A spot-specific range uncertainty framework for robust optimization of proton therapy treatments

Marie Cohilis[1], Edmond Sterpin[1,2,3], Gregory Buti[1], Chih-Wei Chang[4], Liyong Lin[4], John A. Lee[1], Kevin Souris[1]

[1] UCLouvain, Institute of Experimental and Clinical Research, MIRO lab, Brussels, Belgium

[2] KU Leuven, Department of Oncology, laboratory of Experimental Radiotherapy, Leuven, Belgium

[3] Particle Therapy Interuniversity Center Leuven – PARTICLE, Leuven, Belgium

[4] Department of Radiation Oncology, Emory University, Atlanta, GA, USA

Corresponding author: Edmond Sterpin

edmond.sterpin@uclouvain.be

Avenue Hippocrate 54, Bte B1.54.07, 1200 Brussels, Belgium

Running title: Spot-specific range uncertainties for PT

ABSTRACT

Purpose: An accurate estimation of range uncertainties is essential to exploit the potential of proton therapy. According to Paganetti's study, a value of 2.4% (1.5 standard deviation) is currently recommended for planning robust treatments with Monte Carlo dose engines. This number is based on a dominant contribution from the mean excitation energy of tissues. However, it was recently shown that expressing tissues as a mixture of water and "dry" material in the CT calibration process allowed for a significant reduction of this uncertainty. We thus propose an adapted framework for pencil beam scanning robust optimization. First, we move towards a spot-specific range uncertainty (SSRU) determination. Second, we use the water-based formalism to reduce range uncertainties and, potentially, to spare better the organs at risk.

Methods: The stoichiometric calibration was adapted to provide a molecular decomposition (including water) of each voxel of the CT. The SSRU calculation was implemented in MCsquare, a fast Monte Carlo dose engine dedicated to proton therapy. For each spot, a ray-tracing method was used to propagate molecular I-values uncertainties and obtain the corresponding effective range uncertainty. These were then combined with other sources of range uncertainties, according to Paganetti's study of 2012. The method was then assessed on three head-and-neck patients. Two plans were optimized for each patient: the first one with the classical 2.4% flat range uncertainty, the second one with the variable range uncertainty. Both plans were then compared in terms of target coverage and OAR mean dose reduction. Robustness evaluations were also performed, using the SSRU for both plans in order to simulate errors as realistically as possible.

Results: For patient 1, it was found that the median spot-specific range uncertainty was 1.04% (1.5 standard deviation), yielding therefore a very large reduction from the 2.4% flat range uncertainty. All three SSRU plans were found to have a very good robustness level at a 90% confidence interval while sparing OAR better than the classical plan. For instance, in nominal cases, average reductions in the mean dose of 15.7%, 8.4%, and 13.2% were observed in the left parotid, right parotid, and pharyngeal constrictor muscle respectively. As expected, the classical plans showed a higher but unnecessary level of robustness.

Conclusions: Promising results of the SSRU framework were observed on three head-and-neck cases, and more patients should now be considered. The method could also benefit to other tu-

mor sites and, in the long run, the variable part of the range uncertainty could be generalized to other sources of uncertainty in order to move towards more and more patient-specific treatments.

I. INTRODUCTION

Nowadays, proton therapy is widely used for its ability to spare healthy tissues[1,2]. This comes from a characteristic depth-dose deposition profile of protons, ending with a Bragg peak, which is defined by a localized high dose region and a steep distal fall-off at a depth depending on the initial energy of the protons. In particular, the pencil beam scanning technique can produce conformal dose distributions by delivering the dose layer-by-layer, using narrow mono-energetic beams[3,4].

Nevertheless, the position of the Bragg peak also depends on the type of tissues traversed along the beam path, which makes proton therapy very sensitive to treatment uncertainties[5,6]. Therefore, robust optimization is often considered in treatment practice in order to ensure adequate delivered dose distributions even if those uncertainties occur[7,8]. Typically, setup and range errors are considered in current commercial solutions. For moving targets, 4D optimization may also be available[9–11].

Typical commercial implementations of robust optimization follow a worst-case approach[12]: at each iteration of the optimization algorithm, multiple scenarios of uncertainties are simulated and the focus is set on the one leading to the worst value of the objective function. This results in a treatment plan which is robust to the treatment uncertainties that are assumed a priori. Monte Carlo dose engines are now commonly used in treatment practice for their excellent accuracy, with reasonable computation time due to the recent developments[13–16].

Robust optimization, however, inevitably increases the volume of healthy tissue that is irradiated at therapeutic dose as a consequence of the integration of the uncertainties in the planning process. It is thus essential to estimate uncertainties accurately in order to reduce as much as possible a potential toxicity for surrounding organs at risk.

Range uncertainties, in particular, have various causes, many of them related to dose calculation[17,18]. A necessary step in order to compute the dose is to convert the CT patient image into a map of relative stopping powers (SPR). To this end, a scanner-specific calibration curve is typically built to enable a conversion from CT numbers to mass density and elemental

composition. From there, electronic density and mean excitation energy (or I-value), necessary for SPR computation, can be deduced[19]. Electronic densities are computed from mass densities and elemental compositions, while mean excitation energies depend on the elemental composition and tabulated elemental I-values. Uncertainties on values derived from the CT and on elemental I-values propagate towards the SPR (or range) uncertainty. The latter remains also true, although implicit, for a direct calibration of CT numbers into SPRs.

A few studies address the quantification of range uncertainties in the literature. Yang et al., in 2012, decompose the range uncertainty in various contributions and propose a total value between 2.5% and 2.9% (90% confidence level) depending on the tumor site[18]. More recently, a few studies also proposed reduced values for the specific case of dual-energy CT[20,21]. Today, however, many studies and clinics rely on the uncertainty quantification reported in Paganetti's article of 2012[17]. In this paper, the largest contribution to the uncertainty was reported to originate from the image conversion. Uncertainties due to tabulated I-values were estimated to be around 1.5%, which is an important part of the total 2.4% reported uncertainty (in the case of Monte Carlo algorithms and considering a confidence interval of 1.5 standard deviation). However, the study did not consider the correlation between I-values of tissues and water to obtain such a value. In 2017, De Smet et al. introduced an explicit way to take this correlation into account and showed that range uncertainties due to I-values were significantly smaller than 1.5%, especially for the tissues made of a large amount of water[22]. Nevertheless, the variation on the uncertainty was extremely dependent on the type of tissue, making a unique upper bound of the error difficult to determine.

To address this particular issue, we now propose to move towards a patient-specific – and even spot-specific – range uncertainty. The aim of this work is to demonstrate its feasibility. We apply the concept to compute range uncertainties following from I-value uncertainties (Σ_I) for each spot of the plan, according to the type of tissues crossed by protons. As this framework reduces the range uncertainty, the impact on the dose distributions is also investigated through a planning study on three head and neck patients.

II. METHODS

The proposed methodology differs from the regular one by requiring a preliminary spot-specific range uncertainty quantification, which occurs in the Monte Carlo dose engine itself.

Figure 1 depicts a summary of the workflows we will compare, i.e., the flat range uncertainty (FRU) and the spot-specific range uncertainty (SSRU) methodologies. In what follows, we describe in detail each part of the workflow. In a first time, section II A provides the necessary background on the water-based formalism previously introduced in the literature[22]. Secondly, the optimization workflow used in this study for both the FRU and the SSRU methodologies is described (section II B). Then, we focus in section II C on the derivation and the implementation of the spot-specific range uncertainties, which are the core of this study. Finally, section II D describes how the proposed SSRU methodology was assessed and compared to the classical FRU framework.

A. Background

To optimize a treatment plan, the CT image of the patient must be first converted into a SPR map that can be used by the dose engine. The SPR of a tissue can be computed with

$$\mathrm{SPR} = \rho_e \frac{\ln\left(\frac{2m_e c^2\beta^2}{(1-\beta^2)\mathrm{e}^{\beta^2}}\right) - \ln I_\mathrm{t}}{\ln\left(\frac{2m_e c^2\beta^2}{(1-\beta^2)\mathrm{e}^{\beta^2}}\right) - \ln I_\mathrm{w}} \quad , \tag{1}$$

where ρ_e is its relative electron density, I_t is its mean excitation energy, I_w the mean excitation of water, m_e the electron rest mass, c the speed of light and β the proton speed divided by c[19].

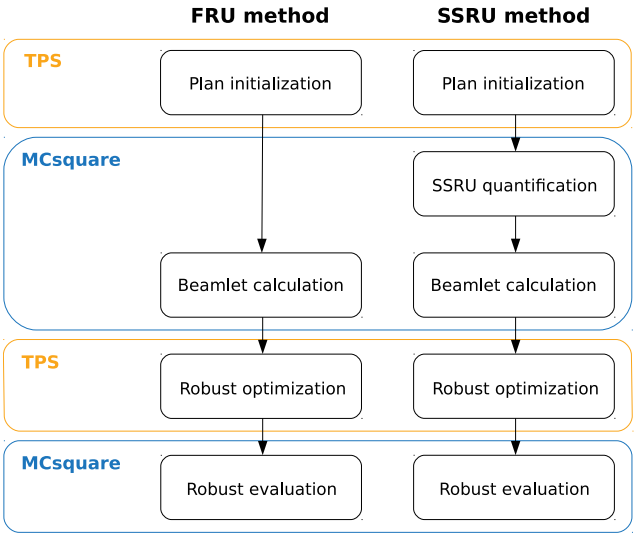

Figure 1: Summary of the implemented workflows: flat range uncertainty (FRU) versus spot-specific range uncertainty (SSRU) planning methodologies.

The I-value of the tissue is generally approximated using Bragg's additivity rule:

$$\ln I_{\mathrm{t}} = \sum_j w_{j,\mathrm{t}} \frac{\left(\frac{Z_j}{A_j}\right)}{\left(\frac{Z}{A}\right)_{\mathrm{t}}} \ln I_j \ , \tag{2}$$

where $w_{j,\mathrm{t}}$ is the mass fraction of element j in the tissue, I_j its I-value, Z_j its atomic number, A_j its atomic mass and $\left(\frac{Z}{A}\right)_{\mathrm{t}}$ is defined as

$$\left(\frac{Z}{A}\right)_{\mathrm{t}} = \sum_j w_{j,\mathrm{t}} \frac{Z_j}{A_j}.$$

The I-value of water, in the denominator of the SPR, is chosen according to the latest recommendations of ICRU. Various values were proposed over time, but the last measurements suggest to use 78 eV, although many TPS still use the previous recommendation of 75 eV. For range estimations in water, this is not a concern as long as consistency is kept with regard to the value used during the commissioning of the proton treatment room. However, the way the stopping power is computed in the numerator remains important. Consistency is key, as mentioned in ICRU report 49[23], from which we copy a paragraph that illustrates the issue for striated muscle: "For mixtures, the additivity rule was applied to a combination of atomic and molecular constituents, specifically including those molecular constituents for which direct experimental I-values are available. For example, striated muscle was treated as containing the constituent water (15.3 percent by weight), and the experimental I-value for liquid water (75.0 eV) was used for this constituent. This provided assurance that such quantities as water/muscle stopping-power ratios are consistent.".

As previously mentioned, De Smet et al.[22] proposed in 2017 a water-based formalism to compute the SPR, in order to better account for the recommendations of ICRU when computing the SPR uncertainties related to I-value uncertainties. This formalism accounts explicitely for the correlation between I-values of tissues and water by applying Bragg's additivity rule on water and "dry" parts of the tissues, i.e.,

$$\ln I_t = w_{\mathrm{w,t}} \frac{\left(\frac{Z}{A}\right)_{\mathrm{w}}}{\left(\frac{Z}{A}\right)_t} \ln I_{\mathrm{w}} + (1 - w_{\mathrm{w,t}}) \frac{\left(\frac{Z}{A}\right)_{\mathrm{dt}}}{\left(\frac{Z}{A}\right)_t} \ln I_{\mathrm{dt}} \ , \tag{3}$$

where $w_{\mathrm{w,t}}$ is the water mass fraction of the tissue, $\left(\frac{Z}{A}\right)_{\mathrm{w}}$ the average number of electrons per

atomic weight of water, $\left(\frac{Z}{A}\right)_{\mathrm{dt}}$ the average number of electrons per atomic weight of the water-free part of the tissue and I_{dt} its I-value. We thus have

$$\left(\frac{Z}{A}\right)_{\mathrm{dt}} = \sum_j w_{j,\mathrm{dt}} \frac{Z_j}{A_j} \tag{4}$$

and

$$\ln I_{\mathrm{dt}} = \sum_j w_{j,\mathrm{dt}} \frac{\left(\frac{Z_j}{A_j}\right)}{\left(\frac{Z}{A}\right)_{\mathrm{dt}}} \ln I_j \ . \tag{5}$$

Using this formalism, I_{w} then appears in both the numerator and the denominator of the SPR. This enables a strong reduction of the global SPR uncertainty. In the extreme case where the considered tissue is water, the SPR uncertainty drops to zero. In the opposite case where the considered tissue would contain no water at all, we fall back onto the high uncertainty of the classical formalism. It is important to point here that only the SPR uncertainty tends to zero, not the range uncertainty since other sources of uncertainties (like beam calibration) may still play a role, even if the material is 100 % water. Moreover, if the fractional weight of water is not perfectly known, this will also impact the uncertainty estimates.

More details on the method can be found in the original paper[22]. However, before continuing, we would like to draw the reader's attention to the fact that this formalism tackles the correlation between numerator and denominator of the SPR, not the correlation between different tissues of the material assignment scheme typically used in the calibration of single-energy CT scanners.

B. Optimization workflow

We describe in this section the global implementation of the compared optimization frameworks.

1. Tools

The Python-based TPS OpenTPS, developed in-house, was used for this study. This tool features the MCsquare dose engine, a multi-core CPU-accelerated Monte Carlo code dedicated to the simulation of proton therapy treatments[24].

2. Robust optimization

A robust optimization framework was implemented within OpenTPS based on the limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm[25] provided with Scipy[26]. A classical worst case approach was chosen. MCsquare was modified to include a spot-specific range uncertainty estimation before the dose calculation. More information on this procedure can be found in subsection II C.

Beamlets are calculated for all considered error scenarios, and the algorithm then optimizes according to a minimax method[12]: the optimizer computes the objective function for each scenario and iterates on the worst case to obtain the best possible solution. Mathematically, the problem can be formulated as

$$w_{\text{opt}} = \arg \min_{w \geq 0} \max_{s \in S} f(w, s) \tag{6}$$

with S the set of considered scenarios, w the spots weights, f the objective function and w_{opt} the optimal weights, i.e., the solution of the problem.

The objective function is a weighted sum of quadratic objectives enforcing a minimum, maximum, or maximum mean dose in given regions of interest (ROIs).

3. Treatment uncertainties

Systematic range and setup errors were here considered in the robust optimization.

For the range, we assumed a 1D Gaussian distribution with zero mean. A constant standard deviation of $\Sigma_r = 1.6\%$ was used in the classical method. This corresponds to the uncertainty recommended by Paganetti due to CT-related issues, as well as I-values and range degradation (for Monte Carlo algorithms)[17]. In the spot-specific implementation, the 1.5% uncertainty on the mean excitation energy was replaced with a variable uncertainty calculated for each spot according to the water content of the materials along the spot track. This variable uncertainty on I_t was then quadratically added to the values accounting for the remaining sources of range uncertainty (see Table 1 in Paganetti's study[17]). More details on the implementation of these spot-specific range uncertainties are given in section II C.

For the setup, isometric normal distribution with zero mean and 1.6 mm standard deviation ($:= \Sigma_s$) was selected, according to the clinical practice in our institution for head and neck

patients. This was assumed to cover all types of possible setup errors occurring during the treatment, including those between fractions. Setup errors were simulated by shifting the CT image according to beam and treatment couch angles. The simulation of setup errors was identical for both the FRU and the SSRU methodologies.

4. Selection of scenarios

The beamlets needed for the optimization of each plan were calculated considering 21 scenarios, aiming at covering 90% of the 1D probability distribution of the range and 90% of the 3D probability distribution of the setup. This complies with what is done in typical commercial TPSs.

For the range, 3 scenarios were considered, with r_{nom} denoting the nominal range: the nominal (r_{nom}), the overshoot ($r_{\mathrm{nom}} + \alpha^{\mathrm{1D}}_{90\%}\Sigma_r$), and the undershoot ($r_{\mathrm{nom}} - \alpha^{\mathrm{1D}}_{90\%}\Sigma_r$) cases, with $\alpha^{\mathrm{1D}}_{90\%} = 1.6$ for a Gaussian distributions. This yielded a 2.6% range error in the classical case, with no spot-specific implementation.

For the setup, 7 scenarios were considered, with (x, y, z) denoting the position of a spot: $(x + \alpha^{\mathrm{3D}}_{90\%}\Sigma_s, y, z)$, $(x - \alpha^{\mathrm{3D}}_{90\%}\Sigma_s, y, z)$, $(x, y + \alpha^{\mathrm{3D}}_{90\%}\Sigma_s, z)$, $(x, y - \alpha^{\mathrm{3D}}_{90\%}\Sigma_s, z)$, $(x, y, z + \alpha^{\mathrm{3D}}_{90\%}\Sigma_s)$, $(x, y, z - \alpha^{\mathrm{3D}}_{90\%}\Sigma_s)$ and the nominal case (x, y, z), with $\alpha^{\mathrm{3D}}_{90\%} = 2.5$ for a Gaussian distribution, leading to a setup error of 4 mm.

Combining range and setup errors yielded in the robust optimization process the 21 aforementioned scenarios.

C. Estimation of spot-specific range uncertainties

In this section, we describe how we implemented the spot-specific range uncertainty within our TPS and its dose engine. We focus first on the contribution of I-value uncertainties. It is important to note here that evaluating stopping power ratios and its associated uncertainties using molecular constituents has already been recommended in ICRU report 49[23] which states: "Improved accuracy can be obtained by applying the additivity rule not to constituent atoms, but to molecular fragments or functional groups.". The Bragg additivity rule applied to elemental composition ignores the effect of molecular bindings on the resulting compound I-value, hence a larger uncertainty. By separating water from the dry tissue, we already diminish the uncertainty

substantially, as described in ICRU report 49[23], Desmet et al[22] and Bar et al[27]. In this work, we want to further diminish the uncertainty by accumulating the uncertainties related to the I-value over the full range of the protons, by computing uncertainties at the most adequate level, ie the molecular fragments or functional groups. If those uncertainties were fully independent voxel-by-voxel, we could simply add them in quadrature, which would diminish the overall uncertainty, roughly with an inverse proportionality to the square root of the number of voxels. However, this will lead to an underestimation of the accumulated uncertainty, because the correlations between molecules of different voxels in the proton path are ignored. Here, we follow a more rigorous approach where we assume that the same protein (or other functional group) in tissue A of a given voxel will have exactly the same I-value (and the same uncertainty) than the same protein (or other functional group) in a different tissue B in another voxel. In such case, the uncertainty remains constant with the number of voxels. The formalism described below implements this conceptual framework.

1. Derivation of the spot-specific contribution to the range uncertainties of I-value uncertainties

As the beam is calibrated in water for proton therapy, the range uncertainty is usually defined as a percentage of the water equivalent range (WER) of protons, which can be written as

$$R_{\mathrm{w}} = \int_0^R \mathrm{SPR}(E(r))\, dr \ , \tag{7}$$

considering a continuous slowing down approximation (CSDA). Here E is the proton energy and R the proton range. A more practical, discretized and approximated form of this equation is the Riemann sum

$$R_{\mathrm{w}} = \sum_{i=0}^{S} \mathrm{SPR}(E_i)\, \Delta r \ , \tag{8}$$

with $S\Delta r = R$ and where Δr should be small.

According to the error propagation theory and assuming independent uncertainties $\sigma_{x_1}, ..., \sigma_{x_N}$ on variables $x_1, ..., x_N$, the uncertainty on a function $f(x_1, ..., x_N)$ is given by

$$\sigma_f = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_{x_i}^2} \ . \tag{9}$$

This formula can thus simply be applied to the range defined in equation (8) to retrieve a range uncertainty, according to the different sources of uncertainty to be considered.

Another possibility would have been to use a Monte Carlo approach and simulate protons transportation many times for each spot, sampling I-value errors according to their assumed distribution at each run. We would then get a distribution of ranges of which we could compute the standard deviation. However, considering the large amount of runs that would be necessary to obtain a stable value along with the significant number of beamlets often necessary to cover a target, this would be highly inefficient.

We thus propagated our different sources of uncertainty through the WER. These were the uncertainty on the dry part of the I-value, the composition of the tissues, and the I-value of water. In the rest of this section, we describe each of these uncertainties and how they affect the global range uncertainty.

The first uncertainty to be considered concerns the dry part I-value I_{dt}. In the original water-based formalism aiming to explicitly take into account the correlation between I-values of water and tissues, De Smet et al. propose to consider an uncertainty on I_{dt} directly[22]. This works well when simulating protons in a single homogeneous tissue. However, when considering heterogeneous media such as patients, sampling I_{dt} independently for each tissue assigned by the CT calibration would lead to neglecting a potentially strong correlation between I_{dt} of similar tissues (notice that this correlation is different and unrelated to the one previously mentioned and tackled by the water-based formalism). For this reason, we consider uncertainties on the I-value of molecules (see section II C 2) rather than on I_{dt} and use Bragg's additivity rule of equation (5) to retrieve I_{dt}. In other words, we transform the water-based formalism into a classical formalism with an I-value calculation that is now based on a molecular decomposition (including water) instead of an elemental decomposition, leading to the following adaptation of Bragg's additivity rule:

$$\ln I_{\mathrm{t}} = \sum_j w_j^{\mathrm{mol}} \frac{\left(\frac{Z}{A}\right)_j^{\mathrm{mol}}}{\left(\frac{Z}{A}\right)_{\mathrm{t}}} \ln I_j^{\mathrm{mol}} \quad , \tag{10}$$

where I_j^{mol} is the I-value of j^{th} molecule, w_j^{mol} its mass fraction in the tissue and $\left(\frac{Z}{A}\right)_j^{\mathrm{mol}}$ its average number of electrons per atomic weight. Using a Bragg additivity rule based on a molecular composition therefore allows us to keep considering the correlation between a tissue and water (*intra-tissue correlation*) while better taking into account the correlation between different

tissues than with the regular water-based formalism as initially proposed by De Smet et al.[22] (*inter-tissue correlation*). The first group of variables inducing a range uncertainty is thus the molecular I-values I_j^{mol}, for $j = 1, ..., M$. The derivative of R_{w} with respect to molecular I-value I_j^{mol} is given by

$$\frac{\partial R_{\mathrm{w}}}{\partial I_j^{\mathrm{mol}}} = \sum_{i=0}^{S} \rho_{e,i} \frac{w_{j,i}^{\mathrm{mol}} \left(\frac{Z}{A}\right)_j^{\mathrm{mol}}}{I_j^{\mathrm{mol}} (C(E_i) - \ln I_{\mathrm{w}}) \left(\frac{Z}{A}\right)_{\mathrm{t},i}} \Delta r \ , \tag{11}$$

where

$$C(E_i) = \ln \left(\frac{2 m_e c^2 \beta^2(E_i)}{(1 - \beta^2(E_i)) \mathrm{e}^{\beta^2(E_i)}} \right) \ .$$

This derivative does not apply to the case of water due to the presence of the water I-value in the denominator of the SPR as well. Notice that index i applied to some of the involved variables identifies the tissue at the current stage of the ray tracing.

The second type of uncertainty to be considered is specific to how we accounted for dry I-value uncertainties with respect to molecular composition. By performing an additional calibration to obtain molecular weights from the usual elemental weights, an additional uncertainty is introduced. The derivative of R_{w} with respect to w_j^{mol} is given by

$$\frac{\partial R_{\mathrm{w}}}{\partial w_{j,k}^{\mathrm{mol}}} = \sum_{i=0}^{S} \chi_i \Delta r \ , \tag{12}$$

with

$$\chi_i = \begin{cases} \rho_{e,i} \dfrac{\left(\frac{Z}{A}\right)_j^{\mathrm{mol}} \left(\ln I_j^{\mathrm{mol}} - \ln I_{\mathrm{t,i}}\right)}{(C(E_i) - \ln I_{\mathrm{w}}) \left(\frac{Z}{A}\right)_{\mathrm{t},i}} & \text{if } i = k \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

The equation (13) realizes the idea that the uncertainties related to the same molecules in different voxels should propagate linearly.

Finally, as previously stated, the water I-value has to be considered as a separate case. The derivative of the range with respect to the water I-value when using the water-based formalism is

$$\frac{\partial R_{\mathrm{w}}}{\partial I_{\mathrm{w}}} = \sum_{i=0}^{S} \rho_{e,i} \frac{(C(E_i) - \ln I_{\mathrm{t},i}) \left(\frac{Z}{A}\right)_{\mathrm{t},i} - w_{\mathrm{w},i} \left(\frac{Z}{A}\right)_{\mathrm{w}} (C(E_i) - \ln I_{\mathrm{w}})}{I_{\mathrm{w}} (C(E_i) - \ln I_{\mathrm{w}})^2 \left(\frac{Z}{A}\right)_{\mathrm{t},i}} \Delta_r \ . \tag{14}$$

Using equation (9) in combination with the above partial derivatives and some assumed uncertainties on I-values and molecular weights then allows the spot-specific range uncertainty to

be calculated. One should note that the uncertainties derived above only relate to the conversion from elemental composition to molecular composition and its influence on the estimation of the cumulative impact of I-value uncertainties. In order to accumulate accurately the range uncertainties due to the I-value uncertainties, we need to properly assign each element to its corresponding molecular fragment. The methods described above take into account how uncertainties linked to the molecular composition translate into spot specific range uncertainties due to I-value uncertainties. The uncertainties related to the conversion from CT numbers to molecular composition will be handled separately in II C 4.

2. CT calibration

As already established, CT calibration is essential when it comes to range uncertainties. Although the considered spot-specific part of the range uncertainty is not directly related to CT calibration, I-values depend on the composition of tissues.

The CT calibration in MCsquare relied on the classical stoichiometric method for Monte Carlo algorithms[28]. It provided us with conversions from CT numbers to mass densities and from CT numbers to elemental compositions.

As mentioned, a CT calibration scheme enabling a conversion to a molecular composition was also necessary. We achieved this based on the works of Woodard and White, which initially reported properties of the main ICRU human tissues often used for CT calibration[29,30]. In this study, the authors also provide the composition of each tissue in terms of water, lipids, proteins, carbohydrates, bone ash and other residual molecules. We used this information to assign a molecular composition to each of the tissues considered by our CT calibration curve. More details on the procedure can be found in Appendix A.

3. Uncertainties assumed for the I-value uncertainty contribution in the SSRU calculation

In order to apply the presented formalism, it is necessary to make some initial assumptions on the uncertainties of the different variables introduced above.

Following the latest guidelines from ICRU, an I-value of 78 eV and a standard deviation of 2 eV were assumed for water while the I-value of air was set to 85.7 ± 1.2 eV[31]. Note that for consistency in the case of water, it is essential to use the same mean excitation energy value as

in the dose calculation performed during the commissioning of the treatment room.

For the uncertainties on the other molecular I-values, we followed recommendations of ICRU Report 37[32]. In this report, the authors mention that for all compounds containing mainly H, C, N, O, F, and Cl, an uncertainty between 5% and 10% on the I-value should be considered. For other molecules, it rises up to 15%. As these values are given with a 90% CI, we thus used in our simulations a standard deviation of 6.1% for all molecules except bone ash that mainly consists of calcium and phosphorus and for which a standard deviation of 9.1% was selected. These two numbers correspond to 10% and 15% uncertainty, respectively, with a 90% CI. Nominal I-values were calculated using the elemental Bragg additivity rule and elemental I-values suggested in ICRU 37 for atomic constituents of compounds[32].

Finally, the uncertainties to be applied to the molecular fractional weights were determined from the initial fit process realized during the calibration (see Appendix A). Note that obtained values were unimportant in terms of resulting range uncertainty when compared to I-values uncertainties.

4. Integration in a Monte Carlo dose engine and computation of the total SSRU

In order to compute spot-specific uncertainties, raytracing through the CT image was implemented in the MCsquare dose engine. For a given spot with a given energy, it allows calculating the water equivalent range and the scoring of applicable uncertainties according to the traversed voxels. Partial derivatives are incremented at each step of the raytracing, and the cumulative range uncertainty related to mean excitation energies is then computed at the end of the raytracing using equation (9). Scattering is ignored during the process. Then, the total range uncertainty (standard deviation) is computed as

$$\Sigma_r = \sqrt{\Sigma_{\mathrm{I}}^2 + \Sigma_{\mathrm{CT1}}^2 + \Sigma_{\mathrm{CT2}}^2 + \Sigma_{\mathrm{CT3}}^2 + \Sigma_{\mathrm{RD1}}^2 + \Sigma_{\mathrm{RD2}}^2} \ , \tag{15}$$

where Σ_I is the range uncertainty due to I-values, Σ_{CT} account for CT-related uncertainties (imaging and calibration, conversion to tissue, grid size) and Σ_{RD} account for range degradation (complex inhomogeneities and local lateral inhomogeneities). Following Paganetti's recommendations, we use in this study $\Sigma_{\mathrm{CT1}} = 0.5\%$, $\Sigma_{\mathrm{CT2}} = 0.2\%$, $\Sigma_{\mathrm{CT3}} = 0.3\%$, $\Sigma_{\mathrm{RD1}} = 0.1\%$ and $\Sigma_{\mathrm{RD2}} = 0.1\%$ (see Table 1 in the original study[17]).

The spot-specific range uncertainty calculation takes place before the MC simulation of the beamlets. It has to be performed once per scenario to correctly model range uncertainties due to setup errors. Then, in the MC simulation, each particle carries the information on the current simulated spot along with its range uncertainty, and the density is locally scaled accordingly.

D. Evaluation of the proposed method

1. Patient cases

Three head and neck patients (oro-pharynx cases) were planned using both the flat and the spot-specific range uncertainty methods (FRU and SSRU, respectively). Beam angles of 60°, 120°, 240° and 300° with a range shifter were used for all cases. The couch angles were set to 10°, 350°, 10° and 350° for each beam, respectively.

The prescription was 70 Gy for both the primary tumor (CTVp_7000) and the nodal region having infiltrated tumor (CTVn_7000) and 54.25 Gy for the prophylactic lymph nodes (CTVn_5425). Minimum and maximum dose objectives on targets were robustly optimized, except for the maximum dose of CTVn_5425 because this CTV also included parts of targets receiving 70 Gy. Robust maximum dose objectives were also defined for sensitive OARs (brainstem, spinal cord) while for other OARs only the nominal mean dose was minimized.

Initial CT resolution was $1.07 \times 1.07 \times 2$ mm^3. A grid of $2.5 \times 2.5 \times 2.5$ mm^3 was used for scoring the dose in all calculations.

2. Calculated uncertainties for the spot-specific range uncertainty method

As a first step, we analyzed the range uncertainties resulting from our variable Σ_I when using the SSRU method. The distribution of range uncertainties over all beamlets of patient 1 was computed, and the range variation in two beamlets was studied after simulating errors in a Monte-Carlo way.

3. Organs at risk

The main goal of the present study is to assess if a decreased range uncertainty can result in a toxicity reduction for the patients. The mean dose or D_2 of OARs were therefore compared between the classical and the spot-specific range uncertainty frameworks. In order to get a fair comparison, DVHs were scaled to provide identical target coverage in the nominal case for both methods. The D50 of the primary CTV was set to the prescription, and the dose was scaled accordingly. OAR dose metrics could this way be compared consistently. These metrics were chosen according to the standard H&N practice at our institution, in order to assess the real clinical impact of the proposed methodology. D_2 was thus reported for serial organs (brainstem and spinal cord) while the mean dose was reported for all other organs.

4. Robustness evaluation

Plan robustness was assessed. Robustness tests were performed using the statistically consistent method proposed in Souris et al.[33] and Sterpin et al.[34]. Instead of simulating extreme scenarios only such as in the optimization process, errors were randomly sampled before each simulated scenario, using the same standard deviation as in robust optimization. The X% best scenarios were then kept to build the dose volume histogram bands (DVH-bands). Their selection was performed as a function of the relative target coverage of the three different kinds of CTVs using the metric

$$M = 0.4\,\frac{\mathrm{D}_{98}^{\mathrm{CTVp\_7000}}}{70} + 0.35\,\frac{\mathrm{D}_{98}^{\mathrm{CTVn\_7000}}}{70} + 0.25\,\frac{\mathrm{D}_{98}^{\mathrm{CTVn\_5425}}}{54.25}\ , \tag{16}$$

with $\mathrm{D}_{98}^{\mathrm{R}}$ the D_{98} of ROI R. Scenarios with the highest M were considered the best ones. The chosen weights aimed at prioritizing infected ROIs. A confidence level of 90% was selected and 300 scenarios were simulated for each plan, which was considered to be a sufficient number of scenarios, allowing for a comprehensive and statistically sound robustness analysis[33]. With this method, we thus perform the scenario selection in the dose space and not in the error space, which is more costly in time but also clinically more relevant.

Spot-specific range uncertainties were used in both plan evaluations to provide realistic robustness simulations. We thus expect that a classical plan is overly robust to the simulated

uncertainties, as the spots specific range uncertainties should be smaller than the flat one used in the classical method.

As in the optimization, a 1.6 mm standard deviation was selected as global setup error. It was considered to be fully systematic and no random setup error was thus considered, making unnecessary the simulation of fractionation. The spot-specific range uncertainty was computed with equation (15). For each scenario, a single error was sampled according to a standard normal distribution $N(0, 1)$ and was then multiplied by each spot-specific range uncertainty to obtain the correct range uncertainty for each spot.

It is worth noting that other methods for statistically consistent robustness evaluation could have been considered, using faster methods such as polynomial chaos expansion[35] or other dose metamodels. Monte Carlo, however, remains the gold standard for accurate dose calculation and the simulation of 300 scenarios was deemed to be sufficient to obtain statistical significance[33].

III. RESULTS

A. Calculated uncertainties for the spot-specific range uncertainty method

Figure 2 shows the distribution of total range uncertainties obtained over all beamlets of patient 1, at one standard deviation. The flat range uncertainty of 1.6% is thus confirmed to be a worst case. The median is 0.69%, corresponding to a Σ_I of 0.28%.

Figure 3 shows the range distribution of two different beamlets of patient 1 when simulating the aforementioned I-values errors. Distributions are very slightly right-skewed (skewness coefficients of 0.04 and 0.06 for 69.7 MeV and 151.9 MeV beamlets, respectively).

B. Dosimetric results

Table I shows the robustness results on CTV targets for each patient. Worst (over the kept scenarios) and nominal D_{95} and D_5 are reported for CTVp_7000 and CTVn_7000. Worst and nominal D_{95} are reported for CTVn_5425 while the D_5 is computed on CTVn_5425 - (CTVp_7000 + CTVn_7000 + 1.25 cm) in the nominal case only as no robust objective was set on the maximum dose of CTVn_5425. As expected, SSRU plans are a little less robust than their corresponding FRU plan, while remaining robust enough to guarantee the safety of the treatments ie

436 within the aforementioned 90% confidence level.

437 Table II reports the nominal mean doses or D_2 in various organs at risk affected by the
438 treatment, for both FRU and SSRU methods. The mean relative gain over all patients (MRG),
439 where the gain is defined as the difference in the reported metric (D_{mean} or D_2) between FRU
440 and SSRU methods, is also reported for each OAR. The impact of the large reduction in range

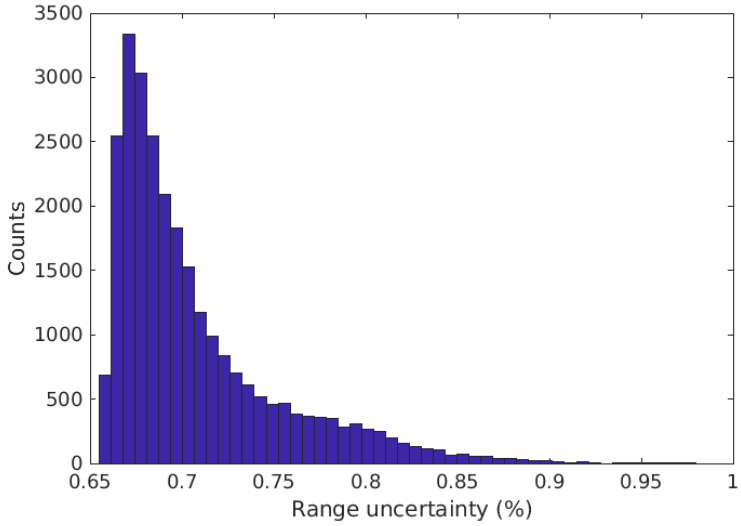

Figure 2: Distribution of range uncertainties (standard deviations) obtained for beamlets of patient 1.

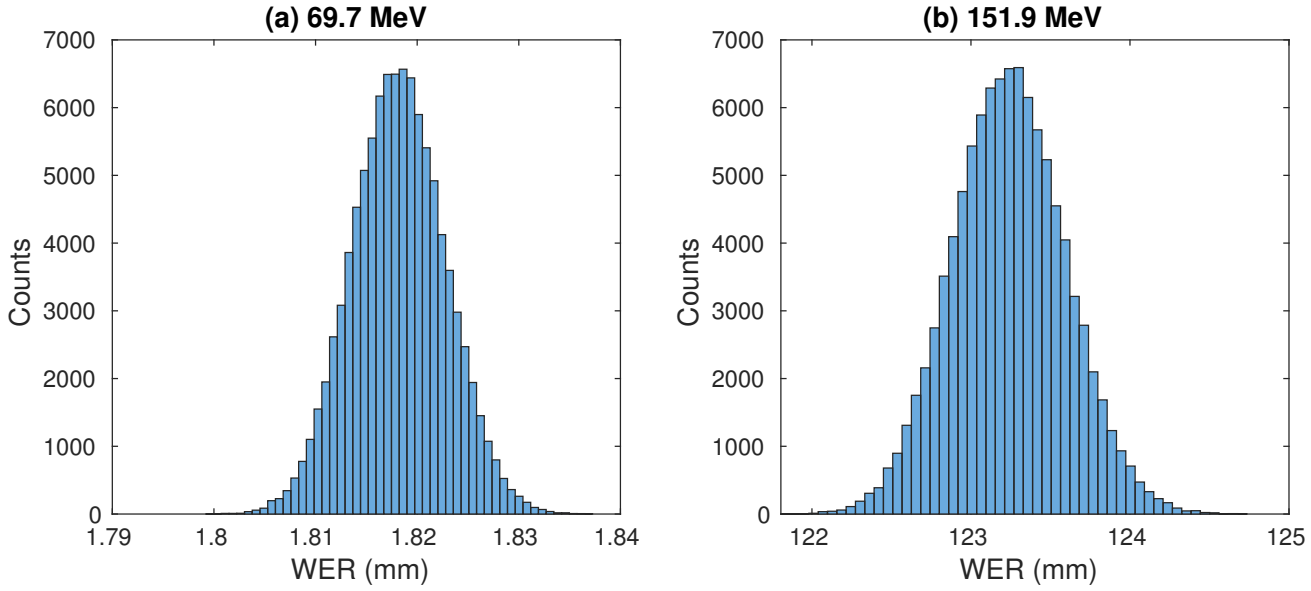

Figure 3: WER (water equivalent range) variation resulting from the simulation of errors in I-values for two different beamlets of patient 1: (a) 69.7 MeV and (b) 151.9 MeV. Notice that a range shifter is used, which explains the small WER values obtained here.

uncertainty is very clear here, with a resulting relative decrease in toxicity for most cases and a positive MRG for all OARs. Because of their location close to the target volume, parotid glands, for instance, are especially well spared using the SSRU method, with mean gains of 15.7% and 8.4% for the left and the right ones, respectively, in the nominal case. Submandibular glands, on the other hand, hardly benefit from the method as a large part of them is typically located within the target volume. Serial organs, finally, do not really benefit from the decreased range uncertainty. This might be due to their location and to the fact that the evaluated metric is a D_2.

Finally, Figures 4 and 5 illustrate our results for patient 1, which was selected arbitrarily for illustration purposes. Figure 4 shows the DVH bands for targets and 3 important OARs (oral cavity, left and right parotids), for both the FRU and the SSRU methods. We see that target DVH bands are slightly larger in the SSRU case. Figure 5 shows the nominal mean dose in organs at risk for patient 1 as well, comparing both methods. The mean dose received by both parotids, for instance, is reduced by about 2 Gy.

Table I: DVH metrics of CTVs for all patients, comparing FRU and SSRU methods. Nominal and worst cases are considered, with the worst case defined at a 90% CI.

Robustness assessment in targets - DVH metrics for a 90% confidence level (Gy)								
			CTVp_7000		CTVn_7000		CTVn_5425 [(a)]	
			D_{98}	D_2	D_{98}	D_2	D_{98}	D_2
P1	Worst	FRU	68.1	72.4	N.A.	N.A.	53.1	N.A.
		SSRU	67.8	72.5	N.A.	N.A.	53.0	N.A.
	Nominal	FRU	68.4	72.0	N.A.	N.A.	53.8	59.9
		SSRU	68.3	72.2	N.A.	N.A.	53.8	59.6
P2	Worst	FRU	68.0	71.9	68.2	72.2	53.1	N.A.
		SSRU	67.9	72.0	67.7	72.2	52.5	N.A.
	Nominal	FRU	68.5	71.6	68.9	72.0	53.9	59.5
		SSRU	68.5	71.7	68.7	71.7	53.8	59.5
P3	Worst	FRU	67.7	72.6	68.1	73.0	53.0	N.A.
		SSRU	67.3	72.6	67.6	73.1	52.7	N.A.
	Nominal	FRU	68.3	72.4	69.1	72.2	53.8	59.6
		SSRU	68.4	72.3	69.1	72.1	53.8	59.3

[(a)] D_2 of CTVn_5425 is calculated on CTVn_5425 - (CTVp_7000 + CTVn_7000 + 1.25 cm)

Table II: Dose metrics in various organs at risk for both the classical (FRU) and the spot-specific range uncertainty (SSRU) methods. The nominal and worst cases (at a 90% CI) are considered, with D_{50} of CTVp_7000 normalized to 70 Gy. OARs are glottic area (GA), oral cavity (OC), left parotid (LP), right parotid (RP), pharyngeal constrictor muscle (PCM), left submandibular (LS), right submandibular (RS), supraglottic larynx (SL), upper esophagus (UE), spinal cord (SC), and brainstem (BS).

			Organs at risk - Mean dose (Gy)										
			GA (D_{mean})	OC (D_{mean})	LP (D_{mean})	RP (D_{mean})	PCM (D_{mean})	LS (D_{mean})	RS (D_{mean})	SL (D_{mean})	UE (D_{mean})	SC (D_2)	BS (D_2)
P1	Worst	FRU	24.9	48.3	29.5	32.0	13.9	57.0	58.1	16.8	4.7	2.6	0.1
		SSRU	20.8	45.6	28.0	29.4	11.1	56.8	57.6	12.0	3.8	2.3	0.1
	Nominal	FRU	18.9	44.0	25.0	28.0	11.0	55.9	56.8	14.7	3.7	1.5	0.1
		SSRU	16.1	42.3	23.1	25.9	8.4	55.8	56.4	10.1	2.7	1.2	0.1
P2	Worst	FRU	32.9	58.2	23.6	31.4	34.3	52.4	66.0	31.2	14.9	22.0	22.2
		SSRU	30.7	56.7	17.8	28.2	30.9	50.4	65.9	27.8	13.0	22.4	21.1
	Nominal	FRU	27.4	55.7	18.7	26.9	30.5	47.7	64.2	23.5	13.0	18.4	19.2
		SSRU	25.9	54.7	13.6	24.6	27.2	46.8	63.9	22.4	11.2	18.9	18.3
P3	Worst	FRU	36.2	42.8	36.7	35.7	46.6	67.3	57.0	26.4	12.6	26.2	9.2
		SSRU	33.8	39.1	32.1	33.6	43.8	67.0	56.9	22.9	11.3	24.8	9.3
	Nominal	FRU	30.0	37.6	33.2	32.2	44.4	66.0	56.0	22.5	10.8	19.2	5.2
		SSRU	28.1	34.2	29.1	29.2	42.1	65.7	55.9	19.6	9.6	18.9	5.3
MRG (%)	Worst		9.9	5.6	14.0	8.1	12.0	1.5	0.4	17.6	14.1	5.0	1.3
	Nominal		8.9	4.9	15.7	8.4	13.2	0.8	0.5	17.3	16.3	6.3	0.9

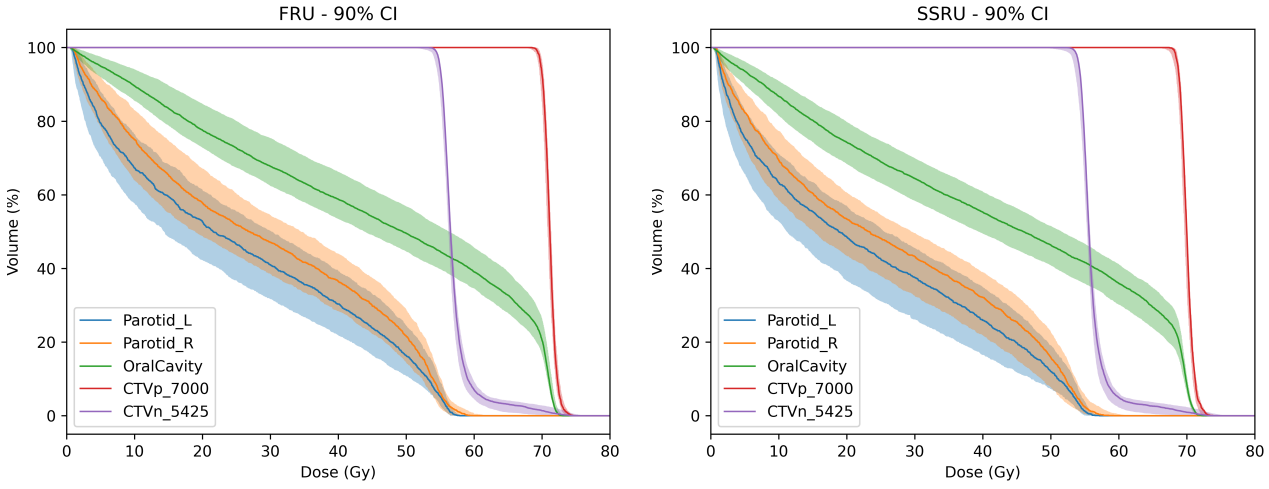

Figure 4: DVH bands for CTVs and three OARs of patient 1 at a 90% confidence level for (a) the FRU method and (b) the SSRU method.

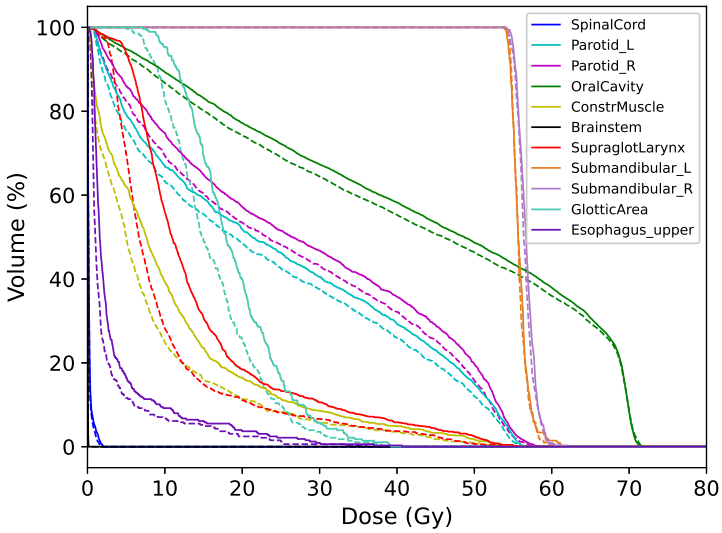

Figure 5: Nominal mean dose in various OARs of patient 1 for the FRU (plain line) and SSRU (dashed line) methods.

IV. DISCUSSION

Accurately quantifying uncertainties involved in particle therapy is essential to benefit the most from that treatment modality. Leading to a reduction of range uncertainties, it could potentially improve patients' disease outcome by allowing to spare healthy tissues better in the

vicinity of the tumor. In 2012, Paganetti studied many sources of range uncertainties, quantified them and proposed a range uncertainty of 2.4% (or 1.6% at one standard deviation) accounting for Monte Carlo dose calculation uncertainties. A very large part of it was assumed to come from uncertainties on I-values (1.5% at one standard deviation)[17].

Taking the correlation between water and tissues I-values into account, however, leads to a more accurate estimation of the impact of I-value uncertainties on the total range uncertainty[22]. The 1.5% value proposed by Paganetti is actually an upper bound and could be much reduced in most cases. The reason is twofold. First, the uncertainty on the water I-value is much lower than that assumed for the I-values of tissues. Second, the SPR is a quotient between stopping powers of tissue and water, which means that the more water in the tissue, the smaller the SPR uncertainty (if this water amount is considered in practice). The exact uncertainty accounting for potential I-value errors is thus very sensitive to the amount of water in tissues. The human body is made of various kinds of tissues, each containing from 10% to 90% of water approximately. According to De Smet et al., this translates into SPR uncertainties from 0.8% to 0.1% at one standard deviation (accounting for I-value uncertainties only)[22]. This wide range of values makes it difficult to come up with a single number to be used in the total range uncertainty calculation. Other studies have attempted to compute range uncertainties according to the tumor site by assuming given proportions of lung, bone and soft tissues[18,21] and given range uncertainties for each of these. Nevertheless, this approach has limitations considering the potential different clinical practices of different proton therapy centers or the differences in patients anatomies. Therefore, we proposed here a spot-specific range uncertainty calculation. Being fully automated and very fast, accurate prediction of how I-values affect the range uncertainty comes at no cost and thus provides a more patient-specific treatment planning.

This methodology was carried out to plan three H&N patient cases retrospectively. Analyzing the distribution of range uncertainties for patient 1, we observed a median of 0.69%, corresponding to a Σ_I of 0.28%. This value seems to be consistent with the recent study of Peters et al.[21], who reported estimated uncertainties of 0.32%, 0.39%, and 0.28% for head, pelvis, and lung regions, respectively, and with the older study of Yang et al. who reported values of 0.2%, 0.2%, and 0.6% for lung, soft tissues and bones, respectively[18]. The shape of the range distribution for given beamlets was also investigated. It was found to be almost symmetrical, although not Gaussian. This distribution, however, only accounts for I-value uncertainties and does not necessarily reflect the distribution of the range when accounting for all treatment uncertainties.

Moreover, Wieser et al. showed that robust optimization remained superior to a PTV-based approach even if incorrect error distributions were assumed[36].

When applied retrospectively to head and neck cancer patients, the SSRU methodology shows improved OAR sparing, owing to its aforementioned important range uncertainty reduction. Average gains of 15.7% and 8.4% are for example obtained for left and right parotid glands in the nominal case, respectively. H&N cases are, however, complicated cases in terms of OAR toxicity. Many organs are located very close to the targets, making a reduced range uncertainty especially useful. Other tumor sites could show less difference in OAR sparing. The study of plan robustness globally shows that FRU plans are indeed more robust than their SSRU counterpart. Nevertheless, in terms of target coverage, the effect is not as strong as what we could have expected, considering the magnitude of the range uncertainty reduction. Differences in dose homogeneity (defined here as D_2 - D_{98}) between FRU and SSRU plans are between 0.2 Gy and 0.6 Gy for our patients. We believe that there are several reasons for this. First, the setup error assumed in the robustness evaluation is 3D, while the range error is only 1D, making it less influencial. Then, FRU and SSRU plans of a same patient do not have exactly equivalent levels of robustness in the first place.

Although good dosimetric results are achieved with the SSRU method in terms of toxicity reduction, some limitations need to be discussed Another limitation relates to the ray-tracing algorithm. Because spot specific uncertainties are computed in a straight line, we ignore multiple Coulomb scattering effects which could impact the range uncertainty for a given proton, since the proton will not necessarily travel in a straight line.

Cautious interpretation is also required here, after observing such a drastic decrease of one of our treatment uncertainties. Studies on range uncertainty quantification, such as Paganetti's, were limited themselves by the initial assumptions they make and the uncertainties they consider. The total range uncertainty coming from dose calculation as recommended by Paganetti is 2.6% at a 90% CI. Yang et al., in 2012, came up with a similar value eventually[18]. When looking into it, however, the main contributing sources of range uncertainty are very different from one study to the other. Some appear in one and not the other, and vice versa. Rightly so, clinicians would rather place more importance on the total recommended value itself than on its justification, especially if it is consistent with the standard values used in clinics. For the present study, however, we had to use intermediate values accounting for other sources of range uncertainty. Starting with Paganetti's values allowed us to observe an important decrease of the total range

uncertainty and, subsequently, of the OAR toxicity. However, using the intermediate values suggested by Yang et al. would have yielded much less decreased uncertainties as these authors had already considered the correlation between water and tissues I-values.

Nevertheless, we believe that, regardless of the used intermediate uncertainties, there would be a benefit to move towards a spot-specific quantification of the range uncertainty. Figure 2 supports that statement. For a given patient, the total range uncertainties vary between 0.65% and 0.98% (0.17%-0.75% for I-values only). Using the SSRU quantification, which comes at a very negligible computational cost, would also ensure that a sufficient range uncertainty was considered for all beamlets.

To further assess its benefit, the spot-specific range uncertainty calculation should now be applied to more patients, with various tumor sites. Here, indeed, only three patients were considered. In addition, other parts of the range uncertainty could be made variable as well, such as CT noise or calibration-related issues. It would mainly be interesting for the most tissue-specific ones. According to Yang et al., the uncertainties due to CT imaging and CT calibration would mostly be concerned, including the deviation of body tissues from ICRU standard tissues[18].

V. CONCLUSIONS

We propose a novel way of accounting for range uncertainties in the robust optimization of proton therapy treatments. By computing a spot-specific range uncertainty, we automatically ensure that the value assumed in the robust optimization is well adapted, regardless of the patient or of the clinical protocols. The spot-specific range uncertainty method is here applied to the uncertainty component coming from I-values by taking advantage of the high amount of water in the human body. It induces uncertainty reductions, compared to the value proposed by Paganetti in 2012[17]. The exact range reduction achieved will depend on the estimation of the other sources of uncertainties. However, there is a potential for substantial reduction of OAR exposure.

VI. ACKNOWLEDGMENTS

Marie Cohilis and Gregory Buti are supported by the Télévie Grant from the Belgian "Fonds National pour la Recherche Scientifique" F.R.S-FNRS (Grant No. 7450517F and 7453918F respectively). Kevin Souris is funded by the Walloon Region (MECATECH/BIOWIN, Grant No. 8090). John A. Lee is a Research Associate with the Belgian fund of scientific research (F.R.S.-FNRS).

VII. CONFLICT OF INTEREST STATEMENT

The authors have no relevant conflicts of interest to disclose.

APPENDIX A

This appendix details how the tissues were assigned a molecular composition.

In their papers of 1986 and 1987, Woodard and White provide a list of human tissues and their recommended properties in the context of medical physics[29,30]. Information such as mass density and elemental composition can be found there for more than 70 soft tissues and bones. These data have been widely used ever since then to perform the CT calibration necessary for radiotherapy.

In that study, soft tissues are also decomposed into a mixture of water, lipid, protein, and residuals. In what follows, we will refer to this as the WLP decomposition. In addition, the main tissues components and their elemental composition are provided. These compounds are listed in Table III, but we refer the reader to the original paper for their elemental compositions[29].

Combining this information with elemental and WLP decompositions allowed us to extract a molecular decomposition for all soft tissues, which were thus expressed as a mixture of the compounds listed in Table 1.

A similar process was carried out for bones. In their study of 1987[30], White and Woodard report a decomposition in osseous tissue, red marrow, yellow marrow, and cartilage. Osseous tissue was found to correspond approximately to hydroxyapatite, one of the main bone components, responsible for the presence of calcium and phosphorus. The knowledge of the WLP decomposition for marrows and cartilage allowed us to assign all bones a molecular decomposition based

on the molecules of Table III, with the addition of an osseous tissue component.

Finally, linear fits were performed over all tissues in order to relate the mass fraction of a molecule in a tissue to its elemental composition. Soft tissues and bones were treated separately. For each molecule M considered for the decomposition, the two most relevant elements E_1 and E_2 were selected to perform the fit and predict the molecular mass fraction w_M using the formula

$$w_M = a + b w_{E_1} + c w_{E_2} \ , \tag{17}$$

with w_{E_1} and w_{E_2} the mass fractions of the first and the second element, respectively, and a, b, and c the fit parameters.

Tables IV and V report, for soft tissues and bones, respectively, which elements E_1 and E_2 were selected for each molecule, along with the obtained coefficients a, b, and c enabling the prediction of molecular mass fractions. Uncertainties on these parameters are given as well. Molecules from Table III not listed in Tables IV and V had non significant weights for all or most tissues and were thus neglected in the decomposition. Sphingomyelin is an exception for soft tissues: no specific trend was detected but its mass fraction was approximated by 0.004, the median weight over all soft tissues. Its uncertainty was set to the standard deviation over all tissues, equal to 0.0152. It should be noted that the uncertainties reported here only represent the additional uncertainty resulting from the conversion from elemental composition to molecular composition, and not from CT numbers to molecular composition.

Table III: List of main molecular components of human soft tissues. Lipids, proteins and water are the most important ones[29].

Main components of the soft tissues
Carbohydrate - chondroitin sulphate
Carbohydrate - monosaccharide
Carbohydrate - polysaccharide
Cholic acid
Lipid - cerebroside
Lipid - cholesterol
Lipid - glycerol trioleate
Lipid - sphingomyelin
Mucin
Protein
Urea
Water

Table IV: List of main molecular components of soft tissues and their parameterization as a function of their elemental composition: E_1 and E_2 are the elements considered in the parameterization; a, b and c the parameters as in Equation (17); σ_a, σ_b and σ_c their standard deviations and σ_{ab}, σ_{ac} and σ_{bc} their correlations.

Molecule	E_1	E_2	a	b	c	σ_a	σ_b	σ_c	σ_{ab}	σ_{ac}	σ_{bc}
Cerebroside	H	O	3.84e-5	1.68e-1	-2.13e-2	2.14e-2	1.87e-1	5.68e-3	-3.96e-3	-6.81e-5	4.45e-4
Cholesterol	C	N	4.52e-3	5.64e-2	5.48e-2	3.91e-3	9.15e-3	1.07e-1	-2.02e-5	-3.49e-4	1.70e-4
Glycerol trioleate	C	N	-1.93e-2	1.22	-3.99	9.99e-3	2.34e-2	2.75e-1	-1.32e-4	-2.81e-3	1.11e-3
Protein	C	N	-1.49e-2	2.15e-2	6.15	4.47e-3	1.05e-2	1.23e-1	-2.65e-5	-4.56e-4	2.22e-4
Water	H	O	-4.69e-1	2.80	1.30	1.84e-2	1.61e-1	4.88e-3	-2.92e-3	-5.03e-5	3.29e-4

Table V: List of main molecular components of bones and their parameterization as a function of their elemental composition: E_1 and E_2 are the elements considered in the parameterization; a, b and c the parameters as in Equation (17); σ_a, σ_b and σ_c their standard deviations and σ_{ab}, σ_{ac} and σ_{bc} their correlations.

Molecule	E_1	E_2	a	b	c	σ_a	σ_b	σ_c	σ_{ab}	σ_{ac}	σ_{bc}
Monosaccharide	H	O	8.81e-2	-7.16e-1	-2.74e-2	5.29e-4	4.78e-3	9.96e-4	-1.5e-6	-4.1e-7	8.6e-8
Cerebroside	H	O	7.13e-3	15.6e-1	-2.83e-2	5.26e-4	4.75e-3	9.90e-4	-1.49e-6	-4.05e-7	-8.5e-8
Cholesterol	H	O	1.91e-2	4.17e-1	-7.56e-2	1.37e-3	1.24e-2	2.58e-3	-1.01e-5	-2.76e-6	5.76e-7
Glycerol trioleate	C	N	-4.36e-2	1.25	-3.59	2.98e-3	4.33e-3	5.76e-2	-1.13e-5	-1.62e-4	1.74e-4
Sphingomyelin	H	O	1.85e-3	7.19e-2	-9.14e-3	9.22e-4	8.34e-3	1.74e-3	-4.58e-6	-1.25e-6	-2.6e-7
Protein	C	N	-2.74e-2	3.54e-2	6.35	3.97e-3	5.77e-3	7.67e-2	-2.01e-5	-2.88e-4	3.08e-4
Water	H	O	-5.93e-1	3.79	1.35	1.90e-2	1.71e-2	3.57e-3	-1.93e-5	-5.26e-6	-1.10e-6
Osseous tissue	N	Ca	4.31e-4	2.45e-3	2.58	6.97e-4	2.51e-2	3.43e-3	-1.49e-5	2.88e-7	5.16e-5

[1] C. H. Clark, A. M. Bidmead, C. D. Mubata, K. J. Harrington, and C. M. Nutting, Intensity-modulated radiotherapy improves target coverage, spinal cord sparing and allows dose escalation in patients with locally advanced cancer of the larynx, Radiother. Oncol. **70**, 189–198 (2004).

[2] X. Zhang, Y. Li, X. Pan, L. Xiaoqiang, R. Mohan, R. Komaki, J. D. Cox, and J. Y. Chang, Intensity-Modulated Proton Therapy Reduces the Dose to Normal Tissue Compared With Intensity-Modulated Radiation Therapy or Passive Scattering Proton Therapy and Enables Individualized Radical Radiotherapy for Extensive Stage IIIB Non-Small-Cell Lung Cancer: A Virtual Clinical Study, Int. J. Radiat. Oncol. Biol. Phys. **77**, 357–366 (2010).

[3] Y. Kase, H. Yamashita, H. Fuji, Y. Yamamoto, Y. Pu, C. Tsukishima, and S. Murayama, A treatment planning comparison of passive-scattering and intensity-modulated proton therapy for typical tumor sites, Journal of radiation research **53**, 272–280 (2012), [DOI:10.1269/jrr.11136]

606 [PubMed:22129564].

607 [4] A. W. Jie and L. Marignol, Pro-con of proton: Dosimetric advantages of intensity-modulation over
608 passive scatter for thoracic malignancies, Technical Innovations & Patient Support in Radiation
609 Oncology **15**, 37–46 (2020), [DOI:10.1016/j.tipsro.2019.11.005] [PubMed:32954018].

610 [5] X. Tian, K. Liu, Y. Hou, J. Cheng, and J. Zhang, The evolution of proton beam therapy:
611 Current and future status, Mol Clin Oncol **8**, 15–21 (2018), [DOI:10.3892/mco.2017.1499]
612 [PubMed:29399346].

613 [6] S. McGowan, N. Burnet, and A. Lomax, Treatment planning optimisation in proton therapy, Br
614 J Radiol **86**, 20120288 (2013), [DOI:10.1259/bjr.20120288] [PubMed:23255545].

615 [7] W. Liu, S. J. Frank, X. Li, Y. Li, P. C. Park, L. Dong, X. Ronald Zhu, and R. Mohan, Effectiveness
616 of robust optimization in intensity-modulated proton therapy planning for head and neck cancers,
617 Medical physics **40**, 051711 (2013), [DOI:10.1118/1.4801899] [PubMed:23635259].

618 [8] J. Unkelbach et al., Robust radiotherapy planning, Physics in Medicine & Biology **63**, 22TR02
619 (2018), [DOI:10.1088/1361-6560/aae659] [PubMed:30418942].

620 [9] S. Dowdell, C. Grassberger, G. Sharp, and H. Paganetti, Interplay effects in proton scanning
621 for lung: a 4D Monte Carlo study assessing the impact of tumor and beam delivery parameters,
622 Physics in Medicine & Biology **58**, 4137 (2013).

623 [10] G. Bosmans, A. van Baardwijk, A. Dekker, M. Öllers, L. Boersma, A. Minken, P. Lambin, and
624 D. De Ruysscher, Intra-patient variability of tumor volume and tumor motion during conven-
625 tionally fractionated radiotherapy for locally advanced non-small-cell lung cancer: a prospective
626 clinical study, International Journal of Radiation Oncology* Biology* Physics **66**, 748–753 (2006).

627 [11] K. Kraus, E. Heath, and U. Oelfke, Dosimetric consequences of tumour motion due to respiration
628 for a scanned proton beam, Physics in Medicine & Biology **56**, 6563 (2011).

629 [12] A. Fredriksson, A. Forsgren, and B. Hårdemark, Minimax optimization for handling range and
630 setup uncertainties in proton therapy, Medical physics **38**, 1672–1684 (2011).

631 [13] R. Kohno, K. Hotta, S. Nishioka, K. Matsubara, R. Tansho, and T. Suzuki, Clinical implemen-
632 tation of a GPU-based simplified Monte Carlo method for a treatment planning system of proton
633 beam therapy, Physics in Medicine & Biology **56**, N287 (2011).

634 [14] A. M. Barragán Montero, K. Souris, D. Sanchez-Parcerisa, E. Sterpin, and J. A. Lee, Performance
635 of a hybrid Monte Carlo-Pencil Beam dose algorithm for proton therapy inverse planning, Medical
636 physics **45**, 846–862 (2018).

[15] P. A. Taylor, S. F. Kry, and D. S. Followill, Pencil beam algorithms are unsuitable for proton dose calculations in lung, International Journal of Radiation Oncology* Biology* Physics **99**, 750–756 (2017).

[16] L. Lin, P. A. Taylor, J. Shen, J. Saini, M. Kang, C. B. Simone, J. D. Bradley, Z. Li, and Y. Xiao, NRG Oncology Survey of Monte Carlo Dose Calculation Use in US Proton Therapy Centers, International journal of particle therapy **8**, 73–81 (2021).

[17] H. Paganetti, Range uncertainties in proton therapy and the role of Monte Carlo simulations, Phys Med Biol **57**, R99–R117 (2012), [DOI:10.1088/0031-9155/57/11/R99] [PubMed:22571913].

[18] M. Yang, X. R. Zhu, P. C. Park, U. Titt, R. Mohan, G. Virshup, J. E. Clayton, and L. Dong, Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration, Physics in Medicine & Biology **57**, 4095 (2012).

[19] U. Schneider, E. Pedroni, and A. Lomax, The calibration of CT Hounsfield units for radiotherapy treatment planning, Physics in Medicine & Biology **41**, 111 (1996).

[20] B. Li, H. Lee, X. Duan, C. Shen, L. Zhou, X. Jia, and M. Yang, Comprehensive analysis of proton range uncertainties related to stopping-power-ratio estimation using dual-energy CT imaging, Physics in Medicine & Biology **62**, 7056 (2017).

[21] N. Peters, P. Wohlfahrt, C. Hofmann, C. Möhler, S. Menkel, M. Tschiche, M. Krause, E. G. Troost, W. Enghardt, and C. Richter, Reduction of clinical safety margins in proton therapy enabled by the clinical implementation of dual-energy CT for direct stopping-power prediction, Radiotherapy and Oncology **166**, 71–78 (2022).

[22] V. De Smet, R. Labarbe, F. Vander Stappen, B. Macq, and E. Sterpin, Reassessment of stopping power ratio uncertainties caused by mean excitation energies using a water-based formalism, Medical physics **45**, 3361–3370 (2018).

[23] ICRU, *Stopping powers and ranges for protons and alpha particles (ICRU Report 49)*, ICRU Publications, Oxford, United Kingdom, 1992.

[24] K. Souris, J. A. Lee, and E. Sterpin, Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures, Med. Phys. **43**, 1700–1712 (2016).

[25] J. Nocedal, Updating quasi-Newton matrices with limited storage, Mathematics of computation **35**, 773–782 (1980).

[26] P. Virtanen et al., SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods **17**, 261–272 (2020).

[27] E. Bär, P. Andreo, A. Lalonde, G. Royle, and H. Bouchard, Optimized I-values for use with the Bragg additivity rule and their impact on proton stopping power and range uncertainty, Physics in Medicine and Biology **63** (2018).

[28] W. Schneider, T. Bortfeld, and W. Schlegel, Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions, Physics in Medicine & Biology **45**, 459 (2000), [DOI:10.1088/0031-9155/45/2/314] [PubMed:10701515].

[29] H. Woodard and D. White, The composition of body tissues, The British journal of radiology **59**, 1209–1218 (1986), [DOI:10.1259/0007-1285-59-708-1209] [PubMed:3801800].

[30] D. White, H. Woodard, and S. Hammond, Average soft-tissue and bone models for use in radiation dosimetry, The British journal of radiology **60**, 907–913 (1987), [DOI:10.1259/0007-1285-60-717-907] [PubMed:3664185].

[31] *Key Data For Ionizing-Radiation Dosimetry: Measurement Standards And Applications (ICRU report 90)*, ICRU Publications, Oxford, United Kingdom, 2014.

[32] ICRU, *Stopping powers for electrons and positrons (ICRU Report 37)*, ICRU Publications, Oxford, United Kingdom, 1984.

[33] K. Souris, A. Barragan Montero, G. Janssens, D. Di Perri, E. Sterpin, and J. A. Lee, Monte Carlo methods to comprehensively evaluate the robustness of 4D treatments in proton therapy, Medical physics **46**, 4676–4684 (2019).

[34] E. Sterpin, S. T. Rivas, F. Van den Heuvel, B. George, J. A. Lee, and K. Souris, Development of robustness evaluation strategies for enabling statistically consistent reporting, Physics in Medicine & Biology **66**, 045002 (2021).

[35] Z. Perkó, S. R. Van der Voort, S. Van De Water, C. M. Hartman, M. Hoogeman, and D. Lathouwers, Fast and accurate sensitivity analysis of IMPT treatment plans using Polynomial Chaos Expansion, Physics in Medicine & Biology **61**, 4646 (2016).

[36] H. Wieser, C. Karger, N. Wahl, and M. Bangert, Impact of Gaussian uncertainty assumptions on probabilistic optimization in particle therapy, Physics in Medicine & Biology **65**, 145007 (2020).