

**Incorporating deliverable monitor unit constraints into spot intensity
optimization in IMPT treatment planning**

5

Wenhua Cao, Gino Lim

Department of Industrial Engineering, University of Houston, Houston, Texas 77204

Xiaoqiang Li, Yupeng Li, X. Ronald Zhu, Xiaodong Zhang^{a)}

Department of Radiation Physics, The University of Texas M. D. Anderson Cancer Center,

10

Houston, Texas 77030

15

20

a) Author to whom all correspondence should be addressed. E-mail: xizhang@mdanderson.org

Purpose: This study is to investigate the feasibility and impact of incorporating deliverable monitor unit (MU) constraints into spot intensity optimization in intensity modulated proton therapy (IMPT) treatment planning.

Methods: The current treatment planning system (TPS) for IMPT at our institution disregards the deliverable MU constraints in the spot intensity optimization (SIO) routine. It performs a post-processing procedure on an optimized plan to enforce deliverable MU values required by the spot scanning proton delivery system. This procedure may create significant dose distribution deviation between optimized and post-processed deliverable plans, especially when small spot spacings were used. In this study, we introduce a two-stage linear programming (LP) approach to optimize spot intensities and constrain deliverable MU values simultaneously, i.e., a deliverable spot intensity optimization (DSIO) model. Thus the post process is eliminated and its deteriorating impact on optimized plan can be avoided. Four prostate cancer cases selected at our institution were studied and two parallel opposed beam angles were planned for all cases. A quadratic programming (QP) based model without MU constraints, i.e., a conventional spot intensity optimization (CSIO) model, was also implemented to emulate the commercial TPS. Plans optimized by both the DSIO and CSIO models were evaluated for five different settings of spot spacing from 3 mm to 7 mm.

Results: Results demonstrated that the DSIO optimized plans yielded better uniformity for the target dose coverage and critical structure sparing than the CSIO optimized ones for all spot spacings. With reduced spot spacings, more significant improvement on target dose uniformity and critical structure sparing was observed comparing the DSIO to the CSIO optimized plans. It is also found that better sparing on the rectum and the bladder were achieved when smaller spacings is used for the DSIO optimized plans.

Conclusions: The proposed deliverable spot intensity optimization approach ensures optimized IMPT plans satisfying deliverable MU constraints. This eliminates the post-processing procedure required by the TPS at our institution as well as the resulting deterioration impact on ultimate dose distributions. Therefore, this approach may allow IMPT plans to adopt all possible spot spacings robustly. Studies on

prostate cancer cases in this work have shown that dosimetric benefits could be achieved by using smaller spot spacings.

Keywords: IMPT, spot scanning, monitor unit constraints, linear programming.

I. INTRODUCTION

The current state-of-the-art intensity modulated proton therapy (IMPT) is achieved by the active spot scanning technique. In this modality, a pencil beam (spot or beamlet) can be magnetically scanned in two-dimensional (2D) directions perpendicular to the beam direction to form an irradiating field. By charging protons with different energies, pencil beams can be used to penetrate with different depths and “scan” the entire designated target volume. The delivery of the sum of all Bragg peaks individually modulated is thus sought to create highly conformal dose distributions to cover the three-dimensional (3D) tumor target^{1,2}. The target volume can be divided into a series of layers from each treatment field. Those layers can be termed as energy layers because the depth of a layer (from the surface of the patient) is controlled by the energy level of a proton beam. The modulation of proton energy allows a spot to reach each of the layers. The scan is performed from the layer with the highest energy to the lowest. The scanning scheme can be either continuous or discrete. The continuous scanning system can continuously sweep a beam at a raster manner, while the discrete scanning system adopts a stop-and-shoot process by which the beam is turned off between spots³. The spot scanning system used at MD Anderson Cancer Center (MDACC) employs the discrete scanning scheme and is capable of generating protons with 94 non-equi-spaced energies from 72.5 MeV to 221.8 MeV using a scanning nozzle (Hitachi, Ltd., Tokyo, Japan and Hitachi America Ltd., Tarrytown, NY)⁴. Proton beams ranges from 4 cm to 30.6 cm in steps of 1 mm for lower energies and up to 6 mm for the higher energies. This study focuses on the discrete scanning system at MDACC. Other discrete scanning systems would share similar principles of this modality.

The arrangement of spatial positions of spots is predetermined to cover the target volume in the current treatment planning system (Eclipse, Varian Medical Systems, Palo Alto, CA) at MDACC. In all energy layers, a set of discrete spots are located with a defined spot spacing between one spot and the next. The default setting of Eclipse assigns the value of spot spacing for each field as a fraction of the spot size for the highest energy used for the field. The spot spacing can range from approximately 6 mm (spot size

in approximately 13 mm) for the 221.8 MeV energy to approximately 16 mm (spot size in approximately 35 mm) for the 72.5 MeV energy. Based on the arranged spots, an optimization process is performed to optimize the intensity of each spot from each treatment field⁵.

Studies have shown that a smaller spot spacing may increase target dose homogeneity and lower organ-at-risk (OAR) dose, but it results in many low-intensity spots and reduces plan robustness⁶. There are minimum monitor unit (MU) constraints for delivering each pencil beam (spot) for the scanning spot system. An MU is defined by a fixed number of output pulses from the main dose monitor ion chamber in the scanning nozzle⁴. Hence an MU value is used to represent spot intensity. The minimum MU (0.005 at MDACC) has to be set to ensure delivery accuracy. However, deliverable minimum MU constraints are not considered in the current treatment planning system. Instead, a post-optimization process is performed to satisfy those constraints. MU values over 0.0025 are rounded up to 0.005 and ones below 0.0025 are rounded down to 0. Rounding errors of the post process can result in significant distortion from optimized dose distributions to delivered ones⁴. The distortion becomes worse if there are more spots with small MU values which can be caused by small spot spacing. Therefore, a threshold value for spot spacing needs to be set in order to resolve the tradeoff between dosimetric advantage and delivery robustness when designing a treatment plan. Using spot spacings less than the threshold value is hence avoided in designing IMPT plans due to the dose distribution deterioration impact of the rounding errors.

Although incorporating MU constraints into inverse treatment planning for IMPT has not been well discussed, the problem of limiting excessive MUs has been extensively studied in inverse treatment planning for intensity modulated radiation therapy with photons (IMXT), despite the distinction of delivery modality between IMPT and IMXT. Most of previous studies focused on including beam segmentation constraints in the IMXT optimization process so that more continuous fluence maps can be generated in an optimized plan⁷⁻¹¹. Coselmon et al specifically discussed a strategy that assigns maximum intensity limits in IMXT optimization to improve delivery efficiency without significantly degrading plan quality¹². Overall, approaches proposed for IMXT optimization have to address the deliverability of

multi-leaf collimator (MLC), hence are not applicable to IMPT in which intensities for all proton scanning spots can be modulated independently. Therefore, only a minimum MU constraint is needed to guarantee the deliverability of IMPT spot scanning. The intensity optimization problem studied in this work focuses on a linear program with a piecewise convex objective function. Similar models have been discussed for both IMXT and IMPT optimization^{13, 14}. One of the most important features of linear optimization, or so-called L1 minimization, is that the optimal solutions tend to be sparse¹⁵. This feature has been greatly used in compressed sensing and IMXT direct aperture optimization¹⁶⁻¹⁸. More recently, Jia et al also introduced an IMXT beam angle optimization method that utilizes L1 minimization to reduce the number of beam angles¹⁹.

10 We approach the deliverable IMPT optimization problem in two stages: 1) optimizing spot intensities with sparse and large values by using L1 minimization; and 2) resolving the minimization problem with imposing strict minimum MU constraints only on non-deliverable spot intensities. Thus, the LP based two-stage linear programming approach optimizes spot intensities and incorporates minimum MU constraints, simultaneously. Unlike conventional optimization used by commercial treatment
15 planning systems, this method allows any setup of smaller spot spacing to achieve dosimetric advantages with delivery robustness. The deliverable optimization model and testing cases are described in Section II. Results on important characteristics of optimized treatment plans are presented in Section III. Insights for implementing the proposed method are discussed in Section IV and the study is summarized in Section V.

II. MATERIALS AND METHODS

20 II.A. Spot intensity optimization with MU constraints

We introduce a two-stage LP model to solve the spot intensity optimization (SIO) problem with incorporating minimum deliverable MU constraints, i.e., a deliverable spot intensity optimization (DSIO) model. A subset of pre-arranged candidate spots is selected in the first-stage LP solve without MU constraints and the intensities of selected spots are optimized in the second-stage LP solve with MU

constraints. The primary LP SIO model is implemented based on our previous study on fluence map optimization^{13,20}.

The influence matrix d_{ij} is calculated based on our in-house dose calculation algorithm²¹⁻²³, where d_{ij} denotes the dose contributed by the j^{th} spot per unit weight and received by voxel i with $j \in J$, $i \in V$.

- 5 Given that V is the set of all voxels in the treatment volume and J is the set of all pre-arranged candidate scanning spots available, the total dose in voxel i is

$$D_i = \sum_j x_j \cdot d_{ij}. \quad (1)$$

Let T , S_k and N denote the set of all voxels in the target, the k^{th} OAR and normal tissues, respectively, then $V = T \cup S_1 \cup S_2 \cup \dots \cup N$. The objective function that penalizes dose deviation on target voxels and

- 10 OAR voxels is formulated as follows,

$$\min_x f(D) = \lambda_T^+ \left\| (D_{i \in T} - \theta^+ \cdot e_T)_+ \right\|_\infty + \lambda_T^- \left\| (\theta^- \cdot e_T - D_{i \in T})_+ \right\|_\infty + \sum_k \frac{\lambda_{S_k} \left\| (D_{i \in S_k} - \phi_{S_k} \cdot e_{S_k})_+ \right\|_1}{|S_k|}, \quad (2)$$

where $(\cdot)_+$ represents $\max\{\cdot, 0\}$, and e_T and e_{S_k} represent the vectors of ones. Given that θ^+ is the overdosing control parameter for target, and θ^- and ϕ_{S_k} are underdosing control parameters for target and OAR k . Also, λ_T^+ , λ_T^- and λ_{S_k} are weighting factors.

- 15 The LP model is generally flexible in adding hard constraints. Dose limits on target and OAR voxels can be assigned by the following constraints:

$$LB_T \leq D_i \leq UB_T, \quad \forall i \in T, \quad (3)$$

$$D_i \leq UB_{S_k}, \quad \forall i \in S_k, \quad (4)$$

$$x_j \geq 0, \quad \forall j \in J. \quad (5)$$

By first solving an LP (1)-(5) with appropriate parameters verified, a tentatively optimized IMPT plan can be created. Meanwhile, only a subset of pre-arranged spots are selected, i.e., optimized spot intensities greater than zero; and the rest of spots are dropped, i.e., optimized spot intensities equal to zero. Assume that J_1 is the set of the selected spots and J_0 is the set of the dropped spots, i.e., $J_1 = \{j : x_j \geq 0, j \in J\}$ and $J_0 = J \setminus J_1$. In order to ensure the minimum MU constraint, a second LP (1)-(7) can be formed by adding two constraints to the first LP:

$$x_j \geq LB_{MU}, \quad \forall j \in J_1. \quad (6)$$

$$x_j \leq \varepsilon, \quad \forall j \in J_0. \quad (7)$$

where LB_{MU} is the lower bound for spot intensity, which is a fraction of the real minimum MU value; and ε is a very small positive value such that the feasibility of the model can be guaranteed. In fact, the feasibility would remain in the second LP mainly because only a minimal number of spots are constrained by (6) and (7), and the majority of spot intensities satisfy the MU constraint preemptively. Note that, although the two-stage LP model is solved consecutively, most of the CPU time is spent on the first stage whereas the second stage runs fast thanks to a warm start strategy.

We compare the DSIO model with an conventional spot intensity optimization (CSIO) model²⁴ defined as

$$\min_x f(D) = \lambda_T^+ \left\| (D_{i \in T} - \theta^+ \cdot e_T)_+ \right\|_2^2 + \lambda_T^- \left\| (\theta^- \cdot e_T - D_{i \in T})_+ \right\|_2^2 + \sum_k \lambda_{S_k} \left\| (D_{i \in S_k} - \phi_{S_k} \cdot e_{S_k})_+ \right\|_2^2. \quad (8)$$

The least-square based objective function penalizes dose deviation on all voxels in the target and the OARs. As an unconstrained nonlinear optimization (NLP) model, it does not contain any constraint such as minimum MU or OAR dose upper bounds. In this CSIO, scanning spots whose MU values violate the MU constraint set to be post-processed to ensure a deliverable IMPT treatment plan as the current TPS does.

II.B. Patient planning setup

We studied four representative prostate cancer cases with prostate glands ranging from small to large volumes (see Table 1). The patients have been treated with IMPT at MDACC previously. The prescribed dose is 78 Gy for 39 fractions for all patients. Two parallel-opposed lateral fields were used in this study.

One is from the right and the other is from the left. Table 1 lists the settings of proton ranges and energy levels. Volumes of the scanning treatment volume (STV)⁴ are also listed in Table 1. Intensities of scanning spots from both fields are simultaneously optimized by both the DSIO model and the CSIO model. Treatment plans with Eclipse default spot spacings (a fraction of spot size, between 6 and 7 mm varying by different patients) and uniform spot spacings (3, 4, 5, 6 and 7 mm) were created to evaluate the proposed methods. In this study, LP models were solved using the interior point method in CPLEX v12.1. The NLP model is solved using the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method that was implemented in our previous study²¹. According to our previous study⁴, the post processing on spot intensities over the maximum MU threshold has no evident impact on deviating spot scanning IMPT plan quality. We only consider minimum MU constraint in this study.

III. RESULTS

III.A. Dose volume measurements on CSIO and DSIO optimized plans

Treatment plans with different settings of spot spacing optimized by both DSIO and CSIO models were compared for all four prostate patients in this study. All plans were normalized with 97.5% prescribed dose for the STV. A comparison of dose volume histograms (DVHs) of plans with the Eclipse

default spot spacing (6.07 mm for the right field and 6.13 for the left field) for the patient case 1 is shown in Figure 1. For the DSIO optimized plans, not only the uniformity on the STV coverage but also the dose sparing to the rectum and the bladder were better than the CSIO optimized plans. For other patient cases and other spot spacings, we observed the dosimetric advantage of DSIO model was consistent.

5 Figure 2 compares the DSIO and CSIO optimized plans with 3 mm spot spacing for the patient case 2. It is clear that the STV dose uniformity and the rectum and bladder sparing were significantly improved by the DSIO method comparing with CSIO. This improvement was increased in the 3 mm spot spacing scenario than the default spot spacing scenario (Figure 1). Plans with all different uniform spot spacings are illustrated in the next section.

10 **III.B. Treatment planning with decreased spot spacing**

 Figure 3 shows DVHs for the STV in both DSIO and CSIO optimized plans with five different uniform spot spacings from 7 mm to 3 mm for one patient (case 3). With decreased spot spacings, the dose uniformity for the STV also decreased for the CSIO optimized plans, while the STV dose uniformity remained with minimal variations for the DSIO optimized plans. Dose volume statistics for the rectum and the bladder in both DSIO and CSIO plans with different spot spacings are shown in Figure 4. Average values over four patient cases are used here. The DSIO model outperformed the CSIO model on rectum and bladder sparing for all dose volume measurements (V30, V40, V50, V70) and all spot spacings. Moreover, with decreased spot spacings, each of those dose volume measurements was improved for DSIO optimized plans, while corresponding ones were degraded for CSIO optimized plans.

20 The deterioration effect of post processing for the CSIO model on final dose distributions is mainly due to the elimination of scanning spots. The number of scanning spots prior- and post-processing and the number of scanning spots associated with rounding-ups and –downs for all patient cases are illustrated in Figure 5. When the spot spacing decreased from 7 mm, numbers of pre-processed spots exponentially increased, but numbers of post-processed spots only increased at a much lower rate until the

spot spacing decreased to 5 mm. For smaller spot spacings, such as 4 and 3 mm, numbers of post processed scanning spots decreased with decreased spot spacings. Meanwhile, the number of rounding-ups and –downs generally demonstrates a positive correlation with the number of post- and prior-processing scanning spots when the spot spacing decreases.

5 Figure 6 demonstrates numbers of scanning spots used in each energy layers in one field for plans with different spot spacings for one patient (case 2). While spot spacing decreases and the number of available scanning spots increases, deliverable scanning spots from DSIO optimized plans consistently are small subsets of default spots. CSIO yields more deliverable scanning spots in all layers than DSIO dose when the spot spacing is relatively big, such as 7 and 6 mm. However, there are more deliverable
10 scanning spots in certain layers (lower energy ones) in the DSIO optimized plans than the CSIO ones, when the spot spacing is relatively small, such as 4 and 3 mm.

III.C. Automatic non-uniform spot arrangement

In our study, treatment plans optimized by the DSIO model exhibited an inhomogeneous distribution of spots spatially. While pre-arranged scanning spots were generally set by a uniform
15 arrangement and deliverable scanning spots in CSIO optimized plans were positioned by the non-optimal post processing, a non-uniform spot arrangement was automatically determined by the DSIO model. As an example, Figure 7 shows spot arrangements in one same energy layer (with an energy of 173.7 MeV from the right field) in pre-optimized and CSIO and DSIO optimized plans with 7, 5, 3 mm spot spacings for one patient (case 2). For uniformly pre-arranged scanning spots prior to spot intensity optimization,
20 they were located more densely when spot spacing decreased. In CSIO optimized plans, more reductions of deliverable scanning posts were seen when spot spacing decreased. A majority of spots were dropped with the rest ones only positioned on the edge of the energy layer when spot spacing is 3 mm. In DSIO optimized plans, with the decreased spot spacing, the number of spots showed no pronounced increase

and the inhomogeneity of spot arrangements also remained unchanged. Note that all energy layers in different patient cases exhibited a similar pattern as shown in Figure 7.

IV. DISCUSSION

Deliverable MU constraints can be robustly included in a two-stage IMPT optimization approach whereas they are not considered within the optimization process in the current mainstream treatment planning system. Thus a post-optimization process for rounding non-deliverable MU values is no longer needed in the proposed approach. The distortion from the optimized to the delivered dose distribution caused by rounding errors can be avoided. It is recommended that the spot spacing needs to be at least 6 mm to prevent excessive rounding errors for prostate cancer patients treated by two lateral fields at MDACC⁴. By applying the proposed deliverable optimization approach, different settings of spot spacing can be realized in IMPT planning to improve plan quality. Studies on prostate cancer patients with the two lateral fields have shown that as the smaller spot spacing was used, improved target dose uniformity and OAR sparing was achieved. Most importantly, more complex dose distributions may require more scanning spots, i.e., smaller spot spacing, to allow more degrees of freedom for modulation. The proposed DSIO approach may provide an important tool to improve IMPT plan quality for more complex patient cases besides prostate studied in this work. Note that optimal spot spacing is still unknown in IMPT treatment planning. A planner currently has to perform a trial and error in search for the best spot spacing to maximize dosimetric benefits for particular patient cases. With the proposed approach, the planner may safely choose smaller spot spacing as long as manageable optimization time is retained.

In IMPT treatment planning, the spatial arrangement of scanning spots or Bragg peaks in a tumor target is critical to treatment plan quality. Different spot intensity optimization, including both mathematical programming models and solution algorithms, may result in different characteristics of optimal solutions. Nonlinear models with least-square or L2 norm based objectives and gradient-based

algorithms are prevailing for optimizing spot intensity in IMPT treatment planning in clinical practices²⁵⁻²⁷, which is a similar application in conventional IMXT treatment planning^{24, 28-31}.

However, we have demonstrated that intensities of all scanning spots pre-arranged with uniform spot spacing are optimized with positive values by such nonlinear approaches both in the commercial treatment planning system⁵ and in an in-house optimization routine²¹. Optimally, these results indicate that all the pre-arranged spots are to be delivered if no MU constraint exists. In contrast, linear programming models, i.e., with L1 norm based objectives as we implemented in this study, and solvers such as interior point method that we have implemented would create comparable dose distributions as of nonlinear approaches with only a fraction of pre-arranged spots required for delivery. Or say that only a fraction of spots are optimized by the linear model with positive intensities and others are with zero intensities. The sparse solution property of L1 optimization has also been applied in regularization of fluence map variations in IMXT planning^{17, 18, 32}. The reduction of delivered spots from the predefined arrangement to the L1 optimized arrangement increases as the uniform spot spacing decreases. Thus the delivery time may be maximally reduced with coupling the optimized scanning path³³ and reduced energy layers³⁴. In addition, the delivery efficiency can also be significantly improved due to a decreased frequency of switching the pencil beam on-and-off between spot irradiations. Hence the machine lifetime can be prolonged.

Another advantage of using an LP based model to optimize spot intensities in IMPT treatment planning may be the automatic and non-uniform arrangement of scanning spots. Without designed extra optimization for spot arrangement, an inhomogeneous distribution of a subset of pre-arranged scanning spots are automatically selected to “represent” the candidate spots with a uniform spacing. Thus unnecessary spots in creating desired dose distributions can be identified and dropped for delivery when L1 optimization model is used whereas those unnecessary spots may not be identified by non-linear optimization models. Based on an assumption that more delivered spots would yield more complex dose distributions and potentially improve plan quality, it is possible to pre-arrange spots in a smaller uniform

spacing, i.e., denser spot arrangement, and to achieve better homogeneity of target dose or sparing of OAR dose. In another word, with the same amount of delivered spots, it is less likely for nonlinear models to achieve the same advantages of dose distributions as can be achieved by linear models. The addition of objectives or constraints on specific spot arrangement requirements into a conventional spot intensity optimization model, such as reducing the number of spots or arranging spots with non-uniform spacing, requires integer variables to form a mixed integer program (MIP) or a mixed integer nonlinear program (MINLP). The model complexity and computational difficulty would increase dramatically. Using a linear model such as the one implemented in this study may be with the least computational cost among different alternative methods.

10 V. CONCLUSION

We have introduced a two-stage optimization approach to optimize spot intensities and incorporate minimum deliverable MU constraints simultaneously in IMPT treatment planning. Compared with the current commercial treatment planning systems, this approach allows treatment planners to use small spot spacings to improve dosimetric performance of IMPT plans without diminishing delivery robustness. Our results demonstrated that the more dosimetric benefits of both target dose uniformity and normal sparing can be achieved, when the smaller spot spacings are used for IMPT plans for the four prostate cancer cases. The proposed approach avoids the troublesome post-processing routine required by current IMPT TPS. More importantly, the trial-and-error step for selecting an appropriate spot spacing can be eliminated. In addition, the deliverable spot intensity optimization can automatically create a non-uniform spot arrangement using only a small fraction of candidate scanning spots. We have demonstrated the potential benefits of the L1 norm based IMPT optimization approach. Future research may need to focus on reducing computational complexity of linear programs for IMPT optimization so that they can be utilized in the routine clinical treatment planning.

ACKNOWLEDGMENTS

The authors thank Radhe Mohan and Wei Liu for their helpful comments. This research is supported through National Cancer institute grant P01CA021239 and MD Anderson Cancer Center support grant CA016672. Part of this work was presented at the 54th American Association of
5 Physicists in Medicine (AAPM) Annual Meeting as an oral presentation.

Table 1 Patient and treatment planning data for four prostate cancer cases.

Patient	STV (cc)	Prescription (Gy)	Right Lateral Field			Left Lateral Field		
			Max Energy (Mev)	Nominal Range (g/cm ²)	No. of Energy Layers	Max Energy (Mev)	Nominal Range (g/cm ²)	No. of Energy Layers
1	90.3	78	198.3	25.2	19	195.6	24.6	18
2	120.1	78	193.0	24.1	20	193.0	24.1	19
3	157.3	78	195.6	24.6	23	198.3	25.2	24
4	233.2	78	203.7	26.4	21	201.0	25.8	23

References

1. A. Lomax, T. Böhringer, A. Bolsi, D. Coray, F. Emert, G. Goitein, M. Jermann, S. Lin, E. Pedroni, H. Rutz, O. Stadelmann, B. Timmermann, J. Verwey and D. C. Weber, "Treatment planning and verification of proton therapy using spot scanning: Initial experiences," *Med. Phys.* 31, 3150-3157 (2004).
2. A. Lomax, "Intensity modulation methods for proton radiotherapy," *Phys. Med. Biol.* 44, 185-205 (1999).
3. A. R. Smith, M. Gillin, M. Bues, X. R. Zhu, K. Suzuki, R. Mohan, S. Woo, A. Lee, R. Komaki, J. Cox, K. Hiramoto, H. Akiyama, T. Ishida, T. Sasaki and K. Matsuda, "The M. D. Anderson proton therapy system," *Med. Phys.* 36, 4068-4083 (2009).
4. X. R. Zhu, N. Sahoo, X. Zhang, D. Robertson, H. Li, S. Choi, A. K. Lee and M. T. Gillin, "Intensity modulated proton therapy treatment planning using single-field optimization: The impact of monitor unit constraints on plan quality," *Med. Phys.* 37, 1210-1219 (2010).
5. Varian Medical Systems, "Proton algorithm reference guide, Chapter 6: Eclipse proton optimizer for modulated scanning," P/N B500299R01C (2007).
6. D. Robertson, X. Zhang, Y. Li, A. Lee, M. Gillin, X. Zhu and R. Mohan, "Optimizing Spot Spacing and Margin for Intensity-Modulated Proton Therapy Planning," *Med. Phys.* 36, 2554 (2009).
7. J. V. Siebers, M. Lauterbach, P. J. Keall and R. Mohan, "Incorporating multi-leaf collimator leaf sequencing into iterative IMRT optimization," *Med. Phys.* 29, 952-959 (2002).
8. C. Cotrutz and L. Xing, "Segment-based dose optimization using a genetic algorithm," *Phys. Med. Biol.* 48, 2987 (2003).
9. H. E. Romeijn, R. K. Ahuja, J. F. Dempsey and A. Kumar, "A Column Generation Approach to Radiation Therapy Treatment Planning Using Aperture Modulation," *SIAM J. on Optimization* 15, 838-862 (2005).
10. J. L. Bedford and S. Webb, "Constrained segment shapes in direct-aperture optimization for step-and-shoot IMRT," *Medical Physics* 33, 944-958 (2006).
11. C. Men, X. Jia and S. B. Jiang, "GPU-based ultra-fast direct aperture optimization for online adaptive radiation therapy," *Physics in Medicine and Biology* 55, 4309 (2010).
12. M. M. Coselmon, J. M. Moran, J. D. Radawski and B. A. Fraass, "Improving IMRT delivery efficiency using intensity limits during inverse planning," *Med. Phys.* 32, 1234-1245 (2005).
13. W. Cao, G. J. Lim, A. Lee, Y. Li, W. Liu, X. R. Zhu and X. Zhang, "Uncertainty incorporated beam angle optimization for IMPT treatment planning," *Med. Phys.* 39, 5248-5256 (2012).
14. H. E. Romeijn, R. K. Ahuja and J. F. Dempsey, "A new linear programming approach to radiation therapy treatment planning problems," *Operation Research* 54, 201-216 (2006).

15. E. Candès, M. Wakin and S. Boyd, "Enhancing sparsity by reweighted l1 minimization," *Journal of Fourier Analysis and Applications* 14, 877-905 (2008).
16. R. Chartrand, "Exact Reconstruction of Sparse Signals via Nonconvex Minimization," *Signal Processing Letters, IEEE* 14, 707-710 (2007).
- 5 17. L. Zhu, T. Niu, K. Choi and L. Xing, "Total-variation regularization based inverse planning for intensity modulated arc therapy," *Technol Cancer Res Treat* 11, 149-162 (2012).
18. L. Zhu and L. Xing, "Search for IMRT inverse plans with piecewise constant fluence maps using compressed sensing techniques," *Med. Phys.* 36, 1895-1905 (2009).
- 10 19. X. Jia, C. Men, Y. Lou and S. B. Jiang, "Beam orientation optimization for intensity modulated radiation therapy using adaptive l_{2,1}-minimization," *Phys. Med. Biol.* 56, 6205 (2011).
20. G. J. Lim and W. Cao, "A two-phase method for selecting IMRT treatment beam angles: Branch-and-Prune and local neighborhood search," *European Journal of Operational Research* 217, 609-618 (2012).
- 15 21. Y. Li, X. Zhang and R. Mohan, "An efficient dose calculation strategy for intensity modulated proton therapy," *Phys. Med. Biol.* 56, N71-84 (2011).
22. Y. Li, R. X. Zhu, N. Sahoo, A. Anand and X. Zhang, "Beyond Gaussians: a study of single spot modeling for scanning proton dose calculation," *Phys. Med. Biol.* 57, 983-997 (2012).
- 20 23. X. Zhang, W. Liu, Y. Li, X. Li, M. Quan, R. Mohan, A. Anand, N. Sahoo, M. Gillin and X. R. Zhu, "Parameterization of multiple Bragg curves for scanning proton beams using simultaneous fitting of multiple curves," *Phys. Med. Biol.* 56, 7725-7735 (2011).
24. X. Zhang, H. Liu, X. Wang, L. Dong, Q. Wu and R. Mohan, "Speed and convergence properties of gradient algorithms for optimization of IMRT," *Med. Phys.* 31, 1141-1152 (2004).
25. W. Liu, X. Zhang, Y. Li and R. Mohan, "Robust optimization of intensity modulated proton therapy," *Med. Phys.* 39, 1079-1091 (2012).
- 25 26. J. Unkelbach, T. C. Y. Chan and T. Bortfeld, "Accounting for range uncertainties in the optimization of intensity modulated proton therapy," *Phys. Med. Biol.* 52, 2755-2773 (2007).
27. F. Albertini, E. B. Hug and A. J. Lomax, "The influence of the optimization starting conditions on the robustness of intensity-modulated proton therapy plans," *Physics in Medicine and Biology* 55, 2863 (2010).
- 30 28. S. V. Spirou and C.-S. Chui, "A gradient inverse planning algorithm with dose-volume constraints," *Med. Phys.* 25, 321-333 (1998).
29. D. M. Aleman, D. Glaser, H. E. Romeijn and J. F. Dempsey, "Interior point algorithms: guaranteed optimality for fluence map optimization in IMRT," *Phys. Med. Biol.* 55, 5467-5482 (2010).

30. T. Bortfeld, "Optimized planning using physical objectives and constraints," *Seminars in Radiation Oncology* 9, 20-34 (1999).
31. Q. Wu and R. Mohan, "Algorithms and functionality of an intensity modulated radiotherapy optimization system," *Med. Phys.* 27, 701-711 (2000).
- 5 32. T. Kim, L. Zhu, T.-S. Suh, S. Geneser, B. Meng and L. Xing, "Inverse planning for IMRT with nonuniform beam profiles using total-variation regularization (TVR)," *Med. Phys.* 38, 57-66 (2011).
33. J. H. Kang, J. J. Wilkens and U. Oelfke, "Demonstration of scan path optimization in proton therapy," *Med. Phys.* 34, 3457-3464 (2007).
- 10 34. J. H. Kang, J. J. Wilkens and U. Oelfke, "Non-uniform depth scanning for proton therapy systems employing active energy variation," *Phys. Med. Biol.* 53, N149-155 (2008).