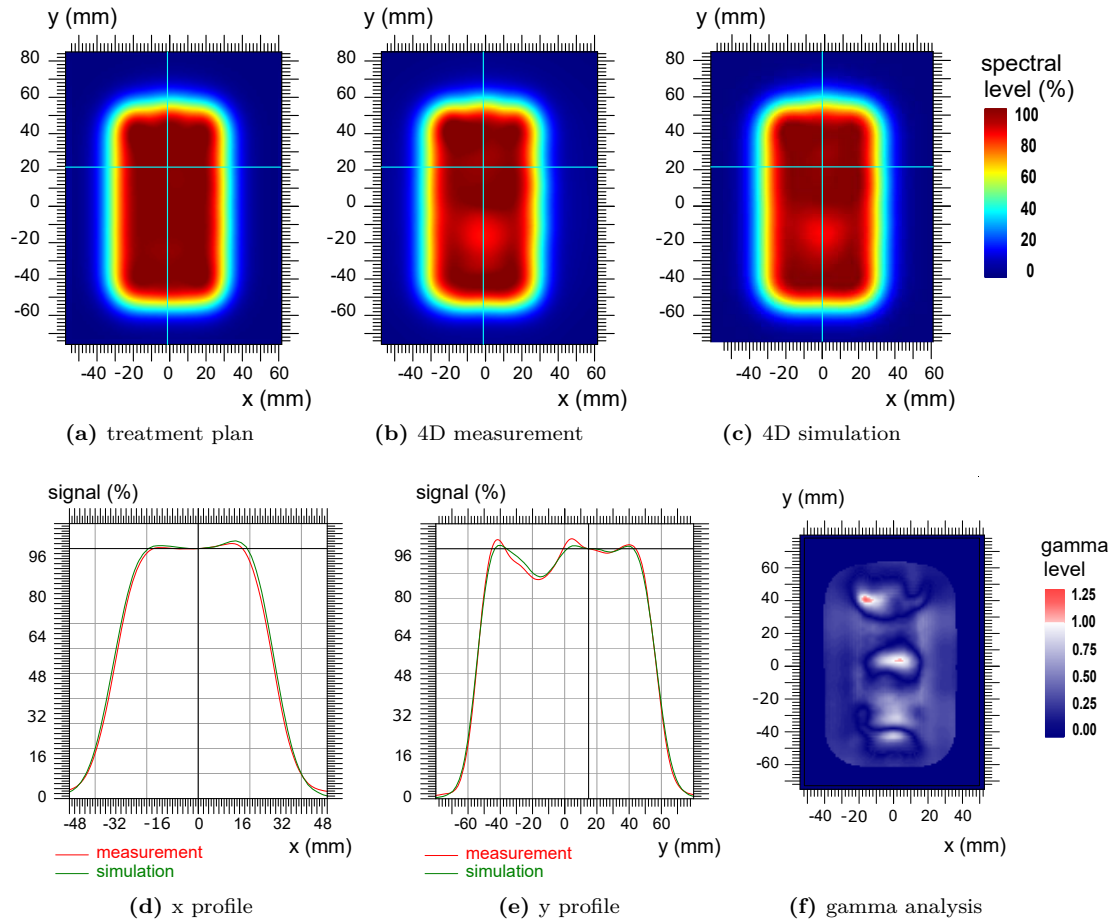


Figure 3.7: Relative dose analysis for the static dose measurement downstream of the dynamic phantom (setup A): (a) static treatment plan, (b) measured 4D dose distribution, (c) 4D reconstructed dose distribution, (d) comparison of the 4D dose profiles in x direction (left-right), (e) comparison of the 4D dose profiles in y direction (superior-inferior), (f) gamma analysis of the 4D dose distribution with a 3 mm/3% passing criterion, gamma pass rate = 99.7%.

3.2.3 Discussion

The relative motion of the different phantom materials to the beam induced changes of the proton range for both, setup A (static detector) and setup B (dynamic detector). The dynamic dose measurement was additionally affected by shifts of the spot positions in relation to the spatial frame of reference of the detector. As a result, larger distortions of the planned dose distribution were expected for setup B. The measurements could confirm this. For setup A and the first test series of setup B, the gamma pass rates gave evidence of the good agreement between measurement and simulation. The fact that the pass rates were above the 95% level and not worse than for the static reference irradiations proved the adequate performance of the 4DDD routine. In some cases, the gamma pass rates were even slightly higher for the 4D experiments than for the two stationary control measurements. A possible reason might be that large density gradients are smoothed in the dynamic case. In addition, inaccuracies due to the dose calculation with the Pencil Beam algorithm and a coarse sampling of dose propagate to the total dose differently than for the stationary case [109].

Reducing the motion period to 5 s (second test series), the gamma pass rate deteriorated from above 95% to about 90% for setup B. This is because the phantom travelled larger distances in a shorter amount of time, which causes motion artefacts in the 4D CT and increases the

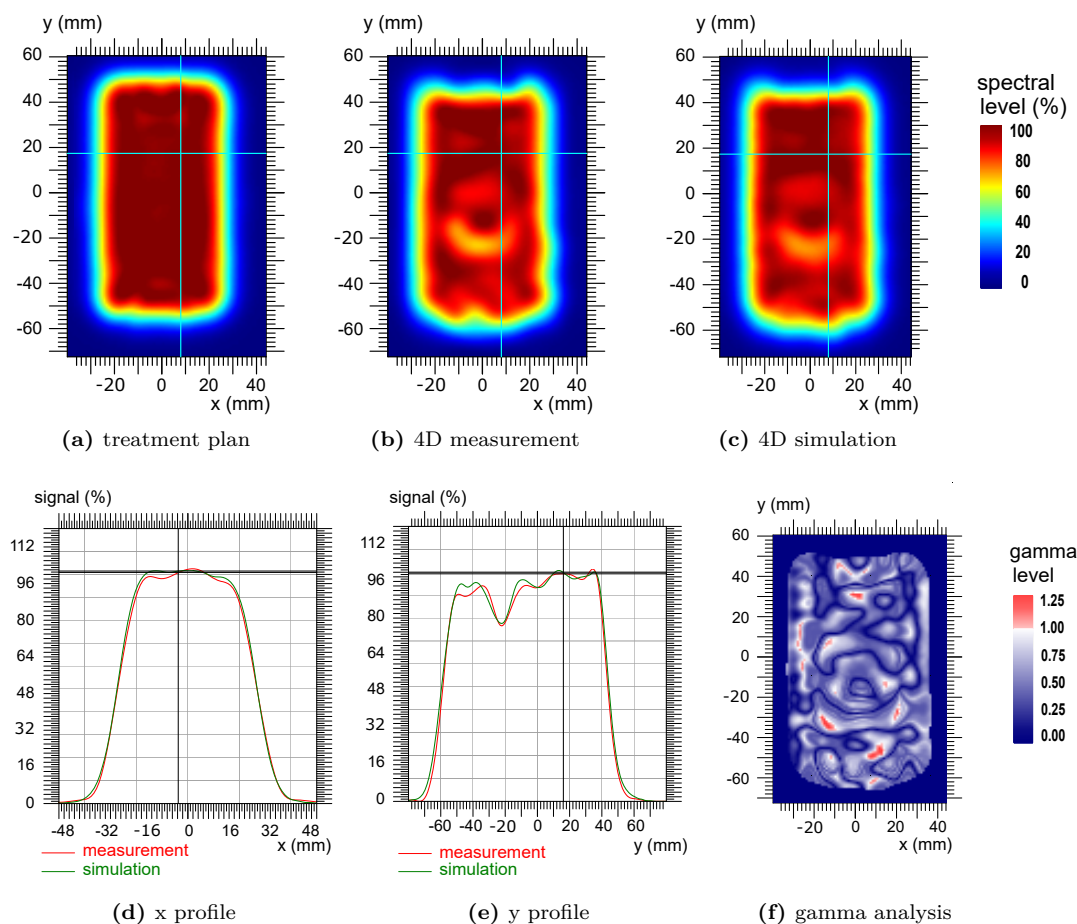


Figure 3.8: Relative dose analysis for the dynamic dose measurement (setup B): (a) static treatment plan, (b) measured 4D dose distribution, (c) 4D reconstructed dose distribution, (d) comparison of the 4D dose profiles in x direction (left-right), (e) comparison of the 4D dose profiles in y direction (superior-inferior), (f) gamma analysis of the 4D dose distribution with a 3 mm/3% passing criterion, gamma pass rate = 98.0% .

impact of small timing errors. Further, the 4D dynamic dose accumulation is more prone to errors for a moving detector (setup B). Even small uncertainties and approximations can have a visible impact on the dose distribution since the measurement points on the matrix are shifted. Contrary, for setup A, the positions of measurement points on the matrix remain unaffected and the same uncertainties and approximations can only cause range changes.

Even though there were some local quantitative deviations, changes in the relative dose distribution due to interplay effects were simulated correctly. Differences between the dynamic simulation and the measurement probably arise from the division of the continuous motion into 10 phases, CT artefacts, residual uncertainties of the DIR and the CT number to density calibration curve, the conversion of the dose grid and slight shifts in the setup for the irradiation compared to the CT acquisition.

The lower gamma pass rate for the second test series of setup B does not primarily indicate shortcomings of the routine but rather points out potential for improvement of the experimental conduct. For example, errors in the time structure could be reduced by determining the time offset of the start phase, an aspect which is addressed in more detail in the context of the sensitivity analysis below.

The differences in the absolute dose of about 8% are the same as for the static control measurement and therefore not attributable to errors in the 4DDD computation. They result more

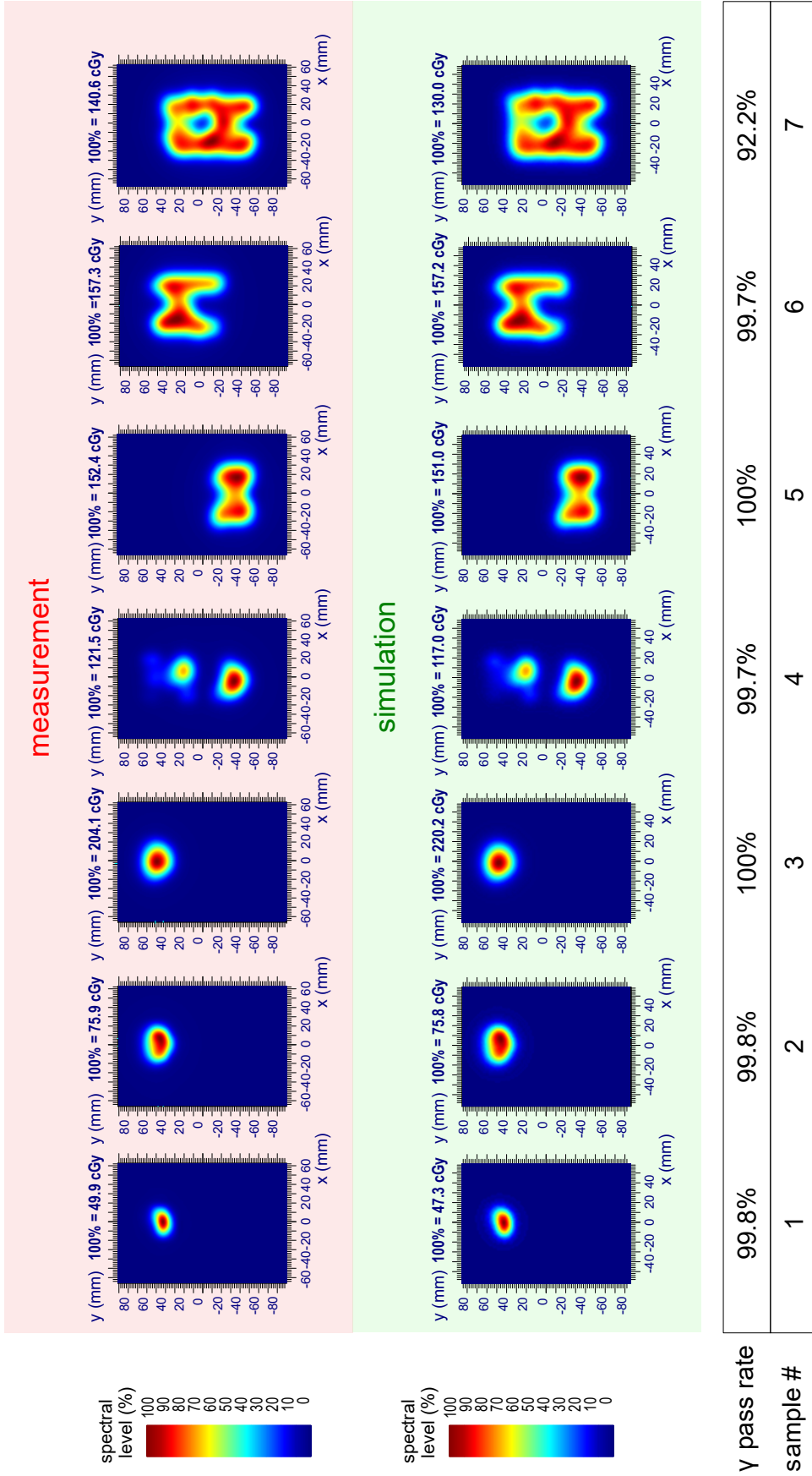


Figure 3.9: Comparison of a time-resolved dose measurement in movie mode with 4D dynamic dose computations for setup A. The sampling time was 2 s per image. Images 1 to 7 correspond to the time interval from 0 s to 14 s.

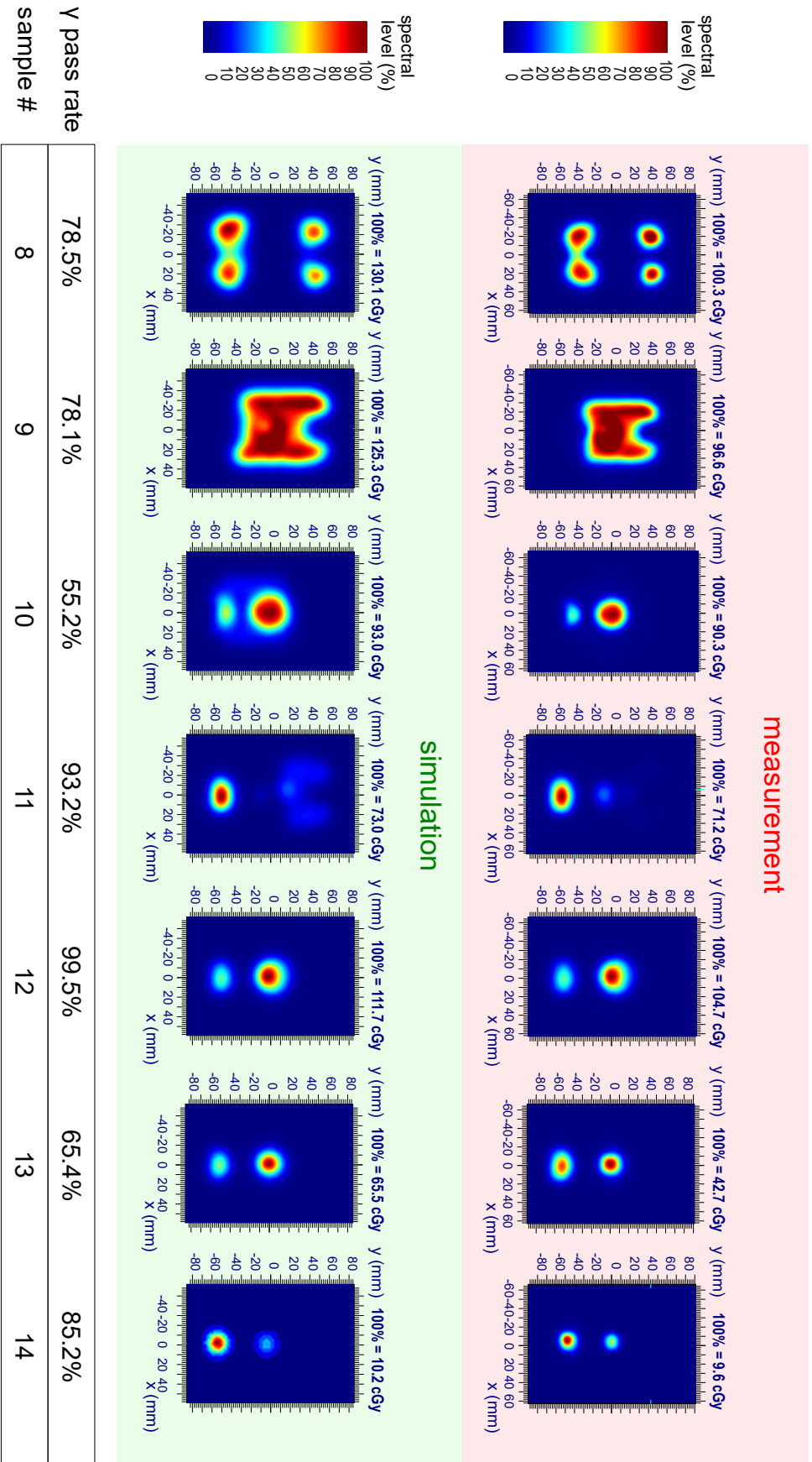


Figure 3.10: Comparison of a time-resolved dose measurement in movie mode with 4D dynamic dose computations for setup A. The sampling time was 2 s per image. Images 8 to 14 correspond to the time interval from 14 s to 28 s.

likely from deficiencies of the Pencil Beam dose algorithm (Sec. 2.4.5) in combination with the chosen phantom geometry which exhibit an air gap of 8.8 cm. Since the transport algorithm does not account for inhomogeneities, it underestimates the lateral spread of the nuclear halo. This is usually not critical for normal patient geometries but may become an issue if a thick compensator or range shifter is used with a large air gap (a setup with large inhomogeneities). The underestimation of the spread of large angle scattered protons lead to an overestimation of the dose on the central axis, particularly for small- and medium-sized fields. The resulting dose difference may amount to several percent [59]. This could be the case for setup A where the phantom has a similar effect as a range shifter or compensator due to the air gap in front of the detector.

Sensitivity towards start phase and motion period

The sensitivity of the routine towards its input parameters could be demonstrated. Differences of the specified start time larger than the length of a motion phase or deviations of the motion period by more than 0.2s led to measurable differences in the dose distributions. Since the motion period of the dynamic platform is highly reproducible and the uncertainty in the experimental determination of the start phase by camera records is less than the length of a motion phase, inaccuracies of the input are not expected to distort the calculation. Nonetheless, there is still potential for improvement. For instance, it would be more accurate to pass the exact delivery start point of the motion curve to the routine instead of the start phase which takes up to 0.6s for a patient (assuming 10 respiratory phases). First attempts showed that the gamma pass rate would improve slightly for cases in which the delivery started in the middle or at the end of a motion phase. However, a more practicable solution than camera records of the motion pattern and the time display should be applied in future. AlignRT (Vision RT Ltd., London, United Kingdom), a system which uses 3D stereo camera units to track a patient's position, could fulfil this function for technical validations and in clinical use.

Time-resolved evaluation

The time-resolved 4DDD computations showed qualitatively a good agreement with movie mode measurements. It was possible to reproduce differences in the dose delivery over time arising from variations in the beam delivery sequence for the same treatment plan. However, despite a gamma pass rate of above 98% for the sum of all partial dose distributions, the pass rate of a single measurement went down to 55% (first trial) and 48% (second trial) in the worst case. This could be caused by limitations of the methodology applied. For instance, the offset between the start of the measurement and the irradiation was determined with an accuracy of up to two decimals but the delivery time of a single spot takes only a few ms (see Sec. 3.3.2). Spots which were delivered at the beginning or end of a time interval might have been assigned to the wrong time interval. The risk of misassignments does not exist if the beginning/end of a time interval coincides with the change of an energy layer when the irradiation is paused. The hypothesis, that some spots were assigned to wrong time intervals by the routine, is supported by the visual examination of the measurements. For example, the upper right point in image number 8 (see Fig. 3.10) exhibits a lower intensity in the simulation than in the measurement. In the subsequent time interval (image number 9), looking at the same area, the simulated intensity is increased compared to the measurement.

Another reason for low gamma pass rates is that the number of spots which contribute to a partial dose distribution is much smaller than for the total dose distribution. Thus, the relative impact of spots which were assigned mistakenly to a wrong respiratory phase increases. The determination of an exact start point within the first respiratory phase as stated previously would prevent incorrect assignments. Another aspect could be the volume averaging effect of the ionisation chambers. Due to their relatively large diameter of 4.2mm, the dose might be perturbed when averaging over the active volume, a well-known problem especially for the dosimetry of small fields [110]. A better agreement between measured and calculated dose distributions is expected when simulating the volume averaging effect by a low-pass filter

function. The function must correspond with the lateral dose response function of an ionisation chamber and therefore has to consider the shape (circular) and size (diameter=4.2mm) of the chamber profile. The dead time of the MatriXX PT could be excluded as a possible cause since the detected dose differences between movie mode and single shot mode measurements were smaller than 0.3%.

The following aspects were investigated in order to explain the differences in the gamma pass rates between different images: the coincidence of the energy layer switching with the start/end of the time interval, contributing respiratory phases, dose differences between measurement and simulation, the number of spots and the amount of dose. The majority of distributions which exhibited a low gamma pass rate started and ended within an energy layer so that spots could have been assigned to the wrong time interval. But there were also some distributions with high gamma pass rates for which that applied. In these cases, the dose rate might have been lower at the interval boundary. It was expected that a wrong allocation of a proton spot to a respiratory phase would have a greater impact on the dose distribution for phases near the mid-ventilation position, where larger organ motions occur. Indeed, most of the distributions with a low gamma pass rate were delivered close to the mid-ventilation phase - but again, a few distributions with high gamma pass rates existed which also included phases near the mid ventilation position. A large difference in the absolute dose between measurement and simulation was found to be a good indication for low gamma pass rates. Conversely, small dose differences did not ensure high gamma rates. A low dose or a small number of spots in combination with a mismatch of proton spots to a wrong time interval or respiratory phase was expected to significantly affect the pass rate. However, a clear correlation could not be observed, except that the average amount of dose which was delivered in the first half of the irradiation was slightly higher.

Both test runs revealed on average a deterioration of the gamma pass rate for the second half of the irradiation. Since the start time of each spot was gained from log files, the worse gamma pass rates in the second half of the irradiation cannot be explained by the summation of time errors. Contributing factors might be differences in the amount of dose and the dose rate or a small deviation between the depth of the exported dose plane and the measurement depth. Such a difference in depth would affect proximal energy layers (second half of the irradiation) to a greater extent because they generate steeper dose gradients in the measurement plane which was located in the middle of the SOBP. However, there are still images with high gamma pass rates in the second half of the irradiation and errors might not be attributable to a single source of error but maybe to a coincidence of several of the above listed factors.

3.3 Development and validation of a beam delivery time model

The implementation of 4DDD computation in RayStation, discussed in the previous section, enables retrospective dose analyses based on irradiation log files, e.g. after a treatment fraction. However, there are some instances in clinical praxis and research, where prospective analyses are demanded. These include for example the evaluation of different treatment approaches for patients or in silico studies on motion mitigation techniques for a huge patient cohort. In such cases, it is unpractical or even not feasible to perform the required number of test irradiations due to the limited manpower and beam time. In order to estimate a 4DDD distribution a priori, access to the scanning control system calculations or an appropriate, facility-specific beam time model are needed. Since the exact logic of the control system is confidential information of IBA, an empirical beam delivery time model had to be developed within the scope of this thesis and will be presented in the following.

3.3.1 Materials and methods

The construction of an empirical beam time model involved the thorough examination of beam delivery parameters, i.e. the spot delivery time, the scanning speed and the energy layer switching time, and their dependencies on the delivered dose per spot, spot distance and beam energy for typical clinical values (see Sec. 2.3.2). Test irradiations were carried out based on dedicated PBS layer definition files (PLD files) to enable systematic analyses. A PLD file is used

to transfer predefined irradiation settings to the treatment machine and contains information on the number of energy layers, the cumulative meter set weight of all spots, the total amount of monitor units to be delivered, the beam energies, the number of rescans, the position in the isocentre plane and the meter set weight of each spot. For the investigation of the spot delivery time, seven PLD files were created with maximum energies ranging from 130 MeV to 192 MeV and minimum energies ranging from 100 MeV to 120 MeV being representative for clinical treatment scenarios. Each energy layer contained spots weights from 0.04 MUs to 2.5 MUs. After the irradiations, the corresponding log files were evaluated. The average delivery time per MU was calculated based on 210 data points for each spot weight. The scanning speed was studied separately for the x and y directions for 19 energies ranging from 102 MeV to 160 MeV. Spot distances of up to 9 mm were included in the PLD files. The measurements were repeated four times in a time interval of about three months and served also for the evaluation of the energy layer switching time. Since the acquired data formed the basis for the beam delivery time model, the standard deviation was an important statistical measure. It was used to assess whether the data were spread out over a wide range or whether they could be well described by average values.

Besides the above-mentioned beam parameters, the order of the spots must be known for the interplay evaluation. A description of the spot sorting for the IBA scanning control system is provided in Sec. 2.3.2. The implementation of the spot sorting in the interplay effect routine was tested based on the reference positions documented in the log files.

The resulting beam delivery time model was validated by comparing 4DDD distributions based on the actual delivery times extracted from log files with 4DDD distributions based on the delivery times calculated by the beam time model. Time differences were assessed for each energy layer and for the total delivery time of each field. This was done for the two treatment plans presented in Section 3.2.1 using the log files which correspond to a motion period of 5 s (three measurements per plan, six log files in total).

3.3.2 Results

Determination of the beam delivery parameters

The dependency of the spot delivery time on the spot weight is shown in Figure 3.11. For spot weights higher than 0.4 MU, the delivery time per MU was approximately constant. Below 0.4 MU, the time needed per MU started to increase strongly with decreasing spot dose since the relative contribution of the turn-off time of the system grows. Depending on the energy of the layer under consideration and the maximum energy of the field, the delivery time varied slightly which is represented by error bars indicating the standard deviation in Figure 3.11. For spot weights above 0.1 MU, the standard deviation remained below 0.3 ms/MU. The maximum standard deviation was about 2.7 ms/MU at 0.04 MU (lowest spot weight). In order to describe the observed time behaviour with an accuracy better than $\pm 2\%$ (with the exception of one data point in the gradient region of the curve) while keeping the number of fit parameters low, a rational function was used for fitting the data with the Nelder–Mead algorithm [111] from GNU Octave, version 4.0.0.

The scan time, here defined as the time needed to change between two spot positions, is depicted in Figure 3.12 as a function of the spot distance. For both scanning directions, x and y, the scan time grew with increasing spot distance. While it took only about 2.25 ms to 2.5 ms to scan in x direction for spot distances in-between 1 mm and 9 mm, it took about 2.25 ms to 4.75 ms for the same distances in y direction. Step functions were defined for the beam time model resembling the trend of the data points as close as possible.

The energy layer switching time did not show a correlation with the energy or the energy difference between layers. It exhibited strong variations even for repeated irradiations of the same PLD file. Outliers due to longer irradiation pauses which can occur after monitoring interlocks were excluded from the evaluation. The measured average value μ was 1.23 s and the standard deviation σ was 0.26 s. For the beam time model, random values were generated from a normal distribution ($\mu = 1.23$ s, $\sigma = 0.26$ s). In addition, values below 0.9 s were set to 0.9 s

and values above 2.5 s were set to 2.5 s.

An algorithm which resorts the proton spots according to the scanning path of the treatment machine was implemented in the interplay effect routine. The calculated spot order matched the one recorded by log files.

Comparison with electronic irradiation protocols

The time structure calculated by the beam time model was compared to the actual beam delivery sequence extracted from log files. In average, the time per energy layer differed by 0.03 s and the total time per field by 1.2% assuming identical energy layer switching times. The impact of such small time differences on the simulation of interplay effects was examined by comparing the resulting 4DDD distributions. Figure 3.13 shows exemplarily one of the 4DDD analyses for setup A. The dose distributions were identical within the tolerances defined in Section 3.2.1. Minor differences could be seen in the 2D profiles (less than 1.5% within the plateau region). The gamma pass rate reached 100%.

3.3.3 Discussion

Since WPE has no access to the algorithms of IBA's scanning control system, the beam time model had to be created on the basis of empirical findings. The model is therefore only valid for the tested clinically relevant spot distances, energies and spot doses.

Determination of beam delivery parameters

In the PBS implementation of the IBA Proteus[®]Plus system, the energy layer switching time dominates the beam delivery process in terms of irradiation time. It is up to three orders of magnitude larger than the spot delivery time or the scan time. Nevertheless, delivery and scan times can add up to several seconds and the irradiation of an energy layer cannot be assumed to be instantaneous. Since the typical respiratory cycle length of an adult is about 3 s to 5 s, spots of one energy layer might be delivered over different respiratory phases. As a consequence, appropriate timing is required. The fit for the spot delivery time describes the measured data with high accuracy for spot weights above 0.1 MU. But even for the highest standard deviation of about 2.7 ms/MU at the minimum spot weight, the absolute time error is negligible with 0.1 ms. The approximation of the scan time by step functions leads mainly to uncertainties for the y direction, where one spot distance cannot always be assigned unambiguously to one scan time. Since the quantisation error is smaller than 1 ms and scanning is predominantly performed in x direction (faster scanning component), these errors will not add up to an unacceptably high value. To cope with the stochastic nature of the energy layer switching time, prospective analyses should include multiple interplay effect evaluations per treatment plan, each with random switching times. In this way, the most probable outcome of the treatment as well as the worst case scenario can be investigated in advance.

Comparison with electronic irradiation protocols

The average time error per energy layer of 0.03 s, caused by the beam time model, is negligible compared to the length of a 4D CT phase of typically 0.3 s to 0.6 s. The total time difference per field, however, reached up to 0.44 s (1.7%). Comparisons of 4DDD distributions based on log files and the beam time model could demonstrate that errors of this order of magnitude do not have a noticeable effect on the dose distribution. Hence, the beam time model was considered to be accurate enough for interplay effect evaluations.

In some in silico studies on interplay effects in proton therapy (for instance Ref. [30]), beam delivery parameter were assumed to be constant. Therefore, the effects of an over-simplified beam time model or an incorrect distribution of proton spots over the different respiratory phases (e.g. only energy layer-wise) shall be demonstrated based on two examples: Assuming

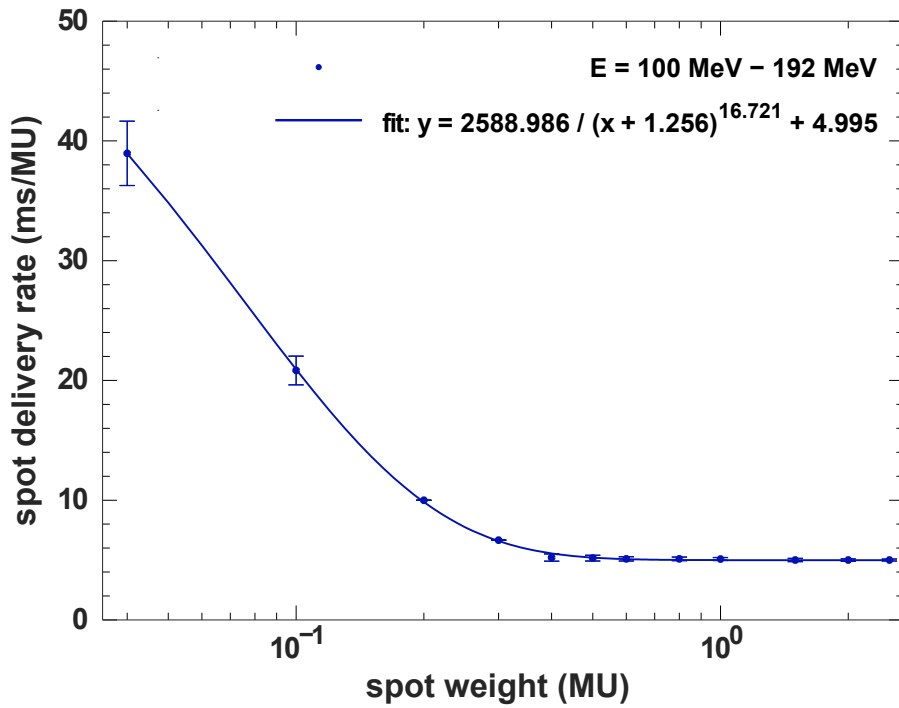


Figure 3.11: Beam delivery time per monitor unit as a function of spot weight: The average time over all beam energies is marked by dots for the investigated spot weights. Error bars indicate the corresponding standard deviation and a solid line represents the fitted curve. Note the logarithmic scale of the x axis.

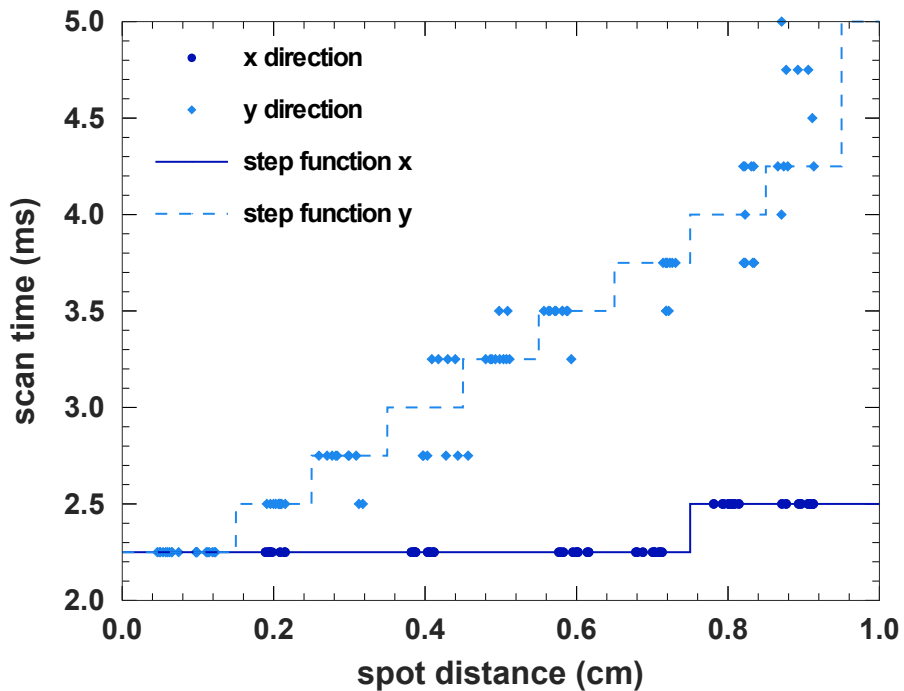


Figure 3.12: Scan time in x and y direction as a function of the spot distance: Measurements were performed for beam energies ranging from 102 MeV to 160 MeV. A solid and a dashed line represent the step functions used to model the scanning speed in x and y direction in the facility-specific beam time model.

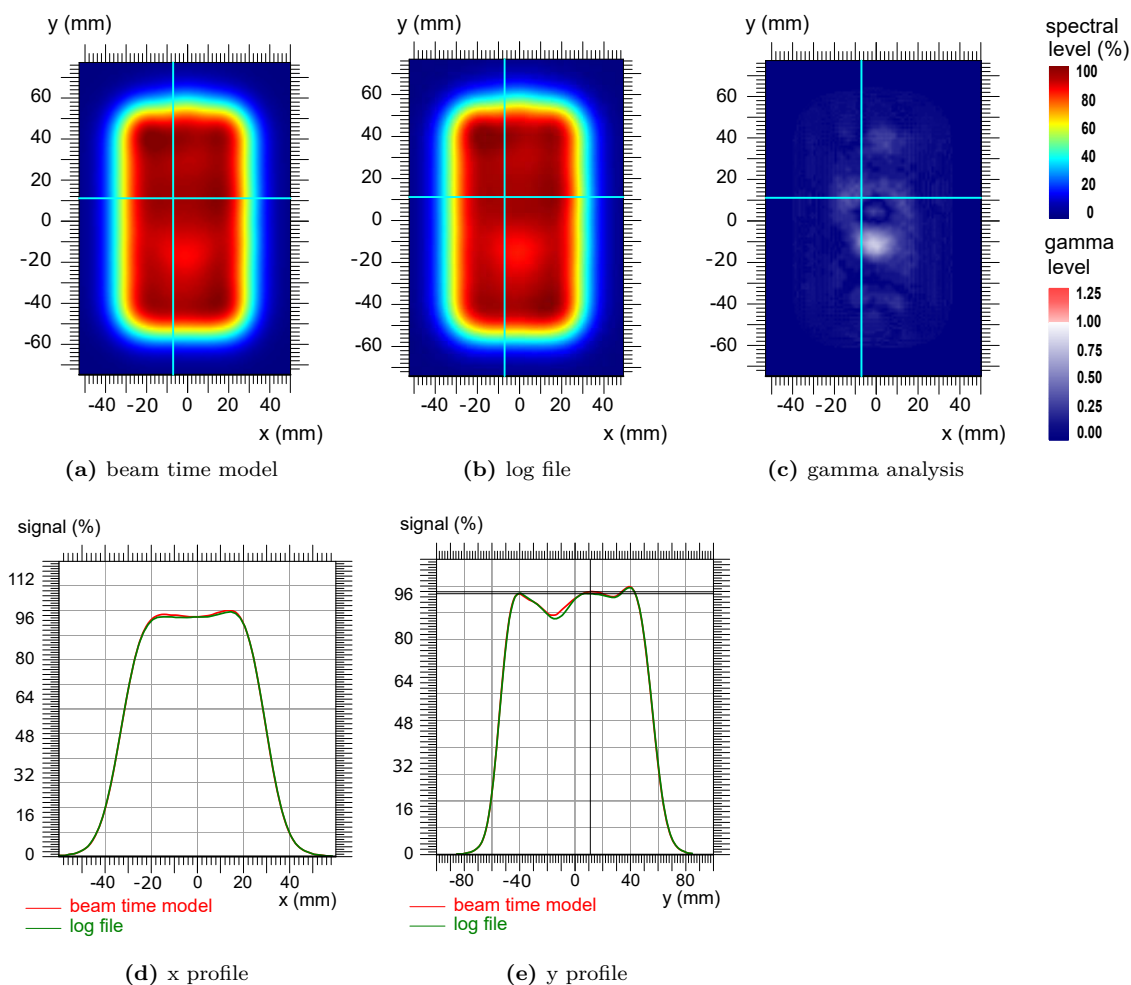


Figure 3.13: Evaluation of the developed beam time model using the treatment plan which was initially created for the irradiation of setup A (Sec. 3.2.1): (a) 4D reconstructed dose distribution based on the beam time model (isocentre plane), (b) 4D reconstructed dose distribution based on irradiation log files (isocentre plane), (c) gamma analysis with a 3 mm/3% passing criterion, gamma pass rate = 100%, (d) comparison of the dose profiles in x direction (left-right), (e) comparison of the dose profiles in y direction (superior-inferior).

a constant scanning speed of 250 cm/s, a spot delivery time of 5 ms/MU and an energy layer switching time of 1.2 s, the gamma pass rate deteriorates from 99.7% to 91.6%. The approximation of an instant dose delivery for each energy layer even leads to a reduction of the gamma pass rate by 10%.

3.4 Clinical aspects

Future challenges in interplay modelling emerge in the areas of 4D imaging, beam delivery and treatment planning. The trend towards the treatment of moving targets with PBS results particularly in a demand for better motion monitoring also during irradiation. For this purpose, imaging modalities such as 4D MRI, fluoroscopy or ultrasound could be suitable candidates. The so gained information on organ motion and deformation has to be transferred to the planning CT in order to calculate the dose distribution for different motion phases. Thus, the developed interplay effect routine can be easily combined with other imaging modalities than CT.

Continuous developments in the beam delivery in the last few years led to shorter energy layer switching times and changes in the scanning speed and delivery time. A development which decreases the treatment time and improves the robustness against organ motion is for example the continuous line scanning investigated by Fattori et al. and Klimpki et al. [112, 113]. But also software and hardware upgrades of the beam delivery system can affect the time structure and thus the interplay pattern. Therefore, it is crucial to keep the beam delivery model implemented in the interplay effect script up to date. For retrospective analyses, the interplay effect routine remains valid even after major changes of the beam dynamics since the timing is read out from log files.

The interplay effect routine can be used not only to analyse the impact of actual changes in the delivery process but also to actively search for the best delivery techniques and optimal beam parameters with regard to the 4DDD distribution. The determination of the optimal pause between two energy layers dependent on the respiratory cycle length or the optimal scanning path could for instance improve the robustness against motion in future applications. The time-resolved version of the interplay effect routine might be a helpful tool for such purposes. However, a methodology to synchronise the OmniPro-ImRT measurement with the irradiation has to be developed before it can be used for dose reconstruction. Otherwise, the uncertainty in the determination of the offset can affect the calculated dose distribution for certain time intervals as discussed above. Another application of the time-resolved interplay routine could be to determine individual gating windows for respiratory-gated irradiations. The comparison of the static and the dynamic dose distribution over time reveals the magnitude of the dose distortion for different time intervals of the respiratory cycle and enables to adjust the gating window accordingly.

In treatment planning, it is expected that the potential benefits of intensity-modulated proton therapy (IMPT) and 4D robust optimisation will be better exploited. Consequently, the standard robustness evaluation including range uncertainties and shifts of the isocentre (Ref. [114]) has to be extended by the interplay effect analysis to test additionally the robustness against organ motion. The dose modulation in IMPT leads to different beam delivery time structures compared to single field uniform dose plans which were used for the experimental validation of the interplay effect routine. This does not pose a problem for prospective evaluations of IMPT plans since the beam time model was created based on the analysis of a wide range of clinical relevant beam delivery parameters including spot delivery times and spot distances used in IMPT. Also the application of multiple fields can be easily simulated with the interplay effect routine as shown in the scope of the hepatocellular carcinoma study in the next chapter. In case of a prospective evaluation, the routine chooses random start phases for each field.

The use of motion mitigation techniques has to be considered in 4DDD evaluations. The presented interplay effect routine already enables the simulation of beam gating and layered rescanning (simplified model). The gating window can be selected through the graphical user

interface based on the choice of respiratory phases for which the beam should be turned on. The number of rescans is automatically taken from the stored treatment plan parameters. Future script versions should take into account the minimum spot weight for rescanning techniques. The consideration of further motion mitigation techniques is possible by small modifications of the script when needed.

The impact of variations in the breathing pattern on the interplay effect evaluation is a clinical aspect which is beyond the scope of the current work and will be investigated in the future (see Sec. 6.4 on future research).

4 Evaluation of 4D robust optimisation for hepatocellular carcinoma patients

The treatment planning study presented in this chapter investigates the ability of 4D robust optimisation to reduce motion induced dose distortions in PBS, while sparing surrounding healthy tissue, for HCC patients. The liver is particularly suited to study motion interplay effects since the treatment region exhibits smaller density gradients and more homogeneous tissue than targets in the thorax, making it less prone to range errors. In case of thoracic targets, for example, the microstructure of lung tissue cannot be resolved by CT images and degrades the accuracy of the dose calculation [21, 22, 23, 24].

Due to the exceptional high mortality rate of HCC patients, an effective local therapy suitable for a wide range of HCC patients is urgently needed as will be explained in Section 4.1. The treatment of HCCs requires very high precision since a sufficient amount of healthy liver tissue must be spared to maintain the liver function. Proton therapy could provide the demanded dose conformality, particularly using PBS. However, contrary to DS, interplay effects in PBS might compromise the target coverage in the presence of organ motion. The motion magnitude of liver tumours depends on the location and ranges from a few millimetres to almost 2 cm [115, 116]. The largest motion magnitudes have been observed for the right upper lobe (adjacent to the diaphragm) [115].

Within this study, it should be examined whether 4D robust optimisation in PBS

1. provides a better 4DDD target coverage than the conventional PTV concept and reduces the disadvantage in dose homogeneity of PBS compared to DS,
2. ensures similar or better sparing of healthy liver tissue than the PTV concept and does not lower the advantage in conformity over DS.

A conservative single field uniform dose approach was applied for PTV-based PBS plans since geometric margins might be less resistant to disturbances which cause position-dependent changes of the water equivalent thickness. In contrast, multi-field optimisation was used for 4D robustly optimised PBS plans to exploit the full potential of intensity-modulated proton therapy.

Partial results of the presented work in this chapter have been submitted to/ published in peer-reviewed journals [117, 118].

4.1 Hepatocellular carcinoma (HCC)

4.1.1 Clinical background

The liver fulfils important functions such as the production of bile, the metabolism of ingested nutrients, the elimination of toxins, the production of hormones, the regulation of glycogen storage and plasma protein synthesis. Since it is a parallel organ which is composed of numerous functional subunits, it can tolerate local damage to a certain extent without losing functionality.

With a proportion of 85% to 90%, HCC represents the most common type of primary liver cancer [19]. The incidence of HCC is highest in East and South-East Asia and Sub-Saharan Africa, where hepatitis B is wide-spread. However, owing to the growing occurrence of hepatitis C infections and, to a lesser degree, alcoholic liver damage, the prevalence of HCC is increasing in Europe and in the United States [119]. HCC is often related to an underlying chronic liver disease such as liver cirrhosis. Consequently, common causes of liver cirrhosis represent key risk factors for HCC. The major risk factors are hepatitis B and C viruses, alcohol abuse and non-alcoholic fatty liver disease. Other factors include aflatoxin intake, diabetes and immune-related diseases of the liver such as autoimmune hepatitis and primary biliary cirrhosis [120, 121]. Symptoms of HCC are often manifestations of a decompensated cirrhosis and usually appear at a late stage of the disease. Non-specific general signs include asthenia (weakness), anorexia (loss of appetite), nausea, weight loss and fever. Furthermore, clinical signs such as jaundice,

gastrointestinal bleeding, peripheral oedemas, abdominal pain or portal vein thrombosis can be an expression of onset of HCC. The disease is associated with a poor prognosis and an exceptionally high mortality rate with a five-year survival of less than 5% [2]. That makes it the third leading cause of cancer-related mortality worldwide even though it is only the fifth most common malignancy in men and the seventh in women [20, 122]. Since symptoms of HCC often do not show up in the early stages, 10% to 20% of the tumours already have a diameter of more than 10 cm at the time of first diagnosis [123]. In order to preserve as much of the surrounding healthy liver tissue as possible local treatment methods are demanded.

4.1.2 Treatment options and the role of proton therapy

The treatment of HCC is often a challenging task since more than 80% of the patients suffer from an underlying liver cirrhosis [124]. A cirrhosis is associated with an impaired liver function and thus represents an exclusion criterion for many treatment options. New therapy recommendations are therefore usually based on the Barcelona Clinic Liver Cancer staging classification which not only considers the number and size of tumours but also the general state of health and the liver function status according to the Child-Pugh score [120]. While a curative approach is followed to treat HCCs in early stages, local control and prolongation of life are clinical goals for advanced HCCs [121].

Commonly applied curative treatment methods are surgical resection, transplantation and radiofrequency ablation. The best long-term survival was demonstrated for surgery (resection or transplantation) with a five-year survival rate of about 70% [125]. However, most patients cannot undergo a surgical resection due to comorbidities such as cirrhosis and the eligibility criteria for liver transplantation are strict. The Milan criteria established itself as the worldwide standard for transplant patient selection to ensure excellent post-transplant survival. They are based on clinical experience and restrict transplantation to patients with a single tumour smaller than 5 cm or two to three tumours smaller than 3 cm each, who do not have an extrahepatic disease or macrovascular involvement [126, 127].

If neither resection nor transplantation can be performed, radiofrequency ablation (RFA) might be a treatment option. RFA, a minimally invasive procedure, uses a needle electrode to create an alternating electric field within the tumour and destroys tumour cells through frictional heat. It can be applied for tumour sizes up to 5 cm but is not suited for locations close to major vessels, the major bile duct, the diaphragm or other abdominal organs. Due to a high local recurrence rate, the long-term survival is worse than for surgery [128].

Other, non-curative local treatment options exist for patients who do not meet the criteria of those approaches or have to wait for the transplantation. In the latter case, the treatment serves as bridging method to prevent tumour growth. If the tumour size exceeds the Milan criteria, non-curative therapies can be applied in some cases to decrease the tumour so that the patient becomes eligible for a transplantation (down-staging) [129].

The most frequently used non-curative therapy for intermediate stages is transarterial chemoembolisation (TACE). The technique is based on the injection of small particles coated with chemotherapeutic drugs in the hepatic artery and kills tumour cells in two ways. Firstly, particles block or slow down the blood supply to the tumour and secondly the chemotherapeutic agents induce cytotoxicity. In contrast to the tumour, which is supplied through the hepatic artery, normal liver tissue gets blood from the portal vein and is thus not affected from cutting off the blood flow. After TACE, a small improvement in local control and overall survival was observed but also significant complications [130, 131]. Contraindications for TACE include main portal vein thromboses.

Further treatment options, which can be applied either individually or in combination with other therapies, are external beam radiation therapy, systemic treatment with the multi-kinase inhibitor sorafenib and internal radiation therapy with radioisotopes [127]. Apart from modern irradiation techniques, each approach is only suited for a small subgroup of patients, not very effective or associated with unacceptable toxicities. An effective local therapy which can be applied for a wide range of patients independent of tumour size, location and underlying liver disease is needed.

In the past, radiotherapy only had a minor role in the treatment of HCC because of the low tolerance dose of normal liver parenchyma³, which is even reduced in case of a chronic liver disease [132]. Since early radiation techniques were not able to spare an adequate amount of normal liver parenchyma, they caused high incidence rates of radiation-induced liver disease (RILD). RILD, occurring typically two weeks to four months after the treatment, is associated with an increase in hepatic enzymes (particularly alkaline phosphatase), hepatomegaly (an enlargement of the liver) and ascites (accumulation of fluids in the peritoneal cavity⁴). It is caused by a venoocclusive disease, i.e. the obstruction of small veins in the liver [133]. In severe cases, RILD can lead to irreversible hepatic failure and death [121]. An effective therapy is still lacking. The incidence rate of RILD is about 5% to 10% for irradiations of the whole liver with 30 Gy_{RBE} to 35 Gy_{RBE} in conventional fractions [134]. In the early days of HCC irradiations, the prescribed dose was therefore limited to 30 Gy_{RBE}. Since higher dose levels are required to achieve a therapeutic effect, HCC was considered to be radioresistant for a long time. Modern irradiation and imaging techniques enable an escalation of the dose to the tumour without unacceptable side effects since large parts of the functional liver tissue can be kept out of the high-dose region [134]. Thus, radiotherapy is no longer only a palliative approach but also an effective treatment for patients who do not qualify for other curative options.

A particular promising approach in radiotherapy of HCC is proton therapy. Compared to photons, protons exhibit a more favourable depth dose profile that allows for very conformal dose delivery and thus better sparing of normal liver tissue (see Sec. 2.1.3). The efficacy and safety of proton therapy for HCC is subject of current research. Encouraging results on local tumour control and survival for HCC patients were already reported by investigators from Japan, where HCC is common and the number of proton therapy facilities comparatively large [124, 135]. Also phase II trials conducted by the Loma Linda University Medical Center (Loma Linda, California, USA) and the Massachusetts General Hospital (Boston, Massachusetts, USA) demonstrated high local control rates and minimal acute toxicities for HCC patients [2, 136]. Korean researchers observed a complete response rate of 100% in a dose escalation study for the highest delivered dose, 72 Gy_{RBE} in 24 fractions corresponding to an equivalent dose in 2 Gy fractions of 78 GyE₁₀ (calculated using the linear quadratic model with an alpha to beta ratio of 10) [137].

So far, the majority of proton therapy studies on HCC has been performed with passive scattering techniques. However, recently built proton therapy centres use active pencil beam scanning. PBS is associated with a lower neutron dose, provides better conformity at the proximal edge of the target and enables IMPT. Thus, it potentially reduces the occurrence of acute toxicities and radiation-induced secondary cancer. On the downside, PBS exhibit a larger lateral penumbra and is more sensitive to motion (Sec. 2.3.2).

Different planning strategies (e.g. beam-specific margins or robust optimisation) and motion mitigation techniques can be applied for the treatment of moving targets in PBS. The effect of these features still needs to be examined for HCC patients.

4.2 Design of the treatment planning study

4.2.1 Patient collective

The patient collective consisted of nine HCC patients with a total of eleven CTVs. Without loss of generality, the CTV was defined as the GTV for the plan comparison. 4D CT data sets of the patients were kindly provided by O. Blanck and M.K. Chan from the University Clinic Schleswig-Holstein (Kiel, Germany) and M. Ayadi from the Centre Léon Bérard (Lyon, France). Seven of the 4D CTs have been acquired within the scope of a previous study [138] and exhibited good tumour contrast due to the selective uptake of lipiodol by the tumour cells after TACE. Tumour motion was restricted by abdominal compression for these patients (ID 1-7),

³Liver parenchyma: functional part of the liver tissue.

⁴Peritoneal cavity: potential space between the parietal peritoneum which surrounds the abdominal wall and the visceral peritoneum which surrounds the abdominal organs

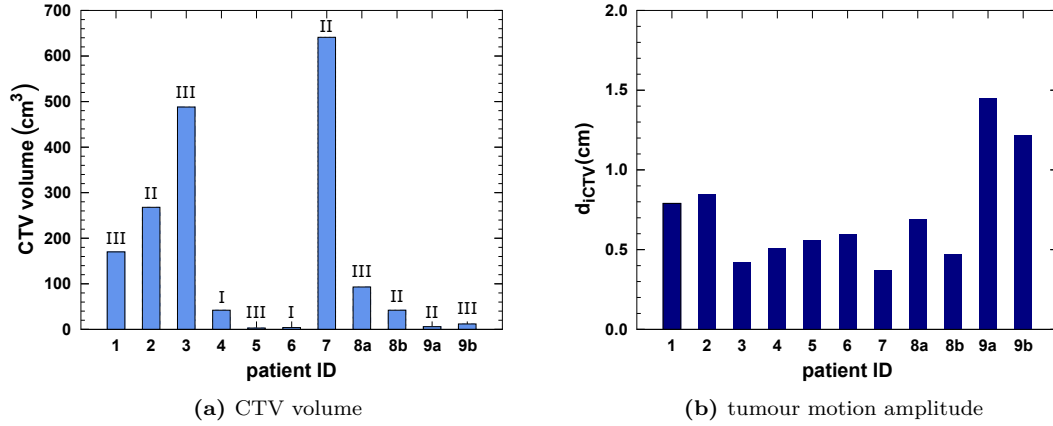


Figure 4.1: CTV volume (a) and tumour motion amplitude (b) as defined in Equation 4.1 for the selected patient collective. CTVs were categorised in groups according to their homogeneity level, where category I includes iCTVs encompassed entirely by liver tissue, category II iCTVs extending into other parts of the abdomen and category III iCTVs adjacent to the ribs or lung. Patient IDs labelled with the letters 'a' or 'b' refer to different CTVs of one patient.

whereas for patient 8 and 9 4D imaging was carried out during free breathing. Patient 8 had two 4D CT scans, one with contrast agent injection to localise the tumour and one without to calculate the treatment plan. Implanted fiducial markers served for the localisation of the tumour for patient 9. All 4D CTs were acquired with a Philips Brilliance Big Bore CT scanner (Philips Medical Systems, Cleveland, OH, USA) using retrospective spiral correlated imaging (see Sec. 2.5.4). Three-dimensional CT data sets were reconstructed from 10% time intervals equally spaced over the respiratory cycle (phase binning), yielding 10 motion phases. The 0% phase corresponded to end-inspiration and the 50% phase to end-expiration. The CT slice thickness was 3.0 mm for patients 1 to 7, 1.0 mm for patient 8 and 2.0 mm for patient 9. The lateral pixel spacing in x and y was 1.2 mm for patients 1 to 7, 1.1 mm for patient 8 and 1.0 mm for patient 9.

The CTV volume, shown in Figure 4.1a, ranged from 3 cm³ to 640 cm³. The tumour motion amplitude d_{iCTV} was defined as the average deformation vector length of the deformable registration in-between end-expiration and end-inspiration [30]:

$$d_{iCTV} = \frac{\sum_{i=1}^N \sqrt{x_i^2 + y_i^2 + z_i^2}}{N}, \quad (4.1)$$

where the internal CTV (iCTV) represents the union of all CTVs on the different 4D CT phases, x_i , y_i and z_i are the deformation vector components of voxel i within the iCTV and N is the total number of iCTV voxels. d_{iCTV} varied between 0.4 cm and 1.5 cm and is displayed in Figure 4.1b.

Depending on the homogeneity level of the iCTV and its surroundings, all CTVs were divided into three categories. Category I comprised all iCTVs surrounded by liver tissue (ID 4 and 6). Category II included iCTVs extending partly into other parts of the abdomen such as the stomach or the kidney (ID 2, 7, 8b and 9a). Category III contained iCTVs adjacent to the ribs or the lung (ID 1, 3, 5, 8a and 9b). In the following, the CTV corresponding to patient ID n is referred to as CTV n .

4.2.2 CT scanner calibration, contouring and deformable image registration

CT numbers were converted into densities using a stoichiometric calibration curve which covers a range from -1000 HU to 2976 HU and has been created within the scope of a Master's thesis at WPE [139]. Material overrides were used to correct for image artefacts and contrast enhanced regions in the liver. To this end, the average liver density was determined by placing a spherical ROI with a radius of 1 cm to 2 cm in a homogeneous region of the normal liver tissue (liver volume minus CTV).

Contours for the CTV, the body and the External had to be delineated on every 4D CT phase for each patient since they were required for deformable image registrations, 4D dose computation and the creation of the iCTV. OARs were only contoured on the planning CT for the reasons explained below. In this study, the liver, the ribs, the heart, the lung, the stomach, the kidneys and the myelon were included as OARs. They were used to define objectives in treatment planning and to evaluate the 4D dynamic dose distribution in the end. A PTV was constructed for non-robust optimised treatment plans by adding margins for setup and range uncertainties to the iCTV (see Sec. 4.2.3). After contouring, all delineated ROIs were controlled and approved by experienced radiotherapists.

The deformable image registration algorithm ANACONDA described in Section 2.4.4 served for the registration of the 4D CT phases. Since contouring is very time-consuming for a 4D CT data set, a compromise was made between DIR accuracy and expenditure of time. The choice of controlling structures fell on the CTV and the body contour. In addition, the body was selected as focusing region. The reason why no further structures were included as controlling ROIs is that OARs in the abdomen typically exhibit similar contrasts as their surroundings and would have to be manually delineated for each motion phase, whereas the outer body contour can be easily created and transformed to other motion phases using automated planning tools. The CTV had to be contoured for each phase anyway since it was needed to create the iCTV. Further, it represents the most important controlling structure with regard to the dynamic dose evaluation.

The planning CT was chosen as reference phase of all DIRs in order to enable dose mapping for the 4D evaluation. Even though the algorithm has been validated before [66, 67, 68, 69], registration errors were assessed on a sample basis for the current patient cohort as recommended by the Task Group No. 132 of the Radiation Therapy Committee of the American Association of Physicists in Medicine (AAPM) [140]. For this purpose, the target registration error and the Dice similarity coefficient (DSC) were determined for randomly selected patients. The target registration error states the distance between an identified point a in phase A and the corresponding point a' mapped from phase B to phase A by means of image registration [140]. The average target registration error was determined based on two to three landmarks per patient considering six patients in total. Landmarks were identified in the 0% phase and 50% phase representative for the maximum deformation. Another four phases were exemplarily examined for two patients.

The DSC is a measure for the overlap between the reference segmentation volume S and the corresponding segmentation volume S' mapped from the deformable registered image. It is defined as double the volume of overlapping pixels divided by the sum of the two single volumes [141]:

$$DSC = \frac{2 \cdot S \cap S'}{S + S'}. \quad (4.2)$$

A DSC value of 1 indicates a perfect match and a value of zero no overlap at all. The DSC value of the liver structure was exemplarily examined for 4 patients considering the 0% and the 50% phase and required additional manual contouring of the liver on the 0% phase.

4.2.3 Treatment plan optimisation in PBS and DS

Treatment plans were created in a research version of RayStation 6 using the Pencil Beam dose algorithm and a dose grid of 2 mm. A uniform dose of 63 Gy_{RBE} (D_{presc}) was prescribed

to the CTV in 15 fractions according to Bush et al. [2]. The 50% phase served as planning CT since it is more reproducible and less affected by organ motion than the other respiratory phases [89, 142].

4D robustly optimised and 3D non-robustly optimised PBS plans

Two types of PBS treatment plans were generated per patient, a 4D robustly optimised plan and a non-robustly optimised, margin-based plan for comparison purposes. These plans are referred to as 4D PBS and 3D PBS plans hereinafter. The same gantry and couch angles were chosen for both techniques using two treatment fields per plan. The initial proton energy ranged from 100 MeV to 195 MeV. Range shifters were applied to reach shallow-seated tumours, when necessary. The air gap between a range shifter and the patient varied between 2 cm and 5.4 cm and was kept as low as reasonably possible. Distances between adjacent energy layers were set automatically with a scale factor of 1.1 with respect to the Bragg peak width (80% dose level). The spot distance was determined by multiplying 1.06 times the radial spread in the Bragg peak (average projected sigma) with a scale factor of 0.8 for each energy layer [77].

In 4D PBS plans, the robust optimisation settings were chosen to achieve plan robustness against 2 mm setup error, 5% range uncertainty and morphological changes in 10 respiratory phases. For the optimisation process, the resolution of the dose grid had to be set to 3 mm or 4 mm in order to prevent the TPS from reaching its limit due to the high data volume. After the optimisation, the final dose was calculated on a 2 mm dose grid based on the optimised plan parameters.

The applied perturbation parameters originate from the setup uncertainties (2 mm) and range uncertainties (2 mm plus 3.5%) clinically used at WPE. For the given energy range, a range uncertainty of 2 mm and 3.5% corresponds to a percentage uncertainty of about 5%. The robust optimisation considered CTV objectives. Objectives referring to OARs were only evaluated for the nominal scenario. The spots of all fields were simultaneously optimised using multi-field optimisation to achieve the best possible conformality through intensity modulation while ensuring robust CTV coverage.

In 3D PBS plans, uncertainties were taken into account by beam-specific PTVs and the whole optimisation was based on the 50% phase. The robustness settings listed above were used as margins in this case. The beam specific PTVs comprised the iCTV, a uniform setup margin and a geometric margin accounting for range uncertainties in beam direction (proximal and distal to the target). Contrary to the 4D PBS plans, a single field uniform dose approach was applied since geometric margins might be less robust against changes in the water equivalent path length.

Double scattering plans

A three-field concept was necessary to achieve adequate dose conformality for DS plans. The same setup and range uncertainties as for PBS were considered. The target was defined as the iCTV plus setup margin (2 mm). The range uncertainty (2 mm and 3.5%) was incorporated into the dose computation. In addition, the smearing radius of the compensator was set to 3 mm. Like for 3D PBS plans, only the 50% phase was considered in the dose computation. The SOBP range varied between 8.0 cm and 19.2 cm and the modulation width of the SOBP ranged from 3.5 cm to 13.6 cm.

4.2.4 CTV coverage and robustness evaluation

All plans had to fulfil the clinical goal $V_{95}(CTV) = 100\%$, where V_{95} is the percentage volume receiving at least 95% of the prescribed dose. In addition, a robustness evaluation was performed for each plan including a density perturbation, a shift of the isocentre and dose computation on different respiratory phases. Since it was not feasible to check all possible perturbation scenarios for every respiratory phase, a set of representative scenarios was chosen and evaluated

using the scripting interface of RayStation. The script is shown in section A.1 in the appendix. The permutation of the following cases

| | |
|------------------------|-----------------------------------|
| isocentre shift : | $dx = dy = dz = \pm 2 \text{ mm}$ |
| density perturbation : | $\pm 5\%$ |
| respiratory phases : | 0%, 50% |

resulted in eight perturbation evaluations. For large tumours, $V_{95}(CTV) > 99\%$ was demanded in each perturbation scenario, whereas for very small tumours with a size smaller than 20 cm^3 $V_{95}(CTV) > 98\%$ was required. Exceptions to this rule were made when it compromised the sparing of close-by OARs. Once a plan passed these criteria and was approved from medical physics experts and physicians, it underwent a detailed interplay effect analysis (PBS plans) or a less complex 4D dose analysis (DS plans), respectively (Sec. 4.2.5).

4.2.5 4D dose accumulation in pencil beam scanning and double scattering

The performance of 4D PBS, 3D PBS and DS plans was evaluated for the selected HCC patients based on 4D dose computations. In the following, a distinction is made between the 4DDD, which considers organ motion and sequential beam delivery, and the 4D accumulated dose (4DD), which assumes a uniform dose delivery over the respiratory cycle.

Pencil beam scanning

In PBS, 4DDD distributions were calculated using the developed interplay effect routine (Sec. 3.1). The beam delivery was simulated with the empirical beam time model presented in Section 3.3. The advantage of the model-based computation over the use of irradiation log files was that possible fluctuations of the beam delivery parameters could be simulated and considered for the planning study and that no beam-time was required. Respiratory start phases were randomly selected for each field by the routine and energy layer switching times were sampled from a normal distribution (see Sec. 3.3.2). For all evaluations, a constant respiratory period of 4 s was assumed and the default beam delivery settings defined in Section 3.3.2 were applied (settling time of the fast/slow scanning magnet = 2.25 ms / 2.5 ms, average / standard deviation of the energy layer switching time = 1.23 s / 0.26 s).

In order to distinguish between motion effects in general and interplay effects in particular, 4DD distributions were computed in addition to the 4DDD distributions for PBS plans. To this end, the original treatment plan was recomputed on all respiratory phases with a weighting factor of one over the number of phases. Subsequently, the calculated dose distributions were mapped to the planning CT using DIRs and summed up.

Double scattering

In DS, the 4DDD was approximated by the 4DD. Since the frequency of IBA's range modulator wheel is much higher than the typical human respiratory rate (600 revolutions per minute versus 10-20 breaths per minute), the exact time structure of the beam delivery could be neglected assuming a uniform distribution of the dose over all motion phases. The computation procedure was the same as for PBS 4DD computations described above.

4.2.6 Evaluation methods and criteria

The outcome of a single treatment fraction was evaluated for 60 simulations per PBS plan with varying start phases and energy layer switching times. The number of scenarios could not be further increased in RayStation 6 because of the long computation times and the high memory consumption which caused data storage problems. The outcome of a whole treatment course consisting of 15 fractions was modelled five times, yielding 75 simulations per plan. Due to the high memory load, the evaluation doses had to be deleted after each run, i.e. after 15 fractions, in order to release memory. The computation time depended on the load of the server and

the number of energy layers of a patient. It took about six to ten hours to complete 60 single fraction evaluations on a server equipped with two Intel Xeon E5 2680 v2 CPUs with a total of 20 cores and 128 GB RAM (Intel Corporation, Santa Clara, USA) when the dose volume histogram plot option was selected in addition to the default analysis of specific dose volume histogram metrics.

The standard deviation (SD) was used as a measure for the variability of interplay effects for a treatment plan due to different start phases and energy layer switching times which can occur during irradiation. The precision of the mean value was exemplarily analysed for one 4D PBS and one 3D PBS plan of patient 2 based on the standard error of the mean (SEM) [143]:

$$SEM = \frac{SD}{\sqrt{n}}, \quad (4.3)$$

where n denotes the number of samples. For this purpose, the single fraction interplay evaluation including 60 scenarios was repeated five times per plan.

4DD distributions only had to be calculated once per plan since they do not depend on the start phase or energy layer switching times.

Based on the results of the 4DDD PBS single fraction evaluations and the 4DD DS evaluations the average value and SD of each technique was determined for the plan comparison. Since there was only one patient cohort, it was not possible to check the reproducibility by means of the SEM.

Evaluation criteria for the CTV coverage

Because of their sensitivity to interplay effects the homogeneity index HI

$$HI = \frac{D_5 - D_{95}}{D_{presc}} \quad (4.4)$$

and the percentage over- and underdosage $V_{107/95}$

$$V_{107/95} = V_{107} + (100 - V_{95}) \quad (4.5)$$

were chosen as evaluation criteria for the plan comparison, where D_v is the dose delivered to $v\%$ of the volume. The statistical significance of differences in HI and $V_{107/95}$ between 4D PBS and 3D PBS plans and between 4D PBS and DS, respectively, was tested using a paired t-test (two-tailed) [144]. The significance level was set to 0.05.

Evaluation criteria for OAR sparing

Representative for close-by OARs, dose volume histogram metrics of the normal liver tissue and ribs were analysed. The risk of RILD was calculated based on the empiric Lyman-Kutcher-Burman (LKB) normal tissue complication probability (NTCP) model [145, 146] using the scripting interface of RayStation. Following the example of Luxton et al. [147], the NTCP was expressed as a function of the gEUD:

$$NTCP = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-x^2/2} dx \quad (4.6)$$

$$u = \frac{gEUD - TD_{50}}{m \cdot TD_{50}}, \quad (4.7)$$

where TD_{50} denotes the tolerance dose leading to a complication probability of 50% for a uniform irradiation of the entire organ and m describes the steepness of the NTCP curve. The used LKB model parameters for RILD, $TD_{50}=39.8 \text{ Gy}_{\text{RBE}}$, $m=0.12$ and $n = 0.97$, were taken from Dawson et al. [134]. The endpoint was defined as Radiation Therapy Oncology Group Grade 3 or higher RILD in the study. Differences in the fractionation scheme compared

to Dawson et al. (fraction dose 1.5 Gy) were considered by a correction of the dose volume histogram according to the LQ model (Sec. 2.2.1) when calculating the gEUD (see Eq. 2.10):

$$gEUD = \left(\sum_i v_i LQED_i^{1/n} \right)^n, \quad (4.8)$$

where $LQED$ is the LQ-based equivalent dose in d_{fx} -Gy fractions for sub-volume (dose volume histogram bin) i [148]:

$$LQED_i = D_i \frac{1 + \frac{D_i/N}{\alpha/\beta}}{1 + \frac{d_{fx}}{\alpha/\beta}}. \quad (4.9)$$

The absorbed dose D_i , corresponding to the i th dose volume histogram bin, was replaced by the biological dose $(D_{RBE})_i$. The total number of fractions N was set to 15. An α to β ratio of 2.5 was assumed for the liver [149]. The reference fraction size d_{fx} was 1.5 Gy.

In addition to the risk of RILD, the mean dose, $V_{42Gy_{RBE}}$, $V_{33Gy_{RBE}}$ and the volume receiving less than 15 Gy_{RBE} were calculated for the normal liver tissue.

The risk for radiation-induced rib fracture after hypofractionated proton beam therapy was determined based on the dose volume histogram metric $V_{60Gy_{RBE}}(rib)$ [150]. Only the three most exposed ribs per patient with $V_{60Gy_{RBE}}(rib) > 0$ were included in the evaluation. Doses to ribs ranked fourth or higher regarding $V_{60Gy_{RBE}}(rib)$ were not considered to be clinically relevant.

The alimentary tract, the heart and the kidney, which were located adjacent to or even within the iCTV for some of the patients, were excluded from the evaluation since sufficient CTV coverage in the static plan was of primary importance to examine motion-induced dose distortions. Compromises in the CTV coverage to spare these OARs would have significantly impacted the magnitude of motion effects and thus the outcome of the study.

The following limits were used as evaluation criteria for the OARs: $V_{42Gy_{RBE}}(normal\ liver) = 33\%$, $V_{33Gy_{RBE}}(normal\ liver) = 67\%$, average normal liver dose = 24 Gy_{RBE} and $V_{60Gy_{RBE}}(rib) = 4.5\text{ cm}^3$ [149, 150, 151].

4.2.7 Supplementary analyses

Two points which might have affected the outcome of the 4D dose comparison are the accuracy of the applied dose engine and the selected number of irradiation fields. These aspects were addressed in supplementary analyses. Details on the investigations are provided in the following.

Monte Carlo calculations

Although the Monte Carlo dose engine is known to be more accurate, the Pencil Beam dose algorithm was explicitly chosen for the treatment planning study since RayStation does currently not support Monte Carlo dose calculations for DS. The Pencil Beam dose algorithm is particularly prone to errors if a treatment plan includes range shifters or a beam traverses lung tissue (Sec. 2.4.5). Therefore, static PBS treatment plans and 4DDD evaluations were reviewed with MC on a sample basis. For this purpose, two targets, which required the use of range shifters and were located adjacent to the bowel (patient ID 2) and the lung (patient ID 3), respectively, were selected. The original 4D PBS plans were recomputed with MC, i.e. the spot positions and weights remained unchanged. No automatic rescaling was applied, if the resulting median dose of the CTV (D_{50}) deviated from the prescription.

For the review of the 4DDD evaluation, new plans were created using the same objectives but MC-based optimisation. It was decided not to use the MC-recomputed plans from the static dose comparison in this case since they could already exhibit cold or hot spots in the nominal scenario which would affect the interplay effect analysis. Thus, the same starting conditions, namely homogeneous target coverage and optimal OAR sparing, were ensured. The MC-optimised 4D PBS plans were evaluated based on 60 4DDD simulations as the original

plans before. Hereinafter, the original plans which were optimised with the Pencil Beam algorithm are referred to as PB plans.

The final dose for both, MC-recomputed and MC-optimised plans, was calculated to 0.5% statistical uncertainty.

Three-field PBS plans

A different number of fields was chosen for PBS and DS due to the differences in dose conformality. In PBS, an adequate level of conformality could be achieved using two fields. Since a third field prolongs the treatment time and increases the amount of normal tissue exposed to irradiation, it would usually not be a clinical option. However, a third field increases the robustness against motion as well as setup and range uncertainties and might have been an advantage of the DS plans over the PBS plans regarding the 4D dose comparison. Therefore, the impact of a third irradiation field was investigated for the test patients used in the MC review (see above). Copies of the original 4D PBS plans were created. The settings remained unchanged, except for the number of fields and the gantry and couch angles, which were taken from the corresponding DS plans. Only if the plans did not fulfil the clinical goal $V_{95}(CTV) = 100\%$, the weights of the optimisation objectives were slightly adjusted. Like the original 4D PBS plans, the three-field 4D PBS plans had to pass the robustness evaluation (Sec. 4.2.4). In the end, they were evaluated based on 60 4DDD simulations and compared to the original PBS and DS plans.

4.3 Results

4.3.1 Evaluation of the deformable image registration

The examined landmarks in the liver revealed an average three-dimensional target registration error of $0.20 \text{ cm} \pm 0.12 \text{ cm}$ (SD) for the maximum deformation (50% phase to 0% phase). The superior-inferior direction, the main direction of motion, contributed most to the error. Smaller registration errors were observed when other phases were included in the evaluation. A test based on two patients with 10 landmarks mapped from the 50% phase to the 20%, 30%, 80% and 90% phase yielded for example an average target registration error of $0.16 \text{ cm} \pm 0.12 \text{ cm}$. The average Dice similarity coefficient of the liver was 0.96 ± 0.03 considering deformations between the 50% and 0% phase.

4.3.2 Perturbation analysis of the nominal treatment plans

With one exception, the treatment plans passed the robustness analysis with $V_{95}(CTV)$ values of at least 99% and 98%, respectively for very small tumours, in all perturbation scenarios. Only the 3D PBS plan for CTV 9b, which is located adjacent to the lung, did not achieve sufficient plan robustness. For four out of eight perturbation scenarios, $V_{95}(CTV)$ was below 98% with a minimum of 83.2%. One of the examined perturbation scenarios is shown by way of example in Figure 4.2 for DS, 4D PBS and 3D PBS.

4.3.3 Analysis of the 4D dynamically accumulated dose distribution

For purpose of illustration, three examples of 4D dose distributions are depicted in Figure 4.3, each representative of one treatment technique. The CT slice in subplot (a) shows no over- or underdosage of the CTV for the selected DS plan. This fact holds true for all other slices and CTVs as will be shown below. In contrast, the corresponding 4D and 3D PBS plans suffer from cold and hot spots, displayed in yellow (90% - 95% of D_{presc}), dark red (107% - 110% of D_{presc}) and pink (>110% of D_{presc}). Another detail which can be seen on the images is that 4D PBS is superior to DS and 3D PBS in sparing OARs as ribs or the normal liver tissue.

Figure 4.4 demonstrates differences in the manifestation of interplay effects based on two dose volume histograms using 4D PBS plans as example. The first dose volume histogram belongs to a large CTV with a small motion amplitude (ID 7) and the second dose volume histogram

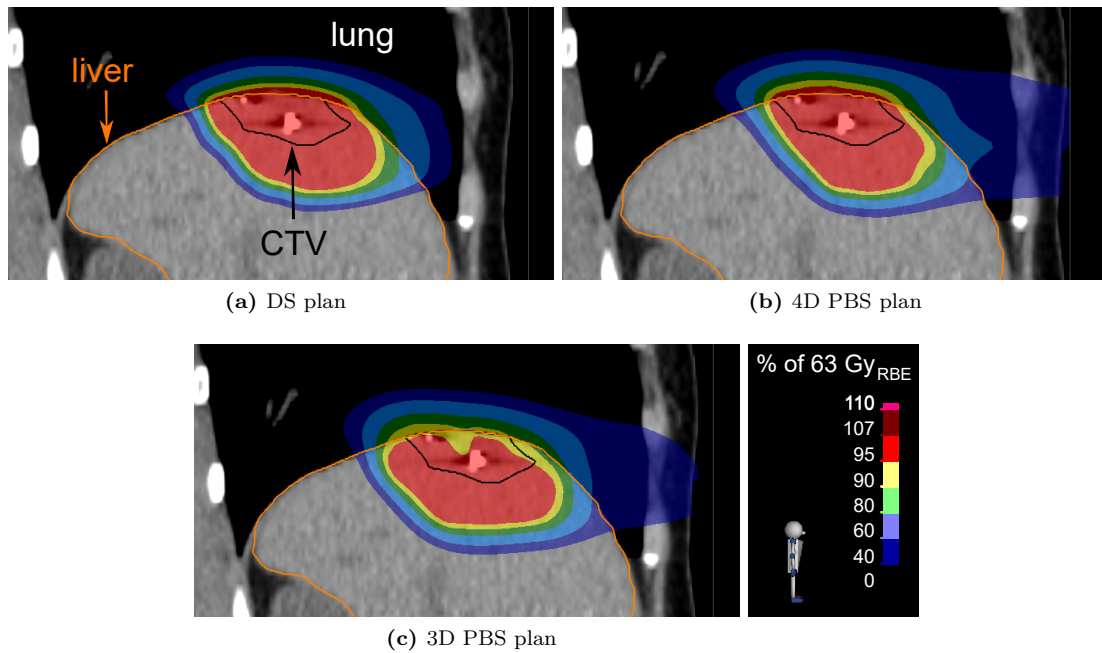


Figure 4.2: Exemplary illustration of a perturbation scenario (sagittal view) for DS, 4D PBS and 3D PBS: The perturbed dose (overlaid as colour wash) is shown for CTV 9b on CT50% for a density perturbation of 5% and an isocentre shift of $dx = dy = dz = -2$ mm. In this scenario, $V_{95}(CTV)$ was 100.0% for DS, 99.3% for 4D PBS and 86.6% for 3D PBS.

to a small CTV with a large motion amplitude (ID 9a). Both dose volume histograms include the static case, the 4DDD evaluations of the whole treatment course (15 fractions) and 60 4DDD single fraction evaluations forming a dose volume histogram band. The small tumour with the large motion amplitude exhibits a much broader CTV dose volume histogram band, an indication for a low robustness level. In addition, the difference in the normal liver dose between the static and the dynamic evaluation is slightly larger for this tumour. Quantitative measures of CTV coverage and OAR sparing are given in the following sections.

CTV coverage

The HI of the CTV was lowest for DS plans looking at the 4D outcome of a single treatment fraction (see Fig. 4.5a). While an average value of $2.0\% \pm 0.9\%$ (SD) was observed for DS, 4D PBS plans exhibited an average HI of $8.1\% \pm 1.3\%$. 3D PBS plans were slightly behind with $8.4\% \pm 1.3\%$. The detected differences in HI between 4D PBS and DS were found to be significant ($P < 0.001$). 4D PBS and 3D PBS, by contrast, did not differ significantly in HI ($P = 0.65$).

For PBS plans, the HI reached up to 7.9% to 18.4% in the worst case scenario of 60 interplay evaluations depending on the patient/ the CTV. On average, 4D robust optimisation led to slightly lower HI values than 3D non-robust optimisation for homogeneity category II and III (see Fig. 4.5a). For category I (patient ID 4 and 6), 3D PBS plans yielded smaller average values.

The HI decreased to below 5% for 4D and 3D PBS simulating a whole treatment course consisting of 15 fractions (Fig. 4.5b). Thus, the differences between DS and PBS plans became much smaller and in three cases the lowest HI was even achieved by a PBS plan. The HI values for 15 fractions almost approached the static plan results computed on the reference CT, which ranged from 0.6% to 3.1% considering all treatment techniques.

$V_{107/95}$ was zero for all 4DDD single fraction evaluations in DS, i.e. no over- or underdosage could be detected. Unlike DS plans, 4D PBS plans exhibited on average a percentage over- and

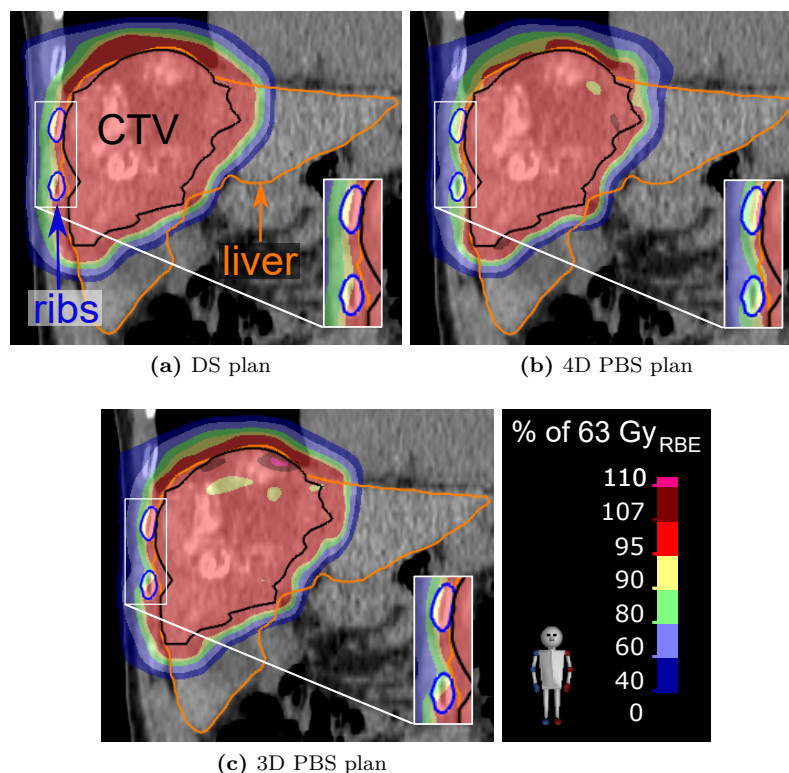


Figure 4.3: Comparison of the 4D dynamically accumulated dose for DS, 4D PBS and 3D PBS considering a single fraction (coronal view): The example shows CTV 3. An enlarged image section illustrating the exposure of close-by ribs is depicted on the bottom right of each screenshot.

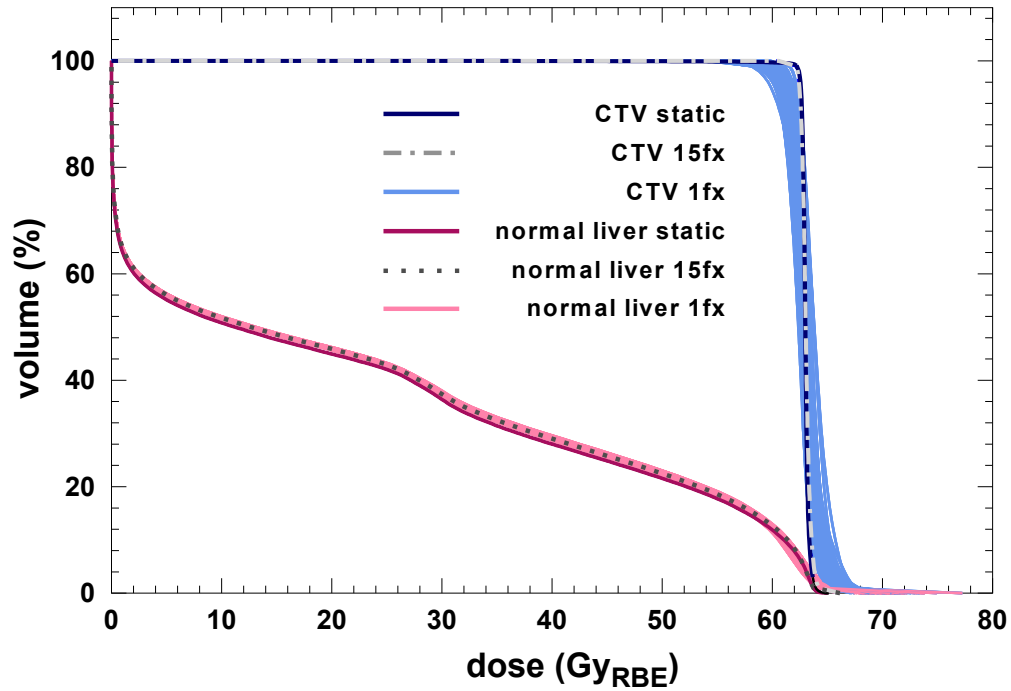
underdosage of $5.9\% \pm 3.4\%$ in a single fraction, closely followed by 3D PBS plans with $6.9\% \pm 5.7\%$. The difference between 4D and 3D PBS plans was not significant ($P=0.60$).

Again, 4D robust optimised plans yielded on average smaller values for homogeneity category II and III than 3D non-robustly optimised plans but performed worse for category I as can be seen in Figure 4.6. In the worst case scenario, $V_{107/95}$ ranged from 5.2% to 26.8% for patient 1 to 8 in 4D and 3D PBS plans. The maximum was even in-between 51.0% and 84.3% in 4D and 3D PBS plans for patient 9 (CTV a and b). When averaging over 15 fractions, however, $V_{107/95}$ approached zero for both, 3D and 4D PBS plans. None of the static plans showed any over- or underdosage.

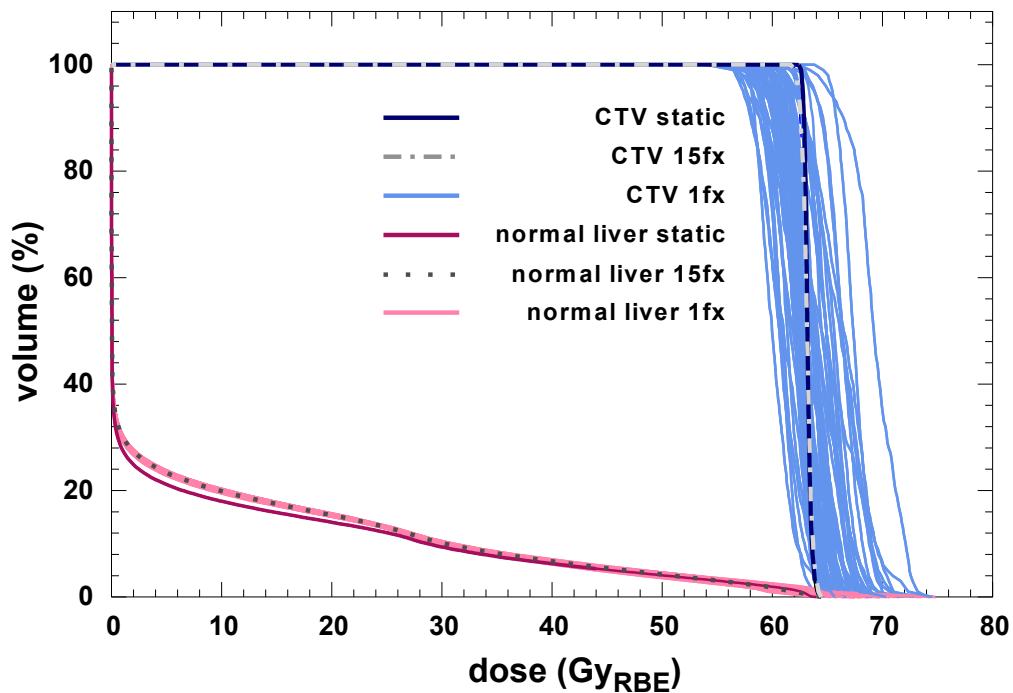
The minimum and the maximum dose per fraction, $D_{99.99}(CTV)$ and $D_{0.01}(CTV)$, remained in-between 95% and 107% of D_{presc} for all DS plans (as is evident from the $V_{107/95}$ analysis). For PBS plans, the average maximum dose was below 114% of D_{presc} considering 60 scenarios per patient. The worst case scenario exhibited dose maxima from 111.6% to 121.8%. The minimum dose per fraction was on average above 90% of D_{presc} for a single fraction in 4D and 3D PBS. But, depending on the patient, $D_{99.99}(CTV)$ could be as low as 77.4% to 91.3% in the worst case.

The impact of tumour motion, CTV volume and tissue homogeneity on interplay effects in PBS is illustrated in Figure 4.7 based on $V_{107/95}$. In both, 4D PBS (subplot (a)) and 3D PBS (subplot (b)), CTV 9a and 9b reached by far the highest $V_{107/95}$ values. The CTVs exhibited comparatively large motion amplitudes, had volumes of less than 15 cm^3 and belonged to the homogeneity categories II and III. Very large CTVs with small motion amplitudes, as ID 3 and 7, led to relatively small $V_{107/95}$ values independent of the tissue homogeneity for 3D and 4D PBS plans. 3D PBS plans performed better than 4D PBS plans, for tumours in homogeneous regions (patient 4 and 6).

The 4DD evaluation of PBS plans revealed that there is no over- or underdosage of the CTV



(a) DVH for patient ID 7



(b) DVH for patient ID 9a

Figure 4.4: Static and 4D dynamically accumulated dose volume histograms (DVHs) of the CTV and the normal liver tissue (liver minus CTV) representative for a large CTV with a small motion amplitude (a) and a small CTV with a large motion amplitude (b) for 4D PBS plans: DVH (a) corresponds to patient ID 7 with a CTV volume of about 640 cm^3 and a motion amplitude of 0.37 cm . DVH (b) corresponds to patient ID 9a with a CTV volume of 6 cm^3 and a motion amplitude of 1.45 cm . The 4DDD outcome of a single fraction (1 fx) was simulated 60 times with varying start phases and energy layer switching times. In addition to the single fraction 4DDD evaluations, the static case and a 4DDD evaluation of the whole treatment course (15 fx) are included in the plot.

if the beam dynamics are neglected, i.e. $V_{107/95}$ is zero. Similar HI values as for the static case were observed ranging from 0.8% to 3.9%. The 4DD distributions showed no significant difference between 4D and 3D PBS plans regarding the HI . In six out of eleven cases, the HI was lower for 3D than for 4D PBS plans.

OAR sparing

The average dose to the normal liver tissue was smallest for 4D PBS plans with $11.1 \text{ Gy}_{\text{RBE}} \pm 6.8 \text{ Gy}_{\text{RBE}}$ (see Fig. 4.8a). 3D PBS plans deposited on average $5.7\% \pm 6.0\%$ more dose than 4D PBS plans and DS even $15.9\% \pm 10.3\%$. In all cases, 4D PBS plans performed either better or at least equivalent to 3D PBS and DS plans in sparing normal liver tissue. The limit of $24 \text{ Gy}_{\text{RBE}}$ for the average normal liver dose was satisfied by all plans except for one DS plan which exceeded the limit by $2.5 \text{ Gy}_{\text{RBE}}$ (patient ID 7).

The percentage normal liver volume which received $42 \text{ Gy}_{\text{RBE}}$ or more is depicted for DS, 4D PBS and 3D PBS in Figure 4.8b. The plot shows that 4D PBS plans were slightly superior to 3D PBS and DS plans with an average $V_{42\text{Gy}_{\text{RBE}}}$ value of $12.7\% \pm 8.4\%$ compared to $15.5\% \pm 10.2\%$ (DS) and $13.8\% \pm 8.9\%$ (3D PBS). All plans remained below the 33% limit, but two of the DS plans (patient ID 3 and 7) almost reached it.

The maximum observed $V_{33\text{Gy}_{\text{RBE}}}$ value for DS, 4D PBS and 3D PBS was 37.0%, which is far below the 67% limit.

The normal liver volume receiving less than $15 \text{ Gy}_{\text{RBE}}$ is shown in Figure 4.9. In five out of eleven cases (patient ID 5, 6, 8b, 9a and 9b), there was almost no difference in-between DS, 4D PBS and 3D PBS. For the rest, the volume receiving less than $15 \text{ Gy}_{\text{RBE}}$ was smallest for DS (minimum: 450 cm^3). 4D and 3D PBS performed equally well.

The calculated risk of RILD was zero for most of the patients. There were only two exceptions. Patient 3 exhibited NTCP values of 19.2% (DS), 1.2% (4D PBS) and 3.4% (3D PBS) and patient 7 of 30.8% (DS), 6.7% (4D PBS) and 7.0% (3D PBS).

For the examined ribs, the volume receiving $60 \text{ Gy}_{\text{RBE}}$ or more was on average $2.3 \text{ cm}^3 \pm 2.2 \text{ cm}^3$ for DS, $0.3 \text{ cm}^3 \pm 0.5 \text{ cm}^3$ for 4D PBS and $1.2 \text{ cm}^3 \pm 1.4 \text{ cm}^3$ for 3D PBS. In two cases, DS led to irradiated volumes larger than the volume constraint of 4.5 cm^3 . The affected ribs belonged to patient 3 and had a $V_{60\text{Gy}_{\text{RBE}}}$ value of 7.3 cm^3 and 4.8 cm^3 , respectively (see Tab. 4.1). For 4D PBS plans, the maximum average value of $V_{60\text{Gy}_{\text{RBE}}}$ was 1.4 cm^3 and for 3D PBS plans 4.0 cm^3 . The differences in the exposure of close-by ribs between DS, 4D PBS and 3D PBS plans are exemplarily illustrated in enlarged image sections in Figure 4.3.

Reproducibility of the results

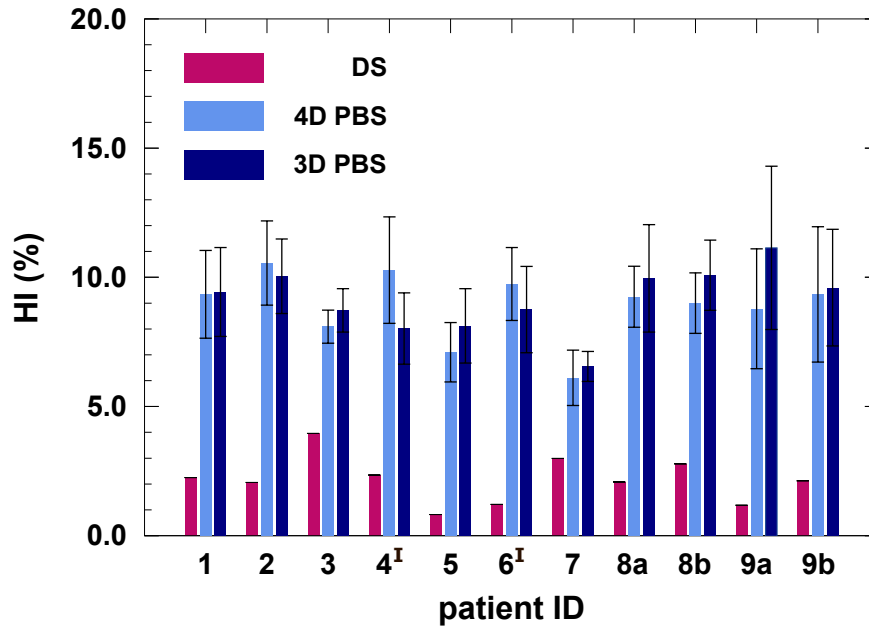
It was tested, exemplarily for one patient, whether the interplay effect analysis yielded reproducible mean values for the chosen number of interplay scenarios (60 single fraction evaluations). Five repetitions with each 60 scenarios revealed an SEM of the HI of 0.1% for both, the 4D and the 3D PBS plan (average values: 10.4% (4D PBS) and 9.6% (3D PBS)). For $V_{107/95}$, the SEM was 0.3% for 4D and 3D PBS (average values: 7.6% (4D PBS) and 5.6% (3D PBS)). A two-tailed, paired t-test demonstrated that the differences in HI and $V_{107/95}$ between the 4D and the 3D PBS plan were significant (HI : $P < 0.001$, $V_{107/95}$: $P < 0.001$).

The average dose of the normal liver tissue exhibited an SEM of only $0.4 \text{ cGy}_{\text{RBE}}$ for both plans (average values: $1377.7 \text{ cGy}_{\text{RBE}}$ (4D PBS) and $1565.5 \text{ cGy}_{\text{RBE}}$ (3D PBS), $P < 0.001$). For $V_{42\text{Gy}_{\text{RBE}}}$, the SEM was even zero. In the case of $V_{<15\text{Gy}_{\text{RBE}}}$, the 4D and the 3D PBS plan had an SEM of 0.1 cm^3 (average values: 795.9 cm^3 (4D PBS) and 766.1 cm^3 (3D PBS), $P < 0.001$).

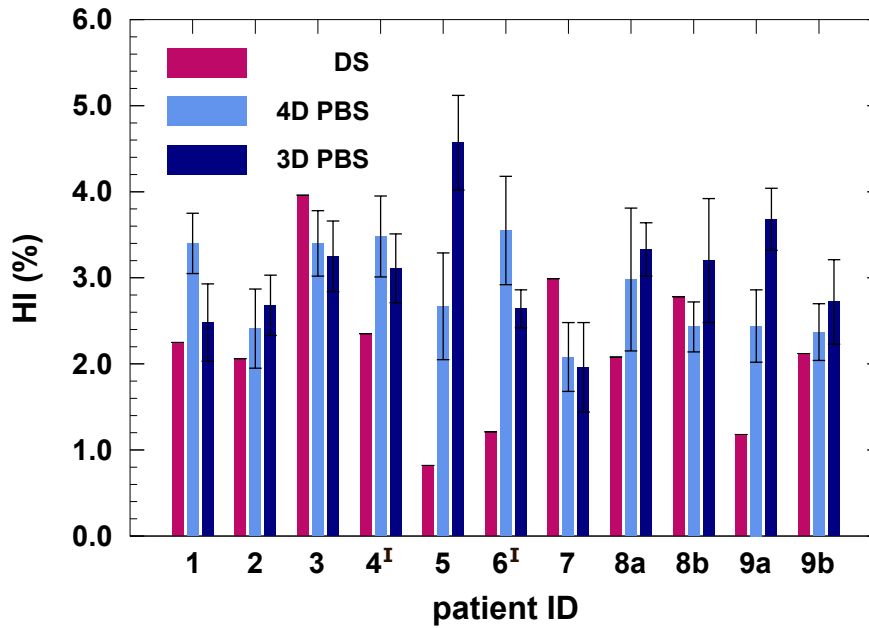
4.3.4 Supplementary analyses

Monte Carlo

For patient 2, the static dose distribution looks very similar for computations with the Pencil Beam algorithm and the MC dose engine as can be seen in Figure 4.10. The dose difference



(a) 1 fraction



(b) 15 fractions

Figure 4.5: Homogeneity index (HI) of the CTV for 4D dynamically accumulated dose distributions of a single treatment fraction (a) and a whole treatment course including 15 fractions (b): Each bar displays the average HI for a certain patient and treatment plan. A comparison is drawn between DS, 4D PBS and 3D PBS. In case of PBS, error bars indicate the standard deviation among 60 single fraction interplay simulations (subplot (a)) and 5 treatment course interplay simulations (subplot (b)), respectively. In DS, only one 4DDD scenario exist per treatment plan since the dose is evenly distributed over all respiratory phases. The patient ID is marked by an 'I' if the iCTV belongs to the homogeneity category I, i.e. it is completely encompassed by liver tissue.

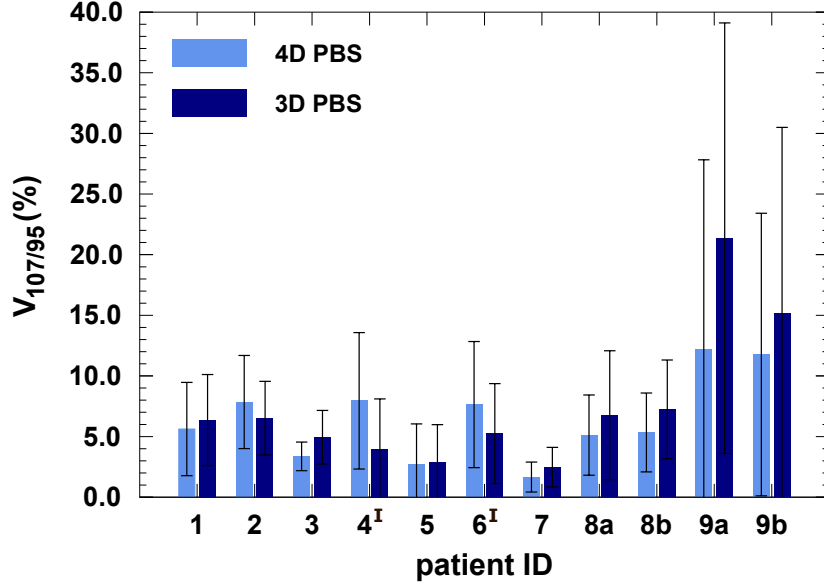
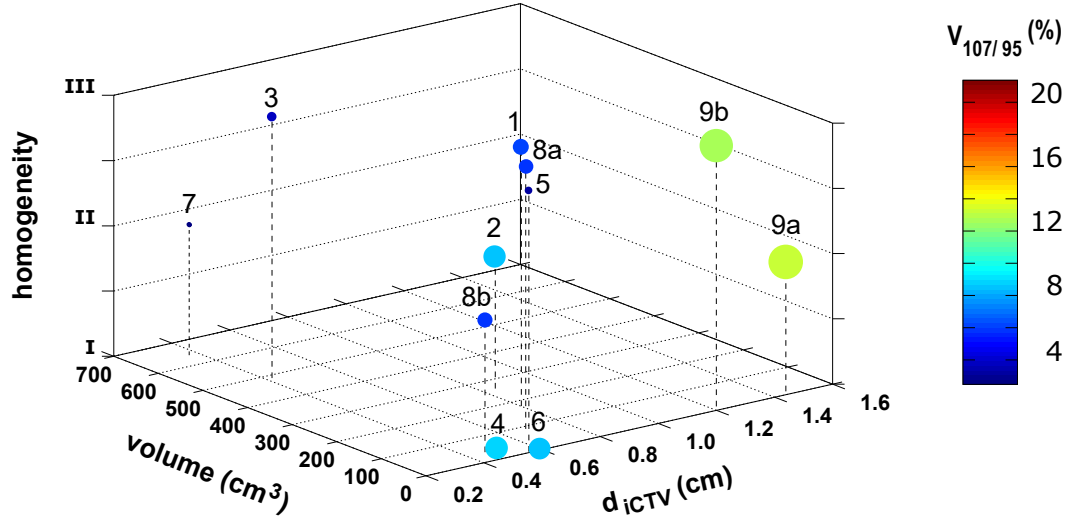


Figure 4.6: Percentage over- and underdosage ($V_{107/95}$) of the CTV for 4D dynamically accumulated dose distributions of a single fraction: The average $V_{107/95}$ is plotted for 4D PBS and 3D PBS plans. Error bars represent the standard deviation among 60 interplay simulations. The patient ID is marked by an 'I' if the iCTV belongs to the homogeneity category I, i.e. it is completely encompassed by liver tissue. Since the 4D dynamically accumulated dose distributions for DS and 15 PBS fractions did not exhibit any over- or underdosage, they are not included in the plot.

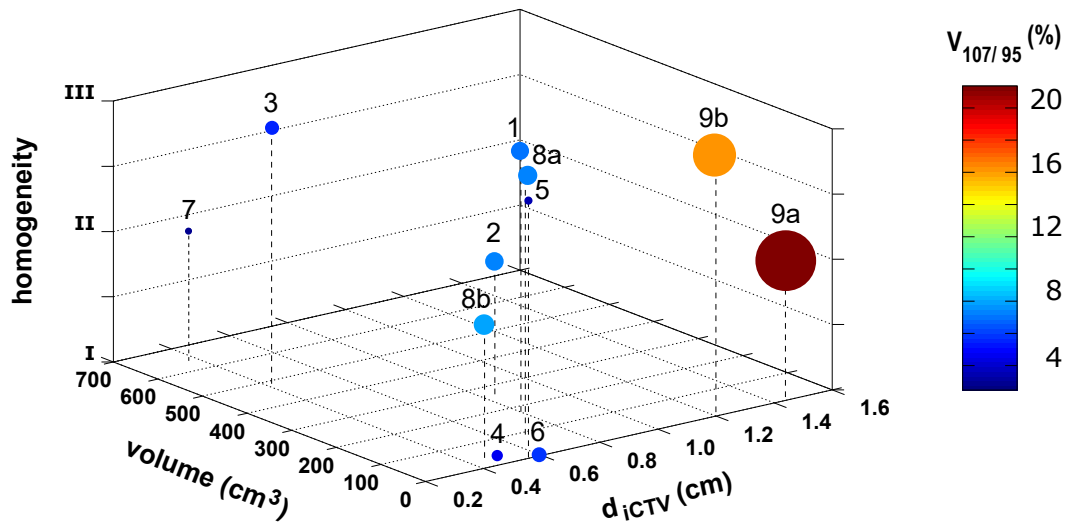
Table 4.1: Comparison of the rib volume receiving 60 Gy_{RBE} or more ($V_{60\text{Gy}_{\text{RBE}}}(\text{rib})$) for 4D dynamically accumulated dose (4DDD) evaluations of DS, 4D PBS and 3D PBS plans: The three most exposed ribs per patient with $V_{60\text{Gy}_{\text{RBE}}} > 0 \text{ cm}^3$ were considered for the evaluation. For PBS, the standard deviation among 60 interplay simulations with varying starting phases and energy layer switching times is stated. In DS, only one 4DDD scenario exist per treatment plan since the dose is evenly distributed over all respiratory phases.

| | $V_{60\text{Gy}_{\text{RBE}}}(\text{rib})$ in cm^3 | | | | | |
|--------|---|---------------|---------------|---------------|---------------|---------------|
| | ID 1 | | | ID 3 | | |
| | rib a | rib b | rib c | rib a | rib b | rib c |
| DS | 3.8 | 1.7 | 0.2 | 7.3 | 4.8 | 1.2 |
| 4D PBS | 1.3 ± 0.3 | 0.2 ± 0.1 | 0.0 ± 0.0 | 1.4 ± 0.2 | 0.4 ± 0.1 | 0.0 ± 0.0 |
| 3D PBS | 4.0 ± 0.3 | 1.5 ± 0.3 | 0.4 ± 0.1 | 3.7 ± 0.4 | 2.2 ± 0.4 | 0.0 ± 0.0 |

| | $V_{60\text{Gy}_{\text{RBE}}}(\text{rib})$ in cm^3 | | | | |
|--------|---|---------------|---------------|---------------|---------------|
| | ID 5 | ID 7 | | | ID 8a |
| | rib a | rib a | rib b | rib c | rib a |
| DS | 0.8 | 2.3 | 1.7 | 0.1 | 3.4 |
| 4D PBS | 0.1 ± 0.1 | 0.2 ± 0.0 | 0.0 ± 0.0 | 0.0 ± 0.0 | 0.4 ± 0.1 |
| 3D PBS | 0.7 ± 0.2 | 0.4 ± 0.1 | 0.4 ± 0.0 | 0.1 ± 0.3 | 1.1 ± 0.4 |

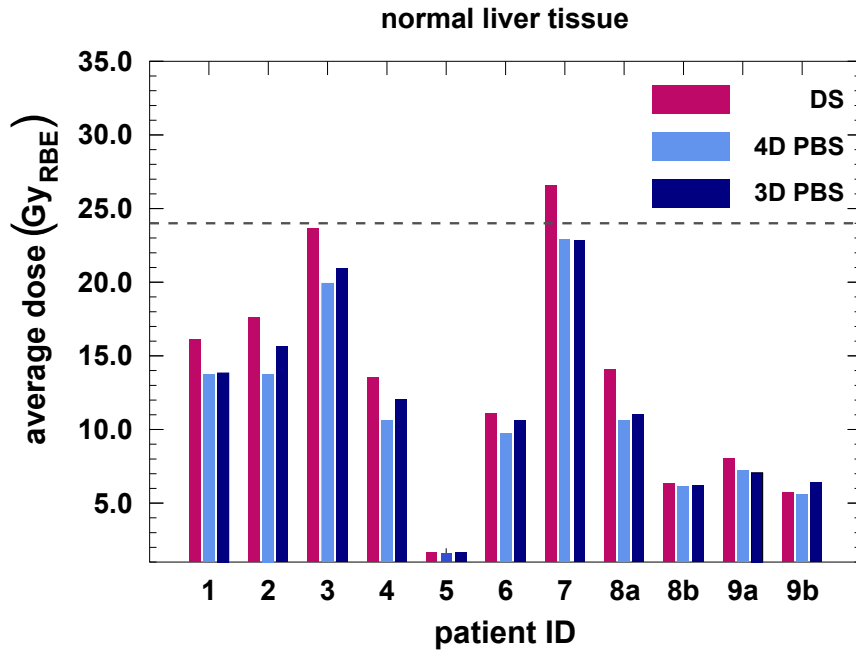


(a) 4D PBS plan

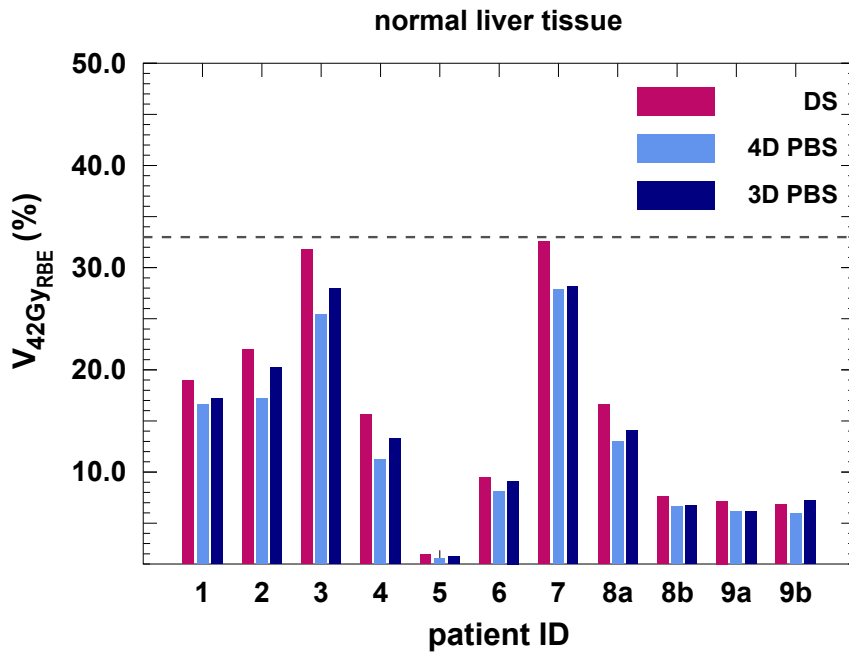


(b) 3D PBS plan

Figure 4.7: Percentage over- and underdosage ($V_{107/95}$) as a function of the tumour motion amplitude d_{iCTV} (see Eq. 4.1), the CTV volume and the tissue homogeneity for 4D PBS (a) and 3D PBS plans (b): Homogeneity category I encompasses targets which are surrounded by homogeneous liver tissue, category II includes targets which are adjacent to other abdominal organs and category III comprises targets next to the lung or ribs. The size and the colour code of the dots represent the magnitude of $V_{107/95}$. The corresponding patient ID is given by a number above each spot.



(a) average dose



(b) $V_{42Gy_{RBE}}$

Figure 4.8: Average normal liver dose (a) and percentage normal liver volume receiving 42Gy_{RBE} or more ($V_{42Gy_{RBE}}$) (b) for 4D dynamically accumulated dose evaluations: Each bar represents the average value for a certain patient ID and treatment plan (DS, 4D PBS or 3D PBS). A dashed line indicates the dose/ dose volume histogram constraint in subplot (a) and (b). The standard deviation between different interplay scenarios in PBS was not plotted since it was less than 0.2 Gy_{RBE} (average dose) and 0.4% ($V_{42Gy_{RBE}}$), respectively, for all IDs. In DS, only one 4DDD scenario exist per treatment plan since the dose is evenly distributed over all respiratory phases.

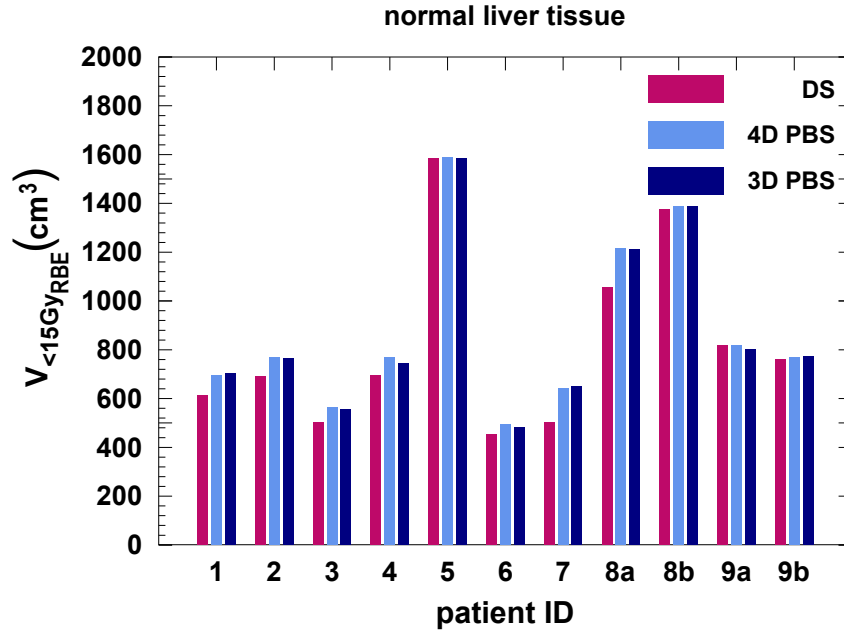


Figure 4.9: Normal liver volume receiving less than 15 Gy_{RBE} for 4D dynamically accumulated dose evaluations of DS, 4D PBS and 3D PBS plans: The standard deviation between the different interplay scenarios was below 2 cm³ for all PBS plans. In DS, only one 4DDD scenario exist per treatment plan since the dose is evenly distributed over all respiratory phases.

map in subplot (c) shows that the plans deviate by less than 2% for the majority of the CTV volume. There are some regions around the CTV which exhibit 2% to 4% more dose in the original PB plan than in the MC-recomputed plan. The effect on the dose statistics was small with regard to the clinical goals. $V_{95}(CTV)$ was for example for both plans 100%, the average normal liver dose decreased from 15.0 Gy_{RBE} to 14.8 Gy_{RBE} and $V_{42}(normal\ liver)$ from 18.9% to 18.7%.

The differences in the static dose distribution between the PB and the MC-recomputed plan were significantly larger for patient 3, who had a tumour adjacent to the lung (see Fig. 4.11). The MC computation revealed a slight underdosage of the CTV (90% - 95% of D_{presc}) for small regions located close to the lung or behind the peritoneal cavity. In addition, the dose to the heart was locally increased by up to 13% for the MC-recomputed plan. Differences could also be seen in the dose statistics. The target coverage ($V_{95}(CTV)$) decreased from 100% to 98.1%, the average normal liver dose from 20.1 Gy_{RBE} to 19.6 Gy_{RBE}, $V_{42}(normal\ liver)$ from 25.4% to 24.7% and $V_{60GyRBE}$ of the most exposed rib from 1.2 cm³ to 0.3 cm³.

The results of the 4DDD evaluation of MC- and PB-optimised 4D PBS plans are summarised in Table 4.2. For both patients, the magnitude of interplay effects measured by $HI(CTV)$ and $V_{107/95}(CTV)$ was higher for MC computations. The average HI deviated by about 5% (ID 2) and 13% (ID 3). The difference between the average values of $V_{107/95}$ was 15% (ID 2) and 46% (ID 3), respectively. Besides the average value, also the standard deviation was slightly increased.

PB plans exhibited a lower normal liver dose in 4DDD computations than MC plans. For patient 2, the average dose, $V_{42GyRBE}$ and $V_{33GyRBE}$ were approximately 10% below the MC values. The normal liver volume receiving less than 15 Gy_{RBE} was about 3.5% higher for the PB plan. For patient 3, the average dose, $V_{42GyRBE}$ and $V_{33GyRBE}$ were 1.2% to 3.7% smaller and $V_{<15GyRBE}$ was 2.9% higher for PB than for MC.

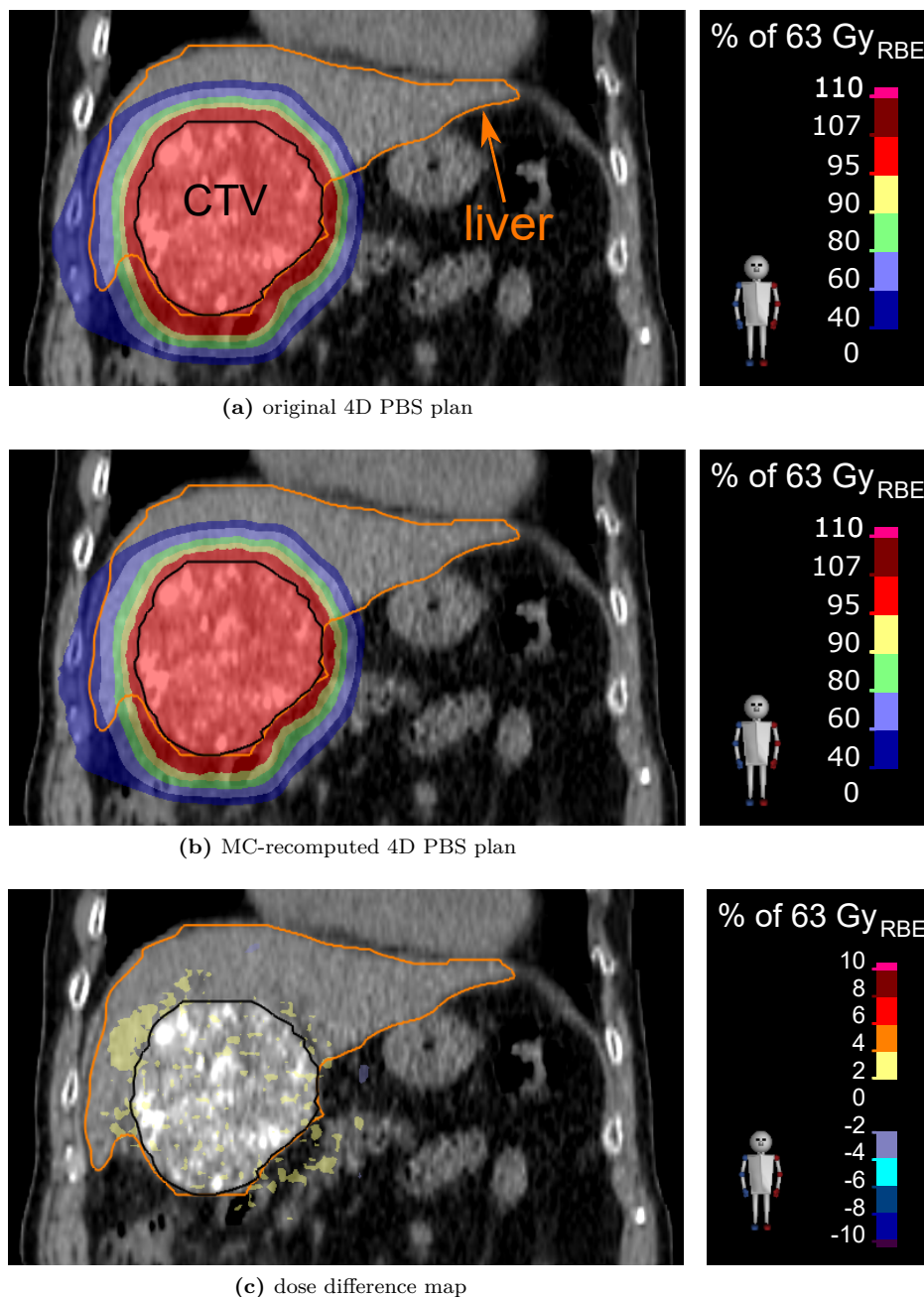


Figure 4.10: Comparison of the static dose distribution of the original 4D PBS treatment plan calculated with the Pencil Beam algorithm (a) and a copy of the treatment plan which was recomputed with the Monte Carlo (MC) dose engine (b) for patient 2: The differences in the dose distributions are displayed as colour wash (c). The colour scale of the dose difference map is the same as for patient 3 (Figure 4.11c).

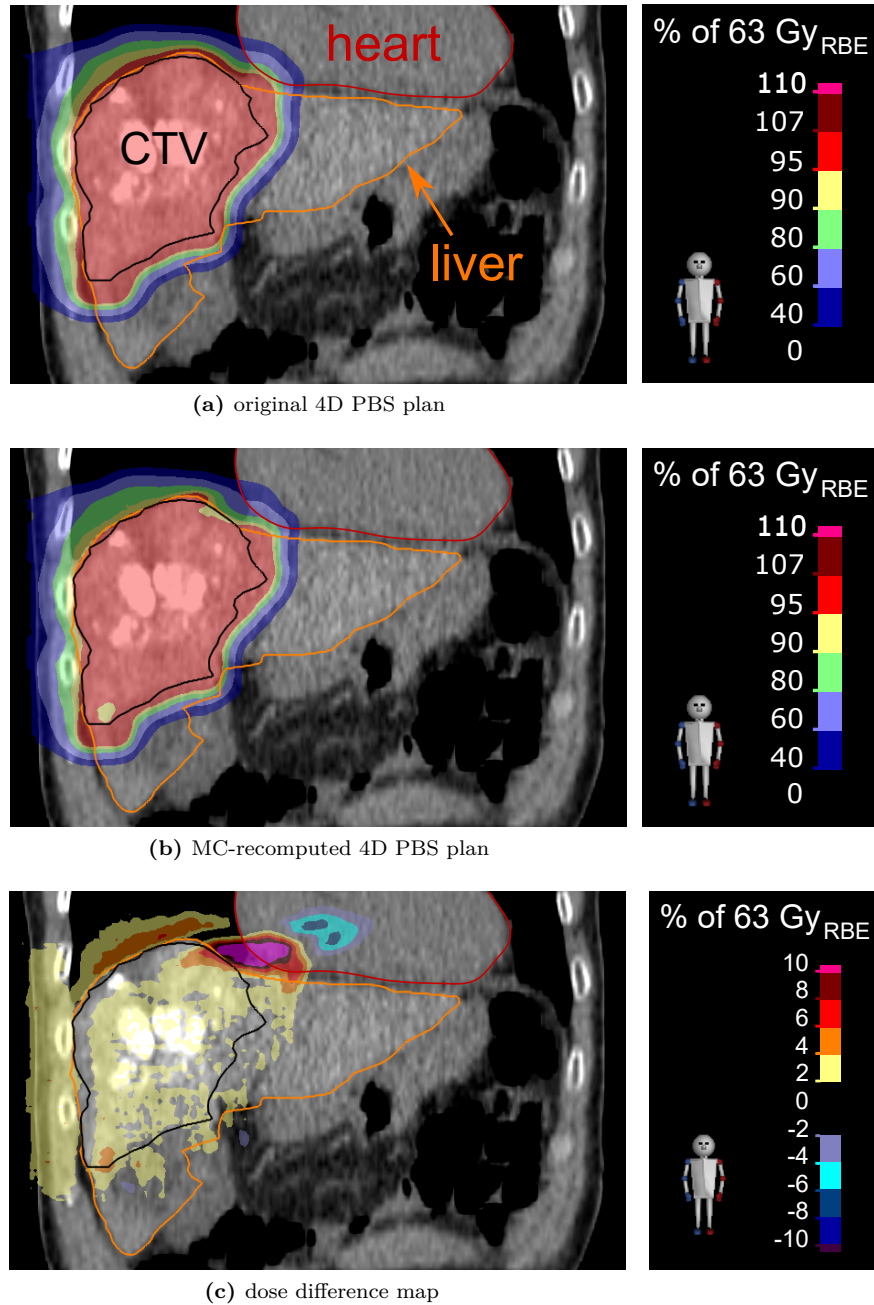


Figure 4.11: Comparison of the static dose distribution of the original 4D PBS treatment plan calculated with the Pencil Beam algorithm (a) and a copy of the treatment plan which was recomputed with the Monte Carlo (MC) dose engine (b) for patient 3: The differences in the dose distributions are displayed as colour wash (c).

Table 4.2: 4D dynamically accumulated dose statistics for 4D PBS plans optimised with the Pencil Beam algorithm (PB) and Monte Carlo (MC): The average values and standard deviations were calculated based on 60 interplay scenarios per plan for two patients (ID 2 and 3).

| | | ID 2 | | ID 3 | |
|--------------|---------------------------------|-----------|-----------|-----------|-----------|
| | | PB | MC | PB | MC |
| CTV | HI (%) | 10.6±1.6 | 11.1±2.2 | 8.1±0.6 | 9.1±0.7 |
| | $V_{107/95}$ (%) | 7.9±3.8 | 9.0±4.7 | 3.4±1.2 | 4.9±1.5 |
| normal liver | average dose ($G_{y_{RBE}}$) | 13.8±0.0 | 15.0±0.0 | 19.9±0.0 | 20.6±0.0 |
| | $V_{42G_{y_{RBE}}}$ (%) | 17.2±0.2 | 19.0±0.1 | 25.4±0.1 | 25.7±0.1 |
| | $V_{33G_{y_{RBE}}}$ (%) | 20.7±0.1 | 22.7±0.2 | 29.7±0.1 | 30.2±0.1 |
| | $V_{<15G_{y_{RBE}}}$ (cm^3) | 796.2±2.2 | 769.4±2.5 | 562.4±0.8 | 546.2±0.8 |

Three-field PBS plans

The static dose distributions of the three-field 4D PBS plans were similar to the ones of the two-field 4D PBS plans with regard to the CTV coverage. The percentage over- and underdosage was zero for all plans. However, the third field increased the dose to the normal liver tissue. The average dose was 22.6% (ID 2) and 2.5% (ID 3) higher than for two fields. The normal liver volume which was exposed to less than 15 $G_{y_{RBE}}$ decreased by 9.7% for patient 2 and by 2.7% for patient 3 compared to the original plan with two fields.

4DDD evaluations demonstrated that the CTV coverage improved in the presence of organ motion when using three fields instead of two. The average HI decreased by about 24.2% (ID 2) and 4.6% (ID 3) as can be seen in Table 4.3. The percentage over- and underdosage could be reduced by 65.6% (ID 2) and 41.2% (ID 3). On the downside, the normal liver was exposed to a higher dose. The differences in the normal liver dose were similar to the differences observed for the static dose distributions.

Table 4.3 shows that a third irradiation field in PBS still does not achieve comparable results to three-field DS plans. With one exception, HI and $V_{107/95}$ of the three-field PBS plans are closer to the results of the two-field PBS plans than to the three-field DS plans. However, the PBS plans are still better in sparing normal liver tissue than DS plans when using three fields.

Table 4.3: 4D dynamically accumulated dose statistics for two-field and three-field 4D PBS plans as well as three-field DS plans: The average values are listed for patient 2 and 3. In addition, standard deviations, considering 60 interplay evaluations per plan, are stated for PBS.

| | | ID 2 | | |
|--------------|---------------------------------|---------------|-----------------|----------------|
| | | two-field PBS | three-field PBS | three-field DS |
| CTV | HI (%) | 10.6±1.6 | 8.0±0.8 | 2.1 |
| | $V_{107/95}$ (%) | 7.9±3.8 | 2.7±1.7 | 0.0 |
| normal liver | average dose ($G_{y_{RBE}}$) | 13.8±0.0 | 16.6±0.0 | 17.6 |
| | $V_{42G_{y_{RBE}}}$ (%) | 17.2±0.2 | 19.8±0.1 | 22.0 |
| | $V_{33G_{y_{RBE}}}$ (%) | 20.7±0.1 | 23.6±0.1 | 25.2 |
| | $V_{<15G_{y_{RBE}}}$ (cm^3) | 796.2±2.2 | 716.2±2.4 | 692.5 |

| | | ID 3 | | |
|--------------|---------------------------------|---------------|-----------------|----------------|
| | | two-field PBS | three-field PBS | three-field DS |
| CTV | HI (%) | 8.1±0.6 | 7.8±0.6 | 4.0 |
| | $V_{107/95}$ (%) | 3.4±1.2 | 2.0±1.0 | 0.0 |
| normal liver | average dose ($G_{y_{RBE}}$) | 19.9±0.0 | 20.5±0.0 | 23.6 |
| | $V_{42G_{y_{RBE}}}$ (%) | 25.4±0.1 | 26.0±0.1 | 31.8 |
| | $V_{33G_{y_{RBE}}}$ (%) | 29.7±0.1 | 30.7±0.1 | 35.9 |
| | $V_{<15G_{y_{RBE}}}$ (cm^3) | 562.4±0.8 | 545.9±0.8 | 501.3 |

4.4 Discussion

4.4.1 Evaluation of the deformable image registration

The DIR evaluation focused on the deformation between the end-exhale and end-inhale phase, representative for the largest errors. The average target registration error is, consequently, even smaller than the stated value of 0.20 cm for the investigated maximum deformations. That could be demonstrated based on 10 additional samples including deformations to other phases. For these samples, the average target registration error was reduced by about 20%. An improvement is also expected for the Dice similarity coefficient when considering all motion phases.

Compared to the average tumour motion amplitude, which ranged from about 0.4 cm to 1.5 cm, the target registration error was small but not insignificant. The largest contribution to the registration error came from the SI direction. One reason is that it corresponds to the main direction of motion. Another reason is that the spatial resolution was higher in the AP and LR direction than in the SI direction for almost all patients (see Sec. 4.2.1).

The target registration error exhibited a much larger standard deviation (SD=60%) than the Dice similarity coefficient (SD=3%). This can be explained by the fact that the target registration error of a single landmark strongly depends on the local CT contrast and the distance to the nearest controlling contour, whereas the Dice similarity coefficient compares the volume overlap of the liver contours which is not affected by errors of the deformation vector field in the middle of the liver.

The comparison with former DIR validation studies [66, 68, 69] showed that ANACONDA performed well for the selected HCC patients. The target registration errors and Dice similarity coefficients were very similar to and in some cases even better than previously reported values. Only when comparing to lung studies, the quality was slightly worse. This was expected due to the higher CT contrast in the case of lung which enables a better DIR quality.

The geometric errors induced by DIR can significantly impact 4DDD evaluations and depend on the DIR method as recently shown for three liver cancer patients by Ribeiro et al. [152]. An underestimation of dose inhomogeneities was reported for all six investigated DIR algorithms by the study. The dosimetric error was smallest for a DIR method which included the liver contour as controlling ROI (mean error of $V_{95}(CTV)$: 0.1% to 2.0%). The magnitude of the error decreased with the number of irradiation fields (one or three) and increased with the motion amplitude (8 mm to 21 mm). For a single field and a DIR method which did not include any controlling or focussing ROI, a mean error of up to 10.6% and an SD of 14.1% were observed. The results cannot be transferred to the current work in the form of error bars due to differences in the DIR method, the 4DDD computation and the number of fields. But they can help to interpret the 4DDD outcome. In this study, treatment plans consisted of multiple fields reducing the dosimetric uncertainties. Since the CTV served as controlling ROI, the registration error is expected to be much smaller within the CTV area than in the rest of the liver. As a consequence, the impact of deformation errors on the used interplay effect evaluation criteria, namely the HI and $V_{107/95}$ of the CTV, is rated comparatively low. Despite all measures, the 4DDD evaluations must be treated with great caution since the exact extent of residual uncertainties is not known for the given conditions.

4.4.2 Perturbation analysis of the nominal treatment plans

DS plans performed superior to 4D and 3D PBS plans in perturbation analyses. There are several reasons why they provided the better CTV coverage in the different scenarios. One decisive reason is that three fields were used instead of two. Further, the amount of dose deposited proximal to the target is higher than for PBS. Last but not least, compensator smearing was applied which leads to additional range margins. These characteristics make DS plans more robust at the expense of OAR sparing.

4D PBS plans achieved comparatively good CTV coverage for inhomogeneous target regions due to multi-field optimisation. For example, spots which were located directly behind a rib from beam's eye view received lower spot weights. In return, the amount of dose deposited by the second field was increased for this region.

In 3D PBS plans, neither a third field nor multi-field optimisation were applied. Density uncertainties were taken into account by adding uniform margins proximal and distal to the target, which cannot adequately compensate for both, range changes in liver and lung tissue, at the same time. As a consequence, it was not possible to create a robust 3D PBS plan for CTV 9b which is directly adjacent to the lung. For all other cases, robustness against the specified density perturbations, shifts of the isocentre and recomputations on the end-inhale phase could be achieved for all three planning techniques. The dynamics of the beam delivery were not considered in the robustness evaluation. Their impact on the dose distribution is discussed in the next section.

4.4.3 Analysis of the 4D dynamically accumulated dose distribution

CTV coverage

On average, 4D robust optimised PBS plans exhibited slightly less dose distortions than non-robustly optimised 3D PBS plans. However, interplay effects were still present and the observed differences of HI and $V_{107/95}$ between 4D and 3D PBS were not found to be significant. Due to the small size of the patient collective, it was not possible to demonstrate whether the slight improvement arose from statistical fluctuations of interplay effects or from an actual decrease of motion effects due to the robust optimisation. One argument against a merely coincidental result is that an improvement was observed for almost all CTVs in inhomogeneous iCTV regions, whereas 3D PBS plans performed better than 4D PBS plans for the two homogeneous target regions (ID 4 and 6). The latter might be explained by the fact that the beam-specific PTV margins in 3D PBS plans were chosen rather conservatively resulting in a larger volume covered by the 95% isodose line than in robustly optimised 4D PBS plans for homogeneous regions (category I). The larger safety margins are expected to reduce the magnitude of motion effects at the edge of the CTV. Besides, a single field uniform dose approach increases the robustness if no major density gradients are involved. In contrast, 4D PBS plans had an advantage over 3D PBS plans for homogeneity category II and III since the shift of high density gradients due to motion, e.g. in the vicinity of ribs, was taken into account by the robust optimisation. The hypothesis that the benefit of 4D robust optimisation depends on the homogeneity category could not be tested for significance because of the small size of the subgroups and the unequal distribution over them (see Fig. 4.1a).

The additional 4DD evaluations, which neglected the time structure of the beam delivery, showed no over- or underdosage of the CTV for any PBS plan. The calculated HI values were very similar to the static case for both, 4D and 3D PBS plans. Therefore, the over- and underdosage and the dose inhomogeneities detected by 4DDD computations must have been caused by interplay effects and not by other motion effects. As a consequence, also the differences in the dose deterioration between 4D and 3D PBS plans are attributable to interplay effects. The small, positive effect of 4D robust optimisation on interplay observed for homogeneity categories II and III could be due to the fact that 4D robust optimisation takes different types of uncertainties into account. Even though 4D robust optimisation does not consider the time structure of the beam delivery, the robustness against organ motion, setup shifts, and range uncertainties might reduce the magnitude of interplay as a side effect. For example, the consideration of motion-induced range changes may guide the optimiser toward a solution with shallower 4D accumulated dose gradients which improves the 4DDD distribution. Besides, multi-field optimisation was exploited by 4D PBS plans, whereas a uniform dose per field was applied in non-robustly optimised 3D PBS plans following a conservative approach for CTV coverage.

A proof of concept study recently published by Engwall et al. [153] investigated the inclusion of uncertainties of the time structure in 4D robust optimisation. For simplicity reasons, only variations in the respiratory pattern were considered in the 4D robust optimisation settings but, in principle, uncertainties in the delivery time could be incorporated as well. For three non-small cell lung cancer patients, Engwall et al. demonstrated an improvement in V_{95} , V_{107}

and HI for the new optimisation approach compared to the standard 4D robust optimisation. Even better results were achieved when 4D robust optimisation including respiratory motion was combined with rescanning. The largest effect was observed for the patient with the highest motion amplitude. The study was limited to evaluations of interplay effects and did not take into account range and setup uncertainties. As a consequence, the number of optimisation scenarios would be much higher for clinical treatment plans. The effect is comparable to an increase of the safety margins in conventional treatment planning. In order to protect normal tissue and realise acceptable computation times, the number of scenarios must be kept as low as reasonable possible. For this purpose, motion mitigation techniques such as gating or abdominal compression which decrease the motion amplitude could be combined with the new 4D robust optimisation approach.

Looking at previous treatment planning studies on 4D robust optimisation, it can be seen that the reported benefit strongly depends on the patient cohort (target location) and the treatment technique used for the plan comparison, i.e. if a multi-field optimisation or single field uniform dose approach is applied and if conventional margins or 3D robust optimisation are utilised for the 3D PBS plan. In a study on non-small lung cancer ($D_{presc}=66 \text{ Gy}_{\text{RBE}}$, 33 fractions), Liu et al. showed a significant improvement of the target dose homogeneity (4D robust optimisation: $D_5=10.3 \text{ Gy}_{\text{RBE}}$, 3D robust optimisation: $D_5=17.7 \text{ Gy}_{\text{RBE}}$; $P=0.002$) and target coverage (4D robust optimisation: $D_{95}=60.6 \text{ Gy}_{\text{RBE}}$, 3D robust optimisation: $D_{95}=55.2 \text{ Gy}_{\text{RBE}}$; $P=0.001$) using 4D robust optimisation [17]. The greater benefit is probably attributable to the larger density variations in the thorax. The results are not directly comparable since Liu et al. used 3D robustly optimised PBS plans as a reference. Yu et al. compared 4D robustly optimised IMPT plans with conventional IMPT plans for distal oesophageal cancer ($D_{presc}=50.4 \text{ Gy}_{\text{RBE}}$, 28 fractions) [18]. They also observed a notable reduction of dose deviations arising from organ motion (4D robust optimisation: $D_{95}(iCTV)_{nominal} - D_{95}(iCTV)_{4DDD}=0.7 \text{ Gy}_{\text{RBE}}$, conventional margins: $D_{95}(iCTV)_{nominal} - D_{95}(iCTV)_{4DDD}=1.4 \text{ Gy}_{\text{RBE}}$; $P=0.008$) and proposed to implement robust optimisation for IMPT in treatment planning.

In this study, PBS 4DDD distributions were additionally compared to DS, the previous standard for moving targets. DS plans are, contrary to PBS plans, not affected by motion interplay effects due to the almost continuous, fast energy changes. In addition, their robustness against organ motion is increased due to the use of three fields and compensator smearing as already mentioned within the context of the perturbation analysis (Sec. 4.4.2). Looking at a single fraction, DS plans clearly outperform 4D and 3D PBS plans with HI values similar to the static case and no over- and underdosage at all. The HI is for DS even slightly better for 4DDD analyses than for the static plan. This is due to the fact that small inhomogeneities in the static plan are compensated by contributions of other 4D CT phases in the 4DDD evaluation. 4D and 3D PBS plans exhibit in contrast an average HI above 8% (4D and 3D PBS) and an average $V_{107/95}$ of about 6% (4D PBS) and 7% (3D PBS), respectively. However, it could be demonstrated that these inhomogeneities were averaged out over the whole treatment course. Local overdoses in single fractions were assumed to be non-critical as long as they were located within the CTV. Worst case scenarios in which the minimum dose of a single fraction was as low as 77% of D_{presc} were not found to be correlated to poor $V_{107/95}$ values. The underdosage occurred rather locally (small cold spots) and thus did not necessarily compromise the coverage of the entire target. Scenarios like that had a low likelihood, i.e. were single outliers among 60 simulations. On average, the minimum dose was above 90% for a single fraction. In stereotactic body radiotherapy (SBRT), it is common to allow underdosage of the GTV and PTV up to a certain degree even within the static plan. Examples for clinical goals from a liver SBRT plan comparison are $V_{90}(GTV) > 98\%$ and $V_{70}(PTV) > 98\%$ [154]. In contrast to SBRT, the over- and underdosage detected in this study is randomly introduced by motion interplay which is why it averages out after some fractions. Therefore, both, PBS and DS, produce acceptable 4DDD distributions for the majority of investigated patients.

It has been shown by several studies that conventional fractionation schemes can mitigate interplay effects to a large extent [30, 98, 155]. However, the current trend in radiotherapy towards hypofractionation reduces the averaging effect [156]. In addition, one may note that the bio-

logical effect of over- and underdosages in single fractions is not yet sufficiently well known. Underdosage in some fractions possibly cannot be compensated by overdosage in other fractions even if the fraction doses sum up exactly to the prescribed dose. This becomes an issue for small CTVs with large motion amplitudes located in inhomogeneous tissue, such as CTV 9a and 9b, for which interplay effects are particularly pronounced.

Since 4D robust optimisation alone cannot eliminate interplay effects in a single fraction, it should be combined with other motion mitigation techniques, such as respiratory gating or rescanning, for these indications. The difficulty is to decide a priori whether a patient can be safely treated using solely 4D robust optimisation. Since the severity of interplay effects depends on various parameters, such as CTV volume, motion amplitude, tissue heterogeneity, the number of fields or field weighting, a very large patient cohort must be studied in order to establish a correlation between the individual parameters and their impact on interplay effects. So far, it is not possible to specify certain limits for each parameter beyond which supplementary motion mitigation techniques become necessary and individual 4D analyses are required for each patient. Prospective evaluations should be complemented by retrospective analyses as the treatment conditions, e.g. the respiratory pattern or beam interruptions due to interlocks, cannot be determined before irradiation.

The 3D scatter plot (Fig. 4.7) revealed low $V_{107/95}$ values for large CTVs with small motion amplitudes independent of the homogeneity category. It is not possible to conclude from the available data whether the small magnitude of interplay effects is solely attributable to the low motion amplitude or whether a large CTV volume is also associated with reduced interplay effects. For instance, motion effects are expected to be more severe at the edge of the target where steep dose gradients and tissue inhomogeneities appear. The impact of over- and underdosage due to effects at the CTV surface might therefore decrease with increasing volume.

ID 4 and 6 exhibited slightly less over- and underdosage for 3D than for 4D PBS plans. Both cases belong to the homogeneity category I and their PTV margins were large compared to their CTV volume. Therefore, the coverage of the 3D PBS plans was more conservative than the coverage of the 4D PBS plans at the expense of OAR sparing.

A systematic investigation of the different variables on the basis of the available data was not possible. This aspect has to be addressed in future studies as described in Section 6.4.

OAR sparing

The question which treatment technique, i.e. DS, 4D PBS or 3D PBS, performs best in sparing OARs cannot be answered in general since it depends on the individual positions of OARs relative to the target and gantry angles. But looking at the normal liver tissue and ribs, PBS plans yielded lower OAR doses than DS, with 4D PBS being better than 3D PBS.

OARs proximal to the target from beam's eye view, such as ribs, could be spared well with 4D PBS, whereas large parts of the OAR volume received more than 95% of the prescribed target dose using DS or 3D PBS (see Tab. 4.1 and Fig. 4.3). A reason is that DS deposit unwanted dose proximal to the target and that the third irradiation field causes a wider overlap of the fields in the entrance region. 3D PBS plans led to higher OAR doses than 4D PBS plans due to the single field uniform dose approach and uniform PTV margins which do not allow for any flexibility of the optimiser to spare those structures. Therefore, 4D robust optimisation could prevent side effects such as rib fractures or RILD for HCC patients.

OARs next to the tumour, such as the kidney or the heart, are expected to benefit from the sharp lateral dose fall-off in DS. However, this advantage is often cancelled out by the third irradiation field. On the one hand, the third field lowers the entrance dose of each field but on the other hand more healthy tissue is exposed to irradiation, especially when the third field angle lead to long path lengths through the body. The influence of the number of fields on the plan comparison is discussed in more detail in Section 4.4.4.

It is possible to further improve the lateral dose fall-off of scanned beams, and thus the normal tissue sparing capabilities of PBS, by mounting supplementary apertures on the nozzle like in DS. This was shown for example by Yasui et al. [157], Moteabbed et al. [158] and Bäumer et al. [159].

The differences in the investigated dose volume histogram metrics were much larger when compared between different patients than when compared between different techniques for one patient due to the wide range of tumour volumes. While the average normal liver dose ranged for example from about 2 Gy_{RBE} (ID 5) to 26 Gy_{RBE} (ID 7) and exhibited a standard deviation of around 60% (DS: 8.20 Gy_{RBE}, 4D PBS: 6.8 Gy_{RBE}, 3D PBS: 7.0 Gy_{RBE}), the maximum detected difference between 4D PBS, 3D PBS and DS in the average dose for one patient was less than 4 Gy_{RBE}. Because of the large standard deviations for the patient collective, it was not possible to demonstrate the significance of OAR exposure differences between 4D PBS and 3D PBS/ 4D PBS and DS using the t-test even though each patient showed better or at least equal results for 4D PBS.

For small tumour volumes, the normal liver volume receiving less than 15 Gy_{RBE} was similar for all treatment techniques. But for large tumour volumes, such as patient ID 7, the additional irradiation field in DS led to comparatively poorer results since the percentage irradiated liver volume increased significantly. A notable difference in the irradiated normal liver volume was observed for DS depending on whether the third field passed mainly through the liver (e.g. patient ID 8a) or went partly through adjacent anatomical structures (e.g. patient ID 8b).

For two patients with large CTV volumes of almost 500 cm³ and 650 cm³, DS exceeded the OAR dose constraints and reached unacceptable high probabilities to develop RILD. The risk of RILD for the other tumours having less than 270 cm³ was zero. Hence, the size of the tumour is an important criterion for the choice of the treatment technique and large CTV volumes might not be safely treated with DS. The correlation between the size of a hepatocellular carcinoma and the risk of RILD has been investigated for non-robustly optimised PBS proton plans and intensity-modulated photon therapy plans by Toramatsu et al. [149]. The study reported much lower risks of RILD for protons than for photons, particularly when the diameter of the tumour was larger than 6.3 cm (average RILD risk: 94.5% for intensity-modulated photon therapy and 6.2% for PBS proton therapy). It is expected that the risk would even be lower applying 4D PBS proton therapy.

In this study, dose limits for close-by OARs, such as the stomach or bowel, were ignored in the optimisation if the OAR was located adjacent to or within the iCTV in order to maintain the target coverage. Otherwise, the interplay effect evaluation would be impacted. In clinic, the PTV (3D PBS and DS) and the robustness settings (4D PBS), respectively, would have to be compromised after some fractions to maintain vital OAR functions. Another solution could be to place a surgical spacer in-between the liver and for example the gastrointestinal tract [160].

Reproducibility of the results

Average dose volume histogram values were better reproducible for the liver than for the CTV. The percentage SEM was 0.0% for all liver dose volume histogram metrics for both plans, whereas it was about 1.0% for *HI* (4D and 3D PBS) and 3.9%/ 5.3% for $V_{107/95}$ (4D PBS/ 3D PBS). This is due to the fact that the average normal liver dose, the normal liver volume receiving at least 42 Gy_{RBE} and the normal liver volume receiving less than 15 Gy_{RBE}, are less sensitive to changes of the dose distribution. For instance, the amount of normal liver tissue exposed to 42 Gy_{RBE} and more remains similar, even though the 42 Gy_{RBE} isodose line might slightly shift for different scenarios. Only a small portion of the normal liver volume is exposed to high doses so that the average normal liver dose and $V_{<15\text{Gy}_{\text{RBE}}}$ do not show a strong dependence on motion effects. In contrast, not only a certain fraction but the entire CTV volume is affected by interplay effects. Nonetheless, it could be demonstrated that the calculated mean values were representative for the respective treatment plan. The differences between 4D and 3D PBS plans were significant even for $V_{107/95}$ which exhibited the highest SEM values. Thus, a number of 60 interplay scenarios per evaluation was found to be sufficient for plan comparison purposes. With regard to the highly reproducible mean values for the normal liver tissue, even less scenarios could be used.

Limitations of the treatment planning study

One of the limitations of the study is the small size of the patient cohort, which restricts the validity of the outcome. At the time of the study, WPE did not treat any movable targets so that 4D CT data sets had to be acquired from other clinics. It was not possible to gain more suitable patient data since the 4D CTs must meet a number of requirements. For instance, the tumour volume has to be visible on all respiratory phases (e.g. by use of contrast agents), the length of the CT scan should be long enough to evaluate relevant OAR dose volume histogram metrics and motion artefacts have to be low/ moderate. Furthermore, the patients must have signed a declaration of consent.

A higher number of patients would also be beneficial with regard to statistical analyses. The significance of differences was tested using a paired t-test which is subject to certain conditions. For instance, the test assumes normality of the variable under investigation in the two subgroups. Due to the small sample size, it was not possible to demonstrate that the data follow a normal distribution. Fortunately, the t-test is very robust toward data which are non-normal distributed and an approximately symmetric distribution of the variable in the population is sufficient to generate a more or less normal sample distribution. In case of a violation of the assumption, the t-test is still valid if the sample sizes are equal and not too small (ideally $n_1 = n_2 > 30$) [161]. Another assumption is that the two subgroups have the same variance. However, if the sample sizes of the subgroups are equal, as it is the case here, the t-test is also highly robust to unequal variances [162].

The availability of different HCC patient collectives would enable to assess the general validity of the study by computing SEMs for 4D PBS, 3D PBS and DS.

As already discussed in Chapter 3, the Pencil Beam dose algorithm has some weak points in comparison to the MC dose engine which might have affect the results. This issue was examined for some test cases and is addressed in detail in Section 4.4.4.

Following the example of a previously published study on scanned proton treatments of HCCs [149], the risk of RILD was calculated based on the LKB-NTCP model parameters determined by Dawson et al. for patients treated with 3D conformal photon beams [134]. However, the use of biological models is prone to uncertainties. Even though the higher biological effectiveness of protons compared to photons was considered in the calculations and dose volume histograms were normalised to correct for differences in the fractionation scheme using the LQ model, the results only represent estimates. Adequate follow-up data for the institution specific patients would be required to assess the validity of a model for deviating conditions (patient cohort, technique and fractionation scheme). Ideally, a NTCP model dedicated to the respective patient population and treatment protocol should be created, but that was beyond the scope of this work. Despite all uncertainties, the calculated NTCP values are well-suited for a relative comparison of 4D PBS, 3D PBS and DS plans.

Due to the high number of interplay relevant parameters, not all influencing factors could be studied within this work. For instance, a fixed respiratory cycle length was assumed. The magnitude of interplay effects, however, may differ for other respiratory periods because of the relative timing of dose application and organ motion. Further, another spot size might be more beneficial regarding interplay effects and OAR sparing [98, 163].

The presented results are only valid for the machine characteristics described in Section 3.3.2. Other institutes must determine their facility-specific beam delivery parameters or have to develop their own beam time model when not using irradiation log files/ having access to the scanning control system. Even then, treatment interruptions due to interlocks could prevent 'accurate' predictions.

Furthermore, all 4DDD computations rely on the organ motion recorded by the 4D CT scan. Any changes in the breathing pattern during treatment can potentially invalidate them. Sieben-thal et al. reported for example liver drifts and deformations of more than 5 mm on a time-scale of tens of minutes after patient positioning [164].

Also the treatment conditions, e.g. the application of contrast agents, could influence the results.

For the clinical implementation, this means that the knowledge gained from 4DDD planning studies must be interpreted with great caution and one must bear in mind that individual 4DDD evaluations for patients might suggest a false sense of security. Nevertheless, such studies are crucial to test the suitability of motion mitigation techniques and to identify which types of tumours and target regions benefit most.

4.4.4 Supplementary analyses

Monte Carlo

The assessment of the accuracy of the Pencil Beam dose algorithm for HCC patients was based on the static plan comparisons with MC-recomputed plans. For patient 2, representative for liver tumours in more or less homogeneous regions, only small dose differences were observed even though range shifters had been applied. Therefore, and since the dose to the surrounding healthy tissue was rather over- than underestimated by the PB plan, the Pencil Beam algorithm was considered not to represent a safety risk for this patient and similar cases.

Patient 3, on the other hand, revealed comparatively large differences between PB and MC-recomputed plans, mainly in regions behind lung tissue or the peritoneal cavity. The causes were already explained previously (Sec. 2.4.5 and 3.2.3). Possible consequences of local overestimations of the target dose and local underestimations of OAR doses by the Pencil Beam algorithm are reduced tumour control and increased normal tissue complications. Thus, in clinic, the MC dose engine should be used for HCC patients with tumours in close proximity to the lung.

The accuracy of interplay effect simulations was evaluated based on MC-optimised plans. The PB plans underestimated the magnitude and standard deviation of interplay effects for both investigated patients. Nonetheless, they enabled a good, qualitative assessment, especially in view of the large variations among different scenarios (visualised by error bars in Fig. 4.5 and 4.6). The differences in the mean values of PB and MC-optimised plans were smaller than (ID 2) or almost equal to (ID 3) the respective standard deviations and did not influence the ranking of the plans.

One should bear in mind that the outcome of a single fraction can differ widely in terms of target coverage and that the average value alone is not a good measure for it. Consequently, a very precise determination of the average value is not required. What is needed instead is a realistic estimate of magnitude and range to allow for 4DDD comparisons of different techniques. The PB plans met this requirement.

If PB plans are used in clinic or treatment planning studies for superficial targets, the applied range shifters should be as thin as possible (dependent on the minimum energy of the machine and the available range shifters). Furthermore, the air gap should be reduced as much as the setup allows. Thus, the errors in the dose distribution of PB plans due to the poor modelling of secondary scattering can be decreased to a minimum.

Three-field PBS plans

The exposure of surrounding healthy tissue due to a third irradiation field in pencil beam scanning depends on the tumour location and beam direction. Accordingly, the increase of the dose to the normal liver tissue was different for patient 2 and 3. For some patients, a third irradiation field might be beneficial since it reduces the entrance dose per beam. This is for example the case for superficial targets where skin protection plays an important role. However, if the third field passes close-by OARs, which could be spared using two fields (e.g. the bowel or kidney), additional normal tissue complications can occur. Often, the path length of beams in the patient increases with the number of fields since also suboptimal beam angles must be chosen. Therefore, a larger volume of healthy tissue is exposed to 'low' doses. Another issue is that it might be necessary to irradiate through the patient couch when using three fields. If that is the case, higher range uncertainties have to be considered in treatment planning by

increasing the safety margins or the uncertainty in the robust optimisation settings.

With regard to organ motion, additional irradiation fields enhance the dose averaging effect similar to rescanning. A considerable improvement in HI and $V_{107/95}$ was observed for both patients. But the averaging effect was too small to compete with DS plans. The normal liver dose in different 4DDD scenarios was worse for three-field PBS plans than for two-field PBS plans but still significantly better than for DS. Thus, the supplementary analysis demonstrated that neither the higher homogeneity of the CTV nor the worse sparing of OARs of DS plans compared to two-field PBS plans can be solely attributed to the difference in the number of irradiation fields. For HCC patients with rather small to medium-sized tumours, which are not located in close proximity to OARs, a third irradiation field might be a good option to decrease interplay effects slightly. However, the third field should only be complementary to, and should not replace, other motion mitigation techniques.

In general, gantry angles have to be chosen carefully to avoid an overlap of the entrance channels and the passage of OARs proximal to the target. Furthermore, the beams should not stop directly in front of a critical structure due to the increased LET at the end of the particle range. Depending on the tumour location, it is not always possible to find three different beam directions that satisfy all these criteria. For such cases, PBS therapy which can also be realised using two fields has a clear advantage over DS which usually needs a third irradiation field due to the lower dose conformity.

4.4.5 Robust optimisation for motion management of hepatocellular carcinomas

4D robust optimisation incorporates motion, range uncertainties and setup uncertainties in the optimisation process to determine optimal spot position and weights. Thus, it creates individual 'margins' which are adapted to the tissue surrounding the HCC and minimises the irradiated volume. Since HCCs are often diagnosed at a late stage (see Sec. 4.1), their size is usually rather large and the probability that they are not only surrounded by liver tissue but border on different anatomical structures, such as ribs, lung tissue, the right kidney or bowel, is high. Consequently, HCC patients benefit particularly from customised 'margins' which consider the local density differences and achieve better sparing of normal tissue than the conventional PTV concept. The superiority of 4D robust optimisation with regard to the dose to normal tissue was shown for the investigated HCC patients in the second part of Section 4.3.3 on OAR sparing. The actual goal of 4D robust optimisation, however, is to render PBS plans less sensitive to uncertainties and motion. An increase of the robustness towards range and setup uncertainties using 4D robust optimisation for HCC patients could be demonstrated based on a case where the tumour was located next to the lung (Sec. 4.3.2). On average, a small improvement of the target coverage in the presence of motion was detected for 4D robustly optimised plans. But, unlike for studies on 4D robust optimisation for lung or oesophageal cancer, the difference was not significant and it seemed that some HCC subgroups (homogeneity category II and III) benefited more than others (Sec. 4.3.3).

Contrary to classic motion mitigation approaches such as gating or rescanning, 4D robust optimisation neither requires additional hardware nor software changes of the proton treatment delivery system. As in many other facilities, rescanning and gating are not yet clinical available at WPE. Besides the technical requirements, the introduction of these techniques involves thorough commissioning tests, i.e. additional manpower. Since 4D robust optimisation is an integrated tool in the currently used TPS version, its implementation in treatment planning for HCC patients is straightforward.

While gating, breath hold or rescanning reduce significantly the efficiency of proton therapy, 4D robust optimisation does not prolong the treatment or setup time. Thus, the workload for the staff is less and more patients could be treated per day. Moreover, shorter treatment times are beneficial for the patients who might suffer from pain and cannot remain in the treatment position for a long time.

4D robust optimisation is neither related to any additional health risks, unlike anaesthesia, nor

does it cause any inconvenience for the patients as it is the case for abdominal compression, breath hold or spirometry, which is used at some facilities as gating signal.

On the downside, 4D robust optimisation is associated with an increased computation time. Considering ten respiratory phases and the range and setup uncertainties specified above (resulting in 210 optimisation scenarios), the optimisation usually takes about one hour for target volumes smaller than 500 cm^3 using the Pencil Beam dose algorithm. However, very large target volumes, such as the CTV 7, represent an exception. Even though the resolution was temporarily reduced to 4 mm, the optimisation time was more than 30 hours for this CTV. This is not suitable for clinical routine where usually several iterations are required to find the objectives and constraints which lead to a reasonable treatment plan. Moreover, the application of the MC engine would further prolong the computation time. In comparison, it took on average only four minutes to optimise 3D PBS plans with the Pencil Beam dose algorithm.

RayStation does not yet offer a GPU-based dose engine or high-performance computing for protons as it does for photons. The computing performance and memory consumption improved in RayStation 7 compared to the TPS versions used in this study resulting in shorter optimisation times and more interplay scenarios which can be computed in a single interplay evaluation. A 4D robust optimisation including 210 optimisation scenarios can be performed within 15 to 30 minutes (target volume $< 500 \text{ cm}^3$). The optimisation time for the CTV 7 is reduced to approximately 70 minutes.

In order to further decrease the optimisation time, it could be tested whether it is sufficient to consider only the end-expiration and end-inspiration phase in the robust optimisation. This would reduce the number of perturbation scenarios considered by the optimizer by a factor of five. If gating is applied, the number of respiratory phases included in the 4D optimisation settings is reduced depending on the size of the gating window.

The workflow of 4D treatment planning is most efficient when standard procedures have been automated. For example, perturbation analyses and interplay effect evaluations of 4D robustly optimised plans can be executed by scripts as done in the current work. But also preliminary work steps such as contouring could be optimised, e.g. using atlas-based automatic segmentation. This was successfully tested within the scope of this thesis: The External and the body contour delineated in only one respiratory phase of the HCC patients served to create a dedicated atlas for the patient collective. Thus, the contours in all other phases could be generated automatically.

In the HCC treatment planning study, it was assumed that the respiratory patterns during the treatment course are identical to the one at the time of the 4D CT scan. However, for real patients, changes in amplitude and breathing period might occur. Small variations of the breathing period might be covered by the evaluation method since random start phases and energy layer switching times between 0.9 s and 2.5 s are sampled for each scenario (see Sec. 3.3.2). But the analysis cannot cover uncertainties in the motion amplitude as it is based on a single 4D CT. Therefore, 4D CT scans should be performed on a weekly basis as recommended for thoracic malignancies [15]. In order to assess whether the plan robustness is still sufficient or replanning is needed, 4DDD computations based on the new 4D CT have to be conducted.

For large motion amplitudes, 4D robustly optimised PBS plans could not provide sufficient target coverage. In such cases, 4D robust optimisation cannot replace conventional motion mitigation techniques and should be used as a supplementary method. Rescanning, breath hold, gating or a combination of them would help to further reduce interplay effects for patients with large motion. If the clinical goals cannot be met, e.g. because appropriate motion mitigation techniques are not available at a centre, PBS therapy should not be offered and DS might be a treatment option for the HCC patients concerned.

5 Usage of computer generated 4D CTs in interplay studies - a proof of concept

This chapter presents a proof of concept study which investigates the application of computer generated 4D CTs for systematic investigations of interplay effects in PBS. The project was carried out in collaboration with the Institute of Medical Physics and Radiation Protection of the TH Mittelhessen (Dr. Ulf Mäder/ Dr. Jörg Wulff). To minimise the effort involved, interplay effects were studied for different tumour sizes while all other parameters such as motion amplitude and tissue characteristics were kept constant. Thus, only a single virtual patient had to be created. Spherical liver tumours with different diameter were delineated in one respiratory phase of the 4D CT and mapped to the other phases using DIR. Even though it was not necessary to generate modified 4D CT data sets for this specific test case, taken a simulated 4D CT instead of real patient image data was still an advantage. Since it was possible to assign unique identification numbers to each structure, a large number of anatomical structures could be easily contoured in different respiratory phases of the simulated 4D CT. These delineated contours were required to improve the quality of the DIR on the basis of which the different tumour contours were mapped to the other phases (see Sec. 5.2.1). For real patient data, it is not possible to accurately map artificial contours if no controlling ROIs exist in close proximity to them. Contrast agents and implanted markers are either missing or not at the right place to indicate several different tumour positions required for such a study.

The magnitude of interplay effects for PBS plans was analysed dependent on the tumour size using 4DDD computation.

In the event of a positive performance, different 4D CT series could be created in future studies in which for example only the motion amplitude or the homogeneity of the liver tissue is altered.

5.1 4D XCAT phantom and CT simulation

There is an urgent need for realistic 4D phantoms in proton therapy to evaluate motion mitigation techniques for PBS therapy of moving targets. An adequate physical, anthropomorphic motion phantom which, firstly, provides similar physical properties as real tissues (regarding proton stopping power, CT contrast and MRI contrast) and, secondly, mimics physiologic motion for different conditions seen in vivo is commercially not available. Due to the wide variety of patient sizes and pathologies, the production of phantoms representing the entire spectrum is anyhow not feasible. A good solution is therefore to use digital phantoms, i.e. realistic computer simulations, for in silico studies.

Computerised phantoms represent a trade-off between realism and flexibility. A distinction is made between voxelised and mathematical phantoms. Voxelised phantoms, which are usually generated from MR or CT data sets, are more realistic but do not allow for exact ray tracing due to their limited resolution. Mathematical phantoms, in contrast, are based on simple equations and can be created with any resolution without interpolation errors. Variations in anatomy or organ motion are easy to model but the reproduction of organ shapes is limited in accuracy [165].

The 4D extended cardiac-torso (XCAT) phantom is a whole-body computer model of human anatomy and physiology [166]. It represents a hybrid approach which combines the advantages of mathematical and voxelised phantoms. Complex mathematical primitives were used to describe human anatomy based on 3D male and female image data sets provided by the National Library of Medicine [167]. Thus, XCAT offers a high level of anatomical detail as well as flexible organ shapes. The flexibility was mathematically realised by non-uniform rational b-spline surfaces [168]. An extension of the original XCAT phantom to four dimensions to simulate cardiac and respiratory motion was achieved through the use of 4D cardiac- and respiratory-gated MR and CT data. The 4D XCAT phantom is very well-suited for research purposes since it can be used to create a realistic patient cohort consisting of males and females as well as children and adults with varying anatomy and physiologic motion.

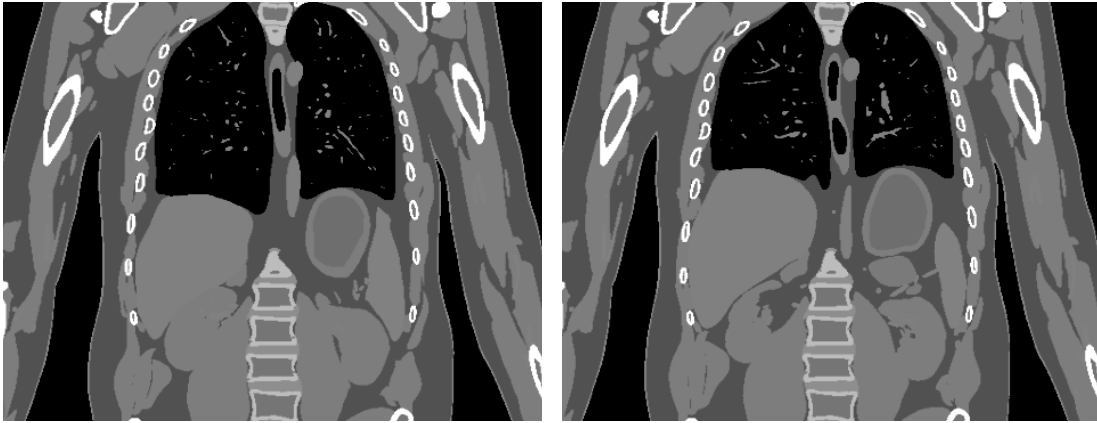


Figure 5.1: Simulated CT data of a 4D extended cardiac-torso (XCAT) phantom at end-inhale (left) and end-exhale phase (right).

CT images were created based on a discretisation of the mathematical phantom geometry assigning unique identifier to the voxels that compose an organ. Thus, they were neither impaired by quantum noise nor any CT artefacts. This is a considerable advantage over real 4D CT patient data which often exhibit motion artefacts due to reconstruction errors in combination with irregular breathing or artefacts due to implants. If the aim is to create more realistic CT images, facility-specific image noise can be simulated by using the 4D XCAT phantom and an x-ray CT projection algorithm which includes a model for quantum noise [165].

5.2 Study design

5.2.1 Simulated 4D CT data and deformable image registration

Figure 5.1 shows exemplarily two coronal views of the computer generated CT data for the end-inhale and end-exhale simulation. In order to facilitate contouring of the different anatomical structures in all respiratory phases two identical 4D CT data sets were created, one with realistic CT numbers and one with artificial, unique identification numbers for each tissue so that the intensity values of different structures did not overlap. Thus, it was possible to use a threshold-based, automatic segmentation tool to contour easily a large number of structures in each respiratory phase (see Fig. 5.2). In total, 48 OARs were created. The ROIs were copied to the original 4D CT data set and served as controlling structures for the DIR. A high quality of the DIR was essential since it had to be used to map artificial tumour volumes from the planning CT to the other phases. Moreover, the quality of the DIR affects the accuracy of 4DDD computations.

The method was tested based on a 'leave one out' analysis. To this end, a second series of deformable registrations was created using the same controlling ROIs except for the gall bladder. Subsequently, the contour of the gall bladder was mapped from the reference phase to the other phases using the new registrations. The mapped and the original contour were compared for each phase based on the Dice similarity coefficient (Eq. 4.2).

Five spherical liver tumours were created in the end-exhale phase with a radius of 1 cm, 2 cm, 3 cm, 4 cm and 5 cm, respectively, using the image segmentation tool of RayStation. All volumes share a common central point, which was selected in such a way that it was surrounded by a maximum amount of liver tissue in all directions. 5 cm was the largest possible radius for which the tumour did not exceed the liver contour. The contours correspond to tumour volumes of about 4 cm^3 to 524 cm^3 . They were mapped to the other phases through DIR including all OARs as controlling ROIs. The tumour motion amplitude defined as in Equation 4.1 was very similar for all CTVs and ranged from 2.17 cm to 2.19 cm.

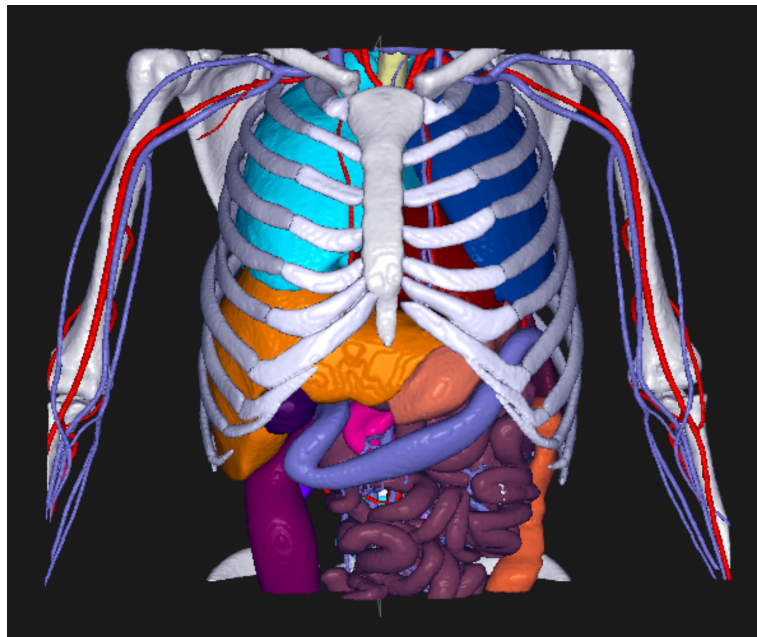


Figure 5.2: Three-dimensional view of the different anatomical structures of the 4D extended cardiothorax (XCAT) phantom contoured in RayStation. The External, the body contour and muscle contours were not displayed since they would obscure the view on the other ROIs. Small, detailed structures are covered by the outer shapes of the organs and are therefore not visible.

5.2.2 Treatment planning

One PBS treatment plan was created for each CTV in the research version of RayStation 7. Without loss of generality, 3D PBS plans were chosen for the feasibility study instead of 4D PBS plans in order to keep the computation time low. As before, beam specific PTVs cover the iCTV plus a 2 mm setup margin and 2 mm plus 3.5% range uncertainty in beam direction. The isocentre of each beam was the centre of the iCTV. The prescription to the CTV was the same as for the HCC treatment planning study (63 Gy_{RBE} in 15 fractions). Two fields were applied for each treatment plan using the same gantry and patient couch angles. For liver irradiations, the patient positioning would be different than simulated. The patient has to raise his arms above his head in order to provide greater freedom of choice between different beam angles. In future studies, this aspect can be taken into account when creating the phantom. Here, a material override with air was used for the right arm. The gantry angles were 20° (an angle from above) and 110° (an angle from the right side), respectively, at a couch angle of 0° each (patient position: head in, supine). Plans were optimised on a 2 mm dose grid using the MC dose engine and identical objectives/ constraints. For the large tumours extending almost to the patient surface, range shifters had to be applied. The second smallest CTV (radius=2 cm) required only a thin range shifter for one beam and the smallest CTV (radius= 1 cm) no range shifter at all. The air gap was minimised by moving the snout (holding the range shifter) as close to the patient's surface as possible. Dependent on the tumour size and whether a range shifter was present, the minimum energy ranged from 100 MeV to 134 MeV and the maximum energy from 127 MeV to 196 MeV.

5.2.3 Evaluation methods

Each plan had to pass the CTV coverage test in the static case and the robustness evaluation as explained in Section 4.2.4. The 4DDD evaluations were carried out in the same way as described in the previous chapter (see Sec. 4.2.5). Since RayStation crashed during the 4DDD computation for some of the MC-optimised plans, the final dose had to be recomputed with

Table 5.1: Homogeneity index HI and percentage over- and underdosage $V_{107/95}$ of the CTV in dependence on the CTV size for 4D dynamically accumulated dose distributions of a single treatment fraction: CTV 1 to CTV 5 correspond to radii from 1 cm to 5 cm.

| | | CTV 1 | CTV 2 | CTV 3 | CTV 4 | CTV 5 |
|------------------|---------|-------|-------|-------|-------|-------|
| HI (%) | average | 11.5 | 13.5 | 15.3 | 17.7 | 22.3 |
| | SD | 3.0 | 2.9 | 2.2 | 1.9 | 2.4 |
| $V_{107/95}$ (%) | average | 24.5 | 22.4 | 22.3 | 26.8 | 39.9 |
| | SD | 17.6 | 11.0 | 7.6 | 5.2 | 5.4 |

the Pencil Beam algorithm for all plans and the evaluations were repeated. Again, HI and $V_{107/95}$ served as evaluation criteria for interplay effects (Eq. 4.4 and 4.5). Each 4DDD plan evaluation included 50 interplay scenarios. Besides the average value, the standard deviation of HI and $V_{107/95}$ was calculated for each plan to reveal interplay effect variations. In order to test the reproducibility of the results the average value and the SEM (Eq. 4.3) of five repeated evaluations were calculated for two CTVs (radius=1 cm and 5 cm).

5.3 Results

5.3.1 Evaluation of the deformable image registration

The Dice similarity coefficient was analysed for the gall bladder which has a very small volume compared to the liver which was used for the test in the previous chapter. The average Dice similarity coefficient of nine deformations for the gall bladder was 0.93 with a standard deviation of 0.03. The highest coefficients were reached for the respiratory phases prior and subsequent to the reference phase (0.98 and 0.97). The lowest value (0.89) was observed for the end-inhale phase (0%) and the 10% phase.

5.3.2 CTV coverage and perturbation analysis of the nominal treatment plans

The static 3D PBS plans fulfilled the target coverage criterium ($V_{95}(CTV) = 100\%$). They showed no over- or underdosage and the HI values ranged from 1.4% (radius=1 cm) to 2.4% (radius=5 cm).

Except for the smallest CTV, all treatment plans exhibited $V_{95}(CTV)$ values in-between 97.6% and 100.0% for the tested perturbation scenarios. But even the CTV with a radius of 1 cm reached in almost all perturbation scenarios values above 98.0%. There was only one case in which $V_{95}(CTV)$ slightly remained under the percentage volume limit of 95% with 94.9%. However, the minimum dose within the CTV was still 92.9% of the prescribed dose.

5.3.3 Analysis of the 4D dynamically accumulated dose distribution

Two 4DDD distributions are shown as examples for a small and a large CTV in Figure 5.3. Table 5.1 summarises the results of the interplay effect evaluation for all investigated CTV radii (1 cm to 5 cm). The HI rose with increasing CTV size from $11.5\% \pm 3.0\%$ (average value \pm SD) to $22.3\% \pm 2.4\%$, whereas $V_{107/95}$ did not follow a clear trend. The average value of $V_{107/95}$ was maximal for the largest CTV with 39.9%. The smallest value was observed for a radius of 3 cm ($V_{107/95}=22.3\%$). $V_{107/95}$ exhibited large SDs compared to the HI ranging from a percentage volume of 5.2% to 17.6%.

Reproducibility of the results

Despite large fluctuations of HI and $V_{107/95}$ among different interplay scenarios of a treatment plan, the average value was reasonably well reproducible when considering 50 interplay scenarios. For the largest CTV (radius=5 cm), the SEM was 0.3% and 0.2% for HI and $V_{107/95}$, respectively (average $HI=22.5\%$, average $V_{107/95}=39.6\%$). The same reproducibility test for

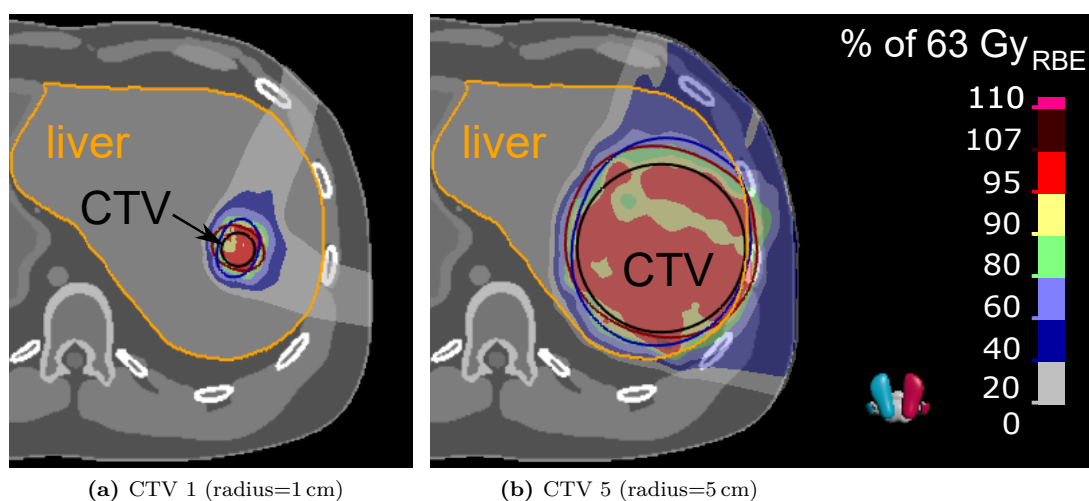


Figure 5.3: 4D dynamically accumulated dose distribution (overlaid as colour wash) of a random interplay scenario for the smallest and the largest CTV (transversal view): The beam-specific PTV margins are coloured dark blue (gantry angle= 20°) and dark red (gantry angle= 110°).

the smallest CTV yielded an SEM of 0.3% for HI and an SEM of 1.6% for $V_{107/95}$ (average $HI= 12.0\%$, average $V_{107/95}=25.6\%$).

5.3.4 Feasibility of interplay studies using simulated 4D CT data

The simulated CT data could be imported into RayStation without any difficulties after adjusting the header of the Digital Imaging and Communications in Medicine (DICOM) files to that of a clinical 4D CT. RayStation tools such as automatic image segmentation could be applied and it was possible to perform deformable image registrations as well as treatment plan optimisations with the Pencil Beam and the MC dose engine. The execution of scripts for the robustness or interplay effect evaluation was successful. The event that the application crashes when performing interplay evaluations of MC-optimised plans was also observed for real patient images and is not attributable to the simulated CT data.

5.4 Discussion

5.4.1 Evaluation of the deformable image registration

The simulated 4D CT did not provide any contrast within the defined anatomical structures since one identification number was assigned to all voxels of each object. It was not possible to identify landmarks which could serve to calculate the target registration error.

The Dice similarity coefficient test was based on the gall bladder. Despite the small size of the gall bladder compared to the liver which was used as test ROI in Chapter 4, the average Dice similarity coefficient over all phases was 0.93 indicating a decent quality of the DIR. Nonetheless, there is still potential for improvement. For example, it would be helpful to incorporate a grid of small markers in the XCAT phantom which can be seen in the different respiratory phases of the simulated 4D CT. Such markers could be used as controlling POIs to further improve the quality of the deformation vector field, especially within large structures. Furthermore, a few markers which do not serve as controlling POIs could be used to assess the target registration error for different regions. One possible way to simulate CT markers are manual overrides of individual voxels. Alternatively, a tumour could be modelled by the 4D XCAT software and transformed in all respiratory phases.

In this approach, the same image values were assigned to the liver tissue and hepatic veins. If more contrast is needed within the liver, different image values could be used even though such a contrast might not be seen in real CT images.

5.4.2 Analysis of the 4D dynamically accumulated dose distribution

Contrary to expectations based on the HCC treatment planning study (Chap. 4, Fig. 4.7), a large target volume did not reduce the relative amount of interplay effects. The assumption that severe interplay effects occur mainly at the edge of the target volume, where steep dose gradients exist, and less commonly in the middle of a large irradiation volume could not be confirmed. Quite the opposite, the highest HI and $V_{107/95}$ values were observed for the largest CTV and the corresponding 4DDD distributions revealed that interplay also causes significant underdosage in central regions (Tab. 5.1 and Fig. 5.3).

The increase of the HI with the size of the CTV is probably related to the number of spots and the treatment time which increase the likelihood that spots are superimposed or drift apart due to interplay. For the five investigated CTVs, the treatment time increased approximately linearly with the volume from about 38 s to 116 s. The exact treatment time depends on the energy layer switching.

Another aspect which affects the amount of interplay are position-dependent motion amplitudes. Indeed, the average motion amplitude was almost equal for all CTV sizes but the span width between minimum and maximum amplitude differed. For the smallest CTV, for example, the motion amplitude ranged from 2.1 cm to 2.2 cm, whereas for the largest CTV, it ranged from 1.7 cm to 2.4 cm. An inhomogeneous deformation vector field enhances interplay effects.

In case of $V_{107/95}$, there is no clear correlation between the average value and the size of the CTV. In contrast to HI , $V_{107/95}$ does not consider small interplay effects leading to less than 7% 'overdosage' and 5% 'underdosage', respectively. Further analyses are required to find out whether a correlation exists.

For the largest CTV, tissue inhomogeneities contributed to the increase of HI and $V_{107/95}$ since the beam specific PTVs encompassed ribs (see Fig. 5.3b).

The standard deviation of $V_{107/95}$ was very high, especially for small CTV volumes. This result shows that the severity of interplay effects can vary considerably for a given CTV and treatment plan due to uncertainties of the time structure. As discussed in detail in Section 4.4.3, supplementary motion mitigation techniques would be necessary to safely treat such patients with PBS.

The proof of concept study indicates that individual variations in anatomy and physiologic motion demand for patient-specific interplay effect evaluations and that a simple parametrisation could be critical. It would be interesting to analyse the impact of the tumour size in more detail, for example for different CTV positions within the liver (only possible for small to medium-sized CTVs) and for different anatomies. The complexity of the different influences and their interference will probably prohibit a division into treatment subgroups solely based on the CTV size, motion amplitude and tissue homogeneity. This has to be confirmed by future 4D XCAT phantom studies focussing also on the dependence on motion amplitude and tissue homogeneity. The studies should be based on a larger database and include 4D robust optimisation.

5.4.3 Feasibility of interplay studies using simulated 4D CT data

All workflow steps from data import to interplay simulations were successfully completed for the computer generated 4D CT. Thus, it is possible to conduct also more complex 4D XCAT phantom studies requiring several 4D CTs with modified motion amplitudes or density distributions. In future studies, it should be ensured that the phantom is positioned according to the tumour site to avoid material overrides, e.g. the arms have to be raised above the head for liver irradiations. Further, simulated CT markers would improve the quality of the DIR and therefore the accuracy of the 4DDD computation.

Although this feasibility study was kept deliberately simple by using a single virtual patient with different targets, the approach allows to simulate a wide range of patient anatomies and physiologic motion.

Using the 4D XCAT phantom, 4D studies can be performed with perfect contouring, minimal DIR uncertainties and realistic image quality without the influence of CT artefacts.

6 Summary and outlook on future research

6.1 Summary - Implementation of 4D dynamic dose accumulation in treatment planning

A 4D dose calculation routine dedicated to the assessment of interplay effects in scanned proton beam therapy was developed for clinical use at WPE based on a rudimentary script provided by RaySearch (E. Engwall, June 2015). Interplay effect modelling relies on the knowledge of the beam delivery sequence and the time structure of organ motion. They can be determined on the basis of irradiation log files (retrospective evaluation), the algorithms of the scanning control systems (subject to access rights) or an empirical beam delivery time model and 4D CT data sets acquired prior to the treatment. The gained information is used to assign the delivery of a proton spot to a respiratory phase/ CT image. Using deformable image registration, the accumulated dose on each phase is mapped to the reference phase (planning CT). The sum of all deformed doses yields the 4D dynamically accumulated dose distribution.

A two-dimensional ionisation chamber array and a 4D phantom made up of different tissue surrogates and a dynamic platform served for the experimental validation of the routine. Respiratory motion was simulated by a sinusoidal motion with an amplitude of 5 mm and a time period of 10 s (first trial) and 5 s (second trial), respectively. The measured dose planes were compared to the corresponding dose distributions simulated by the routine using irradiation log files. Resulting gamma pass rates exceeded the 95% level demanded in patient quality assurance tests and were comparable to static references, with the exception for some measurements at a time period of 5 s. As discussed in Section 3.2.3, the worse gamma pass rates for these measurements are most probably attributable to the design of the study and not to shortcomings of the routine. For instance, the measurement was not synchronised with the irradiation and the exact time offset of the start phase was not considered in the calculation. Nonetheless, gamma pass rates higher than 88% could be reached for fast detector movements. A better agreement is expected when using adequate equipment such as 3D stereo camera units to track the phantom's motion.

The sensitivity of the routine to its input parameters (start phase and motion period) was confirmed. A change of the motion period by more than 0.4 s caused a degradation of the gamma pass rate in the double-digit range. Introducing a phase shift led to a reduction of the gamma pass rate of up to 20%. The uncertainties in the determination of the input parameters were much smaller.

Time resolved 4D dynamic dose calculations reproduced measurements in movie mode well on a qualitative level for a sampling time of 2 s. It was possible to model changes in the dose delivery with time (arising from differences in the irradiation start phase and the energy layer switching times) for repeated irradiations. However, good gamma pass rates were not observed for all time intervals. This might be explained by the fact that proton spots were assigned mistakenly to the previous or subsequent time interval because of uncertainties in the timing of the pencil beam and the phantom motion. Further reasons are stated in Section 3.2.3.

An empirical beam delivery time model was developed to allow for prospective treatment planning studies. To this end, the spot delivery time, the scanning speed and the energy layer switching time of the machine were determined for typical clinical spot doses, spot distances and beam energies. The spot delivery time as a function of the spot dose could be modelled by a rational function. The scan time as a function of the spot distance was approximated by step functions for the x and y direction. The energy layer switching time did not show a correlation with the beam energy and varied strongly. Therefore, random values are generated from a normal distribution based on the measured average value and standard deviation. In addition, an upper and lower limit was introduced for the energy layer switching time. Due to the fluctuations of the energy layer switching time, a certain number of simulations is required to get a realistic estimate of the range of interplay effects for a patient.

The beam delivery time model was validated by comparing actual delivery times extracted from log files with times calculated by the routine. The observed differences were on average

0.03 s per energy layer and 1.2% per field, assuming identical energy layer switching times. The impact of these errors on the dose distribution was negligible and the gamma pass rate was 100%.

Tests using simplified versions of the beam time model showed that an accurate modelling of the beam dynamics as in the current work is a prerequisite for prospective analyses of the impact of interplay effects. The assumption of an instant dose delivery per energy layer, for example, reduced the gamma pass rate by 10%.

6.2 Summary - Evaluation of 4D robust optimisation for HCC patients

The capability of 4D robust optimisation to improve the 4D dynamically accumulated dose distribution was investigated for hepatocellular carcinoma patients which generally have a poor prognosis and would benefit from an effective, local treatment. The study included eleven target volumes with motion amplitudes ranging from 0.4 cm to 1.5 cm. A median dose of 63 Gy_{RBE} was prescribed to the clinical target volume in 15 fractions. Conventional PBS plans based on beam-specific planning target volumes and DS plans served for comparison purposes regarding target coverage and organ at risk sparing. The number of irradiation fields differed for PBS and DS plans since DS plans required three instead of two fields to achieve sufficient dose conformity. Since RayStation was not supporting Monte Carlo calculations for DS at the time of the PhD project, optimisations were performed using the implemented Pencil Beam dose algorithm.

The evaluation of PBS plans was carried out with the developed 4D dynamic dose calculation routine, whereas for DS a uniform dose distribution over all motion phases was assumed due to the almost continuous energy changes. The possible outcome of a single treatment fraction in PBS was analysed based on 60 simulations per treatment plan with varying energy layer switching times and start phases. For the evaluation of the whole treatment course, the sum of 15 single fractions was modelled five times (75 simulations in total).

Looking at a single fraction, 4D robust optimisation did not improve the homogeneity index and the percentage over- and underdosage significantly compared to conventional PBS plans and the results remained far behind those of DS plans. However, for 15 PBS fractions, the over- and underdosage averaged out and the homogeneity index almost approached the static homogeneity index values. The number of patients was too small to draw general conclusions on the potential benefit of 4D robust optimisation for different subgroups (e.g. for different tumour sizes or tissue homogeneity degrees).

An advantage of 4D robust optimisation was demonstrated in terms of normal tissue sparing. The average normal liver dose was 5.7% lower than for conventional PBS plans and 15.9% lower than for DS plans. For large tumours, 4D robust optimisation could considerably reduce the risk to develop radiation-induced liver disease. Moreover, the probability of rib fractures was lowest for 4D robustly optimised plans.

The influence of the dose engine and the different number of irradiation fields for PBS and DS plans was exemplarily examined for two plans. The Pencil Beam dose algorithm underestimated the average magnitude and standard deviation of interplay effects compared to the Monte Carlo dose engine and yielded lower dose values for the normal liver tissue. The differences of the mean values, however, were smaller than or almost equal to the standard deviations for the homogeneity index and the percentage over- and underdosage. Thus, the Pencil Beam dose engine provided a good qualitative estimation and enables comparisons of different treatment plans.

The application of a third irradiation field in PBS, using the same beam directions as in DS, increased the robustness against motion as expected but the averaging effect was still too small to achieve similar results as corresponding DS plans. The dose to the normal liver tissue remained below that for DS plans although a deterioration was detected compared to the two-field PBS plans. These cases show that the better target coverage and the worse organ at risk sparing of DS plans revealed by the hepatocellular carcinoma study are mainly attributable to the delivery technique and only to a small extent to the difference in the number of irradiation fields. In some clinical cases, a third irradiation field might lead to an unacceptable high exposure of organs at risk so that two-field PBS plans represent the only treatment option.

6.3 Summary - Usage of computer generated 4D CTs in interplay studies

The applicability of computer generated 4D CTs for systematic interplay effect analyses in PBS was tested for the 4D XCAT phantom, a whole-body computer model of human anatomy and physiology. The feasibility study investigated the dependence of interplay effects on the target size for one patient with different clinical target volumes. A generated 4D XCAT chest CT was imported to RayStation 7. In addition to a set with typical clinical CT-numbers for treatment planning, a copy with image-values representing organ IDs was imported. The segmentation tools of RayStation thereby allowed to perfectly contour 48 anatomical structures in all phases of the 4D CT copy and map them to the original 4D CT. The structures subsequently guided the deformable image registration. Five spherical liver tumours with radii R of 1 cm to 5 cm were manually delineated as clinical target volumes in the end-exhale phase and mapped to all other phases by deformable image registration. The prescription was the same as for the hepatocellular carcinoma treatment planning study. Conventional PBS plans were used for the proof of concept study instead of 4D robustly optimised PBS plans to save computation time. For comparison purposes, identical gantry and couch angles as well as optimisation objectives and constraints were chosen for all targets. 50 interplay scenarios per treatment plan were evaluated based on the Pencil Beam dose algorithm implemented in RaySearch.

All targets exhibited an average tumour motion amplitude of 2.2 cm. The homogeneity index increased with increasing clinical target volume from $11.5\% \pm 3.0\%$ (average value \pm SD) for $R=1$ cm to $22.3\% \pm 2.4\%$ for $R=5$ cm. This trend might be a result of the growing probability for interplay effects with an increase in the number of spots and the irradiation time. However, the percentage over- and underdosage did not follow the trend. The maximum was detected for $R=5$ cm ($V_{107/95}=39.9\% \pm 5.4\%$), but the minimum was observed for $R=3$ cm ($V_{107/95}=22.3\% \pm 7.6\%$). The results demonstrate the complexity of the manifestation of interplay effects and indicate that it might not be possible to determine the interplay effect level based on parameters like the target size and the motion amplitude. Future investigations are necessary to confirm this hypothesis and will show whether individual, time consuming interplay effect evaluations are still required for each patient. The application of computer generated 4D CTs can facilitate these studies and furthermore allows investigation of motion mitigation strategies in a systematic approach.

The proof of concept study showed that simulated CT data of a 4D XCAT phantom provide realistic image quality without CT artefacts, enable perfect contouring and can minimise deformable image registration uncertainties in 4D studies.

6.4 Outlook on future research

The study on 4D robust optimisation has shown that local cold and hot spots observed in single treatment fractions due to interplay effects were averaged out during the treatment course. The biological effect on tumour control and normal tissue complications, however, still needs to be investigated. For this purpose, future studies should consider biological tumour control probability models for the target and normal tissue complication probability models for all relevant organs at risk in the interplay effect evaluation. A tumour control probability model and two normal tissue complication probability models are provided in RayStation 7 [169]. The tumour control probability model and the first normal tissue complication probability model are based on cell survival curves and Poisson statistics, whereas the Lyman-Kutcher-Burman model presented in Section 4.2.6 forms the basis for the second normal tissue complication probability model. For protons, the tumour control probability and the normal tissue complication probability are functions of the relative biological effectiveness (RBE)-weighted dose. The underlying biological parameters were taken from published clinical and pre-clinical studies. They are only valid for the endpoint and tissue for which they have been determined and it might become necessary to use data from more recent studies. An aspect which should be addressed in future research is whether a minimum dose exists for a single fraction to ensure eradication of tumour cells. If this is the case, the prescribed fraction dose could be adjusted so that typical underdosage resulting from interplay effects does not lead to an violation of the determined limit.

Another issue that needs to be explored is the effect of irregular respiration. The respiratory pattern of a patient during treatment might be different from that at the time the 4D CT scan was acquired. Changes can occur between treatment fractions or even within a fraction for example due to normal fluctuations, acute breathing problems, coughing, anxiety or stress. A variation in amplitude and/or the time period will affect the interplay pattern. The extent could be studied based on the 4D XCAT phantom. To this end, virtual patients with different anatomies should be evaluated. For each motion amplitude, another 4D CT must be created, while the respiratory period can be modified in the user interface of the interplay effect routine. For the study of irregular respiratory cycles within a fraction, minor modifications of the routine are required. For example, a sequence of different cycle lengths could be passed instead of a fixed value and each respiratory cycle could be referred to a specific 4D CT that determines the organ motion and is used to calculate the respective evaluation dose.

In the near future, first hepatocellular carcinoma patients will be treated with DS at WPE. Weekly 4D CTs of these patients could provide valuable information on realistic variations in tumour motion. In addition, irradiation pauses due to interlocks of the machine can be included in the study.

As discussed in the previous chapter, the usage of computer generated 4D CTs will enable systematic investigations of interplay effects. The influence of the motion amplitude, the location of the tumour, the tumour size or the homogeneity of the surrounding tissue can be assessed with the 4D XCAT phantom by altering only a single parameter and keeping the other influencing factors constant. The outcome of such studies will show whether the magnitude of interplay effects can be estimated based on the knowledge of these parameters or whether the optimal treatment option for a patient still has to be determined based on individual interplay effect evaluations.

Future treatment planning system versions and improved hardware might allow to include more scenarios in the 4D robust optimisation so that uncertainties of the time structure can be considered under realistic clinical settings.

Possibilities to decrease the computation time of the routine are not to select the dose volume histogram plot option and to reduce the number of interplay scenarios at the expense of accuracy. If the focus of a study is for example on the exposure of organs at risk, a lower number of interplay scenarios might be sufficient since standard deviations of dose volume histogram parameters were much lower for organs at risk than for the target.

7 Conclusion

Within the scope of this dissertation, a customised routine simulating interplay effects in pencil beam scanning proton therapy was successfully implemented in treatment planning. The routine enables retrospective four-dimensional dose analyses using irradiation log files and, owing to a specially developed beam delivery time model, also prospective evaluations without the need for beam-time. In daily clinical practice, the tool can extend the patient-specific quality assurance into the temporal domain and may serve to assess the treatment plan robustness against intra-fraction motion as well as to reconstruct the 4D dose distribution after a treatment fraction. Thus, adaptations of a treatment plan can be made during the treatment course if the actually delivered dose deviates substantially from the planned dose distribution.

The suitability of the routine for clinical use was thoroughly investigated in an end-to-end test. Dynamic measurements with a two-dimensional ionisation chamber array and the corresponding irradiation log files demonstrated the validity of the 4D calculations and the beam delivery time model. It was shown that an inaccurate modelling of the beam sequence, e.g. assuming constant values for the delivery time, scanning speed and energy layer switching time, can cause considerable deviations between simulation and measurement. Therefore, 4D dynamic dose calculation tools should not be applied for clinical use in a proton therapy facility without making adequate adjustments to the respective beam delivery system and quality assurance tests. Fluctuations of the beam delivery parameters, such as the energy layer switching time, and different phase shifts between patient motion and start of the field application have to be considered in simulations in order to obtain a realistic estimate of the extent of possible interplay effects.

A commercial 4D robust optimisation algorithm was evaluated for the use in scanned proton therapy of hepatocellular carcinomas based on the developed 4D dynamic dose calculation routine. For the investigated hepatocellular carcinoma patients, 4D robust optimisation did not significantly improve the dose homogeneity of the clinical target volume in the presence of organ motion compared to the conventional planning target volume concept, but the same level of robustness was achieved for lower organ at risk doses. Both, conventional and 4D robustly optimised PBS plans, exhibited motion-induced over- and underdosage for large tumour motion amplitudes in a single fraction. Therefore, it is advisable to make use of supplementary motion mitigation techniques for strongly moving tumours.

In contrast to PBS, DS did not cause any over- or underdosage and yielded homogeneous dose distributions comparable to the static case within the clinical target volume. Although DS is the method of choice in terms of target coverage, it is often not a clinical option, particularly for large clinical target volumes, because of the less conformal dose distribution which causes organ at risk dose limits to be exceeded. In addition, most of the recently built particle therapy centres are not equipped with DS technique. Considering that 4D robustly optimised PBS plans yielded (1) homogeneous 4D dynamically accumulated dose distributions looking at the sum dose of all fractions, (2) on average a minimum dose above 90% within a single fraction and (3) the lowest organ at risk doses compared to DS and conventional PBS plans, they represent a good alternative to DS for many hepatocellular carcinoma patients.

Apart from the increased computation time, 4D robust optimisation offers nothing but advantages compared to conventional PBS treatment planning so that it could become a standard in future treatments of moving targets. Since it does not prolong the setup or treatment time, it can be combined with commonly used motion mitigation techniques, such as gating or rescanning, without exceeding a tolerable time limit for the patient or further decreasing the patient throughput. Moreover, no additional hardware is required.

The size of the investigated patient cohort was too small to statistically analyse different patient subgroups and the benefit of 4D robust optimisation could be different for patients treated under other conditions, e.g. without abdominal compression or lipiodol contrast. This needs to be further investigated in a larger patient cohort. Improvements in computing power and software could facilitate such large-scale studies in the next few years. Further, they might allow to increase the number of optimisation scenarios which is important in view of recent 4D robust optimisation approaches including uncertainties of time structures.

In a proof of concept study, the suitability of computer generated 4D CTs for interplay effect evaluations based on 4D XCAT phantoms was confirmed. The generated CT images were artefact-free and enabled perfect contouring due to unique organ IDs. By including a large number of controlling structures, deformable image registration uncertainties could be reduced. The usage of computer generated 4D CTs will allow to model a broad spectrum of clinical cases and systematically investigate different influencing factors in upcoming research projects.

The feasibility study exemplarily examined the magnitude of interplay effects as a function of the target size based on the homogeneity index and the percentage over- and underdosage. While the homogeneity index increased with growing target volume, the percentage over- and underdosage did not reproduce this trend. It could be a first indication that the categorisation of patients according to tumour motion, size or tissue homogeneity into different treatment groups might not be possible. Due to the high complexity, the developed 4D dose calculation routine will most likely still be needed for individual patient evaluations in the future.

A Appendix

A.1 Scripting in RayStation

Within the scope of this thesis, several scripts have been written in Python programming language to facilitate extensive calculations in RayStation, e.g. 4DD and 4DDD calculations. In the following, the plan robustness evaluation script serves as an example for a developed program code since it is shorter and less complex. In addition to the Python file (see A.1.1), an Extensible Application Markup Language (XAML) file (see A.1.2) has been created to provide a graphical user interface (GUI). XAML is a user interface markup language developed by Microsoft to define visible user interface elements, bind data to them and create events. Figure A.1 shows the GUI of the plan robustness evaluation script. The routine (see A.1.1) calculates the perturbed dose distributions for all possible permutations of the specified/selected range uncertainty, setup error and respiratory phases (the planning phase is by default included). However, the same setup error is used for the x, y and z direction (e.g. $dx=dy=dz=+2$ mm and $dx=dy=dz=-2$ mm) to keep the calculation time and the number of scenarios reasonably low. The calculated evaluation doses are stored in RayStation. Additionally, the CTV coverage, in form of V_{95} , the average dose of the CTV, the minimum dose of the CTV, the dose maximum of the CTV, the dose maximum of the patient's body and the dose maximum of the selected organs at risk are displayed by the console.

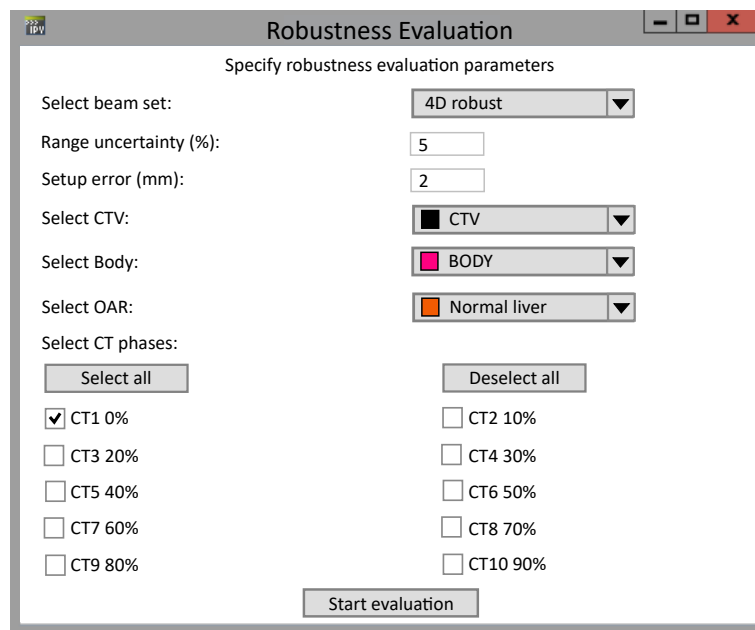


Figure A.1: Graphical user interface of the plan robustness evaluation script including TextBoxes, ComboBoxes, CheckBoxes and event buttons which allow the user to specify parameters, choose items from a predefined selection, select or deselect all items and start the evaluation.

A.1.1 Python file

Comment lines begin with a hashtag and are displayed in green. Numbers are coloured in red, strings in grey, keywords in blue, triple-quotes in orange and the name of a function in pink. The font style of class names is bold.

```

1  # Robustness evaluation considering range uncertainties, setup errors and 4D CT phases
2  # RayStation 5.99 Research
3  # execute with: execfile("R:/RayStation_scripts/Tina/Planning
   studies/RobustnessEvaluation4D_withGUI.py")
4
5  import wpf, time, System, clr, numpy
6  from System.Windows import Application, Window, Thickness, PropertyPath
7  from System.Windows.Input import Key, Keyboard
8  from System.Windows.Controls import *
9  from System.Windows.Data import Binding, BindingOperations, BindingMode
10 from System.Windows.Media import Color, SolidColorBrush
11 from System.Windows.Shapes import Rectangle
12 from connect import *
13
14 ''' GUI implementation '''
15 class MyWindow(Window):
16     def __init__(self, patient, plan, phases_4DCT):
17         wpf.LoadComponent(self, 'R:/RayStation_scripts/Tina/Planning
   studies/RobustnessEvaluation4DWindow.xaml')
18         self.patient = patient
19         self.plan = plan
20
21         # Items for beam set combobox:
22         Beamset_combobox_items = [self.create_beamset_combobox_item(BeamSet) for
   BeamSet in plan.BeamSets]
23         self.BeamsetComboBox.ItemsSource = Beamset_combobox_items
24         self.BeamsetComboBox.SelectedIndex = 0
25
26         # Items for ROI combobox:
27         OAR_combobox_items = [self.create_roi_combobox_item(roi) for roi in
   patient.Cases[case].PatientModel.RegionsOfInterest if
28                             roi.OrganData.OrganType=='OrganAtRisk']
29         CTV_combobox_items = [self.create_roi_combobox_item(ctv) for ctv in
   patient.Cases[case].PatientModel.RegionsOfInterest if
30                             ctv.OrganData.OrganType=='Target']
31         self.CTVComboBox.ItemsSource = CTV_combobox_items
32         self.BodyComboBox.ItemsSource = OAR_combobox_items
33         self.OARComboBox.ItemsSource = OAR_combobox_items
34         self.CTVComboBox.SelectedIndex = 0
35         self.BodyComboBox.SelectedIndex = 0
36         self.OARComboBox.SelectedIndex = 0
37
38         # split 4D CT phases into two columns
39         for idx, e in enumerate(phases_4DCT):
40             if idx % 2:
41                 column_idx = 1
42             else:
43                 column_idx = 0
44             self.MainGrid.RowDefinitions.Add(RowDefinition())
45             row_idx = self.MainGrid.RowDefinitions.Count-2
46             self.add_examination_checkbox(e, row_idx, column_idx)
47
48         self.selected_phases = []
49
50     def key_down_positive_decimal(self, sender, event):
51         if (event.Key >= Key.D0 and event.Key <= Key.D9) or (event.Key >=
   Key.NumPad0 and event.Key <= Key.NumPad9) \
52             or event.Key == Key.Tab or event.Key == Key.Enter:
53             event.Handled = False
54             return
55         if (event.Key == Key.Decimal or event.Key == Key.OemPeriod) and '.' not in
   sender.Text:
56             event.Handled = False
57             return
58         event.Handled = True
59
60     def start_evaluation_clicked(self, sender, event):
61
62         # beam set
63         if type(self.BeamsetComboBox.SelectedItem) == ComboBoxItem:
64             self.beamset_name =
65                 self.BeamsetComboBox.SelectedItem.Content.Children[0].Text

```

```

67         else:
68             self.beamset_name = None
69
70         # robustness parameters
71         self.RangeUncertainty = float(self.RangeTextBox.Text) # in %
72         self.SetupError = float(self.SetupTextBox.Text) # in mm
73
74         # CTV
75         if type(self.CTVComboBox.SelectedItem) == ComboBoxItem:
76             self.CTV_name = self.CTVComboBox.SelectedItem.Content.Children[1].Text
77         else:
78             self.CTV_name = None
79
80         # Body
81         if type(self.BodyComboBox.SelectedItem) == ComboBoxItem:
82             self.Body_name = self.BodyComboBox.SelectedItem.Content.Children[1].Text
83         else:
84             self.Body_name = None
85
86         # OAR
87         if type(self.OARComboBox.SelectedItem) == ComboBoxItem:
88             self.OAR_name = self.OARComboBox.SelectedItem.Content.Children[1].Text
89         else:
90             self.OAR_name = None
91
92         self.DialogResult = True
93
94     def checkbox_selection_changed(self, sender, event):
95         if sender.IsChecked:
96             self.selected_phases.append(sender.Content)
97         else:
98             self.selected_phases.remove(sender.Content)
99
100    def select_button_clicked(self, sender, event):
101        is_checked = sender.Content.startswith('Select') # content: "Select All" or
102        "Deselect All"
103        for c in self.MainGrid.Children:
104            if type(c) == CheckBox:
105                c.IsChecked = is_checked #-> all / no children are selected
106
107    def add_examination_checkbox(self, name, row_idx, column_idx):
108        cb = CheckBox()
109        cb.Content = name
110        cb.Margin = Thickness(15,5,5,5)
111        cb.SetValue(Grid.RowProperty, row_idx)
112        cb.SetValue(Grid.ColumnProperty, column_idx)
113        cb.Checked += self.checkbox_selection_changed
114        cb.Unchecked += self.checkbox_selection_changed
115        self.MainGrid.Children.Add(cb)
116        self.EvaluationButton.SetValue(Grid.RowProperty, row_idx+1)
117
118    def create_beamset_combobox_item(self, BeamSet):
119        cbi = ComboBoxItem()
120        sp = StackPanel()
121        sp.Orientation = Orientation.Horizontal
122        tb = TextBlock()
123        tb.Text = BeamSet.DicomPlanLabel
124        sp.Children.Add(tb)
125        cbi.Content = sp
126        return cbi
127
128    def create_roi_combobox_item(self, roi):
129        cbi = ComboBoxItem()
130        sp = StackPanel()
131        sp.Orientation = Orientation.Horizontal
132        rect = Rectangle()
133        rect.Width = 12
134        rect.Height = 12
135        rect.Fill = SolidColorBrush(Color.FromArgb(roi.Color.A, roi.Color.R,
136        roi.Color.G, roi.Color.B))
137        rect.Stroke = SolidColorBrush(Color.FromRgb(0, 0, 0))
138        rect.Margin = Thickness(0,0,2,0)

```

```

137         sp.Children.Add(rect)
138         tb = TextBlock()
139         tb.Text = roi.Name
140         tb.Margin = Thickness(2,0,0,0)
141         sp.Children.Add(tb)
142         cbi.Content = sp
143         return cbi
144
145     def set_binding(self, source, path, dependency_object, property, converter):
146         binding = Binding()
147         binding.Path = PropertyPath(path)
148         binding.Source = source
149         binding.Converter = converter
150         BindingOperations.SetBinding(dependency_object, property, binding)
151
152     ''' END GUI implementation '''
153
154     patient = get_current('Patient')
155     plan = get_current('Plan')
156     case = 'CASE 1'
157     dose = patient.Cases[case].TreatmentPlans[plan.Name].TreatmentCourse.TotalDose
158     planning_examination = plan.TreatmentCourse.TotalDose.OnDensity.FromExamination.Name
159     #name of the planning CT
160     phases_4DCT = [e.Examination.Name for e in
161                    patient.Cases[case].ExaminationGroups['4DCT'].Items
162                    if e.Examination.Name !=
163                    planning_examination]
164
165     dialog = MyWindow(patient, plan, phases_4DCT)
166     dialog.ShowDialog()
167
168     if dialog.DialogResult:
169         bs = plan.BeamSets[dialog.beamset_name]
170         fractions = bs.FractionationPattern.NumberOfFractions
171         prescribed_dose_cGy = bs.Prescription.DosePrescriptions[0].DoseValue
172         SelectedPhases = dialog.selected_phases
173         print('Prescribed dose (cGy)', prescribed_dose_cGy)
174         print('Number of fractions', fractions)
175         print('Beam set', bs.DicomPlanLabel)
176
177         ru = dialog.RangeUncertainty
178         x = dialog.SetupError
179         y = dialog.SetupError
180         z = dialog.SetupError
181         CTV = dialog.CTV_name
182         Body = dialog.Body_name
183         OAR = dialog.OAR_name
184         print('Range uncertainty (%):', ru, 'Setup error (mm):', x, 'CTV:', CTV,
185              'Body:', Body, 'OAR:', OAR)
186
187     # Range and isocentre shift
188     for pert in [float(ru/100.00), -float(ru/100.00)]: # +/- range uncertainty in
189         decimals
190         for i in [0.1, -0.1]: # shift in x, y and z; mm -> cm
191             x_l=x*i
192             y_l=y*i
193             z_l=z*i
194             isoshift = {'x':x_l, 'y':-z_l, 'z':y_l}
195             # compute perturbed dose on planning examination:
196             plan.BeamSets[0].ComputePerturbedDose(DensityPerturbation = pert,
197             IsocenterShift = isoshift,
198             AllowGridExpansion=False, ExaminationNames= [planning_examination],
199             FractionNumbers=[0])
200             fe = next(f for f in
201                    patient.Cases[case].TreatmentDelivery.FractionEvaluations
202                    if f.FractionNumber == 0)
203             doe = next(d for d in fe.DoseOnExaminations if d.OnExamination.Name ==
204                    planning_examination)
205             perturbed_dose = doe.DoseEvaluations[doe.DoseEvaluations.Count-1]
206             # Maximum dose CTV
207             Max = perturbed_dose.GetDoseStatistic(RoiName=CTV,
208             DoseType="Max")*fractions/float(prescribed_dose_cGy)*100

```

```

199     # Maximum dose body
200     MaxB = perturbed_dose.GetDoseStatistic(RoiName=Body,
201     DoseType="Max")*fractions/float(prescribed_dose_cGy)*100
202     # Maximum dose OAR
203     MaxOAR = perturbed_dose.GetDoseStatistic(RoiName=OAR,
204     DoseType="Max")*fractions
205     # Average dose CTV
206     Ave = perturbed_dose.GetDoseStatistic(RoiName=CTV,
207     DoseType="Average")*fractions
208     # Minimum dose CTV
209     Min = perturbed_dose.GetDoseStatistic(RoiName=CTV,
210     DoseType="Min")*fractions/float(prescribed_dose_cGy)*100
211     # V95 CTV
212     Iso95 = perturbed_dose.GetRelativeVolumeAtDoseValues(RoiName=CTV,
213     DoseValues=
214     [prescribed_dose_cGy/float(fractions)*0.95])[0]*100
215     print(planning_examination, 'range', int(pert *10000)/100.00, 'setup
216     error', int(x_1 *100)/100.00)
217     print('Body', 'Max (%)', MaxB)
218     print('CTV', 'Volume receiving at least 95% dose', int(Iso95
219     *100)/100.00, 'Max (%)', int(Max*10)/10.0,
220     'Min (%)', int(Min*100)/100.00, 'Average', int(Ave))
221     print(OAR, 'Max (cGy)', int(MaxOAR))
222
223     # deform perturbed dose to selected motion phases:
224     for phase in SelectedPhases:
225         # find DIR
226         for reg in patient.Cases[case].Registrations:
227             try:
228                 structure_registration = next(sr for sr in
229                 reg.StructureRegistrations if
230                 sr.FromExamination.Name == phase and
231                 sr.ToExamination.Name == planning_examination)
232             break
233         except:
234             pass
235     patient.Cases[case].MapDose(DoseDistribution=perturbed_dose,
236     StructureRegistration=structure_registration,
237     FractionNumber=0)
238     fe2 = next(f2 for f2 in
239     patient.Cases[case].TreatmentDelivery.FractionEvaluations if
240     f2.FractionNumber == 0)
241     doe2 = next(d2 for d2 in fe2.DoseOnExaminations if
242     d2.OnExamination.Name == phase)
243     mapped_dose = doe2.DoseEvaluations[doe2.DoseEvaluations.Count-1]
244     # Maximum dose CTV
245     Max2 = mapped_dose.GetDoseStatistic(RoiName= CTV,
246     DoseType="Max")*fractions/ float(prescribed_dose_cGy) *100
247     # Maximum dose body
248     MaxB2 = mapped_dose.GetDoseStatistic(RoiName= Body,
249     DoseType="Max")*fractions/ float(prescribed_dose_cGy) *100
250     # Maximum dose OAR
251     MaxOAR2 = mapped_dose.GetDoseStatistic(RoiName= OAR,
252     DoseType="Max")*fractions
253     # Average dose CTV
254     Ave2 = mapped_dose.GetDoseStatistic(RoiName= CTV,
255     DoseType="Average")*fractions
256     # Minimum dose CTV
257     Min2 = mapped_dose.GetDoseStatistic(RoiName= CTV,
258     DoseType="Min")*fractions/ float(prescribed_dose_cGy) *100
259     # V95 CTV
260     Iso95_2 = mapped_dose.GetRelativeVolumeAtDoseValues(RoiName= CTV,
261     DoseValues=
262     [prescribed_dose_cGy/float(fractions) * 0.95])[0] *100
263     print(phase, 'range', int(pert *10000)/100.00, 'setup error',
264     int(x_1 *10)/10.0)
265     print('Body', 'Max (%)', MaxB2)
266     print('CTV', 'Volume receiving at least 95% dose', int(Iso95_2
267     *100)/100.00, 'Max (%)', int(Max2*100)/100.00,
268     'Min (%)', int(Min2*100)/100.00, 'Average', int(Ave2))
269     print(OAR, 'Max (cGy)', int(MaxOAR2))

```

A.1.2 XAML file

In the depicted XAML file, tags are coloured in blue, attributes in red and strings in purple.

```

1  <Window xmlns="http://schemas.microsoft.com/winfx/2006/xaml/presentation"
2  xmlns:x="http://schemas.microsoft.com/winfx/2006/xaml"
3  Title="Robustness Evaluation"
4  Width="520"
5  SizeToContent="Height">
6  <Window.Resources>
7  <BooleanToVisibilityConverter x:Key="b2v" />
8  </Window.Resources>
9  <Grid Name="MainGrid">
10 <Grid.ColumnDefinitions>
11 <ColumnDefinition />
12 <ColumnDefinition />
13 </Grid.ColumnDefinitions>
14 <Grid.RowDefinitions>
15 <RowDefinition />
16 <RowDefinition />
17 <RowDefinition />
18 <RowDefinition />
19 <RowDefinition />
20 <RowDefinition />
21 <RowDefinition />
22 <RowDefinition />
23 <RowDefinition />
24 <RowDefinition />
25 <RowDefinition />
26 <RowDefinition />
27 <RowDefinition />
28 </Grid.RowDefinitions>
29 <TextBlock Grid.ColumnSpan="2"
30 Margin="5"
31 HorizontalAlignment="Center"
32 Text="Specify robustness evaluation parameters" />
33
34 <TextBlock Grid.Row="1"
35 Margin="15,5,5,5"
36 VerticalAlignment="Center"
37 Text="Select beam set:" />
38 <ComboBox Name="BeamsetComboBox"
39 Grid.Row="1"
40 Grid.Column="1"
41 Width="150"
42 Margin="15,5,5,5"
43 HorizontalAlignment="Left" />
44
45 <TextBlock Grid.Row="2"
46 Margin="15,5,5,5"
47 VerticalAlignment="Center"
48 Text="Range uncertainty [%]:" />
49 <TextBox Name="RangeTextBox"
50 Grid.Row="2"
51 Grid.Column="1"
52 Width="50"
53 Margin="15,5,5,5"
54 HorizontalAlignment="Left"
55 KeyDown="key_down_positive_decimal"
56 Text="5.5" />
57 <TextBlock Grid.Row="3"
58 Margin="15,5,5,5"
59 VerticalAlignment="Center"
60 Text="Setup error [mm]:" />
61 <TextBox Name="SetupTextBox"
62 Grid.Row="3"
63 Grid.Column="1"
64 Width="50"
65 Margin="15,5,5,5"
66 HorizontalAlignment="Left"
67 KeyDown="key_down_positive_decimal"
68 Text="2" />
69
70 <TextBlock Grid.Row="4"
71 Margin="15,5,5,5"
72 VerticalAlignment="Center"

```



```

73         Text="Select CTV:" />
74     <ComboBox Name="CTVComboBox"
75         Grid.Row="4"
76         Grid.Column="1"
77         Width="150"
78         Margin="15,5,5,5"
79         HorizontalAlignment="Left" />
80
81     <TextBlock Grid.Row="5"
82         Margin="15,5,5,5"
83         VerticalAlignment="Center"
84         Text="Select Body:" />
85     <ComboBox Name="BodyComboBox"
86         Grid.Row="5"
87         Grid.Column="1"
88         Width="150"
89         Margin="15,5,5,5"
90         HorizontalAlignment="Left" />
91
92     <TextBlock Grid.Row="6"
93         Margin="15,5,5,5"
94         VerticalAlignment="Center"
95         Text="Select OAR:" />
96     <ComboBox Name="OARComboBox"
97         Grid.Row="6"
98         Grid.Column="1"
99         Width="150"
100        Margin="15,5,5,5"
101        HorizontalAlignment="Left" />
102
103     <TextBlock Grid.Row="7"
104         Margin="15,5,5,5"
105         Text="Select CT phases:" />
106     <Button Grid.Row="8"
107         Width="100"
108         Margin="15,5,5,5"
109         HorizontalAlignment="Left"
110         Click="select_button_clicked"
111         Content="Select All" />
112     <Button Grid.Row="8"
113         Grid.Column="1"
114         Width="100"
115         Margin="15,5,5,5"
116         HorizontalAlignment="Left"
117         Click="select_button_clicked"
118         Content="Deselect All" />
119
120     <Button Name="EvaluationButton"
121         Grid.Row="7"
122         Grid.ColumnSpan="2"
123         Width="120"
124         Margin="5"
125         Click="start_evaluation_clicked"
126         Content="Start evaluation" />
127 </Grid>
128 </Window>

```


References

- [1] World Health Organization. Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016. Geneva, Switzerland; 2018. URL: http://www.who.int/healthinfo/global_burden_disease/estimates/en/ (Last accessed: 4 September 2018).
- [2] Bush DA, Kayali Z, Grove R, Slater JD. The safety and efficacy of highdose proton beam radiotherapy for hepatocellular carcinoma: a phase 2 prospective trial. *Cancer* 2011;117:3053–9.
- [3] Bush DA, Cheek G, Zaheer S, Wallen J, Mirshahidi H, Katerelos A, Grove R, Slater JD. High-dose hypofractionated proton beam radiation therapy is safe and effective for central and peripheral early-stage non-small cell lung cancer: results of a 12-year experience at Loma Linda University Medical Center. *Int J Radiat Oncol Biol Phys* 2013;86:964-8.
- [4] Koyama S, Tsujii H. Proton Beam Therapy with High-Dose Irradiation for Superficial and Advanced Esophageal Carcinomas. *Clin Cancer Res* 2003;9:3571-7.
- [5] Foote RL, Stafford SL, Petersen IA, Pulido JS, Clarke MJ, Schild SE, Garces YI, Olivier KR, Miller RC, Haddock MG, Yan E, Laack NN, Arndt CA, Buskirk SJ, Miller VL, Brent CR, Kruse JJ, Ezzell GA, Herman MG, Gunderson LL, Erlichman C, Diasio RB. The clinical case for proton beam therapy. *Radiat Oncol* 2012;7:174.
- [6] Phillips MH, Pedroni E, Blattmann H, Boehringer T, Coray A, Scheib S. Effects of respiratory motion on dose uniformity with a charged particle scanning method. *Phys Med Biol* 1992;37:223-34.
- [7] Lambert J, Suchowerska N, McKenzie DR, Jackson M. Intrafractional motion during proton beam scanning. *Phys Med Biol* 2005;50:4853-62.
- [8] Bert C, Grözinger SO, Rietzel E. Quantification of interplay effects of scanned particle beams and moving targets. *Phys Med Biol* 2008;53:2253-65.
- [9] Seco J, Robertson D, Tromov A, Paganetti H. Breathing interplay effects during proton beam scanning: simulation and statistical analysis. *Phys Med Biol* 2009;54:N283-94.
- [10] Knopf AC, Stützer K, Richter C, Rucinski A, da Silva J, Phillips J, Engelsman M, Shimizu S, Werner R, Jakobi A, Göksel O, Zhang Y, Oshea T, Fast M, Perrin R, Bert C, Rinaldi I, Korevaar E, McClelland J. Required transition from research to clinical application: Report on the 4D treatment planning workshops 2014 and 2015. *Phys Med* 2016;32:874–82.
- [11] De Ruyscher D, Sterpin E, Haustermans K, Depuydt T. Tumour movement in proton therapy: solutions and remaining questions: a review. *Cancers (Basel)* 2015;7:1143-53.
- [12] Bert C and Durante M. Motion in radiotherapy: particle therapy. *Phys Med Biol* 2011;56:R113–44.
- [13] Bert C, Graeff C, Riboldi M, Nill S, Baroni G, Knopf AC. Advances in 4D treatment planning for scanned particle beam therapy - report of dedicated workshops. *Technol Cancer Res Treat* 2014;13:485-95.
- [14] Kubiak T. Particle therapy of moving targets - the strategies for tumour motion monitoring and moving targets irradiation. *Br J Radiol* 2016;89:20150275.
- [15] Chang JY, Zhang X, Knopf A, Li H, Mori S, Dong L, Lu HM, Liu W, Badiyan SN, Both S, Meijers A, Lin L, Flampouri S, Li Z, Umegaki K, Simone CB, Zhu XR. Consensus Guidelines for Implementing Pencil-Beam Scanning Proton Therapy for Thoracic Malignancies on Behalf of the PTCOG Thoracic and Lymphoma Subcommittee. *Int J Radiat Oncol Biol Phys* 2017;99:41-50.

-
- [16] Graeff C. Motion mitigation in scanned ion beam therapy through 4D-optimization. *Phys Med* 2014;30:570-7.
- [17] Liu W, Schild SE, Chang JY, Liao Z, Chang YH, Wen Z, Shen J, Stoker JB, Ding X, Hu Y, Sahoo N, Herman MG, Vargas C, Keole S, Wong W, Bues M. Exploratory Study of 4D versus 3D Robust Optimization in Intensity Modulated Proton Therapy for Lung Cancer. *Int J Radiat Oncol Biol Phys* 2016;95:523-33.
- [18] Yu J, Zhang X, Liao L, Li H, Zhu R, Park PC, Sahoo N, Gillin M, Li Y, Chang JY, Komaki R, Lin SH. Motion-robust intensity-modulated proton therapy for distal esophageal cancer. *Med Phys* 2016;43:1111-8.
- [19] El-Serag HB and Rudolph KL. Hepatocellular Carcinoma: Epidemiology and Molecular Carcinogenesis. *Gastroenterology* 2007;132:2557-76.
- [20] Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;127:2893-917.
- [21] España S and Paganetti H. Uncertainties in planned dose due to the limited voxel size of the planning CT when treating lung tumors with proton therapy. *Phys Med Biol* 2011;56:3843-56.
- [22] Grassberger C, Daartz J, Dowdell S, Ruggieri T, Sharp G, Paganetti H. Quantification of proton dose calculation accuracy in the lung. *Int J Radiat Oncol Biol Phys* 2014;89:424-30.
- [23] Titt U, Sell M, Unkelbach J, Bangert M, Mirkovic D, Oelfke U, Mohan R. Degradation of proton depth dose distributions attributable to microstructures in lung-equivalent material. *Med Phys* 2015;42:6425-32.
- [24] Taylor PA, Kry SF, Followill DS. Pencil Beam Algorithms Are Unsuitable for Proton Dose Calculations in Lung. *Int J Radiat Oncol Biol Phys* 2017;99:750-6.
- [25] Baumann KS, Witt M, Weber U, Engenhardt-Cabillic R, Zink K. An efficient method to predict and include Bragg curve degradation due to lung-equivalent materials in Monte Carlo codes by applying a density modulation. *Phys Med Biol* 2017;62:3997-4016.
- [26] Schätti A, Zakova M, Meer D, Lomax AJ. Experimental verification of motion mitigation of discrete proton spot scanning by re-scanning. *Phys Med Biol* 2013;58:8555-72.
- [27] Schätti A, Zakova M, Meer D, Lomax AJ. The effectiveness of combined gating and re-scanning for treating mobile targets with proton spot scanning. An experimental and simulation-based investigation. *Phys Med Biol* 2014;59:3813-28.
- [28] Kraus KM, Heath E, Oelfke U. Dosimetric consequences of tumour motion due to respiration for a scanned proton beam. *Phys Med Biol* 2011;56:6563-81.
- [29] Zeng C, Plastaras JP, Tochner ZA, White BM, Hill-Kayser CE, Hahn SM, Both S. Proton pencil beam scanning for mediastinal lymphoma: the impact of interplay between target motion and beam scanning. *Phys Med Biol* 2015;60:3013-29.
- [30] Inoue T, Widder J, van Dijk LV, Takegawa H, Koizumi M, Takashina M, Usui K, Kurokawa C, Sugimoto S, Saito AI, Sasai K, Van't Veld AA, Langendijk JA, Korevaar EW. Limited impact of setup and range uncertainties, breathing motion, and interplay effects in robustly optimized intensity modulated proton therapy for stage III non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 2016;96:661-9.
- [31] Gemmel A, Rietzel E, Kraft G, Durante M, Bert C. Calculation and experimental verification of the RBE-weighted dose for scanned ion beams in the presence of target motion. *Phys Med Biol* 2011;56:7337-51.

- [32] Bert C, Richter D, Durante M, Rietzel E. Scanned carbon beam irradiation of moving films: comparison of measured and calculated response. *Radiat Oncol* 2012;7:55.
- [33] Krieger M, Klimpki G, Fattori G, Hrbacek J, Oxley D, Safai S, Weber DC, Lomax AJ, Zhang Y. Experimental validation of a deforming grid 4D dose calculation for PBS proton therapy. *Phys Med Biol* 2018;63:055005.
- [34] Slopsema R. Beam Delivery Using Passive Scattering. In: Paganetti H, editor. *Proton Therapy Physics*. CRC Press, Taylor & Francis Group, Boca Raton London New York; 2012, p.126-53.
- [35] Newhauser WD and Zhang R. The physics of proton therapy. *Phys Med Biol* 2015;60:R155–R209.
- [36] Gottschalk B. Physics of Proton Interactions in Matter. In: Paganetti H, editor, *Proton Therapy Physics*. CRC Press, Taylor & Francis Group, Boca Raton London New York; 2012, p.20-59.
- [37] International Commission on Radiation Units and Measurements. Report 85: Fundamental Quantities and Units for Ionizing Radiation. *J ICRU* 2011;11:1-31.
- [38] Bethe H. Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie. *Annalen der Physik* 1930;5:325-400.
- [39] Bloch F. Zur Bremsung rasch bewegter Teilchen beim Durchgang durch Materie. *Annalen der Physik* 1933;16:285.
- [40] Fano U. Penetration of Protons, Alpha Particles, and Mesons. *Annu Rev Nucl Sci* 1963;13:1-66.
- [41] Andreo P, Wulff J, Burns DT, Palmans H. Consistency in reference radiotherapy dosimetry: resolution of an apparent conundrum when $(60)\text{Co}$ is the reference quality for charged-particle and photon beams. *Phys Med Biol* 2013;58:6593-621.
- [42] Nüsslin F. Meßmethoden für die Dosimetrie. In: Schlegel W and Bille J, editors. *Medizinische Physik 2. Medizinische Strahlenphysik*. Springer, Heidelberg; 2002, p.103-21.
- [43] Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys Med Biol* 2012;57:R99-117.
- [44] Oelfke U. Physikalische Wechselwirkungen ionisierender Strahlung mit Materie. In: Schlegel W and Bille J, editors. *Medizinische Physik 2. Medizinische Strahlenphysik*. Springer, Heidelberg; 2002, p.33-64.
- [45] Pedroni E. Treatment delivery systems: pencil beam scattering. In: Delaney TF and Kooy HM, editors. *Proton and charged particle radiotherapy*. Lippincott Williams & Wilkins, Philadelphia; 2008, p.40-49.
- [46] Gazda MJ and Coia LR. Principles of radiation therapy. In: Pazdur R, editor. *Cancer Management: A Multidisciplinary Approach, Medical, Surgical & Radiation Oncology*. CMP Healthcare Media; 2004, p.9-19.
- [47] Paganetti H and van Luijk P. Biological considerations when comparing proton therapy with photon therapy. *Semin Radiat Oncol* 2013;23:77-87.
- [48] Hall EJ and Giaccia AJ. Cell survival curves. In: Hall EJ and Giaccia AJ, editors. *Radiobiology for the Radiologist*. J.B. Lippincott Williams & Wilkins, Philadelphia; 2006, p.30-46.
- [49] Knedlitschke G and Weibezahn KF. Biologische Grundlagen der Strahlenwirkung. In: Schlegel W and Bille J, editors. *Medizinische Physik 2. Medizinische Strahlenphysik*. Springer, Heidelberg; 2002, p.123-135.

- [50] Münter M and Weber K. Strahlenbiologie. In: Reiser M, Kuhn F, Debus J, editors. *Duale Reihe Radiologie*. 3. Thieme, Stuttgart; 2011, p.37-62.
- [51] Niemierko A. Reporting and analyzing dose distributions: a concept of equivalent uniform dose. *Med Phys* 1997;24:103-10.
- [52] IBA particle therapy. Technical report 10135. Essen PT project interface building document. Louvain-La-Neuve, Belgium; 2006.
- [53] Schippers M. Proton Accelerators. In: Paganetti H, editor, *Proton Therapy Physics*. CRC Press, Taylor & Francis Group, Boca Raton London New York; 2012, p.62-102.
- [54] Schippers M. Proton beam production and dose delivery techniques. In: Das IJ and Paganetti H, editors. *Principles and practices of proton beam therapy*. Madison: Medical Physics publishing, *Med Phys Monograph* 37; 2015, p.129-63.
- [55] IBA particle therapy. The PROTEUS 235 Product Description - M-ID 1889 Rev. A. Louvain-La-Neuve, Belgium; 2009.
- [56] Marchand B, Prieels D, Bauvir B, S epulchre R, G erard M. IBA proton pencil beam scanning: an innovative solution for cancer treatment. *Proceedings of EPAC 2000*, Vienna, Austria.
- [57] B aumer C, Ackermann B, Hillbrand M, Kaiser FJ, Koska B, Latzel H, L uhr A, Menkel S, Timmermann B. Dosimetry intercomparison of four proton therapy institutions in Germany employing spot scanning. *Z Med Phys* 2017;27:80-5.
- [58] Kr amer M and Durante M. Ion beam transport calculations and treatment plans in particle therapy. *Eur Phys J D* 2010;60:195–202.
- [59] RaySearch Laboratories AB. *RayStation 6 Reference Manual RSL-D-RS-6.0-REF-EN-2.0-2017-05-05*. Stockholm, Sweden; 2017.
- [60] RaySearch Laboratories AB. *RayStation 5 Scripting Guideline RSL-DRS- 5.0-SG-EN-2.0-2016-04-19*. Stockholm, Sweden; 2016.
- [61] National Institute of Standards and Technology (NIST). X-ray mass attenuation coefficients, chapter 2. URL: <http://physics.nist.gov/PhysRefData/XrayMassCoef/chap2.html> (Last accessed: 3 August 2018).
- [62] Dowsett DJ, Kenny PA, Johnston RE. Computed tomography. In: Dowsett DJ, Kenny PA, Johnston RE, editors. *The physics of diagnostic imaging*. Hodder Arnold, London, 2nd edition; 2006, p.381-436.
- [63] Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys Med Biol* 1996;41:111-24.
- [64] Brenner DJ, Hall EJ, Phil D. Computed Tomography — An Increasing Source of Radiation Exposure. *N Engl J Med* 2007;357:2277-84.
- [65] International Commission on Radiation Units and Measurements Report 78: Prescribing, recording, and reporting proton-beam therapy. *J ICRU* 2007;7:1602.
- [66] Weistrand O and Svensson S. The ANACONDA algorithm for deformable image registration in radiotherapy. *Med Phys* 2015;42:40-53.
- [67] Garc a-Moll a R, de Marco-Blancas N, Bonaque J, Vidueira L, L opez-Tarjuelo J, Perez-Calatayud J. Validation of a deformable image registration produced by a commercial treatment planning system in head and neck. *Phys Med* 2015;31:219-23.

- [68] Kadoya N, Nakajima Y, Saito M, Miyabe Y, Kurooka M, Kito S, Fujita Y, Sasaki M, Arai K, Tani K, Yagi M, Wakita A, Tohyama N, Jingu K. Multi-institutional validation study of commercially available deformable image registration software for thoracic images. *Int J Radiation Oncol Biol Phys* 2016;96:422-31.
- [69] Zhang L, Wang Z, Shi C, Long T, Xu XG. The impact of robustness of deformable image registration on contour propagation and dose accumulation for head and neck adaptive radiotherapy. *J Appl Clin Med Phys* 2018;19:185-94.
- [70] Gottschalk B, Cascio EW, Daartz J, Wagner MS. On the nuclear halo of a proton pencil beam stopping in water. *Phys Med Biol* 2015;60:5627-54.
- [71] Eyges L. Multiple Scattering with Energy Loss. *Phys Rev* 1948;74:1534-5.
- [72] Soukup M, Fippel M, Alber M. A pencil beam algorithm for intensity modulated proton therapy derived from Monte Carlo simulations. *Phys Med Biol* 2005;50:5089-104.
- [73] Goudsmit S and Saunderson JL. Multiple Scattering of Electrons. *Phys Rev* 1940;57:24-9.
- [74] Goudsmit S and Saunderson JL. Multiple Scattering of Electrons. II. *Phys Rev* 1940;58:36-42.
- [75] International Commission on Radiation Units and Measurements. Report 63: Nuclear Data for Neutron and Proton Radiotherapy and for Radiation Protection. Bethesda, U.S.A.; 2000.
- [76] Fippel M and Soukup M. A Monte Carlo dose calculation algorithm for proton therapy. *Med Phys* 2004;31:2263-73.
- [77] RaySearch Laboratories AB. RayStation 6 User Manual RSL-D-RS-6.0-USM-EN-1.0-2016-12-22. Stockholm, Sweden; 2016.
- [78] RaySearch Laboratories AB. RayStation 6 A Guide To Optimization In RayStation RSL-D-RS-6.0-OPT-EN-1.0-2016-12-22. Stockholm, Sweden; 2016.
- [79] Fredriksson A. Robust optimization of radiation therapy accounting for geometric uncertainty. Doctoral thesis, KTH Engineering Sciences, Stockholm, Sweden, 2013.
- [80] Bokrantz R. Multicriteria optimization for managing tradeoffs in radiation therapy treatment planning. Doctoral thesis, KTH Engineering Sciences, Stockholm, Sweden, 2013.
- [81] Gill PE, Murray W, Saunders MA. SNOPT: An SQP Algorithm for Large-Scale Constrained Optimization. *SIAM REVIEW* 2005;47:99-131.
- [82] Gill PE and Wong E. Sequential quadratic programming methods. In *Mixed Integer Nonlinear Programming*. Springer, New York; 2012, p.147-224.
- [83] Lomax A J. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: The potential effects of calculational uncertainties. *Phys Med Biol* 2008;53:1027-42.
- [84] Lomax A J. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: The potential effects of inter-fraction and inter-field motions. *Phys Med Biol* 2008;53:1043-56.
- [85] Mohan R and Sahoo N. Uncertainties in proton therapy: Their impact and management. In: Das IJ and Paganetti H., editors. *Principles and practices of proton beam therapy*. Madison: Medical Physics publishing, Med Phys Monograph 37; 2015, p.595-622.
- [86] Schaffner B and Pedroni E. The precision of proton range calculations in proton radiotherapy treatment planning: experimental verification of the relation between CT-HU and proton stopping power. *Phys Med Biol* 1998;43:1579-92.

- [87] Unkelbach J, Craft D, Gorissen BL, Bortfeld T. Treatment plan optimization in proton therapy. In: Das IJ and Paganetti H, editors. Principles and practices of proton beam therapy. Madison: Medical Physics publishing, Med Phys Monograph 37; 2015, p.623-46.
- [88] Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. *Med Phys* 2011;38:1672-84.
- [89] Keall PJ, Mageras GS, Balter JM, Emery RS, Forster KM, Jiang SB, Kapatoes JM, Low DA, Murphy MJ, Murray BR, Ramsey CR, Van Herk MB, Vedam SS, Wong JW, Yorke E. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. *Med Phys* 2006;33:3874-900.
- [90] Liang Z, Liu H, Xue J, Hu B, Zhu B, Li Q, Zhang S, Wu G. Evaluation of the intra- and interfractional tumor motion and variability by fiducial-based real-time tracking in liver stereotactic body radiation therapy. *J Appl Clin Med Phys* 2018;19:94-100.
- [91] Sonke JJ, Lebesque J, van Herk M. Variability of four-dimensional computed tomography patient models. *Int J Radiat Oncol Biol Phys* 2008;70:590-8.
- [92] Britton KR, Starkschall G, Tucker SL, Pan T, Nelson C, Chang JY, Cox JD, Mohan R, Komaki R. Assessment of gross tumor volume regression and motion changes during radiotherapy for non-small-cell lung cancer as measured by four-dimensional computed tomography. *Int J Radiat Oncol Biol Phys* 2007;68:1036-46.
- [93] Pan CC, Kavanagh BD, Dawson LA, Li XA, Das SK, Miften M, Ten Haken RK. Radiation-associated liver injury. *Int J Radiat Oncol Biol Phys* 2010;76:S94-100.
- [94] Hu Y, Zhou YK, Chen YX, Zeng ZC. Magnitude and influencing factors of respiration-induced liver motion during abdominal compression in patients with intrahepatic tumors. *Radiat Oncol* 2017;12:9.
- [95] Case RB, Moseley DJ, Sonke JJ, Eccles CL, Dinniwell RE, Kim J, Bezjak A, Milosevic M, Brock KK, Dawson LA. Interfraction and intrafraction changes in amplitude of breathing motion in stereotactic liver radiotherapy. *Int J Radiat Oncol Biol Phys* 2010;77:918-25.
- [96] Kitamura K, Shirato H, Seppenwoolde Y, Shimizu T, Kodama Y, Endo H, Onimaru R, Oda M, Fujita K, Shimizu S, Miyasaka K. Tumor location, cirrhosis, and surgical history contribute to tumor movement in the liver, as measured during stereotactic irradiation using a real-time tumor-tracking radiotherapy system. *Int J Radiat Oncol Biol Phys* 2003;56:221-8.
- [97] Philips Healthcare. White paper. Respiratory motion management for CT. Netherlands; 2013. URL: http://clinical.netforum.healthcare.philips.com/us_en/Explore/White-Papers/CT/Respiratory-motion-management-for-CT (Last accessed: 7 August 2018).
- [98] Grassberger C, Dowdell S, Lomax A, Sharp G, Shackelford J, Choi N, Willers H, Paganetti H. Motion interplay as a function of patient parameters and spot size in spot scanning proton therapy for lung cancer. *Int J Radiat Oncol Biol Phys* 2013;86:380-6.
- [99] Zhang Y, Huth I, Weber DC, Lomax AJ. A statistical comparison of motion mitigation performances and robustness of various pencil beam scanned proton systems for liver tumour treatments. *Radiother Oncol* 2018;128:182-8.
- [100] Pfeiler T, Bäumer C, Engwall E, Geismar D, Spaan B, Timmermann B. Experimental validation of a 4D dose calculation routine for pencil beam scanning proton therapy. *Z Med Phys* 2018;28:121-33.
- [101] IBA Dosimetry GmbH. MatriXX PT User's Guide P-09-005-510-002 01. Schwarzenbruck, Germany; 2012.

- [102] Lin L, Kang M, Solberg TD, Mertens T, Bäumer C, Ainsley CG, McDonough JE. Use of a novel two-dimensional ionization chamber array for pencil beam scanning proton therapy beam quality assurance. *J Appl Clin Med Phys* 2015;16:270–6.
- [103] CIRS Computerized Imaging Reference Systems, Inc. Dynamic Platform 008PL PB 020916. Norfolk, USA; 2013.
- [104] International Electrotechnical Commission (IEC). Technical report IEC 61217:2011. Radiotherapy equipment - coordinates, movements and scales. Geneva, Switzerland; 2011.
- [105] Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys* 1998;25:656–61.
- [106] Low DA, Dempsey JF. Evaluation of the gamma dose distribution comparison method. *Med Phys* 2003;30:2455–64.
- [107] Ezzell GA, Burmeister JW, Dogan N, LoSasso TJ, Mechalakos JG, Mihailidis D, Molineu A, Palta JR, Ramsey CR, Salter BJ, Shi J, Xia P, Yue NJ, Xiao Y. IMRT commissioning: multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. *Med Phys* 2009;36:5359–73.
- [108] Zhu XR, Poenisch F, Song X, Johnson JL, Ciangaru G, Taylor MB, Lii M, Martin C, Arjomandy B, Lee AK, Choi S, Nguyen QN, Gillin MT, Sahoo N. Patient-specific quality assurance for prostate cancer patients receiving spot scanning proton therapy using single-field uniform dose. *Int J Radiat Oncol Biol Phys* 2011;81:552–9.
- [109] Sorriaux J, Testa M, Paganetti H, Orban de Xivry J, Lee JA, Traneus E, Souris K, Vynckier S, Sterpin E. Experimental assessment of proton dose calculation accuracy in inhomogeneous media. *Phys Med* 2017;38:10–5.
- [110] International Atomic Energy Agency (IAEA). Dosimetry of Small Static Fields Used in External Beam Radiotherapy. Technical Reports Series No. 483. Vienna, Austria; 2017.
- [111] Nelder JA and Mead R. A simplex method for function minimization. *Comput J* 1965;7:308–13.
- [112] Fattori G, Klimpki G, Safai S, Weber D, Lomax A, Psoroulas S. TH-CD- 209-07: preliminary experimental comparison of spot- and continuous line scanning with or without rescanning for gated proton therapy. *Med Phys* 2016;43:3887.
- [113] Klimpki G, Zhang Y, Fattori G, Psoroulas S, Weber DC, Lomax A, Meer D. The impact of pencil beam scanning techniques on the effectiveness and efficiency of rescanning moving targets. *Phys Med Biol* 2018;63:145006.
- [114] Pankuch M, Mohammed N, Hecksel D, Laub S, Boyer S, Hartsell WF. Use of Protons for Radiation Therapy. In: Small W Jr, editor. *Clinical Radiation Oncology: Indications, Techniques, and Results*. John Wiley & Sons, Hoboken; 2017, p.115-140.
- [115] Zhao YT, Liu ZK, Wu QW, Dai JR, Zhang T, Jia AY, Jin J, Wang SL, Li YX, Wang WH. Observation of different tumor motion magnitude within liver and estimate of internal motion margins in postoperative patients with hepatocellular carcinoma. *Cancer Manag Res* 2017;9:839-48.
- [116] Liang Z, Liu H, Xue J, Hu B, Zhu B, Li Q, Zhang S, Wu G. Evaluation of the intra- and interfractional tumor motion and variability by fiducial-based real-time tracking in liver stereotactic body radiation therapy. *J Appl Clin Med Phys* 2018;19:94-100.
- [117] Pfeiler T, Ahmad Khalil D, Bäumer C, Blanck O, Chan M, Engwall E, Geismar D, Peters S, Plaude S, Spaan B, Wulff J and Timmermann B. 4D robust optimization in pencil beam scanning proton therapy for hepatocellular carcinoma. *J Phys Conf Ser*. Submitted 2018.

- [118] Pfeiler T, Ahmad Khalil D, Ayadi M, Bäumer C, Blanck O, Chan M, Engwall E, Geismar D, Peters S, Plaude S, Spaan B, Timmermann B, Wulff J. Motion effects in proton treatments of hepatocellular carcinoma - 4D robustly optimised pencil beam scanning plans vs double scattering plans. *Phys Med Biol* 2018;63 235006.
- [119] Altekruse SF, McGlynn KA, Reichman ME. Hepatocellular carcinoma incidence, mortality, and survival trends in the United States from 1975 to 2005. *J Clin Oncol* 2009;27:1485-91.
- [120] Parikh S and Hyman D. Hepatocellular cancer: a guide for the internist. *Am J Med* 2007;120:194-202.
- [121] Dionisi F, Widesott L, Lorentini S, Amichetti M. Is there a role for proton therapy in the treatment of hepatocellular carcinoma? A systematic review. *Radiother Oncol* 2014;111:1-10.
- [122] Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005;55:74-108.
- [123] Sugahara S, Oshiro Y, Nakayama H, Fukuda K, Mizumoto M, Abei M, Shoda J, Matsuzaki Y, Thono E, Tokita M, Tsuboi K, Tokuyue K. Proton beam therapy for large hepatocellular carcinoma. *Int J Radiat Oncol Biol Phys* 2010;76:460-6.
- [124] Kawashima M, Furuse J, Nishio T, Konishi M, Ishii H, Kinoshita T, Nagase M, Nihei K, Ogino T. Phase II study of radiotherapy employing proton beam for hepatocellular carcinoma. *J Clin Oncol* 2005;23:1839-46.
- [125] Tremosini S, Reig M, de Lope CR, Forner A, Bruix J. Treatment of early hepatocellular carcinoma: Towards personalized therapy. *Dig Liver Dis* 2010;42 Suppl 3:S242-8.
- [126] Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, Montalto F, Ammatuna M, Morabito A, Gennari L. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med* 1996;334:693-9.
- [127] Kalogeridi MA, Zygogianni A, Kyrgias G, Kouvaris J, Chatziioannou S, Kelekis N, Kouloulis V. Role of radiotherapy in the management of hepatocellular carcinoma: A systematic review. *World J Hepatol* 2015;7:101-12.
- [128] Lam VW, Ng KK, Chok KS, Cheung TT, Yuen J, Tung H, Tso WK, Fan ST, Poon RT. Risk factors and prognostic factors of local recurrence after radiofrequency ablation of hepatocellular carcinoma. *J Am Coll Surg* 2008;207:20-9.
- [129] Yao FY. Liver transplantation for hepatocellular carcinoma: beyond the Milan criteria. *Am J Transplant* 2008;8:1982-9.
- [130] Shin SW. The current practice of transarterial chemoembolization for the treatment of hepatocellular carcinoma. *Korean J Radiol* 2009;10:425-34.
- [131] Miraglia R, Pietrosi G, Maruzzelli L, Petridis I, Caruso S, Marrone G, Mamone G, Vizzini G, Luca A, Gridelli B. Efficacy of transcatheter embolization/chemoembolization (TAE/TACE) for the treatment of single hepatocellular carcinoma. *World J Gastroenterol* 2007;13:2952-5.
- [132] Komatsu S, Fukumoto T, Demizu Y, Miyawaki D, Terashima K, Sasaki R, Hori Y, Hishikawa Y, Ku Y, Murakami M. Clinical results and risk factors of proton and carbon ion therapy for hepatocellular carcinoma. *Cancer* 2011;117:4890-904.
- [133] Fajardo LF and Colby TV. Pathogenesis of veno-occlusive liver disease after radiation. *Arch Pathol Lab Med* 1980;104:584-8.

- [134] Dawson LA, Normolle D, Balter JM, McGinn CJ, Lawrence TS, Ten Haken RK. Analysis of radiation-induced liver disease using the Lyman NTCP model. *Int J Radiat Oncol Biol Phys* 2002;53:1422.
- [135] Nakayama H, Sugahara S, Tokita M, Fukuda K, Mizumoto M, Abei M, Shoda J, Sakurai H, Tsuboi K, Tokuyue K. Proton beam therapy for hepatocellular carcinoma: the University of Tsukuba experience. *Cancer* 2009;115:5499-506.
- [136] Hong TS, Wo JY, Yeap BY, Ben-Josef E, McDonnell EI, Blaszkowsky LS, Kwak EL, Allen JN, Clark JW, Goyal L, Murphy JE, Javle MM, Wolfgang JA, Drapek LC, Arellano RS, Mamon HJ, Mullen JT, Yoon SS, Tanabe KK, Ferrone CR, Ryan DP, DeLaney TF, Crane CH, Zhu AX. Multi-Institutional Phase II Study of High-Dose Hypofractionated Proton Beam Therapy in Patients With Localized, Unresectable Hepatocellular Carcinoma and Intrahepatic Cholangiocarcinoma. *J Clin Oncol* 2016;34:460-8.
- [137] Kim TH, Park JW, Kim YJ, Kim BH, Woo SM, Moon SH, Kim SS, Koh YH, Lee WJ, Park SJ, Kim JY, Kim DY, Kim CM. Phase I dose-escalation study of proton beam therapy for inoperable hepatocellular carcinoma. *Cancer Res Treat* 2015;47:34-45.
- [138] Chan MK, Lee V, Chiang CL, Lee FA, Law G, Sin NY, Siu KL, Wong FC, Tung SY, Luk H, Blanck O. Lipiodol versus diaphragm in 4D-CBCT-guided stereotactic radiotherapy of hepatocellular carcinomas. *Strahlenther Onkol* 2016;192:92-101.
- [139] Ormeloh T. Erstellung der stöchiometrischen CT-Kalibration für das Westdeutsche Protonentherapiezentrum Essen mit Ausblick auf die Kalibrationsmöglichkeiten von Multi-Energy-CT's. Master's thesis, Hochschule Hamm-Lippstadt, Hamm, Germany, 2017.
- [140] Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. *Med Phys* 2017;44:e43-e76.
- [141] Varadhan R, Karangelis G, Krishnan K, Hui S. A framework for deformable image registration validation in radiotherapy clinical applications. *J Appl Clin Med Phys* 2013;14:4066.
- [142] Berbeco RI, Nishioka S, Shirato H, Jiang SB. Residual motion of lung tumors in end-of-inhale respiratory gated radiotherapy based on external surrogates. *Med Phys* 2006;33:4149-56.
- [143] Koschack J. Standardabweichung und Standardfehler: der kleine, aber feine Unterschied. *Z Allg Med* 2008;84:258-60.
- [144] Eaton JW, Bateman D, Hauberg S, Wehbring R. GNU Octave version 4.2.2 manual: a high-level interactive language for numerical computations. 2018. URL: <https://www.gnu.org/software/octave/doc/v4.2.2/>; <https://octave.sourceforge.io/statistics/index.html> (Last accessed: 23 August 2018).
- [145] Lyman JT. Complication probability as assessed from dose-volume histograms. *Radiat Res Suppl* 1985;8:S13-9.
- [146] Kutcher GJ and Burman C. Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method. *Int J Radiat Oncol Biol Phys* 1989;16:1623-30.
- [147] Luxton G, Keall PJ, King CR. A new formula for normal tissue complication probability (NTCP) as a function of equivalent uniform dose (EUD). *Phys Med Biol* 2008;53:23-36.
- [148] Yorke ED. Modeling the effects of inhomogeneous dose distributions in normal tissues. *Semin Radiat Oncol* 2001;11:197-209.

- [149] Toramatsu C, Katoh N, Shimizu S, Nihongi H, Matsuura T, Takao S, Miyamoto N, Suzuki R, Sutherland K, Kinoshita R, Onimaru R, Ishikawa M, Umegaki K, Shirato H. What is the appropriate size criterion for proton radiotherapy for hepatocellular carcinoma? A dosimetric comparison of spot-scanning proton therapy versus intensity-modulated radiation therapy. *Radiat Oncol* 2013;8:48.
- [150] Kanemoto A, Mizumoto M, Okumura T, Takahashi H, Hashimoto T, Oshiro Y, Fukumitsu N, Moritake T, Tsuboi K, Sakae T, Sakurai H. Dose-volume histogram analysis for risk factors of radiation-induced rib fracture after hypofractionated proton beam therapy for hepatocellular carcinoma. *Acta Oncol* 2013;52:538-44.
- [151] Crane CH and Koay EJ. Solutions that Enable Ablative Radiotherapy for Large Liver Tumors: Fractionated Dose Painting, Simultaneous Integrated Protection, Motion Management and CT Image Guidance. *Cancer* 2016;122:1974–86.
- [152] Ribeiro CO, Knopf A, Langendijk JA, Weber DC, Lomax AJ, Zhang Y. Assessment of dosimetric errors induced by deformable image registration methods in 4D pencil beam scanned proton treatment planning for liver tumours. *Radiother Oncol* 2018;128:174-81.
- [153] Engwall E, Fredriksson A, Glimelius L. 4D robust optimization including uncertainties in time structures can reduce the interplay effect in proton pencil beam scanning radiation therapy. *Med Phys* 2018;45:4020-9.
- [154] Moustakis C, Blanck O, Andratschke N, Boda-Heggemann J, Wilke L, Tanadini- Lang S, Gauer T, Alraun M, Baumer C, Beckers E, Blaschek M, Brunner T, Chan MK, Dierl M, Droege S, Duma MN, Ebrahimi, Eich HT, Fleckenstein J, Ganswindt U, Garbe S, Gkika E, Heinz C, Henkenberens C, Hennig A, Henning S, Köhn J, Kornhuber C, Kretschmer M, Krieger T, Loufti-Kraus B, Mayr M, Mensing T, Pfeiler T, Pollul G, Ramm U, Rieder F, Rieken S, Schmidhalter D, Schmitt D, Schöffler J, Ulm C, Walke M, Weigel R, Wiehle R, Wiezorek T, Wolf U, Zimmer J, Baus W, Guckenberger M. Referenz Planvergleichsstudie für Leber-SBRT – Ergebnisse DEGRO AG/DGMP AK Stereotaxie. *Strahlenther Onkol* 2018;194:1.
- [155] Li Y, Kardar L, Li X, Li H, Cao W, Chang JY, Liao L, Zhu RX, Sahoo N, Gillin M, Liao Z, Komaki R, Cox JD, Lim G, Zhang X. On the interplay effects with proton scanning beams in stage III lung cancer. *Med Phys* 2014;41:021721.
- [156] Paganetti H. Proton Beam Therapy. IOP Publishing; 2017, p.1-23. <https://doi.org/10.1088/978-0-7503-1370-4ch1>
- [157] Yasui K, Toshito T, Omachi C, Kibe Y, Hayashi K, Shibata H, Tanaka K, Nikawa E, Asai K, Shimomura A, Kinou H, Isoyama S, Fujii Y, Takayanagi T, Hirayama S, Nagamine Y, Shibamoto Y, Komori M, Mizoe JE. A patient-specific aperture system with an energy absorber for spot scanning proton beams: Verification for clinical application. *Med Phys* 2015;42:6999-7010.
- [158] Moteabbed M, Yock TI, Depauw N, Madden TM, Kooy HM, Paganetti H. Impact of Spot Size and Beam-Shaping Devices on the Treatment Plan Quality for Pencil Beam Scanning Proton Therapy. *Int J Radiat Oncol Biol Phys* 2016;95:190-8.
- [159] Bäumer C, Janson M, Timmermann B, Wulff J. Collimated proton pencil-beam scanning for superficial targets: impact of the order of range shifter and aperture. *Phys Med Biol* 2018;63:085020.
- [160] Komatsu S, Hori Y, Fukumoto T, Murakami M, Hishikawa Y, Ku Y. Surgical spacer placement and proton radiotherapy for unresectable hepatocellular carcinoma. *World J Gastroenterol* 2010;16:1800–3.

- [161] Rasch B, Frieze M, Hofmann W, Naumann E. Der t-test. In: Rasch B, Frieze M, Hofmann W, Naumann E. Quantitative Methoden 1. Einführung in die Statistik für Psychologen und Sozialwissenschaftler. Springer, Berlin Heidelberg; 2014, p.33-79.
- [162] Posten HO. Robustness of the Two-Sample T-Test. In: Rasch D and Tiku ML, editors. Robustness of Statistical Methods and Nonparametric Statistics. Springer, Dordrecht; 1984, p.92-99.
- [163] Liu C, Schild SE, Chang JY, Liao Z, Korte S, Shen J, Ding X, Hu Y, Kang Y, Keole SR, Sio TT, Wong WW, Sahoo N, Bues M, Liu W. Impact of Spot Size and Spacing on the Quality of Robustly Optimized Intensity Modulated Proton Therapy Plans for Lung Cancer. *Int J Radiat Oncol Biol Phys* 2018;101:479-89.
- [164] Von Siebenthal M, Székely G, Lomax AJ, Cattin PC. Systematic errors in respiratory gating due to intrafraction deformations of the liver. *Med Phys* 2007;34:3620-9.
- [165] Segars WP, Mahesh M, Beck TJ, Frey EC, Tsui BM. Realistic CT simulation using the 4D XCAT phantom. *Med Phys* 2008;35:3800-8.
- [166] Segars WP, Sturgeon G, Mendonca S, Grimes J, Tsui BM. 4D XCAT phantom for multimodality imaging research. *Med Phys* 2010;37:4902-15.
- [167] National Library of Medicine (NLM). Visible human male and female data sets. URL: https://www.nlm.nih.gov/research/visible/visible_human.html (Last accessed: 31 August 2018).
- [168] Piegl L. On NURBS: a survey. *IEEE Comput Graph Appl* 1991;11:55–71.
- [169] RaySearch Laboratories AB. RayStation 7 Reference Manual RSL-D-RS-7.0-REF-EN-1.0-2017-12-08. Stockholm, Sweden; 2017.

Acknowledgment

At this point, I would like to express my deep gratitude to the people and institutions who have supported and helped me throughout my PhD project. The last years at the West German Proton Therapy Centre Essen (WPE) have been an intense period for me full of new experiences and insights, not only on a scientific and professional level but also on a personal level. This work would not have been possible without the support of various people from different disciplines, such as medical physicists, clinicians, technicians and specialists from industry.

I would first like to thank Prof. Dr. Bernhard Spaan for being my first referee and providing an external and independent view on the project. His questions, comments and advices have always been constructive and encouraging. In addition, I would like to thank Prof. Dr. Kevin Kröniger for agreeing to be my second referee on such a short notice. My special thanks go to my supervisor at WPE, Dr. Christian Bäumer. He granted me maximum freedom in the design of the project and the door to his office was always open whenever I had questions regarding my research or academic writing. Furthermore, I am grateful for his critical reading of papers, conference contributions and the first draft of my doctoral thesis. I would also like to thank Xavier Vermeren, chief physicist at WPE, for the opportunity to pursue my research at WPE and to present the results at various international conferences.

My grateful thanks are extended to the whole physics group for their support and the enjoyable working atmosphere. I would particularly like to single out Dr. Jörg Wulff. His valuable and constructive feedback after reviewing papers and the doctoral thesis, respectively, has been an enormous help. Without him, the collaboration with the Institute of Medical Physics and Radiation Protection of the TH Mittelhessen would not have been established. I am also grateful to Dr. Jamil Lambert, a former colleague at WPE, for sharing his experiences in the field of moving targets. Further, I wish to acknowledge the help provided by Sandija Plaude from the physicist's side and from Dr. Dirk Geismar and Dr. Dalia Ahmad Khalil from the physician's side in reviewing treatment plans and contours. There was also great support from Dr. Danny Jazmati who reviewed the medical part of the dissertation on hepatocellular carcinomas. Special thanks go to Ajvar Kern, Nico Verbeek, Isabel Vogt, Sabrina Smyczek, Samantha Kauer and Carina Behrends for creating joyful moments at work, having good conversations and sharing their knowledge. We not only spent a great number of early, late and Saturday shifts together but also shared exciting and fun experiences outside of WPE. Cordial thanks go to the mechanics of our workshop, our IT team and to Dr. Anne-Katrin Nix and Dr. Andrija Matić from the local IBA team for fruitful discussions about the physics of the proton therapy system.

Dr. Erik Engwall from RaySearch Laboratories AB (Stockholm, Sweden) greatly contributed to this work by sharing his initial non-clinical draft for a 4D dose calculation routine and providing advice and assistance during further development. I am thankful to Dr. Mark Chan, Dr. Oliver Blanck and Prof. Dr. Jürgen Dunst from the Department of Radiation Oncology of the University Clinic Schleswig-Holstein (Kiel, Germany) for their support. They and Dr. Myriam Ayadi from the Centre Léon Bérard (Lyon, France) provided the required 4D CT data sets of hepatocellular carcinoma patients. The collaboration with Dr. Ulf Mäder from the Institute of Medical Physics and Radiation Protection of the TH Mittelhessen opened up new opportunities for 4D studies based on computer generated CT data.

Mein ganz besonderer Dank gilt meinen Eltern für ihre Unterstützung während meines Studiums und der Promotion sowie in jeder anderen Lebenslage. Durch ihr Vertrauen, ihre Motivation, Liebe und zuverlässige Hilfe haben sie mir stets den Rücken gestärkt. Abschließend möchte ich mich bei meiner derzeitigen und meiner ehemaligen Mitbewohnerin und meinen langjährigen Freunden bedanken, die mir selbst in arbeitsintensiven und stressigen Zeiten ein Lächeln ins Gesicht zaubern konnten.