IOP Publishing | Institute of Physics and Engineering in Medicine Physics in Medicine & Biology

Phys. Med. Biol. **59** (2014) 7341–7360 doi:10.1088/0031-9155/59/23/7341

FoCa: a modular treatment planning system for proton radiotherapy with research and educational purposes

D Sánchez-Parcerisa, M Kondrla, A Shaindlin and A Carabe

University of Pennsylvania, Department of Radiation Oncology, 3400 Civic Center Bvd, Philadelphia, PA, USA

E-mail: sanchezd@uphs.upenn.edu and a.carabe@uphs.upenn.edu

Received 9 July 2014, revised 18 September 2014
Accepted for publication 13 October 2014
Published 11 November 2014

Abstract

FoCa is an in-house modular treatment planning system, developed entirely in MATLAB, which includes forward dose calculation of proton radiotherapy plans in both active and passive modalities as well as a generic optimization suite for inverse treatment planning.

The software has a dual education and research purpose. From the educational point of view, it can be an invaluable teaching tool for educating medical physicists, showing the insights of a treatment planning system from a well-known and widely accessible software platform. From the research point of view, its current and potential uses range from the fast calculation of any physical, radiobiological or clinical quantity in a patient CT geometry, to the development of new treatment modalities not yet available in commercial treatment planning systems.

The physical models in FoCa were compared with the commissioning data from our institution and show an excellent agreement in depth dose distributions and longitudinal and transversal fluence profiles for both passive scattering and active scanning modalities. 3D dose distributions in phantom and patient geometries were compared with a commercial treatment planning system, yielding a gamma-index pass rate of above 94% (using FoCa's most accurate algorithm) for all cases considered.

Finally, the inverse treatment planning suite was used to produce the first prototype of intensity-modulated, passive-scattered proton therapy, using 13 passive scattering proton fields and multi-leaf modulation to produce a concave dose distribution on a cylindrical solid water phantom without any field-specific compensator.

0031-9155/14/237341+20$33.00 © 2014 Institute of Physics and Engineering in Medicine Printed in the UK & the USA 7341

Keywords: treatment planning, proton therapy, particle therapy, pencil beam, optimization, IMPT

(Some figures may appear in colour only in the online journal)

1. Introduction

Radiation therapy treatments are planned using software specifically designed for this purpose, with the capability of calculating 3D dose distributions on computed tomography (CT) images of the patient. Such treatment planning systems (TPS) must undergo a thorough calibration and commissioning process before they can be used in the clinics. Most centers use commercial TPS for their treatment planning needs. For the sake of robustness, and in order to protect the code from potentially unsafe alterations, these commercial TPS tend to have a closed architecture. Furthermore, due to intellectual property protection the source code of those systems is rarely disclosed. On the other hand, in-house created TPS, such as *Plan-UNC* (Tewell and Adams 2004), ASTROID (Clasie *et al* 2012) or TriP (Krämer *et al* 2000), are by definition flexible and extendable. They were developed to fill gaps where no commercial system was available, but they have become invaluable tools to boost the scientific production of the institutions where they were created.

With this aim in mind, FoCa (which stands for Forward Calculation) was created in 2013, looking for an open treatment planning system for proton radiotherapy that could handle both pencil beam scanning (PBS) and passive scattering ('double scattering', or DS) treatment modalities. FoCa can be used to (i) test out new experimental treatment modalities, currently not supported by existing clinical software, (ii) include radiobiological modeling into treatment planning streamlined process, and (iii) provide a means of training our medical physics students into the modules conforming a TPS, which is considerably harder to do with a closed-architecture system. FoCa will be extended to other radiation types (e.g. carbon ions, electromagnetic, etc.) in order to cover the wide spectrum of clinical modalities available in radiotherapy.

The objective of this work is to provide an insight of FoCa, as well as a description of its modules and capabilities. Section 2 describes the architectures of the dose calculation engine, capable of calculating proton doses in patient geometries for both DS and PBS proton beams, and of the optimization suite. Section 3 describes the commissioning results of the dose calculation engine, and an example implementation of the optimization suite to calculate optimal DS beams for intensity-modulated proton therapy based on passive scattering. At present, clinical DS treatment plans are created using forward calculation, so FoCa's capability of optimizing DS plans is an unprecedented feature that makes the software unique. Finally, section 4 comments on the obtained results obtained and describes the current status of the software and its prospective applications.

2. Description of FoCa

2.1. Architecture

We chose to develop FoCa using MATLAB[1] computing software, taking advantage of its fast prototyping capabilities, its speed at matrix calculations and its extensive geometry libraries. Moreover, recent versions of MATLAB include Object-Oriented (OO) capabilities (MATLAB R2014a 2014), which have allowed us to use a modular design.

[1] www.mathworks.com/products/matlab/.

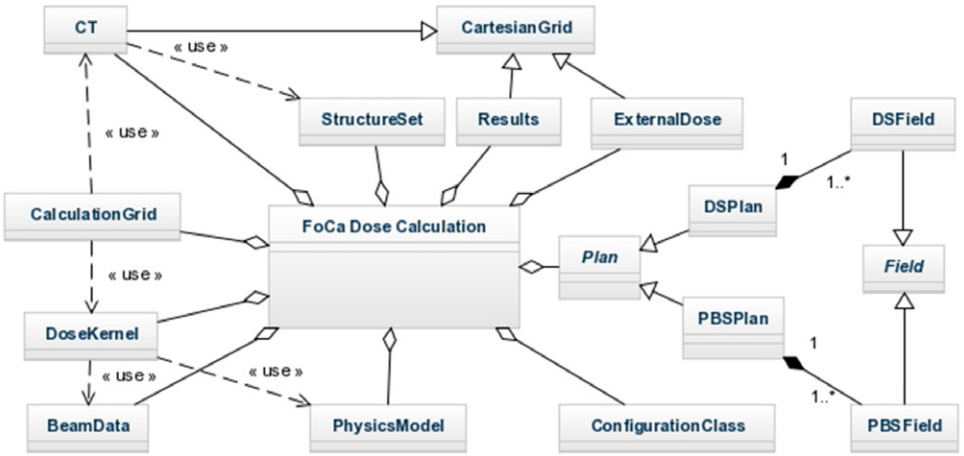

Figure 1. Class diagram for the FoCa dose calculation engine application.

The current implementation of FoCa includes two distinct applications: a forward calculation tool which produces dose distributions of proton plans in patient geometries (figure 1), and an inverse treatment planning framework which can be used to perform optimization of treatment plans based on different objective functions (figure 2). Modularity is the key to FoCa design: each of the components (physics data, CT processing, dose calculation, handling of data structures…) is decoupled from the rest, which allows the code to be extended or adapted to different situations by only overloading certain methods or by using inheritance between classes. The classes can also be combined differently to extend the application at the user's will. The following is a short description of the classes that constitute FoCa.

2.1.1. CartesianGrid class. Base class implementing a 3D Cartesian matrix in space with user-defined voxel sizes. Includes basic geometric and coordinate-conversion methods. It is used as a base class for *CT*, *Results* and *ExternalDose* classes.

2.1.2. CT class. Loads and stores a CT scan of a patient in DICOM format (NEMA PS3 2011). Contains a function to create test CTs of cubic water phantoms for testing and quality assurance purposes. Performs automatic conversion of Hounsfield units into water-equivalent path length (WEPL) (Schaffner and Pedroni 1998) in a voxel-by-voxel basis. Calculates the WEPL between two arbitrary points using the raytracing algorithm described by Krämer *et al* (2000).

2.1.3. StructureSet class. Loads and stores a set of structures in DICOM (RS) format. Contains functions to calculate areas and volumes of the regions of interest (ROIs), to determine whether an arbitrary point in space belongs to one or more ROIs, and to perform raytracing through the ROIs to determine, for example, the range along a given direction.

2.1.4. Configuration class. Loads and stores FoCa configuration files in ASCII format.

2.1.5. Plan class. Abstract class[2] including an array of *Field* objects and basic information about a plan. It has to be implemented in a subclass for each of the considered treatment

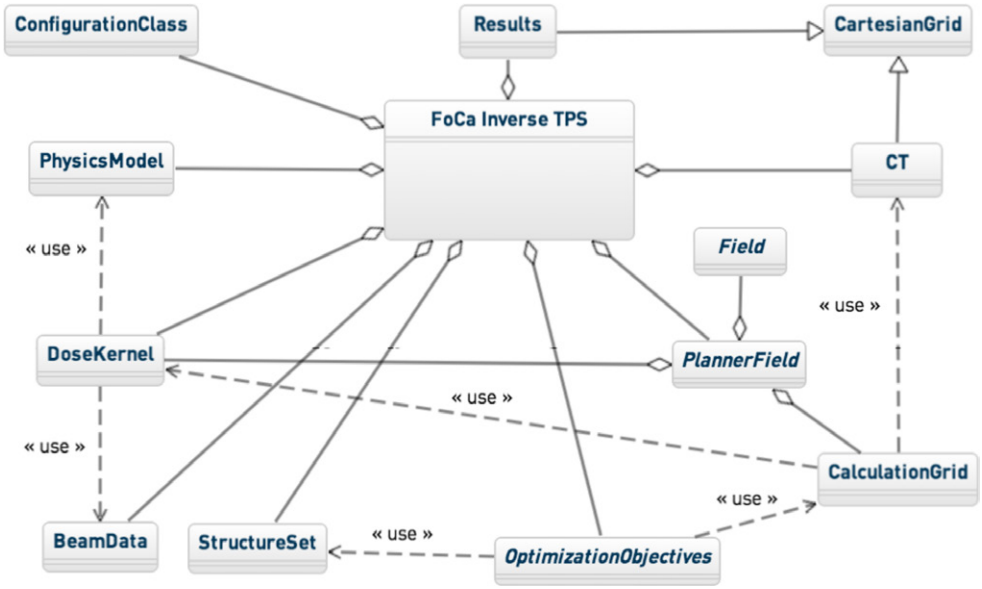

Figure 2. Class diagram for the FoCa inverse TPS framework application.

modalities (DS or PBS). The subclasses must be able to read and store DICOM files of the desired format. They must also implement a *GenerateCalculationGrid* method, which determines the optimal size and granularity of the calculation grid based on the characteristics of the plan.

2.1.6. Field class. Abstract class storing specific machine information for a given field, including patient positioning, isocentre position, couch and gantry angles, machine energy, collimators, etc. Subclasses contain specific DS or PBS attributes.

2.1.7. PhysicsModel, BeamData and DoseKernel classes. These three classes implement the calculation of a proton beamlet using the formulation of Ulmer and Schaffner (2011). The model uses a parameterized formula to describe the depth dose distribution of a proton beam in water, given its energy and range straggling. This longitudinal distribution is combined with a two Gaussian model for lateral scattering. Primary particles, secondaries and recoils are treated separately and combined into a final, single 3D dose beamlet. The beam-independent constants, equations and functions used in the beamlet model are stored in the *PhysicsModel* class. *BeamData* reads and stores all the treatment room-specific, patient-independent data such as beam spot sizes, nominal source-to-axis distances or beam energy spreads, whereas the *DoseKernel* class performs the calculations and stores the final dose kernel.

2.1.8. Calculation Grid class. The dose calculation in FoCa is based on the convolution-superposition algorithm described by Schaffner (2008) (see section 2.2). The *CalculationGrid* class contains overlapping grids storing different physical quantities (dose, fluence,

[2] In the object oriented design paradigm, an abstract class is an *incomplete* class that can only be instantiated by means of a derived class.

water-equivalent path length from origin) over the same geometric points. The class includes not only the data structures, but also the methods implementing the different steps of the convolution-superposition algorithm.

2.1.9. Results class. Stores a Cartesian 3D grid containing the final dose distribution in the patient. Unless otherwise stated, it is created with the same dimensions and voxel size as the CT. The class contains an interpolation method that integrates the dose distribution calculated in the *CalculationGrid* from each of the fields. This is done by iterating through all the voxels in the *Results* grid and filling each of them with the average value of all the neighboring points from the *CalculationGrid*, weighted by the inverse square of their distance to the voxel center. Matlab native method *ScatteredInterpolant* can be applied as well from the class.

2.1.10. External Dose class. Loads and stores a dose distribution matrix imported from a DICOM file. It contains methods for analysis and comparison of the dose distributions, such as the calculation of gamma indexes (Low *et al* 1998).

2.1.11. Optimization Objectives class. Part of the inverse treatment planning system. Stores the optimization objectives, including the relevant ROI, the desired lower and upper dose limits, and the weighting factor for each of the objectives.

2.1.12. Planner Field class. Part of the inverse treatment planning system. As an abstract class, it must be subclassed to implement a concrete optimization problem. It stores and manipulates all the data necessary to produce a treatment field: a *Field* object, a *DoseKernel* object and a *CalculationGrid*, as well as the functions to produce changes in the *Field* properties during iteration.

2.2. FoCa dose calculation engine

The FoCa Dose Calculation engine produces 3D dose distributions in patient geometries, using the formulation described in detail by Schaffner (2008). It is a beam-centered approach, where a divergent calculation grid is created for each of the fields in the plan, matching the beam's natural divergence, and the results are translated to the patient geometry only at the final interpolation step in the *Results* class (see figure 3). This calculation grid for a given field is constructed by selecting an initial plane, perpendicular to the beam, placed at the intersection point between the beam axis and the patient body (entrance point of the beam to the body). Then, a 2D reticule of points is created on this plane, becoming the first layer of the calculation grid. To construct the next layers, rays are traced between the virtual source of the beam and the points of the reticule, and calculation points are placed along the rays with a constant spacing of Δz. The point {I,J,K} of the calculation grid is obtained by translating the {I,J} point of the initial reticule, along the ray originated in the virtual source, a distance of (K − 1) times the longitudinal spacing Δz. A scheme of this procedure is depicted at figure 4(*a*).

FoCa contains two algorithms for dose calculation in the calculation grid, named *Robust* and *Fast*. The *Robust* algorithm uses the standard calculation grid as was just described, and is entirely based on the method detailed by Schaffner (2008). It consists mainly of four steps.

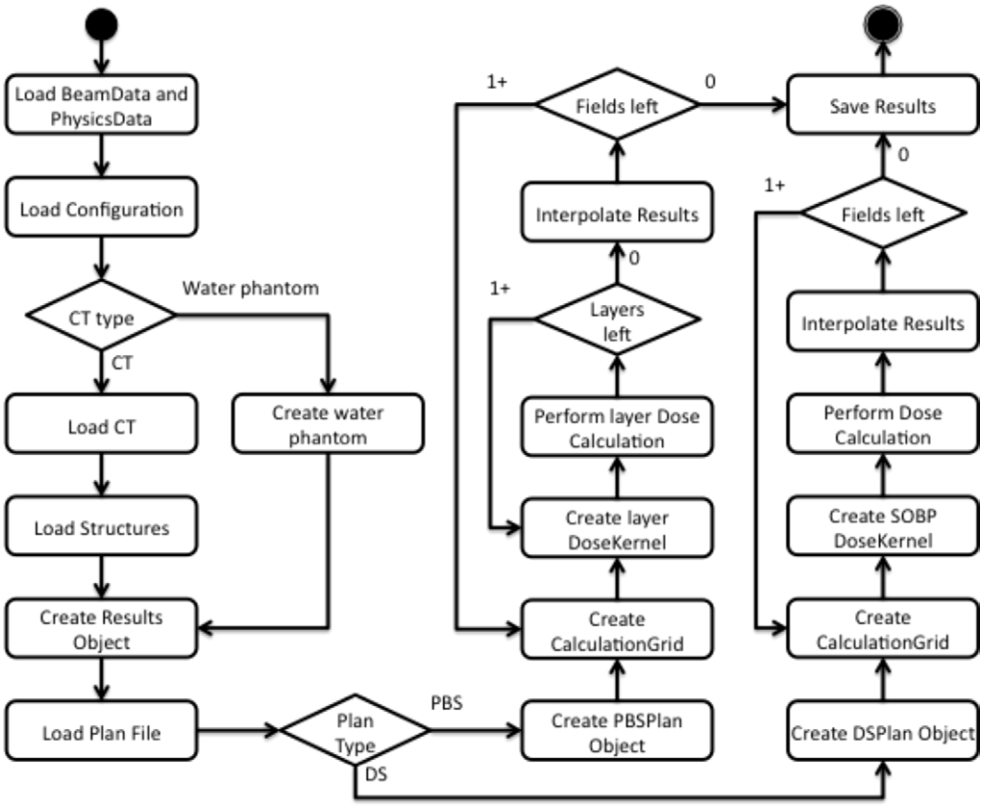

Figure 3. Flowchart diagram for a dose calculation with FoCa.

(a) Calculation of a dose kernel using the formulation of Ulmer and Schaffner (2011).
(b) Calculation of water-equivalent path lengths along each ray of the calculation grid.
(c) Calculation of in-air fluence for each layer of the calculation grid, from the energy and treatment nozzle configuration.
(d) Convolution-superposition of the fluence for each layer with the dose kernel, scaled with the water-equivalent path length at each point of the calculation grid.

The *Fast* algorithm is a variant of it, designed to speed up the dose calculation process to integrate it into the inverse treatment planning system. It is based on a modified calculation grid, as shown in figure 4(*b*). For the standard calculation grid, the geometrical distance between consecutive points of the same ray is always Δz, but the WEPL between consecutive points varies depending on the point of the CT where they lie. This makes it necessary (in step 4 of the *Robust* algorithm) to scale the dose kernel with the WEPL for each point of the calculation grid, which can be a time-intensive task. Therefore, the *Fast* algorithm modifies the calculation grid before the calculation starts, so that the WEPL between consecutive points along the same ray is always equal to Δz, causing the points to be distributed unevenly along the ray. For instance, for high-density tissues such as bone, the calculation grid points will be very close together, whereas for low-density tissues such as lung, the calculation grid points will be further apart from each other. This modification in the calculation grid allows the convolution-superposition of in-air fluence with the dose kernel to be performed faster, since

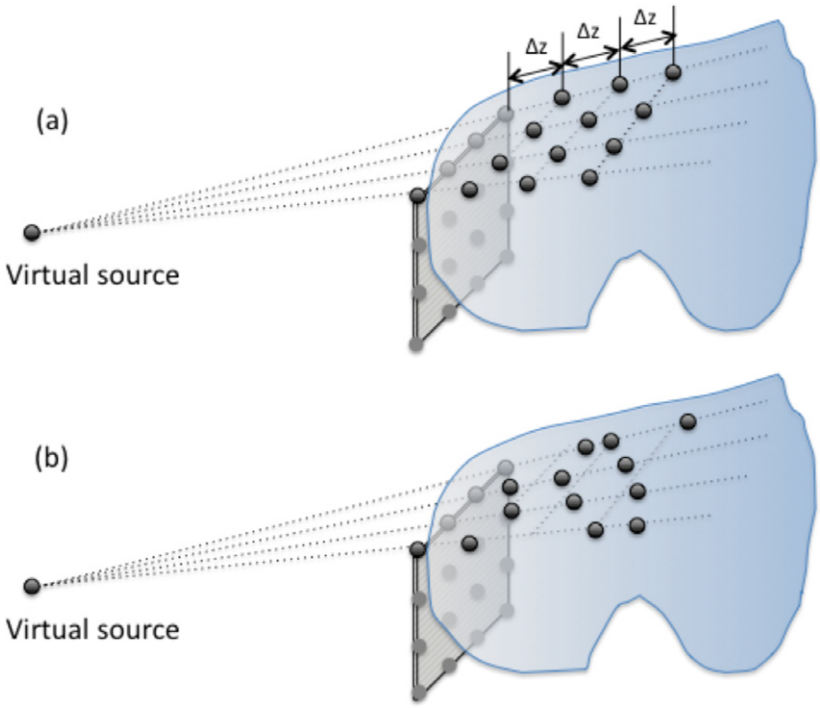

Figure 4. (*a*) Diagram of the standard calculation grid and from the *Robust* FoCa algorithm, for which the calculation points along the same ray are separated by a constant distance, and (*b*) diagram of the adapted calculation grid for the *Fast* FoCa algorithm, for which the calculation points along the same ray are separated by a constant WEPL.

the dose kernel does not need to be re-scaled with the WEPL, at the cost of a less accurate modelling of the lateral scattering. The main steps of the *Fast* algorithm are:

(a) Calculation of a dose kernel using the formulation of Ulmer and Schaffner (2011).
(b) Adaptation of the calculation grid to the patient CT.
(c) Calculation of the lateral distribution of in-air fluence in the first layer of the calculation grid.
(d) Calculation of the longitudinal distribution of in-air fluence, for all points in the calculation grid.
(e) Convolution of the first layer of the calculation grid with the dose-kernel and superposition with the longitudinal fluence distribution.

2.3. FoCa inverse treatment planning suite

The FoCa Inverse TP suite has been created with an open architecture, with the aim of creating a framework for research in protontherapy treatment planning that can adapt any inverse optimization algorithm/method according to the users' needs. As a consequence of this, it is not specific to any treatment modality in particular. Currently the optimization classes implemented (whose flowchart is depicted in figure 5) are loosely based on the Very Fast Adaptive Reannealing method (Ingber 1989), but these classes can be modified to implement any other optimization strategy. Such a software architecture based on Object Oriented programming,

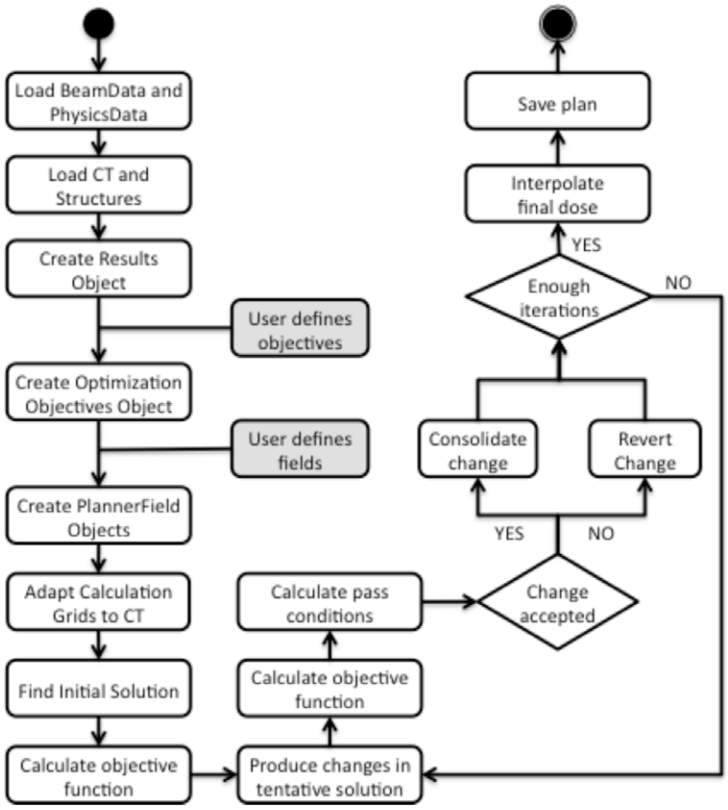

Figure 5. Flowchart diagram for a generic plan optimization with FoCa inverse TPS.

with a number of classes implementing the core of the system and the users taking care of the specifics of their problem, was deployed with great success in the Monte Carlo transport code GEANT4 (Agostinelli *et al* 2003), which is widely used in a number of fields in Physics, including medical physics.

The optimization process starts by loading the patient CT and structure set. Then, the user defines the optimization objectives and creates the fields to be optimized, described by their orientation (gantry and couch angles), isocentre position and snout elevation. From each of these fields, the TPS creates *PlannerField* objects, which will in turn create *CalculationGrid* objects and perform their adaptation to the CT geometry, as required for the *Fast* calculation method of FoCa.

In order to use the FoCa Inverse TPS suite, users must define a series of functions:

- **Cost function**. The function summarizes the difference between a given dose distribution and the required optimization goals into a single number, so the optimization problem can be reduced to minimizing the value of the cost function. A template *calculateOF()* method is provided in the main class. The calculation of the cost function must be fast enough so that it can be included in the iteration process.
- **Subclass of *PlannerField* class**. It must include a definition of the objective plan modality, the plan parameters to be optimized, ranges of variation of the parameters and dose calculation function for the given plan. It can also implement a persistence option, allowing users to export the created plans into DICOM format.

- **Calculation of initial solution**. A *findInitialSolution(…)* function, called from the constructor of the *PlannerField* class, must be implemented. Different variants of this function can coexist at the same time, which allows users to test the sensitivity of the final plan to the initial solution.
- **Temperature scheme**. The concept of 'temperature' in optimization, directly translated from thermodynamics, describes the excitation state of a system and its ability to reach lower energy (cost) states by 'jumping' over potential barriers. In practice, this implies that in a system with a high temperature, changes that result in an increase of the cost function might be accepted, but as the system cools down, the probability of accepting a change that increases the cost function tends to zero. Different temperature variation schemes exist in the literature (Ingber 1996), and the use of one or another affects how fast the optimizer converges to an optimal solution. A constant temperature scheme can be used to implement non-annealing based optimization algorithms.
- **Iteration process**. The method *Iterate()* from the main class must be implemented by the user, defining the variation of the parameters at each iteration. This variation is tightly tied to the temperature scheme: hotter systems will, in general, require larger changes, where cooler systems will require smaller, *fine-tuning* type modifications. It is often advisable to define two or more distinct phases in the optimization phase, in order to minimize the total calculation time. Users must also define the number of changes per iteration and how they are distributed among the different parameters under optimization.
- **Restrictions to the solution**. A Boolean function *meetsRestrictions()* is called after every iteration to ensure that the plans never reach undeliverable states. The users must define that function according to the needs of their system.
- **Termination conditions**. Another Boolean function *terminateOptimization()* is called after each iteration and before starting a new one. The iteration process can be terminated when an acceptable solution is found (below a certain value of the cost function), where successive iterations do not produce improvements in the cost functions, or when the program reaches a certain maximum number of iterations.

This general architecture is in principle valid for both DS and PBS modalities, since the forward dose engine is ready for the two of them. It is up to the users, when defining the *PlannerField* class in a specific implementation, to decide which of the modalities should be used for the planning. It would even be possible to optimize combined radiation fields by defining a dual *PlannerField* class.

3. Results

3.1. Validation of forward dose calculation tool

Software testing can be performed at different levels. In the validation of our tool we considered system tests, which evaluate the performance of the system as a whole from the user's perspective; and unit tests, which evaluate each of its modules independently. During the testing phase of the FoCa treatment planning system, a test plan and a test report based on the IEEE standard 829–2008 (IEEE 2008) were produced, including unit tests for each of the FoCa classes. These tests were not intended to assess whether the solutions produced by the code were correct (i.e. the dose distributions in the patient matching experimental data, or the optimized plans fulfilling the required objectives), but to reveal possible programming bugs and to ensure that the code was performing as intended. Matlab Test Suite (Matlab.unittest.

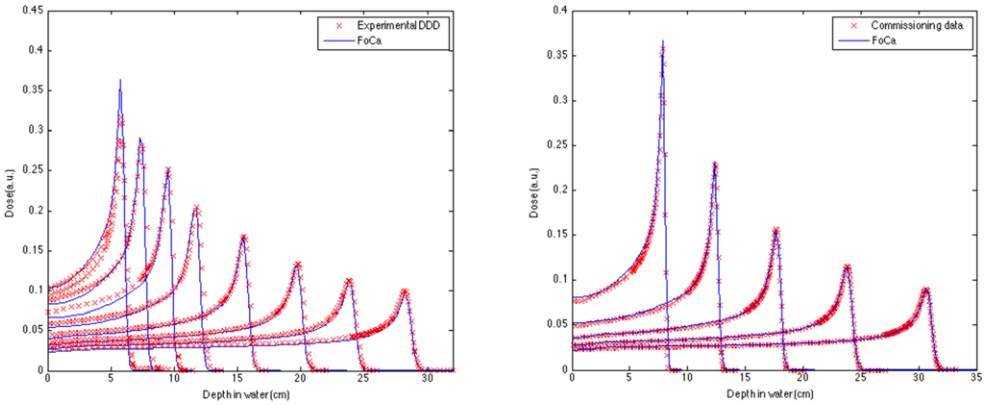

Figure 6. FoCa depth dose distributions for different ranges, in DS mode (left, one peak for each scatterer option 1–8) and PBS mode (right, energies 100–220 MeV in steps of 30 MeV), compared to commissioning data, for the same energy and machine configuration. Commissioning profiles are measured with small ionization chamber at the central axis in the case of DS, and with a broad plane-parallel chamber for PBS. The $1/r^2$ fluence loss effect correction is included where necessary.

TestCase classes) was used to automatize all unit tests and ensure that no new release breaks the previously commissioned results.

In contrast, the validation of the dose calculation was done according to the relevant sections of AAPM TG-53 (Fraass *et al* 1998), but only to a limited extent, since FoCa is not to be used clinically. The results of the dose calculation were compared with experimental data from commissioning where available, and with 3D dose distributions obtained with our commercial TPS software Eclipse (Varian Medical Systems, Palo Alto, CA, USA). Because the validity of the dose calculation tool depends on the concrete implementation of the beam and machine models, the scripts used to produce an automatic comparison of the calculation results with test data are considered part of FoCa. For the sake of brevity, only a selection of the test data is presented here:

3.1.1. Validation of ranges and depth dose profiles against commissioning data. In the formulation of Ulmer and Schaffner (2011), the range of protons depends on their energy by a simple exponential relation, $R = AE^p$. We have derived the coefficients A and p from the commissioning measurements of our treatment planning system, yielding values of $A = 0.00317$ and $p = 1.702$ for DS mode, and $A = 0.00290$ and $p = 1.719$ for PBS mode. With these values, the agreement between FoCa and the commissioning data is excellent, not only in the ranges but also in the shapes of the depth profiles. Figure 6 shows a comparison for both DS (left) and PBS modes (right).

3.1.2. Homogeneity of spread-out Bragg peaks. A number of Spread-out Bragg Peaks (SOBPs) were analyzed for both DS and PBS calculation modes (figure 7), in order to assess the field homogeneity in the target. Using the calculation weights given by Eclipse for the SOBPs, the dose in the 1D peak was within 1% of prescribed for DS peaks and within 3% of prescribed for PBS peaks, and in both cases, very similar to the profiles calculated with Eclipse.

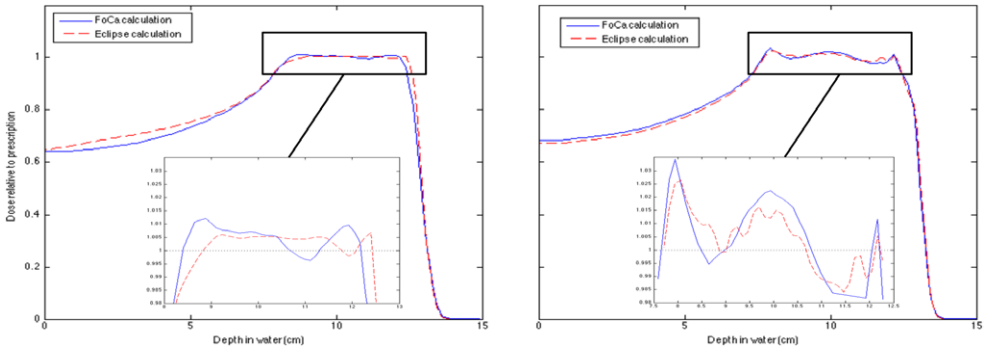

Figure 7. Homogeneity analysis of two SOBPs in water, one in DS mode (left) and one in PBS mode (right), for FoCa and Eclipse calculations.

3.1.3. Validation of fluence profiles against commissioning data. In the proton convolution-superposition algorithm, the fluence calculation encapsulates the dependence of the final dose distribution from the treatment unit configuration. The fluence obtained from the calculation grid of FoCa was compared against experimental in-air profiles for both DS and PBS modalities (Ainsley *et al* 2013, Lin *et al* 2013a, 2014).

In the case of DS, transversal open-field profiles were created for all eight options[3] available in our institution using an MLC collimated square field of 12.6 cm (except for option 8, which required a smaller field of 9.2 cm). Longitudinal fluence profiles, affected by the $1/r^2$ fluence loss effect, were also extracted from FoCa and compared against commissioning data for all eight options. The results of these comparisons, showing a fair agreement, are displayed in figures 8 and 9. Transversal profiles are slightly different for Robust and Fast algorithms (figure 8), with Robust calculation showing a better agreement. Longitudinal profiles (figure 9) are the same for both methods.

For PBS modality the comparison was more straightforward, simply because the beam spot sizes obtained during commissioning (Lin *et al* 2013a, 2014) were directly used in the bi-Gaussian beamlet model of FoCa, so an excellent agreement is expected. The transversal fluence distributions in both X and Y directions were compared with the commissioning data for several equispaced energies, for transversal planes at various depths. Figure 10 shows these spot profile comparison for two representative energies. Due to its limited influence in dose calculation, well below the accuracy of the FoCa dose models, no halo effects (Lin *et al* 2013b) are considered in the current FoCa beam model.

3.1.4. Comparison of dose distributions with commercial TPS in phantom and patient CTs. In order to assess the accuracy of the FoCa dose calculation engine, we compared its output with our commercial TPS. This comparison itself cannot be considered part of the validation of a TPS, since the 3D dose distribution obtained with another TPS is not free of uncertainties either, but the cross comparison can indeed give valuable information on the performance of our system.

On a first step, we created single-field DS and PBS plans in a cylindrical solid water phantom with a 10 cm radius that contained a spherical target of radius of 2.5 cm, with both structures centered at the isocentre. The resulting dose distributions in the phantom were calculated

[3] The treatment unit can be configured according to different options (selection from a set of beamline scatterers and other parameters) for treatment.

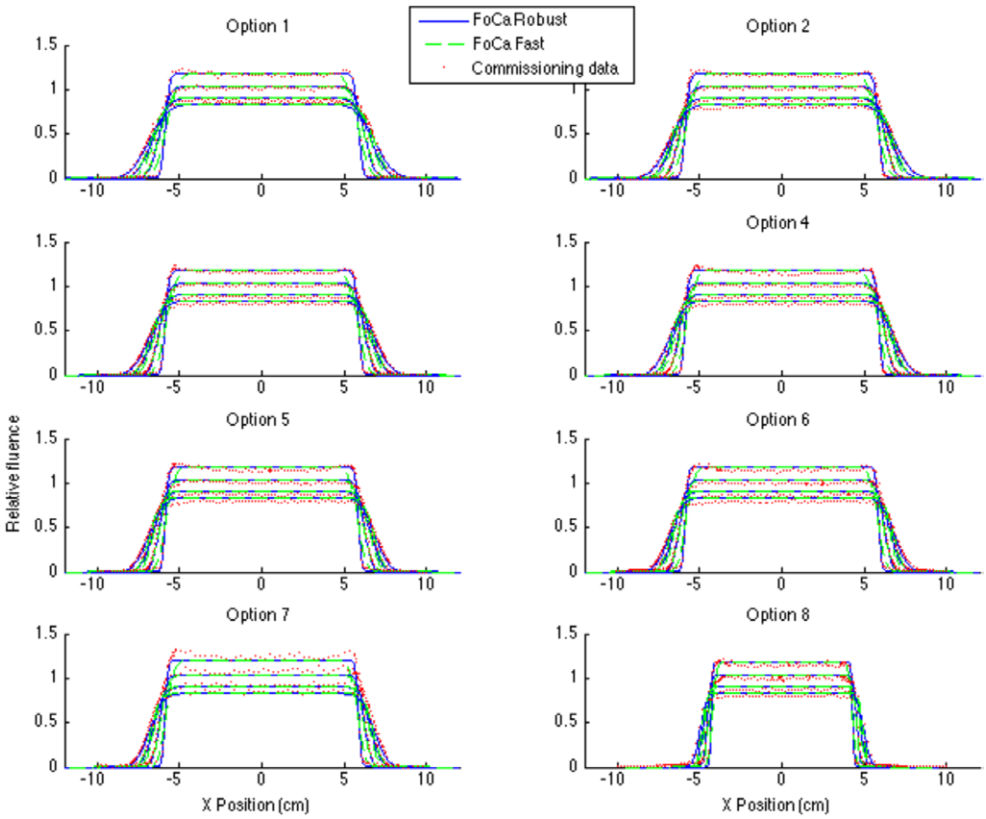

Figure 8. Transversal in-air fluence profiles for DS fields at isocentre plane and at positions +15 cm, –15 cm and –25 cm, where the positive axis pointed towards the treatment nozzle. MLC aperture sizes at isocentre plane projection are 12.6 cm for all options, except for option 8, which is 9.2 cm. Fluences are normalized to 1 at the center of the transversal plane containing the isocentre, so the curves at +15 cm have relative fluences >1 in the central plane, and the curves at –15 cm and –25 cm have central plane relative fluences <1, with the minimum central plane fluence corresponding to the –25 cm curve.

with FoCa using both the *Robust* and *Fast* algorithms, and compared with the dose distributions obtained with our commercial TPS. Because the current implementation of FoCa does not include MU calibration, the comparison was done in terms of relative dose. Dose distributions were normalized so that the mean dose in the target was at 100% of the prescribed dose. Gamma index analysis (for 3% dose and a distance-to-agreement (DTA) of 2 voxels, approximately 5 mm) was performed on the dose matrices for the voxels with at least a 0.5% of the prescription dose. Additionally, we performed the same comparison for three clinical treatment plans corresponding to a brain tumor case. Out of these three plans, two of them were PBS plans (one with multi-field optimization (or Intensity Modulated Proton Therapy, IMPT) and the other one with single-field optimization (or Single Field Uniform Dose, SFUD), and the remaining one was a DS plan without field-specific compensators. All of them were created with the same constraints as the original plan.

The results of these comparisons are shown in table 1 and partially in figures 11 and 12. In general, the pass rates are quite high, with some minor differences (up to 2%) between

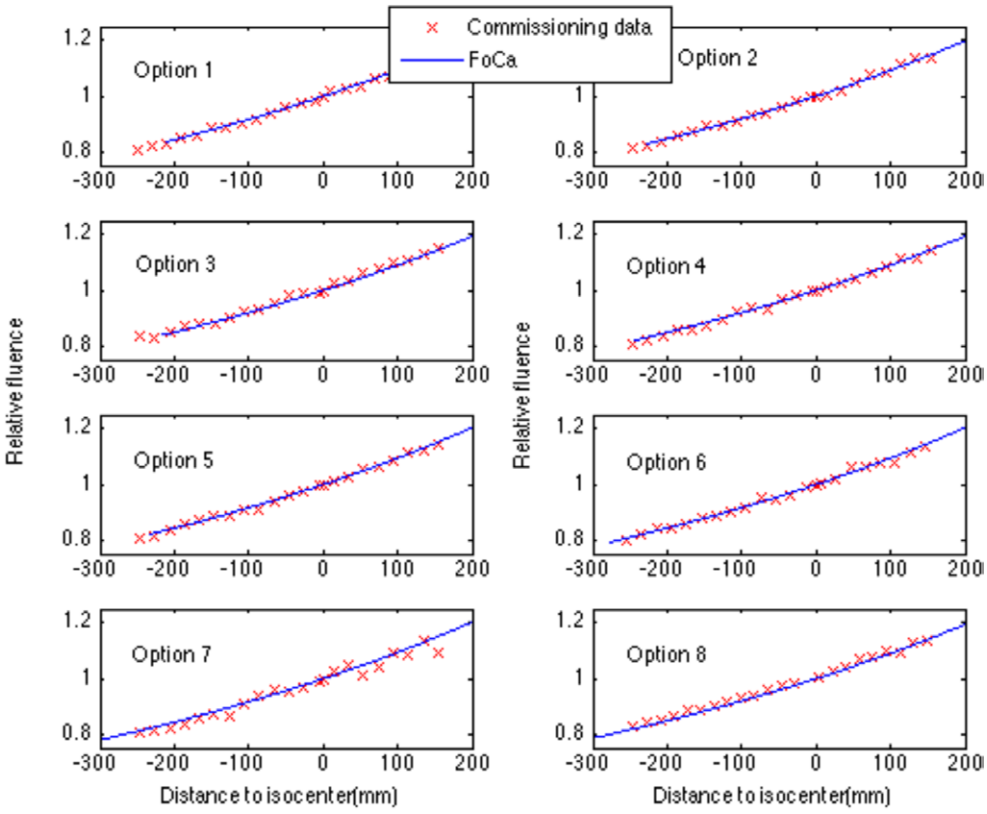

Figure 9. Longitudinal in-air fluence profiles for DS fields, measured at the central axis. The distance to isocentre is positive towards the treatment nozzle. MLC aperture sizes at isocentre plane projection are 12.6 cm for all options, except for option 8, which is 9.2 cm. Fluences are normalized to a value of 1 at isocentre for each option.

the *Robust* (left column) and the *Fast* (right column) dose distributions. The main reasons why the pass rates are not 100% are the inability of the simplified longitudinal dose model to fully describe the longitudinal dose profile of the spread-out Bragg peaks used in the DS mode, as seen clearly in figure 7 and as discussed by Ainsley *et al* (2013), and an insufficiently accurate lateral dose model for the proton kernel in the case of PBS (Lin *et al* 2013b). The effect of the latter is much less pronounced, as reflected in the fact that the PBS calculation for the phantom has a higher pass rate than the DS one. Another source of uncertainty is the interpolation from the divergent dose grid into the Cartesian grid. In our calculations, the global gamma-index pass rates are consistent (within 1%) between the in-house and the Matlab-native dose interpolation methods. The best passing rate is achieved at the patient DS plan, but this is a consequence of the gamma index calculation method: since only the voxels with at least 0.5% of the prescribed dose (normalized so that the mean dose to the target is at 100%) are taken into account for the pass rate, and the discrepancy between the commercial TPS and FoCa is higher in the low-dose areas (penumbrae), it is logical that the patient DS plans, containing three fields, has its penumbra spread across three different volumes, thus lowering the number of potentially non-passing voxels that make it into the 0.5% cut. Finally, it is worth mentioning that the extra level of

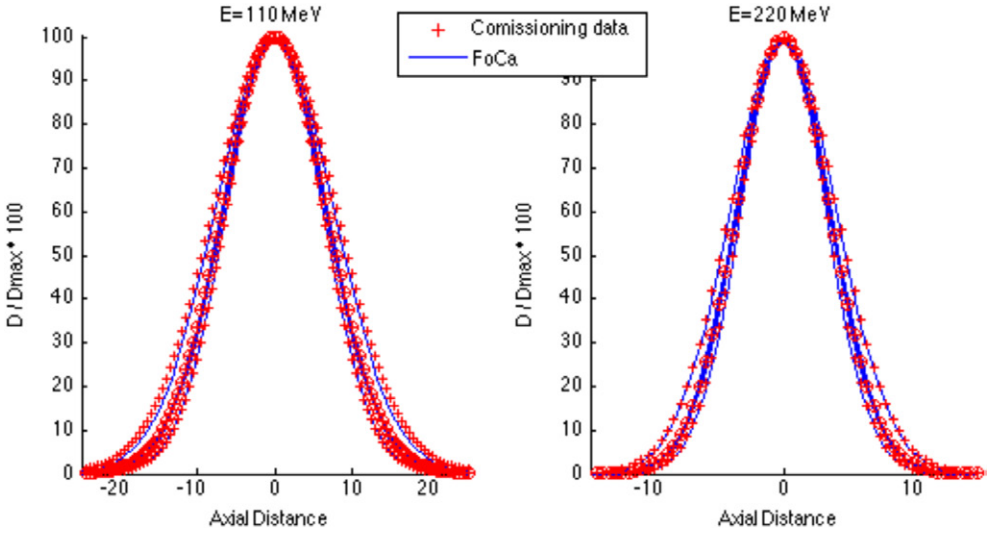

Figure 10. In-air spot profiles in the *X* direction for PBS mode for $E = 110$ MeV (left) and for $E = 220$ MeV (right), for transversal planes at distances +20 cm, 0, and –30 cm from isocentre, where positive distances are directed towards the treatment nozzle. The thicker line and rounded markers indicate the profile at isocentre. The curves above and below correspond to the –30 cm and +20 cm positions, respectively.

Table 1. Gamma-index pass rates for voxels with at least 0.5% of the prescribed dose, for both calculation algorithms and five different treatment plans. Dose interpolation performed with Matlab native *ScatteredInterpolant* method.

Plan	3%, 2 voxel DTA gamma index pass rate (Robust algorithm)	3%, 2 voxel DTA gamma index pass rate (Fast algorithm)
Phantom, DS	94.3%	92.9%
Phantom, PBS	98.7%	98.3%
Patient, DS	99.2%	98.1%
Patient, PBS with SFUD	96.6%	97.4%
Patient, PBS with IMPT	96.7%	97.6%

inhomogeneity found in IMPT plans compared to SFUD plans does not seem to affect the accuracy of the calculations, at least for the case considered.

3.2. Using the inverse treatment planning suite

As an example of the use of the FoCa inverse TPS platform, we implemented a basic DS inverse TPS allowing for multi-field optimization (IMPT). This implementation simply intends to show the capabilities of the framework; however, this work is ongoing and it is in the scope in our development to produce a fully-working DS inverse TPS, able to produce optimal, compensator-free plans.

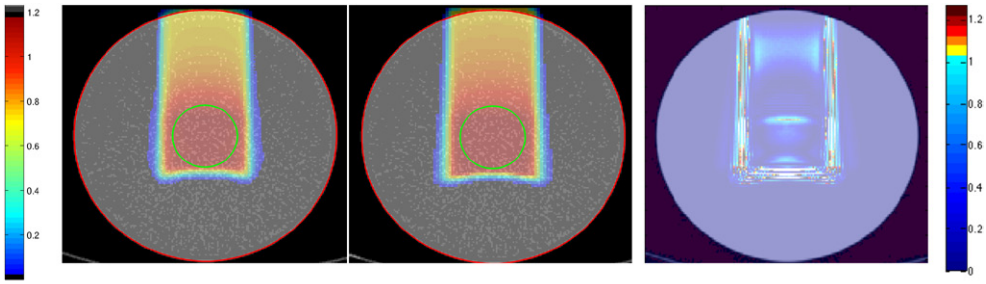

Figure 11. Dose comparison at the *Y* plane containing the isocentre for phantom single-field DS plan using the *Robust* algorithm of FoCa. Colorbar in dose plots indicates dose relative to prescription dose. FoCa dose is normalized so that the mean dose at the target corresponds to a value of 100% of the prescribed dose. Colorbar in gamma-index plot indicates gamma index with 3% dose and DTA = 2 voxels, approximately 5 mm.

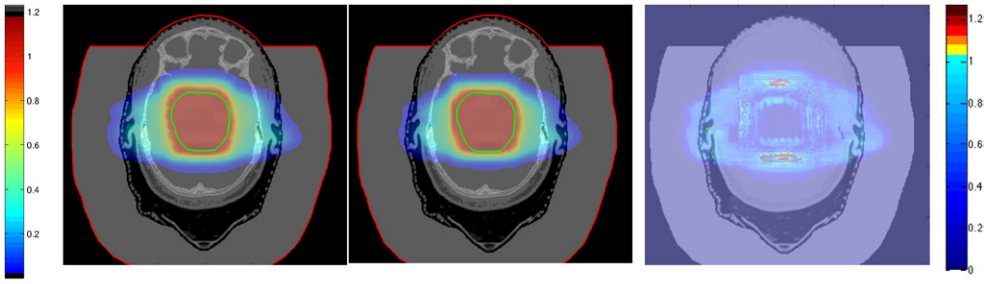

Figure 12. Dose comparisons at the *Y* plane containing the isocentre for a 2-field (SFUD) PBS plan on a patient CT using the *Fast* algorithm of FoCa. Colorbar in dose plots indicates dose relative to prescription dose. FoCa dose is normalized so that the mean dose at the target corresponds to a value of 100% of the prescribed dose. Colorbar in gamma-index plot indicates gamma index with 3% dose and DTA = 2 voxels, approximately 5 mm.

The user-defined functions for the DS inverse TPS prototype have been implemented as follows. The cost function is calculated as the sum of three parts. The first part is a function of the difference between the average dose in the target and the prescribed dose, the second part is a function of the maximum dose in the target (to avoid hot spots), and the main part is the sum of the quadratic dose difference for all structures with a defined objective, weighted by the importance of the objective. The weight that each of this parts take in the final value of the objective function is a parameter of the optimization process.

The *PlannerField* class was subclassed from the FoCa *DSField* class, which stores the position of a treatment nozzle for a given pair of couch and gantry angles at a given distance from the isocentre. The key parameters to be varied in each iteration are the position of each of the leaves in the multi-leaf collimator, the range and modulation of the SOBPs and the relative weights of the different fields. The initial position of the MLC is conformal to the target, with a given lateral margin. The initial range and modulation are derived from the position of

the target relative to the central axis. A database of dose kernels was precalculated in order to allow for fast variations of the field range and modulation during optimization.

A temperature scheme with exponential temperature decay was implemented, where the temperature at the k-th iteration was given by

$$T(k) = T_0 \mathbf{exp}\,(-\gamma k)\,,$$

where γ and T_0 are initial parameters of the calculation, and the acceptance probability of a change with an increase of the cost function ΔE was given by

$$P = \frac{1}{1 + \mathbf{exp}\left(\frac{\Delta E}{T}\right)}\,.$$

The γ temperature parameter also controlled the evolution of the iterations. For each optimizable parameter A_i (be it leaf position, range/modulation or relative weight), the change in each iteration was sampled from a Gaussian distribution with a standard deviation given by

$$\sigma_i = \sigma_i^{\mathbf{min}} + \left(\sigma_i^{\mathbf{max}} - \sigma_i^{\mathbf{min}}\right)\ \mathbf{exp}\,(-\gamma k)\,.$$

Each iteration comprised either one change in the relative weights, one change in the range or modulation of one field, or a finite number of changes (adjustable) on the position of the MLC leaves. The changes violating machine constraints were automatically rejected. The optimizer was run for a fixed number of iterations, the simplest possible implementation of the *terminateOptimization()* function.

The described software was used to create a sample plan, where 13 fields placed every 15° were used to shape the dose on the target on a cylindrical phantom, similar to the one described by Mayo and Urie (2003). The target volume was defined as a cylinder shell with internal radius of 2.5 cm and external radius of 4.5 cm, while the organs at risk (OARs) were an inner cylinder with a radius of 2 cm, and an outer shell with internal and external radii equal to 5 cm and 6 cm, respectively. We chose a concave target (which is not treatable with conventional DS fields) to show the potential of the technique. The acceptance criterion was 98% of the target volume covered by 95% of the dose, with a maximum dose of 107%, and the objective was to minimize the dose to the OARs as much as possible after achieving the target coverage and maximum dose criteria.

The final plan produced by the optimization (after 10 000 iterations) is shown in figure 13, and the resulting dose distribution and dose-volume histogram are displayed in figure 14. The optimizer was configured to minimize the mean dose to the OARs while keeping a good coverage of the target. The plan produced by the prototype TPS achieved a conformity index (defined as V_{95}/V_{PTV}) of 1.34, with the mean doses to the OARs below 78% of the prescribed target dose.

4. Discussion and conclusions

The presented platform is capable of producing accurate dose distributions in patient geometries from existing proton radiotherapy plans, and to implement optimization problems in different modalities of protontherapy. The physical models supporting it, based on well-established algorithms, reproduce faithfully the basic commissioning data from our institution. We still see room for improvement in the accuracy of the current dose calculation algorithms, and the enhancement of the existing models (or the development of new ones) will undoubtedly be favoured by the automatic testing scripts present in the code.

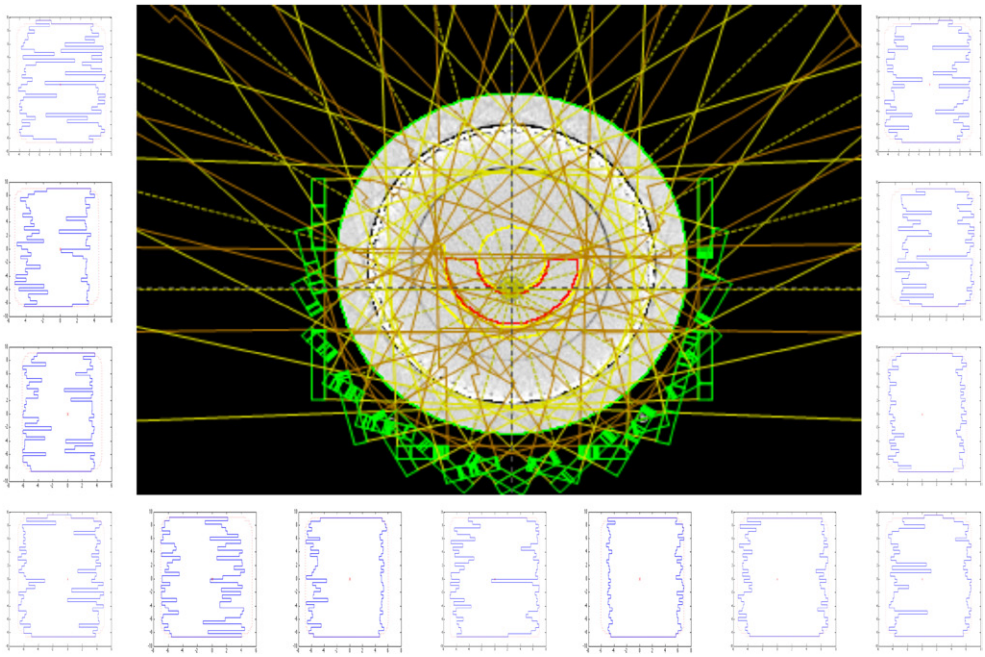

Figure 13. Field positions and optimized MLC apertures for a 13-field DS-IMPT plan on a cylindrical phantom.

The computational times were similar for DS and PBS modalities and ranged between 1–2 min per field in *Fast* mode and 2–3 min per field in *Robust* mode. More than 90% of that time was used in calculating the WEPL for all rays in the divergent grid, and in interpolating the dose from the calculation grid into the Cartesian grid. The non-precalculable, field-position specific time (i.e. the time used to calculate fluence and dose in the divergent grid for a specific field configuration, given the WEPL for all rays) was significantly different for *Robust* (~10 s per field) and *Fast* (~0.1 s per field), which is the reason why the *Fast* method is the one used in the inverse treatment planning system, allowing for complete iterations in under 1 s.

The optimization example presented in section 3.2, along with the wide experience in IMRT, clearly suggest that it is certainly feasible to produce IMPT fields with MLC passive scattering fields. Hypotheses of better dose conformity or reduced dose to the OARs when using IMPT-DS compared with traditional DS plans or even IMRT or VMAT plans can now be tested by using the treatment planning system prototype presented here.

FoCa is currently available within our institution to a selected number of test users (both for educational and research use), but we are currently studying formulas that allow us to make it available to the community once the desired level of maturity is achieved. Moreover, the fact that we use a ubiquous tool such as MATLAB opens the door for integration with other useful MATLAB programs from the radiation oncology world, such as CERR (Deasy *et al* 2003) or HART (Pyakuryal *et al* 2010).

As an ongoing project, the development of FoCa is not finished. Features to be included in the next versions of the code include a full implementation of a PBS inverse TPS, support for patient-specific compensators in the DS mode to account for the full flexibility of the passive

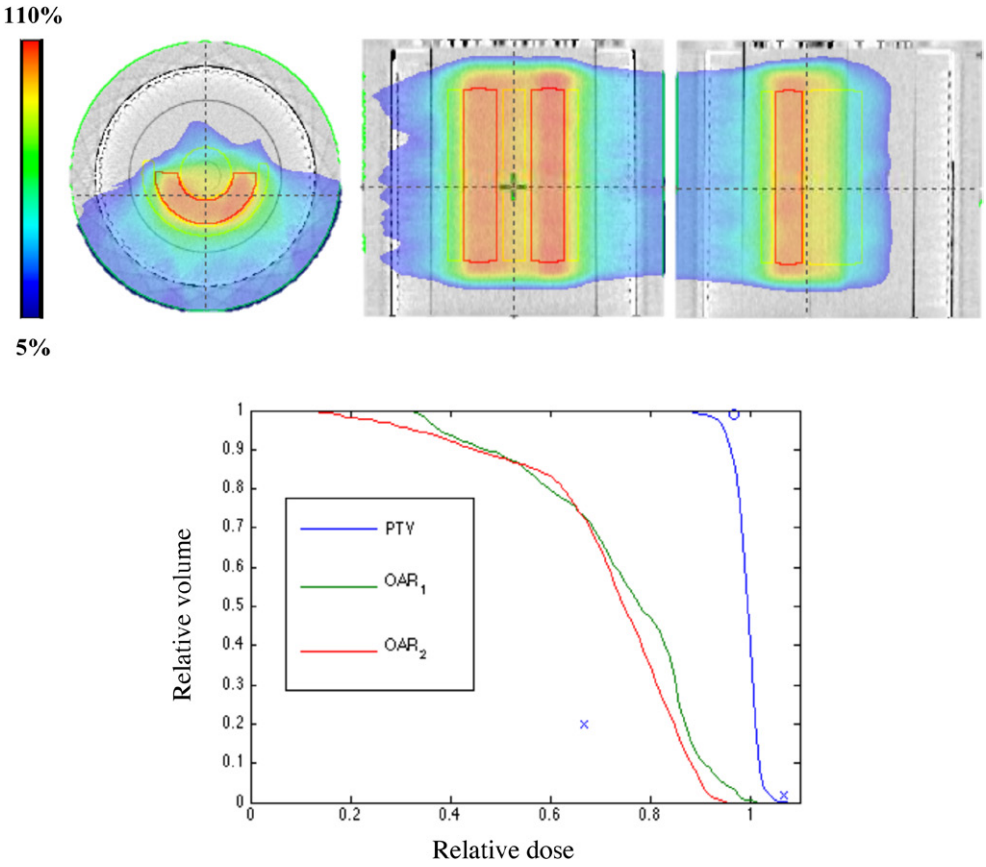

Figure 14. Above: transversal, frontal, and sagital views of the dose distribution produced by the intensity-modulated 13-field proton double-scattering plan on the cylindrical phantom. Below: dose-volume histogram of the shown dose distribution.

scattering technique, extensibility to other radiation types, support for absolute dosimetry and evaluation of plan robustness.

FoCa is currently being used in a number of research projects, including the development of analytical calculations of 3D linear-energy transfer distributions in proton plans based on patient geometries (Wilkens and Oelfke 2003 2004), development and implementation of radiobiological models in TPS-like calculations (Carabe *et al* 2012 2013), inverse planning of DS plans with non-flat SOBPs (Sánchez Parcerisa *et al* 2014) and LET-based inverse treatment planning (Giantsoudi *et al* 2013).

Acknowledgments

The institutional commissioning data used in FoCa and in this article is the result of a collective effort led by Liyong Lin and Chris Ainsley, whom we also thank for their meaningful advice. We also thank Miguel A Cortés-Giraldo for valuable discussions. The authors want to acknowledge minor contributions of Chris Wang at diverse stages of the code development. MATLAB and Eclipse are registered trademarks respectively owned by The Mathworks Inc. and Varian Medical Systems Inc.

References

Agostinelli S *et al* 2003 GEANT4—a simulation toolkit *Nucl. Instrum. Methods Phys. Res.* A **506** 250–303

Ainsley C *et al* 2013 Optimization of the modelling of longitudinal dose distributions for double-scattered proton beams in a commercially-available treatment planning system *Phys. Med. Biol.* **58** N145–55

Carabe A *et al* 2012 Range uncertainty in proton therapy due to variable biological effectiveness *Phys. Med. Biol.* **57** 1159

Carabe A *et al* 2013 Clinical consequences of relative biological effectiveness variations in proton radiotherapy of the prostate, brain and liver *Phys. Med. Biol.* **58** 2103

Clasie B *et al* 2012 Golden beam data for proton pencil-beam scanning *Phys. Med. Biol.* **57** 1147–58

Deasy J O, Blanco A I and Clark V H 2003 CERR: a computational environment for radiotherapy research *Med. Phys.* **30** 979–85

Fraass B *et al* 1998 American Association of Physicists in Medicine Radiation Therapy Committee Task Group 53: quality assurance for clinical radiotherapy treatment planning *Med. Phys.* **25** 1773–829

Giantsoudi D *et al* 2013 Linear energy transfer-guided optimization in intensity modulated proton therapy: feasibility study and clinical potential *Int. J. Radiat. Oncol. Biol. Phys.* **87** 216–22

IEEE 2008 *The Institute of Electrical and Electronics Engineers Standard 829: Standard for Software Test Documentation* (Washington, DC: IEEE Computer Society)

Ingber L 1989 Very fast simulated re-annealing *Math. Comput. Model.* **12** 967–73

Ingber L 1996 Adaptive simulated annealing (ASA): lessons learned *Contr. Cybern.* **25** 33–54

Jäkel O *et al* 2001 Treatment planning for heavy ion radiotherapy: clinical implementation and application *Phys. Med. Biol.* **46** 1101

Krämer M *et al* 2000 Treatment planning for heavy-ion radiotherapy: physical beam model and dose optimization *Phys. Med. Biol.* **45** 3299

Lin L *et al* 2013 Experimental characterization of 2D pencil beam scanning proton spot profiles *Phys. Med. Biol.* **58** 6193

Lin L *et al* 2014 Experimental characterization of 2D spot profiles for two proton pencil beam scanning nozzles *Phys. Med. Biol.* **59** 493

Low D A, Harms W B, Mutic S and Purdy J A 1998 A technique for the quantitative evaluation of dose distributions *Med. Phys.* **25** 656–61

MATLAB R2014a 2014 *Object Oriented Programming* (Natick, MA: The Mathworks, Inc.) (www.mathworks.com/help/pdf_doc/matlab/matlab_oop.pdf)

Mayo C S and Urie M M 2003 A systematic benchmark method for analysis and comparison of IMRT treatment planning algorithms *Med. Dosim.* **28** 235–42

NEMA PS3 2011 *Digital Imaging and Communications in Medicine (DICOM) Standard, Part 3: Information Object Definitions* (Rosslyn, VA: National Electrical Manufacturers Association) (http://medical.nema.org/)

Pyakuryal A, Myint W K, Gopalakrishnan M, Jang S, Logemann J A and Mittal B B 2010 A computational tool for the efficient analysis of dose-volume histograms for radiation therapy treatment plans *J. Appl. Clin. Med. Phys.* **11** 3013 (PMID: 20160690)

Sánchez Parcerisa D *et al* 2014 Fast range switching of passively scattered proton beams using a modulation wheel and dynamic beam current modulation *Phys. Med. Biol.* **59** N19

Schaffner B 2008 Proton dose calculation based on in-air fluence measurements *Phys. Med. Biol.* **53** 1545–62

Schaffner B and Pedroni E 1998 The precision of proton range calculations in proton radiotherapy treatment planning: experimental verification of the relation between CT-HU and proton stopping power. *Phys. Med. Biol.* **43** 1579–92

Schaffner B, Pedroni E and Lomax A 1999 Dose calculation models for proton treatment planning using a dynamic beam delivery system: an attempt to include density heterogeneity effects in the analytical dose calculation *Phys. Med. Biol.* **44** 27–41

Spezi E, Lewis D G and Smith C W 2002 A DICOM-RT-based toolbox for the evaluation and verification of radiotherapy plans *Phys. Med. Biol.* **47** 4223

Tewell M A and Adams R 2004 The PLUNC 3D treatment planning system: a dynamic alternative to commercially available systems *Med. Dosim.* **29** 134–8

Ulmer W 2007 Theoretical aspects of energy-range relations, stopping power and energy straggling of protons *Radiat. Phys. Chem.* **76** 1089–107

Ulmer W and Matsinos E 2010 Theoretical methods for the calculation of Bragg curves and 3D distributions of proton beams *Eur. Phys. J. Spec. Top.* **190** 1–81

Ulmer W and Schaffner B 2011 Foundation of an analytical proton beamlet model for inclusion in a general proton dose calculation system *Radiat. Phys. Chem.* **80** 378–402

Wilkens J J and Oelfke U 2003 Analytical linear energy transfer calculations for proton therapy *Med. Phys.* **30** 806

Wilkens J J and Oelfke U 2004 3D LET calculations for treatment planning of proton therapy *Z. Med. Phys.* **14** 41–6 (PMID: 15104009)