Adapting radiotherapy treatment to deformations in the patient

Claudiu Antonovici (CWI Amsterdam, Leiden University), Tugce Akkaya (TU Delft), Arjen Baarsma (Utrecht University), KaYin Leung (Utrecht University, UMC Utrecht), Corine Meerman (Leiden University), Lisanne Rens (CWI Amsterdam, Leiden University)

Abstract

The physical anatomy of cancer patients who receive radiotherapy treatment may change over time due to e.g. weight loss and tumor deformation. The treatment plan for the patient is based on a CT scan that is made approximately one week prior to treatment. However, when the anatomy of the patient has changed, the treatment plan may no longer be optimal. How can adjustments to the treatment plan be made when changes in the anatomy of the patient are observed? In this paper we report on this 'cancer treatment' problem posed by the Netherlands Cancer Institute. We formulate the problem as an optimization problem. Different deformations are investigated for a toy model and optimization methods are tested on this model. We consider three different optimization algorithms, with a main focus on the Gauss-Newton method. This method turns out to work relatively well for some specific deformations in our simple model. Other methods are also considered and their advantages and disadvantages are described, keeping in mind more realistic situations than we consider here.

keywords: adaptive radiotherapy, dose planning, deformations, optimization, Gauss-Newton, Particle Swarm

1 Introduction

This paper reports on findings of the 'cancer treatment' problem, as posed by the Netherlands Cancer Institute (NKI) at the Study Group Mathematics with Industry 2015, held at Utrecht University. Every year, 469 per 100.000 inhabitants over 18 years in The Netherlands are diagnosed with cancer. Around

fifty percent of cancer patients are still alive five years after the diagnosis and the number of treatment with radiotherapy increases [5]. The Netherlands Cancer Institute (NKI) is a national center dedicated to cancer and plays an important role as a national and international center of scientific and clinical expertise, development and training. One of their ongoing research topics is on improving the efficacy of radiotherapy treatment plans.

If a person is scheduled to receive radiotherapy treatment, a CT scan is made prior to the treatment. The time between the CT scan and the actual treatment is usually between one and two weeks. The CT scan is made to obtain detailed information about the tumor and the surrounding area. Using this information, an accurate though time intensive optimization method is used to acquire the optimal settings of the machine and the corresponding dose distribution. To explain these terms, a closer look at the treatment settings is needed.

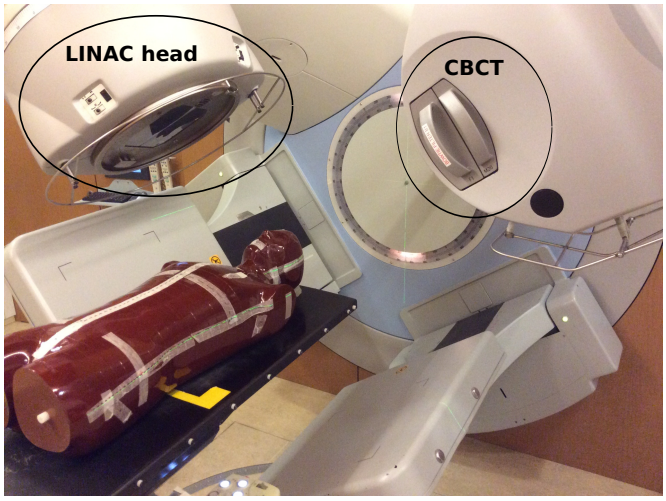

Figure 1: Treatment settings. Image courtesy of Rene Tielenburg.

Figure 1 shows the treatment set-up. In this set-up the patient lies on the treatment table. The radiation source is located in the LINAC (linear accelerators) head. Within this head there are leaves, the settings of which determine the shape of the beam that radiates the patient. This head rotates around the table. At every rotation step, the settings of the leaves within this device, are adjusted such that after a full rotation the optimal dose is delivered. An optimal treatment is defined to mostly affect the tumor, whereas the surrounding tissue, especially the vital organs, should be left alone as much as possible to

avoid damaging healthy tissue.

The actual treatment based on the treatment plan derived from the CT scan ideally starts one (but maximum two) week(s) after the CT scan was made. In this period, the anatomy of the patient may have changed, causing a lot of differences compared to when the CT scan was taken, such as changes in the size of the tumor, weight loss of the patient, posture changes and many other deformations and changes. Thus, applying the original treatment plan to the patient may lead to a sub-optimal dose distribution.

To avoid large amounts of healthy tissue being targeted by radiotherapy, the current procedure is to take another scan of lesser quality of the patient right before the treatment starts. This scan is taken with the imaging device CBCT (cone beam CT) which is part of the treatment device (Figure 1). If the outcome of this scan differs substantially from the data of the CT scan, the treatment is stopped and the entire procedure starts all over, i.e. another CT scan and subsequently a new treatment plan is made. If the deformations are small, the original treatment is applied to the patient.

NKI asked us to developed or propose a method that determines the optimal settings of the leaves using the initial correct dose, the initial data from the CT scan and the data of the CBCT. These new settings have to be obtained within a few minutes because the treatment has to be applied directly after the scan of the CBCT.

In Section 2, we formulate the problem as an optimization problem. In section 3 three different methods to solve this optimization problem are proposed to solve this type of problem: the Gauss-Newton method, evolutionary algorithms (with a focus on the Particle Swarm optimization algorithm), and a gradual deformation based approach. In Section 4, four different tumor-deformation scenarios are considered and the Gauss-Newton is used to solve the optimization problem for these scenarios.[1] The Particle Swarm optimization is applied to one of the four scenarios and compared to the result obtained with Gauss-Newton. Finally, in Section 5 conclusions and recommendations are presented.

[1] After the study week it was discovered that there was a mistake in the code that we had worked with. We corrected this afterwards and the results presented in these proceedings are the results for the corrected code. As a bonus, finding this mistake also greatly improved our results.

2 Formulation of the problem

2.1 The optimization problem

The problem put forward by the NKI is essentially an optimization problem, so we formulated it as such. More specifically, we approached it as a non-linear least squares problem of the form

$$\arg\min_{\phi} \|D(\phi, P_1) - D_{\boldsymbol{\delta}}\|^2. \tag{1}$$

Here ϕ denotes the configuration of the linear accelerator, P_1 is a set of parameters describing the patient during treatment and $D(\phi, P_1)$ is the dose delivered to this patient when the configuration ϕ is used. The dose $D_{\boldsymbol{\delta}}$, which we try to approximate, is similar to the dose absorbed by the undeformed patient under the original plan, but has been transformed using the deformation of the patient

We proceed by providing a more detailed description of the variables ϕ and P, as well as the dose function D and the target dose $D_{\boldsymbol{\delta}}$.

2.2 Variables and parameters

During a treatment session, the LINAC head moves around the patient, stopping at N_{stops} different positions along the way. At each position, the opening formed by the multileaf collimator can be adjusted by changing the positions of its $2\,N_{\text{leaves}}$ leaves. We denote the positions of the j-th pair of leaves while the head is at position i by α_{ij} and β_{ij}, and the time the head stays at position i by τ_i (Figure 2). These variables are subject to the linear constraints $\alpha_{\min} \leq \alpha_{ij} \leq \beta_{ij} \leq \beta_{\min}$ and $\tau_i \geq 0$. For the machine used by the NKI, $N_{\text{stops}} = 90$ and $N_{\text{leaves}} = 40$, which means that the total number of variables is $2\,N_{\text{stops}}\,N_{\text{leaves}} + N_{\text{stops}} = 7290$. This number can be doubled by having the LINAC head make two passes, so that it stops at every position twice.

For our purposes, the patient is effectively described by a density function $\rho(\boldsymbol{r})$ and an attenuation function $\mu(\boldsymbol{r})$. Together with the configuration of the device, these determine how much energy is absorbed at every point in the patient's body. For the sake of brevity, we write $\phi = (\alpha, \beta, \tau)$ for the full configuration of the linear accelerator and $P = (\mu, \rho)$ for the relevant data concerning the patient.

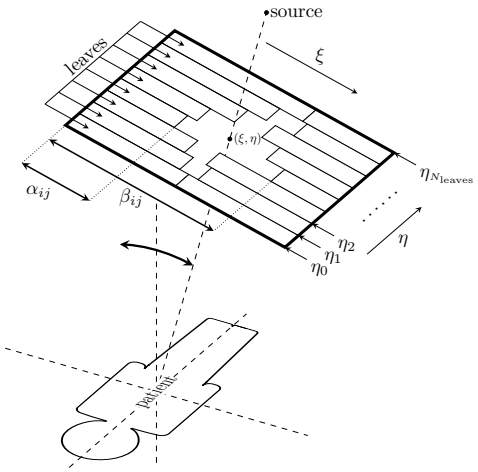

Figure 2: Overview of the leaf configuration.

2.3 The dose function

The total dose absorbed by a point $\boldsymbol{r}$ inside the patient's body during a treatment session is the sum of the contributions of all rays passing through the multileaf collimator.

If we ignore partial transmission through the rounded edges of the leaves, the total dose absorbed at a point $\boldsymbol{r}$ is given by

$$D(\phi, P)(\boldsymbol{r}) = \sum_{i=1}^{N_{\text{stops}}} \tau_i \sum_{j=1}^{N_{\text{leaves}}} \int_{\alpha_{ij}}^{\beta_{ij}} \int_{\eta_{i-1}}^{\eta_i} d_i(\xi, \eta, P)(\boldsymbol{r})\, d\eta\, d\xi, \tag{2}$$

where $d_i(\xi, \eta, P)$ is the dose delivered per unit time by an infinitesimally small (conical) beam going through the point (ξ, η) in the collimator while the emitter is at position i. The dose function is integrated over the opening formed by each pair of leaves, subsequently a sum over all leaf pairs and all emitter positions is taken.

The full expression for the infinitesimal dose $d_i(\xi, \eta, P)(\boldsymbol{r})$ also involves multiple integrals, so finding the full dose profile $D(\phi, P)$ in this way can be quite time consuming. Since this will need to be done at least several times during the optimization process and time is a relevant constraint, the use of approximate methods is in order. Two popular methods are the collapsed cone algorithm,

which is described in detail in [1], and the analytic pencil beam kernel approach detailed in [6].

We have employed the analytic pencil beam kernel approach, which is the simpler of the two. This method approximates $d_i(\xi, \eta; P)$ by a simple analytic expression. The approximation relies on a large number of assumptions, both about the patient and the emitter: the patient is assumed to be homogeneous and flat, and the emitter is assumed to produce a beam which is parallel, uniform and incident normally to the patient. Nevertheless, this approach appears to be accurate to less than 2% when applied to actual treatment plans in regions where the dose gradient is small [6].

2.4 The deformation

The CT scan made prior to treatment is used to determine the parameters $P_0 = (\mu_0, \rho_0)$ that best describe the patient at this time. The attenuation μ_0 is measured directly, and ρ_0 can be inferred from knowledge about human anatomy. This data is used to determine a treatment plan, which in particular involves a configuration ϕ_0 for the linear accelerator such that the applied dose $D_0 = D(\phi_0, P_0)$ is suitable for treatment of the tumor.

The problem is that this preliminary scan may no longer be accurate during treatment because the patients will have naturally undergone some deformations of the types described in the introduction and is now described by a new set of parameters (P_1, μ_1). A new scan can be made on the spot to determine the new attenuation profile μ_1, which can in many situations be used to derive a displacement vector field $\boldsymbol{\delta}$ that describes the deformation that the patient has undergone to go from P_0 to the new configuration P_1. This vector field assigns to every point $\boldsymbol{r}$ in the deformed patient a displacement vector $\boldsymbol{\delta}(\boldsymbol{r})$ such that the corresponding point in the original patient is, at least approximately, $\boldsymbol{r} + \boldsymbol{\delta}(\boldsymbol{r})$. More specifically, it is such that the set of parameters $P_{\boldsymbol{\delta}} = (\mu_{\boldsymbol{\delta}}, \rho_{\boldsymbol{\delta}})$, with

$$\mu_{\boldsymbol{\delta}}(\boldsymbol{r}) = \mu_0(\boldsymbol{r} + \boldsymbol{\delta}(\boldsymbol{r})), \qquad \rho_{\boldsymbol{\delta}}(\boldsymbol{r}) = \rho_0(\boldsymbol{r} + \boldsymbol{\delta}(\boldsymbol{r})),$$

closely approximates P_1.

2.5 The objective function

Since solving the original optimization problem again for the deformed situation is unfeasible due to time constraints, another approach is called for. The approach we took is to use the deformation field $\boldsymbol{\delta}$ to deform the original dose distribution $D_0 = D(\phi_0, P_0)$ along with the patient, and to attempt to find a new configuration ϕ such that the delivered dose $D(\phi, P_1)$ approximates the deformed distribution $D_{\boldsymbol{\delta}}$ as closely as possible. This deformed dose distribution is defined in the same way as the deformed density and attenuation, and is thus given by

$$D_{\boldsymbol{\delta}}(\boldsymbol{r}) = D_0(\boldsymbol{r} + \boldsymbol{\delta}(\boldsymbol{r})).$$

While this distribution may not be optimal for the deformed patient P_1, it should be as acceptable as the original dose distribution D_0 was for the undeformed patient P_0.

Formulated as an optimization problem, we look for a configuration ϕ, close to ϕ_0, that minimizes the distance between the corresponding dose $D(\phi, P_1)$ and the target dose $D_{\boldsymbol{\delta}}$. This is the solution to the non-linear least squares problem

$$\arg\min_{\phi} \|D(\phi, P_1) - D_{\boldsymbol{\delta}}\|^2, \tag{3}$$

which uses the square integral norm

$$\|\Delta\|^2 = \iiint_V \Delta(\boldsymbol{r})^2 \, d^3\boldsymbol{r}.$$

The precise form of the objective function in equation (3) depends on how the dose distribution D is discretized. If a rectangular grid $\Lambda = \{\boldsymbol{r}_{ijk} \mid i, j, k\}$ is used, then

$$\|D(\phi, P_1) - D_{\boldsymbol{\delta}}\|^2 = \sum_{i,j,k} \big(D(\phi, P_1)(\boldsymbol{r}_{ijk}) - D_{\boldsymbol{\delta}}(\boldsymbol{r}_{ijk})\big)^2$$

up to some (immaterial) constant factor.

3 Methods

In this section we discuss three different methods to apply to the optimization problem described in Section 2. The three methods are compared to each

other and pro and cons are discussed. Finally, in Section 3.4, we describe the simplified model that we considered during the study week. This model was used in our investigations in Section 4.

3.1 Gauss-Newton type methods

Since the optimization problem (3) is a non-linear least squares problem, one can try to solve it using the Gauss-Newton algorithm, or similar algorithms such as Levenberg-Marquardt. A description of these algorithms can be found in many textbooks, such as [3] or [4]. These algorithms require the computation of the first order partial derivatives of the dose $D(\phi, P)$ with respect to the components of $\phi = (\alpha, \beta, \tau)$, but no second order derivatives. They are fast when applied to problems which are only mildly non-linear, which one might expect to be the case for very small deformations, but can be slow otherwise.

Without any modifications, the Gauss-Newton algorithm is not very reliable as it is not even guaranteed to converge to a local minimum. Damped version of Gauss-Newton, such as the Levenberg-Marquardt algorithm, avoid this problem by effectively taking smaller steps when necessary. Convergence to an insignificant local minimum is an inherent possibility for all algorithms of this type.

In the Gauss-Newton method the dose function $D(\phi, P_1)$ is approximated around the initial guess ϕ_0 by its first order approximation $D(\phi_0, P_1)+J_D(\phi_0, P_1)(\phi-\phi_0)$ and solving the linear least square problem

$$\arg\min_{\phi} \|D(\phi_0, P_1) + J_D(\phi_0, P_1)(\phi - \phi_0) - D_{\boldsymbol{\delta}}\|^2. \tag{4}$$

This process is repeated, taking the solution from the previous iteration as the new initial guess.

Fortunately, the derivatives required for these algorithms are relatively easy to compute, as the parameters α_{ij} and β_{ij} only appear in the boundaries of the integrals in equation (2) and the dose is manifestly linear in τ_i. The partial derivatives with respect to α_{ij} and β_{ij} are given by

$$\frac{\partial D(\phi, P)}{\partial \alpha_{ij}} = -\tau_i \int_{\eta_{i-1}}^{\eta_i} d_i(\alpha_{ij}, \eta; P)\, d\eta$$

and

$$\frac{\partial D(\phi, P)}{\partial \beta_{ij}} = \tau_i \int_{\eta_{i-1}}^{\eta_i} d_i(\beta_{ij}, \eta, P)\, d\eta$$

respectively. Computing these should be significantly less costly than computing $D(\phi, P)$ itself as the domain of integration in these expressions is 1-dimensional, rather than 2-dimensional.

The derivatives with respect to the variables τ_i are given by

$$\frac{\partial D(\phi, P)}{\partial \tau_i} = \sum_{j=1}^{N_{\text{leaves}}} \int_{\alpha_{ij}}^{\beta_{ij}} \int_{\eta_{i-1}}^{\eta_i} d_i(\xi, \eta, P)\, d\eta\, d\xi.$$

Note that these do not need to be computed separately, as these integrals should already have been evaluated to determine the dose $D(\phi, P)$.

The Jacobian matrix $J_D(\phi, P)$ which contains all of these partial derivatives has $2\,N_{\text{stops}}\,N_{\text{leaves}} + N_{\text{stops}}$ columns (which is 7290 for the NKI set-up), while the number of rows depends on the discretization of D. It will have one row for every voxel if a grid is used, and most of its entries will be zero in this case since $\frac{\partial}{\partial \alpha_{ij}} D(\phi, P)(\boldsymbol{r})$ is only non-zero if the ray from the emitter to $\boldsymbol{r}$ passes near to the edge of the leaf corresponding to the variable α_{ij} (and similarly for $\frac{\partial}{\partial \beta_{ij}} D(\phi, P)(\boldsymbol{r})$). This sparsity can be used to speed up the computation of $J_D(\phi, P)$, as well as any matrix operations applied to it.

Closed leaf pairs should be given extra attention, since the point at which they are closed does not affect the dose and is thus rather arbitrary. When it becomes necessary to open a pair of leaves that was previously closed, these algorithms will only discover this if the point at which an opening should be created is already close to the (arbitrary) leaf edge positions. Closed leaf pairs should therefore always be positioned such that their edges project onto a point in the patient where a high dose is required, or close to neighboring leaf pairs which are already open. It may be necessary to consider different positions for such leaf pairs while running the algorithm and opening them where doing so would produce the greatest improvement to the objective function.

Since the derivatives of D with respect to α_{ij} and β_{ij} add up to zero whenever $\alpha_{ij} = \beta_{ij}$, the Jacobian $J_D(\phi, P_1)$ will not have full column rank for configurations with closed pairs of leaves. While this is not hugely problematic, it can lead to the Gauss-Newton algorithm prescribing large (unnecessary) changes to the position of closed leaf pairs. One way to avoid this is by keeping the

center $\frac{1}{2}(\alpha_{ij} + \beta_{ij})$ fixed (and only varying the size of the opening) whenever $\alpha_{ij} = \beta_{ij}$ at any particular iteration step.

Newton's method, which works by repeatedly making a second order approximation of $\|D(\phi_0 + \Phi, P_1) - D_{\boldsymbol{\delta}}\|^2$ in Φ, might also be worth considering. Unlike the Gauss-Newton algorithm, Newton's method is guaranteed to converge locally and this convergence is significantly faster (quadratic, rather than linear). The major disadvantage to this approach is that it requires taking derivatives of D up to second order. In our situation, at least when using the analytic pencil beam kernel approach, the second order derivative of D should still be manageable. While there are in principle $(2\,N_{\text{stops}}\,N_{\text{leaves}} + N_{\text{stops}})^2$ mixed second order partial derivatives, almost all of these are again zero. Note in particular that $\frac{\partial^2}{\partial\alpha_{ij}\partial\beta_{i'j'}} D(\phi, P_1) = 0$ for all pairs (i, j) and (i', j') and that $\frac{\partial^2}{\partial\alpha_{ij}\partial\alpha_{i'j'}} D(\phi, P_1)$ can be non-zero only if $(i, j) = (i', j')$.

3.2 Evolutionary algorithms

The wide variety of tissues in a patient causes inhomogeneities with respect to attenuation properties and the pencil beam function may be not accurate enough. As a result of these complexities, the surface of the objective function may be very irregular, which may cause methods based on Gauss Newton to end up in local minima. This may lead to inappropriate settings for the treatment. Whether this is the case needs to be investigated.

A possible way to deal with such complexities and avoid local minima, if they occur, is by applying other types of optimization algorithms. One could for instance think of evolutionary computation as a class of optimization methods that are metaheuristic with stochastic elements [2]. These methods can be applied to any type of objective function, since they are used as a black box: no derivatives are needed. Examples of these type of methods are evolutionary algorithms or swarm type algorithms. Evolutionary computation generally consists of a set of individuals, where for each individual the fitness (objective function value) is evaluated. The algorithm then selects/replaces/moves/breeds from (depending on the algorithm) these individuals iteratively, to eventually close in on the optimal solution. In the study week we considered one such evolutionary algorithm in more detail, the so-called Particle Swarm optimization (PSO).

The PSO considers a group of individuals, uniformly distributed around an

initial guess of parameter values. For each individual, the objective function is calculated. Then in each iterative step, the position (parameter values) of each individual is updated. An individual moves in the direction of its current velocity, the best position this individual found so far and the best position the complete group found so far, with some randomness included.

3.3 Gradual deformations

In this project, we also designed another method based on the Gauss-Newton method that might avoid drifts to inappropriate solutions and be nearly as fast as the Gauss-Newton method. The idea is based on the fact that there are two scans of the patient: the first CT scan used for the treatment plan and one obtained directly before treatment. So, between those two time points, the patient has deformed slowly. During the optimization procedure, we can artificially deform the original patient situation $P_0 = (\mu_0, \rho_0)$ to the deformed situation $P_{\boldsymbol{\delta}} = (\mu_{\boldsymbol{\delta}}, \rho_{\boldsymbol{\delta}})$, in N_s steps by using $P_i = (\mu_i, \rho_i)$, with

$$\mu_i(\boldsymbol{r}) = \mu_0(\boldsymbol{r} + \tfrac{n}{N_s}\boldsymbol{\delta}(\boldsymbol{r}))$$
$$\rho_i(\boldsymbol{r}) = \rho_0(\boldsymbol{r} + \tfrac{n}{N_s}\boldsymbol{\delta}(\boldsymbol{r}))$$

for $n \in \{1, 2, ..., N_s\}$. The desired dose distribution D_n at step n is defined as the goal dose distribution at deformation step n:

$$D_n = D_0(\boldsymbol{r} + \tfrac{n}{N_s}\boldsymbol{\delta}(\boldsymbol{r})).$$

Now, every deformation step n, we may use Gauss-Newton (or any another optimization algorithm) iteratively to find the settings ϕ_n that minimize:

$$\|D(\phi_n, P_n) - D_n\|^2. \tag{5}$$

Subsequently, these settings are used as an initial guess for the next deformation step $n + 1$. The concept behind this algorithm is that finding optimal settings for small deformations might be much easier and faster, and we can better take the correct path from original settings to optimal settings, avoiding drifts to an inappropriate solution.

3.4 Toy model

In the study week we focused on a simple model of the optimization problem to test our different methods, in particular the Gauss-Newton algorithm described

in Section 3.1. In this model we assume the following. First of all, the patient is assumed to be a 2D square lattice with a convex-like tumor. The beam is fixed at one point and the beam direction is orthogonal to the patient plane, i.e. the beam is right above the patient. So, we ignore that the LINAC head can travel around the patient with different durations at each stop. The optimization problem then reduces to the optimization of the leaf configuration only. This specific setting that we chose for the problem is visualized in Figure 3 below.

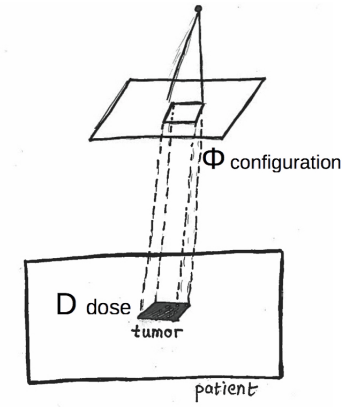

Figure 3: Model setting

Here, for a specific dose requirement D there is a leaf configuration ϕ which ensures that the patient gets the required dose. The goal is then to deform the tumor in a simple way and check whether the algorithm outputs a configuration which gives the same dose distribution, but modified according to the tumor deformation. For example, by simply shifting the tumor to the right as seen in Figure 4, we expect the optimization algorithm to output a configuration for the leaves such that the dose profile follows the same shift.

Aside of the qualitative behavior of the solution, we were also interested in computing the error with respect to some measure. For simplicity, we chose a relative error defined in the following way

$$\mathrm{Err}(\tilde{D}, D) = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\tilde{D}(i,j) - D(i,j)\right)^2}{\sum_{i=1}^{N}\sum_{j=1}^{M}\tilde{D}(i,j)^2}} \tag{6}$$

where $\tilde{D}$ is the desired dose, D is the dose corresponding to a configuration

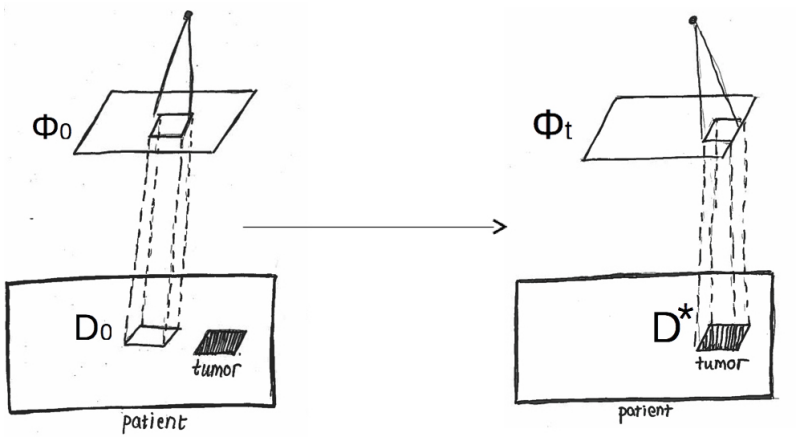

Figure 4: Simple deformation and desired dose profile

given by the optimization algorithm and both N and M correspond to the dimensions of the patient computational domain.

4 Results

In this section we mainly consider the Gauss-Newton algorithm for the optimization problem described in Section 3.4. For the purpose of this project, we chose a 40-by-40 pixels computational domain and a beam emitter with 10 pairs of leaves. Then we assumed for each scenario that at the beginning of the treatment, the patient presents a diamond-shaped tumor. Four qualitatively different deformations of the tumor were considered: vertical shift, lateral shift, diagonal shift, and a shrinkage of the tumor. For each setting, we considered the classical Gauss-Newton algorithms over 20 steps and at each step we recorded the error defined through Equation (6).

Finally, we applied the PSO to the vertical shift of the tumor. For this algorithm we used a matlab code[2] with 10 individuals and 50 iteration steps.

[2] *Particle_Swarm_Optimization.m*© Pramit Biswas. Downloaded from *http://www.mathworks.com/matlabcentral/fileexchange*

4.1 Vertical shift deformation

Figure 5 shows the results for a downward shift of the tumor by 4 pixels. A good solution was found after about seven iterations, which took about 46 seconds to run. After that no further improvements were observed, and the relative error between the desired dose and the dose that is actually realised stabilises at about 19.7%. It is uncertain whether the final solution is a global minimum for the objective function. The speed of convergence definitely seems promising for real-life applications.

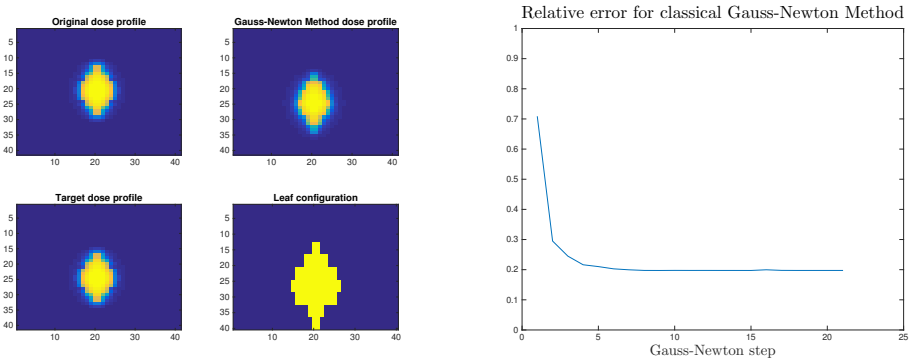

Figure 5: Gauss-Newton algorithm for the vertical shift case of the tumor. The first plot shows the original dose profile, the target profile after the shift, the profile after the last Gauss-Newton step and the configuration of the leaves. The second plot presents the error evolution throughout the iteration steps.

4.2 Lateral shift deformation

For a lateral deformation, we considered a 4 pixel movement to the right. A decent solution is found after about 10 iterations, after a runtime of about 80 seconds. The final error is about 21.9%, as shown in Figure 6. What is interesting to note is that the bottom part of the opening is missing in the final dose leaf configuration because this pair of leaves was closed in the first iteration step at a point which was too far away from the shifted tumor. By moving the closed leaf pair to a point inside the tumor and resuming the algorithm a relative error of about 6.9% is obtained. (This lower error is explained by the fact that a lateral shift of the tumor can be precisely compensated by a corresponding shift of all leaf positions.)

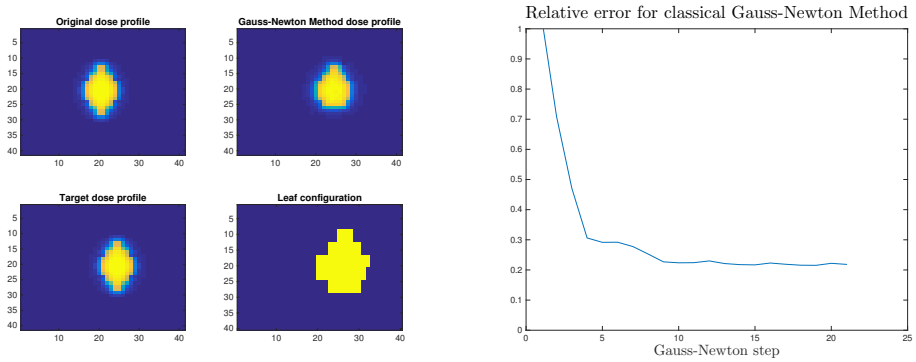

Figure 6: Gauss-Newton algorithm for the lateral shift case of the tumor. The first plot shows the original dose profile, the target profile after the shift, the profile after the last Gauss-Newton step and the configuration of the leaves. The second plot presents the error evolution throughout the iteration steps.

4.3 Diagonal shift deformation

The case of a diagonal shift correspond to a combination of both vertical and lateral changes of the tumor. In our case we chose a 4 pixel deformation in both rightward and downward directions. As expected from the lateral and the vertical case (Figure 7) we see a similar but slightly higher error of 23.8% at the last iteration step after a total runtime of 157 seconds and 20 iterations (the error does not decrease significantly after that). The sudden decrease at the end is due to closed leaf pair that is suddenly opened and could have been found at an earlier stage.

4.4 Size deformation

For the last deformation, we considered a different scenario. Before we only focused on directions in which the tumor can move during the treatment, but in this case we are looking at a size deformation. This particular type of deformation can be justified through the fact that one expects, or hopes, the tumor to shrink during the treatment and thus, the leaf configuration should change accordingly. More precisely, we considered a 50% decrease in size without any lateral or vertical movement involved. After about 7 iterations and 48 seconds of simulation we see the outcome presented in Figure 8. This time the error was about 27.1%.

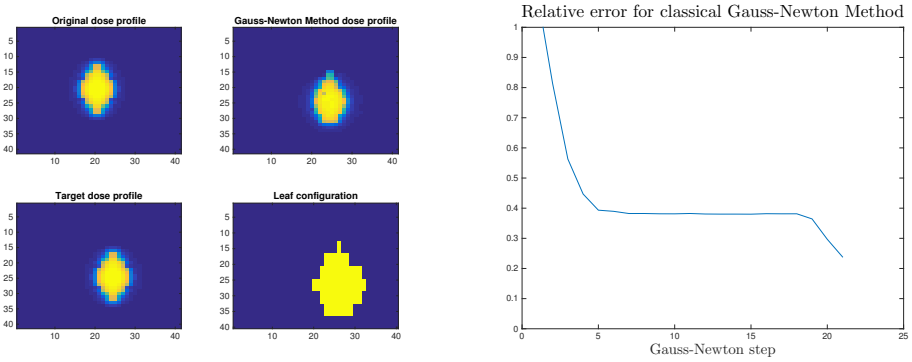

Figure 7: Gauss-Newton algorithm for the diagonal shift case of the tumor. The first plot shows the original dose profile, the target profile after the shift, the profile after the last Gauss-Newton step and the configuration of the leaves. The second plot presents the error evolution throughout the iteration steps.

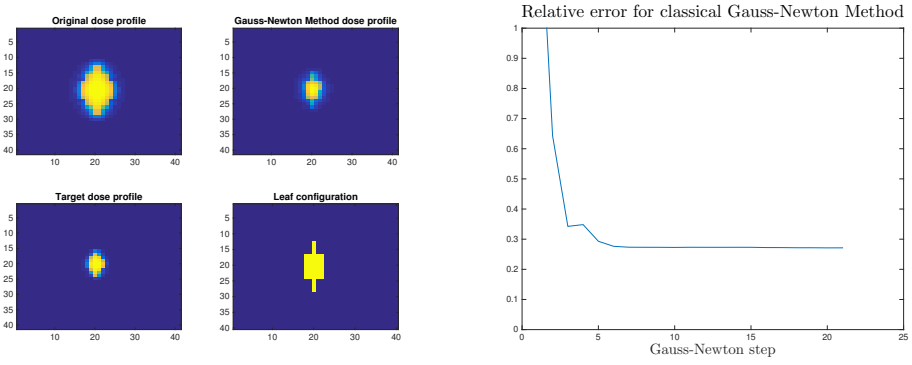

Figure 8: Gauss-Newton algorithm for the shrinking case of the tumor. The first plot shows the original dose profile, the target profile after the shift, the profile after the last Gauss-Newton step and the configuration of the leaves. The second plot presents the error evolution throughout the iteration steps.

4.5 Partical Swarm: vertical shift

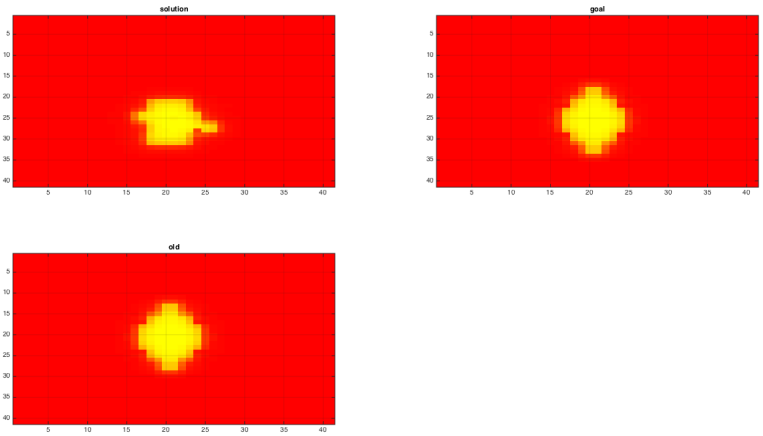

Figure 9: PSO applied to downward shift. Shown is dose profile before optimization (old), after optimization (solution) and shifted dose (goal).

Figure 9 shows that the PSO algorithm works quite well for this shift. However, the solution seems to be less accurate then the one obtained by the Gauss-Newton method (Figure 5). The PSO algorithm with these setting also took approximately half an hour to run, much longer than Gauss Newton. This long duration is caused by the large number of expensive function evaluations, namely for each individual and each iteration. However, it might be profitable to use such black box algorithm, because one does not need to compute more derivatives when the number of parameters increases, but then again, one might need more individuals in PSO. Fortunately, PSO, and other evolutionary computations, can be sped up by employing parallel computing. However, this was outside the scope of the study week.

5 Discussion, conclusions and recommendations

We formulated the problem posed to us by the NKI as an optimization problem. We considered a simplified model: a two-dimensional rectangular patient and tumor and used the pencil beam approach to calculate the dose distribution. Obviously, this is unrealistic but it is a good starting point from a mathematical

point of view. Instead of considering all possible machine settings we only focused on the settings of the leaves. In the model, the deformations that were considered were simple so that the optimal settings were known. We tested several algorithms to see how good they were at finding these optimal settings.

We considered three numerical optimization methods. First we examined the Gauss-Newton method which we applied to four different scenarios: vertical, lateral, diagonal shifts, and tumor shrinkage. As discussed in Section 4, the Gauss-Newton method already appears to work rather well for all types of deformations that were considered. It may nevertheless be advisable to switch to a similar, but more robust, algorithm such as Levenberg-Marquardt. The position of closed leaf pairs is a potential problem for such algorithms.

Another algorithm that we discussed were evolutionary-type algorithms, in particular the PSO. We applied the PSO to the scenario of a vertical shift of the tumor. For our particular test case, the solution obtained through the PSO was less accurate than that of the Gauss-Newton method (compare Figure 5 to Figure 9). It also took a much longer running time, which is not feasible for real life situations. On the other hand, the PSO has the advantage that it can escape from local minima, unlike to the Gauss-Newton method. Faster and eventually more accurate results with the PSO algorithm might be obtained if parallel computing is used (this was outside the scope of the study week).

Finally, we also proposed a method, which we call a gradual deformation method. This method might be much easier and faster than the PSO. An additional advantage of this method is that it might avoid drifts to inappropriate solutions (local minima). Testing this approach was unfortunately outside the scope of the study week.

One of the ideas that came up during the study week was on adding a penalty term to the minimization problem. The rationale behind this is that deformations will typically not be very large and therefore the optimal settings should be somewhat close to original settings as well. More specifically, one could try to add a penalty term $\gamma \|\phi - \phi_0\|^2$ to the objective function, changing the optimization problem (1) to

$$\arg\min_{\phi} \|D(\phi, P_1) - D_{\boldsymbol{\delta}}\|^2 + \gamma \|\phi - \phi_0\|^2.$$

for some number $\gamma > 0$ (where one should normalize both terms so that they become unitless). Doing this will ensure that ϕ does not stray too far from the

original configuration ϕ_0 during the optimization process. Unfortunately there was no time during the study week to work this out further. We considered the minimization problem with $\gamma = 0$, i.e. no penalty term was included in our results. This penalty term deserves further investigation.

Further investigations can be directed at the improvements of our approach by considering a more realistic model than the toy model discussed in Section 3.4. This model should take into account all settings of the machine (such as the larger number of leaves, the rotation steps and the time spent at each angle) and a description of the patient which is more faithful to reality.

Acknowledgements

We thank Simon van Kranen for helpful comments on the manuscript.

References

[1] A. Ahnesjö. Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media. *Med. Phys.*, 16(4):577–592, 1989.

[2] T. Bäck. *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms.* Oxford University Press, 1996.

[3] Å. Björck. *Numerical methods for least squares problems.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996. ISBN 0-89871-360-9.

[4] R. Fletcher. *Practical methods of optimization. Vol. 1.* John Wiley & Sons, Ltd., Chichester, 1980. ISBN 0-471-27711-8. Unconstrained optimization, A Wiley-Interscience Publication.

[5] e. a. Siesling, S.S. *Kankerzorg in beeld.* Integraal kankercentrum Nederland, 2014. ISBN 978-90-72175-00-7.

[6] Y. Watanabee. Point dose calculations using an analytical pencil beam kernel for IMRT plan checking. *Phys. Med. Biol.*, 46:1031–1038, 2001.