Title
A novel optimization framework for VMAT with dynamic gantry couch rotation

Journal
Physics in Medicine and Biology, 63(12)

ISSN
0031-9155

Authors
Lyu, Qihui
Yu, Victoria Y
Ruan, Dan
et al.

Publication Date
2018-06-01

DOI
10.1088/1361-6560/aac704

Peer reviewed

HHS Public Access

Author manuscript

Phys Med Biol. Author manuscript; available in PMC 2019 June 13.

Published in final edited form as:

Phys Med Biol. ; 63(12): 125013. doi:10.1088/1361-6560/aac704.

A novel optimization framework for VMAT with dynamic gantry couch rotation

Qihui Lyu, **Victoria Y Yu**, **Dan Ruan**, **Ryan Neph**, **Daniel O'Connor**, and **Ke Sheng**

Department of Radiation Oncology, University of California Los Angeles, Los Angeles, California

Abstract

Existing Volumetric Modulated Arc Therapy (VMAT) optimization using coplanar arcs is highly efficient but usually dosimetrically inferior to intensity modulated radiation therapy (IMRT) with optimized non-coplanar beams. To achieve both dosimetric quality and delivery efficiency, we proposed in this study, a novel integrated optimization method for non-coplanar VMAT (4πVMAT). 4πVMAT with Direct Aperture Optimization (DAO) was achieved by utilizing a least square dose fidelity objective, along with an anisotropic total variation term for regularizing the fluence smoothness, a single segment term for imposing simple apertures, and a group sparsity term for selecting beam angles. Continuous gantry/couch angle trajectories were selected using the Dijkstra's algorithm, where the edge and node costs were determined based on the maximal gantry rotation speed and the estimated fluence map at the current iteration, respectively. The couch-gantry-patient collision space was calculated based on actual machine geometry and a human subject 3D surface. Beams leading to collision are excluded from the DAO and beam trajectory selection (BTS). An alternating optimization strategy was implemented to solve the integrated DAO and BTS problem. The feasibility of 4πVMAT using one full-arc or two full-arcs was tested on 9 patients with brain, lung, or prostate cancer. The plan was compared against a coplanar VMAT (2πVMAT) plan using one additional arc and collimator rotation. Compared to 2πVMAT, 4πVMAT reduced the average maximum and mean organs-at-risk (OARs) dose by 9.63% and 3.08% of the prescription dose with the same target coverage. R50 was reduced by 23.0%. Maximum doses to the dose limiting organs, such as the brainstem, the major vessels, and the proximal bronchus, were reduced by 8.1 Gy (64.8%), 16.3 Gy (41.5%), and 19.83 Gy (55.5%), respectively. The novel 4πVMAT approach affords efficient delivery of non-coplanar arc trajectories that lead to dosimetric improvements compared with coplanar VMAT using more arcs.

1. Introduction

In 1995, intensity-modulated arc therapy (IMAT) was proposed by Yu (Yu 1995), introducing a new method for rotational IMRT delivery on the widely available C-arm gantry machines. In 2007, Otto (Otto 2007) published a practical arc optimization algorithm that fueled the rapid commercial development and clinical adoption of Volumetric Modulated Arc Therapy (VMAT). Compared to the conventional static beam intensity modulated radiotherapy (IMRT), VMAT significantly improves delivery efficiency (Wolff *et al*. 2009; Rao *et al*. 2010; Verbakel *et al*. 2009) while maintaining comparable dosimetry.

Current VMAT methods utilize one or more arcs, each co-planar by itself, with or without manually selected couch rotations *between* arcs. For clarity, we term it 2πVMAT in this

paper. The 2πVMAT optimization problem has been solved by progressively inserting new control points into coarsely sampled gantry angles, until typically a 2-degree spacing between the gantry angles is achieved. This progressive sampling approach is not only computationally tractable, but also provides deliverable VMAT plans that meet the mechanical constraints of multi-leaf collimator (MLC), including the leaf travelling speed limits. However, this heuristic progressive sampling approach has several limitations. First, the greedy optimization algorithm is susceptible to undesired local minima due to its sensitivity to suboptimal parameter tuning in the early coarse-resolution stage of optimization. Second, the optimization result strongly depends on the parameter tuning history, causing difficulties to exactly reproduce a plan. Third, due to the interpolation that take place in the progressive sampling process, the method tends to spread the photon fluence uniformly throughout the entire arc even if only a fraction of the beam angles is optimal for treating the patient. To overcome these limitations, we developed a level-set based direct aperture optimization (DAO) for 2πVMAT (Nguyen, *et al.* 2016), which solves the entire arc optimization problem in full resolution. This non-progressive sampling approach was shown to generate a single arc coplanar VMAT that outperformed progressive sampling VMAT using two arcs with the same number of control points in each arc.

The improved VMAT method is still limited to coplanar arcs while non-coplanar beams hold unquestionable dosimetric advantages as shown by recent 4π IMRT research, which maximally utilizes the non-coplanar beams for significant dosimetric gains compared to the VMAT plans (Dong, *et al.* 2013; Sheng *et al.* 2015; Dong, *et al.* 2013). The advantage of non-coplanar IMRT beams is further demonstrated by Sharfo *et al.* (Sharfo *et al.* 2017) who combined VMAT with a few non-coplanar IMRT beams and showed improved organs-at-risk (OAR) sparing and dose spillage. However, delivering many isolated non-coplanar beams can be time consuming and laborious to the clinic and an undue burden to the patient. Alternatively, non-coplanar VMAT arcs were investigated.

A straightforward non-coplanar VMAT approach was based on user-defined trajectories (Woods *et al.* 2016; Krayenbuehl *et al.* 2006; Shaitelman *et al.* 2011; Liang *et al.* 2015) that are limited to narrow applications or specific patient anatomies. Several automated VMAT trajectory optimization techniques were proposed to generalize the solution. Smyth *et al.* (Smyth *et al.* 2013) introduced a trajectory optimization for VMAT by choosing a minimal cost trajectory on a cost map, which was computed based on ray-OAR voxel intersections. Yang *et al.* (Yang *et al.* 2011) proposed a hierarchical clustering algorithm to find multiple continuous and extended sub-arcs through a minimum search of a score function containing geometrical information. MacDonald *et al.* (Lee MacDonald & Thomas 2015) also introduced a trajectory optimization algorithm that minimizes the geometric overlap between planning target volumes (PTV) and OARs based on the two-dimensional projection from source to the isocenter plane as a function of gantry and couch angle. VMAT optimization was then performed as a separate step. Due to the separation, these methods were ineffective in solving complex planning problem where clean trajectories not entering one or more OARs do not exist. To overcome the limitation, a gantry/couch trajectory optimization needs to be incorporated in dose optimization.

Due to the difficulty of solving the complete non-coplanar VMAT problem, a mathematically tractable way is to first identify the non-coplanar control points using beam orientation optimization (BOO) and fluence map optimization (FMO), which essentially is the goal of 4π IMRT. Following the idea, Wild *et al.* (Wild *et al.* 2015) utilized a genetic algorithm to solve the combinatorial problem of BOO, and Papp *et al.* (Papp *et al.* 2015) used the gradient norm strategy to heuristically select a few promising beams. Once the static beam positions are determined as nodes, non-coplanar arcs are created to connect them. An intrinsic limitation of these methods is that although the static beam positions are dosimetrically desirable, the arcs connecting them are not. By generating the non-coplanar arc plans, the major workload of dose delivery is shifted to these dosimetrically suboptimal arc trajectories.

Evidently, to fundamentally solve the non-coplanar VMAT problem, not only the nodes, but also entire arc trajectories need to be part of the BOO equation. In this study, we propose a novel optimization framework that simultaneously solves the complete non-coplanar VMAT trajectory optimization and DAO problems for VMAT, while ensuring deliverability by avoiding couch-gantry-patient collision and enforcing mechanical constraints of MLC leaf motion and gantry rotation. We term this method 4πVMAT in contrast to the 2πVMAT methods using only individual coplanar arcs. Videos demonstrating the difference between 2πVMAT and 4πVMAT can be found in the supplement material.

2. Materials and methods

2.1. Formulation

2.1.1. Direct Aperture Optimization and Beam Orientation Optimization—The proposed 4πVMAT DAO and BOO formulation is written as

$$\underset{\{f_{b\alpha}, c_{b\alpha}, u_{b\alpha}\}_{b,\alpha=1}^{n_b, n_\alpha}}{minimize} \underbrace{\frac{1}{2}\left\| W\left(\left(\sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha} A_{b\alpha} f_{b\alpha}\right) - d_0\right)\right\|_2^2}_{fidelity\ term}$$

$$+ \sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\left(\lambda_1\left\|D1_{b\alpha} f_{b\alpha}\right\|_1 + \underbrace{\lambda_2\left\|D2_{b\alpha} f_{b\alpha}\right\|_1}_{anisotropic\ TV\ term\ on\ f}\right)$$

$$+ \frac{1}{2}\sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\left(\gamma_1\left\|\sqrt{diag(u_{b\alpha})}(f_{b\alpha} - c_{b\alpha})\right\|_2^2 + \underbrace{\gamma_2\left\|\sqrt{diag(1-u_{b\alpha})} f_{b\alpha}\right\|_2^2}_{single\ segment\ term}\right)$$

$$+ \sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\left(g_1\left\|D1_{b\alpha} u_{b\alpha}\right\|_1 + \underbrace{g_2\left\|D2_{b\alpha} u_{b\alpha}\right\|_1}_{anisotropic\ TV\ term\ on\ u}\right)$$

$$+ \sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\underbrace{\gamma_3 G_{b\alpha}\left\|f_{b\alpha}\right\|_2}_{group\ sparsity\ term} + \gamma_4 G_{b\alpha}(1 - P_{b\alpha})\left\|f_{b\alpha}\right\|_2 + \underbrace{g_3\left\|D_P u\right\|_1}_{aperture\ continuty\ term}$$

$$subject\ to\ \ f_{b\alpha} \geq 0, c_{b\alpha} \geq 0, u_{b\alpha} \geq 1, b = 1, 2, \ldots, n_b, \alpha = 1, 2, \ldots, n_\alpha$$

$$f_{b\alpha} = 0, c_{b\alpha} = 0, u_{b\alpha} = 0, \quad \forall (b, \alpha) \notin S$$

$$u = \left[u^T_{b=1\ \alpha=1}\ u^T_{b=1\ \alpha=2}\ \cdots\ u^T_{b=1\ \alpha=\alpha_n}\ u^T_{b=2\ \alpha=1}\ u^T_{b=2\ \alpha=2}\ \cdots\ u^T_{b=n_b\ \alpha=n_\alpha}\right]^T, \tag{1}$$

where the notations for the variables and data are summarized in Table 1.

In this formulation, the data fidelity term attempts to find the optimal fluence map $f_{b\alpha}$ such that the total calculated dose from all candidate beams is as close as possible to the optimal dose d_0.

The priorities for structures of interest are controlled by the diagonal weighting matrix W. The anisotropic total variation (TV) regularization on fluence map $f_{b\alpha}$ encourages piecewise continuity of the fluence map within each candidate beam (Zhu *et al.* 2008). λ_1 and λ_2 control the degree of piecewise continuity of $f_{b\alpha}$ in the direction parallel and orthogonal to

the MLC leaf direction, respectively. A balanced tradeoff between dose fidelity and fluence map continuity achieves a high quality and deliverable dose distribution.

The single segment term enforces the final fluence map f_{ba} to contain only one segment per candidate beam. It pushes f_{ba} towards c_{ba} where u_{ba} is 1, and towards 0 where u_{ba} is 0, encouraging the fluence map to be a constant within the aperture, and 0 outside the aperture. By gradually increasing the weighting on this term during the optimization process, the optimizer forces single segment per candidate beam. The anisotropic TV term on aperture variable u_{ba} encourages large segments and penalizes holes in the aperture, with hyperparameters λ_1 and λ_2 controlling the degree of segment continuity in the direction parallel and orthogonal to the MLC leaf direction respectively. The single segment term and the anisotropic TV term address the hardware constraints by enforcing a single deliverable segment within each candidate beam.

The group sparsity term is a $l_{2,1}$ norm penalty. This convex penalty provides a non-greedy approach for BOO in 4πVMAT by promoting group sparsity in the fluence map f_{ba} and encouraging most of candidate beams to be inactive. This term is divided into two components through a parameter P_{ba}, which is defined as 1 for candidate beams that are on the selected trajectory in the beam trajectory selection (BTS) process, and 0 elsewhere. y_3 controls the sparsity level for all candidate beams and y_4 adds additional weighting on candidate beams that are not on the optimal trajectory. The aperture continuity term regulates leaf movement by penalizing apertures differences between adjacent candidate beams on the selected trajectory from BTS. The group sparsity term and the aperture continuity term address the hardware constraints by switching off the off-trajectory beams by the end of the optimization and encouraging aperture continuity along the trajectory respectively.

The flowchart of the proposed 4πVMAT optimization framework is shown in Figure 1. First of all, a patient-specific collision map is generated and candidate beams are divided into a feasible set S and an infeasible set based on an individualized collision model for non-coplanar radiotherapy delivery (Yu et al. 2015). Then P_{ba} is assigned as 1 for all feasible beams, which serves as prior information for the DAO&BOO. The DAO&BOO solves the optimization problem in equation (1) and optimizes fluence map , fluence intensity c_{ba}, and vectorized aperture variable u_{ba} alternatingly, generating a 4πVMAT plan with a small fraction of beams active. Note that these on-beams are not necessarily on a connected trajectory; instead they are separated in the entire feasible space. The optimal fluence map f_{ba} is then utilized to generate the simplified Dijkstra's map (Bollobás & Riordan 1993; Dijkstra 1959), on which the BTS is performed and one or two trajectories are selected based on the tumor region. P_{ba} is assigned as 1 for all candidate beams on the selected trajectories and 0 elsewhere, which is then translated into the DAO&BOO in the next iteration, penalizing heavily on the fluence map development for off-trajectory candidate beams as compared with on-trajectory beams. The optimization alternates between DAO&BOO and BTS until convergence, allowing BTS to fully explore the dose domain before converging to a final trajectory.

2.1.2. Beam trajectory selection—The BTS is formulated as a travelling salesman problem, where the gantry/couch graph G = [*N, E*] is defined as a number of nodes *N* and edges *E* that connects every two nodes. In this 4πVMAT optimization framework, the gantry/couch graph contains n_a by n_b nodes, representing n_a gantry angles and n_b couch angles, and associated with a node cost $NC(b, a)$ for each candidate beam (b, a) and the edge cost $EC((b_1, a_1), (b_2, a_2))$ for the edge that connects (b_1, a_1) and (b_2, a_2). A graph search algorithm is performed on the graph to determine the shortest path from one node to any other nodes. For the patient safety and comfort, this 4πVMAT framework enforces a constant couch rotation direction within each arc and allows the gantry to rotate dynamically. The 4πVMAT arc is represented by a continuous, one-way path on the gantry/couch graph, starting from the first couch angle and ending at the last couch angle, containing only one candidate beam per couch angle.

The node cost $NC(b, a)$ for a feasible candidate beam with couch angle b and gantry angel a depends on f_{ba}, the fluence map from DAO&BOO at the current iteration, which is deliverable for each candidate beams, though they may not be connected by a trajectory. The node cost (b, a) for an infeasible candidate beam is infinity, which enforces trajectory selection from feasible beams only. The formulation of (b, a) is given by

$$Cost_{max} = max\left\{\frac{1}{2}\left\|W(A_{b\alpha}f_{b\alpha} - d_0)\right\|_2^2 \middle| \forall(b,\alpha) \in S\right\}$$
$$NC(b,\alpha) = \begin{cases} -log\left(Cost_{max} - \frac{1}{2}\left\|W(A_{b\alpha}f_{b\alpha} - d_0)\right\|_2^2\right), & if (b,\alpha) \in S \\ \infty, & if (b,\alpha) \notin S \end{cases}. \quad (2)$$

Candidate beams that correspond to dose contribution closer to the ideal dose distribution are assigned with lower node costs. By finding the shortest path on the Dijkstra's graph, the optimal trajectory prefers candidate beams that are more dosimetrically promising.

The edge costs between every two nodes enforce constraints on trajectory selection. The edge cost between candidate beams (b_1, a_1) and (b_2, a_2) is defined as

$$EC\big((b_1,\alpha_1),(b_2,\alpha_2)\big) = \begin{cases} 0 & if\ b_2 - b_1 = 1\ and\ \left\|\alpha_1 - \alpha_2\right\| < \alpha_0 \\ \infty & otherwise \end{cases}, \quad (3)$$

where a_0 is calculated as $a_0 = \omega_g \cdot \tau$, with ω_g representing the maximum angular rotation speed of the gantry and τ the time frame between each control point. Equation (3) ensures that the selected trajectory is a continuous one-way path that contains only one gantry angle per couch angle, and that the gantry angles of adjacent beams on the selected trajectory are within gantry rotation speed limits (controlled by a_0). For simpler cases such as the brain tumor, a single 4πVMAT arc can be adequate to achieve satisfactory dose profile. For more complex patients and body sites where non-coplanar angles are restricted by collision, two arcs may be necessary to substantially improve the dose distribution. In this case, the node

cost (b, a) is assigned to be infinity for any candidate beams on the first selected trajectory, and then the second trajectory is chosen by finding the shortest path on the updated graph.

2.2. Algorithm

Section 2.1 illustrates the entire optimization framework, including the DAO&BOO module and the BTS module. This section provides a solution to the optimization framework. The DAO&BOO module was solved using FISTA, with details discussed in the section 2.2.1. The BTS module was solved using Dijkstra's algorithm, shown in the section 2.2.2. Together the optimizer solves the entire problem iteratively and alternatingly between the two modules.

2.2.1. DAO&BOO using FISTA—Akin to our previous study (Nguyen, *et al.* 2016), the DAO&BOO module consists of three submodules, each of which solves the optimization problem in equation (1) with respect to one optimization variable while holding the other two variables constant, and the whole optimization problem is solved in an alternating block fashion. In the submodule 2, the algorithm solves equation (1) with respect to the fluence intensity c_{ba}, and results in a closed form equation for c_{ba}. In the submodule 1 and submodule 3, the optimization problem solves for the vectorized fluence map f_{ba} and the vectorized aperture variable u_{ba}, and generates a convex optimization problem that can be efficiently solved through an accelerated proximal gradient method, known as the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) (Beck & Teboulle 2009).

FISTA is capable of solving any optimization problem that can be formulated as

$$minimize\ F(x) + G(x), \quad (4)$$

where F is a differentiable convex function with a Lipschitz continuous gradient, and G is a convex function which has a proximal operator that can be evaluated efficiently. The proximal operator (Parikh & Boyd 2013) of a function G with step size t is defined by

$$prox_{tG}(x) = \underset{z}{argmin}\left(G(z) + \frac{1}{2t}\left\|z - x\right\|_2^2\right).$$

The pseudocode for the FISTA with line search algorithm is summarized in Algorithm 1, where evaluation of the gradient of F and the proximal operator of G are required at each iteration, and the function value of F is also assessed to obtain the optimal step size through a line search method.

Algorithm 1

FISTA with line search

Pseudocode for FISTA with line search
Initialize x_0: = 0, v_0: = x_0 , $t_0 > 0$, $r_1 > 1$, $r_2 > 1$
for $k = 1, 2, \ldots$ *do*

Pseudocode for FISTA with line search

$t \coloneqq r_1 t_{k-1}$

Repeat

$$\theta := \begin{cases} 1 & if\ k = 1 \\ positive\ root\ of\ t_{k-1}\theta^2 = t\theta_{k-1}^2(1-\theta) & if\ k > 1 \end{cases}$$

$y \coloneqq (1-\theta)x_{k-1} + \theta v_{k-1}$

$x \coloneqq prox_{tG}(y - t\nabla F(y))$

$$break\ if\ F(x) \le F(y) + < \nabla F(y), x - y > + \frac{1}{2t}\|x-y\|_2^2$$

$t \coloneqq t/r_2$

$t_k \coloneqq t$

$\theta_k \coloneqq \theta$

$$v_k := x_k + \frac{1}{\theta_k}(x - x_k)$$

$$break\ if \frac{\|x - x_k\|}{\|x_k\|} \le \varepsilon$$

$x_k \coloneqq x$

end for

return x

2.2.1.1. Algorithm Submodule 1: Update f_{ba}: Submodule 1 updates the fluence map vector f_{ba} while holding the aperture variable u_{ba} and fluence intensity c_{ba} constant. Omitting terms that do not depend on f_{ba}, we see that solving equation (1) with respect to the vector f_{ba} is equivalent to solving the reduced problem

$$\underset{\{f_{b\alpha}\}_{b,\alpha=1}^{n_b,n_\alpha}}{minimize}\ \frac{1}{2}\left\|W\left(\left(\sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}A_{b\alpha}f_{b\alpha}\right)-d\right)\right\|_2^2+\sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\left(\lambda_1\left\|D_{1_{b\alpha}}f_{b\alpha}\right\|_1+\lambda_2\left\|D_{2_{b\alpha}}f_{b\alpha}\right\|_1\right) \quad (5)$$

$$+\frac{1}{2}\sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\left(\gamma_1\left\|\sqrt{diag(u_{b\alpha})}(f_{b\alpha}-c_{b\alpha})\right\|_2^2+\gamma_2\left\|\sqrt{diag(1-u_{b\alpha})}f_{b\alpha}\right\|_2^2\right)$$

$$+\sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\left(\gamma_3 G_{b\alpha}\left\|f_{b\alpha}\right\|_2+\gamma_4 G_{b\alpha}(1-P_{b\alpha})\left\|f_{b\alpha}\right\|_2\right)$$

$$subject\ to\ \ f_{b\alpha}\geq 0,\quad b=1,2,\ldots,n_b,\alpha=1,2,\ldots,n_\alpha$$

$$f_{b\alpha}=0,\quad \forall(b,\alpha)\notin S$$

This subproblem finds the optimal fluence map $f_{b\alpha}$ that minimizes the difference between the calculated dose and the ideal dose, regularizes $f_{b\alpha}$ to approach the aperture $u_{b\alpha}$ and intensity $c_{b\alpha}$ through a total variation term and a single segment term, while at the same time encouraging most candidate beams to be inactive by utilizing the group sparsity term. As shown in the appendix, equation (5) can be written in the canonical FISTA form, and the optimization problem could be solved through FISTA with evaluation of the gradient of F_1 and the proximal operator of G_1.

2.2.1.2. Algorithm Submodule 2: Update $c_{b\alpha}$: Submodule 2 updates $c_{b\alpha}$ while holding $f_{b\alpha}$ and $u_{b\alpha}$ constant. The optimization problem in equation (1) with respect to $c_{b\alpha}$ is equivalent to

$$\underset{\{f_{b\alpha}\}_{b,\alpha=1}^{n_b,n_\alpha}}{minimize}\ \sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\left(\left\|\sqrt{diag(u_{b\alpha})}(f_{b\alpha}-c_{b\alpha})\right\|_2^2\right)$$

$$subject\ to\ \ c_{b\alpha}\geq 0,\quad b=1,2,\ldots,n_b,\alpha=1,2,\ldots,n_\alpha.$$

$$c_{b\alpha}=0,\quad \forall(b,\alpha)\notin S.$$

The optimized $c_{b\alpha}$ has a closed form

$$c_{b\alpha} = \frac{u_{b\alpha}^T f_{b\alpha}}{u_{b\alpha}^T \vec{1}} \quad b = 1, 2, .., n_b, \quad \alpha = 1, 2, .., n_\alpha, \quad (b, \alpha) \in S.$$

This optimal fluence intensity $c_{b\alpha}$ is the average of all beamlet intensities within the aperture for each candidate beam.

2.2.1.3. Algorithm Submodule 3: Update $u_{b\alpha}$: Submodule 3 updates $u_{b\alpha}$ while holding $f_{b\alpha}$ and $c_{b\alpha}$ constant. The optimization problem in equation (1) with respect to $u_{b\alpha}$ is equivalent to

$$\underset{\{u_{b\alpha}\}_{b,\alpha=1}^{n_b,n_\alpha}}{minimize} \quad \frac{1}{2}\sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\left(\gamma_1\left\|\sqrt{diag(u_{b\alpha})}(f_{b\alpha}-c_{b\alpha})\right\|_2^2+\gamma_2\left\|\sqrt{diag(1-u_{b\alpha})}f_{b\alpha}\right\|_2^2\right) \quad (6)$$

$$+\sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\left(g_1\left\|D_{1_{b\alpha}}u_{b\alpha}\right\|_1+g_2\left\|D_{2_{b\alpha}}u_{b\alpha}\right\|_1\right)+g_3\left\|D_P u\right\|_1$$

$$subject\ to \quad 0 \le u_{b\alpha} \le 1, \quad b = 1, 2, \ldots, n_b, \alpha = 1, 2, \ldots, n_\alpha$$

$$u_{b\alpha} = 0, \quad \forall (b, \alpha) \notin S$$

This optimization problem could also be efficiently solved through FISTA, which is described in the appendix. The whole DAO&BOO is solved in an alternating block fashion that repeatedly runs through submodule 1 to submodule 3 until the fluence map converges.

2.2.2. Beam Trajectory Selection: Dijkstra's Algorithm—As illustrated in section II. 1.2, the gantry/couch graph contains n_α by n_b nodes with node costs and edge costs defined in equation (2) and equation (3). The edge costs in this problem is special, since all connections between two candidate beams are assigned as infinity except for the path that goes directly to the next couch angle with a moderate gantry angle variation. It ensures that the selected trajectory starts from the first couch angle and ends at the last couch angle, with only one candidate beam selected for each couch angle. This setup not only guarantees a practicable trajectory for 4πVMAT delivery that addresses patient safety and comfort, but also makes the travelling salesman problem more straightforward and computationally inexpensive to solve. This special graph optimization problem is solved using a simplified Dijkstra's algorithm (Bollobás & Riordan 1993; Dijkstra 1959) as shown in Algorithm 2. a_{short} is a vector with n_b elements denoting the candidate beams on the selected trajectory, with couch angle b and gantry angle $a_{short}(b)$. For lung and prostate patient, a secondary trajectory a_{short2} is obtained on top of a_{short}. Once the trajectories are chosen, $P_{b\alpha}$ is

assigned to be 1 for any selected candidate beams, and 0 elsewhere, and is utilized in DAO&BOO to assign different group sparsity penalties for selected and unselected candidate beams.

Algorithm 2

Pseudocode for Simplified Dijkstra's algorithm

Simplified Dijkstra's algorithm

Initialize the graph with each node $G(b, \alpha) \coloneqq \infty$

$G(1, \alpha) \coloneqq NC(1, \alpha)$

for $b = 1, 2, \ldots, n_b - 1$

  for $\alpha = 1, 2, \cdots, n_\alpha$

$$G(b+1,\alpha) := \min_{\alpha'} G(b,\alpha') + NC(b+1,\alpha) + EC((b,\alpha'),(b+1,\alpha))$$

$$\alpha_t(b+1,\alpha) := \operatorname*{argmin}_{\alpha'} G(b,\alpha') + NC(b+1,\alpha) + EC((b,\alpha'),(b+1,\alpha))$$

  end for

end for

$$\alpha_{short}(n_b) := \operatorname*{argmin}_{\alpha'} G(n_b,\alpha')$$

for $b = n_b - 1, n_b - 2, \ldots, 1$

  $\alpha_{short}(b) \coloneqq \alpha_{short}(b + 1, \alpha_1(b + 1))$

end for

if planning for lung and prostate cancer patient

  $NC(b, \alpha_{short}(b)) \coloneqq \infty$

  $G_2(1, \alpha) \coloneqq NC(1, \alpha)$

  for $b = 1, 2, \cdots, n_b$

    for $\alpha = 1, 2, \cdots, n_\alpha$

$$G_2(b+1,\alpha) := \min_{\alpha'} G_2(b,\alpha') + NC(b+1,\alpha) + EC((b,\alpha'),(b+1,\alpha))$$

$$\alpha_{t2}(b+1,\alpha) := \operatorname*{argmin}_{\alpha'} G_2(b,\alpha') + NC(b+1,\alpha) + EC((b,\alpha'),(b+1,\alpha))$$

    end for

  end for

$$\alpha_{short2}(n_b) := \operatorname*{argmin}_{\alpha'} G_2(n_b,\alpha')$$

  for $b = n_b - 1, n_b - 2, \cdots, 1$

    $\alpha_{short2}(b) \coloneqq \alpha_{t2}(b + 1, \alpha_{short2}(b + 1))$

  end for

End

2.3. Implementation Details

The feasible beam set is based on the collision model from our previous study (Yu *et al.* 2015), which provides an individualized collision prediction model for the purpose of non-

coplanar beam delivery in IMRT. In the 4πVMAT study, all candidate beams that are predicted to be undeliverable isocentrically are excluded, and a 10-degree margin is added to both upper bound and lower bound of the infeasible gantry angle range for each couch angle to avoid couch-gantry-patient collision during the gantry/couch rotation between feasible beams.

Parameter tuning can be tricky in this 4πVMAT framework, since there are a number of hyper-parameters in addition to the regular structure weightings that help ensure the deliverability of 4πVMAT plan. Nevertheless, the optimization results are insensitive to most of the hyper-parameters that control the deliverability. Once a suitable set of these parameters are found, our experiments show robust performance even if the structure weightings are changed or applied to different patient cases with comparable sizes of PTV and OARs.

To achieve a satisfactory local minimum that makes much of the data fidelity term, some heuristics are set up for parameters $y_1 - y_4$. All of them increase as the optimization progresses, which allow the optimization to focus on the convex dose fidelity term and ensure a good local minimum with high dosimetric quality at the initial stages of the optimization. During the optimization, the single segment term and group sparsity term are gradually emphasized to ensure plan deliverability of the DAO result. y_1 and y_2 were updated when each round of three submodules in DAO is completed

$$\gamma_1 = \gamma_{10} \cdot \left(1 + (1000)^{\frac{i}{N}}\right) \quad \gamma_2 = \gamma_{20} \cdot \left(1 + (1000)^{\frac{i}{N}}\right),$$

where i is the number of rounds from submodule 1 to submodule 3 within the DAO module, N is the total number of rounds within the DAO module, which we chose to be 6 in this study empirically. y_{10} and y_{20} are the initial values of y_1 and y_2 respectively. y_3 and y_4 enforces the number of active candidate beams to be within certain ranges, which is designed to decrease with the iteration of BTS process. In submodule 1 where the group sparsity term controls the sparsity level of , the number of active beams is evaluated every 20 iterations. y_3 and y_4 are automatically increased or decreased by 20% if there are too many or insufficient active beams. All the hyperparameters were tuned based on the optimization performance. For example, the λ_1 and λ_2 would be increased if the fluence map was too rough, or decreased if it was too smooth. We also increase the values of y_{10} and y_{20} if there are multiple segments in one candidate beam. The range of these hyper parameters in our study are listed in Table 2.

A video showing the alternating optimization process can be found in the supplement material. The video shows the fluence map f_{ba}, fluence intensity c_{ba}, aperture variable u_{ba}, and selected trajectory P_{ba} alternatingly, labelled by the iteration number $N_t \cdot N_D \cdot N_x$, where N_t is the iteration number of performing the DAO&BOO module and BTS module, N_D is the iteration number within the DAO&BOO module, and N_x is the iteration number within the submodules of the DAO&BOO that optimizes for f_{ba}, c_{ba}, and u_{ba} respectively.

2.4. Evaluation

The feasibility of the optimization algorithm was tested on three glioblastoma multiforme retreatment patient (GBM), three lung cancer patient (LNG), and three prostate cancer patient (PRT). The GBM retreatment planning followed an internal protocol to minimize dose to previously irradiated critical organs (Yu et al. 2018). The centrally located lung and the prostate stereotactic body radiotherapy (SBRT) plans follow (Chang et al. 2008) and (King et al. 2012), respectively. Table 3 summarizes the number of feasible beams, the prescription doses, and PTV volumes for all patients. patients

Using a convolution/superposition code with a 6 MV x-ray polyenergetic kernel as described in our previous publications (Neylon *et al.* 2014; Dong, *et al.* 2013), the beamlet dose was calculated for all feasible beams among 2400 candidate beams in 4πVMAT. Since the gantry rotation angles are more likely to be collision-free when the couch stays close to central position, the couch angles are more densely sampled for couch angles ranging from 15° to 0° and from 360° to 345°, following the International Electrotechnical Commission (IEC 61217) convention. The 2400 candidate beams correspond to 30 gantry angles with 12-degree spacing from 0 to 359 degrees, and 80 couch angles, separated by 1 degree for the central 30 degree range and 3 degrees elsewhere, from 270 to 90 degrees. For 2πVMAT, the beamlet dose was calculated for 80 gantry angles with 4.5 degree of separation for each arc, from 0 to 359 degrees. The beamlet resolution was 0.5×0.5 cm^2, and the dose array resolution was 0.25×0.25×0.25 cm^3.

All arcs contain 80 control points in this study. The 4πVMAT plan contains one arc for the GBM patient, 80 control points in total, which is compared against a 2πVMAT plan with two coplanar arcs with collimator angles at 45° and 135°, 160 control points in total. For the lung cancer patient and prostate cancer patient, the 4πVMAT plan contains two arcs with 160 control points in total, while the 2πVMAT contains three arcs with collimator angles at 30°, 90°, and 150°, 240 control points in total. The 2πVMAT and 4πVMAT utilize the same DAO model except that the 2πVMAT has a predefined coplanar trajectory. The travelling time between each beam is 2s, allowing 12° of gantry rotation.

PTV statistics including PTV D95, D98, D99, D2, (defined as the dose which is received by at least 95%, 98%, 99%, and 2% of the volume respectively), and PTV homogeneity $\left(\text{defined as}\frac{D_{95}}{D_5}\right)$ were evaluated. For OAR, the Dmax and Dmean were assessed. Maximum dose is defined as D2, recommended by the ICRU-83 report (Grégoire & Mackie 2011). The dose conformity defined as the ratio between the PTV volume receiving 100% or more of the prescription dose and the PTV volume, and the Integral Dose defined as the volume integral of the dose deposited in the patient, were also obtained. The R50, defined as the 50% isodose volume divided by the target volume, was evaluated to quantify the amount of high dose spillage in the patient body.

3. Results

Figure 2 shows the fluence map of the GBM #1 patient as an example of the optimization result on the Gantry/Couch graph. The red regions on the Gantry/Couch graph indicate the

candidate beams that cause collision. Notice that only those candidate beams that are on the selected trajectory have nonzero fluence weights. To deliver the 4πVMAT plan, the couch rotates from 90° to 270° and the gantry rotates accordingly. Figure 3 shows the selected beam angles on the Gantry/Couch graph with corresponding 3D view for all patients. The 4πVMAT algorithm optimizes for all candidate beams simultaneously and finds one or two optimal trajectories that go through the safe region, from the first couch angle to the last couch angle, indicated by the green blocks, affording efficient delivery of non-coplanar arc. The maximum allowance of the gantry rotation is 12° between two adjacent beams. With the gantry rotation speed limit at 6° per second, 2 seconds is sufficient for delivery at one control point. For the GBM patient, the 4πVMAT plan takes around 3 minutes in estimation to deliver the single arc 4πVMAT plan, and for the lung cancer patient and prostate cancer patient, the calculated time is approximately 5 minutes to deliver the two arcs 4πVMAT plan. Videos demonstrating the 4πVMAT delivery process can be found in the supplement material.

Figure 4 shows the dose distribution of 4πVMAT and 2πVMAT for all patients. By using non-coplanar beams, 4πVMAT has the flexibility to distribute the dose in any non-colliding direction within the 4π spherical space, depending on the benefits of OAR sparing and target coverage. A greater separation of these non-coplanar beams also reduces the high dose spillage in 4πVMAT. In the GBM cases, the geometry is simpler, the single arc 4πVMAT plans were able to significantly avoid dose spillage to the brainstem as well as other critical structures, such as the chiasm and right optic nerve in the GBM #1 case, and the left cochlea in the GBM #2. In all the LNG cases, by utilizing beams with greater separation, the two-arc 4πVMAT plans substantially reduced the high dose spillage compared with the three-arc 2πVMAT plan, including those to the critical organs, such as the chest wall, proximal bronchus, and spinal cord in the LNG #1, the heart in LNG #2, and the aorta and pulmonary vessel in LNG #3. For all three prostate cases, the two-arc 4πVMAT plans resulted in a more desirable asymmetric dose in the anterior/posterior direction to substantially better spare the rectum as compared with the three-arc 2πVMAT plan, while achieving similar bladder high dose volumes. In all cases, by utilizing these OAR-sparing angles, 4πVMAT achieved consistently better OARs sparing at the same time maintaining or improving PTV coverage.

Figure 5 shows the DVHs of 4πVMAT and 2πVMAT for all OARs of the patients. Complete DVH plots with PTV included are shown in the supplement materials. The 4πVMAT was able to markedly reduce dose to OARs while achieving comparable or better PTV statistics across all patients, especially for the dose limiting organs, such as the brainstem in the GBM #2 and GBM #3, the proximal bronchus in all three LNG patients, the major vessels in LNG #2, and LNG #3, and the seminal vesicle and the rectum in all PRT patients. In the GBM #3 case, the one-arc 4πVMAT plan reduced the maximum doses to the brainstem by 8.1 Gy (64.8%) compared with the two-arc 2πVMAT plan. In the LNG #1 case and the LNG #2 case, the two-arc 4πVMAT plan reduced the dose to the proximal bronchus and the major vessels by 16.3 Gy (41.5%) and 19.83 Gy (55.5%) compared with the three-arc 2πVMAT plan. Across all LNG case, the 4πVMAT plan substantially reduced the hot spots in the critical structures that were covered by high dose up to 30Gy.

The quantitative statistics for the PTV and OARs are shown in Table 4 and Table 5. The PTV statistics, Dose Conformity, and Integral Dose are comparable across 4πVMAT and 2πVMAT, but the 4πVMAT is able to substantially reduce R50, indicating a remarkable improvement in dose compactness that is consistent with previously reported 4π IMRT-VMAT comparison (Tran et al. 2017). On average, the 4πVMAT plan reduced the OARs max and mean doses by 9.63% and 3.08% of the prescription dose. The single largest sparing in OARs max and mean dose is up to 39.7% and 14.9% of the prescription dose.

4. Discussion

Our previous 2πVMAT method (Nguyen, *et al.* 2016) solves the VMAT optimization problem by considering all beams simultaneously instead of progressive sampling, making it possible to incorporate BOO into non-coplanar VMAT. The integrated 4πVMAT optimization framework further develops the 2πVMAT method by simultaneously solving both the DAO and BTS problems. By expanding the solution to the non-coplanar space, 4πVMAT was able to greatly improve the dose compactness as indicated by R50, reduce dose spillage to OARs and subsequently achieve better dosimetry than the 2πVMAT with more arcs. Considering that the non-progressive sampling 2πVMAT was already superior to the existing progressive sampling VMAT algorithm that is widely employed in clinics (Nguyen, *et al.* 2016), 4πVMAT holds the strong promise of substantially improving state-of-the-art radiotherapy without sacrificing delivery efficiency.

4πVMAT takes an alternating approach between DAO&BOO and BTS. From the dosimetric point of view, the DAO&BOO achieves high plan quality by activating a set of dosimetrically promising beam, and the BTS finds the optimal trajectory that tends to include the most of these active beams by minimizing the selected node costs. On the other hand, to address the mechanical constraints, the regularization in the DAO&BOO enforces the plan deliverability, and the BTS meets the delivery/treatment time constraint and selects only the trajectories that does not require a substantial gantry rotation between adjacent beams, by minimizing the edge cost. The selected trajectory is then translated into DAO&BOO to further explore the dose domain through a weighted group sparsity term, where off-trajectory candidate beams are penalized more heavily. In the DAO&BOO process, promising off-trajectory beams could still be turned on to provide better dose distribution, and on-trajectory beams might be rejected if they are dosimetrically undesirable. By alternatingly optimizing between the DAO&BOO and BTS, the algorithm makes an integrated final decision based on both the dosimetry and mechanical constraints. With the gantry rotation speed limit at 6° per second, 2 seconds is sufficient for delivery at one control point of the current 4πVMAT plan, where the largest gantry rotation angle is 12 degrees between adjacent beams. A more rapid delivery may be achieved with faster gantry rotation.

Compared with the previous non-coplanar VMAT study with user-defined heuristic trajectories (Krayenbuehl *et al.* 2006; Shaitelman *et al.* 2011; Liang *et al.* 2015) and trajectory optimization techniques using geometric information (Smyth *et al.* 2013; Yang *et al.* 2011; Lee MacDonald & Thomas 2015), this 4πVMAT optimization framework is able to thoroughly and automatically search the entire non-coplanar trajectory space for various

patient anatomies and dosimetric requirements. Compared with other non-coplanar VMAT optimization algorithms using optimized static IMRT beams as anchoring nodes for non-coplanar arcs that may not be dosimetrically optimal (Wild *et al.* 2015; Papp *et al.* 2015), our 4πVMAT integrates BTS into VMAT optimization, and encourages the fluence map to be developed natively into 4πVMAT arcs. Compare with the current commercial solution HyperArc (Ohira et al. 2018), which is restricted to using a few predefined trajectory templates and is currently only applied to the brain treatment, the proposed 4πVMAT framework fully explores the 4π space with a site-specific collision model, and hence holds the promise of substantially improving the dose profile and being more broadly applicable to other body sites.

4πVMAT in its current form is computationally intensive due to the optimization problem size and alternating optimization between DAO and BTS despite our effort to accelerate the computation. In our previous study on FMO for IMRT (Nguyen *et al.* 2015; Nguyen, *et al.* 2016) and VMAT (Nguyen, *et al.* 2016), a first-order primal-dual algorithm, the Chambolle-Pock (Chambolle & Pock 2011) algorithm, has shown a robust performance for treatment planning optimization and is computationally inexpensive compared to other first-order methods such as alternating direction method of multipliers (ADMM) (Boyd 2011), since ADMM requires to solve a linear equation involving the system matrix at each iteration. In this study, the DAO modules are solved by FISTA (Beck & Teboulle 2009), a fast proximal gradient method. On one hand, FISTA shares the same merits with the Chambolle-Pock algorithm in that it only requires the multiplication of the system matrix and its transpose at each iteration. On the other hand, it achieves a convergence rate of $O(\frac{1}{k^2})$, which is substantially faster than the $O(\frac{1}{k})$ convergence rate of the Chambolle-Pock algorithm. Even with the accelerated algorithm, the optimization run time on a single desktop using MATLAB implementation still took from 1 hour for the GBM case to 9 hours for the LNG case and PRT case. To further speed up optimization for clinical implementation, aside from switching to a higher performance language such as C, the most computationally expensive matrix multiplication step in FISTA is parallelizable and can therefore be moved to graphic processing units (GPUs) for higher computation efficiency. The current 4πVMAT framework was evaluated with 12 degrees of gantry angle spacing to reduce the computation cost. With future acceleration techniques, we may be able to perform the optimization with a denser sampling of the 4π space. Another limitation of the 4πVMAT optimization framework is its complexity in tuning the hyperparameters, compounded with its heavy computation costs, can be time consuming and laborious. One way to get around with this problem is to simplify the dataset when tuning the hyperparameters, either by downsampling the structure of interest or truncating the dose calculation matrix. From our experience, the algorithm is insensitive to this procedure, hence the same set of hyperparameters apply to the full sampled data.

5. Conclusion

This study presents a novel non-coplanar 4πVMAT method, which solves the direct aperture optimization (DAO) and beam trajectory selection (BTS) in an alternating approach using

FISTA and Dijkstra's algorithms. Without sacrificing deliverability and treatment efficiency, the novel 4πVMAT method substantially improves dose compactness and OAR sparing with the same PTV coverage, as compared with coplanar VMAT (2πVMAT) with more arcs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research is supported by DOE Grants No. DE-SC0017057 and DE-SC0017687, and NIH Grant R44CA183390, R01CA188300 and R43CA183390.

Appendix

A1 Smooth approximation to facilitate the application of FISTA

FISTA requires the objective function to be the sum of a smooth term and a simple term, which makes the nondifferentiable l_1 norm in the TV regularization terms difficult to deal with. To avoid this problem, the l_1 norm is replaced by the Huber penalty (Noe 1930) defined as

$$H(x) = \left\|x\right\|_1^{(\mu)} = \sum_i \left|x_i\right|^{(\mu)},$$

where

$$\left|x_i\right|^{(\mu)} = \begin{cases} \frac{1}{2\mu} x_i^2, & \left|x_i\right| \le \mu \\ \left|x_i\right| - \frac{\mu}{2}, & \left|x_i\right| > \mu \end{cases}.$$

The Huber penalty provides a convex, differentiable approximation to the l_1 norm, with a smoothing parameter μ. The gradient of the Huber penalty function (x) is Lipschitz continuous with Lipschitz constant $L = \frac{1}{\mu}$.

A2 Apply FISTA on Submodule 1

To apply FISTA on the subproblem formulated in equation (5), it is necessary to formulate the problem in the canonical FISTA form shown in equation (4). Here we define functions F_1 and G_1 as:

$$F_1(f) = \frac{1}{2}\|W(Af-d)\|_2^2 + \lambda_1\left\|D_1 f\right\|_1^{(\mu)} + \lambda_2\left\|D_2 f\right\|_1^{(\mu)}$$

$$+ \frac{1}{2}\left(\gamma_1\|\sqrt{diag(u)}(f-c)\|_2^2 + \gamma_2\|\sqrt{diag(1-u)}f\|_2^2\right)$$

$$G_1(f) = \sum_{b=1}^{n_b}\sum_{\alpha=1}^{n_\alpha}\left(\gamma_3 G_{b\alpha}\left\|f_{b\alpha}\right\|_2 + \gamma_4 G_{b\alpha}(1-P_{b\alpha})\left\|f_{b\alpha}\right\|_2\right) + I_+(f_{b\alpha})$$

where

$$I_+(f_{b\alpha}) = \begin{cases} 0 & if\ f_{b\alpha} \geq 0\ and\ (b,\alpha) \in S, \\ \infty & otherwise \end{cases},$$

$$A = \left[A_{b=1\ \alpha=1}\ A_{b=1\ \alpha=2}\ \cdots\ A_{b=1\ \alpha=n_\alpha}\ A_{b=2\ \alpha=1}\ A_{b=2\ \alpha=2}\ \cdots\ A_{b=n_b\ \alpha=n_\alpha}\right],$$

$$D_1 = \begin{bmatrix} D_{1\,b=1\ \alpha=1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & D_{1\,b=n_b\ \alpha=n_\alpha} \end{bmatrix},\ D_2 = \begin{bmatrix} D_{2\,b=1\ \alpha=1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & D_{2\,b=n_b\ \alpha=n_\alpha} \end{bmatrix},$$

$$f = \left[f^T_{b=1\ \alpha=1}\ f^T_{b=1\ \alpha=2}\ \cdots\ f^T_{b=1\ \alpha=n_\alpha}\ f^T_{b=2\ \alpha=1}\ f^T_{b=2\ \alpha=2}\ \cdots\ f^T_{b=n_b\ \alpha=n_\alpha}\right]^T,$$

$$u = \left[u^T_{b=1\ \alpha=1}\ u^T_{b=1\ \alpha=2}\ \cdots\ u^T_{b=1\ \alpha=n_\alpha}\ u^T_{b=2\ \alpha=1}\ u^T_{b=2\ \alpha=2}\ \cdots\ u^T_{b=n_b\ \alpha=n_\alpha}\right]^T,$$

$$c = \left[c_{b=1\ \alpha=1}\vec{1}^T\ c_{b=1\ \alpha=2}\vec{1}^T\ \cdots\ c_{b=1\ \alpha=n_\alpha}\vec{1}^T\ c_{b=2\ \alpha=1}\vec{1}^T\ c_{b=2\ \alpha=2}\vec{1}^T\ \cdots\ c_{b=n_b\ \alpha=n_\alpha}\vec{1}^T\right].^T$$

The gradient of F_1 and the proximal operator of G_1 can be evaluated efficiently through

$$\nabla F_1(f) = A^T W^2(Af-d) + \frac{\lambda_1}{\mu}D_1^T P_{[-\mu,\mu]}(D_1 f) + \frac{\lambda_2}{\mu}D_2^T P_{[-\mu,\mu]}(D_2 f) + \gamma_1 diag(u)(f-c) + \gamma_2 diag(1-u)f$$

$$[prox_{tG_1}(f)]b\alpha = f_{b\alpha} - P_B(max(f_{b\alpha}, 0))$$

$$prox_{tG_1}(f)$$

$$= \left[prox_{tG_1}(f)^T_{b=1\ \alpha=1}\ prox_{tG_1}(f)^T_{b=1\ \alpha=2}\ \cdots\ prox_{tG_1}(f)^T_{b=2\ \alpha=1}\ \cdots\ f^T_{b=n_b\ \alpha=n_\alpha}\right]^T$$

where B is the Euclidean ball of radius $(y_3 G_{b\alpha} + y_4 G_{b\alpha}(1 - P_{b\alpha}))t$ centered at the origin and P_B is the function that projects onto B. The maximum is interpreted componentwise. With these formulas for the gradient of F_1 and the proximal operator of G_1, the submodule 1 subproblem can be solved using FISTA.

A3 Apply FISTA on Submodule 3

To apply FISTA on submodule 3, equation (6) is formulated in the canonical FISTA form (4) by defining F_2 and G_2 as:

$$F_2(u) = \frac{\gamma_1}{2}\langle q, u\rangle + \left(g_1\left\|D_1 u\right\|_1^{\mu} + g_2\left\|D_2 u\right\|_1^{\mu}\right) + g_3\left\|D_P u\right\|_1^{\mu},$$

$$G_2(u) = I_{[0,1]}(u),$$

where

$$I_{[0,1]}(u) = \begin{cases} 0 & if\ 0 \le u \le 1 \\ \infty & otherwise \end{cases},$$

$$u = \left[u^T_{b=1\ \alpha=1}\ u^T_{b=1\ \alpha=2}\ \cdots\ u^T_{b=1\ \alpha=n_\alpha}\ u^T_{b=2\ \alpha=1}\ u^T_{b=2\ \alpha=2}\ \cdots\ u^T_{b=n_b\ \alpha=n_\alpha}\right]^T,$$

$$q = \left[q^T_{b=1\ \alpha=1}\ q^T_{b=1\ \alpha=2}\ \cdots\ q^T_{b=1\ \alpha=n_\alpha}\ q^T_{b=2\ \alpha=1}\ q^T_{b=2\ \alpha=2}\ \cdots\ q^T_{b=n_b\ \alpha=n_\alpha}\right]^T,$$

$q_{b\alpha}$ is the vector defined as

$$diag(q_{b\alpha}) = diag(f_{b\alpha} - c_{b\alpha})^2 - diag(f_{b\alpha})^2.$$

The gradient of F_2 and the proximal operator of $G_2$are straightforward to evaluate

$$\nabla F_2(u) = \frac{\gamma_1}{2}q + \frac{g_1}{\mu}D_1^T P_{[-\mu,\mu]}(D_1 u) + \frac{g_2}{\mu}D_2^T P_{[-\mu,\mu]}(D_2 u).$$

$$prox_{tG_2}(u) = P_{[0,1]}(u)$$

where $P_{[0,1]}$ is the projection onto the feasible set $\{u \mid 0 \le u_{b\alpha} \le 1\ for\ all\ b, \alpha\}$.

With these formulas for the gradient of F_2 and the proximal operator of G_2, it is straightforward to solve the submodule 3 subproblem using FISTA.

References

Beck A, Teboulle M. A Fast Iterative Shrinkage-Thresholding Algorithm. Society for Industrial and Applied Mathematics Journal on Imaging Sciences. 2009; 2(1):183–202.

Bollobás B, Riordan O. Dijkstra â€™ s Algorithm. Network. 1959; 69:36114. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21282851.

Boyd S. Alternating Direction Method of Multipliers. Proceedings of the 51st IEEE Conference on Decision and Control. 2011; 3(1):1–44. Available at: http://www.nowpublishers.com/product.aspx?product=MAL&doi=2200000016.

Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision. 2011; 40(1):120–145.

Chang JY, et al. Stereotactic Body Radiation Therapy in Centrally and Superiorly Located Stage I or Isolated Recurrent Non-Small-Cell Lung Cancer. International Journal of Radiation Oncology Biology Physics. 2008; 72(4):967–971.

Dijkstra EW. A note on two problems in connexion with graphs. Numerische Mathematik. 1959; 1(1): 269–271.

Dong P, Lee P, Ruan D, Long T, Romeijn E, Yang Y, et al. 4?? non-coplanar liver SBRT: A novel delivery technique. International Journal of Radiation Oncology Biology Physics. 2013; 85(5): 1360–1366.

Dong P, Lee P, Ruan D, Long T, Romeijn E, Low DA, et al. 4π Noncoplanar Stereotactic Body Radiation Therapy for Centrally Located or Larger Lung Tumors. International Journal of Radiation Oncology*Biology*Physics. 2013; 86(3):407–413. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0360301613001636 [Accessed August 24, 2017].

Ee236C. Fast proximal gradient methods. 2013:1–32.

Grégoire V, Mackie TR. State of the art on dose prescription, reporting and recording in Intensity-Modulated Radiation Therapy (ICRU report No. 83). Cancer/Radiotherapie. 2011; 15(6–7):555–559.

King CR, et al. Long-term outcomes from a prospective trial of stereotactic body radiotherapy for low-risk prostate cancer. International Journal of Radiation Oncology Biology Physics. 2012; 82(2): 877–882.

Krayenbuehl J, Davis JB, Ciernik IF. Dynamic intensity-modulated non-coplanar arc radiotherapy (INCA) for head and neck cancer. Radiotherapy and Oncology. 2006; 81(2):151–157. [PubMed: 17055095]

Lee MacDonald R, Thomas CG. Dynamic trajectory-based couch motion for improvement of radiation therapy trajectories in cranial SRT. Medical Physics. 2015; 42(5):2317–2325. Available at: DOI: 10.1118/1.4917165 [PubMed: 25979026]

Liang J, et al. Trajectory Modulated Arc Therapy: A Fully Dynamic Delivery with Synchronized Couch and Gantry Motion Significantly Improves Dosimetric Indices Correlated with Poor Cosmesis in Accelerated Partial Breast Irradiation. International Journal of Radiation Oncology Biology Physics. 2015:1148–1156.

Neylon J, et al. A nonvoxel-based dose convolution/superposition algorithm optimized for scalable GPU architectures. Medical Physics. 2014; 41(10):101711. Available at: http://doi.wiley.com/10.1118/1.4895822 [Accessed August 24, 2017]. [PubMed: 25281950]

Nguyen D, Lyu Q, et al. A comprehensive formulation for volumetric modulated arc therapy planning. Medical Physics. 2016; 43(7):4263–4272. Available at: http://scitation.aip.org/content/aapm/journal/medphys/43/7/10.1118/1.4953832. [PubMed: 27370141]

Nguyen D, Ruan D, et al. A novel software and conceptual design of the hardware platform for intensity modulated radiation therapy. Medical Physics. 2016; 43(2):917–929. Available at: DOI: 10.1118/1.4940353%5Cnhttp://scitation.aip.org/content/aapm/journal/medphys/43/2/10.1118/1.4940353 [PubMed: 26843252]

Nguyen D, et al. Dose domain regularization of MLC leaf patterns for highly complex IMRT plans. Medical physics. 2015; 42(4):1858–70. Available at: http://scitation.aip.org/content/aapm/journal/medphys/42/4/10.1118/1.4915286%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/25832076. [PubMed: 25832076]

Noe M. The Annals of Mathematical Statistics. The Annals of Mathematical Statistics. 1930; 1(1):1–2.

Ohira S, et al. HyperArc VMAT planning for single and multiple brain metastases stereotactic radiosurgery: A new treatment planning approach. Radiation Oncology. 2018; 13(1)

Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. Medical Physics. 2007; 35(1): 310–317. Available at: http://doi.wiley.com/10.1118/1.2818738 [Accessed August 22, 2017].

Papp D, Bortfeld T, Unkelbach J. A modular approach to intensity-modulated arc therapy optimization with noncoplanar trajectories. Physics in Medicine and Biology. 2015; 60(13):5179–5198. Available at: http://stacks.iop.org/0031-9155/60/i=13/a=5179?key=crossref.2625947ada2395160c7ff0f2eef8a9d2. [PubMed: 26083759]

Parikh N, Boyd S. Proximal Algorithms. Foundations and Trends in Optimization. 2013; 1(3):123–231.

Rao M, et al. Comparison of Elekta VMAT with helical tomotherapy and fixed field IMRT: plan quality, delivery efficiency and accuracy. Medical physics. 2010; 37(3):1350–1359. [PubMed: 20384272]

Shaitelman SF, et al. Continuous arc rotation of the couch therapy for the delivery of accelerated partial breast irradiation: A treatment planning analysis. International Journal of Radiation Oncology Biology Physics. 2011; 80(3):771–778.

Sharfo AWM, et al. VMAT plus a few computer-optimized non-coplanar IMRT beams (VMAT+) tested for liver SBRT. Radiotherapy and Oncology. 2017; 123(1):49–56. [PubMed: 28341061]

Sheng K, Shepard DM, Orton CG. Noncoplanar beams improve dosimetry quality for extracranial intensity modulated radiotherapy and should be used more extensively. Medical Physics. 2015; 42(2):531–533. Available at: DOI: 10.1118/1.4895981 [PubMed: 25652473]

Smyth G, et al. Trajectory optimization for dynamic couch rotation during volumetric modulated arc radiotherapy. Physics in Medicine and Biology. 2013; 58(22):8163–8177. Available at: http://stacks.iop.org/0031-9155/58/i=22/a=8163?key=crossref.bc2995836dfb531b0d4154da1bd70c05. [PubMed: 24200876]

Tran A, et al. Treatment planning comparison of IMPT, VMAT and 4Π radiotherapy for prostate cases. Radiation Oncology. 2017; 12(1)

Verbakel WFAR, et al. Volumetric Intensity-Modulated Arc Therapy Vs. Conventional IMRT in Head-and-Neck Cancer: A Comparative Planning and Dosimetric Study. International Journal of Radiation Oncology Biology Physics. 2009; 74(1):252–259.

Wild E, et al. Noncoplanar VMAT for nasopharyngeal tumors: Plan quality versus treatment time. Medical Physics. 2015; 42(5):2157–2168. Available at: DOI: 10.1118/1.4914863%5Cnhttp://doi.wiley.com/10.1118/1.4914863 [PubMed: 25979010]

Wolff D, et al. Volumetric modulated arc therapy (VMAT) vs. serial tomotherapy, step- and-shoot IMRT and 3D-conformal RT for treatment of prostate cancer. Radiotherapy and Oncology. 2009; 93(2):226–233. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0167814009004472 [Accessed August 22, 2017]. [PubMed: 19765846]

Woods K, et al. Viability of Noncoplanar VMAT for liver SBRT compared with coplanar VMAT and beam orientation optimized 4π IMRT. Advances in Radiation Oncology. 2016; 1(1):67–75. [PubMed: 27104216]

Yang Y, et al. Choreographing couch and collimator in volumetric modulated arc therapy. International Journal of Radiation Oncology Biology Physics. 2011; 80(4):1238–1247.

Yu CX. Intensity-modulated arc therapy with dynamic multileaf collimation: an alternative to tomotherapy. Physics in medicine and biology. 1995; 40(9):1435–1449. [PubMed: 8532757]

Yu VY, et al. The development and verification of a highly accurate collision prediction model for automated noncoplanar plan delivery. Medical Physics. 2015; 42(11):6457–6467. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4608969/. [PubMed: 26520735]

Yu VY, et al. A Prospective 4π Radiation Therapy Clinical Study in Recurrent High-Grade Glioma Patients. International Journal of Radiation Oncology • Biology • Physics. 2018; 101(1):144–151. Available at: DOI: 10.1016/j.ijrobp.2018.01.048

Zhu L, et al. Using total-variation regularization for intensity modulated radiation therapy inverse planning with field-specific numbers of segments. Physics in Medicine and Biology. 2008; 53(23):6653–6672. Available at: http://stacks.iop.org/0031-9155/53/i=23/a=002?key=crossref.0031cfaebafb43bc3cf895696db4929d. [PubMed: 18997262]

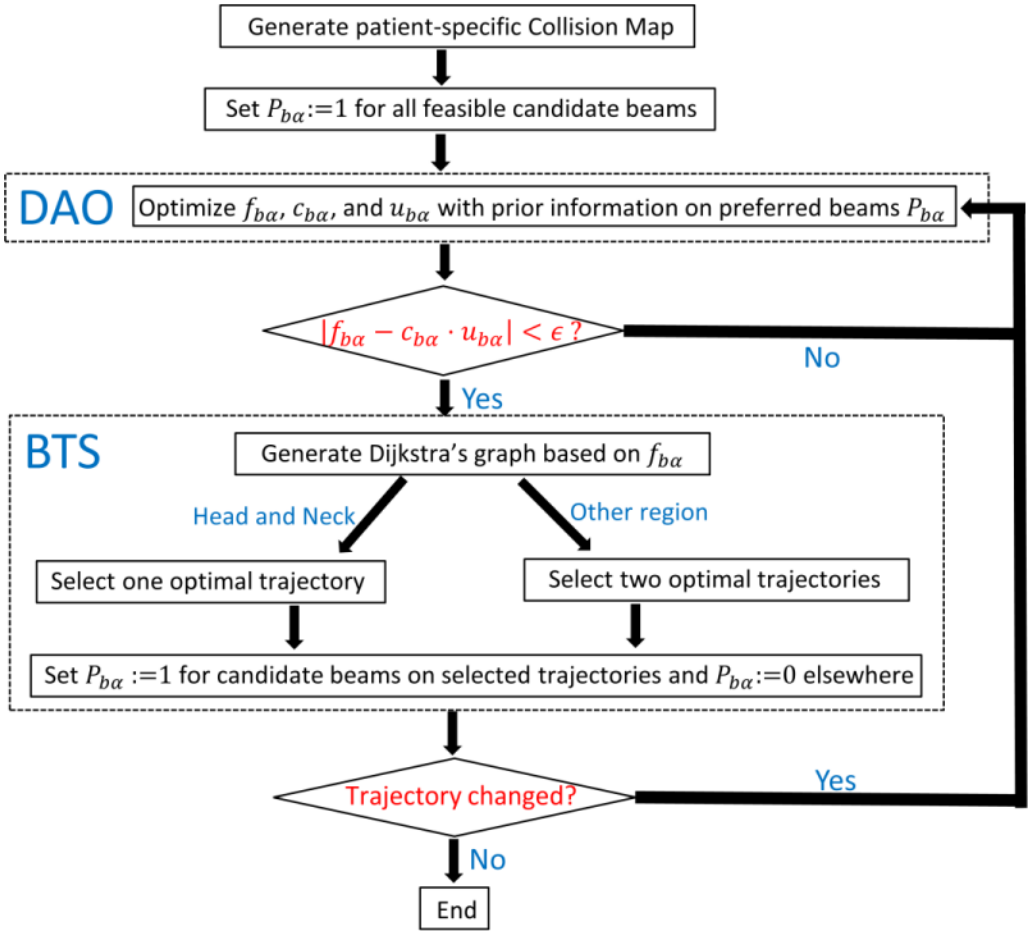

Figure 1.
Flowchart of 4πVMAT optimization

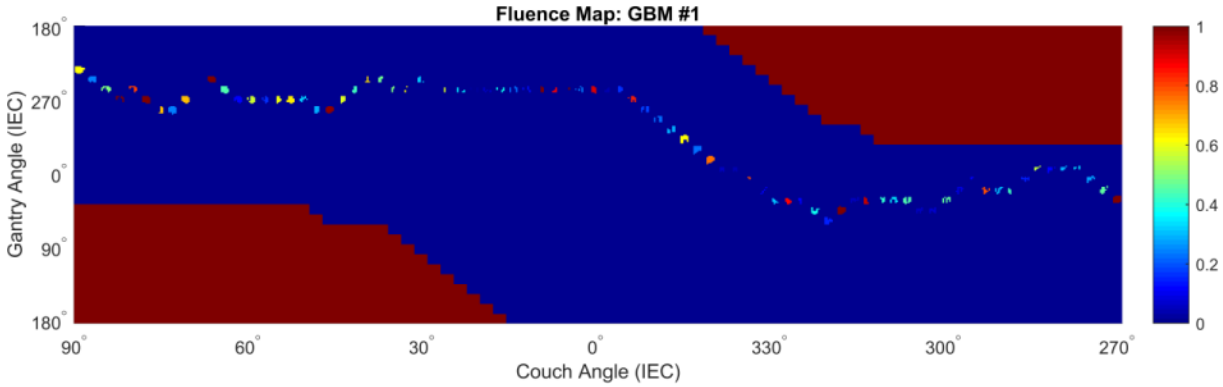

Figure 2.

Normalized fluence map on the Gantry/Couch graph (GBM #1). Only those candidate beams that are on the selected trajectory have nonzero fluence weights. The couch rotates from 90° to 270° and the gantry rotates accordingly. MLC leaf direction is vertical for this diagram. The red regions denote the gantry/couch angles that cause collision. The relative intensities of apertures are indicated by the colorbar.

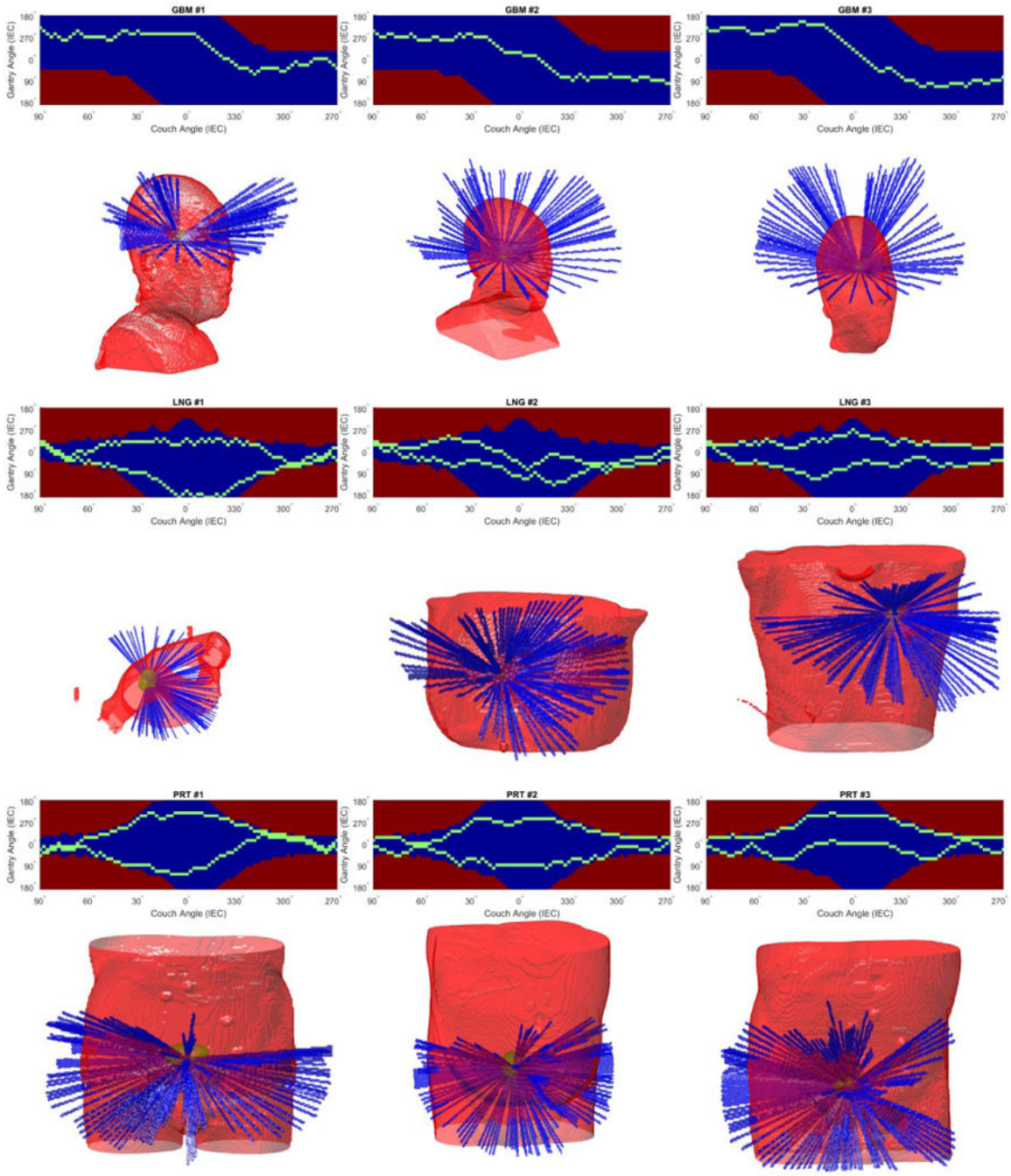

Figure 3.

The selected beam angles on the Gantry/Couch graph with corresponding 3D view for all patients. The red regions on the Gantry/Couch graph indicate the candidate beams that cause collision. The green blocks show the selected trajectories.

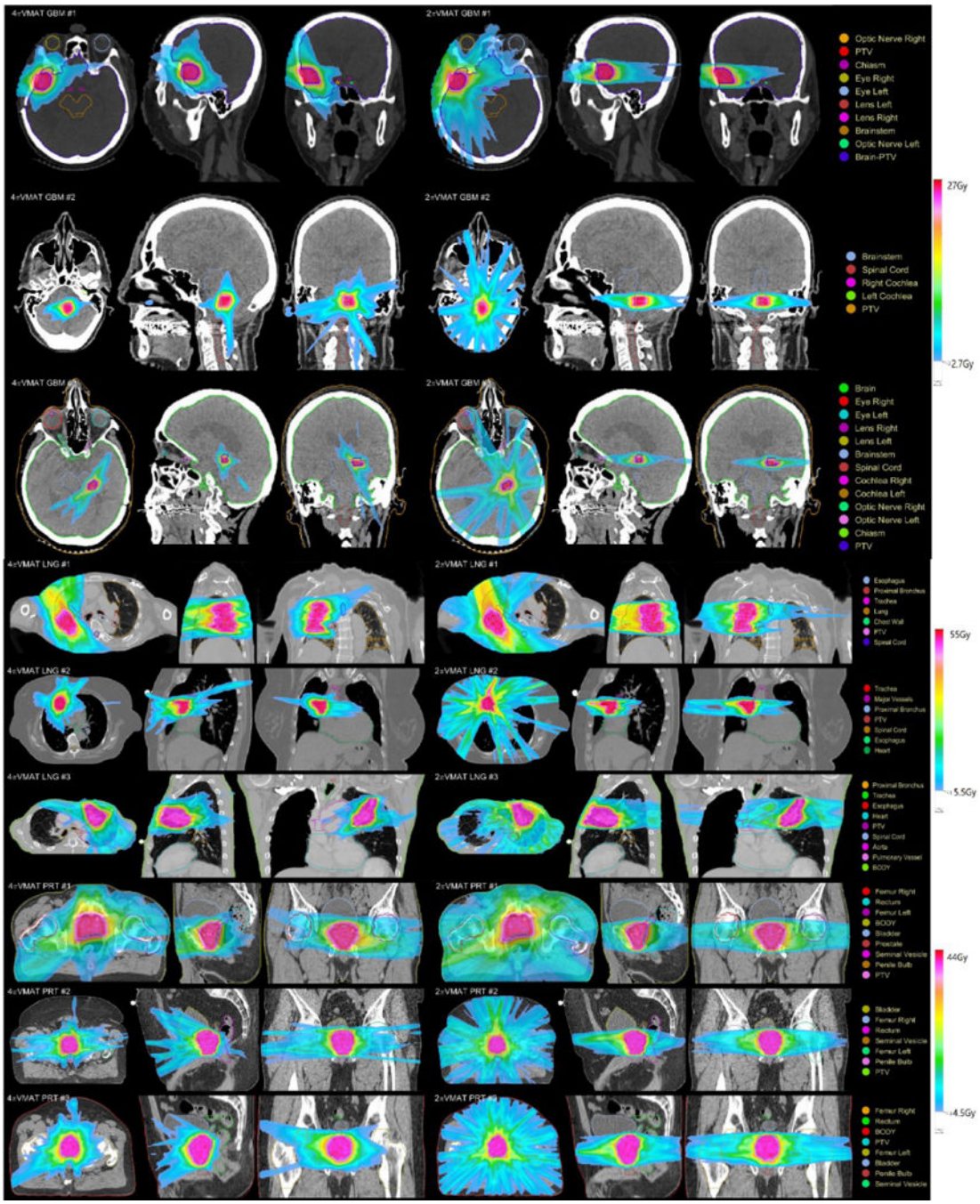

Figure 4.
Isodose colorwash comparison for all patients.

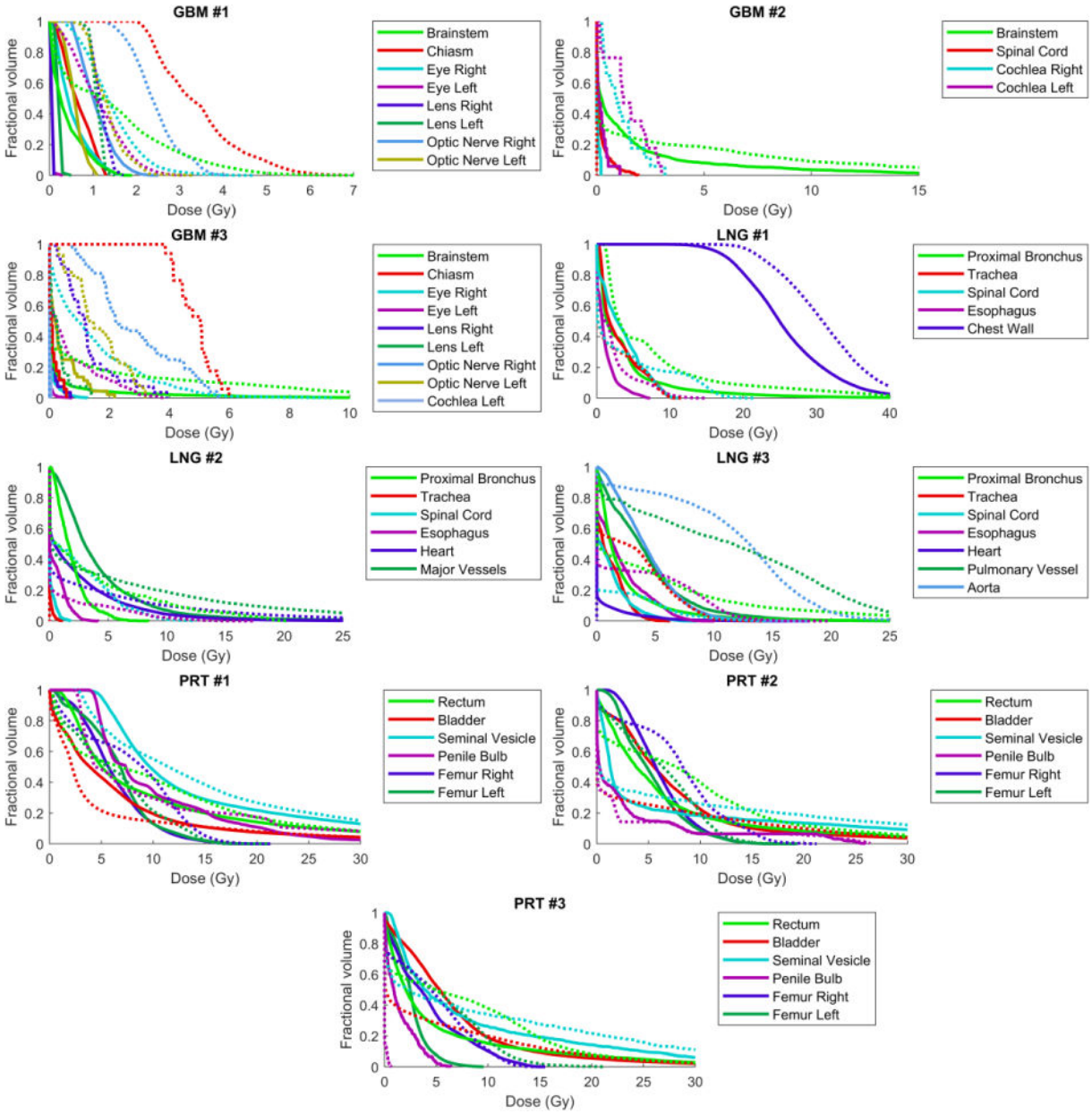

Figure 5.

DVH comparison of the 4πVMAT (solid) and the 2πVMAT (dotted) for all OARs of the patients.

Table 1

Notations and data structures used in this study

Notation	Type	Description
Indices		
b	Index	Index for couch angle, b = 1,2, . ., n_b
α	Index	Index for gantry angle, α = 1,2, . ., n_α
Optimization Variables		
$f_{b\alpha}$	Vector	Vectorized fluence map of the (b_{th}, α_{th}) candidate beam at couch angle b and gantry angle α
$c_{b\alpha}$	Vector	Fluence intensity that $f_{b\alpha}$ approaches within the aperture of the (b_{th}, α_{th}) candidate beam
$u_{b\alpha}$	Vector	Aperture variable for the (b_{th}, α_{th}) candidate beam Defined as 1 where the aperture exists and 0 elsewhere
u	Vector	Aperture variable that indicates MLC leaf positions for all candidate beams $u = \left[u_{b=1}^{T}{}_{\alpha=1} \; u_{b=1}^{T}{}_{\alpha=2} \cdots u_{b=1}^{T}{}_{\alpha=n_\alpha} \; u_{b=2}^{T}{}_{\alpha=1} \; u_{b=2}^{T}{}_{\alpha=2} \cdots u_{b=n_b}^{T}{}_{\alpha=n_\alpha}\right]^T$
Other data		
S	Set	A set of all feasible candidate beams
$A_{b\alpha}$	Matrix	Fluence to dose transformation matrix for the (b_{th}, α_{th}) candidate beam
W	Matrix	Diagonal weighting matrix, with weightings for structures of interest as diagonal elements
d_0	Vector	Ideal dose with the prescription dose at the PTV and zero elsewhere
$D_{1_{b\alpha}}$	Matrix	Derivative matrix in the direction parallel to the MLC leaf movement for the (b_{th}, α_{th}) candidate beam
$D_{2_{b\alpha}}$	Matrix	Derivative matrix in the direction orthogonal to the MLC leaf movement for the (b_{th}, α_{th}) candidate beam
$P_{b\alpha}$	Scalar	$P_{b\alpha}$ is 1 for candidate beams on the selected trajectory from BTS and 0 elsewhere
I_{PTV}	Matrix	Indicator diagonal matrix for PTV, with its diagonal elements equal to 1 for voxels in PTV and 0 elsewhere
$n_{b\alpha}$	Scalar	Number of beamlets with a trajectory that intersects PTV in the (b_{th}, α_{th}) candidate beam Weightings of the group sparsity term for each feasible candidate beam to compensate for
$G_{b\alpha}$	Scalar	unfair penalization on candidate beams $G_{b\alpha} = \frac{sum(I_{PTV} A_{b\alpha} \vec{1})}{\sqrt{n_{b\alpha}}}$
D_p	Matrix	Derivative matrix in the direction of the selected trajectory, which calculates the difference in MLC leaf positions between adjacent beams on the selected trajectory.

Table 2

The range of the hyper parameters used in this study

λ_1	λ_2	y_{10}	y_{20}	y_3	g_1	g_2	g_3
10–1000	0.2–20	0.05–50	0.05–50	50–2000	$10^{-5} - 10^{-3}$	$10^{-5} - 10^{-3}$	$10^{-5} - 10^{-3}$

Table 3

Number of feasible beams, prescription doses for each fraction, and PTV volumes for all patients

	Number of feasible beams	Prescription Dose (Gy/fx)	PTV Volume (cc)
GBM1	1824	25/5	6.23
GBM2			2.34
GBM3			0.77
LNG1			139
LNG2	1174	50/4	10.2
LNG3			116
PRT1			111
PRT2	1200	40/5	127
PRT3			85

Table 4

PTV statistics for all patients

Patient Case	PTV Statistics						Dose Conformity		R50		Integral Dose (10^3 *Gy·cm*^{3})	
	Homogeneity		D95	D98	D99	Dmax						
	4π VMAT	2π VMAT	4πVMAT – 2πVMAT (Gy)				4π VMAT	2π VMAT	4π VMAT	2π VMAT	4π VMAT	2π VMAT
GBM1	0.97	0.96	−0.01	0.16	0.31	−0.11	0.95	0.95	3.84	5.02	2.34	2.06
GBM2	0.97	0.95	0.02	0.03	0.09	−0.56	0.95	0.94	3.52	5.20	2.13	1.49
GBM3	0.94	0.94	0.00	0.05	0.04	−1.06	0.94	0.95	4.74	6.84	0.85	0.87
LNG1	0.96	0.95	0.00	−0.05	−0.20	−0.33	0.95	0.95	5.36	6.98	74.60	83.13
LNG2	0.97	0.94	−0.01	0.33	0.82	−1.50	0.95	0.95	3.77	6.77	15.17	15.99
LNG3	0.94	0.94	−0.01	−0.08	−0.27	−0.04	0.95	0.95	2.37	2.87	36.01	43.86
PRT1	0.95	0.95	0.00	−0.16	−0.26	−0.06	0.95	0.95	3.25	3.72	47.12	47.26
PRT2	0.94	0.93	0.00	−0.20	−0.28	−0.21	0.95	0.95	2.19	2.51	57.24	53.48
PRT3	0.94	0.95	−0.01	−0.27	−0.66	0.11	0.95	0.95	1.45	1.62	34.42	31.28

Table 5

OAR mean and maximum dose sparing differences between 4πVMAT and 2πVMAT for all patients.

OAR dose sparing: 2πVMAT - 4πVMAT (Gy)				
	Dmax		Dmean	
	Largest Value	Average Value	Largest Value	Average Value
GBM1	3.21 Brainstem	1.88	2.33 R Opt Nrv	1.17
GBM2	6.68 Brainstem	2.51	1.25 L Cochlea	0.77
GBM3	8.10 Brainstem	3.63	2.85 R Opt Nrv	1.49
LNG1	16.3 Proximal Bronchus	6.39	5.29 Chest wall	1.81
LNG2	19.83 Major Vessels	10.11	2.13 Spinal Cord	0.88
LNG3	15.48 Proximal Bronchus	7.19	7.46 Aorta	2.77
PRT1	4.20 Penile Bulb	1.09	1.72 R Femur	0.14
PRT2	3.80 R Femur	1.09	2.30 Rectum	0.56
PRT3	8.22 L Femur	0.70	2.70 L Femur	0.53

The columns labelled “Largest Values” represents the largest amount of dose sparing difference achieved among all OARs, and the corresponding OAR. “Average values” represents the average sparing difference among all OARs