

Methodology

Open Access

On the visualization of universal degeneracy in the IMRT problem

Markus L Alber*¹ and Gustav Meedt*²

Address: ¹Section for Biomedical Physics, Clinic for Radiooncology, University of Tübingen, Germany and ²Computerized Medical Systems, St. Louis, USA

Email: Markus L Alber* - markus.alber@med.uni-tuebingen.de; Gustav Meedt* - gustav.meedt@cms-euro.com

* Corresponding authors

Published: 18 December 2006

Received: 01 June 2006

Radiation Oncology 2006, **1**:47 doi:10.1186/1748-717X-1-47

Accepted: 18 December 2006

This article is available from: <http://www.ro-journal.com/content/1/1/47>

© 2006 Alber and Meedt; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: In general, the IMRT optimisation problem possesses many equivalent solutions. This makes it difficult to decide whether a result produced by an IMRT planning algorithm can be further improved, e.g. by adding more beams, or whether it is close to the globally best solution.

Results: It is conjectured that the curvature properties of the objective function around any globally optimum dose distribution are universal. This allows an assessment of optimality of dose distributions that are generated by different beam arrangements in a complementary manner to the objective function value alone. A tool to visualize the curvature structure of the objective function is devised.

Conclusion: In an example case, it is demonstrated how the assessment of the curvature space can indicate the equivalence of rival beam configurations and their proximity to the global optimum.

Background

Because of the dependence on optimisation algorithms, treatment planning for intensity modulated photon radiotherapy (IMRT) frequently evades intuition. This is particularly true for the issue of beam placement in complex cases that may even require non-coplanar beam directions, where it is hard to decide whether both the number and the directions of beams are optimum in the sense that the resulting dose distribution cannot be improved further with reasonable effort. The purpose of this note is to give an intuitive picture and a tool for visualization of a central property of the IMRT optimisation problem, which is the degeneracy of the solution space.

It is a practical wisdom of the field, that there exists a very large number of virtually equivalent solutions for an IMRT optimisation problem. Due to this degeneracy of the solution space, it is not necessary to find one solitary

best set of beam directions and fluence weight profiles, but merely a set that is equivalent (i.e. in practice: close enough) to the global optimum. It is fair to say that the true optimum of an IMRT problem is virtually never sought: usually, a few dozen iterations of a gradient algorithm are deemed sufficient, while a non-linear optimisation problem of a few thousand parameters would typically require thousands of iterations. Eventually, it is degeneracy that makes beamlet-based IMRT optimisation tractable.

The degeneracy of the objective function is inherent to the problem and does not arise spuriously from any particular mathematical formulation. In a previous paper [1], a conjecture was made about the universality of the curvature properties in optimisation problems that differed in the number and arrangement of beams. It is helpful to imagine the situation in 2D as a long, narrow valley that is

absolutely flat along its floor, but rises sharply perpendicular to it. The steep slopes on either side correspond to the high costs for increasing the dose, imposed by some normal tissue objective, and for decreasing dose, imposed by some target objective. The minimum in this direction is a finely tuned balance between these objectives. Since these objectives oppose each other for all reasonable beam arrangements, it can be assumed that the balance is sought in a very similar way for all of them. This means, that the solution space for all beam arrangements should show the same pattern of directions of great curvature, which is characteristic for the given case. It would be intriguing if this universal curvature pattern could be visualized.

In order to compare the curvature properties of the optimum dose distributions of different beam arrangements, the matrix of second derivatives for all beamlets of the union of the two beam arrangements would have to be computed. This can quickly become an onerous task. In this manuscript, we try to overcome this obstacle by introducing an independent set of arbitrary, but well chosen *probing fluence elements* which can be shared between different beam arrangements and are solely used for visualization. These probing fluence elements define a *curvature map*. In case the objective function of the optimisation problem is convex, it can be shown that all globally optimum solutions share the same subspace of zero curvature. Further, it is conjectured that the subspace of non-vanishing curvature is also an invariant. An example of this curvature map and its application in the context of manual and computerized beam direction optimisation is given for a head and neck tumour case.

Methods

The notation of this development is introduced as follows. Let $F(D)$ be a twice continuously differentiable objective function of the dose distribution $D = (D_j)_{j=1..m}$ in the patient volume V consisting of m voxels, whose minimum defines the desired solution. Further, let

$$\frac{\partial^2 F(D)}{\partial D_i \partial D_j} = 0 \text{ for } i \neq j. \text{ All the treatment goals with respect}$$

to target coverage, target dose homogeneity and normal tissue sparing are assumed to be incorporated in this objective function. For simplicity, let F be a sum of local objective function densities $f(D_j)$, so that

$F = \sum_{j=1}^m f(D_j)$. This does not restrict generality unduly, since most objective functions proposed for IMRT can be written in this form and fulfill the condition on the second derivative. For a proof see Romeijn et al. [2]. For example, the classical one- and two-sided quadratic pen-

alty functions $f(D) = (D - D_0)^2 \Theta(D, D_0)$, where $\Theta(D, D_0)$ determines whether the penalty applies to doses greater and/or smaller than a threshold D_0 , are of this shape (Here, it is convenient to postulate that Θ is not a step function, but a steep yet differentiable sigmoid so as to ensure the existence of a second derivative). Also, biological indices based on sums of local dose-response functions like EUD ($f = D_j^k, k > 0$), TCP ($f = \exp(-\alpha D_j)$) and NTCP (e.g. $f = D_j^k, k \geq 1$), as well as DVH penalties (e.g. $f = 1/(1 + (D_0/D)^k), k > 1$) can be transformed into this 'standard form' without loss of generality [3]. The dose distribution is generated by a fluence distribution, which is by definition composed of a weighted sum of elementary fluence distributions. Let \mathcal{B} be a basis set of elementary fluence distributions (fluence elements) $(\eta_i)_{i=1..n}$ and $(\phi_i)_{i=1..n} \geq 0$ the associated weights (A fluence element can, but need not be a beamlet in the common understanding as a constituent of an intensity modulated field. A beamlet is an elementary fluence distribution with the property that a finite superposition yields a deliverable intensity modulated field). For example, \mathcal{B} can be the set of all beamlets of an arrangement of intensity modulated beams, or the set of all arclets of radiosurgery arcs. The basis set defines the parameter space of the optimisation as the space of the weights that are associated with the fluence elements. By means of the basis \mathcal{B} , a subset of the abstract space of all physically feasible fluence distributions is mapped onto $\mathbb{R}_0^{|\mathcal{B}|}$. This is in analogy to geometry, where e.g. a point in space can be identified by its coordinates relative to a set of basis vectors.

Let $(T_j)_{j=1..m}$ be the (linear) dose deposition operator that associates the fluence element η_i with its dose distribution $(T_j)_{j=1..m} \eta_i = (T_{ji})_{j=1..m}$. Hence,

$$D_j = \sum_{i=1}^n T_{ji} \phi_i. \tag{1}$$

The solution of the IMRT problem is a vector of fluence weights $(\phi_i^*)_{i=1..n}$ which obtains from the minimization of F under the constraints $\phi_i \geq 0$. Given that F is assumed to be twice continuously differentiable, the necessary optimality conditions

$$\frac{\partial F(D^*)}{\partial \phi_i} = 0 \text{ or } \phi_i^* = 0 \quad \forall \eta_i \in \mathcal{B} \tag{2}$$

hold. If all f are convex, F is also convex, and by virtue of the convexity of the feasible set $\mathcal{F} = \{\phi_i \in \mathbb{R}^{|\mathcal{B}|} : \phi_i \geq 0\}$, only a single global minimum exists (generally, the dose limiting objectives ensure that F is bounded from below and $F \rightarrow \infty$ for all $\phi_i \rightarrow \infty$). However, this does *not* mean that the minimum is attained in a single point. In the case of degeneracy, as commonly present in the IMRT setting, the *set of minimizers* is a sizeable, high dimensional, closed and connected subset of the parameter space. Any such point in this solution space is indistinguishable from any other with respect to its objective function value. The uniqueness of the minimizer cannot be guaranteed even if the objective densities f are strictly convex, because $F(\phi)$ is a composition of monotonously increasing and decreasing convex functions. In practice, this causes IMRT optimisation algorithms to terminate at different solutions if started from slightly different points, which may be mistaken for local, isolated minima. The situation is different in case either f or \mathcal{F} are non-convex, which can occur when dose-volume objectives are employed or MLC constraints need to be considered. In this case the multiple global minimizers can lie in disconnected and non-convex sets, or local minima may exist. Notice that these are possibilities, not necessities.

If two different dose distributions, possibly generated by different beam arrangements, have the same objective function value, they may be degenerate to the global solution, or merely be local solutions that share the same objective function value by coincidence. By virtue of the linearity of the dose operator T , there exists an easy test to investigate this situation. Let $\mathcal{B}_1, \mathcal{B}_2$ be two fluence basis sets, for example the sets of all beamlets of two arrangements of beam directions, and D_1^*, D_2^* be the associated optimum dose distributions. Further, let $F(D_1^*) = F(D_2^*)$. For $0 \leq \lambda \leq 1$, we define

$$F(\lambda) = F(\lambda D_1^* + (1 - \lambda) D_2^*) \quad (3)$$

and compute $F(\lambda)$ for a small number of $\lambda \in [0, 1]$. This test does not entail more than the weighted addition of two dose distributions and repeated evaluations of the objective function, but no operations in fluence weight space.

If there exists one λ' with $F(\lambda') > F(D_1^*)$, then obviously both dose distributions do not belong to the same mini-

mum. This situation can only occur if F is non-convex. If F is convex, and D_1^*, D_2^* are degenerate to the global minimum, then

$$F(D_1^*) = F(D_2^*) = F(\lambda D_1^* + (1 - \lambda) D_2^*). \quad (4)$$

By definition, F is always locally convex in a certain environment of the global minimum (otherwise it would not be a minimum), which makes the following argumentation applicable to some extent also to the globally non-convex case. In the following, we use eq. 4 to motivate a conjecture about universal curvature properties of all solutions that are degenerate to the global optimum. In all but the most contrived cases, the comprehensive basis \mathcal{B}^* of all globally optimum dose distributions, i.e. the union of all basis sets that generate a global optimum, is very large. Simultaneously, a very large number of degenerate and near-degenerate solutions exist. Assume that some dose distribution D_1^* is degenerate to the global optimum D^* , i.e. $F(D_1^*) = F(D^*)$. By the above definition, $\mathcal{B}_1 \subset \mathcal{B}^*$. The optimality conditions have to be satisfied for all elements of \mathcal{B}^* . We apply the optimality conditions eq. 2 to the right hand side of eq. 4

$$\begin{aligned} \frac{\partial f(\lambda D_1^* + (1 - \lambda) D_2^*)}{\partial \phi} = \\ \sum_{j=1}^m \frac{\partial f(\lambda D_1^* + (1 - \lambda) D_2^*)}{\partial D_j} \cdot T_{ji} = 0 \quad \forall \eta_i \in \mathcal{B}^*. \end{aligned} \quad (5)$$

We expand the left hand side to first order in λ , using that

$$\begin{aligned} \frac{\partial^2 f(D)}{\partial D_i \partial D_j} = 0 \\ \text{lhs} = 0 + \lambda \sum_{j=1}^m (D_{1,j}^* - D_j^*) \frac{\partial^2 f(D^*)}{\partial D_j^2} T_{ji} + \mathcal{O}(\lambda^2) \quad \forall \eta_i \in \mathcal{B}^*. \end{aligned} \quad (6)$$

Since D_1^* is degenerate to D^* , the left hand side vanishes, i.e.

$$\sum_{j=1}^m (D_{1,j}^* - D_j^*) \frac{\partial^2 f(D^*)}{\partial D_j^2} T_{ji} = 0 \quad \forall \eta_i \in \mathcal{B}^*. \quad (7)$$

In order to interpret this condition, it is helpful to observe the Hessian matrix of second derivatives with respect to the fluence weights

$$H_{ij}(D) = \sum_{k=1}^m T_{ki}^T \frac{\partial^2 f(D)}{\partial D_k^2} T_{kj}. \quad (8)$$

Notice that H_{ij} is a symmetric, but not diagonal matrix, while the inner second derivative of the objective function with respect to the local dose is diagonal by definition. Commonly, the vast majority of eigenvalues of this matrix is zero, owing to the high degeneracy of the IMRT problem, while the solution is characterized by the (smaller) subspace of non-vanishing curvature [1]. By comparison with eq. 7, it becomes apparent that this criterion evaluates whether the difference between two degenerate global solutions $(D_1^* - D^*)_j = \Sigma_i(\phi_i^* - \phi_i)T_{ji}$ lies entirely in the subspace of zero curvature of the global optimum. Since this has to hold for any two globally optimum dose distributions, their curvature matrices $H_{ij}(D^*)$ share the same subspace of zero curvature (To see this, assume there exists a direction $\eta\Delta = D_1^* - D_2^*$ for which the curvature around D_1^* is zero, but around D_2^* is greater than zero. In a convex setting, it is possible to go from D_1^* to D_2^* along a straight line without leaving the solution space, but the gradient of F in the direction of $\eta\Delta$ has to increase when going from D_2^* to D_1^* . This is in contradiction with the assumption that both dose distributions are globally optimum).

In [1], it was conjectured that because all globally degenerate optima share the same modes of solving the fundamental conflicts of the problem between target and normal tissue objectives, the structure of the subspace of high curvature around them is universal. Together with the universality of the subspace of zero curvature, this earlier conjecture would be a consequence of the following, more general formulation: Let \mathcal{D}^* be the set of all globally optimum dose distributions and let Φ^* be the associated set of all fluence weight distributions with respect to the basis \mathcal{B}^* . Let the set Φ_0 span the subspace of zero curvature of $F(D^*)$, $D^* \in \mathcal{D}^*$, which is assumed to be locally convex around \mathcal{D}^* . *It is conjectured that all third and higher order derivatives of F with respect to any fluence vector in Φ_0 vanish for all dose distributions in \mathcal{D}^* .* The first and second order derivatives vanish by virtue of the optimality conditions and the above considerations following eq. 7.

A direct consequence of this conjecture is, that the subspace of non-vanishing curvature of $H_{ij}(D^*)$ is also an invariant and the matrix $H_{ij}(D^*)$ is constant for all dose distributions in \mathcal{D}^* . The parameter space \mathcal{F} becomes a direct product of the space of vanishing curvature \mathcal{F}_0 and the space of positive curvature \mathcal{F}_+ . In practice, three obstacles to validating this conjecture for a given example exist. Firstly, the basis set \mathcal{B}^* can only be approximated by a finite basis set. Secondly, the computation of the Hessian matrix is extremely time consuming for large basis sets, since it grows quadratically with the number of basis fluence elements. Thirdly, the conjecture cannot be proven numerically because the size of the derivative tensors grows exponentially with the order of the derivative.

In order to visualize the putatively universal curvature properties with a minimum of computational effort, the concept of the *curvature map* is introduced, which avoids these problems. Let D^* be a globally optimum dose distribution and let \mathcal{B}_p be any set of arbitrarily chosen *probing fluence elements*, then the *curvature map* (CM) is defined as

$$M(\eta_i \in \mathcal{B}_p) := \sum_{j=1}^m D_j^* \frac{\partial^2 f(D^*)}{\partial D_j^2} T_{ji} \eta_i. \quad (9)$$

The finite set of probing fluence elements is assumed to be a subset of the fictitious basis set \mathcal{B}^* and is prompted by practical reasons only. This set can be chosen freely. Notice that due to the (local) convexity of F , the CM cannot be negative. If the conjecture holds, the curvature maps of all dose distributions that are degenerate to the global optimum are equivalent.

The globally optimum solution is essentially determined by those volumes in the patient where it is hard to meet target volume or normal tissue objectives. If a probing fluence element η passes through these volumes where the dose distribution is a forced compromise, its corresponding CM-value $M(\eta)$ will be high. In contrast, if a probing fluence element passes only through volumes where the target and normal tissue objectives can be met, the CM-value will be close to zero.

One advantageous choice of probing fluence elements for beam direction optimisation could be the 'generalized Gamma knife basis set', a set of conical beams of a certain diameter of 15–20 mm say, which impinge on the isocentre from about 5000 directions distributed uniformly across the unit sphere. The curvature map would then show beam directions which pass through or avoid areas

of conflicts between dose prescriptions. There are no limits to the choice of \mathcal{B}^* other than practicality and visualization.

Results and Discussion

The example case presented here is a paranasal sinus tumour, see figure 1. The obvious obstacle for achieving the target dose is the proximity of the optical pathway and to a much smaller degree the brain. The setup of the optimisation problem included hard constraints on the maximum equivalent uniform dose (EUD) [3,4] of the organs at risk (optical pathways, brain, brainstem, unspecified tissue), and a hard constraint for the maximum dose to the target. For the target, the objective was to maximize EUD under the given constraints. The sum of these constituent objective functions, together with their appropriate Lagrange multipliers, yields the objective function F . With this combination of physical and EUD constraints, the objective function is convex. The Lagrange multipliers are fixed such that the constraints are obeyed for the solution D^* , with the consequence that the sole floating figure of merit is the EUD of the target volume. In figure 2, curvature maps for several beam configurations are shown. The CM value is colour coded, starting from zero in blue to the common maximum in red, i.e. all CMs cover the same range of curvatures. Beam directions are indicated as dots on the sphere.

Several observations can be made.

- The rows two through four show CMs of a five, six and seven IMRT beam configuration created by a beam direction optimisation (BDO) algorithm [5]. The first row

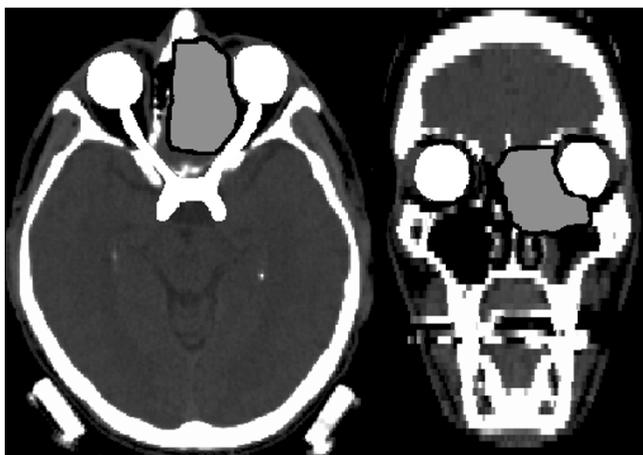


Figure 1
Geometry of the example case. Transversal and coronal sections of the example case, a paranasal sinus tumour. Target volume outlined in black/dark grey, optical pathway light grey.

shows a beam configuration which was chosen manually (6 MAN). The difference between 5 BDO and 6 BDO is somewhat greater than between 6 BDO and 7 BDO, but the general distribution of curvatures is rather constant, with a noticeable shrinkage of the total area covered in red and yellow. This impression of increasing proximity to the global optimum is also supported by their final objective function values, which were 97 per cent and 99 per cent respectively of the EUD of the 7 BDO configuration. Notice, that each configuration has no more than one beam in common. The difference between the BDO plans would be greater if they were scaled to their common maximum, rather than the maximum of 6 MAN.

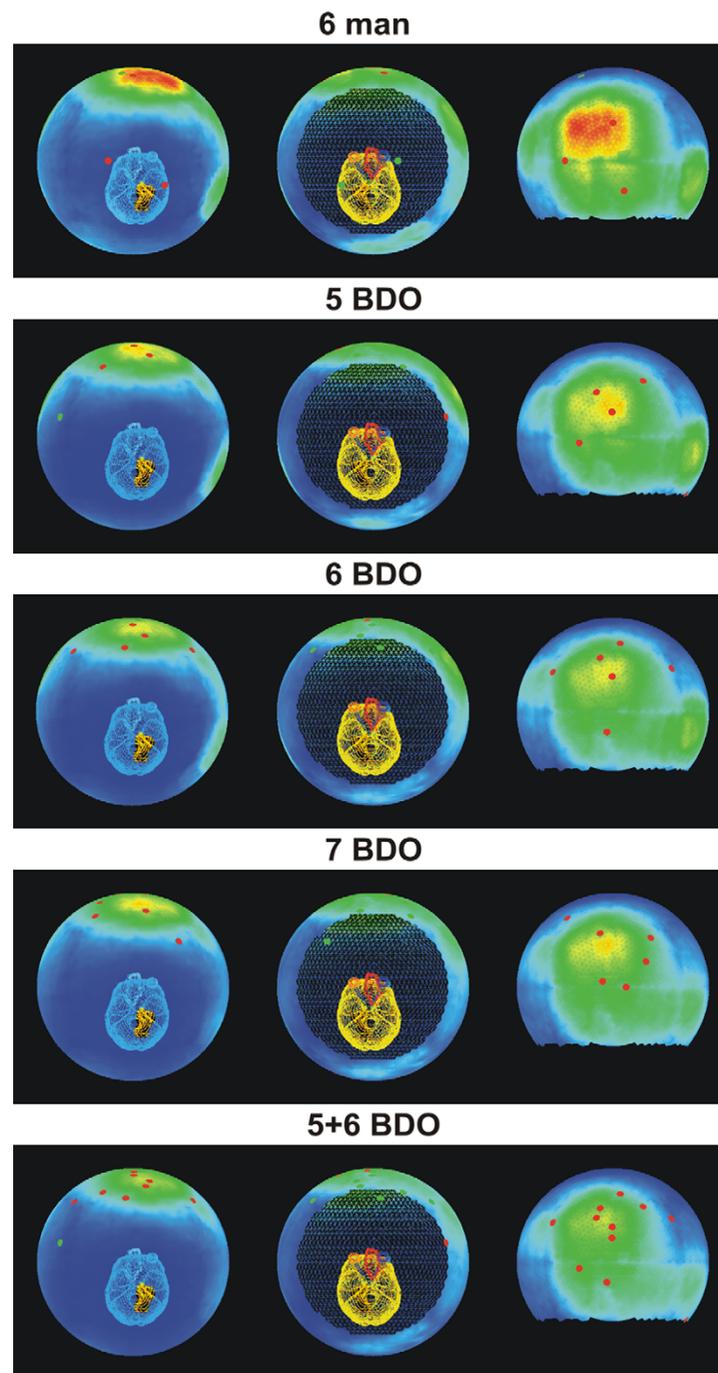
- The first row shows the CM of a manually selected 6 field configuration. The CM differs noticeably from the 7 BDO CM. At the same time, only 94 per cent of the EUD of 7 BDO could be reached. The frontal red region is more extensive in the 6 MAN configuration, which shows that a larger sector of the unit sphere is affected by conflict volumes in the patient. There is one beam direction which originates from a blue region. Blue sectors indicate that directions lie in the subspace of vanishing curvature and thus can be easily substituted. If possible, the BDO algorithm eliminates these directions. In contrast, the human expert picked this direction on the basis of intuition, which was misleading in this case.

- The fifth row shows the CM for the combined 5 BDO and 6 BDO configuration (i.e. an optimized dose distribution of 11 beams). Notice that there is virtually no difference to the 7 BDO configuration. The combined configuration yields the same EUD in the target as 7 BDO. This indicates, that about 7 beams suffice to generate a globally optimum dose distribution for this case, and that a 11 beam arrangement can contain several superfluous beams.

- Notice that in any case, since the objective function in this example is convex, at least the blue areas of the CM of two globally optimum dose distributions have to agree.

Naturally, the basis of probing fluence elements is never complete. In this example, it was chosen such that the full extent of the patient is covered. For more extensive target volumes, different choices may be more appropriate. Notice, that according to the above conjecture, the CM of globally optimum dose distributions should be equivalent for any choice of basis.

If the CM is applied in beam direction optimisation, it provides complementary information to the objective function value. Iterative beam selection has to continue as long as the CM changes between successive configurations. If two rival configurations differ in their CMs, nei-

**Figure 2**

Curvature Maps for the example case. Curvature maps of a manually optimised 6 beam configuration, a 5, 6, and 7 beam direction optimised configuration, and a 11 field configuration consisting of the union of the 5 BDO and 6 BDO configurations. The CM values were mapped onto a sphere for 5000 rays which sample all feasible directions of incidence. Left column: cranial view, second column: caudal view, third column: frontal view. The black excision corresponds to beam angles that were excluded due to couch/gantry collisions. Blue sectors correspond to a CM-value of zero, red and yellow regions display zones of conflicts between target objectives and normal tissue constraints. The colour scale is equivalent for all rows. Beam directions are given as red dots.

ther of them is truly equivalent to the global optimum. By virtue of the above conjecture, the CM contributes supplemental information in case the beam selection process gets stuck with two rival beam arrangements which produce the same objective function value, yet could be improved further. Once a final configuration was found, beams originating from blue zones are most likely redundant and can be removed from the configuration. Notice, that the CM cannot be used to guide the search for additional beam directions at intermediate stages: since it is subject to significant change while the beam configuration evolves, the indication of redundant beams is not stable. The first derivative with respect to the weight of a probing fluence element may offer more information for this purpose.

Conclusion

It is conjectured for the IMRT optimisation problem, that in the proximity of the global optimum, the second and all higher order derivatives vanish in a subspace of sizeable dimension. As a consequence, the curvature of the objective function for any globally optimum dose distribution is equivalent. In order to verify this property, the tool of a curvature map was introduced which relies on a freely choosable set of probing fluence elements. This allows to compare the curvature properties of dose distribution which have no beams in common. These considerations about curvature are intended to highlight the special properties of the IMRT optimisation problem, which may further its understanding or be used to advantage in algorithm design.

Authors' contributions

MA and GM developed the maths and prepared the manuscript. GM coded the software tools and prepared the example.

Acknowledgements

This research was supported by Computerized Medical Systems, St. Louis, USA.

References

1. Alber M, Meedt G, Reemtsen R, Nüsslin F: **On the degeneracy of the IMRT optimisation problem.** *Med Phys* 2002, **29(11)**:2584-2589.
2. Romeijn HE, Dempsey JF, Li JG: **A unifying framework for multi-criteria fluence map optimization models.** *Phys Med Biol* 2004, **49(10)**:1991-2013.
3. Alber M, Nüsslin F: **An objective function for radiation treatment optimization based on local biological measures.** *Phys Med Biol* 1999, **44(2)**:479-493.
4. Niemierko A: **Reporting, and analyzing dose distributions: A concept of equivalent uniform dose.** *Med Phys* 1997, **24**:103-110.
5. Meedt G, Alber M, Nüsslin F: **Non-coplanar beam direction optimization for intensity modulated radiotherapy.** *Phys Med Biol* 2003, **48(18)**:2999-3019.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

