Beamlet selection and energy layer reduction in IMPT by sparsity induced optimisation

Bachelor Thesis

by

A.F.J.G. Janssen

to obtain the degree of Bachelor of Science
at the Delft University of Technology,
to be defended publicly on Wednesday June 26, 2019 at 1:00 PM.

Student number: 4584171
Project duration: November 12, 2018 – June 26, 2019
Thesis committee: Dr. ir. S. Breedveld, Erasmus MC, supervisor
Dr. ir. M. Keijzer, TU Delft, supervisor
Dr. ir. D. Lathouwers, TU Delft, supervisor
Dr. D. C. Gijswijt, TU Delft
Dr. Z. Perkò, TU Delft

An electronic version of this thesis is available at http://repository.tudelft.nl/.
Abstract

In Intensity Modified Proton Therapy, the number of energy layers and the number of beamlets determine the radiation time and the plan calculation time. The purpose of this research project is to test the $\ell_1$-norm and the $\ell_1/\ell_2$-norm on the reduction of beamlets and energy layers.

A data set for a prostate cancer patient, that contains 18777 beamlets in 40 energy layers, and a reference plan, generated by 1705 beamlets of a regular grid in 37 energy layers, were included in this project. Both norms were added to the weighted sum. The weights of the objectives were extracted from a sequential $\epsilon$-constraint optimisation. After deletion of beamlets and energy layers, sequential $\epsilon$-constraint optimisation was performed to calculate the output plan. The Dose Volume Histogram (DVH) of the output plan and reference plan are compared to evaluate the plan quality.

The addition of an $\ell_1$-norm is a useful tool for selecting beamlets. Using a weight for the norm, that is smaller than the lowest weight in the weighted sum, an output plan was generated by 1260 beamlets in 37 energy layers. If the weight of the norm was too large, it was not possible to generate an acceptable output plan with the beamlets that were selected.

The selection of 28 up to 40 energy layers using the $\ell_1/\ell_2$-norm did not differ much from the selection without the norm. The selection of fewer energy layers was different which resulted in a difference in the DVH for the selection of 21 energy layers. In all these selections on energy layers, the plans were generated by more than 14900 beamlets. The $\ell_1/\ell_2$-norm gave a good result for deleting both beamlets and energy layers, although more than double the number of the beamlets in the reference plan needed to be selected.

Using both the $\ell_1$-norm and $\ell_1/\ell_2$-norm an almost equivalent plan to the output plan of the $\ell_1/\ell_2$-norm could be achieved with a reduction of another thousand beamlets.

The results encourage additional testing of the sparsity inducing norms on a data set with more energy layers and implementation of the $\ell_1$-norm for the reduction of beamlets.
In autumn last year I started this research project within the Erasmus University Medical Centre in Rotterdam. At the start I was a bit overwhelmed, "What could I do, I am just a student?". Sometimes it felt like searching for a needle in a haystack, when a minor mistake in one of the so many scripts caused my scripts not to work. In the end, I absolutely loved working on a project and seeing all the results. This project has shown me how much you can do with mathematics and physics in an area that I would not have expected it.

Writing this preface I have mixed feelings, I am proud of the result but it also feels strange. For almost 8 months I have been working hard on this project, but it is also the end of three amazing years in Delft.

Now I would like to thank all three of my supervisors for reading this report and correcting me when I was wrong. Sebastiaan, for the meetings whenever I was in Rotterdam and helping me right away when I got stuck. Marleen besides putting me in contact with Sebastiaan and kindly asking me for chapters, for writing all the recommendation letters for my masters degree.

I am ready for the next chapter.

Anne-Fleur Janssen
Delft, June 2019
## Contents

Abstract iii  
Preface v  

### 1 Proton therapy  
1.1 Proton radiation 1  
1.2 Treatment 2  
1.2.1 Treatment plan 3  
1.2.2 Treatment device 4  
1.3 Movement of the patient and the PTV 5  
1.3.1 Movement of organs in between fractions 5  
1.3.2 Movement of the patient during treatment 5  
1.3.3 Reduction of energy layers and beamlets 6  

### 2 Tumour coverage 7  
2.1 Spread-out Bragg peak 7  
2.1.1 Proton interaction mechanisms 7  
2.1.2 Stopping power 9  
2.1.3 Bragg curve 9  
2.2 Active scanning 10  

### 3 Treatment plan optimisation 13  
3.1 Discretisation 13  
3.2 Optimisation 14  
3.2.1 Sequential $\epsilon$-constraint optimisation 15  
3.2.2 Weighted sum optimisation 16  
3.2.3 Comparison weighted sum and $\epsilon$-constraint optimisation 16  

### 4 Sparsity inducement using norms 17  
4.1 $\ell_1$-norm sparsity term 18  
4.1.1 $\ell_q$ in a simple 2D problem 18  
4.1.2 Extension of $\ell_q$ to a larger problem 18
4.2 $\ell_1$-grouped sparsity term ........................................... 20

5 Method 23
  5.1 Data set and scripts ..................................................... 23
  5.2 Process of making a sparse treatment plan ......................... 24
  5.3 $\ell_1$-norm ............................................................... 25
  5.4 Energy layer reduction .................................................. 26
    5.4.1 Energy layers ...................................................... 26
    5.4.2 $\ell_1$/\ell_2$-norm ............................................... 26
  5.5 Selection of beamlets and energy layers ............................ 28
    5.5.1 Selection of beamlets ........................................... 28
    5.5.2 Selection of energy layers .................................... 29

6 Results 31
  6.1 $\ell_1$-norm .............................................................. 32
    6.1.1 Sparsity of the intermediate solutions ......................... 34
    6.1.2 Spot reduction to 1705 beamlets ................................ 35
    6.1.3 Further reduction of beamlets .................................. 38
  6.2 $\ell_1$/\ell_2$-norm ..................................................... 41
    6.2.1 Selection of energy layers .................................... 43
    6.2.2 Beamlet and energy layer reduction ........................... 44
  6.3 $\ell_1$-norm and $\ell_1$/\ell_2$-norm .................................. 46
    6.3.1 Improvement on Figure 6.4 .................................... 46
    6.3.2 Difference in number of beamlets ............................. 48

7 Conclusion 51

A Scripts 53

B Table of energy layers 55

Bibliography 59
Radiotherapy is used in the treatment of cancer, sometimes in addition to surgery, chemotherapy or hyperthermia therapy. The goal of radiotherapy is to irradiate the tumour cells, whilst sparing the healthy cells as much as possible. The introduction of proton therapy and the problem definition will be given in this chapter.

1.1. Proton radiation

When protons travel through matter, they ionise the material and leave a dose behind. The radiation damages and eventually kills cells in the body. Hence, radiation is an effective way of treating cancer. A typical dose deposition as a function of depth is visualised in the Bragg curve in Figure 1.1. The dose a proton deposits during its travel is relatively low in comparison to the dose it deposits at its maximum. After reaching its peak, the dose deposition goes to zero. This limits the negative effects on the healthy tissue and organs. The range depends on the energy of the proton and the electron density of the tissue [1, 2]. This phenomenon is visualised in the Bragg curve. A detailed explanation of the Bragg curve and different interaction methods of the proton will be given in Chapter 2.
1. Proton therapy

Figure 1.1: Schematic representation of a single Bragg curve. Adaptation from [1]

1.2. Treatment

As healthy cells repair much quicker than tumour cells and to avert side-effects, the treatment is spread out over multiple treatment fractions.

Figure 1.2: Normal tissue cells repair much quicker than cancerous cells. The cancerous cells will be killed whilst sparing the healthy cells. [3]

Each fraction usually takes twenty minutes to one hour of which most of the time is used for positioning the patient. The patient is irradiated for less than ten minutes. [4]

During the treatment, the patient should lie as still as possible. When the patient moves, the actual dose that is delivered might not correspond to the dose that was planned.
1.2. Treatment

1.2.1. Treatment plan

Before the treatment can be performed, a personal treatment plan is created, since each patient is anatomically unique. The plan is based on various different constraints, such as a minimum dose for the volume of the tumour, the Planning Target Volume (PTV), and objectives, such as minimising the maximum doses for different organs, the organs-at-risk (OARs). The PTV of a prostate cancer patient is depicted in Figure 1.3. This makes it a multi-criteria problem. The objectives and constraints are contained in a wish list with their priorities.

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Name</th>
<th>Type</th>
<th>Limit (Gy)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PTV-high</td>
<td>Minimum</td>
<td>71.78</td>
</tr>
<tr>
<td></td>
<td>PTV-intermediate</td>
<td>Minimum</td>
<td>54.45</td>
</tr>
<tr>
<td></td>
<td>PTV-low</td>
<td>Minimum</td>
<td>54.45</td>
</tr>
</tbody>
</table>

Table 1.1: Wish list with all constraints and objectives. The priority of the objective indicates in which order the sequential $\epsilon$-constraint optimisation optimises. A low number will be optimised first. PTV is planning targeting volume and Gp is gigaprotons

<table>
<thead>
<tr>
<th>Objectives</th>
<th>Name</th>
<th>Type</th>
<th>Goal (Gy)</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1$</td>
<td>PTV-high</td>
<td>Maximum</td>
<td>79.18</td>
<td>1</td>
</tr>
<tr>
<td>$f_2$</td>
<td>PTV-intermediate</td>
<td>Maximum</td>
<td>79.18</td>
<td>1</td>
</tr>
<tr>
<td>$f_3$</td>
<td>PTV-low</td>
<td>Maximum</td>
<td>58.85</td>
<td>1</td>
</tr>
<tr>
<td>$f_4$</td>
<td>Conformity ring PTV-high</td>
<td>Maximum</td>
<td>79.18</td>
<td>2</td>
</tr>
<tr>
<td>$f_5$</td>
<td>Conformity ring PTV 0-10mm</td>
<td>Maximum</td>
<td>58.85</td>
<td>2</td>
</tr>
<tr>
<td>$f_6$</td>
<td>Conformity ring PTV 10-15mm</td>
<td>Maximum</td>
<td>49.5</td>
<td>2</td>
</tr>
<tr>
<td>$f_7$</td>
<td>Femural heads</td>
<td>Maximum</td>
<td>50</td>
<td>3</td>
</tr>
<tr>
<td>$f_8$</td>
<td>Rectum</td>
<td>Mean</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>$f_9$</td>
<td>Large and small intestines</td>
<td>Mean</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>$f_{10}$</td>
<td>Bladder</td>
<td>Mean</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>$f_{11}$</td>
<td>Femural heads</td>
<td>Mean</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>$f_{12}$</td>
<td>All conformity rings</td>
<td>Mean</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>$f_{13}$</td>
<td>Rest of the conformity rings</td>
<td>Maximum</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>$f_{14}$</td>
<td>Total spot weight</td>
<td>Mean</td>
<td>1 Gp</td>
<td>9</td>
</tr>
</tbody>
</table>

A sequential $\epsilon$-constraint optimisation as described by Breedveld [5] is used in the computation of a dose distribution that satisfies the requirements in the wish list as good as possible. The optimisation method is explained briefly in Section 3.2.1. The goal of the optimisation is to keep the dose in the OARs as low as possible whilst satisfying the minimum dose in the tumour to eradicate all malignant cells. If the plan is clinically acceptable, a treatment planning software computes the plan that is used within the treatment device.
1. Proton therapy

Figure 1.3: The PTV-high is an expansion of the prostate. The PTV-low consists of the expanded seminal vesicles and lymph nodes, that is not contained in PTV-intermediate, the 15 mm transition region between PTV-low and PTV-high. The conformity ring of PTV-high is the PTV-high with a 15 mm expansion. The conformity ring of PTV-full is the area within 10 mm outside the PTV-full, the PTV-high conformity ring and PTV-low. [6]

1.2.2. Treatment device

In External Beam Radiotherapy (EBRT) protons exit the treatment device through the head. The head contains magnets to steer a bundle of protons to different positions in the patient. The energy of the proton determines the range of the Bragg peak. The positioning of the proton using magnets will be explained in more detail in Chapter 2.

The treatment device most commonly used in EBRT can be seen in Figure 1.4.

Figure 1.4: The machine most commonly used for radiotherapy. The gantry can rotate 360° to irradiate the patient from different angles. The table can move and rotate. Adapted from [7]
The machine can irradiate the patient from different directions, as the gantry can rotate around the patient and the table can turn. In this report we will only consider two opposite beam directions.

### 1.3. Movement of the patient and the PTV

Treatment with protons is very precise and the majority of the dose ends up in the tumour, which spares the surrounding tissues quite well. However, this precision has a downside when the tumour or OAR are displaced from their positions on the scan. As a consequence, the dose at the edge of the tumour might not be sufficient to prevent the tumour from growing back. Another consequence is that the dose in the surrounding organs might be too high.

The shift can be caused by several reasons. One of them is a natural movement as a cause of expanding organs such as the stomach or bladder in between the different treatment fractions. Another cause is the slight movement of the patient during the procedure even though the patient is held as still as possible. The longer the procedure, the higher risk of movement.

#### 1.3.1. Movement of organs in between fractions

If the movements of organs is to be taken into account, a CT scan should be made before each treatment session. Using this CT scan a modified plan can be made. However, making a new plan for each session is time consuming, as the optimisation contains a large number of variables.

Reducing the number of variables enables us to make personalised plans for each treatment.

#### 1.3.2. Movement of the patient during treatment

In order to reduce the risk of body movement during the treatment, the radiation time should be reduced. This does not include repetitive movements such as breathing. The length of the session depends among others on the number of different energies that are used. These energies can be ordered into layers; all beamlets with the same energy are in the same layer. Currently switching between energy layers takes two to five seconds. Most of this time is used for calibrating the magnets in the head of the collimator. As the time of radiation is usually less than ten minutes, having a lot of energy layers is a major contribution on the total radiation time of the treatment. Besides the lower risk on movement, more patients can be treated due to the decrease in radiation time.
1.3.3. Reduction of energy layers and beamlets
The aim of this research project is to reduce the number of beamlets and energy layers used in the optimisation using sparsity inducing norms. The sparsity inducing norms will be introduced in Chapter 4 and in Chapter 5 it is described how the sparsity inducing terms are added.
The proton pencil beam can be modified to fully irradiate the tumour. This can be achieved by changing the energy of the proton, to choose the range of the proton, the intensity, to modify the size of the dose and the magnetic deflection, to steer the proton. [2] Protons of different energies will create a spread-out Bragg peak (SOBP). The spread-out Bragg peak, range calculation and magnetic deflection will be covered in this chapter.

2.1. Spread-out Bragg peak

Intensity Modulated Proton Therapy, (IMPT), allows the alteration of the intensity of the proton beam. In order to get a sufficient dose in the tumour, the tumour is irradiated with protons of different energies and intensities. This results in the spread-out Bragg peak pictured in Figure 2.1.

2.1.1. Proton interaction mechanisms

The proton loses energy due to three main interactions, which are depicted in Figure 2.2. Inelastic Coulomb scattering between the proton and atomic electrons causes energy loss and determines the range of the proton. The deflection of the proton is negligible as the mass of the proton is much greater than the mass of electrons.
As the loss per interaction is small, the proton is continuously slowing down. Elastic Coulomb scattering changes the trajectory of the proton. The proton experiences a repulsive force from the nucleus of the atom, which also has a positive charge, and is deflected from its original path. A non-elastic nuclear interaction is less predictable. When the proton enters the nucleus, a deuteron, secondary proton, triton or heavier ion or one or more neutrons can be emitted. [8, 9]
2.1. Spread-out Bragg peak

2.1.2. Stopping power

The rate of energy loss of the projectile is defined as the quotient of the change in energy $dE$ and the travelled path $dx$. This is the stopping power, $S$.

$$S(E) = -\frac{dE}{dx} \quad (2.1)$$

The Bethe formula appears to be an accurate formula to approximate the stopping power for radiotherapy.

$$S = \frac{4\pi n z^2}{m_e e^2 \beta^2} \left( \frac{e^2}{4\pi\epsilon_0} \right)^2 \left[ \ln \frac{2m_e c^2 \gamma \beta^2}{I} - \beta^2 \right] \quad (2.2)$$

$n$ is the electron density of the material, $z$ the atomic number of the projectile, $m_e$ the mass of an electron, $c$ the speed of light, $e$ the electric charge and $\epsilon_0$ the permittivity of vacuum. $\beta = v/c$, where $v$ the velocity of the projectile. $\gamma = (1 - \beta^2)^{-\frac{1}{2}}$ is the Lorentz factor and $I$ the mean excitation potential of the absorbing material. [8, 10]

2.1.3. Bragg curve

Using Einsteins famous relationship for kinetic energy, it becomes clear that the velocity is high for high energy particles.

$$E_k = (\gamma - 1) m_p c^2 \quad (2.3)$$

Where $m_p$ is the mass of the proton and $\gamma$ the Lorentz factor. The Lorentz factor, $\gamma = (1 - \beta^2)^{-\frac{1}{2}}$, is larger for large $v$.

As the particle has a certain velocity, the stopping power is not zero and energy is deposited to the tissue. With the decrease of energy, the velocity and thus $\beta$ decreases as well. This results in a larger stopping power. At low energy, as $\beta \to 0$, $S$ increases causing a peak in the deposition of energy to occur. This results in the Bragg-Peak that is depicted in Figure 2.1. Under the assumption that the path of the proton is straight and that protons lose energy continuously the range can be calculated.

$$R(E) = \int_0^E \frac{1}{S(E')} dE' \quad (2.4)$$

$R$ is the range in m. [8]
2.2. Active scanning

To get a full tumour coverage, active scanning is used. Magnets in a magnetic scanner deflect the path of the pencil beam to specific positions in a vertical plane. A sketch of different paths of the pencil beam is given in Figure 2.3. [1, 11]

![Active Scanning](image)

Figure 2.3: Pencil beam active scanning. The pencil beam is deflected to points on a vertical plane perpendicular to the initial pencil beam. [11]

The deflection of the particle will be explained with the Lorentz-force.

$$\mathbf{F}_L = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}$$  \hspace{1cm} (2.5)

All bold variables are vectors. $\mathbf{F}_L$ is the Lorentz force on the particle with charge $q$ and velocity $\mathbf{v}$. $\mathbf{E}$ and $\mathbf{B}$ are the electric field and magnetic field respectively. The force on the proton is summation of a force perpendicular on both the velocity of the particle and direction of the magnetic field and a force in the direction of the electric field.

Consider the situation where $\mathbf{E} = 0$ and $\mathbf{B} = B\hat{y}$ is constant in the y-direction. The direction of the velocity is in the z-direction. This is sketched in Figure 2.4.

In this situation, the Lorentz force simplifies to a simple equation.

$$\mathbf{F}_L = -qB(v_z \hat{x} - v_x \hat{z})$$  \hspace{1cm} (2.6)

$\hat{x}$ and $\hat{z}$ are the unit vectors in the x-direction and z-direction respectively.

Using Newton’s law of motion, the centripetal force is given as

$$\mathbf{F}_c = \frac{m_p v^2}{r}$$  \hspace{1cm} (2.7)

With $m_p$ the mass of the proton and $r$ the radius of the circle. Solving
2.2. Active scanning

Figure 2.4: The incident particle is deflected in a circular path towards the negative x-direction under the influence of a uniform magnetic field in the y-direction. [12]

These equations for the initial velocity of $v_x = 0$, $r$ can be calculated. The proton will travel in a circular path with radius $r$.

Changing the strength and direction of the magnetic field the path of the proton can be modified.
T
reatment plan optimisation

In the optimisation of treatment plans, discretised values are used. In this chapter the discretisation of the patient’s volume and the beam positioning is covered. Next, the sequential $\epsilon$-constraint optimisation method will be compared with the weighted sum optimisation method for the multi-criteria problem.

3.1. Discretisation

A beam source of high energy protons exits the head of the treatment device as shown in Figure 1.4. From the beam source, the protons travel through the head to the patient. The modulation device is divided into beamlets, $x_i$, and the volume of the patient into voxels, $d_i$. A simple figure of the discretisation of the patient and the modulation device is pictured in the following figure.
3. Treatment plan optimisation

Figure 3.1: The modulation device is discretised into beamlets, \( x_i \). A beam source sends out ionising radiation, that is deflected through the grid of the modulation device. The patient’s volume is discretised into voxels, \( d_j \). The higher the intensity of the beam, the higher the resulting dose in the patient. Adapted from [7]

The numerical value of \( x_i \) is the intensity of the pencil beam that passes through the grid element. Using the dose influence matrix \( A \), the discretised dose in the patient, \( d_j \), can be calculated. There is a different dose influence matrix for each organ in order to optimise different organs separately. \( d \) is a vector with all doses \( d_j \) and \( x \) of all beamlets \( x_i \).

\[
\mathbf{d}(\mathbf{x}) = A\mathbf{x}
\]  

The total dose distribution is pictured in a 3D dose distribution, overlayed on the CT. The 2D representation can be given in a Dose Volume Histogram (DVH). To keep the problem as small as possible, the most important volumes are sampled with a higher sampling resolution than less important volumes. [7]

3.2. Optimisation

In radiotherapy various multi-criteria optimisation methods can be used, such as weighted sum optimisation and sequential \( \epsilon \)-constraint optimisation. The two methods are interesting, because it is possible to switch between these. The switch between the two methods is useful, as it is easier to add, modify or delete objectives and constraints in a weighted sum than in a sequential \( \epsilon \)-constraint optimisation. This way all treatment objectives are optimised simultaneously while preserving patient specific trade-offs.

In this section the sequential \( \epsilon \)-constraint optimisation and the weighted sum
3.2. Optimisation are introduced. The objectives are denoted as $f_i$ where $i \in [1, \ldots, s]$ and $s$ the number of objectives and the constraints on the tumour as $g_j$ where $j \in [1, \ldots, r]$ and $r$ the number of constraints. The constraints $g_j$ are combined in $g(x)$, such that $g(x) \leq 0$. [5]

An example of a wish list can be found in Table 1.1.

3.2.1. Sequential $\epsilon$-constraint optimisation

The method of sequential $\epsilon$-constraint optimisation, starts by minimising the dose in the OAR with the highest priority, $f_1$, whilst satisfying all constraints $g(x)$. This will generate a plan that has a minimal dose everywhere in the tumour and the lowest dose possible in the OAR with the highest priority. When this optimisation is finished, the result will be relaxed by a value $\epsilon_1$ and added as a constraint in the next steps in the optimisation. Relaxation is necessary to allow clinically interesting trade-offs.

In the next step the OAR with the second to highest priority will be minimised satisfying all the new and old constraints. This will generate a plan that has a minimal dose everywhere in the tumour and a dose in the OAR with the highest priority that is smaller than $\epsilon_1$.

These steps are repeated until all OARs have been processed. The optimisation problem of the $i$’th step is given below.

\[
\begin{align*}
\text{minimise} \quad & f_i(x) \\
\text{subject to} \quad & g(x) \leq 0 \\
& f_k(x) \leq \epsilon_k, \quad k \in [1, \ldots, i - 1]
\end{align*}
\]

All the objectives $f_k$, with a higher priority than $f_i$, are added as constraints with value $\epsilon_k$:

\[
\epsilon_k = \begin{cases} 
 b_k & f_k(x^*) \delta \leq b_k \\
 f_k(x^*) \delta & f_k(x^*) \delta > b_k
\end{cases}
\]

$\epsilon_k$ is the goal set for the minimisation of the objectives and $\delta$ the relaxation value. In most cases the relaxation value is $1.03$ (3%). Without relaxation of the result, it might occur that further calculations do not give a solution. A detailed explanation can be found in Breedveld [5].
3.2.2. Weighted sum optimisation

In the weighted sum method, the objective function is obtained by summing the objectives \( f_i \) with a specific weight \( w_i \) for each objective.

\[
\begin{align*}
\text{minimise} & \quad \sum w_i f_i(x) \\
\text{subject to} & \quad g(x) \leq 0
\end{align*}
\] (3.4)

3.2.3. Comparison weighted sum and \( \epsilon \)-constraint optimisation

To switch between the two optimisation methods, it needs to be confirmed that the methods give similar results. The optimal solutions of both methods will be calculated and compared.

First the optimal solution is calculated using the \( \epsilon \)-constraint method. In Breedveld it is proven that if particular weights are chosen, the weighted sum optimisation results in an optimal solution identical with the optimal solution of the \( \epsilon \)-constraint optimisation. [5]

Therefore, these particular weights will be extracted from the plan generated with the sequential \( \epsilon \)-constraint method and used in the weighted sum method.

![Figure 3.2: DVH of a treatment plan generated by \( \epsilon \)-constraint optimisation and by the weighted sum optimisation.](image)

It appears that the solutions are quite similar. Numerical and rounding errors can be a cause for the differences in the solutions.
In this chapter the cost function $\Omega(x)$ for sparsity-induced optimisation is described. The goal of a sparsity-inducing cost function is to make the solution more sparse. A vector or matrix is sparse when more elements are zero than non-zero. Such a criterion should be active during optimisation of each objective, making inclusion in the sequential $\epsilon$-constraint method impractical. As an alternative, the weighted-sum formulation is used, that is described in Section 3.2.2. The sparsity-induced cost function is added by some small weight $\lambda$.

$$\begin{align*}
\text{minimise} & \quad \sum w_i f_i(x) + \lambda \Omega(x) \\
\text{subject to} & \quad g(x) \leq 0
\end{align*}$$

(4.1)

It is suggested that the sparsity is induced by an $\ell_1$-norm or a mixed $\ell_1/\ell_p$-norm. [13, 14]

In Chapter 5 the inclusion of the sparsity-induced cost functions on the optimisation of a treatment plan in radiotherapy is explained.
4. Sparsity induction using norms

4.1. \( \ell_1 \)-norm sparsity term

At first glance, it is not immediately clear why an addition of an \( \ell_1 \)-norm induces sparsity. In order to understand this, the \( \ell_1 \)-norm is sketched and compared to the general \( \ell_q \)-norms. The \( \ell_q \)-norm is defined as \( \|x\|_q = (\sum_{i=1}^{n} |x_i|^q)^{\frac{1}{q}} \).

4.1.1. \( \ell_q \) in a simple 2D problem

In Figure 4.1a \( \|x\|_q = x_1^q + x_2^q = 1 \) is sketched for \( q = 1, 2 \) and 10. In this figure an important result can be seen. The curve of \( \|x\|_q = 1 \) is a diamond if \( q = 1 \), a circle if \( q = 2 \) and tends to look like a square if \( q = 10 \). For \( q \to \infty \) the curve approaches a square. For all other \( q \), where \( a \leq q \leq b \), the curve of \( \|x\|_q = 1 \) is positioned outside the curve of \( \|x\|_a = 1 \) and inside the curve of \( \|x\|_b = 1 \).

In Figure 4.1b only the positive quadrant is taken into account. A simple optimisation problem with only one constraint, \( x_2 + \frac{3}{2} x_1 \geq 2 \), is considered. Then all points on and above the red line satisfy the constraint. The objective is the minimisation of \( x_2 + \frac{3}{2} x_1 \). Then all points on the red line are possible solutions. Now, \( \Omega(x) = \|x\|_q \) is added with weight \( \lambda = 1 \) to the problem with the single constraint. Thus the goal of the optimisation is to choose the smallest \( c \), such that \( \|x\|_q = c \) and the red line intersect. This can be visualised as an expanding balloon, where the balloon is inflated until it touches the line of the constraint.

As the curve of \( \|x\|_1 = c \) is a diamond, the point of intersection is very likely to be on one of the axes, which results to a sparse solution. The curve of \( \|x\|_{10} = c \) resembles a square. Therefore, the intersection is very likely to be in the middle of the line and the solution is not sparse at all. The \( \ell_2 \)-norm is not as straightforward. In that case, the probability of getting a sparse or almost sparse solution is quite low. We say that the solution is almost sparse if \( x_1 \) or \( x_2 \) is small in comparison to the other.

Visually, it is very clear that the addition of an \( \ell_1 \)-norm increases sparsity and the addition of an \( \ell_q \)-norm where \( q \geq 2 \) does not.

4.1.2. Extension of \( \ell_q \) to a larger problem

In the previous section a problem with two variables and one constraint is considered. This can be extended to a problem with more constraints and variables. The extension to a problem with multiple variables is rather straightforward. Then, the surface can be seen as a hypervolume and the minimisation of the norm corresponds with expanding the volume until it intersects with the constraint.
4.1. $\ell_1$-norm sparsity term

Figure 4.1: a) $\|x\|_q = 1$ for $q = 1, 2$ and 10. b) $\|x\|_q = c$ and constraint $x_2 + \frac{2}{3} x_1 \geq 2$ for positive values $x_1$ and $x_2$. The point of intersection is indicated with dots.

If there is more than one constraint in the problem, the same thought process can be used. Even though the intersection of $\|x\|_1$ with the constraints might not be exactly at one of the axis, the probability of getting a sparse or almost sparse solution is higher.
4.2. \( \ell_1 \)-grouped sparsity term

A grouped sparsity term, \( \ell_q/\ell_p \)-norm, is used if the goal is to induce sparsity in a problem, where all variables of a group in a partition should either be selected or ignored. The \( \ell_q/\ell_p \) is defined as

\[
\Omega(x) = \left( \sum_{g \in G} \|x_g\|_p^q \right)^{\frac{1}{q}}
\]  

(4.2)

\( G \) is a partition of the set \( S = \bigcup x_i \), the set with all \( x_i \). \( G \) is an element of partition \( G \) and therefore a subset of \( S \). \( \|x_g\|_p \) denotes the \( \ell_p \)-norm performed on all \( x_i \) that are in set \( g \).

The grouped sparsity term will be clarified using a simple example.

Consider Figure 4.2 with set \( S = \{x_i : i \in [1,6]\} \) and assume that the variables of set \( S \) are either red, blue or green. Each subset \( S_1, S_2 \) and \( S_3 \) corresponds to one of these colours. \( G = \{[x_1], [x_2, x_3], [x_4, x_5, x_6] = [S_1, S_2, S_3] \) is a partition of the set with elements \( S_1, S_2 \) and \( S_3 \) based on the different colours.

![Figure 4.2: Set S can be divided in three disjoint sets S1, S2 and S3 based on the colour of xi.](image)

The \( \ell_p \)-norm of each group is

\[
\begin{align*}
S_1 : \quad & \|x_{S_1}\|_p = \sqrt[\ell]{x_1^p} \\
S_2 : \quad & \|x_{S_2}\|_p = \sqrt[\ell]{x_2^p + x_3^p} \\
S_3 : \quad & \|x_{S_3}\|_p = \sqrt[\ell]{x_4^p + x_5^p + x_6^p}
\end{align*}
\]  

(4.3)

Calculating the \( \ell_q/\ell_p \)-norm results in

\[
\Omega(x) = \sqrt[q]{\sqrt[\ell]{x_1^p} + \sqrt[\ell]{x_2^p + x_3^p} + \sqrt[\ell]{x_4^p + x_5^p + x_6^p}}
\]  

(4.4)
Logically, the $\ell_1$-norm is used on the disjoint groups, as the goal is to select or ignore whole groups. Another norm will be used on the variables of the groups. The norm on the variables within the group depends on the preferred result. If the variables need to be equal, it is advised to use an $\ell_p$-norm with high $p$, whereas a low $p$ is advised if the sparsity of the variables within the group need to be induced even more. These norms are combined in a mixed $\ell_1/\ell_p$-norm.

$$\Omega(x) = \sum_{g \in G} \| x_g \|_p \quad (4.5)$$

Now the $\ell_1$-norm is calculated over the $\ell_p$-norm of each group, $\| x_g \|_p$.

The $\ell_1/\ell_2$-norm is defined on the coloured groups of set $S$ as,

$$\Omega(x) = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2} \quad (4.6)$$
The goal of this research project is to reduce the number of energy layers and the number of beamlets, that are used in a radiotherapy treatment plan. In this chapter it is described how to add the sparsity inducing terms to the weighted sum and to find a solution with fewer beamlets and energy layers.

First the data set and scripts that are used are described. Second, the overall process is explained using a workflow. Then the $\ell_1$-norm and $\ell_1/\ell_2$-norm are written in forms such that these can be included in the optimisation problem. Finally the method that is used for the selection of beamlets is explained.

5.1. Data set and scripts

The data set includes the details of two laterally opposing beams, the wishlist (see Table 1.1) of a prostate cancer patient, a CT scan of the patient and other information that is required to visualise the data. In total there were 18777 beamlets in 40 energy layers, 18 in the first beam and 22 in the second beam.

All plans and calculations are made in the house-build program iCycle using already existing routines such as "cons2weights", to extract the weights for the weighted sum optimisation, "primaldual" to solve an optimisation problem by a weighted sum and "mcopt" to calculate a plan using sequential $\epsilon$-constraint optimisation. The script "cons2weights" is based on the theory in Breedveld. [5]. The scripts "primaldual" and "mcopt" are able to optimise for objectives and
constraints that have a specific cost function. For each function the function value, gradient and Hessian need to be defined given a matrix or vector that is specific for this function.

Take for example the minimisation of the dose in an OAR. The function is defined as \( d = Ax \), with \( A \) dose influence matrix. If the matrix \( A \) is known, the function value, gradient and Hessian can be calculated.

### 5.2. Process of making a sparse treatment plan

The workflow sketched below will be used to reduce the number of beamlets and energy layers.

![Workflow Diagram](image)

Figure 5.1: Workflow of the reduction of spots, energy layers or reduction of spots and energy layers.

The input plan is a plan created with sequential \( \epsilon \)-constraint optimisation on all 18777 beamlets in the data set. In all cases the weighted sum method is used with the addition of one or two sparsity inducing terms as described in Chapter 4. The plans with the added norm(s) will be called "intermediate plans".
These intermediate plans will be compared with the plan calculated with the weighted sum method without the addition of a norm. This comparison is used to analyse the influence of the norms on the weighted sum.

The selection of beamlets and energy layers is performed on the intermediate plans. Using the remaining beamlets and energy layers a new (output) plan is created that satisfies the minimum doses in the PTVs, using the sequential $\epsilon$-constraint optimisation.

The output plan is then compared to the reference plan in Figure 5.2. The reference plan is created with beamlets on a "regular grid" with a lateral spacing of 6 mm and a relative energy spacing of 2. All beamlets of this reference plan were included in the data set.

![Figure 5.2: The reference plan of a prostate cancer patient. The plan is generated by 1705 beamlets in 37 different energy layers.](image)

### 5.3. $\ell_1$-norm

As all $x \geq 0$, the $\ell_1$-norm simplifies to $\|x\|_{l=1} = \sum_{i=1}^{n} x_i$, where $n$ is the number of beamlets. The norm results in a linear cost function and is denoted in matrix form as

$$\|x\|_1 = cx$$

where $c$ is a $1 \times n$ matrix.

The function value $f$, the $i$'th element of the gradient $\nabla f$, and element $h_{i,k}$
of the Hessian, \( H = \nabla^2 f \), of a linear function are defined as,

\[
f = \sum_{i=1}^{n} c_i x_i \tag{5.2}
\]

\[
\frac{\partial f}{\partial x_i} = c_i \tag{5.3}
\]

\[
h_{i,k} = 0 \tag{5.4}
\]

Where \( c_i \) is the \( i \)’th element of the cost function vector \( c \).

5.4. Energy layer reduction

In the previous chapter it is stated that an \( \ell_1 \)-grouped sparsity term is a useful tool to induce sparsity in a problem where there is a partition of all variables. In this part the \( \ell_1 \)-grouped sparsity term is applied to energy layers and a mixed \( \ell_1/\ell_2 \)-norm is written in a form that can be added to the optimisation problem.

5.4.1. Energy layers

As mentioned in Chapter 1, protons of different energies are used during treatment. Additionally to the spatial information, each beamlet also contains the information of the energy of the protons passing through the head. So each beamlet comes from a specific position and is in a certain energy level. Therefore, a partition, \( E \), of beamlets based on the energy can be constructed. \( E \) is a set with elements \( E_j \), where \( j \in \{1, \ldots, m\} \) and \( m \) the number of energy layers. \( E_j \) is the set containing all beamlets that are in energy layer \( j \) such that the union of all \( E_j \) is an exact cover of the set with all beamlets, \( S \). Using this exact cover, a grouped sparsity term can be added.

\[
S = \bigcup_{j=1}^{m} E_j \tag{5.5}
\]

The mixed \( \ell_1/\ell_2 \)-norm is an \( \ell_1 \)-norm on all \( \| x_{E_j} \|_p \) for \( E_j \in E \). The grouped sparsity term will result in the fact that whole energy layers can be neglected.

5.4.2. \( \ell_1/\ell_2 \)-norm

The \( \ell_1/\ell_2 \)-norm does not have a simple function value that can be written in a linear or other simple matrix form. First, the function value \( f \) will be calculated.
5.4. Energy layer reduction

\[ f = \sum_{j=1}^{m} z_j \]  

(5.6)

\[ z_j \text{ is the } \ell_2\text{-norm of layer } j. \]

\[ z_j = \sqrt{\sum_{i=1}^{n} x_{E_j}^2} \]  

(5.7)

\( f \) can be divided into a part that is dependent on a specific \( x_i \) that is in energy layer \( E_j \) and a part that is not.

\[ f = z_j(x_i) + g(x_k) \]  

(5.8)

Where \( g(x_k) \) is a function on all \( x_k \)'s that are not in the same energy layer as \( x_i \), \( x_k \notin E_j \). The gradient of the \( i \)'th element that is in the \( j \)'th energy layer is the derivative of \( z_j(x_i) \).

\[ \frac{\partial f}{\partial x_i} = \frac{x_i}{z_j} \]  

(5.9)

The Hessian has the following elements \( h_{i,k} \) where \( i, k \in \{1, \ldots, n\} \) with \( n \) the number of beamlets and \( j \in \{1,2,\ldots,m\} \) with \( m \) the number of energy layers.

\[ h_{i,k} = \begin{cases}  
-\frac{x_i^2}{z_j} + \frac{1}{z_j} & i = k \\
-\frac{x_i x_k}{z_j^2} & x_i, x_k \in E_j \\
0 & \text{otherwise} 
\end{cases} \]  

(5.10)

Using the script that is given in Appendix A, the function value, gradient and Hessian matrix can be calculated with matrix \( B \). \( b_{j,i} \), where \( i \in \{1,2,\ldots,n\} \) with \( n \) the number of beamlets and \( j \in \{1,2,\ldots,m\} \) with \( m \) the number of energy layers, is an element of \( B \).

\[ b_{j,i} = \begin{cases}  
1 & x_i \in E_j \\
0 & \text{otherwise} 
\end{cases} \]  

(5.11)

\( B \) is \( m \times n \) matrix, where \( m \) indicates the number of energy layers and \( n \) the total number of beamlets. Each \( b_{j,i} \) indicates whether beamlet \( x_i \) with \( i \in \{1,2,\ldots,n\} \) is in energy layer \( E_j \) with \( j \in \{1,2,\ldots,m\} \).
5.5. Selection of beamlets and energy layers

The plan that is generated by the weighted sum with either the $\ell_1$-norm and/or $\ell_1/\ell_2$-norm contains all of the beamlets and energy layers that were used in the input plan. Depending on the weight of the norm, a number of the numerical values of the beamlets or energy layers will be small or zero. Selecting all beamlets and energy layers that have a large numerical value (or deleting the lower ones), hopefully a new clinically acceptable plan can be made with fewer beamlets and energy layers.

5.5.1. Selection of beamlets

All beamlets $x_i \geq b$ will be selected, where $b$ is the lower bound. In the results there will be looked at different values of $b$ for different weights for the sparsity terms. Deleting $x_i < b$ will delete two kinds of beamlets; beamlets with a numeric value that is zero or that is non-zero.

The deletion of the beamlets with zero contribution can be done without any problem if the plan of the weighted sum is clinically acceptable. In that case the resulting $x_i$’s still satisfy the constraints and a plan without the norm will result in an equivalent or better plan. In the final optimisation, using the $\epsilon$-constraint optimisation, the optimisation is performed on all the beamlets and energy layers that were selected without the extra term. Thus the DVH might result in a better plan, as the minimisation of the norm(s) is not one of the objectives anymore. The minimisation of the norms caused that different trade-offs were made.

However, after deleting the beamlets with a non-zero numeric value the intermediate plan is insufficient to satisfy the minimal dose in the tumour and the numeric value of other beamlets need to change. To explain this consider the following case.

Suppose that the dose in the tumour is sufficient after the deletion of beamlets and the numerical value of the beamlets is non-zero. Then, the deletion will only influence the maximum doses in organs and in the tumour. However the dose in all organs and in the tumour is minimised. Thus these beamlets would have been zero or the dose in the tumour is not sufficient anymore.

As a result of an insufficient dose, the numerical values of other beamlets need to increase. This will cause the DVH to change negatively for the healthy organs.
5.5. Selection of beamlets and energy layers

The lower the numerical value of the beamlets, the less the numerical values of other beamlets need to change. Thus ideally the beamlets with a low numerical value are deleted given that the DVH is clinically acceptable.

5.5.2. Selection of energy layers

The same thought process and selection process as in the previous subsection can be used with energy layers instead of beamlets. Vector $\mathbf{y} = B\mathbf{x}$ will be introduced, the vector of $y_j$. The numerical value of $y_j = \sum_{x_i \in E_j} x_i$ is the sum of all $x_i$ in the $j$'th energy layer. $B$ is the matrix that is defined in Section 5.4.2. Then, the energy layer can be deleted in full if $y_j$ is small.

If the goal is to have a plan with $m'$ energy layers, the $m'$ energy layers with the largest value for $y_j$ will be selected and the other layers deleted.
Results

In this chapter the intermediate results and output plans after spot and/or energy layer reduction using different sparsity-inducing terms will be discussed. In all cases the goal of the output plans is to approximate the reference plan in number of beamlets, energy layers and in the 2D dose distribution. First the results of the $\ell_1$-norm will be analysed, followed by the mixed $\ell_1/\ell_2$-norm. Finally, the results of the addition of both of these norms will be shown. The weighted sum optimisation was performed with the weights given in Table 6.1.
Table 6.1: Weights for all the objectives in the weighted sum. PTV is planning targeting volume and Gp is gigaprotons

<table>
<thead>
<tr>
<th>Objectives</th>
<th>Name</th>
<th>Type</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1$</td>
<td>PTV-high</td>
<td>Maximum</td>
<td>3.30e-02</td>
</tr>
<tr>
<td>$f_2$</td>
<td>PTV-intermediate</td>
<td>Maximum</td>
<td>3.60e-03</td>
</tr>
<tr>
<td>$f_3$</td>
<td>PTV-low</td>
<td>Maximum</td>
<td>3.75e-01</td>
</tr>
<tr>
<td>$f_4$</td>
<td>Conformity ring PTV-high</td>
<td>Maximum</td>
<td>1.58e-02</td>
</tr>
<tr>
<td>$f_5$</td>
<td>Conformity ring PTV 0-10mm</td>
<td>Maximum</td>
<td>6.18e-02</td>
</tr>
<tr>
<td>$f_6$</td>
<td>Conformity ring PTV 10-15mm</td>
<td>Maximum</td>
<td>1.82e-03</td>
</tr>
<tr>
<td>$f_7$</td>
<td>Femoral heads</td>
<td>Maximum</td>
<td>4.08e-06-4.73e-03</td>
</tr>
<tr>
<td>$f_8$</td>
<td>Rectum</td>
<td>Mean</td>
<td>1.42e-01</td>
</tr>
<tr>
<td>$f_9$</td>
<td>Large and small intestines</td>
<td>Mean</td>
<td>3.45e-08-2.13e-01</td>
</tr>
<tr>
<td>$f_{10}$</td>
<td>Bladder</td>
<td>Mean</td>
<td>9.89e-02</td>
</tr>
<tr>
<td>$f_{11}$</td>
<td>Femoral heads</td>
<td>Mean</td>
<td>1.93e-02-2.09e-02</td>
</tr>
<tr>
<td>$f_{12}$</td>
<td>All conformity rings</td>
<td>Mean</td>
<td>3.96e-08-2.03e-03</td>
</tr>
<tr>
<td>$f_{13}$</td>
<td>Rest of the conformity rings</td>
<td>Maximum</td>
<td>1.75e-03-2.81e-03</td>
</tr>
<tr>
<td>$f_{14}$</td>
<td>Total spot weight</td>
<td>Mean</td>
<td>2.26e-07</td>
</tr>
</tbody>
</table>

6.1. $\ell_1$-norm

To examine the effect of the $\ell_1$-norm, the intermediate solutions using different weights for the sparsity inducing term will be compared to the result of the weighted sum without an added sparsity term. In this section addition of the $\ell_1$-norm with weights $\lambda = 1e-4$, $\lambda = 1e-6$, $\lambda = 1e-8$ and $\lambda = 1e-10$ will be compared. The DVHs are given in Figure 6.1.
6.1. $\ell_1$-norm

In the intermediate result with weight $\lambda = 1e-4$ large differences in the DVH can be seen. Adding the norm with this weight a different trade-off is made, as the minimisation of the $\ell_1$-norm weighed against minimising some maximum doses, which were not hard constraints in this formulation. This happens if the weight of the norm is higher than the weights of the objectives of the weighted sum. This, however, does not imply that the solution after the reduction of the beamlets and re-optimisation is not acceptable. But, it might also be possible
that the norm with a higher weight $\lambda$ caused the wrong beamlets to be selected. The intermediate plans with weights $\lambda = 1e-6$, $\lambda = 1e-8$ and $\lambda = 1e-10$ have very similar DVHs as the weighted sum. This is not surprising, as the weights $\lambda = 1e-8$ and $\lambda = 1e-10$ are lower than the lowest weight of the objectives and $\lambda = 1e-6$ is smaller than most weights in the weighted sum. As weight $\lambda = 1e-6$ is larger than some of the objectives it was expected that the dose distribution changed a bit more. The only noticeable difference is the change of dose in the femurs. Adding the sparsity term with one of these weights, the optimisation can be performed with enough freedom to minimise the maximum doses whilst trying to make the $\ell_1$-norm as low as possible.

### 6.1.1. Sparsity of the intermediate solutions

Now remains the question if the intermediate solutions are sparser than the weighted sum. In the table on the next page the number of beamlets that have a higher numeric value than $b = 0.01, 0.1, 1$ and 10, is shown, that is $|\{x_i : x_i \geq b\}|$. The lower the numbers, the sparser the solution. Ideally $|\{x_i : x_i \geq b\}|$ is small when $b$ is small and the DVH is almost identical. Then, all these beamlets can be deleted and since the beamlets already have a low value, the DVH will not change a lot.

Table 6.2: Indicates the number of beamlets that have a numerical value higher than $b = 0.01, 0.1, 1$ and 10, $|\{x_i : x_i \geq b\}|$. The total number of beamlets is 18777. $\lambda$ indicates the weight, that is used to add the $\ell_1$-norm to the weighted sum. "Weighted sum" and "$\epsilon$-constraint" are plans without a sparsity term calculated using a weighted sum and the sequential $\epsilon$-constraint optimisation.

|        | $b = 0.01$ | $b = 0.1$ | $b = 1$  | $b = 10$
|--------|------------|-----------|----------|---------
| $\lambda = 1e-4$ | 18757     | 5512      | 2363     | 1271    |
| $\lambda = 1e-6$ | 18777     | 16973     | 11193    | 2874    |
| $\lambda = 1e-8$ | 11868     | 5375      | 2149     | 1771    |
| $\lambda = 1e-10$ | 12137     | 5915      | 2228     | 1775    |
| weighted sum | 18777     | 17328     | 11850    | 4178    |
| $\epsilon$-constraint | 10698     | 3931      | 2042     | 1758    |

As can be seen in Table 6.2, the $\ell_1$-norm does indeed increase the sparsity in the solution as all values of the plans with a sparsity term are lower than or close to the values of the weighted sum without an added $\ell_1$-norm. It appears that the solution of the $\epsilon$-constraint optimisation is already quite sparse in comparison to the other plans, especially in comparison to the weighted sum method. By optimising over a different set of beamlets it appeared that this does not always
need to be the case.

As expected, the sparsity increases for weights $\lambda = 1e-10$ and $\lambda = 1e-8$ whilst keeping the DVH almost similar. Then for $\lambda = 1e-6$ the solution is less sparse in comparison to the plans with a lower weight, which is not what was expected. Then the sparsity is again increased for $\lambda = 1e-4$. In this case, the number of beamlets that is larger than $b = 0.01$, thus $|x_i| \geq 0.01$, is high, whereas the number of beamlets that is larger than $b = 10$ is lower than in every other method. Thus a lot more beamlets have a value in the interval $0.01 \leq x_i < 10$. It can be concluded that in order to reduce some beamlets, other beamlets cannot be reduced to almost zero. This is different from with weights $\lambda = 1e-8$ and $\lambda = 1e-10$, where a lot more beamlets could be reduced to a value of $x_i < 0.01$.

As a large number of beamlets have a low value in the intermediate results of $\lambda = 1e-8$, $\lambda = 1e-10$ and the $\epsilon$-constraint optimisation and the DVHs are similar, these plans will very likely result in an acceptable output plan. $\lambda = 1e-4$ does have a large number of beamlets that have a value $x_i < 1$ as well, however the corresponding DVH is not acceptable and it is not certain that re-optimisation will return an acceptable plan.

### 6.1.2. Spot reduction to 1705 beamlets

Now the results of the second sequential $\epsilon$-constraint optimisation with 1705 beamlets is shown. The beamlets are selected using the method described in Section 5.5 using different values of $b$ for each weight $\lambda$. By comparing the output plan and the reference plan both with the same number of beamlets, the quality of the norms with different weights can be discussed. Besides that the different values for $b$, that are chosen for the same reduction, can be compared.

<table>
<thead>
<tr>
<th>$b$</th>
<th>$\lambda = 1e-4$</th>
<th>$\lambda = 1e-6$</th>
<th>$\lambda = 1e-8$</th>
<th>$\lambda = 1e-10$</th>
<th>weighted sum</th>
<th>$\epsilon$-constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>2.3</td>
<td>47.7</td>
<td>22.0</td>
<td>21.6</td>
<td>60.6</td>
<td>17.0</td>
</tr>
</tbody>
</table>
$b$ for $\lambda = 1e^{-4}$ is very low. This means that all beamlets that are deleted already had a small numeric value. Due to the high weight of the norm present in the weighted sum, the quality of the plan calculated with $\epsilon$-constraint optimisation might increase. This is different from $b = 60.6$ in the weighted sum and $b = 47.7$ in the addition with weight $\lambda = 1e^{-10}$, where $b$ is quite large and the beamlets are selected more randomly. Then, the numeric values of the other beamlets need to change more and it is not certain this will result in an acceptable output plan. The DVHs of the plans with 1705 beamlets in comparison to the reference plan are given in the figures below.
The plan which was calculated after the selection based on the $\ell_1$-norm with weight $\lambda = 1e^{-4}$ differs from the reference plan quite a lot. The plan did not get better after the re-optimisation without the sparsity-inducing term. Looking at the DVH, the selection from the intermediate plan with $\lambda = 1e^{-4}$ did not result in an admirable selection as the maximum doses increased. The $\ell_1$-norm with a higher weight caused the wrong choices in the selection of beamlets to be made as suggested in the previous section.
The rest of the DVHs are nearly identical to the reference plan. In almost all plans the maximum dose of the left thighbone (Femur l) is increased. Besides that, there is a trade-off between the two thighbones (Femur l and Femur r). As the DVHs are nearly identical, it can be concluded that is possible to achieve a similar plan only using the result of the sequential $\epsilon$-constraint optimisation or the weighted-sum with or without the addition of an $\ell_1$-norm. But, it might be a lucky case that the solution of the $\epsilon$-constraint optimisation is sparse.

In the table below the number of similar selected beamlets of two intermediate plans is shown.

Table 6.4: Number of similar beamlets, that were selected from the intermediate plans. $\lambda$ indicates the weight, that is used to add the $\ell_1$-norm to the weighted sum. "WS" and "$\epsilon$-constr" are plans without a sparsity term calculated using a weighted sum and the sequential $\epsilon$-constraint optimisation.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>WS</th>
<th>$\epsilon$-constr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1e-4</td>
<td>342</td>
<td>308</td>
</tr>
<tr>
<td>1e-6</td>
<td>1376</td>
<td>1426</td>
</tr>
<tr>
<td>1e-8</td>
<td>1468</td>
<td>1542</td>
</tr>
<tr>
<td>1e-10</td>
<td>1469</td>
<td>1546</td>
</tr>
</tbody>
</table>

As suggested before, the plan with $\lambda = 1e-4$ selected very different beamlets than the rest of the plans. Only 300-400 beamlets out of the 1705 beamlets were similar as the selected beamlets of other plans. On the other hand, the beamlets that were selected from the plan with $\lambda = 1e-8$ and $\lambda = 1e-10$ are very similar and only 32 beamlets were not the same.

As the output plan with $\lambda = 1e-4$ is significantly worse than the other plans and is not clinically acceptable, this plan will not be considered in further reduction of the beamlets.

6.1.3. Further reduction of beamlets

It appears that a good plan with 1705 beamlets can be made using different intermediate plans even though according to Table 6.4, the beamlets that are selected are not the same. Thus it is interesting to reduce the beamlets even more and see if any differences between the plans arise. In this subsection the number of beamlets is reduced to 1260 beamlets. All beamlets larger than $b$ are selected with $b$ in Table 6.5.
Table 6.5: $b$ for each intermediate plan such that $\|x_i : x_i \geq b\| = 1260$ rounded off to one decimal.

<table>
<thead>
<tr>
<th>$\ell_1$, $\lambda$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1e-6$</td>
<td>121.0</td>
</tr>
<tr>
<td>$1e-8$</td>
<td>149.9</td>
</tr>
<tr>
<td>$1e-10$</td>
<td>145.5</td>
</tr>
<tr>
<td>weighted sum</td>
<td>126.0</td>
</tr>
<tr>
<td>$c$-constraint</td>
<td>149.0</td>
</tr>
</tbody>
</table>

As all $b$'s are quite large, deviations from the reference plan are expected. In the following figures, the DVHs are shown of the plans generated by 1260 beamlets.
Figure 6.3: Solution of sequential $\epsilon$-constraint optimisation using 1260 beamlets in comparison with the reference plan with 1705 beamlets. The selection of beamlets was performed on the solution of the weighted sum with addition of an $\ell_1$-norm with weight $\lambda$. (a) $\lambda = 1e-6$, (b) $\lambda = 1e-8$ and (c) $\lambda = 1e-10$, (d) the weighted sum without an extra term and (e) the sequential $\epsilon$-constraint optimisation of all 18777 beamlets.

Even with 1260 beamlets, the DVHs are quite similar. A difference in maximum dose appears in the DVH of the weighted sum and a small difference in the other DVHs. The weighted sum does not have enough freedom to keep the dose in PTV-low at 60 Gy whilst satisfying the dose constraints in the tumour. In all cases further reduction of the beamlets increases the maximum doses of PTV-low and
eventually all other volumes. This is shown in the next figure.

![Figure 6.4: DVHs of the \( \epsilon \)-constraint optimisation with 1260, 1150 and 1000 beamlets. The beamlets are selected from the result of the weighted sum with an \( \ell_1 \)-norm with weight \( \lambda = 1e^{-8} \).](image)

Even though \( b \) is larger for \( \lambda = 1e^{-8} \) than \( \lambda = 1e^{-10} \) and the weighted sum, this does not result in a worse plan. As stated before, the numerical value of some beamlets need to increase in order to decrease other beamlets. In case of weight \( \lambda = 1e^{-8} \), more beamlets have a smaller numerical value and therefore the numerical value of other beamlets have already changed. This way, the re-optimisation with the selected beamlets will result in a better plan. The next example will clarify this.

Consider beamlets \( x_1, x_2 \) and \( x_3 \) that have numerical values \( x_1 = 4, x_2 = 6 \) and \( x_3 = 12 \). Due to the sparsity term with a higher weight the new values are \( x_1 = 0, x_2 = 13, x_3 = 9 \) and with the lower weight \( x_1 = 3, x_2 = 7 \) and \( x_3 = 10 \). Then assume \( b = 10 \) for the higher weight and \( b = 8 \) for the lower weight. In the first case, \( x_1 \) and \( x_3 \) are deleted, whereas in the second case \( x_1 \) and \( x_2 \) are deleted. In both cases \( x_2 \) needed to increase to allow the other beamlets to become smaller. In the second case however the weight was not high enough to make \( x_3 < x_2 \). As a result when two \( x_j \)’s are deleted, \( x_2 \) is deleted instead of \( x_3 \) in the case of the lower weight.

### 6.2. \( \ell_1/\ell_2 \)-norm

Another goal is to reduce the number of energy layers that are used in the treatment of the patient. To reduce the number of energy layers it was suggested to add an \( \ell_1/\ell_2 \)-norm to the weighted sum. The norm is added with weights \( \lambda = 1e^{-4}, \lambda = 1e^{-6}, \lambda = 1e^{-8} \). The solutions of these optimisations are compared to the
solution of the weighted sum without the added norm. In the previous section it was concluded that the solution of the sequential $\epsilon$-constraint optimisation was a good plan to reduce the spots if the plan is already sparse. This might also be the case with energy layers. Therefore this plan will also be used to reduce the energy layers.

Figure 6.5: DVHs of solution of the weighted sum without a sparsity term and of the intermediate result of the weighted sum method with addition of an $\ell_1/\ell_2$-norm with weight $\lambda$. (a) $\lambda = 1e-4$, (b) $\lambda = 1e-6$ and (c) $\lambda = 1e-8$. 
6.2. \( \ell_1/\ell_2 \)-norm

Just like the \( \ell_1 \)-norm, the plan with the addition with weight \( \lambda = 1e^{-4} \) has the largest difference in the DVH with the weighted sum without the sparsity term. The other plans look quite similar to the weighted sum without the added \( \ell_1/\ell_2 \)-norm.

6.2.1. Selection of energy layers

Using matrix \( B \), introduced in Section 5.4.2, vector \( y \) is calculated. Recall that \( y \) is the vector of all \( y_j = \sum_{x_i \in E_j} x_i \), the sum of all \( x_i \) in every energy layer. The values for \( y_j \) of the solutions of the intermediate results are given in Appendix B. These values are sorted from lowest to highest for each intermediate plan. In the table an interesting result can be seen. The order of the first eight layers is the same for all plans and the values of the solution with weight \( \lambda = 1e^{-8} \) is quite similar to the solution of the \( \epsilon \)-constraint optimisation. In the layers 9 to 12 a minor change in order can be observed. The only large difference can be seen in layer 39. That layer became very large in the weighted sum with \( \lambda = 1e^{-4} \).

As the order of the energy layers is the same for most layers, deleting an arbitrary number of layers up to the eighth layer or all the lowest twelve layers without deleting spots will result in the same DVH. In the figures below the DVHs with the highest 37, 32 and 28 energy layers will be compared with the reference plan.
6. Results

Figure 6.6: DVHs with the reference plan and plans with (a) 37, (b) 32 and (c) 28 energy layers generated by sequential $\epsilon$-constraint optimisation.

The plan with 37 energy layers is very similar with the reference plan. Deleting more than these three layers does not change the DVH as much, only the dose in both the femurs is changed and the dose in the bladder increases as more energy layers are deleted. In all plans, more than 15000 beamlets are used in the optimisation, thus having more than enough freedom to make a plan. 19 or even more more layers can be deleted, when there is no selection on beamlets.

6.2.2. Beamlet and energy layer reduction

As the plans in the previous section have a large number of beamlets, the number of beamlets will be reduced as well. In Section 6.1 it was shown that the $\ell_1$-norm is not always necessary to reduce the number of beamlets. It was already covered that the deletion of the first 8 layers will be similar if the deletion was performed on the input result or the intermediate result with any $\ell_1/\ell_2$-norm. This is useful if the outcome plans after deletion of spots is compared, as the deletion of different energy layers cannot be the cause of different DVHs.
In Figure 6.7 the DVHs of plans with 5000 beamlets and 32 energy layers will be compared with the reference plan.

In contrast to the $\ell_1$-norm, the norm with weight $\lambda = 1e-4$ results in the best plan and all other plans have large deviations. This might be the result of the shift in numerical values of the beamlets in this plan. As the weight is higher, a larger number of numerical values of the beamlets became smaller. As a result the beamlets that should take over the contribution of the beamlets that are in the layers that will be deleted, are larger and thus not deleted.
6. Results

Figure 6.7: Solution of sequential $\epsilon$-constraint optimisation with 5000 beamlets and 32 energy layers in comparison with the reference plan with 1705 beamlets and 37 energy layers. The selection of beamlets was performed on the solution of the weighted sum with addition of an $\ell_1/\ell_2$-norm with weight $\lambda$. (a) $\lambda = 1e^{-4}$, (b) $\lambda = 1e^{-6}$ and (c) $\lambda = 1e^{-8}$ and (d) the sequential $\epsilon$-constraint optimisation of all 18777 beamlets.

6.3. $\ell_1$-norm and $\ell_1/\ell_2$-norm

The $\ell_1$-norm is a useful tool to reduce beamlets and the $\ell_1/\ell_2$-norm to select energy layers. However if both reductions are wanted, these norms individually are not satisfactory. In this section an improvement will be made on the plan with the selection on the intermediate plan of the $\ell_1/\ell_2$-norm with weight $\lambda = 1e^{-4}$.

6.3.1. Improvement on Figure 6.4

In Section 6.2.2 a plan was shown with 5000 beamlets in 32 energy layers. In this section the $\ell_1$-norm and $\ell_1/\ell_2$-norm will be used to find an equivalent plan with less beamlets.

For the removal of beamlets, an $\ell_1$-norm with weights $\lambda = 1e^{-6}$ and $\lambda = 1e^{-8}$ seem to give a good result. For the selection of energy layers it is better to use a higher weight on the $\ell_1/\ell_2$-norm, for example $\lambda = 1e^{-4}$. Therefore a selection was made from the intermediate result with $\lambda = 1e^{-8}$ for the $\ell_1$-norm and $\lambda = 1e^{-4}$ for the grouped norm.

The addition of the $\ell_1$-norm with $\lambda = 1e^{-8}$ to the $\ell_1/\ell_2$-norm with weight $\lambda = 1e^{-4}$ did however not change the number of beamlets that were needed for the optimisation. Looking at the beamlets that were selected of the intermediate plans it could be concluded that almost 100 percent of the beamlets that were selected were similar. Then $\lambda = 1e^{-6}$ was used for the $\ell_1$-norm. This did not improve the DVH as well. In this case 97 percent of the beamlets were similar to
6.3. $\ell_1$-norm and $\ell_1/\ell_2$-norm

the selection from the intermediate plan of the $\ell_1/\ell_2$-norm alone.
A smaller weight than $\lambda = 1e-4$ for the $\ell_1/\ell_2$-norm in combination with the $\ell_1$-norm with weight $\lambda = 1e-8$ or $\lambda = 1e-6$ gave worse results. Just like in the addition of the $\ell_1/\ell_2$-norm alone, the numeric values of the beamlets in a specific energy layer did not change enough that a good selection was made (see Section 6.2.2). Using a weight for the $\ell_1/\ell_2$-norm that is higher than $\lambda = 1e-4$ was to aggressive on the energy layers. For example with $\lambda = 1e-3$, if 5000 beamlets are selected, these beamlets were only in 17 different layers.

It seems as if the addition of a higher weight for the $\ell_1$-norm is better to influence the selection of beamlets when the $\ell_1/\ell_2$-norm is also present. Using a weight for the $\ell_1$-norm that was bigger than $\lambda = 1e-6$ did give some interesting results, for example in the case of $\lambda = 1e-5$. In that case, the number of beamlets that were small increased a lot and due to the $\ell_1/\ell_2$-norm the numbers of layers that had beamlets with a small numeric value increased even more. This resulted in the fact that, for example, the selection 4000 beamlets immediately included that only 30 energy layers were used instead of 37 energy layers with the $\ell_1$-norm alone or 34 with the $\ell_1/\ell_2$-norm alone. The DVH of the output plan is given below in comparison to both the reference plan and the output plan of the selection on the $\ell_1/\ell_2$-norm with weight $\lambda = 1e-4$. In order to keep the comparison fair, the plan is compared with a plan that contains 4000 beamlets in 30 energy layers.

The result is better than the result of only the $\ell_1/\ell_2$-norm in the doses of the
PTVs. However, the dose in the bladder increased in both plans a lot in comparison with the reference plan. The addition of both norms made a trade-off between the dose in the PTVs and the bladder in order to reduce the numeric values in the beamlets. Thus the weights of the norms should be closer to each other to have an effect and not too high such that too many energy layers are deleted.

6.3.2. Difference in number of beamlets

$\lambda = 1e^{-5}$ for the $\ell_1$-norm and $\lambda = 5e^{-5}$ for the $\ell_1/\ell_2$-norm are chosen. The DVHs of the output plans will be compared to both the reference plan and the output plan of the selection on the $\ell_1/\ell_2$-norm with weight $\lambda = 1e^{-4}$.

![Figure 6.9: DVH of the reference plan, selection of the intermediate result with an $\ell_1/\ell_2$-norm, weight $\lambda = 5e^{-5}$, and an $\ell_1$-norm, weight $\lambda = 1e^{-5}$ and selection from the intermediate result with an $\ell_1/\ell_2$-norm, weight $\lambda = 1e^{-4}$. The plans from the intermediate results have 3500 beamlets in 32 energy layers.](image)

The result is better in almost organs and PTVs. The dose in the bladder is slightly increased, but the maximum dose is the same. Further reduction of beamlets while keeping the number of energy layers the same was not possible. In the following figure the DVH of the output plan with the $\ell_1/\ell_2$-norm and both norms are shown containing 32 energy layers and a different number of beamlets, so that the DVHs are quite similar. The DVHs in the figure are quite similar even though the plan with both norms contains 3000 beamlets and the plan with only the $\ell_1/\ell_2$-norm 4000. Thus the combination of both norms indeed increases the ability to reduce the beamlets even further given a number of energy layers.

A similar result can be seen with 21 energy layers. Now, the plan with the selection on the intermediate plan with the $\ell_1/\ell_2$-norm
6.3. $\ell_1$-norm and $\ell_1/\ell_2$-norm

Figure 6.10: $\epsilon$-constraint optimisation of different selections of beamlets and the reference plan with 1705 beamlets in 37 energy layers. Selection of 3000 beamlets in 32 energy layers from the intermediate result with an $\ell_1/\ell_2$-norm, weight $\lambda = 5e^{-5}$, and an $\ell_1$-norm, weight $\lambda = 1e^{-5}$. Selection of 4000 beamlets in 32 energy layers from the intermediate result with an $\ell_1/\ell_2$-norm, weight $\lambda = 1e^{-4}$.

Figure 6.11: $\epsilon$-constraint optimisation of different selections of beamlets and the reference plan with 1705 beamlets in 37 energy layers. Selection of 3500 beamlets in 21 energy layers from the intermediate result with an $\ell_1/\ell_2$-norm, weight $\lambda = 5e^{-5}$, and an $\ell_1$-norm, weight $\lambda = 1e^{-5}$. Selection of 4500 beamlets in 21 energy layers from the intermediate result with an $\ell_1/\ell_2$-norm, weight $\lambda = 1e^{-4}$.

contains 4500 beamlets and the plan with the selection on both norms 3500 beamlets. In both cases the addition of both norms reduced the number of beamlets with 1000 beamlets.

From these figures it can be concluded, that there is a trade-off between the reduction of energy layers and beamlets. The more energy layers that are deleted, the more beamlets should be selected. It would be interesting to see, what combination of number of energy layers and beamlets is the most time efficient.

In all cases, the weights that are chosen might not give the best results, thus it might be possible that better output plans can be achieved by choosing differ-
ent weights. As the combination of the two norms has two weights that can be modified, these results can most likely be improved.
Conclusion

The numeric values of the beamlets could be influenced with the $\ell_1$-norm, so that after selecting beamlets with the highest values, the DVH of the output plan was similar to that of the reference plan. In this data set the input plan was already quite sparse. Thus the same result could be achieved by selecting beamlets on the solution of the initial $\epsilon$-constraint optimisation. Using the $\ell_1$-norm with a weight that is slightly lower than the weights of the objectives, the best result can be achieved with an output plan of 1260 beamlets. If a weight is chosen that is too high, there might be a selection of the wrong beamlets. When the weight is too low, the numeric values are not changed enough, resulting in a more random selection of beamlets.

The $\ell_1/\ell_2$-norm did not change the order of the energy layers in most cases. However it did influence the optimisation in such a way that the numeric values of the beamlets were shifted and, besides energy layers, beamlets could be deleted because of this shift. The addition of both norms resulted in a plan with less energy layers but still a larger number of beamlets than the reference plan, namely 5000 instead of 1705.

With the addition of both norms, the numeric values of the beamlets and energy layers can be reduced. However there is a trade-off between energy layers and beamlets. This resulted in an output plan with less beamlets than with the addition of the $\ell_1/\ell_2$-norm alone.

It might be possible to refine the exact weights to improve the output plans even
more. Furthermore, it can be interesting to investigate the radiation time of all plans so that the trade-off between the number of energy layers and beamlets can be made.
% calculate L2-norm of each group
L2_group = sqrt(B*(x.*x)); % vector m(groups)x1

% calculate function value, f and gradient, grad
f = sum(L2_group);
grad = x./(B'*L2_group); % vector n(beamlets)x1

% calculate Hessian
hess = diag(-x.*x); % matrix n(beamlets)xn

for i = 1:size(B,1)
    group = (-any(B(:,1:(end-i))~=B(:,(i+1):end)))';
    if sum(group) == 0 % then x_i and x_k are
        break % not in the same
    end % group anymore
    D = group.*(-x(1:(end-i)).*x((i+1):end));
    hess = hess + diag(D, i) + diag(D, -i);
end
```
hess = hess ./ (B'*(L2_group.^3)) + diag(grad./x);
```
Table of energy layers
Table B.1: \( \lambda \) indicates the weight, that is used to add the \( \ell_1/\ell_2 \)-norm to the weighted sum. "Weighted sum" and "\( \epsilon \)-constraint" are plans without a sparsity term calculated using a weighted sum and the sequential \( \epsilon \)-constraint optimisation. The number of "Layer" is the index of the beam. All numbers 1 to 18 are in the first beam and numbers 19 to 40 in the second beam. The lowest value of each beam corresponds to the layer with the lowest energy.

<table>
<thead>
<tr>
<th>#</th>
<th>( \lambda = 1e-4 )</th>
<th>( \lambda = 1e-6 )</th>
<th>( \lambda = 1e-8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0 40</td>
<td>0.1 40</td>
<td>0.1 40</td>
</tr>
<tr>
<td>2</td>
<td>2.9 19</td>
<td>4.5 19</td>
<td>1.4 19</td>
</tr>
<tr>
<td>3</td>
<td>12.3 1</td>
<td>7.6 1</td>
<td>2.0 1</td>
</tr>
<tr>
<td>4</td>
<td>20.1 20</td>
<td>299.0 20</td>
<td>318.6 20</td>
</tr>
<tr>
<td>5</td>
<td>167.9 2</td>
<td>707.6 2</td>
<td>669.6 2</td>
</tr>
<tr>
<td>6</td>
<td>1315.4 21</td>
<td>2322.9 21</td>
<td>2087.2 21</td>
</tr>
<tr>
<td>7</td>
<td>1661.3 22</td>
<td>3226.2 22</td>
<td>3078.8 22</td>
</tr>
<tr>
<td>8</td>
<td>2544.5 3</td>
<td>3754.1 3</td>
<td>3385.6 3</td>
</tr>
<tr>
<td>9</td>
<td>3394.8 23</td>
<td>5553.6 23</td>
<td>5625.2 24</td>
</tr>
<tr>
<td>10</td>
<td>3895.3 4</td>
<td>5848.5 4</td>
<td>5993.6 23</td>
</tr>
<tr>
<td>11</td>
<td>4035.1 5</td>
<td>6792.4 24</td>
<td>6254.5 4</td>
</tr>
<tr>
<td>12</td>
<td>4531.5 24</td>
<td>7675.0 5</td>
<td>8272.4 5</td>
</tr>
<tr>
<td>13</td>
<td>6305.6 25</td>
<td>8794.1 39</td>
<td>10468.1 25</td>
</tr>
<tr>
<td>14</td>
<td>7548.8 6</td>
<td>10712.6 25</td>
<td>10651.6 39</td>
</tr>
<tr>
<td>15</td>
<td>8152.7 27</td>
<td>10820.8 6</td>
<td>10877.0 6</td>
</tr>
<tr>
<td>16</td>
<td>8421.8 26</td>
<td>12464.7 27</td>
<td>12203.7 27</td>
</tr>
<tr>
<td>17</td>
<td>10238.3 28</td>
<td>13500.9 26</td>
<td>13015.1 14</td>
</tr>
<tr>
<td>18</td>
<td>11518.9 7</td>
<td>13562.5 14</td>
<td>13258.0 26</td>
</tr>
<tr>
<td>19</td>
<td>12381.2 29</td>
<td>14676.7 35</td>
<td>14773.5 35</td>
</tr>
<tr>
<td>20</td>
<td>13361.6 14</td>
<td>15145.2 28</td>
<td>15080.4 28</td>
</tr>
<tr>
<td>21</td>
<td>13580.7 8</td>
<td>17248.0 8</td>
<td>16762.5 18</td>
</tr>
<tr>
<td>22</td>
<td>14729.8 39</td>
<td>17395.1 18</td>
<td>16778.3 8</td>
</tr>
<tr>
<td>23</td>
<td>14948.8 18</td>
<td>18083.5 7</td>
<td>16860.0 7</td>
</tr>
<tr>
<td>24</td>
<td>16895.4 35</td>
<td>18660.6 29</td>
<td>18990.6 29</td>
</tr>
<tr>
<td>25</td>
<td>19610.6 30</td>
<td>19696.4 15</td>
<td>19020.7 15</td>
</tr>
<tr>
<td>26</td>
<td>22617.7 9</td>
<td>22812.0 30</td>
<td>23111.6 30</td>
</tr>
<tr>
<td>27</td>
<td>22866.9 31</td>
<td>24442.6 34</td>
<td>25804.8 34</td>
</tr>
<tr>
<td>28</td>
<td>27088.2 34</td>
<td>26780.5 9</td>
<td>27757.4 9</td>
</tr>
<tr>
<td>29</td>
<td>28650.4 15</td>
<td>27657.6 31</td>
<td>28188.5 31</td>
</tr>
<tr>
<td>30</td>
<td>28999.1 10</td>
<td>31398.2 36</td>
<td>32072.6 10</td>
</tr>
<tr>
<td>31</td>
<td>33114.0 13</td>
<td>33013.1 10</td>
<td>32715.4 36</td>
</tr>
<tr>
<td>32</td>
<td>35797.1 36</td>
<td>37562.7 13</td>
<td>37840.3 13</td>
</tr>
<tr>
<td>33</td>
<td>36916.5 32</td>
<td>39600.5 32</td>
<td>39311.2 32</td>
</tr>
<tr>
<td>34</td>
<td>40028.8 11</td>
<td>42813.7 16</td>
<td>40556.2 16</td>
</tr>
<tr>
<td>35</td>
<td>42152.8 33</td>
<td>43153.7 33</td>
<td>43921.4 11</td>
</tr>
<tr>
<td>36</td>
<td>49224.5 12</td>
<td>43555.0 11</td>
<td>43979.6 33</td>
</tr>
<tr>
<td>37</td>
<td>54991.0 38</td>
<td>46865.9 17</td>
<td>45252.7 38</td>
</tr>
<tr>
<td>38</td>
<td>57214.0 16</td>
<td>46973.6 38</td>
<td>46615.7 17</td>
</tr>
<tr>
<td>39</td>
<td>61394.5 17</td>
<td>54980.2 37</td>
<td>54310.7 37</td>
</tr>
<tr>
<td>40</td>
<td>73182.4 37</td>
<td>56759.8 12</td>
<td>57407.3 12</td>
</tr>
<tr>
<td>#</td>
<td>(\lambda = 1e-10)</td>
<td>(\text{Weighted sum})</td>
<td>(\epsilon)-constraint</td>
</tr>
<tr>
<td>----</td>
<td>----------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>y_j</td>
<td>y_j</td>
<td>y_j</td>
<td>y_j</td>
</tr>
<tr>
<td>1</td>
<td>0.2</td>
<td>40</td>
<td>11.0</td>
</tr>
<tr>
<td>2</td>
<td>3.0</td>
<td>19</td>
<td>178.0</td>
</tr>
<tr>
<td>3</td>
<td>4.3</td>
<td>1</td>
<td>285.1</td>
</tr>
<tr>
<td>4</td>
<td>322.0</td>
<td>20</td>
<td>708.8</td>
</tr>
<tr>
<td>5</td>
<td>672.8</td>
<td>2</td>
<td>991.7</td>
</tr>
<tr>
<td>6</td>
<td>2105.1</td>
<td>21</td>
<td>2446.9</td>
</tr>
<tr>
<td>7</td>
<td>3036.7</td>
<td>22</td>
<td>3616.5</td>
</tr>
<tr>
<td>8</td>
<td>3371.0</td>
<td>3</td>
<td>3790.3</td>
</tr>
<tr>
<td>9</td>
<td>5638.3</td>
<td>24</td>
<td>6220.9</td>
</tr>
<tr>
<td>10</td>
<td>5996.4</td>
<td>23</td>
<td>6434.3</td>
</tr>
<tr>
<td>11</td>
<td>6249.7</td>
<td>4</td>
<td>6485.7</td>
</tr>
<tr>
<td>12</td>
<td>8281.8</td>
<td>5</td>
<td>8662.2</td>
</tr>
<tr>
<td>13</td>
<td>10224.8</td>
<td>25</td>
<td>10119.4</td>
</tr>
<tr>
<td>14</td>
<td>10682.1</td>
<td>39</td>
<td>11182.7</td>
</tr>
<tr>
<td>15</td>
<td>10843.2</td>
<td>6</td>
<td>11369.7</td>
</tr>
<tr>
<td>16</td>
<td>12053.2</td>
<td>27</td>
<td>12900.4</td>
</tr>
<tr>
<td>17</td>
<td>12979.4</td>
<td>14</td>
<td>13430.4</td>
</tr>
<tr>
<td>18</td>
<td>13714.5</td>
<td>26</td>
<td>13814.0</td>
</tr>
<tr>
<td>19</td>
<td>14833.6</td>
<td>35</td>
<td>15442.0</td>
</tr>
<tr>
<td>20</td>
<td>15084.6</td>
<td>28</td>
<td>15571.4</td>
</tr>
<tr>
<td>21</td>
<td>16452.4</td>
<td>8</td>
<td>16276.8</td>
</tr>
<tr>
<td>22</td>
<td>16738.2</td>
<td>18</td>
<td>17684.5</td>
</tr>
<tr>
<td>23</td>
<td>16837.0</td>
<td>7</td>
<td>18329.5</td>
</tr>
<tr>
<td>24</td>
<td>19021.8</td>
<td>15</td>
<td>19763.9</td>
</tr>
<tr>
<td>25</td>
<td>19025.6</td>
<td>29</td>
<td>19880.1</td>
</tr>
<tr>
<td>26</td>
<td>23136.5</td>
<td>30</td>
<td>24446.5</td>
</tr>
<tr>
<td>27</td>
<td>25834.2</td>
<td>34</td>
<td>27322.8</td>
</tr>
<tr>
<td>28</td>
<td>28168.6</td>
<td>31</td>
<td>28609.8</td>
</tr>
<tr>
<td>29</td>
<td>28196.9</td>
<td>9</td>
<td>29151.1</td>
</tr>
<tr>
<td>30</td>
<td>32176.6</td>
<td>10</td>
<td>33099.8</td>
</tr>
<tr>
<td>31</td>
<td>32827.2</td>
<td>36</td>
<td>34044.3</td>
</tr>
<tr>
<td>32</td>
<td>37763.4</td>
<td>13</td>
<td>39194.1</td>
</tr>
<tr>
<td>33</td>
<td>39281.2</td>
<td>32</td>
<td>40333.3</td>
</tr>
<tr>
<td>34</td>
<td>40580.3</td>
<td>16</td>
<td>41197.1</td>
</tr>
<tr>
<td>35</td>
<td>43918.1</td>
<td>11</td>
<td>44027.5</td>
</tr>
<tr>
<td>36</td>
<td>43978.0</td>
<td>33</td>
<td>45198.3</td>
</tr>
<tr>
<td>37</td>
<td>45257.9</td>
<td>38</td>
<td>45233.3</td>
</tr>
<tr>
<td>38</td>
<td>46636.0</td>
<td>17</td>
<td>45402.3</td>
</tr>
<tr>
<td>39</td>
<td>54220.7</td>
<td>37</td>
<td>52043.5</td>
</tr>
<tr>
<td>40</td>
<td>57458.9</td>
<td>12</td>
<td>58704.7</td>
</tr>
</tbody>
</table>
Bibliography


