Stochastic formulation of patient positioning using linac-mounted cone beam imaging with prior knowledge

W. Hoegele[a)]
Department of Radiation Oncology, Regensburg University Medical Center,
93053 Regensburg, Germany and Department of Computer Science and Mathematics,
University of Applied Sciences, 93053 Regensburg, Germany

R. Loeschel
Department of Computer Science and Mathematics, University of Applied Sciences,
93053 Regensburg, Germany

B. Dobler
Department of Radiation Oncology, Regensburg University Medical Center, 93053 Regensburg, Germany

J. Hesser
Medical Faculty of Mannheim, Heidelberg University, 68167 Mannheim, Germany

O. Koelbl
Department of Radiation Oncology, Regensburg University Medical Center, 93053 Regensburg, Germany

P. Zygmanski
Department of Radiation Oncology, Brigham and Women's Hospital and Harvard Medical School,
Boston, Massachusetts 02115

Purpose: In this work, a novel stochastic framework for patient positioning based on linac-mounted CB projections is introduced. Based on this formulation, the most probable shifts and rotations of the patient are estimated, incorporating interfractional deformations of patient anatomy and other uncertainties associated with patient setup.

Methods: The target position is assumed to be defined by and is stochastically determined from positions of various features such as anatomical landmarks or markers in CB projections, i.e., radiographs acquired with a CB-CT system. The patient positioning problem of finding the target location from CB projections is posed as an inverse problem with prior knowledge and is solved using a Bayesian maximum *a posteriori* (MAP) approach. The prior knowledge is three-fold and includes the accuracy of an initial patient setup (such as in-room laser and skin marks), the plasticity of the body (relative shifts between target and features), and the feature detection error in CB projections (which may vary depending on specific detection algorithm and feature type). For this purpose, MAP estimators are derived and a procedure of using them in clinical practice is outlined. Furthermore, a rule of thumb is theoretically derived, relating basic parameters of the prior knowledge (initial setup accuracy, plasticity of the body, and number of features) and the parameters of CB data acquisition (number of projections and accuracy of feature detection) to the expected estimation accuracy.

Results: MAP estimation can be applied to arbitrary features and detection algorithms. However, to experimentally demonstrate its applicability and to perform the validation of the algorithm, a water-equivalent, deformable phantom with features represented by six 1 mm chrome balls were utilized. These features were detected in the cone beam projections (XVI, Elekta Synergy®) by a local threshold method for demonstration purposes only. The accuracy of estimation (strongly varying for different plasticity parameters of the body) agreed with the rule of thumb formula. Moreover, based on this rule of thumb formula, about 20 projections for 6 detectable features seem to be sufficient for a target estimation accuracy of 0.2 cm, even for relatively large feature detection errors with standard deviation of 0.5 cm and spatial displacements of the features with standard deviation of 0.5 cm.

Conclusions: The authors have introduced a general MAP-based patient setup algorithm accounting for different sources of uncertainties, which are utilized as the prior knowledge in a transparent way. This new framework can be further utilized for different clinical sites, as well as theoretical developments in the field of patient positioning for radiotherapy. © *2011 American Association of Physicists in Medicine*. [DOI: 10.1118/1.3532959]

I. INTRODUCTION

Accurate patient positioning in the framework of image guided radiotherapy (IGRT) is an important component in the clinical application of radiation oncology with linear accelerators.[1] Its aim is to determine the optimal position of the patient directly before, and sometimes during, the actual treatment fraction to fit optimally to the treatment plan, which is based on a high-quality computed tomography (CT), obtained during simulation. In order to bring the patient to treatment position, the patient can be moved rigidly, translated, and also slightly rotated with the treatment couch. To get information about the internal structures of the patient, photon transmission imaging techniques are typically utilized with either a cone beam CT (CB-CT) mounted orthogonally to the gantry of the MV beamline for kV applications or based on electronic portal imaging device for MV applications. An intrinsic problem of patient positioning is that interfractional changes inside the body of the patient are, in general, of nonrigid nature, showing the necessity of a well weighted compromise for a suitable patient setup.

Typically, one can differentiate between two principle approaches: Patient positioning based on reconstructed images or based directly on projection images (radiographs) measured directly without reconstruction. In the first approach, a large set of projections is utilized to reconstruct the 3D volume of the current patient's anatomy first and in a second step, this 3D data set is matched to the 3D volume of the planning CT, leading to a two-step approach: Reconstruction followed by image registration.[2–4] In the second approach, typically a few radiographs are directly used to find the target location, such as the classical orthogonal radiography or related techniques.[5–8]

Both of these approaches have inherent problems, such as reconstruction artifacts (e.g., metal or motion artifacts or artifacts due to the undersampling of projections in limited view reconstructions) in the first approach or the presence of background structures and scatter in the radiographs in the second one. Therefore, both lead to challenging image registration problems which are a wide field of study and developments.[9–11]

Classically, image registration algorithms are divided into feature based (working with the identification of certain visible features) and intensity based techniques (working with the intensity levels acquired). The types of mappings to match an image onto the reference data set are further classified as rigid, affine, and elastic matching.[10,12] For example, a standard case for image registration in medicine is point-pair-matching, which solves the problem based on features with a rigid matching algorithm, leading to direct analytical results based on geometrical considerations.[13] This approach typically utilizes markers or anatomical landmarks as features, thus leading to robust and fast methods. In contrast with this, intensity based image registration is also often applied since it does not depend on artificially introduced markers and valuable information is contained in the acquired or reconstructed intensity levels.

The clinical availability of linac-mounted cone beam imaging systems is widespread and their application represents a standard approach to patient positioning in many clinics. Accurate patient setup can be performed by image registration of the fully reconstructed CB-CT with the planning CT. With the increased number of CB-CT scans per patient and sometimes multiple scans per single treatment, there is a justified concern about dose from CB-CT imaging. There are also equally valid concerns about the time per scan and problems with gantry clearance, especially for tumor sites such as breast, extremities, or soft tissue tumors (sarcomas), which are located further away from the patient craniocaudal axis.

One way to overcome these difficulties is to perform a partial CB-CT scan (with angles less than 210°) or a very short scan, such as CB tomosynthesis[14] (CB-TS) or, similarly, digital tomosynthesis (DTS) (with angles less than 40°). However, partial and short scans result in undersampled CT reconstructions and lead to reconstruction artifacts that may prohibit their efficient use for the registration needed for patient setup.[14–16] Because of these reconstruction artifacts, in short CB-CT, CB-TS, and DTS scans, the registration of the reference (planning) data is much more difficult than in full-rotation CB-CT. The reconstruction methods for short scans are not established clinically and are currently the focus of research in the patient setup area.

Furthermore, in the mentioned approaches, prior knowledge is not fully utilized. Specifically, the prior knowledge, which can be derived from multiple CT, fluoroscopy, multiple treatment CB-CT, or other imaging studies, performed on a given patient or obtained from literature on the subject for a given treatment site. This prior knowledge includes (a) the range of possible setup errors which are known for a given treatment site and patient immobilization and (b) the possible range of deformations of patient anatomy surrounding the tumor. Furthermore, (c) additional knowledge about specific reconstruction and registration algorithms might be available, such as detection accuracy of features of interest used for the patient setup. All this prior knowledge may provide valuable information, which in principle should lead to better quality reconstruction and registration, but it is difficult to consistently account for it in the present algorithms.

To introduce prior knowledge to inverse problems with intrinsic uncertainties, a stochastic Bayesian framework can typically be utilized. This is especially true in image registration for medical applications, where a whole set of prior information is available from clinical practice (e.g., Ref. 17). But also in other fields, such as in machine learning, this concept is applied broadly, e.g., in Gaussian processes for regression purposes.[18]

In the following, we present a rigid, feature based patient positioning approach, utilizing either all or only a few radiographs (CB projections). This approach has its roots in the estimation theory and especially applies the Bayesian formulation for the inclusion of the aforementioned prior knowledge [(a)–(c)]. In consequence, this method can be interpreted as a direct extension of projection based registration techniques, such as orthogonal projections. The formulation is independent of any particular reconstruction or registration algorithm and is based on a general stochastic-geometrical

model of the patient. Thanks to this formulation, we derive formulas for most probable shifts and rotations of the patient required to align the patient before treatment.

Specifically, the mathematical formulation is based on the Bayesian maximum *a posteriori* (MAP) estimation. The specific use of MAP techniques is widely spread in medical physics, e.g., to increase image quality in image reconstruction (e.g., Ref. 19) or emission tomography (e.g., Refs. 20 and 21) and in segmentation algorithms (e.g., Refs. 22 and 23), but to our knowledge it has not been used for patient setup posed herein. In this context, the main reason for the utilization of estimation theory is to account for the intrinsic uncertainties (noise, spatial uncertainties, etc.) of an inverse problem on a solid basis. The idea of utilizing MAP estimation for the particular problem at hand is to include prior knowledge in a transparent way in order to increase the accuracy of the determination of rigid transformation parameters for patient positioning. It includes a clear stochastical model of the uncertainties, having an intuitive understandable objective function, delivering a physically interpretable weighting between prior knowledge [(a)–(c)] and actual measurements in the radiographs and, therefore, eventually leading to a robust algorithm. The latter can be understood by the fact that MAP can be regarded as a direct extension of maximum likelihood (ML) estimation including prior knowledge by additive regularization terms (similar to Tikhonov regularization), which increase the robustness.

II. THEORY

II.A. Statement of the problem

In this work, the patient setup problem is essentially defined as a target localization problem. The assumption is that a patient is initially positioned based on an in-room laser system and then a sequence of cone beam projections is acquired along a short or longer arc or along multiple arcs (such as in Fig. 2 for the experimental verification in Sec. III A). From these images and from the prior knowledge, which is formulated here in stochastic terms, the most probable shifts and rotations are determined that bring the patient into the desired treatment position. Since the target itself may not be identifiable in the radiographs, other anatomical features have to be considered for localization, representing an inverse problem.

The features considered here are as general as possible, markers, regions containing bones or soft tissue-interfaces, with the number of features ranging from a few (e.g., markers) to large (anatomical features in the whole imaged area). For instance, for *nondeforming regions* (such as for cranial structures), it is sufficient to localize three positions in two orthogonal projections to determine the six-parameter transformation.[5] Therefore, having more than three features or using more projections in this case creates a certain redundancy. On the other hand, not all features may be accurately localizable in all projections or some of their coordinates may not be determinable, which is the case, for instance, for prostate-rectum interface along the inferior-superior direction.
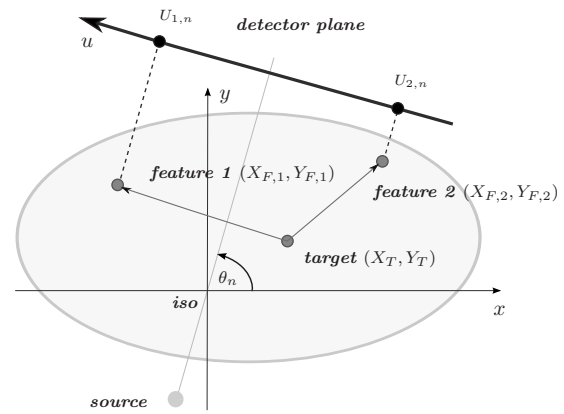

FIG. 1. Schematical view of scanned object with the random variables introduced.

Because of the possibility of *interfractional nonlinear deformations* of the patient anatomy, the geometrical relation of the features to the target must contain these uncertainties, which are best described by probability distributions [as prior knowledge (b)]. Furthermore, the expected patient *setup error* after the initial laser alignment is typically known for a given setup technique and treatment site (e.g., from literature or from clinical experience) and therefore also can be described by their respective probability distributions [as prior knowledge (a)]. In addition, the positions of the anatomical features are detectable in the projections only with certain detection errors due to the low quality of the radiographs, which are also described by respective probability distributions [as prior knowledge (c)]. With such a three-fold stochastic prior knowledge and with the projections acquired in pretreatment position, the desired treatment target position is determined with respect to the position in the reference (planning) data.

First, we provide a simplified analytical exposition of the proposed MAP method, focusing on a 2D parallel beam geometry for the estimation of the target shift only. The more general formulation in 3D including *translations* and *rotations* of the patient anatomy for cone beam geometry is presented in Appendix A. The derivations of the 2D and 3D problem use analogical arguments.

II.B. Patient setup variables

Determination of target position will be achieved working with features, such as implanted markers, in projections that have deformable geometrical relations to the target. We assume that these features are detectable with a certain probability.

All probability distributions of the stochastic-geometrical model are *Gaussians* in this work, due to a straightforward introduction to the theoretical framework and the possibility of direct analytic solutions (besides reasoning from clinical patient positioning studies, e.g., Ref. 24). In Fig. 1, the essential variables of this model are illustrated in a simple overview. In general, we will denote *random variables* with capital letters and their *realizations* (their concrete measur-

able or determinable values) with small letters.

We define the detected feature positions in the projections as our *measurements*. This includes the acquisition of the projections and an (at this point) unspecified detection algorithm that is able to identify the position of the features with a certain accuracy. Out of these measurements, the target position will then be estimated in an optimal MAP fashion.

The unknown target position after the initial patient setup will be described by the two random variables for each coordinate

$$X_T \sim \mathcal{N}(\mu_{X_T}, \sigma_T^2), \quad Y_T \sim \mathcal{N}(\mu_{Y_T}, \sigma_T^2), \tag{1}$$

where $\mathcal{N}(\mu, \sigma^2)$ represents the Gaussian distribution, with expectation value μ and standard deviation σ (index T for target). Furthermore, the relative anatomical relation between the target and the features (independent of the current target position in the coordinate system) will be described with the probability distributions

$$X_{F,l} \sim \mathcal{N}(\mu_{X_{F,l}}, \sigma_F^2), \quad Y_{F,l} \sim \mathcal{N}(\mu_{Y_{F,l}}, \sigma_F^2), \tag{2}$$

for all features $l=1,\ldots,L$ (index F for feature). For both abovementioned classes of random variables, we are assuming constant standard deviations σ_T and σ_F for every coordinate direction due to simplicity of the derivation (in Appendix A, these assumptions will be lifted to account for more realistic situations). We will assume that there are $N \geq 1$ projections at angles θ_n, $n=1,\ldots,N$. In every of these n projections, we define additionally a random variable describing the detection error of the feature detection algorithm in the projections with

$$\Delta U_{l,n} \sim \mathcal{N}(0, \sigma_{\Delta U}^2). \tag{3}$$

Therefore, we are expecting that the detection of the features in the projection does not work perfectly, representing an additional, independent source of uncertainty, due to the low quality of the radiographs.

In consequence, the random variable corresponding to the feature positions in the projection at angle θ_n in the treatment coordinate system for the 2D parallel beam model is determined by

$$U_{l,n} := -(X_T + X_{F,l}) \cdot \sin \theta_n + (Y_T + Y_{F,l}) \cdot \cos \theta_n + \Delta U_{l,n} \tag{4}$$

for feature l at projection n (see Fig. 1). Since the distributions for all these random variables are defined, the distribution of the resulting measurement random variables $U_{l,n}$ is also known.

The three different sources of prior knowledge are then identified by their corresponding random variables (and their probability distributions).

(a) The knowledge about the range of realistic target positions (X_T, Y_T) determined by the initial setup technique (such as laser alignment by skin markers), whose expectation values μ_{X_T} and μ_{Y_T} typically are the target position in the planning CT and σ_T is its expected standard deviation of the setup by laser alignment.

(b) The knowledge about the relative (anatomical) distances between the lth feature and the target position $(X_{F,l}, Y_{F,l})$, which means that $\mu_{X_{F,l}}$ and $\mu_{Y_{F,l}}$ are the expected distances between the (fixed) target position and feature l (either just determined out of planning CT or the arithmetic mean of several imaging modalities, see Sec. II E 2) and σ_F represents the plasticity of the body (which range of relative shifts, i.e., deformations inside the body, seem possible).

(c) The knowledge about the detection error $\Delta U_{l,n}$ for the lth feature in projection n, which should have the expectation value 0, since it is expected that the detection algorithm works in average, and the standard deviation $\sigma_{\Delta U}$ which essentially describes the range of possible misdetections.

To provide simple analytical derivations in the following introduction to the MAP estimator, we will additionally assume that the distribution of imaging angles $\{\theta_1, \ldots, \theta_N\}$ fulfills the following conditions:

$$\sum_{n=1}^{N} \sin^2 \theta_n = \frac{N}{2}, \quad \sum_{n=1}^{N} \cos^2 \theta_n = \frac{N}{2}, \quad \sum_{n=1}^{N} \sin \theta_n \cos \theta_n = 0. \tag{5}$$

These conditions look arbitrary at first sight, but they are fulfilled exactly, for example, with either an equidistant angular distribution over the range of 180° *or* by $N/2$ projections at 0° and $N/2$ projections at 90° (for N even). These conditions can also be approximated reasonably by other angular distributions. As a special example the case $N=2$ represents two *orthogonal projections*.

It is pointed out that these assumptions for the angular distribution are not essential and are lifted in the general estimator in Appendix A.

II.C. Target localization using MAP estimation

The MAP argumentation combines prior knowledge with current measurements to estimate parameters of interest. In our 2D parallel beam application (3D case in Appendix A), we specifically want to estimate the *position* of the treatment target in a nonrigid body geometry during the course of fractionation. The starting point for our MAP approach is the conditional posterior probability

$$\arg\max_{(x_T, y_T)} \mathrm{P}(X_T = x_T \cap Y_T = y_T | \bigcap_{l=1}^{L} \bigcap_{n=1}^{N} U_{l,n} = u_{l,n}), \tag{6}$$

which can be translated into words by: Find the most likely parameters (x_T, y_T) (for the current treatment target position) under the assumption that the detection algorithm measured the feature positions at $u_{l,n}$ in the projections. This formula can be rearranged to the equivalent expression

$$\arg\max_{(x_T, y_T)} \mathrm{P}(\bigcap_{l=1}^{L} \bigcap_{n=1}^{N} U_{l,n} = u_{l,n} | X_T = x_T \cap Y_T = y_T)$$

$$\cdot \mathrm{P}(X_T = x_T \cap Y_T = y_T) \tag{7}$$

using *Bayes' Theorem*. The first factor represents the *likelihood*, which is the basis of the closely related maximum likelihood approach. The second factor contains information about probable values for (x_T, y_T) after the initial patient setup, which is our *prior knowledge* (a). By inserting the mathematical derivations of Appendix B and taking the logarithm, this leads to the following final minimization problem:

$$
\arg \min_{(x_T, y_T)} \left(\frac{x_T - \mu_{X_T}}{\sigma_T} \right)^2 + \left(\frac{y_T - \mu_{Y_T}}{\sigma_T} \right)^2
$$
$$
+ \sum_{l=1}^{L} \frac{(x_T + \mu_{X_{F,l}})^2 + (y_T + \mu_{Y_{F,l}})^2}{\sigma_F^2}
$$
$$
- \frac{2\sigma_F^2}{\sigma_{\Delta U}^2 (N\sigma_F^2 + 2\sigma_{\Delta U}^2)} \left[\left(\sum_{n=1}^{N} - u_{l,n} \sin \theta_n + \frac{x_T + \mu_{X_{F,l}}}{\sigma_F^2} \right)^2 \right.
$$
$$
\left. + \left(\sum_{n=1}^{N} u_{l,n} \cos \theta_n + \frac{y_T + \mu_{Y_{F,l}}}{\sigma_F^2} \right)^2 \right], \tag{8}
$$

which we call the *MAP estimator* for the 2D parallel beam case, estimating the target position. This objective function is essentially a quadratic function in the variables (x_T, y_T) and therefore can be solved *analytically*, which is not essential but important for the derivation of the rule of thumb in Sec. II D 1. We denote the estimation values then with $(\hat{x}_T, \hat{y}_T)$. It can be shown that this estimator is stochastically *efficient*, meaning it fulfills the *Cramér–Rao-bound* and therefore this estimator is optimal.

II.D. A practical rule of thumb formula

II.D.1. Derivation

One of the major advantages of utilizing the estimation theory is that the accuracy of the estimator can be investigated in a theoretical way. In this section, a rule of thumb formula will be presented that shows an intuitive relation of the central parameters of acquisition and prior knowledge with the maximal achievable estimation accuracy. Although this formula is derived under the simplifying assumptions of Sec. II B, it has numerically been proved to be valid also for the translational setup error of the 3D cone beam case in Sec. IV B. A typical quantity to investigate the performance of estimators is the mean-squared-error (MSE)

$$
\mathbb{MSE}[\hat{X}_T|_{x_T, y_T}] = \mathbb{VAR}[\hat{X}_T|_{x_T, y_T}] + (\mathbb{BIAS}[\hat{X}_T|_{x_T, y_T}])^2, \tag{9}
$$

with $\mathbb{BIAS}[\hat{X}_T|_{x_T, y_T}] := \mathbb{E}[\hat{X}_T|_{x_T, y_T}] - x_T$ as the bias of the estimator and $\mathbb{VAR}[\hat{X}_T|_{x_T, y_T}]$ its variance. The MSE determines the expected quadratic error in between *estimated values* $(\hat{x}_T, \hat{y}_T)$ and the *real values* (x_T, y_T).

In our application, the MAP estimator is biased, meaning that $\mathbb{BIAS}[\hat{X}_T|_{x_T, y_T}] \neq 0$. In Sec. II D 2, we will show that the unbiased version (which corresponds to the maximum likelihood case) has a larger MSE and is therefore overall not as accurate. General investigations and examples on the profits of biased estimations are illustrated by Eldar.[25]

To get a physically intuitive measure for the *estimation*

error, we take the square root of the expectation value of the MSE, denoted as the Bayesian root-mean-squared-error (RMSE)

$$
\mathbb{RMSE} := \sqrt{\mathbb{E}_{X_T, Y_T}[\mathbb{MSE}[\hat{X}_T | X_T, Y_T]]}, \tag{10}
$$

(analog for $\hat{Y}_T$). Numerically, one can evaluate this expression by simulations for every type of geometry as we do it in Sec. IV B. For our specific MAP estimator in patient positioning with the assumptions mentioned, this leads to

$$
\mathbb{RMSE} = \sqrt{\frac{(\sigma_F^2 N + 2\sigma_{\Delta U}^2)\sigma_T^2}{\sigma_F^2 N + 2\sigma_{\Delta U}^2 + LN\sigma_T^2}} \tag{11}
$$

(the same result for *y*-direction), which we will denote as the *rule of thumb formula* for the expected *estimation error*. This formula has a central meaning for the interpretation of the estimation situation and can be regarded as a rather general rule of thumb for translational setup error determination by projections with an x-ray source on a circular trajectory for a deforming body. This formula shows the relation of the expected patient setup accuracy, represented by the RMSE, in relation to the number of detectable features L, the number of projections N, the uncertainty by the initial patient setup σ_T, the uncertainty by interfractional organ motion between features and target σ_F (essentially the plasticity of the body), and the detection error of features in the projections $\sigma_{\Delta U}$.

II.D.2. Asymptotic behavior

In the following, we will investigate the asymptotic behavior of the rule of thumb of Sec. II D 1 to investigate the estimation accuracy. In consequence, we will gain insight in the use of different factors in patient positioning in general.

Case $N \to \infty$. By using a large number of projections in which we can detect L features we get

$$
N \to \infty \Rightarrow \mathbb{RMSE} \to \sqrt{\frac{\sigma_F^2 \sigma_T^2}{\sigma_F^2 + L\sigma_T^2}}. \tag{12}
$$

This leads to the conclusion that having a lot of projections but not being able to adequately interpret the data (not detecting enough features with a spatial relation to the target) will lead to the maximal achievable limit of certainty of the estimation as in Eq. (12). Furthermore, the estimation error $\mathbb{RMSE}$ is independent of $\sigma_{\Delta U}$: Being able to detect the features in several projections reduces the influence introduced by the detection error, as one would expect.

This formula can be regarded as a lower limit for the patient setup error one can achieve independently of the acquisition, which is only depending on the in-principle anatomical parameters L and σ_F, and the prior knowledge about the initial setup σ_T.

Case $L \to \infty$. By being able to identify an infinite number of features, we see that

$$
L \to \infty \Rightarrow \mathbb{RMSE} \to 0. \tag{13}
$$

Therefore, identifying as many features as possible should be an essential task of every detection algorithm for an accurate patient setup. It has to be pointed out that $\mathbb{RMSE} \to 0$ inde-

pendently of the other quantities (e.g., amount of uncertainties, number of projections, etc.). Since it is not possible to identify the position for an infinite number of features, a trade-off is necessary, which we will discuss in Sec. IV B.

Case $\sigma_T \to \infty$. Decreasing the certainty (increasing the standard deviation) of the prior knowledge about the initial patient setup leads to the corresponding ML estimation (without any prior knowledge about the initial patient setup). For such a situation, we get

$$\sigma_T \to \infty \Rightarrow \mathbb{RMSE} \to \sqrt{\frac{\sigma_F^2}{L} + 2\frac{\sigma_{\Delta U}^2}{LN}}. \tag{14}$$

To investigate the *gain* in accuracy of estimation by using prior knowledge about the accuracy of the initial setup technique with MAP in comparison with the case with not utilizing this prior knowledge with ML, we regard the fraction

$$\frac{\mathbb{RMSE}_{ML}}{\mathbb{RMSE}_{MAP}} = \sqrt{1 + \frac{\sigma_F^2}{L\sigma_T^2} + 2\frac{\sigma_{\Delta U}^2}{LN\sigma_T^2}} > 1. \tag{15}$$

This means, as long as there is any prior knowledge about the initial setup accuracy ($\sigma_T < \infty$) in a deformable body ($\sigma_F > 0$), the estimation is expected to benefit from it. The third summand has the detection error $\sigma_{\Delta U}$ in the numerator (which is typically rather small in comparison to the other uncertainties) and if we additionally use a high N, the third summand disappears completely. Therefore, we should regard the second summand as the more essential component of the gain.

II.E. Clinical application of MAP estimation for patient positioning

II.E.1. General workflow

After the patient is positioned with an initial setup technique (such as in-room lasers and skin marks), CB projections are acquired from different imaging angles. In the acquired projections, the position of features, such as metallic fiducials, is determined by a detection algorithm. It has to be pointed out that it is not very important what specific detection algorithm is actually used, since after about 20 projections, according to Eq. (11), the influence of the algorithms on estimation accuracy is strongly reduced (see Sec. IV B).

Based on the position of features and the prior knowledge [(a)–(c)] (see Sec. II E 2), we are performing MAP estimation for the target position and its rotation by solving the minimization problem, whose full exposition is in Appendix A. The results of minimization are the estimates of the three coordinates of the target position in the treatment coordinate system and the three angles of rotation. It has to be pointed out that the objective function is almost of quadratical shape in the translational parameters and the minimum can therefore be calculated efficiently. The actual transformation of the patient is then determined by Eq. (A1) in Appendix A.

During the course of fractionation, systematic errors can be separated from random errors of the setup technique and the MAP estimator can be adapted accordingly. For example, Bortfeld *et al.*[26] proposed when systematic errors should be

corrected (almost best after the fourth fraction). For this reason, the arithmetic mean of the differences of the reference target position and the estimated target position of the last fractions can be calculated. In consequence, the parameters $\{\mu_{X_T}, \mu_{Y_T}, \mu_{Z_T}\}$ can be updated during the course of the treatment.

Moreover, a general concept of prior knowledge in the form of images acquired prior to treatment, fully reconstructed CB-CT during fractionation or tomosynthetic scans might be utilized to improve the model of the prior knowledge continuously and being adaptive to systematic long-term changes inside the body, such as the shrinking of the target, during the course of fractionation. We suggest that these different approaches for determining the setup error should work together for better patient positioning. In consequence, an individualized patient positioning method can be derived easily with the aim of an even more accurate patient setup (based on the presented model) than using one standard technique for all patients and fractions. In consequence, this method might be useful for adaptive radiotherapy, which focuses on interfractional changes.

II.E.2. Determination of the prior knowledge

In the following, we describe how the necessary prior knowledge can be determined out of different sources of clinical data and imaging modalities for the general MAP estimator in Appendix A. Different imaging techniques acquired prior to treatment might be utilized for target definition (for example, multiple planning CTs, MRI, PET-CT, or 4D-CT). This large amount of available data could be directly utilized for the positioning by MAP estimation. For example, the anatomical plasticity of the body of the patient (i.e., the variability of feature positions relative to the target) can be determined if in each of those volume data sets the positions (x, y, z) of the target and the features are detected manually or with an automatic algorithm.

The expected interfractional relative distances $\{\mu_{X_{F,l}}, \mu_{Y_{F,l}}, \mu_{Z_{F,l}}\}$ can then be estimated by the arithmetical mean of the difference of the position of the target and the feature along the x-, y-, and z-directions of all those imaging modalities, denoted with $\Delta x_{F,l}$, $\Delta y_{F,l}$, and $\Delta z_{F,l}$. By having J imaging modalities (indexed with $j = 1, \ldots, J$), the formulas for the mean are

$$\mu_{X_{F,l}} = \frac{1}{J}\sum_{j=1}^{J} \Delta x_{F,l,j}, \quad \mu_{Y_{F,l}} = \frac{1}{J}\sum_{j=1}^{J} \Delta y_{F,l,j},$$

$$\mu_{Z_{F,l}} = \frac{1}{J}\sum_{j=1}^{J} \Delta z_{F,l,j}. \tag{16}$$

Also, the expected deviations (which corresponds to σ_F for the simplified estimator) and its directionality can be estimated prior by the symmetric sample covariance matrix with

$$(\Sigma_{F,l})_{x,x} = \frac{1}{J-1} \sum_{j=1}^{J} (\Delta x_{F,l,j} - \mu_{X_{F,l}})^2,$$

$$\text{analog for} \quad (\Sigma_{F,l})_{y,y} \quad \text{and} \quad (\Sigma_{F,l})_{z,z},$$

$$(\Sigma_{F,l})_{x,y} = \frac{1}{J-1} \sum_{j=1}^{J} (\Delta x_{F,l,j} - \mu_{X_{F,l}}) \cdot (\Delta y_{F,l,j} - \mu_{Y_{F,l}}),$$

$$\text{analog for} \quad (\Sigma_{F,l})_{x,z} \quad \text{and} \quad (\Sigma_{F,l})_{y,z}. \tag{17}$$

The same method can be utilized to calculate the expected mean position and deviation of the position of the target after the initial setup if several patient setups are performed in the same treatment coordinate system and setup technique (e.g., multiple planning CTs with laser setup). Therefore, one can calculate $\{\mu_{X_T}, \mu_{Y_T}, \mu_{Z_T}\}$ and the sample covariance matrix Σ_T (which corresponds to σ_T for the simplified estimator) in an analogous way.

Beside the images at hand, one could use values of the relevant quantities reported in literature (e.g., Refs. 27 and 28) to get typical accuracies of different setup techniques for several treatment sites. Moreover, values for $\Sigma_{F,l}$ and Σ_T can be determined for specific setup and imaging techniques used at a radiation oncology department based on clinical experience and retrospective studies.

The detection accuracy of the features in the projections depends on the detection algorithm utilized, the image quality of the radiographs, and the feature type and should be validated for the specific application. For example, in an experiment, the true position of an object could be known in the projections and the standard deviation $\sigma_{\Delta U}$ to the actually detected position could be calculated.

III. EXPERIMENT

III.A. Experimental idea

As an application of the MAP-based stochastic formulation of the patient setup problem, we will consider a few potential imaging schemes presented in Fig. 2, starting with a simple [two orthogonal radiographs, in case a] to more complex [cases b–f]. Taking two radiographs separated by a smaller than 90° angle, as in case b, may be beneficial because of clearance (e.g., evaluation of seed positions during brachytherapy), but the problem is that with the decreased angle the triangulation accuracy becomes worse along the midline between the two views. In case c, we will investigate a new scheme with two orthogonal short arcs to evaluate it against the classical orthogonal projections. It is expected that the accuracy of feature localization is increased in 3D when we compare case a to case c because more information is provided. Similar as in case b due to clearance, we want to evaluate a short arc scan in case d (as typical for CB-TS). Eventually, we will compare the results to cases e and f, to investigate if a further gain in accuracy can be achieved by acquiring more projections.
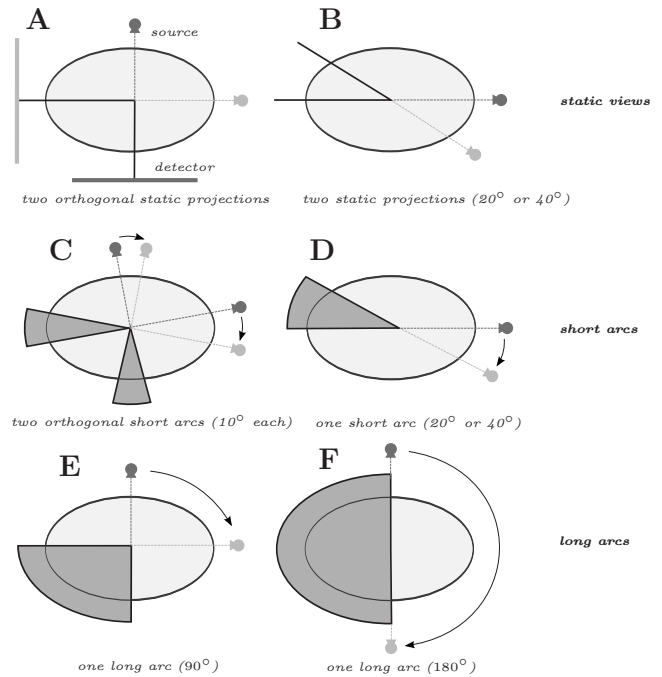

FIG. 2. Different scenarios for the acquisition of CB projections. (a) Two orthogonal projections, (b) two projections with less than 90°, (c) two small orthogonal arcs, (d) one small arc, (e) full 90°, and (f) full 180° arc with a dense set of projections.

The aim of this verification is to assure that the method works in principle and to compare cases a–f of Fig. 2 to each other. Additionally, the rule of thumb of Eq. (11) will be confirmed for real experimental data.

For verification, a deformable phantom is utilized, which we will describe in Sec. III B in detail. The most important information is that in this phantom, six small 1 mm chrome balls are inserted as features. This phantom is then aligned to the in-room laser system for the initial setup by surface marks. Successively, in 20 scans, the position of these chrome balls is shifted relative to each other according to a random table, representing the plasticity of the body. Additionally, in between each scan, the position of the whole phantom is also changed in a translational way. This represents then the phantom deformation with the additional initial setup error.

For each phantom setup, a full set of cone beam projections, as it is necessary for reconstruction, is acquired. From this full data set, projections for shorter arcs or other subsets, as in the cases of Fig. 2, are selected for estimation.

For applying the MAP estimation of Appendix A, we need the position of the features in the projections as our measurements. Therefore, we utilize a simple detection method based on a local threshold method as a feature detection algorithm utilizing standard routines of the MATLAB system. Other algorithms can be used for this purpose, but as we see in Sec. IV B, this might only have limited influence on the overall accuracy.

After the measurements are determined, the MAP estimation is applied to all 20 CB projection data sets. For evaluation, we compare the estimated target position $(\hat{x}_T, \hat{y}_T, \hat{z}_T)$ to
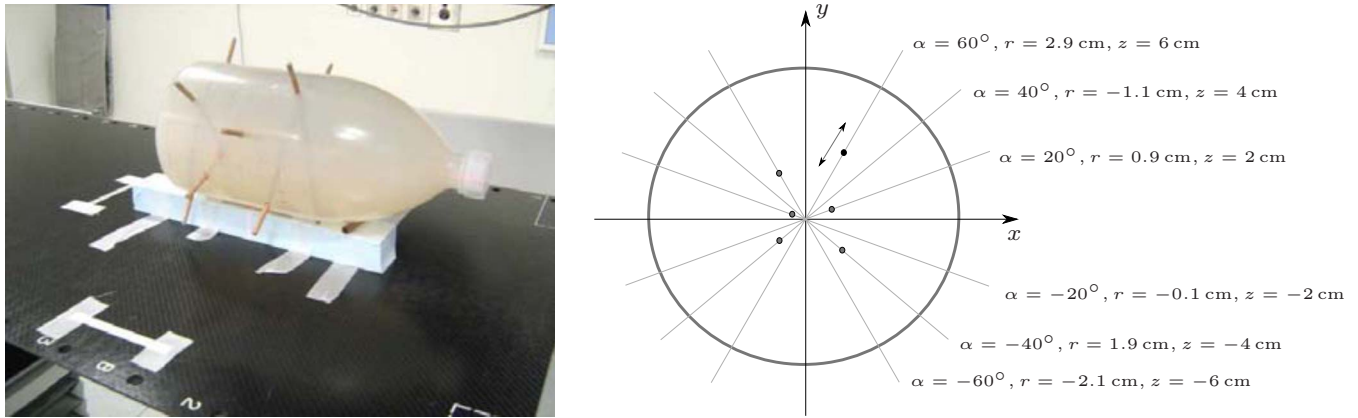
FIG. 3. The phantom for the experimental verification. Left: The phantom utilized for the measurements. Right: The schematical overview for the positions of the 1 mm chrome balls. The positions of the balls are calculated with $(x,y,z)^T = (r \cos \alpha, r \sin \alpha, z)^T$. The balls are shifted along radial r-direction to represent the plasticity of the phantom.

the setup error, which was introduced artificially by moving the couch with the phantom. In consequence, the closer the estimated target coordinates are to the phantom shifts, the more accurate the estimation works.

In the statistical evaluation of the (in our case 20) data sets, we will compare the empirical RMSEs of the cases a–f of Fig. 2. The full evaluation of this data and its comparison to the RMSEs out of the rule of thumb are presented in Sec. IV A.

III.B. Phantom setup

As a deformable phantom, we utilized a cylindrical phantom of a diameter 13.72 cm filled with water-equivalent gelatin (see Fig. 3 left). In the walls of the bottle, bore holes opposite to each other were drilled so that six wooden sticks could be inserted going through the rotational axis of the bottle. In Fig. 3 on the right, the data for these bore holes are illustrated in a schematical overview. In these wooden sticks, 1 mm chrome balls were inserted as features that we want to detect in the projections. Moving the wooden sticks leads to a change of the position of the chrome balls along radial direction. With different relative shifts of these features for every phantom setup, the plasticity of the body in the x-y-plane is simulated in a reproducible fashion.

In each of the 20 setups, the phantom is aligned to the in-room laser and then the sticks are shifted; additionally the position of the treatment couch is shifted in the x-, y-, and z-directions. The shifts are determined by a Gaussian random number table generated by MATLAB, fulfilling the standard deviations of $\sigma_{X_T} = \sigma_{Y_T} = \sigma_{Z_T} = 1.0$ cm and $\sigma_r = 0.5$ cm with zero mean. In the computation of the results, the known shifts of the couch are the gold standard (i.e., the true setup errors) to which the MAP estimated setup errors based on feature detections are compared. In detail, σ_r represents the standard deviation for the radial shifts of the chrome balls, which can be interpreted as σ_F in the rule of thumb in Eq. (11). In the estimation algorithm, we will utilize then in consequence $\sigma_{X_{F,l}} = \sigma_{Y_{F,l}} = \sigma_{Z_{F,l}} = 0.5$ cm. In a clinical application with patient data, these deviation parameters would be deter-

mined out of prior imaging data or literature, as described in Sec. II E 2.

The acquisition of the projections was performed with the XVI System of the Elekta Synergy® linear accelerator. The arc range of the full data set was from $282.0°$ to $81.0°$ with a step size of $-0.51°$ leading to 393 projections (angles referred to the coordinate system in Fig. 6). In the projections, the whole phantom was visible from all angles.

The detection of the features in the projections happened with standard routines of the MATLAB system utilizing a local threshold method with an initial manual identification of the features in the first projection. For this simple phantom, the detection worked very well ($\sigma_{\Delta U} \approx 0.03$ cm at the limit of the pixel width) since the 1 mm chrome balls can be identified in the projections with high accuracy. To simulate a more realistic scenario with less quality of detection, we introduced an additional detection error by adding Gaussian noise with standard deviation of $\sigma_{\Delta U} = 0.25$ cm to both projection coordinates (u, v) (see Fig. 6 for the definition).

The aim of the phantom experiment is to verify that the estimation technique works with clinical data and equipment (XVI, Elekta Synergy®). In this context, the influence of different imaging directions on the estimation accuracy is investigated. Moreover, the rule of thumb of Sec. II D 1 is confirmed as a reasonable measure of certainty. Based on the rule of thumb, we derive clinically interesting factors, such as the necessary number of projections N for different detection error ranges and the necessary number of detectable features L for different plasticities of the body

IV. RESULTS

IV.A. Comparing different CB acquisition schemes

For the evaluation of the experiment, we are comparing the imaging directions of cases a–f of Fig. 2. The results for the empirical $\mathbb{R}$MSEs along each coordinate direction and of the rule of thumb (when reasonable) are presented in Table I. In Fig. 4, an illustration of the detected feature positions for the 20 phantom setups at an exemplary imaging direction is

TABLE I. Statistic evaluation of the accuracy of different imaging schemes of Fig. 2, cases a–f, for 20 setups. Shown are the empirical RMSEs as difference between estimated target position $(\hat{x}_{T,k}, \hat{y}_{T,k}, \hat{z}_{T,k})$ and the target position determined by CB-CT reconstruction $(x_{T,k}, y_{T,k}, z_{T,k})$ along each coordinated for setup $k = 1, \ldots, 20$. They are defined with $\mathrm{RMSE}_x = \sqrt{1/20 \Sigma_{k=1}^{20} (\hat{x}_{T,k} - x_{T,k})^2}$ and analog for RMSE_y and RMSE_z. Additionally, the theoretical value out of the rule of thumb $\mathrm{RMSE}_{\mathrm{ROT}}$ is presented for comparison (with NA due to high asymmetry of imaging directions).

Case	Specification	N	RMSE_x (cm)	RMSE_y (cm)	RMSE_z (cm)	$\mathrm{RMSE}_{\mathrm{ROT}}$ (cm)
A	90° separation angle	2	0.24 ± 0.01	0.14 ± 0.01	0.07 ± 0.01	0.22
B	20° separation angle	2	0.31 ± 0.01	0.14 ± 0.01	0.08 ± 0.01	NA
B	40° separation angle	2	0.24 ± 0.01	0.14 ± 0.01	0.08 ± 0.01	NA
C	$2 \times 10°$ arc range	10 (2×5)	0.19 ± 0.01	0.14 ± 0.01	0.06 ± 0.01	0.20
D	20° arc range	10	0.27 ± 0.01	0.14 ± 0.01	0.07 ± 0.01	NA
D	40° arc range	10	0.20 ± 0.01	0.13 ± 0.01	0.07 ± 0.01	NA
E	90° arc range	89	0.19 ± 0.01	0.13 ± 0.01	0.06 ± 0.01	NA
F	180° arc range	177	0.20 ± 0.01	0.13 ± 0.01	0.06 ± 0.01	0.20

presented. In this figure, detection clusters of the feature positions occur due to the superposition of the initial setup error of each phantom setup, the internal deformations by displacing each marker for every setup, and the artificially introduced detection error of every detection. To gain a deeper insight, the detected feature positions of the first and second setup are presented. By comparing these two data sets, one can observe a general trend of displacement (indi-
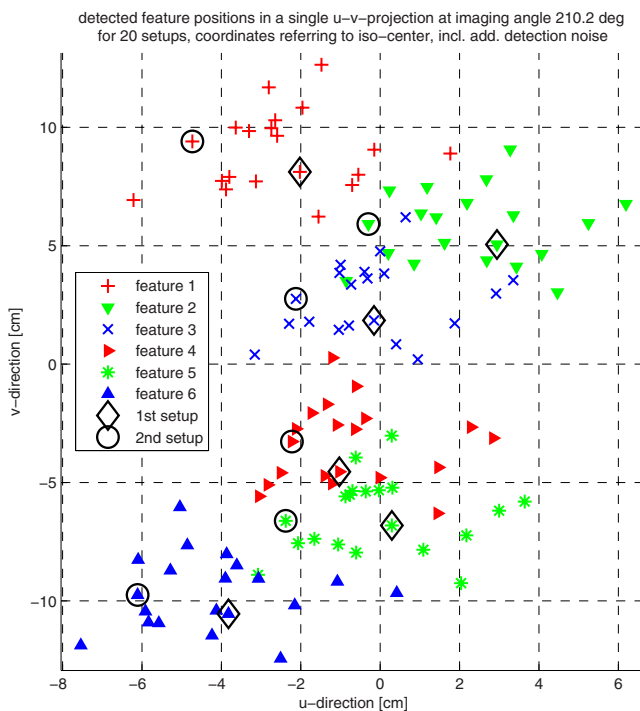

FIG. 4. The detected feature positions in a projection at the imaging angle of 210.2° for the 20 phantom setups. The plot is in the *u-v*-coordinates of the projection coordinate system with the imaging isocenter at its origin, as it is defined in Appendix A (Fig. 6). Each feature is present by a unique symbol, eventually presenting six detection clusters of the 20 phantom setups. Additionally, the 6 feature positions corresponding to the first phantom setup are framed by diamonds and for the second setup by circles. The detected feature positions of one phantom setup for all imaging angles are the input of the MAP estimation.

cating an allover setup error) and on top of this individual displacements (corresponding to the internal deformations and to the additional detection errors).

A strong difference in the estimation accuracy along the *x*-, *y*-, and *z*-directions for all acquisition schemes can be observed. This can be explained by the way the phantom was built. At first, no plasticity deformations of the features along *z*-direction are introduced by the phantom setup and therefore a low estimation error should be expected, as it can be confirmed. Second, the difference between the *x*- and *y*- directions lies in the directionality of the wooden sticks along the features can be shifted in radial direction. As it is shown in Fig. 3 on the right, we had to spare the angular range around the *y*-axis to be able to position the phantom on a couch. Therefore, most deformations happen along the *x*-axis and only a smaller fraction along the *y*-axis, leading to lower estimation accuracy along the *x*-axis than the *y*-axis. In consequence, the rule of thumb (which is evaluated with the parameters $\sigma_T = 1.0$ cm, $\sigma_F = 0.5$ cm, $\sigma_{\Delta U} = 0.25$ cm, and $L = 6$) should be compared to the *x*-direction.

In general, the resulting estimation errors do not vary strongly among cases a–f due to the robustness of the estimator derived, but they give constructive indications which data acquisition scenarios are more favorable.

At first, comparing the orthogonal projections of case a, which is a classical setup technique for rigid body structures, to smaller arcs in between the projections in case b, shows that using just a 20° separation angle leads to a loss in estimation accuracy. On the other hand, using only a 40° separation angle can compensate these inaccuracies due to imaging directions, so that for patient positioning also a smaller arc than 90° can be suggested.

Second, by comparing the orthogonal projections of case a to the two small orthogonal arcs of case c (10° arcs with five projections each), one can see the benefit in using more projections and, in consequence, in eliminating the influence of the detection error.

Third, by following the argumentation of using only short arcs for patient positioning, we are showing in case d a 20°
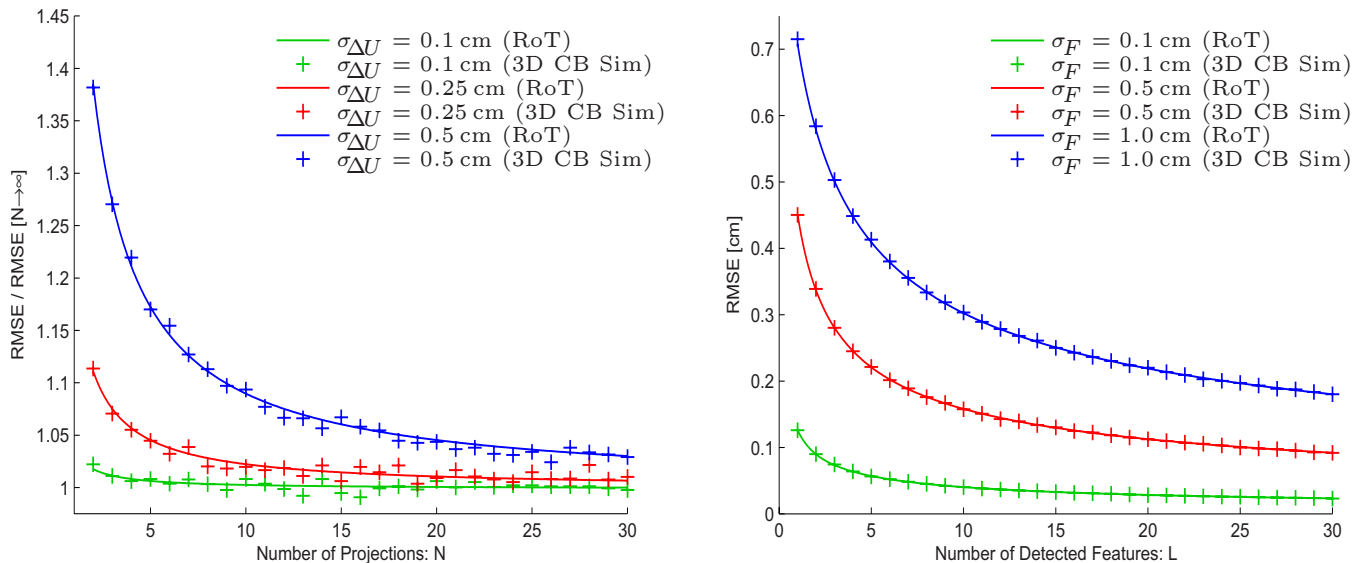
FIG. 5. The results for the RMSE according to the rule of thumb formula (RoT, solid lines) and to 10 000 numerical simulations for 3D cone beam geometry (3D CB Sim, crosses). Left: The estimation error RMSE depending on the number of projections N for different detection errors $\sigma_{\Delta U}$ in reference to the maximal achievable RMSE for $N \rightarrow \infty (=0.20$ cm$)$. The other parameters are $L=6$, $\sigma_F=0.5$ cm, and $\sigma_T=1$ cm. Right: The estimation error RMSE depending on the number of detectable features L for different plasticity parameters of the body σ_F. The other parameters are $N=20$, $\sigma_{\Delta U}=0.25$ cm, and $\sigma_T=1$ cm.

and 40° arc range with ten projections each. Comparing these results to the corresponding case b, an increase in the estimation accuracy can be observed, which corresponds to the elimination of the influence of the detection error, analog to above. Moreover, with a 40° arc range, already the accuracy of the two orthogonal short arcs of case c can be achieved, leading to equivalent setups.

Fourthly, by comparing cases C and D 40° utilizing ten projections, with a high number of projections for a full 90° arc in case E with 89 projections and a full 180° arc in case F with 177 projections, no essential further gain in accuracy can be observed. Therefore, utilizing case C or D 40° might lead to a strong reduction of projections one needs to acquire for patient positioning compared to a full cone beam reconstruction with 393 projections in the CB-CT System XVI and, in consequence, a reduction of dose and time. The only assumption for estimation is the possibility of detecting features in the radiographs at all, and even high detection errors can be compensated by using more projections.

IV.B. Evaluation of the rule of thumb

For practical purposes, we are evaluating the behavior of the rule of thumb formula for different ranges for the parameters of the prior knowledge. We will show, therefore, the results of the rule of thumb as solid lines and, additionally, the empirical RMSEs for 10 000 numerical simulations (for each parameter set) of a 3D cone beam geometry as crosses. These empirical RMSEs based on simulations refer to the translational error in the *x*- and *y*-directions, utilizing the same parameters and distribution of imaging angles as for the rule of thumb. In consequence, we verify that the change from 2D parallel beam to 3D cone beam geometry does not change the derived rule of thumb.

At first, we want to illustrate how many projections are necessary to make the estimation itself sufficiently independent of the feature detection algorithm in the projections. Therefore, we essentially have to investigate the influence of the number of projections N on the RMSE. In Fig. 5 (left), different scenarios are presented, leading to the conclusion that using about $N=20$ projections reduces the residual estimation error to about 5% for an even poorly working feature detection algorithm.

Second, the behavior of the estimation accuracy for variations of the number of detectable features in the projections will be illustrated. In Fig. 5 (right), a strong dependence of the estimation accuracy on the plasticity of the body (represented by σ_F) and the number of detectable features L can be observed. This means that more flexible geometries (such as $\sigma_F=1.0$ cm) do need a high L (such as 30 features for a RMSE of 0.2 cm) and almost rigid geometries (such as $\sigma_F=0.1$ cm) only need $L \approx 1-5$ features for good setup accuracy (RMSE of 0.1 cm), which is in agreement to the high accuracy in the work of Gall[5] for cranial structures.

V. DISCUSSION

Patient positioning techniques are an important part of fractionated radiotherapy in the framework of IGRT and the focus of current research. Different approaches are developed for this task, such as the broadly applied reconstruction based techniques using cone beam CT,[2–4] optical cameras,[29–31] ultrasound methods,[32] or positioning techniques based on just a few radiographs.[5–8] Typically, these methods provide submillimeter accuracy for rigid bodies and good contrast but lead to problems in interpretation for determining the rigid transformation parameters when interfractional deformations occurred inside the patient. Addition-

ally, some of these methods have inherent problems, such as metal and motion artifacts in CB-CT or systematically low image quality in ultrasound and projection based methods.

In this work, we introduce a patient setup technique based on prior knowledge, which utilizes feature positions detected in CB projections (recorded by a CB-CT system), without the need of 3D CB-CT reconstruction. A novel part of this formulation is the utilization of prior knowledge in a stochastic-geometrical model about the initial patient setup accuracy (e.g., by an in-room laser system and skin markers), the relative shifts about the features inside the body (essentially the plasticity of the body), and a model for the detection error of the features in the projections. Therefore, this work can also be regarded as a general introduction to a *stochastic-geometrical framework* for patient positioning in which further developments might be achieved.

In this context, the technique presented can also be interpreted as an extension of classical feature based patient setup by orthogonal projections for rigid body structures (such as in Ref. 5) to deformable bodies, incorporating prior knowledge.

The estimation of the translational and rotational interfractional setup errors is based on the maximum *a posteriori* argumentation in the Bayesian framework, which is typically utilized to introduce prior knowledge into estimators. The estimator for a 2D parallel beam is derived under simplifying assumptions for which it can be shown that the Cramér–Rao bound is fulfilled (and that the estimator is therefore efficient). The result for the 3D cone beam model for the clinical application is presented in Appendix A.

One of the major results is a rule of thumb that relates the parameters of prior knowledge (such as the initial setup error σ_T, plasticity parameter of the body σ_F, etc.) and data acquisition (such as number of projections N, feature detection accuracy $\sigma_{\Delta U}$, etc.) to the expected resulting estimation error. One main consequence is that for a finite number of features, an inherent limit of accuracy can be achieved depending on plasticity and initial setup accuracy, which cannot be increased by acquiring more projections. Another point is that detecting as many features as possible (depending on the plasticity of the body) is a central challenge for increasing estimation accuracy. It is also shown theoretically that utilizing prior knowledge about the initial setup error will always benefit for the estimation.

Besides the theoretical derivations, the clinical application of the estimator and the determination of the prior knowledge out of multimodality imaging techniques are also discussed in Sec. II E 2. It is pointed out that utilizing a model for prior knowledge also allows individualizing the patient setup by updating this prior knowledge successively and essentially separating long-term systematic errors from short-term random errors during the course of fractionation. This might be useful for applications in the adaptive radiotherapy.

In Sec. III A, an experimental verification of the estimation is presented. It utilizes a deformable phantom which is filled with gelatin and in which six chrome balls with 1 mm

diameter as features can be moved in radial directions relative to each other, simulating the plasticity of a body in a reproducible fashion. In total, 20 setups with different relative shifts of the features to each other and a global initial setup error by shifting the whole phantom along the x-, y-, and z-directions are performed. In every setup, cone beam projections are acquired and the features are determined in these projections with a semiautomatic detection algorithm. After that, the corresponding setup errors are estimated with the MAP estimator for different imaging directions starting from orthogonal projections to a full set of cone beam projections in a 180° arc.

The evaluation of the phantom experiment suggests that two orthogonal small arcs (10° each) or a small arc with 40° lead to the same accuracy of about 0.2 cm of the patient setup as acquisition schemes for the full data set of projections. We want to point out that for setup techniques involving, first, CB-CT reconstruction and then reconstructed image registration with the planning CT, accuracies of down to about 0.1 cm are reported.[2] Current studies suggest that the overall positioning accuracies of CB-TS algorithms (without internal deformations) are about 0.1–0.2 cm for small arcs.[14,33,34] Therefore, the experimental accuracy of our estimation method can be roughly expected to be in the range of CB-TS based registration. However, a detailed comparison

TABLE II. Nomenclature of extended and newly introduced estimates and parameters of the prior knowledge. For clarification of the measured values, see Fig. 6.

Estimates	Description
$\hat{x}_T, \hat{y}_T, \hat{z}_T$	Estimated position of the target in x-, y-, and z- directions
$\hat{\varphi}_x, \hat{\varphi}_y, \hat{\varphi}_z$	Estimated rotation of the target around x-, y-, and z- axes
Prior knowledge	**Description**
$\mu_{X_T}, \mu_{Y_T}, \mu_{Z_T}$	Expected position of the target (typically from planning CT)
Σ_T	Covariance matrix of the deviations of the target position (depending on initial setup technique)
$\mu_{\Phi_x}, \mu_{\Phi_y}, \mu_{\Phi_z}$ $\sigma_{\Phi_x}, \sigma_{\Phi_y}, \sigma_{\Phi_z}$	Expected rotation of the target (typically 0°) Standard deviations of the target rotation (depending on initial setup technique)
$\mu_{X_{F,l}}, \mu_{Y_{F,l}}, \mu_{Z_{F,l}}$	Expected anatomical distances between target and feature l (determined by multiple planning CT or other imaging modalities)
$\Sigma_{F,l}$	Covariance matrix of the deviations of the anatomical distances between the target and feature l (based on multiple imaging and/or anatomical reasoning)
Measurements	**Description**
$u_{l,n}, v_{l,n}$	Position of feature l in projection n in u-direction (parallel to the x-y-plane) and v-direction (parallel to the z-axis)

of all of these methods, especially for deformable phantoms or clinical cases, is recommended for the future research.

VI. CONCLUSION

In this work, a feature based patient positioning method is presented utilizing cone beam projections (or radiographs). The stochastic formulation of this problem includes three sources of uncertainties, such as the initial patient setup error after laser-based adjustment of the patient position, the deformability of the patient anatomy compared to the reference data, and measurement noise for the detection of the features in the radiographs.

The principles of derivation of the maximum *a posteriori* approach are presented for a 2D parallel beam model and the final estimator for the 3D cone beam model is given in Appendix A, which can be applied in clinical practice. Additionally, a rule of thumb for the estimation accuracy is derived, which provides further insight in the behavior of the estimator depending on deformation, detection, and CB data acquisition parameters.

The approach presented can also be interpreted as an extension of previous developed deterministic methods, which relied only on pure geometrical considerations. In consequence, the methodology can also be understood as a stochastic-geometrical formulation of the inverse problem of patient positioning and an introduction to a theoretical framework for future applications.

APPENDIX A: THE 3D CONE BEAM ESTIMATOR FOR TARGET POSITION AND ROTATIONS

For 3D cone beam geometry with the estimation of target translation and rotation and lifting the assumptions of Sec. II, new parameters have to be introduced as presented in Table II (please note that all quantities are in reference to the treatment coordinate system and Gaussians are utilized as probability density functions to describe different sources of prior knowledge).

In this approach, an optimal rotation and shift is estimated although nonlinear deformations may have occurred inside the body by utilizing probability distributions. The final transformation of the patient with the treatment couch [from coordinates $(x,y,z)^T$ to $(x',y',z')^T$] is then described with

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \Omega_z(\hat{\varphi}_z)\Omega_y(\hat{\varphi}_y)\Omega_x(\hat{\varphi}_x)\begin{pmatrix} x - \hat{x}_T \\ y - \hat{y}_T \\ z - \hat{z}_T \end{pmatrix} + \begin{pmatrix} \hat{x}_T \\ \hat{y}_T \\ \hat{z}_T \end{pmatrix}, \qquad (A1)$$

with the rotation matrices Ω_x, Ω_y, and Ω_z of the Euler angles around the x-, y-, and z-axes (e.g., see Ref. 5) and the estimated values $\{\hat{x}_T, \hat{y}_T, \hat{z}_T, \hat{\varphi}_x, \hat{\varphi}_y, \hat{\varphi}_z\}$.

With these definitions, the optimal MAP estimates for the 3D cone beam geometry are then determined by Eq. (A2). The objective function in this expression is almost quadratic in the translational parameters and therefore can be solved numerically efficiently with iterative optimization algorithms.

$$\arg\min_{(x_T,y_T,z_T,\varphi_x,\varphi_y,\varphi_z)} \begin{pmatrix} x_T \\ y_T \\ z_T \end{pmatrix}^T \Sigma_T^{-1} \begin{pmatrix} x_T \\ y_T \\ z_T \end{pmatrix} - 2\begin{pmatrix} \mu_{X_T} \\ \mu_{Y_T} \\ \mu_{Z_T} \end{pmatrix}^T \Sigma_T^{-1} \begin{pmatrix} x_T \\ y_T \\ z_T \end{pmatrix} + \left(\frac{\varphi_x - \mu_{\Phi_x}}{\sigma_{\Phi_x}}\right)^2 + \left(\frac{\varphi_y - \mu_{\Phi_y}}{\sigma_{\Phi_y}}\right)^2 + \left(\frac{\varphi_z - \mu_{\Phi_z}}{\sigma_{\Phi_z}}\right)^2$$

$$+ \sum_{l=1}^{L} c - \left[b + \begin{pmatrix} \mu_{X_{F,l}} \\ \mu_{Y_{F,l}} \\ \mu_{Z_{F,l}} \end{pmatrix}^T \Sigma_{F,l}^{-1}\right] \cdot (M^{-1} + \Sigma_{F,l}) \cdot \left[b + \begin{pmatrix} \mu_{X_{F,l}} \\ \mu_{Y_{F,l}} \\ \mu_{Z_{F,l}} \end{pmatrix}^T \Sigma_{F,l}^{-1}\right]^T, \qquad (A2)$$

with the arrays

$$M = \begin{pmatrix} \displaystyle\sum_{i=1}^{N} \frac{\omega_{u1}^2 + \omega_{v1}^2}{\sigma_{\Delta U}^2} + \frac{1}{\sigma_{X_{F,l}}^2} & \displaystyle\sum_{i=1}^{N} \frac{\omega_{u1}\omega_{u2} + \omega_{v1}\omega_{v2}}{\sigma_{\Delta U}^2} & \displaystyle\sum_{i=1}^{N} \frac{\omega_{u1}\omega_{u3} + \omega_{v1}\omega_{v3}}{\sigma_{\Delta U}^2} \\ \displaystyle\sum_{i=1}^{N} \frac{\omega_{u1}\omega_{u2} + \omega_{v1}\omega_{v2}}{\sigma_{\Delta U}^2} & \displaystyle\sum_{i=1}^{N} \frac{\omega_{u2}^2 + \omega_{v2}^2}{\sigma_{\Delta U}^2} + \frac{1}{\sigma_{Y_{F,l}}^2} & \displaystyle\sum_{i=1}^{N} \frac{\omega_{u2}\omega_{u3} + \omega_{v2}\omega_{v3}}{\sigma_{\Delta U}^2} \\ \displaystyle\sum_{i=1}^{N} \frac{\omega_{u1}\omega_{u3} + \omega_{v1}\omega_{v3}}{\sigma_{\Delta U}^2} & \displaystyle\sum_{i=1}^{N} \frac{\omega_{u2}\omega_{u3} + \omega_{v2}\omega_{v3}}{\sigma_{\Delta U}^2} & \displaystyle\sum_{i=1}^{N} \frac{\omega_{u3}^2 + \omega_{v3}^2}{\sigma_{\Delta U}^2} + \frac{1}{\sigma_{Z_{F,l}}^2} \end{pmatrix}, \qquad (A3)$$

$$b = \left(-\sum_{i=1}^{N} \frac{\alpha \omega_{u1} + \beta \omega_{v1}}{\sigma_{\Delta U}^2} + \frac{\mu_{X_{F,l}}}{\sigma_{X_{F,l}}^2}, -\sum_{i=1}^{N} \frac{\alpha \omega_{u2} + \beta \omega_{v2}}{\sigma_{\Delta U}^2} \right.$$

$$\left. + \frac{\mu_{Y_{F,l}}}{\sigma_{Y_{F,l}}^2}, -\sum_{i=1}^{N} \frac{\alpha \omega_{u3} + \beta \omega_{v3}}{\sigma_{\Delta U}^2} + \frac{\mu_{Z_{F,l}}}{\sigma_{Z_{F,l}}^2} \right), \tag{A4}$$

and

$$c = \sum_{i=1}^{N} \frac{\alpha^2 + \beta^2}{\sigma_{\Delta U}^2}. \tag{A5}$$

The appearing auxiliary expressions are

$$\omega_u := \begin{pmatrix} \sin\theta_n + \dfrac{u_{l,n}\cos\theta_n}{\text{SAD}+\text{ADD}} \\ -\cos\theta_n + \dfrac{u_{l,n}\sin\theta_n}{\text{SAD}+\text{ADD}} \\ 0 \end{pmatrix}^T \cdot \Omega_z(\varphi_z)\Omega_y(\varphi_y)\Omega_x(\varphi_x), \tag{A6}$$

$$\omega_v := \begin{pmatrix} \dfrac{v_{l,n}\cos\theta_n}{\text{SAD}+\text{ADD}} \\ \dfrac{v_{l,n}\sin\theta_n}{\text{SAD}+\text{ADD}} \\ -1 \end{pmatrix}^T \cdot \Omega_z(\varphi_z)\Omega_y(\varphi_y)\Omega_x(\varphi_x), \tag{A7}$$

$$\alpha := \left(\sin\theta_n + \frac{u_{l,n}\cos\theta_n}{\text{SAD}+\text{ADD}} \right) \cdot x_T$$

$$+ \left(-\cos\theta_n + \frac{u_{l,n}\sin\theta_n}{\text{SAD}+\text{ADD}} \right) \cdot y_T + \frac{u_{l,n}\text{SAD}}{\text{SAD}+\text{ADD}}, \tag{A8}$$
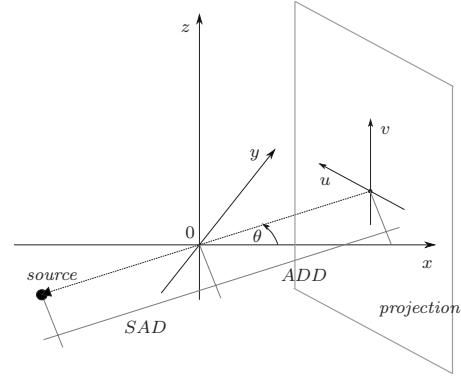
and

FIG. 6. Visualization of the coordinate systems used for the derivation of the MAP estimator.

$$\beta := \frac{v_{l,n}}{\text{SAD}+\text{ADD}}(\cos\theta_n \cdot x_T + \sin\theta_n \cdot y_T + \text{SAD}) - z_T. \tag{A9}$$

In these formulas SAD represents the *source-axis-distance* and ADD the *axis-detector-distance* (see Fig. 6).

APPENDIX B: TECHNICAL INTERPRETATION OF INTERSECTION OF EVENTS

By assuming that the geometrical shifts just *in between different features* do not contain information about the position of the treatment target, expression (7) is equivalent to

$$\arg\max_{(x_T, y_T)} \prod_{l=1}^{L} P\left(\bigcap_{n=1}^{N} U_{l,n} = u_{l,n} \middle| X_T = x_T \cap Y_T = y_T\right)$$

$$\cdot P(X_T = x_T \cap Y_T = y_T). \tag{B1}$$

Essentially, we have to insert the definition and apply the law of total probability to the likelihood expression

$$P\left(\bigcap_{n=1}^{N} U_{l,n} = u_{l,n} \middle| X_T = x_T \cap Y_T = y_T\right) \tag{B2}$$

$$= P\left(\bigcap_{n=1}^{N} -(x_T + X_{F,l}) \cdot \sin\theta_n + (y_T + Y_{F,l}) \cdot \cos\theta_n + \Delta U_{l,n} = u_{l,n}\right) \tag{B3}$$

$$= \int\int_{\mathbb{R}^2} P\left(\bigcap_{n=1}^{N} \Delta U_{l,n} = u_{l,n} + (x_T + s) \cdot \sin\theta_n - (y_T + t) \cdot \cos\theta_n\right) P(X_{F,l} = s \cap Y_{F,l} = t)\,\mathrm{d}(s,t) \tag{B4}$$

to being able to interpret all occurring probabilities. After inserting the Gaussian probability distributions, this integral can be calculated analytically, leading to

$$= \text{const} \cdot \exp\left[\sum_{n=1}^{N} \frac{u_{l,n}}{\sigma_{\Delta U}^2} + \frac{(x_T + \mu_{X_{F,l}})^2 + (y_T + \mu_{Y_{F,l}})^2}{\sigma_F^2} \right. \tag{B5}$$

$$\left. - \frac{2\sigma_F^2}{\sigma_{\Delta U}^2(N\sigma_F^2 + 2\sigma_{\Delta U}^2)}\left[\left(\sum_{n=1}^{N} -u_{l,n}\sin\theta_n + \frac{x_T + \mu_{X_{F,l}}}{\sigma_F^2}\right)^2 + \left(\sum_{n=1}^{N} -u_{l,n}\cos\theta_n + \frac{y_T + \mu_{Y_{F,l}}}{\sigma_F^2}\right)^2 \right] \right], \tag{B6}$$

with a positive constant const.

a)Electronic mail: wolfgang@hoegele.de

[1] D. Verellen, M. De Ridder, N. Linthout, K. Tournel, G. Soete, and G. Storme, "Innovations in image-guided radiotherapy," Nat. Rev. Cancer **7**(12), 949–960 (2007).

[2] M. Guckenberger, J. Meyer, D. Vordermark, K. Baier, J. Wilbert, and M. Flentje, "Magnitude and clinical relevance of translational and rotational patient setup errors: A cone-beam ct study," Int. J. Radiat. Oncol., Biol., Phys. **65**(3), 934–942 (2006).

[3] U. Oelfke, T. Tücking, S. Nill, A. Seeber, B. Hesse, P. Huber, and C. Thilmann, "Linac-integrated kV-cone beam CT: Technical features and first applications," Med. Dosim. **31**(1), 62–70 (2006).

[4] D. Létourneau, J. W. Wong, M. Oldham, M. Gulam, L. Watt, D. A. Jaffray, J. H. Siewerdsen, and A. A. Martinez, "Cone-beam-CT guided radiation therapy: Technical implementation," Radiother. Oncol. **75**(3), 279–286 (2005).

[5] K. P. Gall, L. J. Verhey, and M. Wagner, "Computer-assisted positioning of radiotherapy patients using implanted radiopaque fiducials," Med. Phys. **20**(4), 1153–1159 (1993).

[6] G. C. Sharp, S. Kollipara, T. Madden, S. B. Jiang, and S. J. Rosenthal, "Anatomic feature-based registration for patient set-up in head and neck cancer radiotherapy," Phys. Med. Biol. **50**(19), 4667–4679 (2005).

[7] D. B. Russakoff, T. Rohlfing, K. Mori, D. Rueckert, A. Ho, J. R. Adler, and C. R. Maurer, "Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration," IEEE Trans. Med. Imaging **24**(11), 1441–1454 (2005).

[8] J. Wu, M. Kim, J. Peters, H. Chung, and S. S. Samant, "Evaluation of similarity measures for use in the intensity-based rigid 2D-3D registration for patient positioning in radiotherapy," Med. Phys. **36**(12), 5391–5403 (2009).

[9] J. B. Maintz and M. A. Viergever, "A survey of medical image registration," Med. Image Anal. **2**(1), 1–36 (1998).

[10] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," Phys. Med. Biol. **46**(3), R1–R45 (2001).

[11] B. Zitova and J. Flusser, "Image registration methods: A survey," Image Vis. Comput. **21**, 977–1000 (2003).

[12] L. G. Brown, "A survey of image registration techniques," ACM Comput. Surv. **24**, 325–376 (1992).

[13] *Imaging Systems for Medical Diagnostics: Fundamentals, Technical Solutions and Applications for Systems Applying Ionizing Radiation, Nuclear Magnetic Resonance and Ultrasound*, edited by A. Oppelt (Wiley-VCH, New York, 2005).

[14] A. Buehler, S.-K. Ng, Y. Lyatskaya, D. Stsepankou, J. Hesser, and P. Zygmanski, "Evaluation of clip localization for different kilovoltage imaging modalities as applied to partial breast irradiation setup," Med. Phys. **36**(3), 821–834 (2009).

[15] B. Winey, P. Zygmanski, and Y. Lyatskaya, "Evaluation of radiation dose delivered by cone beam CT and tomosynthesis employed for setup of external breast irradiation," Med. Phys. **36**(1), 164–173 (2009).

[16] D. J. Godfrey, F.-F. Yin, M. Oldham, S. Yoo, and C. Willett, "Digital tomosynthesis with an on-board kilovoltage imaging device," Int. J. Radiat. Oncol., Biol., Phys. **65**(1), 8–15 (2006).

[17] J. Ashburner, P. Neelin, D. L. Collins, A. Evans, and K. Friston, "Incorporating prior knowledge into image registration," Neuroimage **6**(4), 344–352 (1997).

[18] C. E. Rasmussen and C. K. I. William, *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, 2006).

[19] J.-B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," Med. Phys. **34**(11), 4526–4544 (2007).

[20] H. Zaidi, "Correction for image degrading factors is essential for accurate quantification of brain function using PET," Med. Phys. **31**(3), 423–425 (2004).

[21] A. R. De Pierro, "Fast EM-like methods for maximum 'a posteriori' estimates in emission tomography," IEEE Trans. Med. Imaging **20**, 280–288 (2001).

[22] H. Keller, W. Tomé, M. A. Ritter, and T. R. Mackie, "Design of adaptive treatment margins for non-negligible measurement uncertainty: Application to ultrasound-guided prostate radiation therapy," Phys. Med. Biol. **49**(1), 69–86 (2004).

[23] X. Li, L. Li, H. Lu, and Z. Liang, "Partial volume segmentation of brain magnetic resonance images based on maximum a posteriori probability," Med. Phys. **32**(7), 2337–2345 (2005).

[24] V. Rudat, P. Schraube, D. Oetzel, D. Zierhut, M. Flentje, and M. Wannenmacher, "Combined error of patient positioning variability and prostate motion uncertainty in 3D conformal radiotherapy of localized prostate cancer," Int. J. Radiat. Oncol., Biol., Phys. **35**(5), 1027–1034 (1996).

[25] Y. C. Eldar, "Rethinking biased estimation: Improving maximum likelihood and the Cramer–Rao bound," Foundations and Trends in Signal Processing **1**, 305–449 (2008).

[26] T. Bortfeld, M. van Herk, and S. B. Jiang, "When should systematic patient positioning errors in radiotherapy be corrected?," Phys. Med. Biol. **47**(23), N297–N302 (2002).

[27] C. W. Hurkmans, P. Remeijer, J. V. Lebesque, and B. J. Mijnheer, "Set-up verification using portal imaging; review of current clinical practice," Radiother. Oncol. **58**(2), 105–120 (2001).

[28] M. van Herk, "Errors and margins in radiotherapy," Semin. Radiat. Oncol. **14**(1), 52–64 (2004).

[29] A. Brahme, P. Nyman, and B. Skatt, "4D laser camera for accurate patient positioning, collision avoidance, image fusion and adaptive approaches during diagnostic and therapeutic procedures," Med. Phys. **35**(5), 1670–1681 (2008).

[30] M. Krengli, S. Gaiano, E. Mones, A. Ballarè, D. Beldì, C. Bol-chini, and G. Loi, "Reproducibility of patient setup by surface image registration system in conformal radiotherapy of prostate cancer," Radiat. Oncol. **4**, 9–9 (2009).

[31] J. Talbot, J. Meyer, R. Watts, and R. Grasset, "A method for patient set-up guidance in radiotherapy using augmented reality," Australas. Phys. Eng. Sci. Med. **32**(4), 203–211 (2009).

[32] M. Wang, R. Rohling, C. Duzenli, B. Clark, and N. Archip, "Evaluation of targeting errors in ultrasound-assisted radiotherapy," Ultrasound Med. Biol. **34**(12), 1944–1956 (2008).

[33] J. Zhang, Q. J. Wu, D. J. Godfrey, T. Fatunase, L. B. Marks, and F.-F. Yin, "Comparing digital tomosynthesis to cone-beam CT for position verification in patients undergoing partial breast irradiation," Int. J. Radiat. Oncol., Biol., Phys. **73**(3), 952–957 (2009).

[34] B. A. Winey, P. Zygmanski, R. A. Cormack, and Y. Lyatskaya, "Balancing dose and image registration accuracy for cone beam tomosynthesis (CBTS) for breast patient setup," Med. Phys. **37**(8), 4414–4423 (2010).