- -

Liver-ultrasound-guided lung tumour tracking for scanned proton therapy: a feasibility study

- -

Krieger, Miriam; Giger, Alina; Jud, Christoph; Duetschler, Alisha; Salomir, Rares Vincent; Bieri, Oliver; Bauman, Grzegorz; Nguyen, Damien; Cattin, Philippe C; Weber, Damien C; Lomax, Antony J; Zhang, Ye

Liver-ultrasound-guided lung tumour tracking for scanned proton therapy: a feasibility study

Miriam Krieger[a,b*], **Alina Giger**[c,d*], **Christoph Jud**[c,d],
Alisha Duetschler[a,b], **Rares Salomir**[e,f], **Oliver Bieri**[c,g],
Grzegorz Bauman[c,g], **Damien Nguyen**[c,g], **Philippe C Cattin**[c,d],
Damien C Weber[a,h,i], **Antony J Lomax**[a,b], **Ye Zhang**[a]

[a] Paul Scherrer Institute (PSI), Center for Proton Therapy, Villigen PSI, Switzerland
[b] Department of Physics, ETH Zurich, Zurich, Switzerland
[c] Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland
[d] Center for medical Image Analysis & Navigation, University of Basel, Allschwil, Switzerland
[e] Image Guided Interventions Laboratory (GR-949), Faculty of Medicine, University of Geneva, Geneva, Switzerland
[f] Radiology Division, University Hospitals of Geneva, Geneva, Switzerland
[g] Department of Radiology, Division of Radiological Physics, University Hospital Basel, Basel, Switzerland
[h] Department of Radiation Oncology, University Hospital Zurich, Zurich, Switzerland
[i] Department of Radiation Oncology, Inselspital Bern, Bern, Switzerland
* Both authors contributed equally

E-mail: miriam.krieger@psi.ch, alina.giger@unibas.ch

September 2020

Abstract. Pencil beam scanned (PBS) proton therapy of lung tumours is hampered by respiratory motion and the motion-induced density changes along the beam path. In this simulation study, we aim to investigate the effectiveness of proton beam tracking for lung tumours both under ideal conditions and in conjunction with a respiratory motion model guided by real-time ultrasound imaging of the liver. Multiple-breathing-cycle 4DMRIs of the thorax and abdominal 2D ultrasound images were acquired simultaneously for five volunteers. Deformation vector fields extracted from the 4DMRI, referred to as ground truth motion, were used to generate 4DCT(MRI) data sets of two lung cancer patients, resulting in 10 data sets with variable motion patterns. Given the 4DCT(MRI) and the corresponding ultrasound images as surrogate data, a patient-specific motion model was built. The model consists of an autoregressive model and Gaussian process regression for the temporal and spatial prediction, respectively. Two-field PBS plans were optimised on the reference CTs, and 4D dose calculations (4DDC) were used to simulate dose delivery for (a) unmitigated motion, (b) ideal 2D and 3D tracking (both beam adaption and 4DDC based on ground truth motion), and (c) realistic 2D and 3D tracking (beam adaption based on motion predictions, 4DDC on ground truth motion). Model-guided tracking retrieved clinically acceptable target dose homogeneity, as seen in a substantial reduction of the D5-D95 % compared to the non-mitigated simulation. Tracking in 2D and 3D resulted in a similar improvement of the dose homogeneity, as did ideal and realistic tracking simulations. In some

Ultrasound-guided lung tumour tracking 2

cases, however, the tracked deliveries resulted in a shift towards higher or lower dose levels, leading to unacceptable target over- or under-coverage. The presented motion modelling framework was shown to be an accurate motion prediction tool for the use in proton beam tracking. Tracking alone, however, may not always effectively mitigate motion effects, making it necessary to combine it with other techniques such as rescanning.

1. Introduction

In the treatment of thoracic or abdominal tumours, the advantages of pencil beam scanned (PBS) proton therapy are challenged by respiratory motion. Not only lateral target miss but also interplay between the delivery dynamics and the moving tumour can strongly deteriorate the quality of the initially optimised 3D treatment plan (Phillips, Pedroni, Blattmann, Boehringer, Coray & Scheib 1992, Bert & Durante 2011). In order to ensure a safe and effective treatment for tumours affected by respiratory motion employing appropriate motion mitigation strategies is compulsory (Zhang, Huth, Weber & Lomax 2018). Indeed, scanned proton deliveries lend themselves to tumour tracking by nature, because each beam position could be adapted in real-time to follow the tumour motion. This probably being the most elegant motion mitigation concept, it is also the most challenging technique to be implemented clinically (Rietzel & Bert 2010). From a motion modelling point of view, the full 3D deformation of the patient's anatomy needs to be known in real-time; motion trajectories of the tumour's centre of mass are not sufficient due to the sensitivity of the proton beam with respect to any changes in tissue density (Bertholet, Knopf, Eiben, McClelland, Grimwood, Harris, Menten, Poulsen, Nguyen, Keall et al. 2019). From a technical point of view, the delivery system needs to support fast changes in the spot position and the proton energy (Grözinger, Bert, Haberer, Kraft & Rietzel 2008).

A technical implementation of beam tracking in a particle scanning system was presented and validated by Bert, Gemmel, Saito, Chaudhri, Schardt, Durante, Kraft & Rietzel (2010). The authors showed that their system is capable of adapting the carbon ion beam in real-time both laterally and in energy and with high accuracy. This was achieved by using the scanning magnets and including a down-stream wedge degrader. In order to extract the target motion during treatment, Prall, Kaderka, Saito, Graeff, Bert, Durante, Parodi, Schwaab, Sarti & Jenne (2014) and Schwaab, Prall, Sarti, Kaderka, Bert, Kurz, Parodi, Günther & Jenne (2014) indicated the potential of ultrasound (US) guidance for such a tracking system. They performed tracking deliveries for scanned ion beams and found a substantial improvement of target coverage. All of the above studies, however, only considered rigidly moving geometries in a water phantom. Deformable respiratory motion of a real patient geometry has been addressed by Zhang, Knopf, Tanner, Boye & Lomax (2013) and Zhang, Knopf, Tanner & Lomax (2014), where online fluoroscopy in beam's-eye-view in combination with a principal component analysis (PCA) based motion model were used to estimate liver deformations. Motion of either implanted markers or the diaphragm were extracted to infer the full patient geometry in form of dense deformation vector fields (DVFs). The first study showed that both the implanted markers and the diaphragm are appropriate surrogates for estimating the applied 4D dose distribution with high dosimetric accuracy (Zhang et al. 2013). In a further study, the same setup was used to simulate tumour tracking of the liver for proton PBS (Zhang et al. 2014). It was found that indeed, full DVFs are necessary to safely apply proton beam tracking to liver tumours. Moreover, abdominal US imaging was proposed as

Ultrasound-guided lung tumour tracking 4

surrogate signal before both for respiratory motion modelling of the liver (Preiswerk, De Luca, Arnold, Celicanin, Petrusca, Tanner, Bieri, Salomir & Cattin 2014) and for lung tumour tracking (Mostafaei, Tai, Gore, Johnstone, Haase, Ehlers, Cooper, Lachaine & Li 2018). While the former infers dense DVFs of the liver from a population-based model, the latter predicts the tumour trajectory in superior-inferior direction on a patient-specific basis. However, both studies do not investigate the dosimetric impact of the motion predictions.

Recently, we have introduced a respiratory motion model that uses Gaussian process regression (GPR, Williams & Rasmussen (2006)) to estimate dense DVFs, i.e. defined on every voxel, of the lung using abdominal US images as input (Giger, Krieger, Jud, Duetschler, Salomir, Bieri, Bauman, Nguyen, Weber, Lomax, Zhang & Cattin 2020). This framework was used to retrospectively reconstruct the delivered 4D dose distributions by taking into account variable motion patterns. In order to bring tracking treatments closer to clinical implementation, a temporal prediction of the patient motion is imperative. It is important to consider the system latency induced by the time required for US image processing, motion inference, and proton beam adaption. In conventional radiotherpay, the system latency ranges from 50 ms and 115 ms for the VERO system (Depuydt, Verellen, Haas, Gevaert, Linthout, Duchateau, Tournel, Reynders, Leysen, Hoogeman et al. 2011) and the Cyberknife (Hoogeman, Prévost, Nuyttens, Pöll, Levendag & Heijmen 2009), respectively, up to several hundreds of milliseconds for multileaf collimator tracking systems (Krauss, Nill, Tacke & Oelfke 2011, Poulsen, Cho, Ruan, Sawant & Keall 2010). Similarly for proton beam therapy, fast energy changes in range of 80 ms for typical range steps of 5 mm were reported (Pedroni, Meer, Bula, Safai & Zenklusen 2011, Safai, Bula, Meer & Pedroni 2012). For the US-based tracking system mentioned above (Prall et al. 2014), a latency of 200 ms was reported. However, the major portion of the latency was introduced by non-optimised image processing steps. To compensate system latencies, various motion prediction models were proposed and compared, among which are linear predictors, support vector regression, atlas-based approaches, Kalman filters, and neural networks (Arnold, Preiswerk, Fasel, Salomir, Scheffler & Cattin 2011, Krauss, Nill & Oelfke 2011, Ernst, Dürichen, Schlaefer & Schweikard 2013, Preiswerk et al. 2014). While one comparison study reported relatively small differences in prediction accuracy for the different approaches (Krauss, Nill & Oelfke 2011), another study suggests that machine learning methods tend to outperform linear models for increased prediction horizons (Ernst et al. 2013). Yet, the differences between the various prediction models have primarily been discussed from a mathematical point of view. The benefits of more advanced motion predictors on the resulting dose distribution is not clear (Knopf, Stützer, Richter, Rucinski, da Silva, Phillips, Engelsman, Shimizu, Werner, Jakobi et al. 2016).

In this study, we extend our previously presented motion model (Giger et al. 2020) with a linear autoregressive (AR) model (Giger, Jud, Nguyen, Krieger, Zhang, Lomax, Bieri, Salomir & Cattin 2019) which allows forecasting of the motion information to account for

Ultrasound-guided lung tumour tracking   5

system latency. We simulate various treatment scenarios for proton beam tracking and investigate the performance of the spatio-temporal motion prediction model in terms of dose restoration. We compare the resulting dose distributions based on the predicted motion to ideal tracking where perfect knowledge of the respiratory motion is assumed. As such, this is the first study which quantifies the feasibility of the proposed US-guided respiratory motion model when applied to PBS proton tracking of lung tumours in terms of 4D dose distributions. The presented simulation study is based on comprehensive and realistic motion patterns taking into account dense anatomical deformations.

2. Methods

The motion modelling pipeline presented in this simulation study follows a typical approach (McClelland, Hawkes, Schaeffter & King 2013):

(i) In a pretreatment phase, motion information and surrogate data are acquired simultaneously to ensure temporal correspondence between the two modalities.

(ii) The motion model is formulated and its parameters estimated to approximate the relationship between the motion estimates and the surrogate signal.

(iii) During the simulated treatment delivery, only the surrogate data is available and the motion model predicts the corresponding motion.

In order to investigate the dosimetric effects of the motion model, further image processing tools are required. In this study in particular we use:

- time-resolved volumetric MRI (4DMRI) of the lungs to extract the respiratory motion in form of dense DVFs,
- 2D abdominal imaging to extract the respiratory surrogate signal, and
- CT volumes of lung cancer patients to incorporate the relative proton stopping power for the dosimetric analysis.

The main image processing tools and data sets used here have been described in detail in our preceding study (Giger et al. 2020). Although the principal concepts and methods are presented below, certain information has been omitted for conciseness. Interested readers are referred to Giger et al. (2020) for more details.

2.1. Image data sets

Simultaneous abdominal US and 4DMRI of the lungs were acquired of five healthy volunteers under free respiration using two different 4DMRI sequences (Celicanin, Giger, Bauman, Cattin & Bieri 2017, Jud, Nguyen, Sandkühler, Giger, Bieri & Cattin 2018). Intercostal US imaging of the liver was performed with the US probe held in place by means of a strap and a custom made casting (Santini, Gui, Lorton, Guillemin, Manasseh, Roth, Bieri, Vallée, Salomir & Crowe 2020). B-mode US imaging was performed at 3.3 MHz and case-specific acoustic power. The hybrid data acquisition was performed for

Ultrasound-guided lung tumour tracking   6

up to 11 min with frame rates of 1.25 to 2.25 Hz and 15 Hz for 4DMRI and US imaging, respectively. The two imaging modalities were temporally aligned using the optical output trigger of the MRI scanner and signal post-processing. The proposed approach does not require a spatial co-registration of the US and the MRI images. Voxel-wise DVFs were extracted from the 4DMRI data sets using deformable image registration of all motion frames with respect to a reference full-exhale frame (www.plastimatch.org and Sandkühler, Jud, Andermatt & Cattin (2018)).

Synthetic 4DCT(MRI) data sets (Krieger, Giger, Weber, Lomax & Zhang 2020, Boye, Samei, Schmidt, Székely & Tanner 2013) were generated by warping two lung cancer patient CTs with the DVFs of the volunteers, resulting in 10 combinations of CT geometry and 4DMRI motion patterns, hereinafter referred to as geometry / motion cases. The patient CTs were chosen to differ considerably in terms of position, size, and motion characteristics as shown in figure 1 and figure 3. Figure 1 shows the motion patterns of the clinical target volume (CTV) in superior-inferior (SI) direction for all data sets. In this study, the motion amplitude in SI direction tends to be larger for CT geometry 1 (smaller tumour) when compared to CT geometry 2 (larger tumour). Further information on the motion characteristics for each geometry / motion case can be found in Appendix A. The differences in signal properties, such as duration, missing US frames, or number of states needed for treatment delivery, are induced by the different 4DMRI sequences and synchronisation setups for motions 1 and 2 on the one hand (Celicanin et al. 2017), and motions 3 to 5 on the other hand (Jud et al. 2018).

2.2. Motion modelling

It is imperative to predict the motion vectors into the future since any tracking system inevitably experiences latency, e.g. due to signal processing and transport or time needed to change the spot position and energy. For this reason, the motion model employed in this study is split into two parts: first, the US-based surrogate signal is forecast to some time points ahead using an AR model (Giger et al. 2019); second, a voxel-wise DVF of the lungs is predicted using the forecast surrogate signal and a correspondence model based on GPR (Giger et al. 2020). Principal component analysis (PCA) of both the US images and the DVFs was performed in order to reduce the dimensionality of the motion model. By retaining a reduced number of principal components only, the model is less prone to over-fitting and is rendered computationally more efficient, which is important for the online application of such a model. A schematic of the complete motion modelling framework is shown in figure 2.

The data sets were split into three parts as shown in figure 1. The first 300–750 US frames were used to train and validate the AR model, while the last 33–270 US/MR image pairs represented the test set and, thus, were excluded from any model training. The size of the test set was governed by the number of states needed to deliver the respective treatment plan. The remaining data was used for GPR modelling and, for this purpose, split into a training set and a fixed-sized validation set consisting of the

Ultrasound-guided lung tumour tracking   7
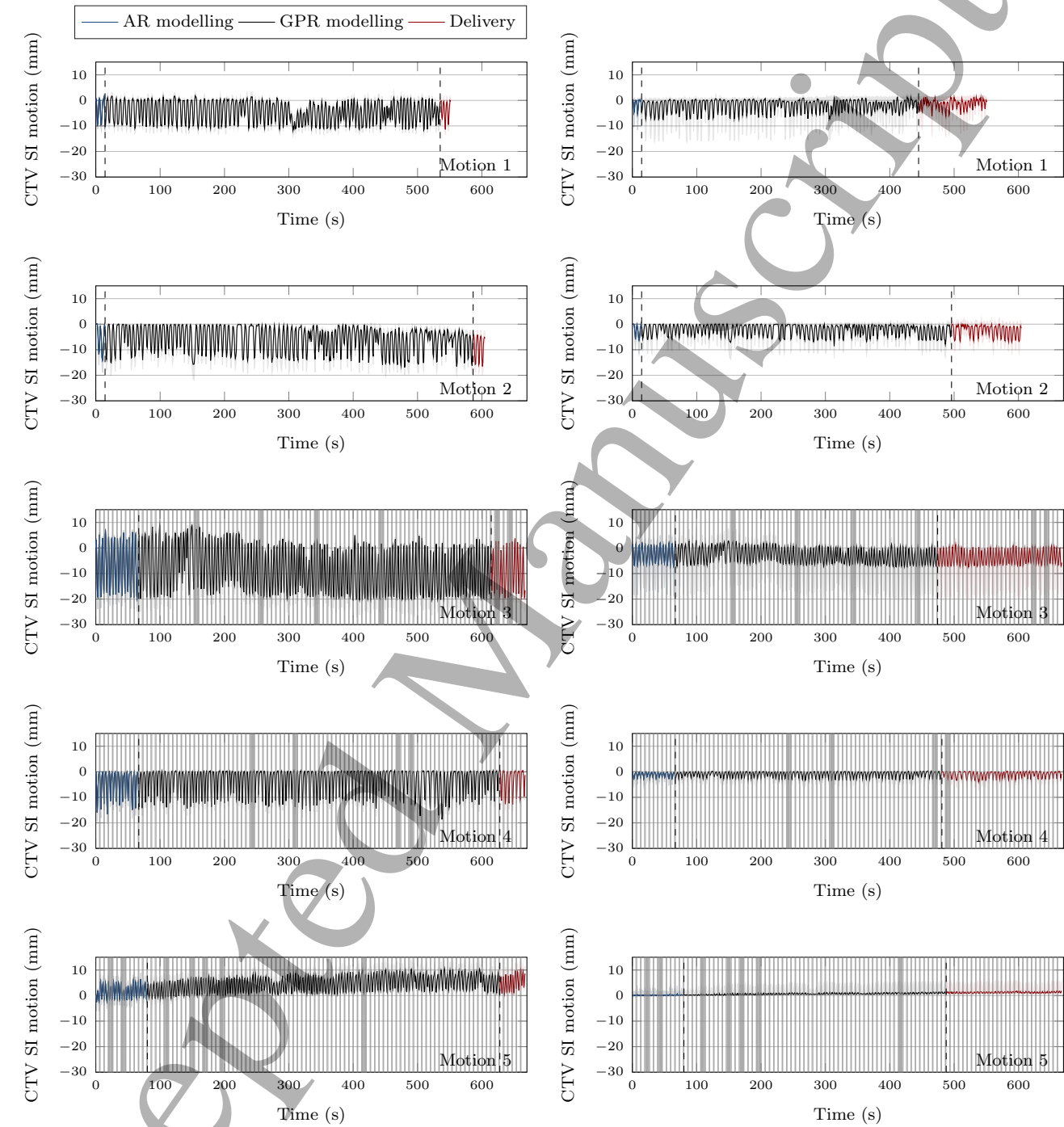
(a) CT geometry 1   (b) CT geometry 2

Figure 1: SI motion of the CTV for all geometry / motion cases. The data sets are split into three parts as highlighted by the vertical dashed lines: AR training and validation sets (blue), GPR training and validation sets (black), and treatment delivery (red). The evaluation of the motion model and tracking algorithms was performed for the last part only (red). Solid line: median, shades: 10th to 90th percentile of all CTV voxels. For motions 3 to 5 some US images were lost due to the synchronisation setup as indicated by the gray areas.
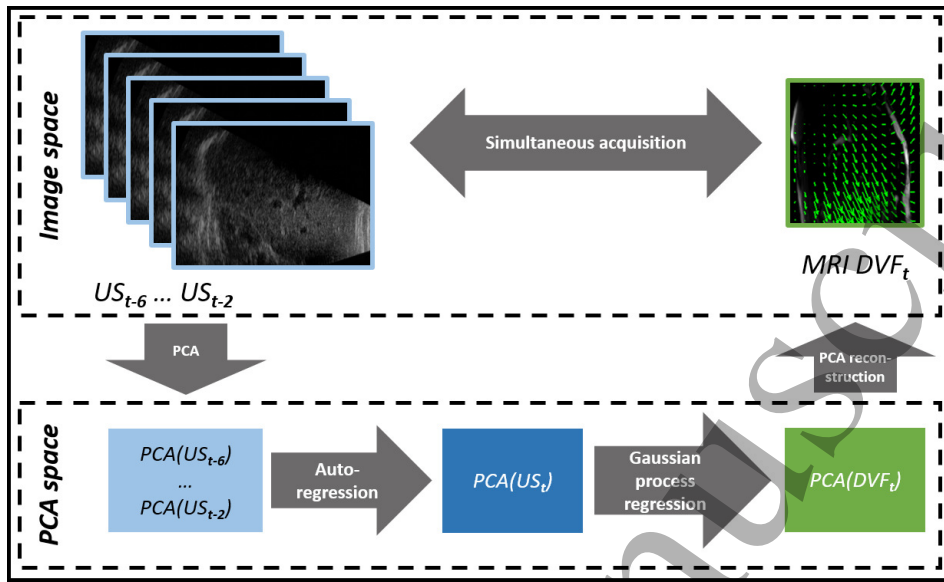
Figure 2: Illustration of the respiratory motion model.

last 50 US/MR image pairs. A detailed overview of all data sets is given in table 1. The DVFs predicted by this model framework are referred to as *predicted motion*, while the DVFs as extracted from the 4DMRI are referred to as the *ground truth motion*. It is important to note that our ground truth DVFs do not necessarily represent the real patient motion patterns. Due to uncertainties in image acquisitions as well as deformable image registrations, the actual patient motion is not known exactly. However, the DVFs extracted from the 4DMRI which were used for both model training and the 4D dose calculations represent a valid ground truth to evaluate the model performance and the resulting dose distributions.

The frame rate of the US acquisition was an order of magnitude higher when compared to the frame rate of the 4DMRI ($15\,\mathrm{Hz}$ vs. 1.25 to $2.25\,\mathrm{Hz}$). Since the AR model used for the temporal prediction only depends on US images, substantially more training data was available. The GPR model for the spatial prediction relies on US/MR image pairs and, therefore, the data available for training is limited by the frame rate of the 4DMRI. However, for the online application of the correspondence model, the frame rate of the surrogate signal is relevant.

Autoregressive model The autoregressive model used in this study has been described previously (Giger et al. 2019) and, therefore, is only briefly summarised here. Let $\boldsymbol{\alpha}_t \in \mathbb{R}^u$ describe the vector of the u most dominant principal components of the US image acquired at time t. Furthermore, let α_t^j denote the jth element of the vector $\boldsymbol{\alpha}_t$. The AR model of order p is applied element-wise and its parameters are estimated using ordinary least squares:

$$\alpha_t^j = \vartheta_0^j + \sum_{i=1}^{p} \vartheta_i^j \alpha_{t-i}^j + \epsilon_t \quad \forall j \in \{1, ..., u\},$$

Ultrasound-guided lung tumour tracking   9

Table 1: Overview over the different data sets used for motion modelling. For AR, all US frames are considered, whereas for GPR, only US/MR image pairs are used.

		AR		GPR		
		Training	Validation	Training	Validation	Test
Geometry 1	Motion 1	280	20	595	50	33
	Motion 2	280	20	659	50	33
	Motion 3	600	150	730	50	60
	Motion 4	600	150	750	50	60
	Motion 5	600	150	720	50	60
Geometry 2	Motion 1	280	20	482	50	146
	Motion 2	280	20	546	50	146
	Motion 3	600	150	520	50	270
	Motion 4	600	150	540	50	270
	Motion 5	600	150	710	50	270

with model parameters $\boldsymbol{\vartheta}^j = \begin{bmatrix} \vartheta_0^j & \ldots & \vartheta_p^j \end{bmatrix}^T$ and noise ϵ_t. In order to perform a n-step-ahead forecasting, the model is repeatedly applied.

In this study, the US images were reduced to u principal components such that the explained variance ratio was 10%. The autoregressive model was of order $p = 5$ and the principal components $\tilde{\boldsymbol{\alpha}}_t$ were predicted $n = 2$ steps ahead, resulting in a prediction horizon of $133\,\text{ms}$ for the given acquisition parameters.

Gaussian process regression The correspondence model based on GPR has been described in detail in (Giger et al. 2020). Analogously, let $\boldsymbol{\beta}_t \in \mathbb{R}^v$ denote the vector of the v most dominant principal components of the DVF at time t. However, in contrast to the previous work, here the input to the correspondence model is the forecast surrogate signal, $\tilde{\boldsymbol{\alpha}}_t$. The number of principal components v was chosen such that 75% of the explained variance was retained. The parameters of the applied Gaussian covariance function were empirically chosen to be $\theta_0 = 30$, $\theta_1 = 35$, and $\sigma_n^2 = 1$ for all data sets as introduced in Giger et al. (2020). The computation time needed for performing the PCA and estimating the model parameters was in the range of 52 to $108\,\text{min}$ on an Intel Xeon CPU E5-2620 v3 @ $2.40\,\text{GHz}$ for the given data sets. Given the $p = 5$ preceding US images, motion prediction was performed on average in $(126.9 \pm 36.3)\,\text{ms}$.

2.3. Treatment planning and 4D dose calculations

For each CT geometry, a 2-field SFUD (single field, uniform dose) plan was optimised on the reference full-exhale CT scan using the in-house planning system described by Pedroni, Bearpark, Böhringer, Coray, Duppich, Forss, George, Grossmann, Goitein, Hilbes, Jermann, Lin, Lomax, Negrazus, Schippers & Kotrle (2004). This gantry features a fast, upstream energy switching of $80\,\text{ms}$ and lateral scanning dead times of $3\,\text{ms}$, see Zhang et al. (2018) (model 4) for more details. As optimisation target,

Ultrasound-guided lung tumour tracking   10

(a) CT geometry 1, CTV volume: $18.8\,\mathrm{cm}^3$
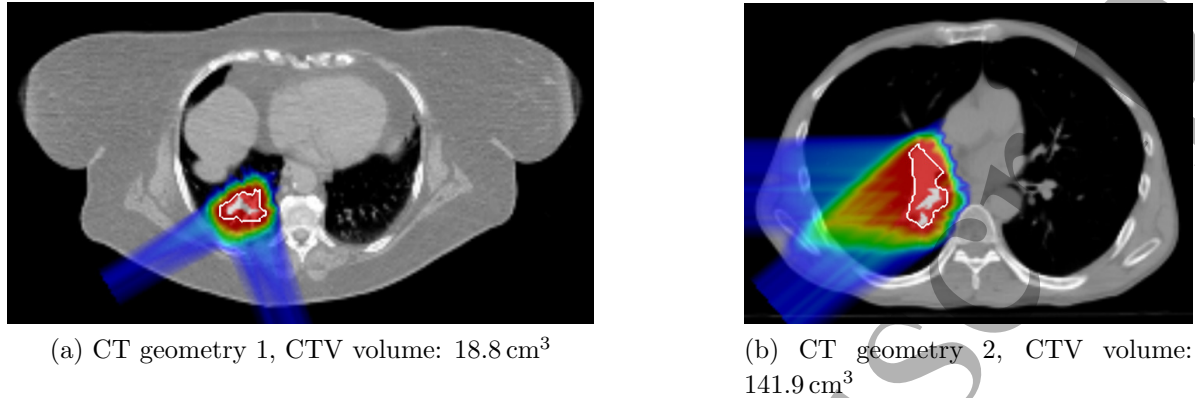

(b) CT geometry 2, CTV volume: $141.9\,\mathrm{cm}^3$

Figure 3: Illustration of the two CT geometries and the corresponding chosen field arrangements.

Table 2: Summary of the simulated dose deliveries. $\mathrm{M_{adapt}}$ describes the motion according to which the beam is adapted, $\mathrm{M_{calc}}$ describes the motion that is used for the actual dose calculation.

Name	$\mathrm{M_{adapt}}$	$\mathrm{M_{calc}}$	Simulation code
Reference dose	–	–	4D dose calculation
4DDC	–	Ground truth	4D dose calculation
Ideal 2D tracking	Ground truth	Ground truth	2D tracking code
Ideal 3D tracking	Ground truth	Ground truth	3D tracking code
Realistic 2D tracking	Prediction	Ground truth	2D tracking code
Realistic 3D tracking	Prediction	Ground truth	3D tracking code

the respective CTV plus a $2\,\mathrm{mm}$ margin was used. Since the motion and density variations are taken care of by the tracking simulation, they were not included in the plan optimisation. The field arrangements are illustrated in figure 3: For CT geometry 1, gantry angles of $160°$ and $240°$ were used, whereas for CT geometry 2, the angles were $230°$ and $270°$.

Several delivery scenarios were simulated, as summarised in table 2. First, the static reference dose distribution was calculated by running a 4D dose calculation (4DDC, see Boye, Lomax & Knopf (2013) and Krieger, Klimpki, Fattori, Hrbacek, Oxley, Safai, Weber, Lomax & Zhang (2018)) without motion input. Second, the unmitigated 4D dose distributions were calculated as 'worst case reference' by using the ground truth DVFs as respiratory motion. Finally, four different tracking scenarios were simulated by considering different motion compensation approaches. Simulations were performed for both 2D tracking (only adapting the beam laterally) and 3D tracking (adapting the beam laterally and in energy). For both 2D and 3D tracking simulations, *ideal* and *realistic* tracking were distinguished. Ideal tracking assumed perfect motion information for the beam adaptation, meaning that the motion according to which the beam is

Ultrasound-guided lung tumour tracking 11

adapted, M_{adapt}, is equal to the ground truth motion used for the 4DDC, M_{calc}. It thus serves as a reference for the best-case capability of tracking with perfect knowledge of the deformable motion. Under more realistic conditions, the two motion patterns M_{adapt} and M_{calc} are not equal but rather M_{adapt} is predicted by the model. This scenario is referred to as realistic tracking. All simulations were performed as single deliveries, $9\times$ volumetric rescanning, and $9\times$ volumetric re-tracking (van de Water, Kreuger, Zenklusen, Hug & Lomax 2009).

In clinical practice, the start of a delivery is usually not synchronised with the respiration of the patient. To cover the whole range of potential delivery/respiration interplays, the simulations of each treatment field were started on arbitrary, but different, respiratory states of the first breathing cycle of M_{calc}. The number of respiratory states was in the range of 5–11 for one field, resulting in 25–121 combinations for a full treatment plan with two fields.

2.4. Analysis

Geometrical error   The geometrical prediction error for a motion state t is defined as the Euclidean norm of the voxel-wise vector difference between the ground truth and the predicted DVFs. For a quantitative analysis, the 50th and 95th percentiles of this prediction error within the CTV were calculated for each respiratory state. The results are reported as a distribution of these percentiles over all motion states.

Dosimetric analysis   The dosimetric analysis was performed by comparing the dose volume histograms (DVHs) of the different delivery simulations and by evaluating the target dose homogeneity in terms of CTV D5-D95 %.

3. Results

3.1. Geometrical error

Figure 4 summarises the Euclidean geometrical prediction error of the model, as defined in section 2.4, for all geometry / motion scenarios. The boxes describe the 50th and the 95th percentile of all voxels within the CTV and include all motion states that were predicted by the model and thus used for the dose calculations. It can be observed that, except for a few outlying states, the 50th percentile error is smaller than 4 mm for all geometry / motion scenarios. For motions 3–5, it is even smaller than 2 mm for almost all motion states. A similar pattern is seen for the 95th percentile error: except for a few outliers, all motion states present an error smaller than 8 mm for all geometry / motion combinations. For motions 3–5, it is smaller than 4 mm for all motion states except some outliers.

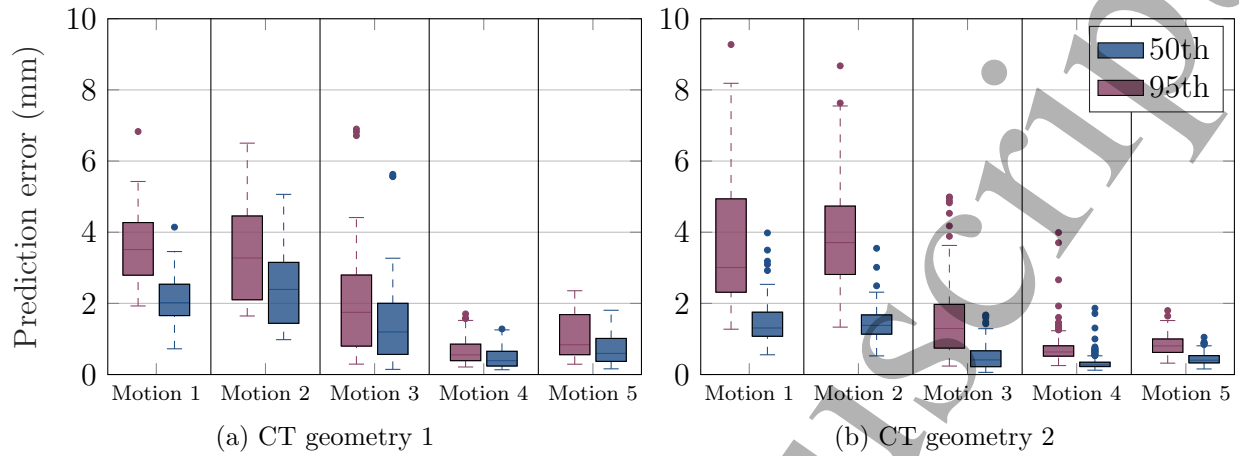*Ultrasound-guided lung tumour tracking*   12

Figure 4: Spatio-temporal prediction errors for all geometries and motions. The boxplots indicate the error percentiles of all voxels within the CTV and include all predicted motion states. The whiskers expand to the most extreme values within 1.5 times the inter-quartile range.

3.2. Motion mitigation efficacy of ideal tracking

DVHs for all geometry / motion scenarios for ideal tracking are shown in figures 5 and 6 for CT geometry 1 and CT geometry 2, respectively, comparing ideal 2D and 3D tracking to either the scenario with no motion mitigation or the reference dose. Figure 7 additionally compares the CTV dose homogeneity in terms of D5-D95 % for the different tracking scenarios.

It can be seen that for geometry 1, there is a clear under-dosage and severely compromised dose homogeneity for non-mitigated motion cases. Tracking brings the CTV dose homogeneity much closer to the reference dose for both 2D and 3D tracking, with a slightly better homogeneity for 2D tracking. However, for motions 3 and 4 there is a clear under-dosage of the CTV. In addition, rescanning leads to a smaller dependence of the dose on the starting phase of the simulations, indicated by the more narrow band plots. For motions 1, 2 and 4, re-tracking additionally achieves a result closer to the reference dose compared to the single-tracking simulations. Motions 3 and 5 do not show an improvement if rescanning is added. For geometry 2, the dose degradation due to missing motion mitigation is much more limited, because the tumour motion amplitudes are much smaller than for geometry 1. For this reason, tracking brings no major improvement compared to non-mitigated simulations. Again, 2D and 3D tracking are very similar, with a slightly improved homogeneity for 2D tracking. Rescanning slightly reduces the dependence on the starting phase.

Ultrasound-guided lung tumour tracking   13

(a) No rescanning   (b) With 9× rescanning
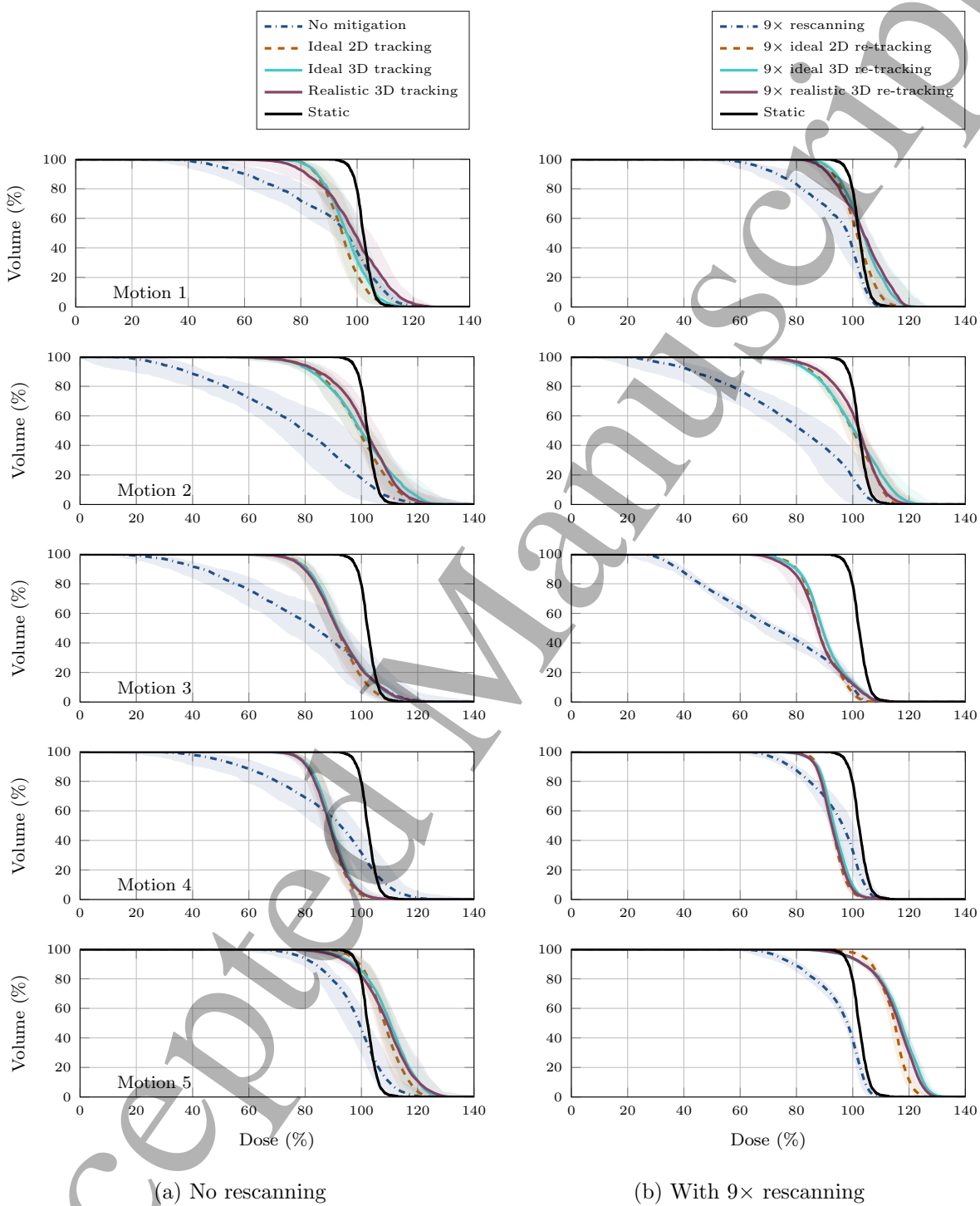
Figure 5: CTV DVHs for all motion scenarios for CT geometry 1 for ideal and realistic tracking. Blue, dashed-dotted: no tracking, orange, dashed: ideal 2D tracking, cyan, solid: ideal 3D tracking, purple, solid: realistic 3D tracking, black, solid: reference dose (no motion included). The shaded bands include all starting phase combinations, whereas the solid lines represent their median. (a) No rescanning, (b) with 9× rescanning.

Ultrasound-guided lung tumour tracking 14

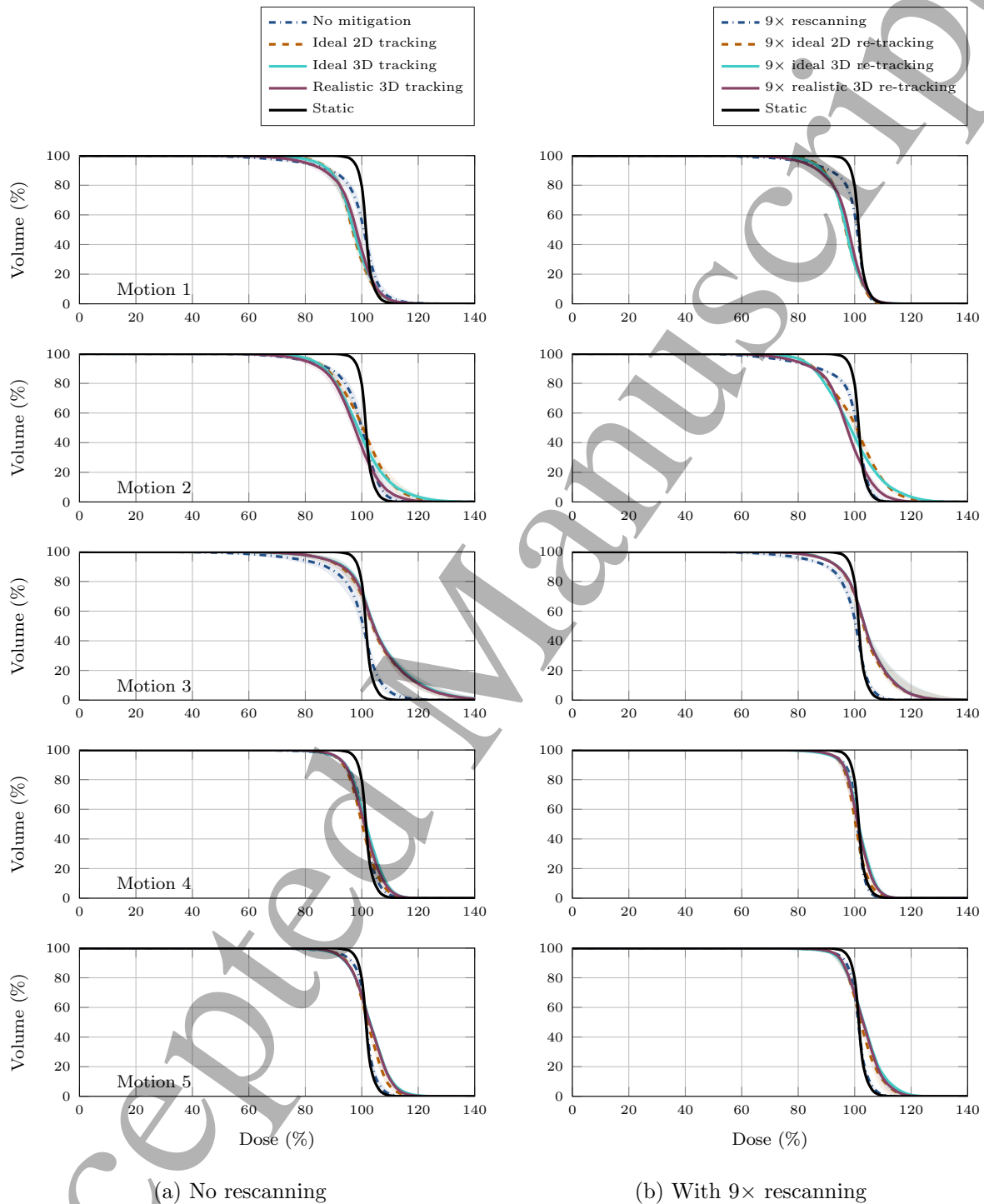(a) No rescanning

(b) With 9× rescanning

Figure 6: CTV DVHs for all motion scenarios for CT geometry 2 for ideal and realistic tracking. Blue, dashed-dotted: no tracking, orange, dashed: ideal 2D tracking, cyan, solid: ideal 3D tracking, purple, solid: realistic 3D tracking, black, solid: reference dose (no motion included). The shaded bands include all starting phase combinations, whereas the solid lines represent their median. (a) No rescanning, (b) with 9× rescanning.

Ultrasound-guided lung tumour tracking   15

3.3. Ideal versus realistic tracking

Figures 5 and 6 additionally compare the DVHs for ideal and realistic 3D tracking for no rescanning as well as $9\times$ rescanning / re-tracking for CT geometry 1 and CT geometry 2, respectively. It is seen that there are some differences between the two scenarios, originating from the different motion patterns used for beam adaptation. For motion 1, realistic tracking results in lower dose homogeneity due to the larger prediction error already seen in figure 4. This is also confirmed by figure 7. For most of the scenarios however, there are only slight differences in homogeneity between ideal and realistic tracking, meaning that the model provides a suitable motion prediction for tracking, which is also reflected in the geometrical prediction error in figure 4.

4. Discussion

In this study, we presented a framework for proton beam tracking based on abdominal US and a patient-specific respiratory motion model. We made use of simultaneously acquired 4DMRI and 2D US images to train a motion model consisting of two parts: an autoregressive model to forecast the US information ahead in time to cope with system latency, and a Gaussian process regression model to estimate the corresponding dense DVFs inside the lungs. These deformation fields were used to adapt the proton beam laterally and in energy, and thus simulate various tracking scenarios. We have shown that for large tumour motion amplitudes, 2D as well as 3D tracking can substantially improve the target dose homogeneity as compared to a non-mitigated dose delivery (figure 7, top row). This is true not only for ideal tracking where perfect motion information is assumed, but also for realistic tracking which uses the model's prediction as input for beam position and energy adaptation. Additionally, combining tracking with rescanning (re-tracking) may further improve the target coverage, as seen for cases with motions 1, 2 and 5. Generally, rescanning reduces the dose variations due to different delivery starting phases. For small motion amplitudes, the detrimental effects of the motion were much smaller, and thus tracking as well as re-tracking brought only minor improvements with respect to the unmitigated delivery (figure 7, bottom row).

The prediction horizon of 133 ms is in a reasonable range for tracking systems in radiotherapy. Although no closed loop system latency for proton tracking systems is reported here, we expect the time needed for energy adaption in the range of 80 ms to contribute a major portion of the overall latency for 3D tracking. Thus, we further expect the overall latency to be substantially lower for 2D tracking. It is important to note that the current imaging processing pipeline as well as the motion model used for the presented simulation study has not been optimised for real-time applications. Further improvements and dedicated software solutions are required for the presented motion modelling approach to be real-time capable. Moreover, the linear AR model used here is only one possible prediction approach among many and without doubt the most straightforward one (Krauss, Nill & Oelfke 2011, Ernst et al. 2013). The
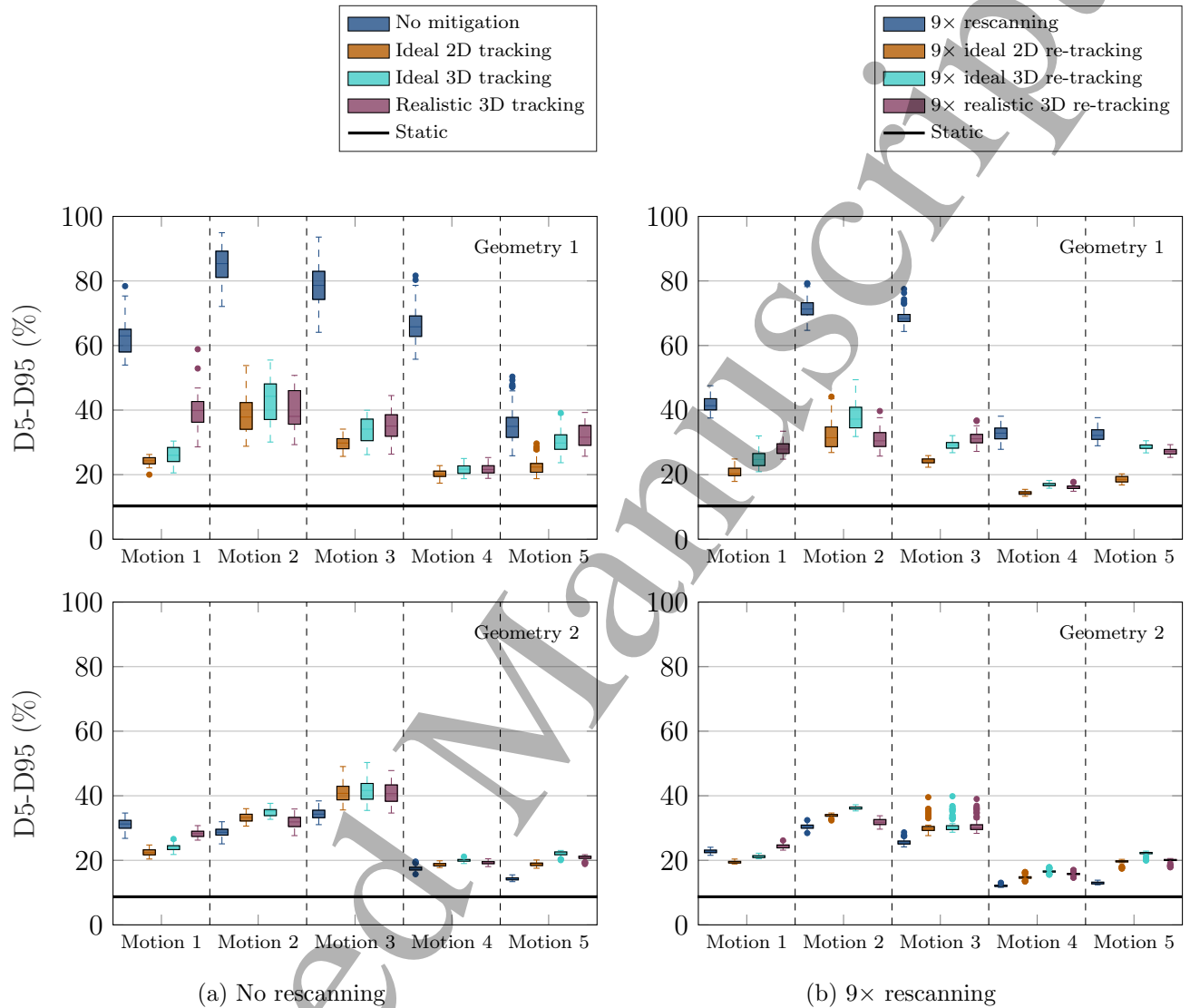
Figure 7: CTV dose homogeneity in terms of D5-D95 % for all geometry / motion scenarios, (a) without and (b) with rescanning, in comparison to the reference dose. Top: CT geometry 1 (larger motion amplitude), bottom: CT geometry 2 (smaller motion amplitude). The box plots include simulation results for all starting phases.

performance of the motion predictions might in fact be further improved by enhancing the AR model. However, the analysis of the obtained dose distributions suggests that the linear AR model in combination with GPR substantially improves the dose homogeneity for tumour tracking in PBS proton therapy.

Both parts of the motion model are patient-specific, meaning that they were trained for each geometry / motion pattern separately. This enables the models to estimate patient-specific variability better than a population-based model. However, in order to make such models more suitable for clinical practice, the same parameters were used for all geometry / motion scenarios. Nevertheless, the models were shown to predict

the DVFs with reasonable accuracy. This demonstrates the strength of the applied respiratory motion model.

However, one needs to keep in mind that all US and MRI data sets were acquired in one session per volunteer, without any re-positioning of the volunteer or the US probe. In order to apply this tracking framework in clinical practice, the motion models will have to be trained on a pretreatment imaging data set, while the actual treatment, and thus the test set of the model, will happen days or even weeks after the imaging session and will be fractionated over several weeks. It is unclear, therefore, whether and how the motion of a patient will change on an inter-fractional basis, rendering the pretrained motion models less accurate. Additionally, it is impracticable to achieve the same imaging plane of the US probe for every set-up of the patient, posing another challenge to the model. A feasibility study has recently been presented by Giger et al. (2019) to investigate the influence of inter-fractional positioning changes. In this study, the volunteers were asked to stand up between two MRI scans and the US probe was removed. Afterwards, the volunteers and the US probe were repositioned as reproducibly as possible. The motion model trained on the first session resulted in larger prediction errors for the second session, showing the adverse influence of the repositioning. But even in this study, the data sets were not truly inter-fractional since they were acquired within a few minutes after each other. A longitudinal study is therefore planned where the volunteers will be scanned over a time frame of weeks or months. These data sets will lead to more clinically relevant results. Also, for future data acquisition, we will use an improved set-up to synchronise the US and MRI acquisitions, enabling longer imaging sessions without missing US frames.

Since this work is based on the same data as our previous study (Giger et al. 2020), it inherently has similar limitations. In particular, the analysis was performed on 4DCT(MRI) data sets whose motion is restricted to the lung volume only. Consequently, density heterogeneities proximal to the target, such as ribs, were static in all our calculations. A previous simulation study on simple heterogeneous motion phantoms suggests that tumour tracking in heterogeneous targets may have detrimental effects on the dose distribution (van de Water et al. 2009). In our analysis, such effects were considered by the tracking simulations which altered the spot positions relative to the rib cage. Moreover, the motion extracted from the 4DMRI of healthy volunteers, although realistic, might not perfectly represent the motion characteristics of lung cancer patients. As such, for CT geometry 1 and motions 3 and 4, a clear under-dosage of the CTV was observed. Similarly, for motion 5, an over-dosage was found, even though for all cases tracking improved the dose homogeneity substantially. These changes in the dose level are a result of tumour stretching / compression with respect to the reference state during the delivery. If a tumour is stretched during delivery, the anatomical parts of the tumour move further apart and thus, by applying tracking, the pencil beams will be placed further apart as well. Since the number of protons per spot is unchanged, this will result in a lower proton fluence and thus a lower dose. The analogue explanation holds for tumour compression. It is, however, questionable whether such

Ultrasound-guided lung tumour tracking 18

pronounced stretching or compression of a lung tumour represents reality. In this study, we extracted the motion from healthy volunteers and applied it to a lung cancer case. It is not given that a lung lesion moves in the same way as a healthy lung without any lesion, thus the above described effect may not be an issue for a real lung tumour motion. For this reason, it will be important to extend such a study to real patient data sets, once they will be available. It should be pointed out, however, that the model performance is expected to be comparable although the motion patterns have been extracted from healthy volunteers instead of lung cancer patients since the model parameters are subject-specific. If the respiratory motion in the pre-treatment imaging stage is representative for the motion during the treatment, the model is expected to be equally predictive for lung cancer patients as it is for healthy volunteers, even if the respiratory motion patterns might differ.

Due to residual motion effects, tracking alone may not be able to reproduce the dose conformality of the reference plan. On the one hand, deformations of the target volume cannot be precluded in real patient data, leading to effects similar to the ones described above. On the other hand, changes of density heterogeneities proximal to the target will distort the shape of the pencil beams, leading to non-nominal dose distributions which in turn can reduce the homogeneity of the total delivered dose. In addition, non-rigid motion of the geometry will lead to inverse interplay effects. In a tracking framework, the beam position is typically adjusted according to the target displacement at Bragg peak depth. However, the anatomy proximal to the Bragg peak does not necessarily move in the same way, particularly in a lung treatment where the rib cage typically exhibits a much smaller motion than a lung tumour. This effect leads to additional deformations of the dose distribution with respect to the patient geometry. The above described residual motion effects could be greatly reduced by combining beam tracking with 4D optimisation methods, as presented for example by Eley, Newhauser, Lüchtenborg, Graeff & Bert (2014). There, the authors include a patient 4DCT in the plan optimisation procedure and optimise the pencil beam weights by assigning each to a certain motion phase of the 4DCT. This method leads to an additional degree of freedom for the optimisation, potentially allowing for better target coverage and healthy tissue sparing. Their approach however relies on a periodic respiration, which is not realistic, as we have found in this study. Alternatively, the weight of each pencil beam could be re-optimised in an online fashion to correct for motion effects of the previously delivered pencil beams (Lüchtenborg, Saito, Durante & Bert 2011).

The results found in this study suggest that 2D tracking can mitigate the motion effects similarly well as full 3D tracking, even though it does not correct for changes in water-equivalent path length. This is of interest for the clinical implementation of tracking, since one of the major challenges is the fast energy change needed for full 3D tracking. If further studies confirm that 2D tracking is similarly effective as 3D tracking, this would simplify the technical implementation of the tracking framework.

Ultrasound-guided lung tumour tracking   19

5. Conclusion

In this paper, we demonstrated that the accuracy of the US-guided spatio-temporal motion model was sufficient for PBS proton beam tracking, with comparable plan quality achieved for realistic and ideal tracking. However, the dosimetric advantage of 3D tracking in comparison to rescanning is limited, indicating the necessity of applying online plan optimisation and/or the combination with other motion mitigation strategies.

Acknowledgments

References

Arnold, P., Preiswerk, F., Fasel, B., Salomir, R., Scheffler, K. & Cattin, P. C. (2011). 3D organ motion prediction for MR-guided high intensity focused ultrasound, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, pp. 623–630.

Bert, C. & Durante, M. (2011). Motion in radiotherapy: particle therapy, *Physics in Medicine & Biology* **56**(16): R113.

Bert, C., Gemmel, A., Saito, N., Chaudhri, N., Schardt, D., Durante, M., Kraft, G. & Rietzel, E. (2010). Dosimetric precision of an ion beam tracking system, *Radiation Oncology* **5**(61).

Bertholet, J., Knopf, A., Eiben, B., McClelland, J., Grimwood, A., Harris, E., Menten, M., Poulsen, P., Nguyen, D. T., Keall, P. et al. (2019). Real-time intrafraction motion monitoring in external beam radiotherapy, *Physics in Medicine & Biology* **64**(15): 15TR01.

Boye, D., Lomax, A. J. & Knopf, A. C. (2013). Mapping motion from 4D-MRI to 3D-CT for use in 4D dose calculations: A technical feasibility study, *Med. Phys.* **40(6)**: 061702–1 – 061702–11.

Boye, D., Samei, G., Schmidt, J., Székely, G. & Tanner, C. (2013). Population based modeling of respiratory lung motion and prediction from partial information, *Medical Imaging 2013: Image Processing*, SPIE. doi: 10.1117/12.2007076.

Celicanin, Z., Giger, A., Bauman, G., Cattin, P. & Bieri, O. (2017). Temporally-resolved volumetric imaging (4DMRI) of the lungs, *Proceedings of the ISMRM 25th Annual Meeting and Exhibition*, Honolulu, Hawaii, USA.

Depuydt, T., Verellen, D., Haas, O., Gevaert, T., Linthout, N., Duchateau, M., Tournel, K., Reynders, T., Leysen, K., Hoogeman, M. et al. (2011). Geometric accuracy of a novel gimbals based radiation therapy tumor tracking system, *Radiotherapy and Oncology* **98**(3): 365–372.

Eley, J. G., Newhauser, W. D., Lüchtenborg, R., Graeff, C. & Bert, C. (2014). 4D optimization of scanned ion beam tracking therapy for moving tumors, *Physics in Medicine and Biology* **59**(13): 3431–3452.

Ernst, F., Dürichen, R., Schlaefer, A. & Schweikard, A. (2013). Evaluating and comparing algorithms for respiratory motion prediction, *Physics in Medicine & Biology* **58**(11): 3911.

Giger, A., Jud, C., Nguyen, D., Krieger, M., Zhang, Y., Lomax, A. J., Bieri, O., Salomir, R. & Cattin, P. C. (2019). Inter-fractional Respiratory Motion Modelling from Abdominal Ultrasound: A Feasibility Study, *International Workshop on PRedictive Intelligence In MEdicine*, Springer, pp. 11–22.

Ultrasound-guided lung tumour tracking   20

Giger, A., Krieger, M., Jud, C., Duetschler, A., Salomir, R., Bieri, O., Bauman, G., Nguyen, D., Weber, D. C., Lomax, A., Zhang, Y. & Cattin, P. C. (2020). Liver-ultraound based motion modelling to estimate 4D dose distributions for lung tumours in scanned proton therapy, *Phys. Med. Biol.* p. Accepted for publication. doi: 10.1088/1361-6560/abaa26.

Grözinger, S. O., Bert, C., Haberer, T., Kraft, G. & Rietzel, E. (2008). Motion compensation with a scanned ion beam: a technical feasibility study, *Radiation Oncology* **3**(1): 34.

Hoogeman, M., Prévost, J.-B., Nuyttens, J., Pöll, J., Levendag, P. & Heijmen, B. (2009). Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files, *International Journal of Radiation Oncology* Biology* Physics* **74**(1): 297–303.

Jud, C., Nguyen, D., Sandkühler, R., Giger, A., Bieri, O. & Cattin, P. C. (2018). Motion Aware MR Imaging via Spatial Core Correspondence, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, pp. 198–205.

Knopf, A.-C., Stützer, K., Richter, C., Rucinski, A., da Silva, J., Phillips, J., Engelsman, M., Shimizu, S., Werner, R., Jakobi, A. et al. (2016). Required transition from research to clinical application: report on the 4D treatment planning workshops 2014 and 2015, *Physica Medica* **32**(7): 874–882.

Krauss, A., Nill, S. & Oelfke, U. (2011). The comparative performance of four respiratory motion predictors for real-time tumour tracking, *Physics in Medicine & Biology* **56**(16): 5303.

Krauss, A., Nill, S., Tacke, M. & Oelfke, U. (2011). Electromagnetic real-time tumor position monitoring and dynamic multileaf collimator tracking using a Siemens 160 MLC: Geometric and dosimetric accuracy of an integrated system, *International Journal of Radiation Oncology* Biology* Physics* **79**(2): 579–587.

Krieger, M., Giger, A., Weber, D. C., Lomax, A. J. & Zhang, Y. (2020). Impact of internal target volume definition for pencil beam scanned proton treatment planning in the presence of respiratory motion variability, *Radiotherapy and Oncology* **145**: 154–161.

Krieger, M., Klimpki, G., Fattori, G., Hrbacek, J., Oxley, D., Safai, S., Weber, D. C., Lomax, A. J. & Zhang, Y. (2018). Experimental validation of a deforming grid 4D dose calculation for PBS proton therapy, *Physics in Medicine & Biology* **63**(5): 055005.

Lüchtenborg, R., Saito, N., Durante, M. & Bert, C. (2011). Experimental verification of a real-time compensation functionality for dose changes due to target motion in scanned particle therapy, *Medical Physics* **38**(10): 5448–5458.

McClelland, J. R., Hawkes, D. J., Schaeffter, T. & King, A. P. (2013). Respiratory motion models: a review, *Medical image analysis* **17**(1): 19–42.

Mostafaei, F., Tai, A., Gore, E., Johnstone, C., Haase, W., Ehlers, C., Cooper, D. T., Lachaine, M. & Li, X. A. (2018). Feasibility of real-time lung tumor motion monitoring using intrafractional ultrasound and kV cone beam projection images, *Medical physics* **45**(10): 4619–4626.

Pedroni, E., Bearpark, R., Böhringer, T., Coray, A., Duppich, J., Forss, S., George, D., Grossmann, M., Goitein, G., Hilbes, C., Jermann, M., Lin, S., Lomax, A., Negrazus, M., Schippers, M. & Kotrle, G. (2004). The PSI Gantry 2: a second generation proton scanning gantry, *Zeitschrift für Medizinische Physik* **14**: 25–34.

Pedroni, E., Meer, D., Bula, C., Safai, S. & Zenklusen, S. (2011). Pencil beam characteristics of the next-generation proton scanning gantry of PSI: design issues and initial commissioning results, *The European Physical Journal Plus* **126**(7): 66.

Phillips, M. H., Pedroni, E., Blattmann, H., Boehringer, T., Coray, A. & Scheib, S. (1992). Effects of respiratory motion on dose uniformity with a charged particle scanning method, *Physics in Medicine & Biology* **37**(1): 223.

Poulsen, P. R., Cho, B., Ruan, D., Sawant, A. & Keall, P. J. (2010). Dynamic multileaf collimator tracking of respiratory target motion based on a single kilovoltage imager during arc radiotherapy, *International Journal of Radiation Oncology* Biology* Physics* **77**(2): 600–607.

Prall, M., Kaderka, R., Saito, N., Graeff, C., Bert, C., Durante, M., Parodi, K., Schwaab, J., Sarti, C. & Jenne, J. (2014). Ion beam tracking using ultrasound motion detection, *Medical Physics*

Ultrasound-guided lung tumour tracking   21

41(4): 041708.

Preiswerk, F., De Luca, V., Arnold, P., Celicanin, Z., Petrusca, L., Tanner, C., Bieri, O., Salomir, R. & Cattin, P. C. (2014). Model-guided respiratory organ motion prediction of the liver from 2D ultrasound, *Medical image analysis* **18**(5): 740–751.

Rietzel, E. & Bert, C. (2010). Respiratory motion management in particle therapy, *Medical physics* **37**(2): 449–460.

Safai, S., Bula, C., Meer, D. & Pedroni, E. (2012). Improving the precision and performance of proton pencil beam scanning, *Transl Cancer Res* **1**(3).

Sandkühler, R., Jud, C., Andermatt, S. & Cattin, P. C. (2018). AirLab: Autograd image registration laboratory, *arXiv preprint arXiv:1806.09907* .

Santini, F., Gui, L., Lorton, O., Guillemin, P. C., Manasseh, G., Roth, M., Bieri, O., Vallée, J.-P., Salomir, R. & Crowe, L. A. (2020). Ultrasound-driven cardiac MRI, *Physica Medica* **70**: 161–168.

Schwaab, J., Prall, M., Sarti, C., Kaderka, R., Bert, C., Kurz, C., Parodi, K., Günther, M. & Jenne, J. (2014). Ultrasound tracking for intra-fractional motion compensation in radiation therapy, *Physica Medica* **30**(5): 578 – 582.

van de Water, S., Kreuger, R., Zenklusen, S., Hug, E. & Lomax, A. (2009). Tumour tracking with scanned proton beams: assessing the accuracy and practicalities, *Phys. Med. Biol.* **54**: 6549–6563.

Williams, C. K. & Rasmussen, C. E. (2006). *Gaussian processes for machine learning*, Vol. 2, MIT press Cambridge, MA.

Zhang, Y., Huth, I., Weber, D. C. & Lomax, A. J. (2018). A statistical comparison of motion mitigation performances and robustness of various pencil beam scanned proton systems for liver tumour treatments, *Radiotherapy and Oncology* **128**(1): 182–188.

Zhang, Y., Knopf, A., Tanner, C., Boye, D. & Lomax, A. J. (2013). Deformable motion reconstruction for scanned proton beam therapy using on-line x-ray imaging, *Physics in Medicine & Biology* **58**(24): 8621.

Zhang, Y., Knopf, A., Tanner, C. & Lomax, A. (2014). Online image guided tumour tracking with scanned proton beams: a comprehensive simulation study, *Physics in Medicine & Biology* **59**(24): 7793.

Appendix A. Motion characteristics

Figure A1 shows an overview of the CTV motion amplitude in all three directions as well as the respiratory periods for all CT geometry/motion scenarios. The boxplots include all breathing cycles included in this study.

Ultrasound-guided lung tumour tracking   22

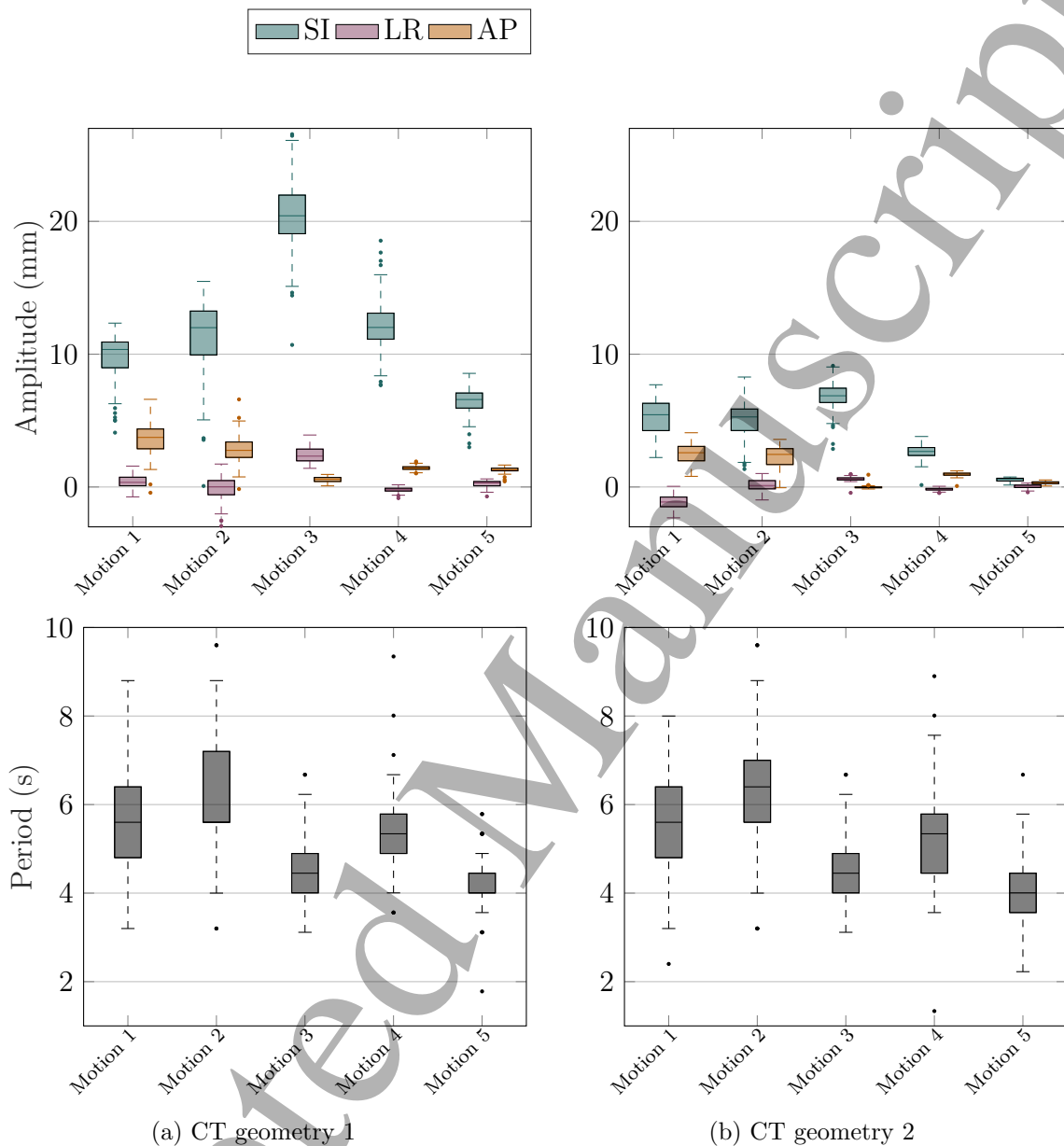(a) CT geometry 1   (b) CT geometry 2

Figure A1: Top row: summary of the CTV motion amplitudes in superior-inferior (SI), left-right (LR), and anterior-posterior (AP) direction, bottom row: motion periods.