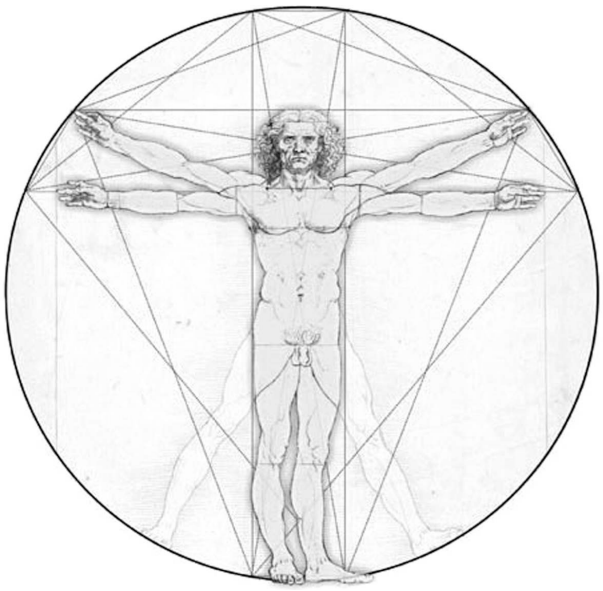
Radiation Therapy Treatment Planning with Uncertainties in Patient Position



Sandro Cerri
August 2006

TU Delft
Delft University of Technology

Erasmus MC
University Medical Center Rotterdam

Acknowledgments

Several people, in different ways, helped me writing this thesis.

My family, Dennis, my friends, Mitcy. And of course prof. Arnold Heemink and dr. Marleen Keijzer, who gave me the opportunity to work on this subject. Prof. Ben Heijmen, dr. Pascal Storchi and Sebastiaan Breedveld, who introduced me to this new world. Prof. Roger Cooke, dr. Dorota Kurowicka and the other students, for making me part of the MSc program "Risk and Environmental Modelling".

Many *grazie* to you all.

Contents

List of Figures

Glossary and Abbreviations

CTV	Clinical Target Volume
$D(x)$	vector of total doses delivered to the patient
$D_i(x)$	total dose delivered to voxel i
$D_{i,l}(x)$	dose delivered to voxel i in fraction l
DV	Dose-Volume
DVC	Dose-Volume Constraint
DVH	Dose-Volume Histogram
fraction	each of the session that makes a treatment
Gy	gray, the international system unit of absorbed dose
IMRT	Intensity-Modulated Radiation Therapy
IP	Interior-Point
LP	Linear Programming
m	number of voxels
M	number of beamlets (size of x)
n	number of scenarios (random or systematic)
N	number of fractions in the treatment
OAR	Organ At Risk (the healthy structures to be spared)
p	vector of scenario probabilities
PTV	Planning Target Volume
s	number of scenarios (random and systematic)
SOC	Second-Order Cone
x	vector of beamlet energies (fluence profile)

Chapter 1

Introduction

1.1 Cancer

Cancer develops when cells in a part of the body begin to grow out of control. Although there are many kinds of cancer (over 200 types), they all start because of out-of-control growth of abnormal cells, which results in serious health problems: at the moment, cancer is the leading cause of death among Americans under the age of 85. To get an idea, one in three people will develop cancer, and one in four will actually die of cancer.

Malignant tumors (figure 1.1) are collections of those cancerous cells that are growing faster or dividing faster than the normal cells around them.

The three modalities currently used for the treatment of cancer are surgery, chemotherapy and radiation therapy.

1.2 Radiation therapy

It is estimated that more than 50% of cancer patients will receive radiation therapy at some point during their treatment. Radiation therapy (also called radiotherapy, x-ray therapy, or irradiation) specifically acts against cells that are reproducing rapidly. Normal cells are programmed to stop dividing (reproducing) when they come into contact with other cells. In the case of a tumor, this stop mechanism is missing, causing cells to continue to divide over and over. It is the DNA of the cell that makes it capable of reproducing. Radiation therapy uses a certain type of high energy rays (x-rays or gamma rays) to damage the DNA of cells, thereby killing the cancer cells,

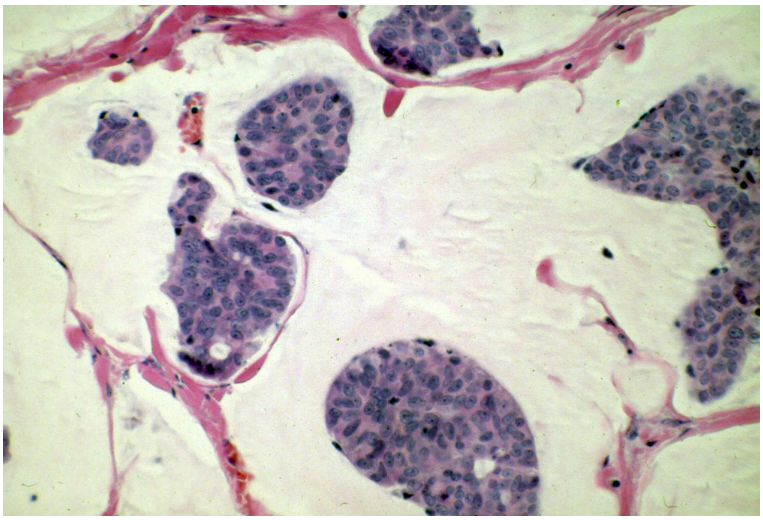


Figure 1.1: An enlarged picture of breast cancer cells.

or at least stopping them from reproducing. Radiation also damages normal cells, but as normal cells grow more slowly, they are better able to repair this radiation damage than cancer cells. In order to give normal cells time to heal and to reduce side effects, radiation treatments are typically given in small daily doses, five days a week, over a six- or seven-week period. Each of these treatment sessions is called a *fraction*. A treatment is therefore usually made of 30 up to 45 fractions.

As a result of new imaging and computer technology, the outcomes for radiotherapy have steadily improved over the last twenty years.

Radiation therapy is considered to be a "local" therapy, meaning it treats a specific localized area of the body. This is in contrast to systemic therapies, such as chemotherapy, which affect the whole body. There are two main types of radiation therapy: external radiation therapy, where a beam of radiation is directed from outside the body, and internal radiation therapy, also called brachytherapy or implant therapy, where a source of radioactivity is surgically placed inside the body near the tumor.

Figure 1.2: An Elekta (Philips) Linear Accelerator.

1.3 IMRT

Intensity-modulated radiation therapy (IMRT) is one of the latest and most advanced techniques of external radiation therapy that uses radiation beams (usually x-rays). The technology allows for the delivery of higher doses of radiation within the tumor and lower doses to nearby healthy tissue.

In IMRT, x-ray beams are aimed at a tumor from several (usually three or five) angles. During treatment, the radiation intensity of each beam is controlled. As a result, IMRT allows for the delivery of higher doses within the tumor, while sparing (as much as possible) important healthy tissues in a way that is impossible with other techniques.

Currently, IMRT is being used to treat cancers of the prostate, head and neck, breast, thyroid and lung, as well as in gynecologic, liver and brain tumors, lymphomas and sarcomas. IMRT is also beneficial for treating pediatric malignancies.

Radiotherapy treatment is usually given using a machine called linear accelerator, or LINAC (figure 1.2).

A medical linear accelerator generates the photons (x-rays) used in IMRT. The patient lies on the treatment table, while the linear accelerator delivers beams of radiation to the tumor from various directions. The intensity of each beam's radiation dose is dynamically varied according to a treatment plan

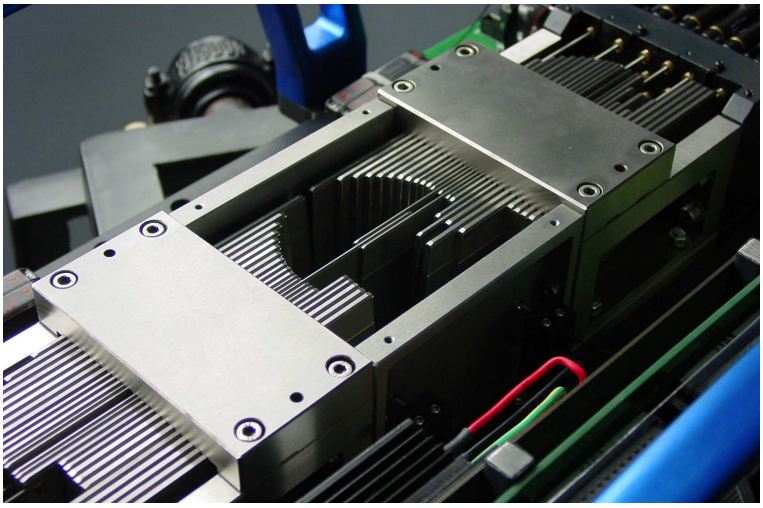


Figure 1.3: A multi-leaf collimator with 54 leaves: this device is used in radiotherapy for defining the shape of a treatment field. It also allows to vary the intensities in the field.

by a multi-leaf collimator, also known as an MLC. The multi-leaf collimator is an apparatus in the head of the treatment machine with thin leaves that automatically extend or retract to shape the beam (figure 1.3).

1.4 Treatment planning

The goal of radiation therapy treatment planning is to determine the intensity of the radiation for a certain patient, maximizing the dose to the tumor, while minimizing the dose to the surrounding healthy tissue.

Medical teams use CT (Computed Tomography, see figure 1.4) and MRI (Magnetic Resonance Images) scans of the patient, in order to develop a radiation therapy plan.

These scan images are usually made only once: based on these images the physicians locate the CTV (Clinical Target Volume, the tumor mass) and the OARs (Organs At Risk, the healthy tissues to be spared).

In order to account for both internal tumor motion and patient set-up error, the classical approach to IMRT involves the use of a PTV (Planning Target Volume): this is achieved placing a planning margin around the physician-defined CTV. For PTV margins that are in proximity to critical

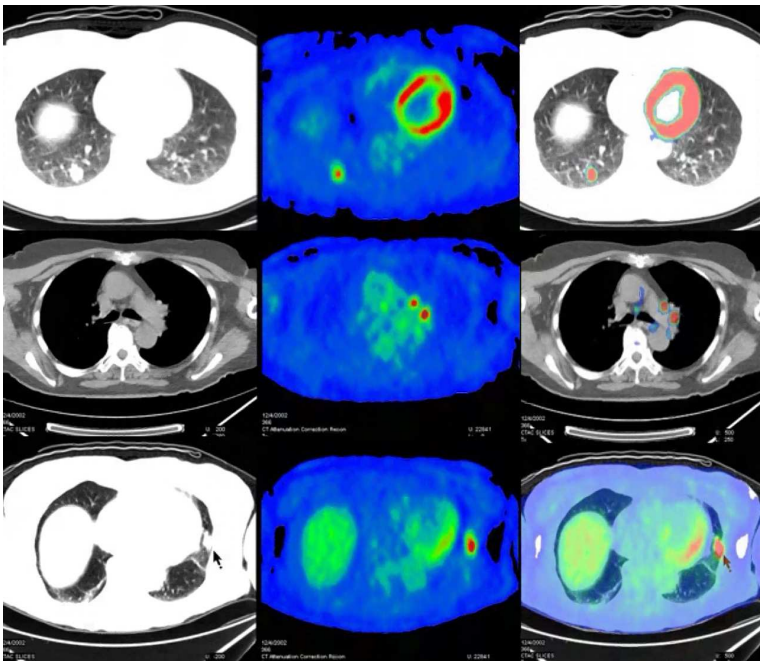


Figure 1.4: A computed tomography (CT) scan uses X-rays to make detailed pictures of structures inside the body. The word "tomography" is derived from the Greek "tomos" (slice) and "graphia" (describing).

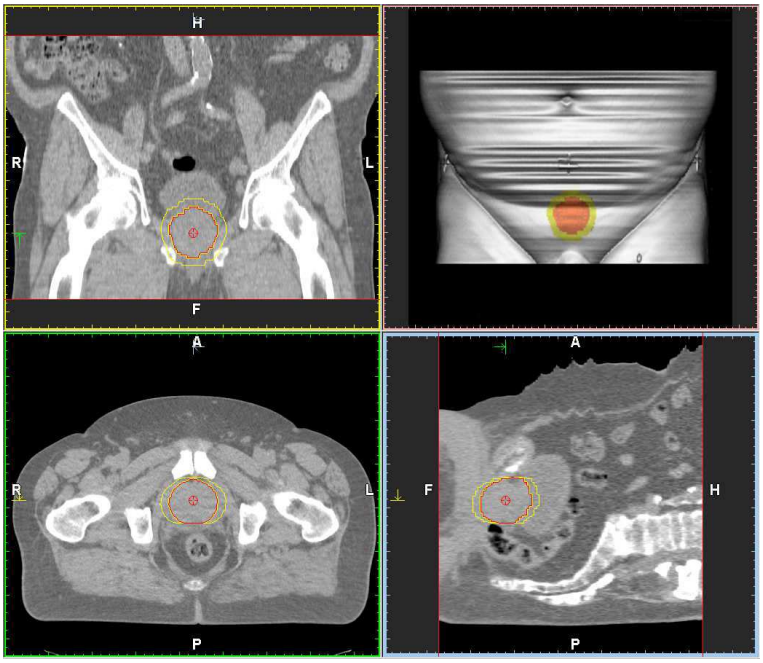


Figure 1.5: Automatic extraction of the PTV from the delineated CTV by applying user defined clinical margins. The extracted PTV and CTV are shown in different colors in the three major planes and in the 3D view. The margins used in this case are: 10 mm in the lateral and 10 mm in the cranio-caudal directions but 0 mm (that is, no margin) in the ventro-dorsal direction.

OARs, somewhat smaller margins are selected a priori in an effort to better spare the healthy organs, sometimes even at the expense of a lower PTV dose. Figure 1.5 shows an example of CTV and PTV definition on a real scan.

The ideal objective is a very high dose in the PTV and no dose outside the PTV. This is physically not possible. Determining an IMRT plan by trial and error is also not possible due to the complexity and the degrees of freedom of the system. Instead of the physicians defining beam directions, beam energy profile, their margins, etc., and then computing and displaying dose distributions to assess whether the treatment plan will lead to an acceptable outcome ("forward" planning), they do the opposite ("inverse" planning). They state their clinical objectives mathematically and let the

Structure	Function	Prescription	Importance
PTV	min dose	70 Gy	100
PTV	max dose	75 Gy	50
Bladder	max dose	60 Gy	80
Bladder	max dose-volume	40 Gy in 50%	70

Table 1.1: Example of treatment plan constraints.

IMRT optimization process determine the beam parameters that will lead to the desired solution.

The physicians usually set the following requirements for a plan:

- Minimum dose for the PTV. This is usually the most important constraint: the tumor is required to get a minimum prescribed dose so that the cancerous cells are killed.
- Maximum dose for the PTV. This constraint is set in order to prevent high radiation peaks in the PTV that may result in unnecessary damage, such as cells burning and the occurrence of holes.
- Maximum dose for the OARs. These are set in order to spare, as much as possible, healthy organs located in proximity of the tumor.
- Dose-volume constraints for the OARs, of the form "no more than a certain percentage of a certain structure may receive more than a certain dose".

Doses are measured in grays (abbreviated as Gy). One gray is equal to an absorbed dose of 1 joule per kilogram of absorber tissue. To give an idea of the magnitude of prescribed doses in real treatments, minimum doses for the PTV are often in the order of 60 up to 75 Gy; in comparison, the average cumulative radiation doses on humans coming from natural, medical, and occupational sources are estimated to be about 0.002 Gy per year.

For each of the aforementioned constraints, the physicians also set a weight (or priority) that measures the relative importance given to the constraint.

Table 1.1 reports an example of treatment plan constraints.

Chapter 2

Models

2.1 Forward model

In order to be able to optimize the dose delivered to a patient as a function of the fluence profile, a discretization of the system has to be performed. The patient's body is discretized into m "voxels". A voxel (a combination of the words "volume" and "pixel") is a volume element in a regular grid in the three-dimensional space. It is analogous to a pixel, which represents 2D image data. All voxels are grouped into a single m-by-1 (column) vector. To get an idea, voxels have in some cases size of 0.4
$textcm^3$, and for some treatments m may be in the order of 10,000.

We will work in a two-dimensional geometry[1], even if the results can be easily extended to the three-dimensional case. Suppose for the moment that only one beam is used in the treatment. The fluence profile of the beam is also discretized; we will indicate by x the M-by-1 (column) vector expressing the energy of the M beamlets of the beam.

The dose delivered to each voxel when fluence profile x is used, is indicated with the m-by-1 (column) vector $D(x)$. It can be shown that this dependence is linear in the sense that

$$D(x) = Hx,$$

where H is a m-by-M matrix. We will not explain in details why this relation is linear and on how matrix H can be calculated. The interested reader is referred, as a starting point, to [2]. It is worth mentioning however that H

[1]Therefore we will consider only one "slice" of the patient; in this case, to be precise, we should use the term "pixel" rather than "voxel".

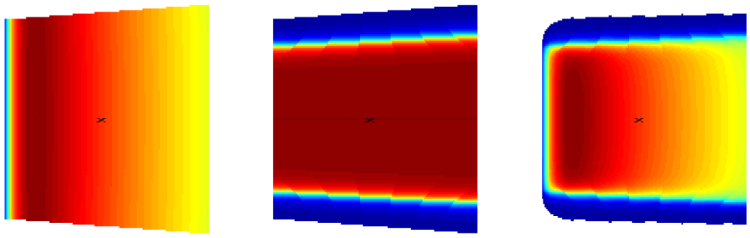


Figure 2.1: Depth (left) and off-axis (right) component of the dose hold structural properties that are somehow lost in the composition (right).

is not explicitly calculated in the algorithm, as we have

$$H = P \cdot K.$$

In practice, only P and K are calculated. Matrix P is related to the "depth dose", the component of the dose describing the dose decay along the beam axis; matrix K is, on the other hand, related to the "off-axis dose", the component of the dose describing the dose profile perpendicular to the beam axis. Figure 2.1 illustrates this idea.

The strong structure of each component translates into highly-structured, sparse matrices P and K: this is the reason why it is (computationally) convenient to keep these two matrices separate.

In the case of multiple beams (from different angles), matrices P and K will have a block-structure (one block for each beam direction), but all the results still apply.

2.2 Sources and models of errors

Measurements of the location of the tumor and OARs are prone to errors. There are many reasons why errors occur, but they are usually divided into two categories: random and systematic errors.

2.2.1 Systematic error

A systematic error is any biasing effect in the environment, methods of observation or instruments used, which introduces errors into an experiment and is such that it always effects the results of an experiment in the same

direction. Systematic error may appear, for example, because there is something wrong with the instrument (CT, LINAC, etc.) or its data handling system, or because the instrument is wrongly used by the physician. The so called "set-up error" is the most common systematic error appearing in IMRT: remember that usually only one CT scan is made at the beginning of the treatment. If the tumor is "dislocated" for a few centimeters from its real position, the treatment is misplanned. (The CTV/PTV has to be drawn by the physician, and the planning is made on the basis of his/her definition.) Figure 2.2 illustrates the problem.

2.2.2 Random error

Random error on the other hand is an error that does not result from a measurement method that is inherently wrong, but that occurs due to natural variation in the process. In our case, random error occurs during each fraction because, for example, the patient is not always perfectly placed on the table. Another source of random error are, for example, the breathing movements of the patient, or the level of filling of his/her bladder: these phenomena could shift some organs from their "average" position.

2.2.3 Effect of errors

It is intuitive that random errors tends to cancel each other out, while systematic errors are more difficult to handle. As an example, suppose that the position of the patient during each fraction is affected (only in one direction, as a first example) by random and/or systematic error; moreover suppose that both errors have standard normal distribution. Figure 2.2 displays an example of the effect of each error, and the combination of the two type of error.

2.3 Probabilistic model

As mentioned before, the classic approach to take into account the uncertainty in the patient position is to enlarge the area to be irradiated (from CTV to PTV). This approach, however, does not take into account OARs location uncertainty, but most importantly it does not take into account the

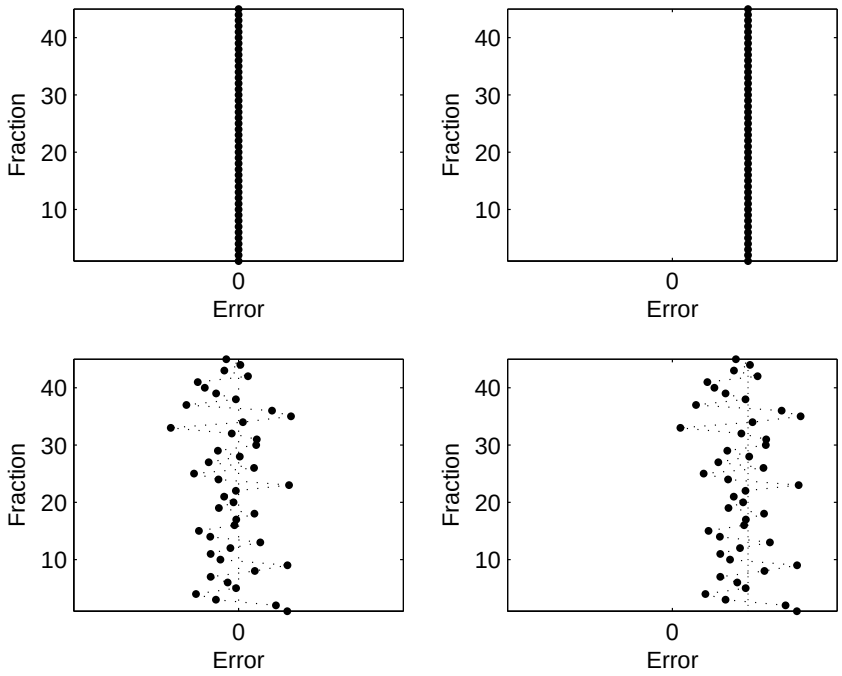


Figure 2.2: Top-left: no error. Top-right: systematic error. During each fraction the error is the same. Bottom-left: random error. During the course of the treatment, the average of the error tends to be zero, especially if the number of fractions is high. Bottom-right: combination of systematic and random error.

j_r or j_s	Direction	Shift (in cm)	Probability
0	-	0	0.4
1	Left	1	0.2
2	Right	1	0.2
3	Anterior	1	0.1
4	Posterior	1	0.1

Table 2.1: Scenarios (and their probabilities) of body dislocation.

correlated motions of CTV and OARs. Therefore, overlaps between PTV and OAR may occur, resulting in optimization problems.

The probabilistic approach that we investigate here tries, on the contrary, to incorporate such correlated uncertainty directly into the optimization. We follow the work proposed in [1], trying to extend that model in order to take into account both random and systematic error.

In this approach we discard the definition of PTV: the physician will therefore set minimum and maximum dose constraints directly for the CTV rather than for the PTV. The physician is also required to assess the scenarios of the random/systematic errors that may occur during the treatment. The number of scenarios is indicated with n, and an example of scenarios are given in table 2.1. Random scenarios are indexed by j_r, while systematic scenarios by j_s. Both j_r and j_s take value in $\{0, \dots, n-1\}$, where index 0 refers to the scenario with no dislocation.

If the n scenarios are supposed to independently model both random and systematic scenarios, then the total number s of scenarios that may occur during different patient treatments is at most n^2. For example, considering the scenarios given in table 2.1 to be valid both (but separately) for random and systematic errors, we have $s = 13$ ($< n^2 = 25$). See figure 2.3 for a graphical representation of this idea.

2.3.1 A tempting idea

We assume, from now on, that the number of fractions (indicated with N) is determined in advance by the physicians, and therefore is an input of the optimization process. Also the number and the angle of the beams are supposed to be determined in advance. The variable to be optimized is therefore the (column) vector of beamlet intensities, indicated with x. The

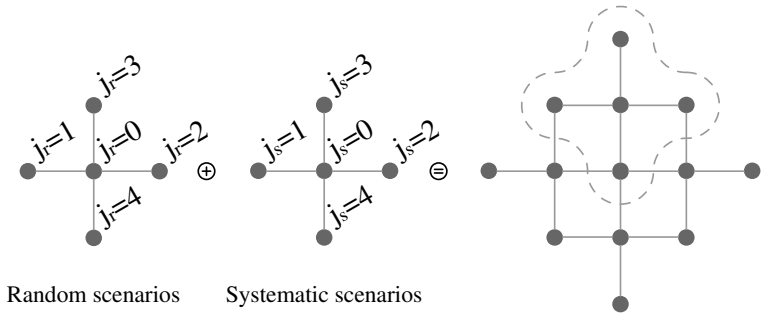


Figure 2.3: If the scenarios of table 2.1 model both (independently) random and systematic errors, the actual scenarios that may occur in the course of *several* treatments are the ones given on the right. Note that during one *particular* treatment (made of several fractions) not all 13 scenarios can occur, but only 5 of them (for example, if $j_s = 3$, the ones within the dashed line).

size of x is indicated by M (number of beamlets).

Indicate with $a^{(i)}_{j_s,j_r}$ the 1-by-M (row) vector, indexed by beamlets, giving the (deterministic) dose delivered to voxel i in systematic-scenario j_s and random-scenario j_r if the beamlets have unit intensity. Doses $a^{(i)}_{j_s,j_r}$ can be calculated using the forward model described in section 2.1. Remember that because of linearity, the dose delivered to voxel i in systematic-scenario j_s and random-scenario j_r if the beamlets have profile x, would simply be $a^{(i)}_{j_s,j_r}x$.

At this point one would be tempted to do the following reasoning: if we want to impose, for example, the constraint "the CTV should receive a minimum (total) dose of m_{Tmin} Gy", we could impose the following linear constraints:

$$a^{(i)}_{j_s,j_r}x \geq \frac{m_{\mathrm{Tmin}}}{N} \qquad (j_s, j_r = 0, \ldots, n-1; \forall i \in \mathrm{CTV}). \tag{2.1}$$

Translating into words: we impose the fraction dose (no matter the systematic and random scenario occurring) to exceed one Nth of the minimum total dose. Doing so, the total dose delivered after N fractions will definitely be greater than m_{Tmin}, as required.

To a closer look, however, these constraints are too restrictive, as this approach completely discards the information we have about the probabilities of occurrence of each scenario: we are imposing constraints on the "marginal" fraction dose rather than on the "joint" total dose.

Having such constraints could, however, be important from a medical point of view, since they would introduce some control over the homogeneity of the delivered doses among fractions. For the moment we will follow another route in order to model the "minimum/maximum organ dose", but we will come back to this idea in section 2.4.3.

2.3.2 Expected value and variance of the fraction dose

Suppose that the systematic error (for one particular treatment) is known, and therefore that the n (out of the s) scenarios that are likely to occur are determined. In other words, fix j_s: looking back at figure 2.3, this means that we know that we are dealing, for example, with only the five scenarios within the dashed line.

Now, for each voxel i and for each systematic-scenario j_s, it is convenient to build the n-by-M matrix

$$A_{j_s}^{(i)} = \begin{pmatrix} a_{j_s,0}^{(i)} \\ \vdots \\ a_{j_s,n-1}^{(i)} \end{pmatrix}.$$

Indicate with $S(l)$ the random variable that returns the index of the random-scenario occurring in fraction l ($l = 1, \ldots, N$), according to the distribution specified by the physician:

$$S(l) = \begin{cases} 0 & \text{with probability } p_0 \\ \vdots & \vdots \\ n-1 & \text{with probability } p_{n-1}. \end{cases}$$

Indicate[2] with $D_{i,l}(x)$ the dose delivered to voxel i during fraction l if the beamlets have profile x; in our probabilistic formulation (contrary to the classic static formulation) this is a random variable, and we can write

$$D_{i,l}(x) = e_{S(l)}^T A_{j_s}^{(i)} x = \begin{cases} e_0^T A_{j_s}^{(i)} x & \text{with probability } p_0 \\ \vdots & \vdots \\ e_{n-1}^T A_{j_s}^{(i)} x & \text{with probability } p_{n-1} \end{cases} \tag{2.2}$$

[2]Here we will omit, for better readability, the dependence on j_s.

where e_k indicates the $(k+1)$-th standard basis (column) vector of $\mathbb{R}^n$. From this, it is easy to write an expression for the expected value of this random variable:

$$\mathbb{E}[D_{i,l}(x)] = \sum_{j_r=0}^{n-1} p_{j_r} e_{j_r}^T A_{j_s}^{(i)} x = p^T A_{j_s}^{(i)} x,$$

where p is the n-by-1 (column) vector of scenario probabilities $(p_0, \ldots, p_{n-1})^T$.

It is also easy to write an expression for the variance of this dose:

$$\begin{aligned}
\mathbb{V}\mathrm{ar}[D_{i,l}(x)] &= \mathbb{E}[(D_{i,l}(x) - \mathbb{E}[D_{i,l}(x)])^2] \\
&= \sum_{j_r=0}^{n-1} p_{j_r} (e_{j_r}^T A_{j_s}^{(i)} x - p^T A_{j_s}^{(i)} x)^2 \\
&= \sum_{j_r=0}^{n-1} (\sqrt{p_{j_r}} (e_{j_r}^T - p^T) A_{j_s}^{(i)} x)^2 \\
&= ||R A_{j_s}^{(i)} x||^2,
\end{aligned}$$

where

$$R = \begin{pmatrix} \sqrt{p_0} & 0 & \cdots & 0 \\ 0 & \sqrt{p_1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{p_{n-1}} \end{pmatrix} \begin{pmatrix} 1-p_0 & -p_1 & \cdots & -p_{n-1} \\ -p_0 & 1-p_1 & & \vdots \\ \vdots & & \ddots & -p_{n-1} \\ -p_0 & \cdots & -p_{n-2} & 1-p_{n-1} \end{pmatrix}.$$

2.3.3 Approximation of the distribution of the total dose

Indicate with $D_i(x)$ the dose delivered to voxel i during the whole treatment if the beamlets have profile x. Again, this is a random variable, and we assume that fraction doses are additive:

$$D_i(x) = \sum_{l=1}^{N} D_{i,l}(x).$$

The doses $D_{i,l}(x)$ delivered during the various fractions ($l = 1, \ldots, N$) are assumed to be independent and are identically distributed. Therefore

the expected value of the total delivered dose is simply[3]:

$$\begin{aligned}
\mathbb{E}[D_i(x)] &= \mathbb{E}\left[\sum_{l=1}^{N} D_{i,l}(x)\right] && (2.3)\\
&= \sum_{l=1}^{N} \mathbb{E}[D_{i,l}(x)] && (2.4)\\
&= \sum_{l=1}^{N} pA_{j_s}^{(i)} x && (2.5)\\
&= NpA_{j_s}^{(i)} x. && (2.6)
\end{aligned}$$

For the variance we have[4]:

$$\begin{aligned}
\mathbb{V}\mathrm{ar}[D_i(x)] &= \mathbb{V}\mathrm{ar}\left[\sum_{l=1}^{N} D_{i,l}(x)\right] && (2.7)\\
&= \sum_{l=1}^{N} \mathbb{V}\mathrm{ar}[D_{i,l}(x)] && (2.8)\\
&= \sum_{l=1}^{N} ||RA_{j_s}^{(i)} x||^2 && (2.9)\\
&= N||RA_{j_s}^{(i)} x||^2. && (2.10)
\end{aligned}$$

It is well-known (see figure 2.4 for an example) that the sum of (many) independent and identically distributed (i.i.d.) random variables approximates a normal random variable. In our framework, N is in the order of 45, so this approximation is plausible.

A normal random variable is completely determined given its expected value and standard deviation. Therefore, using (2.6) and (2.10), we can write

$$D_i(x) \approx Z(NpA_{j_s}^{(i)} x, \sqrt{N}||RA_{j_s}^{(i)} x||), \qquad (2.11)$$

where $Z(\mu, \sigma)$ indicates a normal random variable with mean μ and standard deviation σ.

[3]Remember that the expected value operator $\mathbb{E}[\cdot]$ is linear in the sense that $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ for any two random variables X and Y (which may or may not be independent) and any real numbers a and b.

[4]Here we use the fact that if X and Y are independent random variables, then $\mathbb{V}\mathrm{ar}[X + Y] = \mathbb{V}\mathrm{ar}[X] + \mathbb{V}\mathrm{ar}[Y]$.

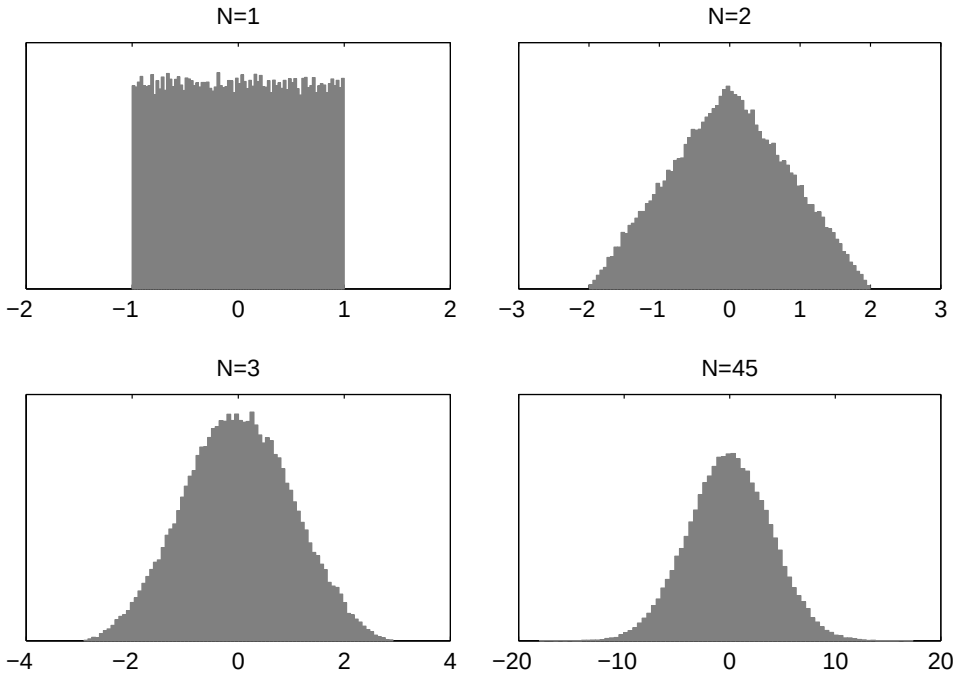


Figure 2.4: The sum of many independent identically distributed random variables (of finite variance) will tend to be distributed according to a Normal random variable. As an example, we present the distribution of the sum of 2 (top-right), 3 (bottom-left) and 45 (bottom-right) independent copies of the random variable with distribution given in the top-left subfigure.

2.4 Constraint models

2.4.1 Minimum and maximum dose constraints

Suppose that we want to mathematically express: "the CTV should receive a total minimum dose of m_{Tmin} Gy" in our probabilistic framework. A first idea would be to express this in terms of expected total dose:

$$\mathbb{E}[D_i(x)] \geq m_{\text{Tmin}} \qquad (\forall i \in \text{CTV}).$$

Using expression(2.6) this would result in a linear inequality constraint:

$$NpA_{j_s}^{(i)} x \geq m_{\text{Tmin}} \qquad (\forall i \in \text{CTV}).$$

We can actually do better than this: when dealing with lives, thinking "on average" is not enough. A treatment plan that is good "on average" is *not* a good treatment plan. If we have, for example, $\mathbb{E}[D_i(x)] = m_{\text{Tmin}}$, then half of the times, that treatment would not deliver enough dose to voxel i.

Remember that we not only have an expression for the expected total dose, but also an approximation of the distribution of the total dose. Therefore, by using expression (2.11), we can express the aforementioned constraint in the following probabilistic terms:

$$P\left(D_i(x) \geq m_{\text{Tmin}}\right) \geq \delta \qquad (\forall i \in \text{CTV}), \tag{2.12}$$

where δ is a predefined constant, say 0.95, expressing our desired "confidence" in delivering the prescribed dose. Note that because of (2.11), we approximate (2.12) by

$$P\left(Z(NpA_{j_s}^{(i)} x, \sqrt{N}||RA_{j_s}^{(i)} x||) \geq m_{\text{Tmin}}\right) \geq \delta \qquad (\forall i \in \text{CTV}).$$

Ideally we would require $\delta = 1$; because of the previous normal approximation, that would however result in a unfeasible formulation, since any normal random variable has a strictly positive density throughout the whole $\mathbb{R}$.

Remembering that if $Z(\mu, \sigma)$ is a normal random variable with mean μ and standard deviation σ, then $\frac{Z(\mu,\sigma)-\mu}{\sigma}$ is a standard normal random variable, we can rewrite

$$P\left(Z(0,1) \geq \frac{m_{\text{Tmin}} - NpA_{j_s}^{(i)} x}{\sqrt{N}||RA_{j_s}^{(i)} x||}\right) \geq \delta \qquad (\forall i \in \text{CTV}),$$

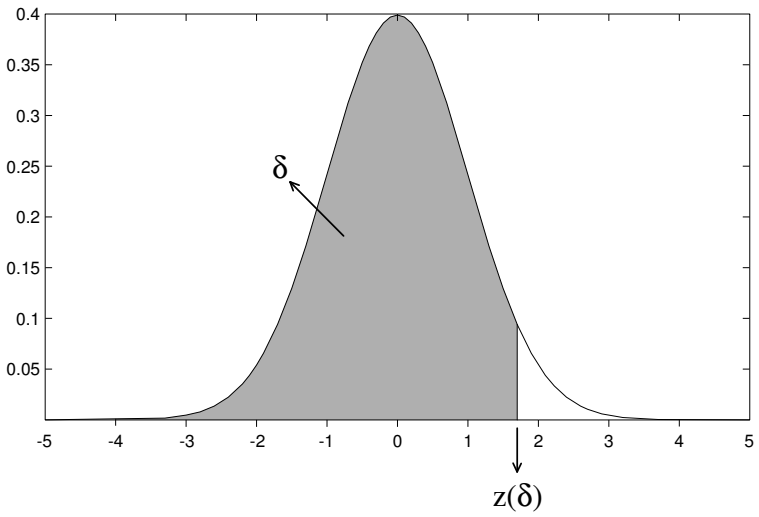


Figure 2.5: Graphical interpretation of $z(\delta)$.

which is equivalent to

$$\frac{NpA_{j_s}^{(i)}x - m_{\mathrm{Tmin}}}{\sqrt{N}||RA_{j_s}^{(i)}x||} \geq z(\delta) \qquad (\forall i \in \mathrm{CTV}), \tag{2.13}$$

where $z(\delta)$ is the number such that $P(Z(0,1) < z(\delta)) = \delta$ (see figure 2.5).

Re-shuffling the terms, we get the following inequality:

$$||RA_{j_s}^{(i)}x|| \leq \frac{NpA_{j_s}^{(i)}x - m_{\mathrm{Tmin}}}{z(\delta)\sqrt{N}} \qquad (\forall i \in \mathrm{CTV}), \tag{2.14}$$

which is a second-order constraint that can be handled in the SOCP (Second-order Cone Programming) framework (see section 2.6). Note that in order to get the inequality (2.14), we had to divide each side of inequality (2.13) by $z(\delta)$ which, for δ close to 1 (to be precise for any $\delta > 0.5$) is a positive number.

Similar calculations allow us to express the constraint "maximum dose for organ-at-risk OAR_k" and "maximum dose for the CTV" respectively like

$$||RA_{j_s}^{(i)}x|| \leq \frac{m_k - NpA_{j_s}^{(i)}x}{z(\delta)\sqrt{N}} \qquad (\forall i \in \mathrm{OAR}_k) \tag{2.15}$$

and

$$||RA_{j_s}^{(i)}x|| \leq \frac{m_{\mathrm{Tmax}} - NpA_{j_s}^{(i)}x}{z(\delta)\sqrt{N}} \qquad (\forall i \in \mathrm{CTV}), \tag{2.16}$$

where m_k and m_{Tmax} are, respectively, the total maximum dose for organ OAR_k and total maximum dose for the CTV. For each constraint a different value for δ could be chosen; however, for now, we have used $\delta = 0.95$ everywhere.

2.4.2 DV constraints

The other kind of constraints that we need to express are the dose-volume constraints, of the form: "no more than $100{\cdot}v_k\%$ of organ-at-risk OAR_k may receive more than d_k Gy". These constraint, in mathematical terms, would read

$$\frac{1}{|\mathrm{OAR}_k|} \sum_{i \in \mathrm{OAR}_k} I_{\{D_i(x) > d_k\}} \leq v_k \qquad (k = 1, \ldots, h; \forall i \in \mathrm{OAR}_k), \tag{2.17}$$

where $|\mathrm{OAR}_k|$ indicates the number of voxels of organ OAR_k.

But remember that in our probabilistic framework, $D_i(x)$ is a random variable, therefore the term $D_i(x) \geq d_k$ should be replaced (as done for the minimum and maximum constraints, see previous section) by a probability statement. However, in order to also keep the complexity of the optimization as low as possible, in this case we prefer to replace $D_i(x) \geq d_k$ with $\mathbb{E}[D_i(x)] \geq d_k$, rather than with $P(D_i(x) \geq d_k) \geq \delta$. Therefore, using again (2.6) we approximate (2.17) by

$$\frac{1}{|\mathrm{OAR}_k|} \sum_{i \in \mathrm{OAR}_k} I_{\{NpA_{j_s}^{(i)}x > d_k\}} \leq v_k \qquad (k = 1, \ldots, h; \forall i \in \mathrm{OAR}_k).$$

Unfortunately indicator functions are discontinuous functions that are not very well suited for optimization problems. An idea might be to approximate the indicator function with a continuous piece-wise linear function (see also figure 2.6.):

$$I_{\{x>0\}} \approx \min(\max(x, 0), 1).$$

This approach, however, introduces non-convex constraints.

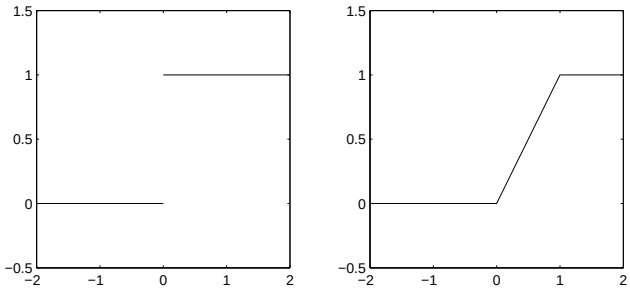


Figure 2.6: An approximation of the indicator function.

Another way to handle DV constraint (used for example in [1]), is by imposing

$$\frac{1}{|\mathrm{OAR}_k|}\sum_{i\in\mathrm{OAR}_k}\max(NpA^{(i)}_{j_s}x-d_k,0)\le v_k(m_k-d_k)$$
$$(k=1,\dots,h;\forall i\in\mathrm{OAR}_k). \quad (2.18)$$

Luckily it is possible to transform such constraints into convex standard linear constraints.

2.4.3 Additional constraints

Minimum scenario dose

As mentioned before, a physician may require some homogeneity control over the fraction dose delivered to the CTV: recycling the idea presented in section 2.3.1, this could be done imposing

$$a^{(i)}_{j_s,j_r}x\ge\alpha\frac{m_{\mathrm{Tmin}}}{N}\qquad(j_s,j_r=0,\dots,n-1;\forall i\in\mathrm{CTV}).$$

where the α is a number smaller than 1 (say 0.8) introduced in order not to "overrule" constraint (2.14).

Fluence profile constraints

The vector x of fluence energy profile should obviously be positive:

$$x\ge 0.$$

If this is the only constraint on the fluence, however, the solution given by the optimization will most likely show high peaks and large energy differences between adjacent beamlets. This is not acceptable from a medical point of view. To solve this problem, extra constraints are imposed. One idea is to limit to ϵ the difference that the energy that adjacent beamlets are allowed to have. One way to express this is by defining the M-by-M matrix

$$T = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 \\ -1 & 1 & 0 & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix},$$

the M-by-1 (column) vector

$$e = \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{pmatrix}$$

and by imposing

$$-\epsilon \cdot e \leq Tx \leq \epsilon \cdot e.$$

Another idea is to make the fluence profile smooth by imposing that the (discrete) second derivative is close to zero.

The aforementioned constraints are valid in case of a single beam. In case of multiple beams, the matrices involved will be block matrices, where the number of blocks equals the number of beam angles and each block has the structure explained before. Figure 2.4.3 gives a comparison of fluence profile solutions with and without "smoothing constraints" imposed.

2.5 Objective function and formulation

The objective function we will minimize is

$$\sum_{j_s=0}^{n-1} w_{\mathrm{Tmin}}(j_s) r_{\mathrm{Tmin}}(j_s) + \sum_{j_s=0}^{n-1} w_{\mathrm{Tmax}}(j_s) r_{\mathrm{Tmax}}(j_s) + \sum_{k=1}^{h} w_k r_k + \sum_{k=1}^{h} \overline{w}_k q_k \quad (2.19)$$

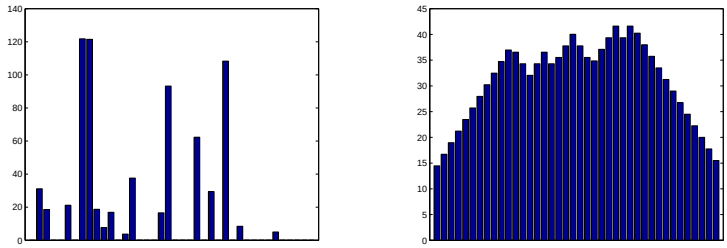


Figure 2.7: On the left, an example of an optimal fluence profile that we would get without imposing the smoothing constraints. On the right, an example of an optimal fluence profile obtained imposing the smoothing constraints.

subject to the following constraints[5]:

$$||RA_{j_s}^{(i)}x|| \leq \frac{NpA_{j_s}^{(i)}x - u_{\text{Tmin}}(j_s)}{z(\delta)\sqrt{N}} \qquad (\forall j_s; \forall i \in \text{CTV}) \qquad (2.20)$$

$$m_{\text{Tmin}} - u_{\text{Tmin}}(j_s) \leq r_{\text{Tmin}}(j_s) \qquad (\forall j_s) \qquad (2.21)$$

$$r_{\text{Tmin}}(j_s) \geq 0 \qquad (\forall j_s) \qquad (2.22)$$

$$||RA_{j_s}^{(i)}x|| \leq \frac{u_{\text{Tmax}}(j_s) - NpA_{j_s}^{(i)}x}{z(\delta)\sqrt{N}} \qquad (\forall j_s; \forall i \in \text{CTV}) \qquad (2.23)$$

$$u_{\text{Tmax}}(j_s) - m_{\text{Tmax}} \leq r_{\text{Tmax}}(j_s) \qquad (\forall j_s) \qquad (2.24)$$

$$r_{\text{Tmax}}(j_s) \geq 0 \qquad (\forall j_s) \qquad (2.25)$$

$$||RA_{j_s}^{(i)}x|| \leq \frac{u_k - NpA_{j_s}^{(i)}x}{z(\delta)\sqrt{N}} \qquad (j_s = 0; \forall k; \forall i \in \text{OAR}_k) \qquad (2.26)$$

$$u_k - m_k \leq r_k \qquad (\forall k) \qquad (2.27)$$

$$r_k \geq 0 \qquad (\forall k) \qquad (2.28)$$

$$Na_{j_s,j_r}^{(i)}x \geq \alpha m_{\text{Tmin}} \qquad (\forall j_s, j_r; \forall i \in \text{CTV}) \qquad (2.29)$$

$$\frac{1}{|\text{OAR}_k|}\sum_{i \in \text{OAR}_k} \max(NpA_{j_s}^{(i)}x - d_k, 0) \leq f_k \qquad (j_s = 0; \forall k; \forall i \in \text{OAR}_k) \qquad (2.30)$$

$$f_k - v_k(m_k - d_k) \leq q_k \qquad (\forall k) \qquad (2.31)$$

$$q_k \geq 0 \qquad (\forall k) \qquad (2.32)$$

$$-\epsilon \cdot e \leq Tx \leq \epsilon \cdot e \qquad (2.33)$$

$$x \geq 0 \qquad (2.34)$$

[5]Here, to keep the notation more compact, when we write $\forall j_s$, $\forall j_r$ and $\forall k$ we mean, respectively, $j_s = 0, \ldots, n-1$, $j_r = 0, \ldots, n-1$ and $k = 1, \ldots, h$.

Note that constraint (2.14) has been replaced by constraints (2.20), (2.21) and (2.22): u_{Tmin} has been introduced as an intermediate variable, so that r_{Tmin} (present also in the objective function) represents the penalizing excess dose $m_{\text{Tmin}} - u_{\text{Tmin}}$. The same applies to (2.16) (replaced by (2.23), (2.24) and (2.25)), (2.15), (replaced by (2.26), (2.27) and (2.28)) and (2.18) (replaced by (2.30), (2.31) and (2.32)).

Moreover, note that in constraints (2.26) and (2.30) (the ones regarding the OARs) we are working only on the systematic scenario $j_s = 0$, namely the scenario that implies no systematic error (check again, for example, figure 2.3). This is done to keep the computational burden as low as possible. However, if the resulting solution is not satisfactory, the above formulation can be easily extended to include all systematic scenarios not only in the CTV constraints, but also in the OAR constraints.

The w's in the objective function (2.19) are weight factors related to the importance weights described in section 1.4 (see, for an example, table 1.1). Because of the formulation we used for the CTV, however, w_{Tmin} and w_{Tmax} now depends on j_s. Also note that until now we are treating each systematic scenario equally, as the information we have on p (the vector of scenario probabilities) has only been included in the conic constraints (the ones modeling *random* scenarios). Therefore one idea is to choose:

$$w_{\text{Tmin}}(j_s) \propto p_{j_s}.$$

For example, using the scenarios given in table 2.1 and starting from the importance weights of table 1.1, we would get:

$$\begin{aligned} w_{\text{Tmin}}(0) &= 100 \\ w_{\text{Tmin}}(1) &= 50 \\ w_{\text{Tmin}}(2) &= 50 \\ w_{\text{Tmin}}(3) &= 25 \\ w_{\text{Tmin}}(4) &= 25 \end{aligned}$$

and

$$\begin{aligned} w_{\text{Tmax}}(0) &= 50 \\ w_{\text{Tmax}}(1) &= 25 \\ w_{\text{Tmax}}(2) &= 25 \\ w_{\text{Tmax}}(3) &= 12.5 \\ w_{\text{Tmax}}(4) &= 12.5. \end{aligned}$$

2.6 Second-order cone programming

In a second-order cone program (SOCP), a linear function is minimized over the intersection of an affine set and the product of second-order (quadratic) cones. SOCPs are nonlinear convex problems that include linear and (convex) quadratic programs as special cases, but are less general than semidefinite programs (SDPs). Several efficient primal-dual interior-point methods for SOCP have been developed in the last few years.

2.6.1 Standard SOCP

The standard form of a SOCP problem is

$$\begin{aligned} \text{(SOCP) minimize} \quad & f^T x & (2.35)\\ \text{subject to} \quad & ||A_i x + b_i|| \leq c_i^T x + d_i \qquad (i = 1, \ldots, m), & (2.36) \end{aligned}$$

where $x \in \mathbb{R}^n$, $f \in \mathbb{R}^n$, $A_i \in \mathbb{R}^{k_i \times n}$, $b_i \in \mathbb{R}^{k_i}$ and $d_i \in \mathbb{R}$. Here, as usual, $||\cdot||$ indicates the euclidean norm: $||u|| = (u^T u)^{1/2}$.

Constraints of the form $||Ax + b|| \leq c^T x + d$ are called second-order cone constraints because the affinely defined variables $u = Ax + b \in \mathbb{R}^k$ and $t = c^T x + d \in \mathbb{R}$ are constrained to belong to $\mathcal{C}_{k+1}$, the second-order cone (which is also called ice-cream or Lorentz cone) of dimension $k + 1$ defined by:

$$\mathcal{C}_{k+1} = \{(u, t) | u \in \mathbb{R}^k, t \in \mathbb{R}, ||u|| \leq t\}. \tag{2.37}$$

Moreover we define

$$\mathcal{C}_1 = \{t \in \mathbb{R} | 0 \leq t\}.$$

Figure 2.8 illustrates the surface of $\mathcal{C}_3$ ($k = 2$). Note that second-order cone constraints can be used to represent several convex constraints. For example, if $k_i = 0$ for all i's, then the SOCP reduces to the linear program:

$$\begin{aligned} \text{(LP) minimize} \quad & f^T x \\ \text{subject to} \quad & 0 \leq c_i^T x + d_i \qquad (i = 1, \ldots, m). \end{aligned}$$

Therefore the constraints set in our formulation (SOC constraints and linear constraints, see again section 2.5) are compliant with the SOCP framework.

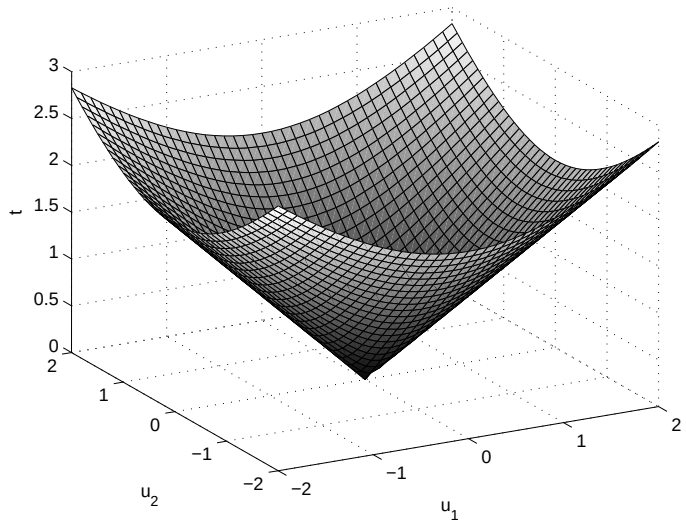


Figure 2.8: The surface of the second-order cone $\mathcal{C}_3$ (see equation (2.37)) for $k = 2$.

2.6.2 Interior point method

Methods for finding optimum points for a LP problem usually start and remain on the boundary of the feasible region: the well-known Simplex method belongs to this class of algorithms.

SOCP problems can be efficiently solved, on the other hand, via specialized Interior Point (IP) methods[6]. An IP method is a linear or nonlinear programming method that achieves optimization by going through the middle of the solid defined by the problem rather than around its surface.

One of the most interesting types of IP methods is the path-following algorithm, which combines excellent behavior in theory and practice. A member of the this category known as the primal-dual path-following algorithm has become the method of choice in large-scale implementations.

[6]Enhancement of the Simplex method to solve SOCP problems is an area of active research, but commercial quality implementations do not exist yet.

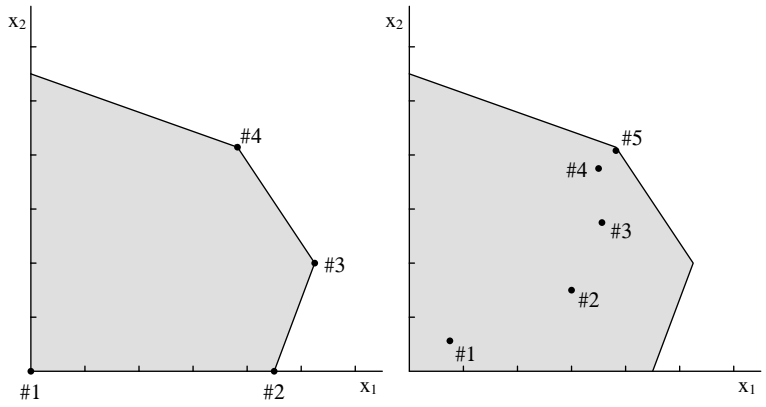


Figure 2.9: The simplex algorithm (left) begins at a starting vertex and moves along the edges of the polytope until it reaches the vertex of the optimum solution. An IP method (right) goes through the middle of the polytope.

2.6.3 Conic duality

Assume $m = 1$ for the ease of presentation; the constraint of the SOCP problem given in (2.35) can be rewritten as

$$||Ax + b|| \leq c^T x + d \Leftrightarrow \begin{pmatrix} A \\ c^T \end{pmatrix} x + \begin{pmatrix} b \\ d \end{pmatrix} \in \mathcal{C}_{k+1}.$$

Therefore, an SOCP problem can be formulated as

$$\begin{aligned} \text{(CP) minimize} \quad & f^T x && (2.38) \\ \text{subject to} \quad & \hat{A}x - \hat{b} \in \mathcal{C}_{k+1} && (2.39) \end{aligned}$$

where

$$\hat{A} = \begin{pmatrix} A \\ c^T \end{pmatrix}$$

and

$$\hat{b} = -\begin{pmatrix} b \\ d \end{pmatrix}.$$

This is referred to as the primal formulation. The dual formulation is:

$$\begin{aligned} \text{(CD) maximize} \quad & b^T y && (2.40) \\ \text{subject to} \quad & \hat{A}^T y = f && (2.41) \\ & y \in \mathcal{C}_* && (2.42) \end{aligned}$$

where

$$\mathcal{C}_* = \{z \in \mathbb{R}^{k+1} | z^T x \geq 0 \ \ \forall x \in \mathcal{C}\}.$$

Theorem 1 (Conic Duality) *Consider the conic problem (CP) and its dual (CD). Then:*

- *the dual to (CD) is equivalent to (CP);*
- *for any x feasible for (CP) and y feasible for (CD), we have $f^T x \geq b^T y$;*
- *if either (CP) or (CD) is bounded and strictly feasible, then any primal-dual pair (x, y) is an optimal solution if and only if $f^T x = b^T y$.*

Proof 1 *See, for example, [3] and [4].*

These results can be extended to the case in which the feasible region $\mathcal{C}_{k+1}$ of (CP) is replaced by K, the intersection of an affine set and the direct product of quadratic cones. Because of the previous theorem, if x is feasible for (CP) and y is feasible for (CD), then the so-called duality gap $f^T x - b^T y$ is always positive and zero in the optimal solution. The primal-dual path following algorithm operates simultaneously on the primal and dual problems, searching for a path for which the duality gap decreases in each step. Points of this path are kept inside the feasible region using so-called barrier functions. A barrier function $F(x) : \text{int}(K) \to \mathbb{R}$ is a function such that

$$F(x) \to +\infty \text{ as } x \to \partial K.$$

Such barrier functions are smartly added to the objective function of (CD), in such a way that they introduce a growing penalty as the path approaches the boundary of the feasible region. More details on how these algorithms work can be found, for example, in [5] and [6].

We have used a Matlab-based tool called SeDuMi, which is an efficient implementation of a primal-dual interior point method for solving SOCP problems.

Chapter 3

Computational results

For the computational tests, we have used a modified version of the software developed in [2], including, in the optimization process, the package SeDuMi. We have used the two-dimensional phantom presented in figure 3.1. It consists of 1,257 voxels. The number of voxels of each organ is reported in table 3.1.

Note the organ at risk "OAR1" positioned in proximity of the CTV: this is the organ that will most likely "suffer" the most, due to the high dosage prescribed for the nearby tumor.

Three beams have been used, positioned at 0, 110 and 250 degrees (see again figure 3.1). Each beam consists of 10 beamlets, so that $M = 30$.

The scenarios of body dislocation considered are the ones presented in table 3.2. The parameters used in the optimization are given in table 3.3.

The other parameters used are:

- $N = 45$;
- $\delta = 0.95$ (so that $z(\delta) \approx 1.64$);
- $\alpha = 0.8$;
- $\epsilon = 0.1$.

Treatment plannings are usually evaluated by looking at Dose-Volume Histograms (DVH): in this type of histograms, the vertical axis represents the percent of total tissue volume that receives a dose greater than or equal to a specified dose.

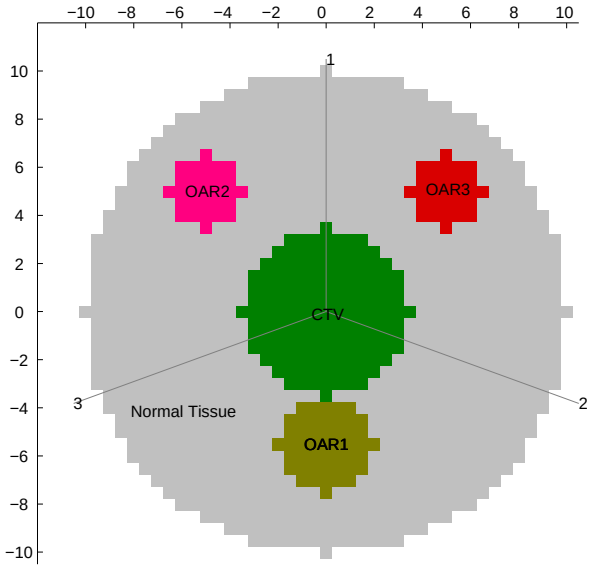


Figure 3.1: The phantom and the beam directions used in the test.

Structure	Number of voxels
CTV	148
OAR1	49
OAR2	29
OAR3	29
Normal tissue	1002

Table 3.1: Number of voxels for each structure.

j_r or j_s	Direction	Shift (in voxels)	Probability
0	-	0	0.5
1	Left	1	0.125
2	Right	1	0.125
3	Anterior	1	0.125
4	Posterior	1	0.125

Table 3.2: Scenarios (and their probabilities) of body dislocation used.

Structure	**Prescription**	**Weight**
CTV	$m_{\mathrm{Tmin}} = 70$	$\begin{cases} w_{\mathrm{Tmin}}(0) = 100 \\ w_{\mathrm{Tmin}}(1) = 25 \\ w_{\mathrm{Tmin}}(2) = 25 \\ w_{\mathrm{Tmin}}(3) = 25 \\ w_{\mathrm{Tmin}}(4) = 25 \end{cases}$
CTV	$m_{\mathrm{Tmax}} = 75$	$\begin{cases} w_{\mathrm{Tmax}}(0) = 1 \\ w_{\mathrm{Tmax}}(1) = 0.25 \\ w_{\mathrm{Tmax}}(2) = 0.25 \\ w_{\mathrm{Tmax}}(3) = 0.25 \\ w_{\mathrm{Tmax}}(4) = 0.25 \end{cases}$
OAR1	$m_1 = 60$	$w_1 = 80$
OAR1	$v_1 = 0.4;\ d_1 = 40$	$\overline{w}_1 = 10$
OAR2	$m_2 = 60$	$w_2 = 80$
OAR2	$v_2 = 0.4;\ d_2 = 40$	$\overline{w}_2 = 10$
OAR3	$m_3 = 60$	$w_3 = 80$
OAR3	$v_3 = 0.4;\ d_3 = 40$	$\overline{w}_3 = 10$
Normal tissue	$m_4 = 60$	$w_4 = 80$

Table 3.3: Parameters used for the tests.

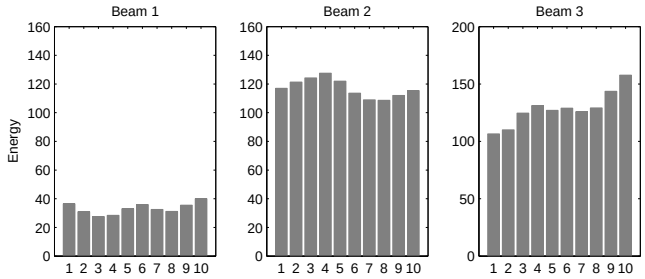


Figure 3.2: Optimal fluence profile (x^*) for each beam.

The optimal fluence profile x^* has been calculated solving the optimization problem described in section 2.5. It took 208 seconds (and 42 iterations) for the problem to be solved. The plot of x^* is given in figure 3.2. Note how beam 1 has a considerably lower energy with respect to beam 2 or 3: this is because beam 1 points directly to OAR1, and high energies would result in a high dose delivered to that organ-at-risk.

The actual dose distribution (for scenario $j_s = 0$ and $j_r = 0$) is reported in figure 3.4.

Once that the solution has been obtained, 30 treatments (each made of 45 fractions) have been sampled according to the random/systematic scenario distributions of table 3.2. The DVH of these treatments are given in figure 3.3. Note how the histograms relative to the CTV are close to each other (in other words, they are less sensitive to the uncertainty): this can be explained in part by the additional and stronger constraints imposed to this organ. Note how OAR1 (as foreseen) is, on the other hand, the most sensitive organ, as it suffers, in some scenarios, from the high dose delivered to the CTV.

As a comparison, we also present the DVH and the dose distribution obtained by solving the same problem, but using the classical approach of PTV definition (see again section 1.4). In this case, the PTV has been defined so that it covers the CTV in every possible random/systematic scenario. This means, for our particular case (check again table 3.2), enlarging the CTV by two voxels in every direction. The formulation we used in this case, is the one given in the Appendix ("Static case").

Note how, especially for OAR3, the proposed SOCP approach outperforms the classical approach, as this organ receives, in general, a lower dose of radiation. Note moreover how a big percentage of the CTV, in the classical-approach solution, receives a dose that exceeds the maximum prescribed dose

(75 grays).

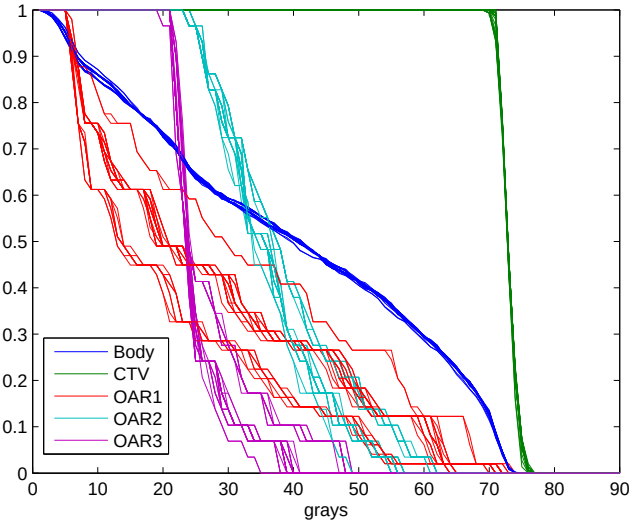


Figure 3.3: SOCP approach: DVH resulting from the optimal solution.

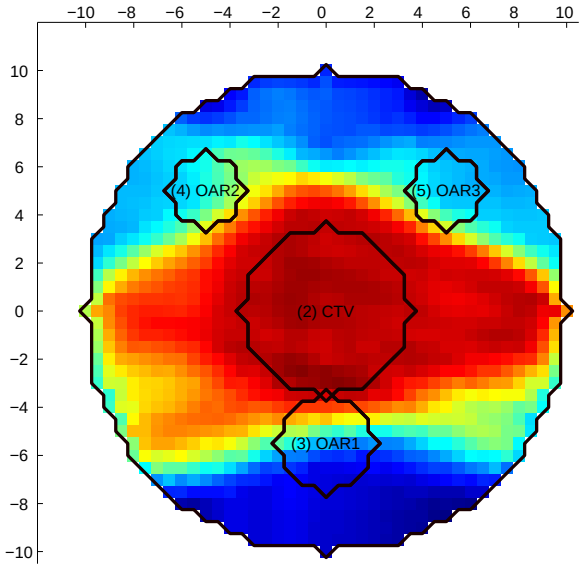


Figure 3.4: SOCP approach: dose distribution resulting from the optimal solution.

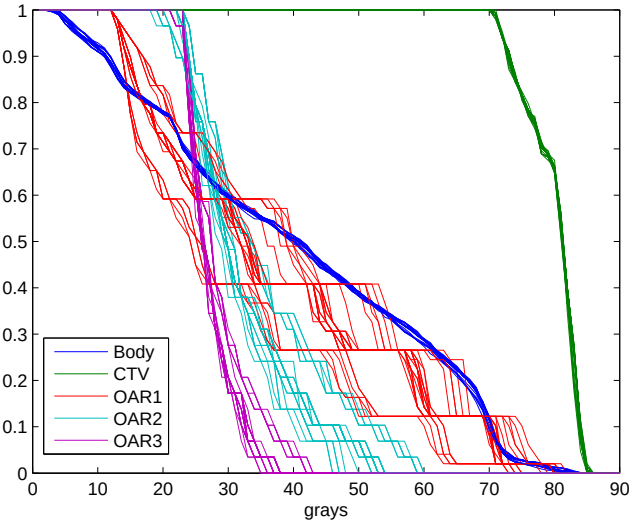


Figure 3.5: Classical approach: DVH resulting from the optimal solution.

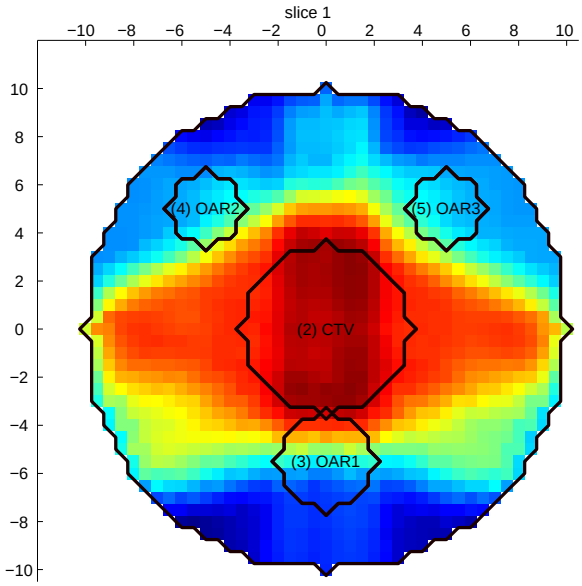


Figure 3.6: Classical approach: dose distribution resulting from the optimal solution.

Chapter 4

Conclusions

We have described, implemented and tested a method that uses second order cone programming for the resolution of the problem of radiation therapy treatment planning. This SOCP approach is a promising alternative to the classical method, especially when dealing with uncertainties in the patient position due to both random and systematic error.

Further developments offer the prospect of up to 10% improvement in cure rates for patients having radiotherapy. Possible extensions of the proposed approach include the optimizaztion of the number and location of the beams and, ideally, the implementation of a method that makes use of continuous distributions of voxel dislocations.

Appendix

We present three adaptations of the problem formulation given in section 2.5 in case of presence of random error only, in case of presence of systematic error only, and in case of presence of no error.

Random-error-only case

If random error is the only type of error affecting the position of the patient, then the formulation of the problem is simply the one given in section 2.5, where we replace $\forall j_s$ in (2.20), (2.21), (2.22), (2.23), (2.24), (2.25) and (2.29) with $j_s = 0$. Accordingly, in the objective function (2.19), we replace

$$\sum_{j_s=0}^{n-1} w_{\mathrm{Tmin}}(j_s) r_{\mathrm{Tmin}}(j_s)$$

and

$$\sum_{j_s=0}^{n-1} w_{\mathrm{Tmax}}(j_s) r_{\mathrm{Tmax}}(j_s)$$

with, respectively

$$w_{\mathrm{Tmin}}(j_0) r_{\mathrm{Tmin}}(j_0)$$

and

$$w_{\mathrm{Tmax}}(j_0) r_{\mathrm{Tmax}}(j_0).$$

The resulting formulation is very similar to the one proposed in [1].

Systematic-error-only case

Remember that if systematic error is the only type of error present, then, for one particular patient treatment, the doses delivered during each fraction

are equal to each other (see again figure 2.2). Therefore we will:

- remove the conic constraints, as they were originally introduced to model fraction-to-fraction stochasticity due to random error;
- utilize, for each organ, constraints of type (2.1).

This latter constraints will be, in this case, introduced with intermediate variables that allow us to weight them in the objective function (similarly to what we did in 2.5).

Summing up, the formulation becomes:

$$\sum_{j_s=0}^{n-1} w_{\mathrm{Tmin}}(j_s) r_{\mathrm{Tmin}}(j_s) + \sum_{j_s=0}^{n-1} w_{\mathrm{Tmax}}(j_s) r_{\mathrm{Tmax}}(j_s) + \sum_{j_s=0}^{n-1} \sum_{k=1}^{h} w_k(j_s) r_k(j_s) + \sum_{j_s=0}^{n-1} \sum_{k=1}^{h} \overline{w}_k(j_s) q_k(j_s)$$

subject to the following constraints:

$$\begin{aligned}
Na_{j_s,0}^{(i)} x &\geq u_{\mathrm{Tmin}}(j_s) && (\forall j_s; \forall i \in \mathrm{CTV}) \\
m_{\mathrm{Tmin}} - u_{\mathrm{Tmin}}(j_s) &\leq r_{\mathrm{Tmin}}(j_s) && (\forall j_s) \\
r_{\mathrm{Tmin}}(j_s) &\geq 0 && (\forall j_s) \\
Na_{j_s,0}^{(i)} x &\leq u_{\mathrm{Tmax}}(j_s) && (\forall j_s; \forall i \in \mathrm{CTV}) \\
u_{\mathrm{Tmax}}(j_s) - m_{\mathrm{Tmax}} &\leq r_{\mathrm{Tmax}}(j_s) && (\forall j_s) \\
r_{\mathrm{Tmax}}(j_s) &\geq 0 && (\forall j_s) \\
Na_{j_s,0}^{(i)} x &\leq u_k(j_s) && (\forall k; \forall j_s; \forall i \in \mathrm{OAR}_k) \\
u_k(j_s) - m_k &\leq r_k(j_s) && (\forall k; \forall j_s) \\
r_k(j_s) &\geq 0 && (\forall k; \forall j_s) \\
\frac{1}{|\mathrm{OAR}_k|} \sum_{i \in \mathrm{OAR}_k} \max(Na_{j_s,0}^{(i)} x - d_k, 0) &\leq f_k(j_s) && (\forall k; \forall j_s; \forall i \in \mathrm{OAR}_k) \\
f_k(j_s) - v_k(m_k - d_k) &\leq q_k(j_s) && (\forall k; \forall j_s) \\
q_k(j_s) &\geq 0 && (\forall k; \forall j_s) \\
-\epsilon \cdot e \leq Tx &\leq \epsilon \cdot e \\
x &\geq 0
\end{aligned}$$

Static case

Finally, if we consider the case in which the patient is supposed to always be perfectly positioned (no random nor systematic error involved), we simply would solve the problem described in the systematic-error-only case, replacing each "$\forall j_s$" with "$j_s = 0$". Note that in these two latter cases, the problem to be solved is a standard Linear Programming (LP) problem, a problem in which the objective and all of the constraints are linear functions of the decision variables.

Bibliography

[1] M. Chu, Y. Zinchenko, S. Henderson and M. Sharpe, "Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty" (Phys. Med. Biol. 50 5463-5477), 2005.

[2] Sebastiaan Breedveld, "Inverse planning with uncertainty in organ body position", 2005.

[3] Fernando Ordonez, "Robust optimization - Examples of Conic Programming", 2004.

[4] A. Ben-Tal and A. Nemirovski, "Convex optimization in engineering: Modeling, analysis, algorithms" (Technion, Israel Institute of Technology), 1999.

[5] Paul A. Jensen and Jonathan F. Bard, "Operations Research Models and Methods" (John Wiley and Sons), 2000.

[6] E. D. Andersen, C. Roos, and T. Terlaky, "On implementing a primal-dual interior-point method for conic quadratic optimization" (Mathematical Programming Ser. B, 95, pp. 249-277), 2000.

[7] Jos F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones" (Optimization Methods and Software, 11-12, pp. 625-653), 1999.