International Symposium on Operational Research, ISOR'08
Algiers, Algeria : November 2-6, 2008

Discrete Optimization Problems for Radiation Therapy Planning

Konrad ENGEL, Thomas KALINOWSKI and Antje KIESEL

University of Rostock, 18051 Rostock, Germany

konrad.engel@uni-rostock.de

Abstract: A well-studied problem in intensity modulated radiation therapy (IMRT) is the representation of a given intensity matrix, i.e. a matrix of nonnegative integers, as a nonnegative linear combination of special 0-1-matrices, called segments. These segments can be practically realized by multileaf collimators (MLC). One important aim is the minimization of the sum of the coefficients of the linear combination, i.e. the delivery time. This paper gives an introduction into this subject and surveys recent results for related problems that arise in practical applications. In particular, segmentation problems are discussed where not exactly the given matrix but a matrix with small deviations has to be segmented. Moreover, restrictions on the segments like the rectangular constraint, the interleaf collision constraint and the tongue-and-groove constraint are considered.

Keywords: IMRT planning, multileaf collimator, intensity matrix, approximated segmentation

Résumé : Un problème bien étudié dans la thérapie par intensité d'un rayonnement modulé (IMRT) est la représentation d'une matrice d'intensité donnée, c.à.d. une matrice d'entiers non négatifs, comme une combinaison linéaire non négative de (0,1)-matrices particulières, appelées segments. Ces segments peuvent être pratiquement réalisés par des collimateurs multifeuilles (MLC). Un objectif important est la minimisation de la somme des coefficients de la combinaison linéaire, c.à.d. le durée de livraison. Ce document donne une introduction à ce sujet et passe en revue les résultats récents pour des problèmes similaires qui se posent dans des applications pratiques. En particulier, des problèmes de segmentation sont discutés lorsque ce n'est pas exactement cette matrice donnée, mais une matrice avec de petits écarts qui doit être segmentée. De plus, les restrictions sur les segments comme la contrainte rectangulaire, la contrainte de collision inter-feuille et la contrainte rainure et languette sont considérées.

Mots clés : plan IMRT, collimateur multifeuille, matrice d'intensité, segmentation approximée

1 Introduction

Intensity modulated radiation therapy (IMRT) has become an important method in cancer treatment. High energetic radiation is used to destroy the tumor (target volume), but also affects healthy cells in the surrounding organs. Therefore, the treatment must be planned in such a way that the cancerous cells receive a clinically prescribed dose while the surrounding organs (organs at risk) are protected from the radiation. The design of a treatment plan realizing this aim consists of two main steps. In the first step, a set of beam directions has to be chosen and the amount of intensity modulated radiation for each direction has to be determined. For each beam direction, this amount is given by the entries of an intensity matrix with nonnegative integral entries. Commonly, a multileaf collimator (MLC) is used for the realization of the fluence given by this intensity matrix. An MLC consists of a number of metal leaf pairs, each corresponding to one row of the intensity matrix, that can be shifted towards each other. The region between the left and the right leaves receives radiation while the area that is covered by the leaves is protected. Each choice of the leaf positions generates a homogeneous field. The second step, called segmentation, consists of the determination of a number of fields whose superposition realizes the given intensity matrix. We focus on this segmentation procedure and on the minimization of the delivery time.

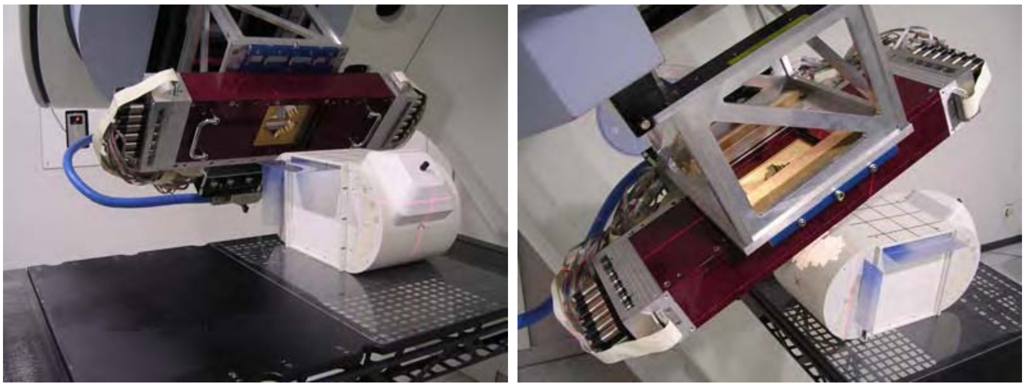


Figure 1: A multileaf collimator

2 The determination of the intensity matrices

Because the intensity matrices are the main objects for this paper we briefly sketch how they are determined. We consider the (abstract) *patient* as a finite set V of voxels which are geometrically small cuboids in the body. The kind of region yields a partition of the patient V into blocks. Usually, we have one block T which is called the *planning target volume* (PTV), a family $\mathcal{R}$ of blocks R which are the *organs at risk* (OAR), and one block S which consists of the remaining voxels of V, i.e. the rest of body (ROB),

$$V = \left(\bigcup_{R \in \mathcal{R}} R \right) \cup S \cup T.$$

In practice, these regions are given by contours drawn by the physician or physicist on each CT–slice. In Figure 2 there is given one CT–slice containing contours of the PTV and the OARs. This slice is illustrated on the left by means of the Hounsfield–values and on the right by means of the Electron density values for dose calculation. The voxels (on the right) are represented here by two-dimensional squares though they are in fact three-dimensional cuboids. The physician fixes a value b_T, i.e. the prescribed dose to

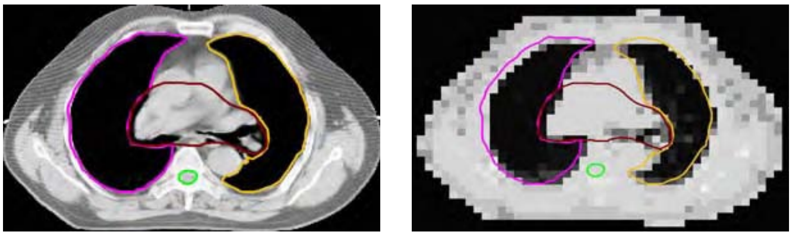


Figure 2: CT–slices representing Hounsfield values (left) and Electron density values of voxels (right)

the target, and the aim is to find a treatment plan such that the accumulated dose after the treatment is very near to b_T for each $v \in T$ and that it is as small as possible for each other voxel, but in particular for the voxels in the OARs. The deviation from this aim is modeled by a penalty function and various optimization algorithms lead to a set of beam directions, cf. [7]. In conformal radiation therapy, the radiation is delivered for each beam direction in a homogeneous way, i.e. one has a fixed open region O of the collimator and the radiation is delivered through O for a certain time. The region O is usually determined in such a way that the projection of the PTV onto the collimator plane is encircled by O. One can obtain better results if the fluence of radiation differs from point to point in O. This *intensity modulation* is realized after a discretization in the following way: One starts with a rectangle surrounding O and partitions it into small squares, called *bixels*. For each bixel b, one introduces a variable x_b which is the still unknown fluence of radiation through the bixel. Then one determines the dose values of the voxels under the supposition that only the bixel b is open and the rest of the rectangle is closed by the leaves. The penalty function and methods from nonlinear optimization provide a fixed value x_b and, after some rounding, a corresponding value a_b for each bixel b. If b is positioned in row i and column j we use the notation a_{ij} instead of a_b. The matrix $A = (a_{ij})$ is then the optimized intensity matrix.

3 Mathematical formulation of the segmentation problem

For any $k, l, r \in \mathbb{N}$ let $[k] = \{1, \ldots, k\}$ and $[l, r] = \{i \in \mathbb{N} : l \leq i \leq r\}$. We call $[l, r]$ an *interval*. Let A be a matrix of dimension $m \times n$ with nonnegative integral entries. As above, it represents the given intensity matrix and has to be decomposed into a number of segments that correspond to the leaf positions of the MLC. A matrix $S = (s_{ij})$ is called a *segment* if there are intervals $I_1, \ldots, I_m$ in $[n]$ such that

$$s_{ij} = \begin{cases} 1, & \text{if } j \in I_i = [l_i, r_i] \\ 0, & \text{otherwise} \end{cases} \qquad i \in [m],\ j \in [n]. \tag{1}$$

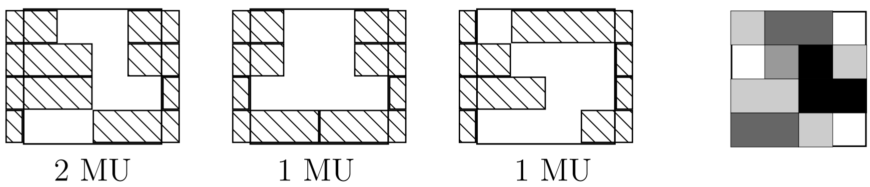
The intervals I_i represent the region between the left and right leaves in row i of the MLC, i.e. $l_i - 1$ and $r_i + 1$ are the positions of the i-th left and right leaf, respectively. A *segmentation* of A is a decomposition

$$A = \sum_{i=1}^{k} u_i S_i$$

with segments S_i and positive integers u_i $(i = 1, \ldots, k)$. The *delivery time (DT)* of the segmentation is

$$DT = \sum_{i=1}^{k} u_i.$$

The coefficients u_i represent the time of radiation for the fixed segment, i.e. the fixed leaf positions. They are usually given in a certain scale where the units are called *monitor units* (MU). Hence we require everywhere that they are integers. It turns out that this additional integrality constraint often is not really a restriction. In Figure 3 there is given a segmentation with a DT of 4 monitor units.



$$2 \cdot \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 3 & 0 \\ 0 & 2 & 4 & 1 \\ 1 & 1 & 4 & 4 \\ 3 & 3 & 1 & 0 \end{pmatrix}$$

$$DT = 2 + 1 + 1 = 4$$

Figure 3: A segmentation with 4 MUs

A standard problem is the following:

Problem 1 (MIN-DT) *Given a matrix A of nonnegative integers, find a segmentation such that its DT is minimal. Let $c(A)$ be this minimal DT.*

From the practical point of view it is not necessary that exactly the intensity matrix A is realized by the segmentation. The reason is that the determination of A in the first step depends on several procedures that are not completely exact, e.g. the determination of the regions of interest (PTV, OAR), the dose calculation, the objective function, and the optimization algorithm. Hence corresponding approximation problems are of interest: Let, as usual,

$$\|\boldsymbol{A}\|_1 = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}| \quad \text{and} \quad \|\boldsymbol{A}\|_\infty = \max\{|a_{ij}| : i \in [m], j \in [n]\}.$$

Problem 2 (MIN-DT-APP) *Given a matrix A of nonnegative integers and a nonnegative integer δ, find a matrix B of nonnegative integers such that $\|B-A\|_\infty \leq \delta$ and $c(B)$ is minimal. Let $c_\delta(A)$ be this minimal DT.*

Given two matrices A, B, we call $\|B - A\|_1$ the *total change (TC).*

Problem 3 (MIN-DT-TC-APP) *Given a matrix A of nonnegative integers and a nonnegative integer δ, find a matrix B of nonnegative integers with the property that $\|B - A\|_\infty \leq \delta$, $c(B) = c_\delta(A)$ and $\|B - A\|_1$ is minimal.*

Further related problems are discussed in the last two sections.

4 Solution of the problem MIN-DT

Given a matrix A with nonnegative integer entries, let $\boldsymbol{a}_i$ denote the i-th row of A for $i \in [m]$. Ignoring machine-dependent constraints, we can solve the problem MIN-DT independently for each row of A, i.e. m problems MIN DT-ROW. The problem MIN-DT-ROW is defined exactly as the problem MIN-DT with the only exception that the matrix A is replaced by a vector $\boldsymbol{a}$ (with $a_0 = a_{n+1} = 0$) and the segments S are given as vectors $\boldsymbol{s}(l, r)$ where

$$s(l,r)_i = \begin{cases} 1 & \text{if } i \in [l,r], \\ 0 & \text{otherwise.} \end{cases}$$

Let the *complexity* of a vector $\boldsymbol{a}$ be defined by

$$c(\boldsymbol{a}) = \sum_{j=1}^{n} \max\{0, a_j - a_{j-1}\}.$$

The solution of the problem MIN-DT-ROW is given by the following theorem, cf. [4] and the surveys [9], [11] containing several other references. The first algorithm that yields an optimal solution was presented by Bortfeld, Kahler, Waldron, and Boyer [3].

Theorem 1 *The minimal DT of a segmentation of a vector $\boldsymbol{a}$ is given by its complexity $c(\boldsymbol{a})$.*

Proof. We present here a very elementary proof: First of all note that we may assume that all coefficients u are equal to 1 since an item $u\boldsymbol{s}$ can be written as a sum $\boldsymbol{s}+\cdots+\boldsymbol{s}$ with u items. Let $f(\boldsymbol{a})$ be the minimal DT for $\boldsymbol{a}$. So we have to prove that $f(\boldsymbol{a}) = c(\boldsymbol{a})$. We use induction on $\|\boldsymbol{a}\|_1$. The case $\|\boldsymbol{a}\|_1 = 0$, i.e. $\boldsymbol{a} = \boldsymbol{0}$, is trivial. First we show that $f(\boldsymbol{a}) \geq c(\boldsymbol{a})$. Let $\boldsymbol{s}(l,r)$ be an item of an optimal segmentation of $\boldsymbol{a}$ and let $\boldsymbol{a}' = \boldsymbol{a} - \boldsymbol{s}(l,r)$. Note that $\|\boldsymbol{a}'\|_1 < \|\boldsymbol{a}\|_1$. Since, for $j \in [n]$,

$$s(l,r)_j - s(l,r)_{j-1} = \begin{cases} 1 & \text{if } j = l, \\ -1 & \text{if } j = r+1, \\ 0 & \text{otherwise} \end{cases}$$

we have

$$\begin{aligned} c(\boldsymbol{a}) - c(\boldsymbol{a}') = \max\{0, a_l - a_{l-1}\} - \max\{0, a_l - 1 - a_{l-1}\} \\ + \max\{0, a_{r+1} - a_r\} - \max\{0, a_{r+1} - (a_r - 1)\}). \end{aligned}$$

Hence,

$$c(\boldsymbol{a}) - c(\boldsymbol{a}') \begin{cases} = 1 & \text{if } a_{l-1} < a_l \text{ and } a_r > a_{r+1}, \\ < 1 & \text{otherwise.} \end{cases}$$

By induction hypothesis, $f(\boldsymbol{a}') = c(\boldsymbol{a}')$. Clearly, $f(\boldsymbol{a}) = f(\boldsymbol{a}') + 1$. Accordingly, $f(\boldsymbol{a}) = c(\boldsymbol{a}') + 1 \geq c(\boldsymbol{a})$. Now we show that $f(\boldsymbol{a}) \leq c(\boldsymbol{a})$. If $\boldsymbol{a} \neq \boldsymbol{0}$ we can find a first index l such that $a_{l-1} < a_l$ and then a first index $r \geq l$ such that $a_r > a_{r+1}$. By the choice of l and r, $a_j \geq 1$ for all $j \in [l,r]$. Hence $\boldsymbol{a}' = \boldsymbol{a} - \boldsymbol{s}(l,r)$ is a vector of nonnegative integers and $c(\boldsymbol{a}) - c(\boldsymbol{a}') = 1$. In view of the induction hypothesis, $\boldsymbol{a}'$ can be segmented with $c(\boldsymbol{a}')$ segments and together with the segment $\boldsymbol{s}(l,r)$ this gives a segmentation of $\boldsymbol{a}$ with $c(\boldsymbol{a}') + 1 = c(\boldsymbol{a})$ segments. Consequently, $f(\boldsymbol{a}) \leq c(\boldsymbol{a})$. □

Corollary 1 *The minimal DT $c(A)$ of a segmentation of a matrix A is given by $c(A) = \max\{c(\boldsymbol{a}_i) : i \in [m]\})$, i.e. by the maximal complexity of its rows, and it can be determined in time $O(mn)$.*

5 Solution of the approximation problems for one row

Here we only present the results, complete proofs can be found in [6]. Using the results from the previous section we come to the following problems:

Problem 4 (MIN-DT-APP-ROW) *Given a vector $\boldsymbol{a} = (a_1, \ldots, a_n)$ with nonnegative integral entries, find a vector $\boldsymbol{b}$ with nonnegative integral entries such that $\|\boldsymbol{b} - \boldsymbol{a}\|_\infty \leq \delta$ and $c(\boldsymbol{b})$ is minimal. Let $c_\delta(\boldsymbol{a})$ denote this minimum.*

Problem 5 (MIN-DT-TC-APP-ROW) *Given a vector $\boldsymbol{a} = (a_1, \ldots, a_n)$ with nonnegative integral entries, find a vector $\boldsymbol{b}$ with nonnegative integral entries such that $\|\boldsymbol{b} - \boldsymbol{a}\|_\infty \leq \delta$, $c(\boldsymbol{b}) = c_\delta(\boldsymbol{a})$ and $\|\boldsymbol{b} - \boldsymbol{a}\|_1$ is minimum.*

In the following all intervals are subsets of $[0, n+1]$. Let $\epsilon > 0$. For any $j \in [0, n+1]$ we define its *lower ϵ-level interval* $I_\epsilon(j)$ and its *upper ϵ-level interval* $I^\epsilon(j)$ to be the maximal interval containing j such that

$$\begin{aligned} a_i &\le a_j + \epsilon \quad \forall i \in I_\epsilon(j), \\ a_i &\ge a_j - \epsilon \quad \forall i \in I^\epsilon(j), \end{aligned}$$

respectively. Now let $\boldsymbol{a} = (a_1, \dots, a_n)$ be a vector of nonnegative integral entries and let always $a_0 = a_{n+1} = 0$. An element $j \in [0, n+1]$ is called an *ϵ-local minimum for $\boldsymbol{a}$* if

$$a_j \le a_i \quad \forall i \in I_\epsilon(j),$$

see Figure 4, and it is called an *ϵ-local maximum for $\boldsymbol{a}$* if

$$a_j \ge a_i \quad \forall i \in I^\epsilon(j).$$

We say that j is an *ϵ-local extremum* if it is an ϵ-local minimum or an ϵ-local maximum.

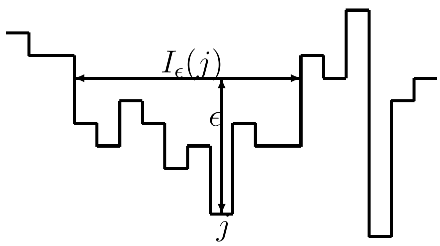


Figure 4: An ϵ-local minimum

Two ϵ-local extrema i and j where $0 \le i < j \le n+1$ are called *consecutive* if there is no $k \in [i+1, j-1]$ that is an ϵ-local extremum.

In the following we set $\epsilon = 2\delta$. For each maximal sequence of consecutive 2δ-local minima and for each maximal sequence of consecutive 2δ-local maxima we pick the first and the last one. In such a way we obtain an alternating sequence

$$\boldsymbol{s} = (0 = m_1, m^1, M_1, M^1, m_2, m^2, M_2, M^2, \dots, M_t, M^t, m_{t+1}, m^{t+1} = n+1)$$

of pairs of 2δ-local maxima and minima. We call $\boldsymbol{s}$ the *2δ-min-max sequence.* We note that $m_l = m^l$ and $M_l = M^l$ is allowed - this is the case if the corresponding sequence of consecutive 2δ-local extrema contains only a single element. Moreover, as an exception, we set $m^1 = m_1 = 0$ and $m_{t+1} = m^{t+1} = 0$.

We say that a vector $\boldsymbol{b}$ is *conform* to the 2δ-min-max sequence if

$$\begin{aligned} b_j &= a_{m_l} + \delta \text{ for all } j \in [m_l, m^l], \quad l = 2, \dots, t, \\ b_j &= a_{M_l} - \delta \text{ for all } j \in [M_l, M^l], \quad l = 1, \dots, t, \\ b_j &\ge b_{j-1} \text{ for all } j \in [m^l + 1, M_l], \quad l = 1, \dots, t, \\ b_j &\ge b_{j+1} \text{ for all } j \in [M^l, m_{l+1} - 1], \quad l = 1, \dots, t. \end{aligned}$$

Hence, if $\boldsymbol{b}$ is conform to the 2δ-min-max sequence, then $\boldsymbol{b}$ is constant in the intervals $[m_l, m^l], [M_l, M^l]$, increasing in the intervals $[m^l, M_l]$, and decreasing in the intervals $[M^l, m_{l+1}]$, see Figure 5.

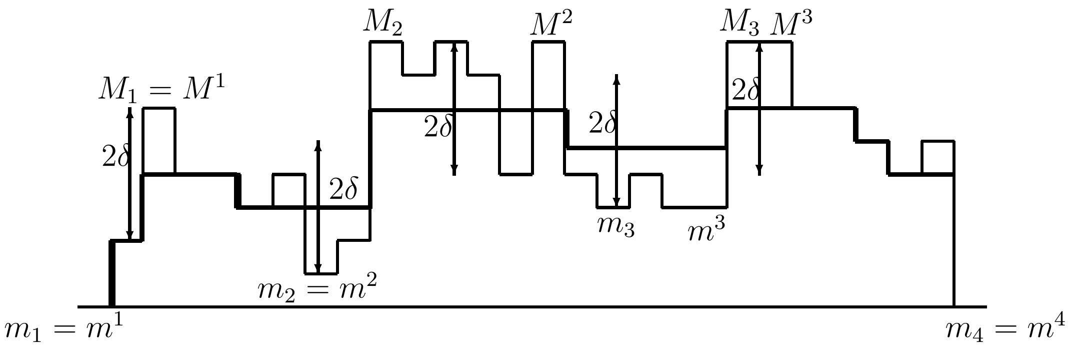


Figure 5: A vector that is conform to the 2δ-min-max sequence

We construct two vectors $\underline{\boldsymbol{b}}$ and $\overline{\boldsymbol{b}}$ which turn out to be conform to the 2δ-min-max sequence as follows: We set

$$\begin{aligned}
\underline{b}_0 = \overline{b}_0 &= 0,\\
\underline{b}_{n+1} = \overline{b}_{n+1} &= 0,\\
\underline{b}_j = \overline{b}_j &= a_{m_l} + \delta, \quad j \in [m_l, m^l], \quad l = 2, \dots, t,\\
\underline{b}_j = \overline{b}_j &= a_{M_l} - \delta, \quad j \in [M_l, M^l], \quad l = 1, \dots, t.
\end{aligned}$$

Further, we set

$$\begin{aligned}
\underline{b}_j &= \max\{\underline{b}_{j-1}, a_j - \delta\}, \quad j = m^l + 1, \dots, M_l - 1, \quad l = 1, \dots, t,\\
\underline{b}_j &= \max\{\underline{b}_{j+1}, a_j - \delta\}, \quad j = m_{l+1} - 1, \dots, M^l + 1, \quad l = 1, \dots, t.
\end{aligned}$$

and

$$\begin{aligned}
\overline{b}_j &= \min\{\overline{b}_{j+1}, a_j + \delta\}, \quad j = M_l - 1, \dots, m^l + 1, \quad l = 1, \dots, t,\\
\overline{b}_j &= \min\{\overline{b}_{j-1}, a_j + \delta\}, \quad j = M^l + 1, \dots, m_{l+1} - 1, \quad l = 1, \dots, t.
\end{aligned}$$

Theorem 2 *Let $\|\boldsymbol{a}\|_\infty > 2\delta$. The vectors $\underline{\boldsymbol{b}}$ and $\overline{\boldsymbol{b}}$ are optimal solutions for the problem MIN-DT-APP-ROW. A vector $\boldsymbol{b}$ is an optimal solution for the problem MIN-DT-APP-ROW iff $\boldsymbol{b}$ is conform to the 2δ-min-max sequence and $\underline{\boldsymbol{b}} \le \boldsymbol{b} \le \overline{\boldsymbol{b}}$. We have*

$$c_\delta(\boldsymbol{a}) = (a_{M_1} - \delta) + \sum_{l=2}^{t} (a_{M_l} - a_{m_l} - 2\delta).$$

Corollary 2 *The problem MIN-DT-APP-ROW can be solved in time $O(n)$.*

Corollary 3 *The problem MIN-DT-TC-APP-ROW can be solved in time $O(\delta n)$.*

6 Solution of the approximation problems for matrices

As the problem MIN-DT in Section 4, the problem MIN-DT-APP can be solved independently for each row. This immediately yields:

Theorem 3 *The problem MIN-DT-APP can be solved in time $O(mn)$.*

For the solution of MIN-DT-TC-APP it is not necessary to realize the individual minimal DT for each row. We only have to realize the bound $c_\delta(A)$ for each row. This leads us immediately to the following *Constrained DT and Min TC problem for single rows*:

Problem 6 (CON-DT-TC-ROW-APP) *Given a vector $\boldsymbol{a} = (a_1, \ldots, a_n)$ with nonnegative integral entries and a bound C, find a vector $\boldsymbol{b}$ with nonnegative integral entries such that $\|\boldsymbol{b} - \boldsymbol{a}\|_\infty \leq \delta$, $c(\boldsymbol{b}) \leq C$ and $\|\boldsymbol{b} - \boldsymbol{a}\|_1$ is minimum.*

The use of dynamic programming methods [6] leads to the following result:

Theorem 4 *The problem CON-DT-TC-ROW-APP can be solved in time $O(\delta^3 n^2)$.*

Another approach to the problem CON-DT-TC-ROW-APP [10] is based on the observation that an LP-dual of this problem is essentially a min-cost-circulation problem for the following network N. The node set is

$$V = \{q, s\} \cup \{(j, 0),\ (j, 1)\ :\ j \in [0, n+1]\},$$

and the arc set is $E = E_0 \cup E_1 \cup E_2$, where

$$\begin{aligned} E_0 &= \{(q, (0,0)),\ (q, (0,1)),\ ((n+1, 0), s),\ ((n+1, 1), s),\ (s, q)\} \\ &\quad \cup \{((j, k), (j+1, k))\ :\ j \in [0, n],\ k \in \{0, 1\}\}, \\ E_1 &= \{((j, 0), (j+1, 1))^{(\lambda)}\ :\ j \in [n],\ \lambda \in \{0, 1\}\}, \\ E_2 &= \{((j, 1), (j+1, 0))^{(\lambda)}\ :\ j \in [n],\ \lambda \in \{0, 1\}\}. \end{aligned}$$

The arc capacities and costs are collected in Table 1, where we put

$$\underline{a_j} = \max\{0, a_j - \delta\}, \quad \overline{a_j} = a_j + \delta \qquad (j \in [n]).$$

All arcs that do not appear in the table have infinite capacity and zero cost. The definition of the network is illustrated in Figure 6.

A standard primal-dual algorithm can be used to obtain a circulation φ of minimum cost together with a corresponding potential π. We obtain an optimal solution $\boldsymbol{b}$ for the problem CON-DT-TC-ROW-APP by setting, for all $j \in [n]$,

$$b_j = \begin{cases} \pi(j+1, 0) - \pi(j, 1) & \text{if } \varphi((j, 1), (j+1, 0))^{(0)} > 0, \\ \pi(j, 0) - \pi(j+1, 1) & \text{if } \varphi((j, 0), (j+1, 1))^{(0)} > 0, \\ a_j & \text{otherwise.} \end{cases}$$

Arc	Capacity	Cost
$((j,0),(j+1,1))^{(0)}$	1	$-a_j$
$((j,0),(j+1,1))^{(1)}$	∞	$-\underline{a_j}$
$((j,1),(j+1,0))^{(0)}$	1	a_j
$((j,1),(j+1,0))^{(1)}$	∞	$\overline{a_j}$
(s,q)	∞	C

Table 1: The arc capacities and costs in our network

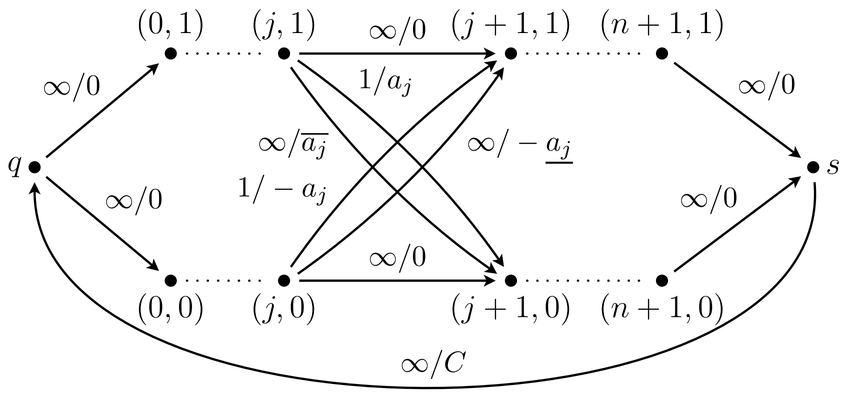


Figure 6: The network N for the problem CON-DT-TC-ROW-APP. The arc labels have the format “capacity/cost”.

By this transformation we obtain an alternative estimate for the time-complexity.

Theorem 5 *The problem CON-DT-TC-ROW-APP can be solved in time $O(n^2 \log^2 n)$.*

The independent solution for each row finally provides the following theorem:

Theorem 6 *The problem MIN-DT-TC-APP can be solved in time* $\min(O(m\delta^3 n^2), O(mn^2 \log^2 n))$.

7 Restricted shapes: Rectangles

Hitherto all segments from the definition (1) have been allowed for a segmentation. In practice, several other restrictions are necessary or desirable, e.g. one can require for dosimetric reasons that the open region should be connected and as regular as possible. The easiest variant is the case if the open region has to be a rectangle. We call this constraint the rectangle constraint (RC). The problem MIN-DT for this case is solved in [5]. Up to now we only have some specialized integer programming method that provides a solution for arbitrary matrices A. But the case of two rows can be treated in a satisfactory way:

With the two-row matrix A we associate the following *segmentation network* $N = (V, E, q, s, c)$, where

$$V = \{q, s, 0, 1, \ldots, n\}, \qquad E = E_1 \cup E_2 \cup E_3,$$

q is the source, s is the sink, and

$$\begin{aligned}
E_1 &= \{(j-1, j) : j \in [n]\},\\
E_2 &= \{(q, j-1) : a_{1,j-1} < a_{1,j} \text{ and } a_{2,j-1} < a_{2,j}, j \in [n]\},\\
E_3 &= \{(j, s) : a_{1,j} > a_{1,j+1} \text{ and } a_{2,j} > a_{2,j+1}, j \in [n]\}.
\end{aligned}$$

The capacities are given by

$$\begin{aligned}
c(j-1, j) &= \min\{a_{1,j}, a_{2,j}\} \text{ for } (j-1, j) \in E_1,\\
c(q, j-1) &= \min\{a_{1,j} - a_{1,j-1}, a_{2,j} - a_{2,j-1}\} \text{ for } (q, j-1) \in E_2,\\
c(j, s) &= \min\{a_{1,j} - a_{1,j+1}, a_{2,j} - a_{2,j+1}\} \text{ for } (j, s) \in E_3.
\end{aligned}$$

Figure 7 contains the segmentation network for the matrix

$$\begin{pmatrix} 0 & 4 & 3 & 1 & 4 & 5 & 2 & 3 \\ 1 & 3 & 6 & 5 & 7 & 6 & 7 & 4 \end{pmatrix}.$$

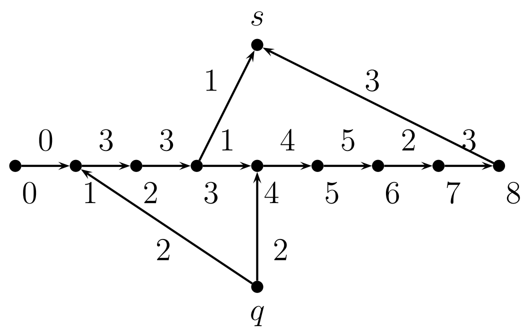


Figure 7: The segmentation network

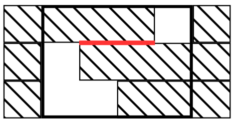
Let $w(A)$ be the maximal value of a flow through N. We still use the notation $\boldsymbol{a}_1, \boldsymbol{a}_2$ for the two rows of A.

Theorem 7 *The minimal DT of a rectangular segmentation of a matrix A with two rows is given by $c(\boldsymbol{a}_1) + c(\boldsymbol{a}_2) - w(A)$ and it can be computed in time $O(n^2)$.*

8 Restricted shapes: The interleaf and the tongue-and-groove constraints

Other restrictions are given by machine dependent constraints. The interleaf collision constraint (ICC) forbids an overlap of the left leaf in row i and the right leaf in row $i \pm 1$, see Figure 8. Formally, we have as condition

$$\text{(ICC)} \qquad l_i \leq r_{i+1} + 1,\ r_i \geq l_{i+1} - 1 \quad (i \in [m-1]).$$



$$\Rightarrow \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \text{ is not admitted.}$$

Figure 8: A forbidden segment by (ICC)

The tongue-and-groove constraint (TGC) should be considered in order to prevent under-dosage effects due to the tongue-and-groove design of the MLC, see Figure 9. A significant

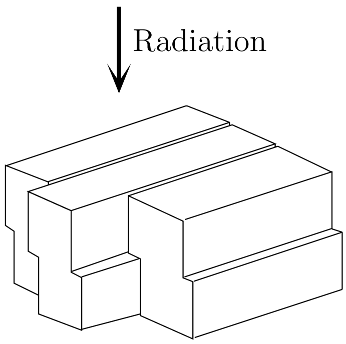


Figure 9: The tongue-and groove design

area between two neighboring leaves is still covered if one of the leaves is open and the other one is closed. Hence such situations have to be avoided as far as possible. So we require that the bixel corresponding to the matrix entry $a_{i\pm1,j}$ is not covered by a leaf if the bixel corresponding to the matrix entry a_{ij} is not covered by a leaf and if $a_{ij} \leq a_{i\pm1,j}$. Formally, we have the condition

$$\text{(TGC)} \qquad \begin{cases} a_{ij} \leq a_{i-1,j} \text{ and } s_{ij} = 1 \Rightarrow s_{i-1,j} = 1 & (i \in [2,m], j \in [n]), \\ a_{ij} \leq a_{i+1,j} \text{ and } s_{ij} = 1 \Rightarrow s_{i+1,j} = 1 & (i \in [m-1], j \in [n]). \end{cases}$$

Again the problems MIN-DT, MIN-DT-APP, MIN-DT-TC-APP can be considered under the additional restrictions (ICC) or/and (TGC). The case (ICC) is relatively well studied. For the problem MIN-DT there are several algorithms. Combinatorial algorithms have been designed by Baatar, Hamacher, Ehrgott, and Woeginger [1] and by Kamath, Sahni, Palta, and Ranka [15], see also [11]. Moreover, Boland, Hamacher, and Lenzen [2] solved this problem by a reduction to a min-cost-circulation problem with side constraints. Because of these additional constraints, this approach finally needs an ILP-solver, a min-cost-circulation algorithm is not sufficient. The algorithm in [8] is based on a duality approach and on a min-max theorem for the following weighted digraph $G = (V, E)$ which is called the ICC-digraph:

$$
\begin{aligned}
V &= \{q, s\} \cup [m] \times [0, n+1], \\
E &= \{(q, (i,0)) \ : \ i \in [m]\} \cup \{((i, n+1), s) \ : \ i \in [m]\} \\
&\quad \cup \{((i,j),(i,j+1)) \ : \ i \in [m], \ j \in [0,n]\} \\
&\quad \cup \{((i,j),(i+1,j)) \ : \ i \in [m-1], \ j \in [n]\} \\
&\quad \cup \{((i,j),(i-1,j)) \ : \ i \in [2,m], \ j \in [n]\},
\end{aligned}
$$

where the arc weights are given by

$$
\begin{aligned}
w(q,(i,0)) = w((i,n+1),s) &= 0 && (i \in [m]) \\
w((i,j-1),(i,j)) &= \max\{0, a_{ij} - a_{i,j-1}\} && (i \in [m], \ j \in [n+1]) \\
w((i,j),(i+1,j)) &= -a_{ij} && (i \in [m-1], \ j \in [n]) \\
w((i,j),(i-1,j)) &= -a_{ij} && (i \in [2,m], \ j \in [n]).
\end{aligned}
$$

Figure 10 shows the ICC-digraph for the matrix

$$
A = \begin{pmatrix} 4 & 5 & 0 & 1 & 4 & 5 \\ 2 & 4 & 1 & 3 & 1 & 4 \\ 2 & 3 & 2 & 1 & 2 & 4 \\ 5 & 3 & 3 & 2 & 5 & 3 \end{pmatrix}.
$$

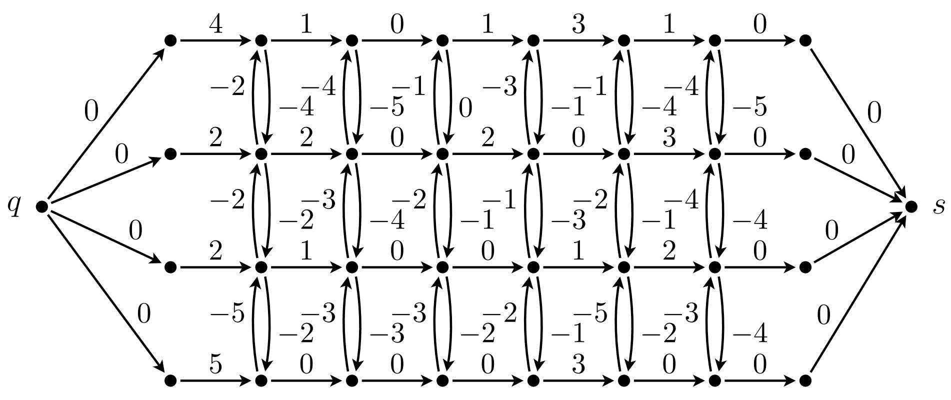


Figure 10: The digraph for the matrix A

Theorem 8 *The minimal DT of a segmentation of a matrix A under (ICC) equals the maximal weight of a q-s-path in the ICC-digraph.*

The problem MIN-DT-APP is solved in [13]. Using an extension of the ICC-digraph one obtains:

Theorem 9 *The problem MIN-DT-APP under (ICC) can be solved in time* $O(\delta^2 m^2 n)$.

In [10] this result is further extended to the problem MIN-DT-TC-APP. Combining the ICC-digraph method with the min-cost-circulation approach from Section 6 the following result is proved:

Theorem 10 *The problem MIN-DT-TC-APP under (ICC) can be solved in time* $O((mn)^2 \log(mn)^2)$.

The ICC-digraph method can also be applied to a combination of (ICC) and (TGC), see [12] for the problem MIN-DT. The situation is more difficult if only the tongue-and-groove condition (TGC) has to be satisfied. Using the method of Boland, Hamacher, and Lenzen [2], the problem MIN-DT under (TGC) is solved in [14]. Because there an ILP-solver has to be used, the problem size cannot be too large. Hence there is still the need of a purely combinatorial algorithm (perhaps including network-flow algorithms). Moreover, it is still unknown whether the corresponding LP-relaxation of the problem (omit the condition that all coefficients u_i have to be integral) always has integral optimal solutions.

References

[1] D. Baatar, H.W. Hamacher, M. Ehrgott, and G.J. Woeginger. Decomposition of integer matrices and multileaf collimator sequencing. *Discrete Appl. Math.*, 152(1-3):6–34, 2005.

[2] N. Boland, H. W. Hamacher, and F. Lenzen. Minimizing beam-on time in cancer radiation treatment using multileaf collimators. *Networks*, 43(4):226–240, 2004.

[3] T.R. Bortfeld, D.L. Kahler, T.J. Waldron, and A.L. Boyer. X–ray field compensation with multileaf collimators. *Int. J. Radiat. Oncol. Biol. Phys.*, 28:723–730, 1994.

[4] K. Engel. A new algorithm for optimal multileaf collimator field segmentation. *Discrete Appl. Math.*, 152(1-3):35–51, 2005.

[5] K. Engel. Optimal matrix–segmentation by rectangles. submitted to *Discrete Appl. Math.*, 2006.

[6] K. Engel and A. Kiesel. Approximated matrix decomposition for IMRT planning with multileaf collimators. submitted to *OR Spectrum*, 2008.

[7] K. Engel and E. Tabbert. Fast simultaneous angle, wedge, and beam intensity optimization in inverse radiotherapy planning. *Optimization and Engineering*, 6(4):393–419, 2005.

[8] T. Kalinowski. A duality based algorithm for multileaf collimator field segmentation with interleaf collision constraint. *Discrete Appl. Math.*, 152(1-3):52–88, 2005.

[9] T. Kalinowski. Realization of intensity modulated radiation fields using multileaf collimators. In R. Ahlswede et al., editor, *Information Transfer and Combinatorics*, volume 4123 of *LNCS*, pages 1010–1055. Springer-Verlag, 2006.

[10] T. Kalinowski. A min cost network flow formulation for approximated MLC segmentation. submitted to *Networks*, 2008.

[11] T. Kalinowski. Multileaf collimator shape matrix decomposition. In G.J. Lim, editor, *Optimization in Medicine and Biology*, pages 253–286. Auerbach Publishers Inc., 2008.

[12] T. Kalinowski. Reducing the tongue-and-groove underdosage in MLC shape matrix decomposition. to appear in *Algorithmic Operations Research*, 2008.

[13] T. Kalinowski and A. Kiesel. Approximated MLC shape matrix decomposition with interleaf collision constraint. submitted to *Algorithmic Operations Research*, 2008.

[14] T. Kalinowski and A. Kiesel. MLC shape matrix decomposition without tongue-and-groove underdosage. submitted to *Networks*, 2008.

[15] S. Kamath, S. Sahni, J. Li, J. Palta, and S. Ranka. Leaf sequencing algorithms for segmented multileaf collimation. *Phys. Med. Biol.*, 48(3):307–324, 2003.