Minimizing the Number of Apertures in Multileaf Collimator Sequencing with Field Splitting

Davaatseren Baatar
School of Mathematical Sciences,
Monash University, Clayton
Victoria 3800, Australia
davaatseren.baatar@monash.edu

Matthias Ehrgott
Department of Management Science
Lancaster University Management School
Bailrigg, Lancaster LA1 4YX, United Kingdom
m.ehrgott@lancaster.ac.uk

Horst W. Hamacher, Ines M. Raschendorfer
Department of Mathematics
University of Kaiserslautern
67663 Kaiserslautern, Germany
{*hamacher,raschendorfer*}*@mathematik.uni-kl.de*

October 9, 2015

Abstract

In this paper we consider the problem of decomposing a given integer matrix A into a positive integer linear combination of consecutive-ones matrices with a bound on the number of columns per matrix. This problem is of relevance in the realization stage of intensity modulated radiation therapy (IMRT) using linear accelerators and multileaf collimators with limited width. Constrained and unconstrained versions of the problem with the objectives of minimizing beam-on time and decomposition cardinality are considered. We introduce a new approach which can be used to find the minimum beam-on time for both constrained and unconstrained versions of the problem. The decomposition cardinality problem is shown to be $\mathcal{NP}$-hard and an approach is proposed to solve the lexicographic decomposition problem of minimizing the decomposition cardinality subject to optimal beam-on time.

Keywords: intensity modulated radiation therapy, multileaf collimator sequencing, field splitting, beam-on time, decomposition cardinality.

1 Introduction

In *intensity modulated radiation therapy (IMRT)*, linear accelerators (*linacs*) (Figure 1) are used to deliver radiation to a target volume (the tumor tissue). The linac is mounted on a gantry which is able to rotate along a central axis while the patient is positioned on a couch that can rotate as well. In this way, it is possible to irradiate the patient from almost any angle. A number of radiation beams is selected and optimal fluence profiles for each beam are determined, which are represented as integer *intensity matrices* (IMs). The entries of an intensity matrix represent exposure times for particular bixels or beamlets of a radiation beam.

Figure 1: Medical linear accelerator from outside and inside.

Source: http://varian.mediaroom.com/index.php?s=13&cat=12&mode=gallery

Radiation passes through a *multileaf collimator (MLC)* (Figure 2) which realizes the fluence profile. The MLC consists of several pairs of identical tungsten alloy leaves. The leaves are positioned in opposing pairs and can move towards the opposing leaf or away from it to block or open the radiation beam. Thereby, the intensity of radiation can be individually controlled for each bixel, which is defined by an area of the radiation field the size of which is equal to the width of a leaf times the length of a minimal feasible move of the leaf. A beam shaping region can thus be created as shown in Figure 2. In this beam-shaping region, all areas not covered are irradiated with the same intensity. Because the dose delivered to the patient body is proportional to exposure time, by overlaying several beam shaping regions (or apertures) it is possible to form any intensity matrix. For more details on the planning process of IMRT please see Schlegel and Mahr [2002] and Ehrgott et al. [2008] and references therein.

Example 1.1 shows how a multileaf collimator is used to create an IM of different intensities. In Figure 3 the darker cells indicate a higher intensity.

Example 1.1. If each of the light grey cells in Figure 3 corresponds to a radiation intensity

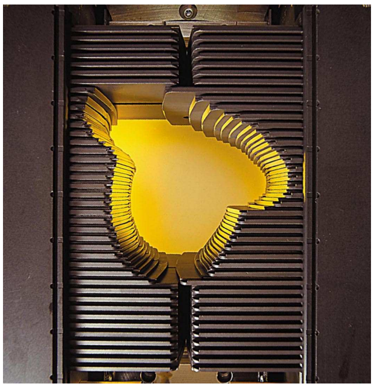

Figure 2: Multileaf Collimator.

Source: http://varian.mediaroom.com/index.php?s=13&cat=22&mode=gallery

of value 1, and each of the dark grey cells corresponds to an intensity of value 2 then the overall intensity distribution can be modeled by the integer intensity matrix

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

The planning process of intensity modulated radiation therapy involves three optimization problems: the optimal selection of the number and angle of the beam directions to be used (the beam angle or geometry optimization problem), the optimization of the fluence maps or intensity matrices for each chosen direction (the fluence map or intensity optimization problem) and finally, the collimator sequencing or realization problem. For an overview of optimization techniques used in IMRT planning we refer to Ehrgott et al. [2008]. In this paper we only discuss the *realization problem*. Therefore we assume that the number and directions of the beams from which the patient is going to be irradiated are already fixed and that optimal intensity matrices for each of these beams are known. The realization problem is to find an efficient delivery sequence, i.e., a sequence of beam shaping regions via MLC adjustments to deliver the corresponding intensity matrix ensuring the best possible treatment. Throughout this paper we will consider *step-and-shoot* static IMRT where the radiation is turned off during the leaf adjustments, i.e., leaves do not move during irradiation.

Depending on the design of MLCs, there may be several technical constraints that have to be respected in the realization problem. In this paper, we consider the *maximum leaf spread constraint* and the *interleaf collision constraint*. The maximum leaf spread constraint restricts the maximum distance between opposing leaves. In other words, the mechanical design of MLCs restricts the beam-shaping region since no leaf can have a larger distance

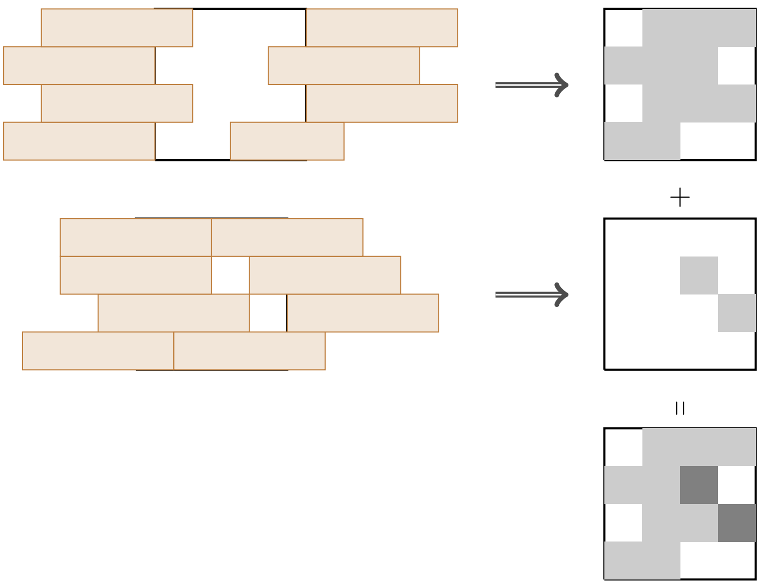

Figure 3: Leaf positions of an MLC and intensity profiles.

from the vertical center line of the MLC than a certain threshold value. For example, size limits for Elekta and Varian MLCs are 12.5 cm and 15 cm, respectively [Chen et al., 2011]. Therefore, large intensity matrices (radiation fields) need to be split into several (adjacent) subfields, where the width of each subfield is not allowed to be larger than a given threshold value. There are two versions of this problem as stated by Chen et al. [2011]:

1. Splitting using vertical lines without overlapping of the subfields,
2. splitting using vertical lines, allowing overlapping of the resulting subfields. In the literature this is often referred to as *field splitting with feathering* [Wu et al., 2000, Liu and Wu, 2010].

In this paper, we focus on field splitting with feathering since the former can be considered as a special case of the latter.

Example 1.2. Consider the intensity matrix A from Example 1.1,

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

Suppose that the maximum field width is 3. Then, in order to realize the intensity profile we need to split it into at least two subfields. For example, one can split the intensity matrix A

into two subfields

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 1 & 2 \\ 0 & 0 \end{pmatrix},$$

such that no overlapping of the subfields occurs, i.e., each entry of the matrix A is covered by only one of the subfields:

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 \end{pmatrix},$$

where the light grey part represents A_1 and the dark grey part represents A_2.

On the other hand, if overlapping is allowed the desired intensities in the feathering region are represented by the sum of subfields in the feathering region. Consider the following split of A into two subfields

$$A_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 2 & 0 \\ 1 & 2 \\ 0 & 0 \end{pmatrix}.$$

Then the desired intensity profile is achieved as

$$A = \begin{pmatrix} 0 & 1 & 1+0 & 1 \\ 1 & 1 & 0+2 & 0 \\ 0 & 1 & 0+1 & 2 \\ 1 & 1 & 0+0 & 0 \end{pmatrix},$$

where the matrices A_1 and A_2 overlap in the third column of A (colored grey) which is represented as the sum of the third and first column of the matrices A_1 and A_2, respectively.

Some commercial MLCs restrict leaf positions in the beam shaping region. More precisely, a leaf is not allowed to be positioned further than the opposing leaves in the adjacent rows. This restriction is referred as *interleaf collision constraint* and extensively studied in Kalinowski [2005] and Baatar et al. [2005]. For example, leaf collision occurs in the last row of the second beam shape in Figure 3, where the right leaf is positioned further than the opposing left leaf in the third row.

The realization problem has a great impact on the quality of the radiation treatment. The quality of the segmentation can be characterized by several features of the segmentation (see, e.g., Ehrgott et al. [2008], Chen et al. [2011], Lim and Lee [2008], Pardalos and Romeijn [2009]). In this paper we consider the total beam-on time and total number of shape matrices. The total beam-on time represents the total amount of time a patient is exposed to radiation, whereas the number of shape matrices represents the total number of adjustments of the leaves (beam shapes) of the MLC required to deliver the IM. Although the realization problem is a multi-objective optimization problem, the algorithms that have been developed for sequencing with field splitting consider only beam-on time (see, e.g., the exact algorithms introduced by Kamath et al. [2004] and Chen et al. [2011]). Our paper will address the cardinality objective function in the sequencing problem with field splitting. This has, to the best of our knowledge, never been discussed in the literature. We also consider the field splitting problem as a lexicographic optimization problem. Moreover, we extend our approach to MLCs with interleaf collision constraints, which also has not been covered in the existing literature. We would like to mention that some of this research originated in the Diploma thesis of Raschendorfer [2011].

The rest of the paper is organized as follows. Single field decomposition problems without field splitting are reviewed in Section 2. The decomposition problem with field splitting is discussed in Section 3, where we also propose our lexicographic optimization approach. We address the complexity of the problems with single objectives and introduce new formulations which can be used for both constrained and unconstrained versions of the problems. Section 4 presents numerical results. In Section 5 we summarize the contributions made by this article and give suggestions for further work.

2 MLC sequencing without field splitting

In this section we review the most relevant results in the literature on MLC sequencing without field splitting. We will follow the notation used in Baatar et al. [2005].

Definition 2.1. An $m \times n$ matrix $Y = (y_{i,j})$, $i = 1, ..., m$, $j = 1, ..., n$ is called a *consecutive ones matrix* or a *C1 matrix*, if for each row i, $i = 1, ..., m$, there exists an integer pair $[l_i, r_i)$, $l_i, r_i \in \{1, ..., n+1\}$, such that

$$y_{i,j} = \begin{cases} 1 & \text{if } l_i \leq j < r_i, \\ 0 & \text{otherwise,} \end{cases}$$

i.e., the ones occur consecutively in a single block in each row.

Obviously, any beam shape can be represented as a C1 matrix [Ahuja and Hamacher, 2004, Baatar et al., 2005, Ehrgott et al., 2008, Neumann, 2009] where ones and zeros represent the bixels where radiation is allowed to pass through or is blocked, respectively. The intervals $[l_i, r_i)$ can be interpreted as the left and right leaf positions, respectively, for the ith pair of leaves. Totally blocked rows can be represented by any of the intervals $[l_i, r_i)$ with $l_i = r_i$. However, it is worth mentioning that they represent different leaf configurations. Some of the presentations might not be valid for MLC's with interleaf collision constraint. For example, the second beam shaping region (leaf configuration) shown in Figure 3 is not valid for such MLCs since collision occurs between the left leaf in the third row and the right leaf in the fourth row. Hereafter, we refer to a C1 matrix as a *shape matrix* if it represents a valid leaf configuration.

Let us denote the set of all C1 matrices as $\mathcal{C}$. For the sake of brevity, we do not specify the dimension of the matrices which will be clear from the context.

Definition 2.2. Let $A \in \mathbb{Z}_{\geq 0}^{m \times n}$ and $\mathcal{C}' \subseteq \mathcal{C}$. Then a *C1 decomposition with respect to* $\mathcal{C}'$ is defined by non-negative integers α_k and C1 matrices Y_k such that

$$A = \sum_{Y_k \in \mathcal{C}'} \alpha_k Y_k.$$

Indeed, the realization problem is a decomposition problem – an integer matrix A is decomposed into a positive integer linear combination of C1 matrices [Ahuja and Hamacher, 2004, Baatar et al., 2005, Ehrgott et al., 2008]. Coefficients α_k represent the beam-on time corresponding to the shape matrices Y_k and are measured in monitor units (MU). Then the problem of minimizing the total beam-on time *(BOT)* can be formulated as

$$
\begin{aligned}
\text{(BOT)} \qquad\qquad BOT(A) \;=\; \min \sum_{k=1}^{|\mathcal{C}'|} \alpha_k & \\
\text{s.t.} \sum_{k=1}^{|\mathcal{C}'|} \alpha_k Y_k \;&=\; A, \\
\alpha_k \;&\in\; \mathbb{Z}_{\geq 0}, \qquad k = 1, \ldots, |\mathcal{C}'|, \\
Y_k \;&\in\; \mathcal{C}', \qquad k = 1, \ldots, |\mathcal{C}'|,
\end{aligned}
$$

where $\mathcal{C}'$ is the set of all admissible shape matrices and *BOT(A)* is the minimum total beam-on time for a C1 decomposition of the matrix A. This formulation can represent both versions of the problem, i.e., the problem with or without interleaf collision constraints. In the first case, the subset $\mathcal{C}'$ corresponds to the set of all C1 matrices which can represent beam shaping regions without violating the constraint. In the latter case, any C1 matrix is a shape matrix, i.e., $\mathcal{C}' = \mathcal{C}$. From now on, to be short, we say the problem is *unconstrained* if there is no interleaf collision constraint and *constrained* otherwise.

In both versions of the problem, we have an exponential number of possible shape matrices. Thus, *(BOT)* is a large scale integer program. However, this problem can be solved efficiently in linear time. There are different constructive exact algorithms available in the literature, see, for example, Baatar et al. [2005] and Engel [2005] for the beam-on time problem without interleaf collision constraint as well as Baatar et al. [2005] and Kalinowski [2005] for the constrained case.

For the unconstrained problem, the minimum beam-on time can be obtained directly from the intensity matrix.

Theorem 2.3. [Engel, 2005, Baatar et al., 2005] For the unconstrained problem, i.e., $\mathcal{C}' = \mathcal{C}$, the minimum total beam-on time is

$$BOT(A) \;=\; \max_{i=1,\ldots,m} \sum_{j=1}^{n+1} \max\left\{0, a_{i,j} - a_{i,j-1}\right\}, \tag{1}$$

where $a_{i,0} = a_{i,n+1} = 0$ for all rows $i = 1, \ldots, m$.

For the constrained problem, the relationship between the total beam-on time and shape matrices can be characterized using a pair of integer matrices:

Theorem 2.4. [Baatar et al., 2005] A matrix $A \in \mathbb{Z}_{\geq 0}^{m \times n}$ has a C1 decomposition w.r.t. $\mathcal{C}'$ with total beam-on time β if and only if there exist $m \times (n+1)$ matrices $L = (l_{i,j})$ and

$R = (r_{i,j})$ with non-negative entries such that

$$
\begin{aligned}
l_{i,j} - r_{i,j} &= a_{i,j} - a_{i,j-1}, && i = 1, \ldots, m,\ j = 1, \ldots, n, && (2)\\
\beta &= \sum_{j=1}^{n+1} l_{i,j} = \sum_{j=1}^{n+1} r_{i,j}, && i = 1, \ldots, m, && (3)\\
\sum_{j=1}^{k} l_{i-1,j} &\leq \sum_{j=1}^{k} r_{i,j}, && i = 2, \ldots, m,\ k = 1, \ldots, n+1, && (4)\\
\sum_{j=1}^{k} l_{i,j} &\leq \sum_{j=1}^{k} r_{i-1,j}, && i = 2, \ldots, m,\ k = 1, \ldots, n+1, && (5)
\end{aligned}
$$

where $a_{i,0} = a_{i,n+1} = 0$ for all rows $i = 1, \ldots, m$.

Constraints (4) and (5) represent the interleaf collision constraints. Note that Theorem 2.4 is valid for MLCs without interleaf collision constraints, in which case we neglect constraints (4) and (5). Matrices L and R represent a set of C1 decompositions and a decomposition can be extracted in linear time (for more details see Baatar et al. [2005]).

The minimization of the number of shape matrices can be formulated as

$$
\begin{aligned}
\text{(DC)} \qquad DC(A) &= \min \sum_{k=1}^{|\mathcal{C}'|} \gamma_k \\
\text{s.t.} \sum_{k=1}^{|\mathcal{C}'|} \alpha_k Y_k &= A, \\
\alpha_k &\leq M\gamma_k, && k = 1, \ldots, |\mathcal{C}'|, \\
\alpha_k &\in \mathbb{Z}_{\geq 0}, && k = 1, \ldots, |\mathcal{C}'|, \\
\gamma_k &\in \mathbb{B}, && k = 1, \ldots, |\mathcal{C}'|, \\
Y_k &\in \mathcal{C}', && k = 1, \ldots, |\mathcal{C}'|,
\end{aligned}
$$

where M is a sufficiently large number. In the literature, the problem (DC) is commonly referred to as *minimum decomposition cardinality problem.* We denote by $DC(A)$ the minimum number of shape matrices required in a C1 decomposition of an integer matrix A.

Obviously, both the (BOT) and (DC) problems are feasible for any positive integer matrix A and a feasible solution can be obtained easily.

Theorem 2.5. The minimum decomposition cardinality problem is strongly $\mathcal{NP}$-hard. In particular, the following results hold.

1. The (DC) problem is strongly $\mathcal{NP}$-hard for matrices with a single row [Baatar et al., 2005].

2. The (DC) problem is strongly $\mathcal{NP}$-hard for matrices with a single column [Collins et al., 2007].

However, some special cases of the (DC) problem can be solved in polynomial time. The following theorem holds for both constrained and unconstrained problems.

Theorem 2.6. [Baatar et al., 2005] If $A = pB$ is a positive integer multiple of a binary matrix B, then the minimum decomposition cardinality problem can be solved in polynomial time.

More generally, considering both beam-on time and decomposition cardinality as objectives to be minimized, the field segmentation problem can be presented as the following multicriteria optimization problem:

$$
\begin{aligned}
\min & \begin{pmatrix} \sum_{k=1}^{|\mathcal{C}'|} \alpha_k \\ \sum_{k=1}^{|\mathcal{C}'|} \gamma_k \end{pmatrix} & & \\
\text{s.t.} \sum_{k=1}^{|\mathcal{C}'|} \alpha_k Y_k &= A, & & \\
\alpha_k &\leq M\gamma_k, & & k = 1, \dots, |\mathcal{C}'|, \\
\alpha_k &\in \mathbb{Z}_{\geq 0}, & & k = 1, \dots, |\mathcal{C}'|, \\
\gamma_k &\in \mathbb{B}, & & k = 1, \dots, |\mathcal{C}'|, \\
Y_k &\in \mathcal{C}', & & k = 1, \dots, |\mathcal{C}'|,
\end{aligned}
$$

where the objectives represent the total beam-on time and the number of shape matrices. In the literature, a lexicographic optimization approach is proposed to find a Pareto optimal solution of the problem (see, for example, Baatar et al. [2005] and Kalinowski [2005]).

3 The matrix decomposition problem with field splitting

Analogously to Section 2, MLC sequencing with field splitting can in general be formally presented as the multicriteria optimization problem (FS):

$$
\text{(FS)} \qquad \begin{aligned}
\min & \begin{pmatrix} \sum_{k=1}^{d} \sum_{t=1}^{|\mathcal{C}'|} \alpha_{kt} \\ \sum_{k=1}^{d} \sum_{t=1}^{|\mathcal{C}'|} \gamma_{kt} \end{pmatrix} \\
\text{s.t.} & \sum_{k=1}^{d} [\ A_k\ ]_{s_k} \ = \ A
\end{aligned}
$$

$$
\begin{array}{rcll}
A_k & = & \sum_{t=1}^{|\mathcal{C}'|} \alpha_{kt} Y_t, & k = 1, \ldots, d, \\
\alpha_{kt} & \leq & M \gamma_{kt}, & k = 1, \ldots, d,\ t = 1, \ldots, |\mathcal{C}'|, \\
A_k & \in & \mathbb{Z}_{\geq 0}^{m \times w}, & k = 1, \ldots, d, \\
s_k & \in & \{1, \ldots, n\}, & k = 1, \ldots, d, \\
\gamma_{kt} & \in & \mathbb{B}, & k = 1, \ldots, d,\ t = 1, \ldots, |\mathcal{C}'|, \\
\alpha_{kt} & \in & \mathbb{Z}_{\geq 0}, & k = 1, \ldots, d,\ t = 1, \ldots, |\mathcal{C}'|, \\
Y_t & \in & \mathcal{C}', & t = 1, \ldots, |\mathcal{C}'|,
\end{array}
$$

where $d = \lceil \frac{n}{w} \rceil$ and $[\ A_k\ ]_{s_k}$ represents a $m \times n$ matrix with columns from s_k to $s_k + w - 1$ represented by the matrix A_k and the remaining columns all being 0. Here w is the maximum leaf spread. In other words, the matrix A is split into d submatrices with w columns each, such that the C1 decompositions of the submatrices yield an as small as possible total beam-on time and decomposition cardinality. Note that the column indices s_k are unknown and submatrices A_k can be overlapping.

In the literature, the number of subfields is usually defined as $d = \lceil \frac{n}{w} \rceil$ (see, for example, Chen et al. [2011]). However, we were not able to find any reference to explain why this number of subfields was chosen. We propose the following Proposition 3.1 to provide the answer to this question.

Proposition 3.1. Splitting increases the beam-on time and cardinality, i.e., for any feasible split of the matrix $A = \sum_{k=1}^{q} [\ A_k\ ]_{s_k}$ it holds, that

$$
\begin{aligned}
BOT(A) &\leq \sum_{k=1}^{q} BOT(A_k), \\
DC(A) &\leq \sum_{k=1}^{q} DC(A_k).
\end{aligned}
$$

The statement is true since any shape matrix of any subfield A_k can be presented as a shape matrix of A.

Example 3.2. Let us consider the following matrices:

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 \end{pmatrix}, A_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

One can easily check that the matrices A_1 and A_2 provide a feasible solution solution to (FS) that minimizes both objectives of (FS) (also called an ideal solution in multicriteria optimization) with the following C1 decompositions.

$$\begin{aligned} A_1 &= \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \\ A_2 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \end{aligned}$$

which can be represented as C1 decomposition of the matrix A by adding all 0 columns to the shape matrices, i.e.,

$$A = \begin{pmatrix} 0 & 1 & 1 & \mathbf{0} \\ 1 & 1 & 0 & \mathbf{0} \\ 0 & 1 & 0 & \mathbf{0} \\ 1 & 1 & 0 & \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{0} & 0 & 0 & 0 \\ \mathbf{0} & 0 & 1 & 0 \\ \mathbf{0} & 0 & 1 & 1 \\ \mathbf{0} & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} \mathbf{0} & 0 & 0 & 1 \\ \mathbf{0} & 0 & 1 & 0 \\ \mathbf{0} & 0 & 0 & 1 \\ \mathbf{0} & 0 & 0 & 0 \end{pmatrix}.$$

In other words, from an ideal solution to (FS) we constructed a feasible C1 decomposition of the matrix A. Thus, the statement of Proposition 3.1 holds. Indeed, an ideal C1 decomposition of the matrix A is

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The field splitting problem (FS), to the best of our knowledge, has never been considered as a multi-objective optimization problem or even as a single objective optimization problem with the cardinality objective even though both objectives are important in IMRT [Ehrgott et al., 2008]. Moreover, algorithms minimizing the beam-on time often produce a large number of shape matrices [Ehrgott et al., 2008]. In the next sections, we develop a lexicographic optimization approach to find a Pareto optimal solution of the problem (FS). This Pareto optimal solution has a practical significance which can be interpreted as prioritization of reducing the total time a patient is exposed to radiation and then decreasing the treatment time by minimizing the number of shape matrices within the given beam-on time.

Kamath et al. [2004] and Chen et al. [2011] consider the field splitting problem with a single objective – total beam-on time. They introduce constructive exact algorithms for the unconstrained (BOT) problem. In the next sections we develop a linear programming based approach which can be used to exactly solve both the constrained and unconstrained field splitting problem with feathering.

Theorem 3.3. The minimization of the number of shape matrices with field splitting is a strongly $\mathcal{NP}$-hard problem even for a single row intensity matrix and field splitting without feathering.

Proof. Let us consider a row intensity matrix

$$A \;=\; (a_1, a_2, \ldots, a_w, 0, \ldots, 0, a_{2w}) \in \mathbb{Z}^{2w}$$

with the last w entries being 0 except for the very last entry. Obviously, $d = 2$ and the matrix must be split as

$$A \;=\; [(a_1, a_2, \ldots, a_w)]_1 + [0, \ldots, 0, a_{2w}]_{w+1}.$$

The second matrix can be realized using a single shape matrix. Thus, the minimization of the number of shape matrices is equivalent to the minimization of the number of shape matrices in the C1 decomposition of the row matrix $(a_1, a_2, \ldots, a_w)$, which is strongly $\mathcal{NP}$-hard due to Theorem 2.5. □

3.1 Minimization of beam-on time

In this section, we consider the minimization of beam-on time with field splitting. There are several algorithms available in the literature, for example, see Kamath et al. [2004] or Chen et al. [2011]. However, those algorithms are for the unconstrained version of the problem. In this section we develop a new approach which can be used for both the constrained and unconstrained versions of the problem. The minimization of beam-on time with field splitting can be formally formulated as (FSBOT):

$$\text{(FSBOT)} \qquad \min \sum_{k=1}^{d} BOT(A_k)$$

$$\begin{aligned} \text{s.t.} \sum_{k=1}^{d} [\, A_k \,]_{s_k} \;&=\; A, & \\ A_k \;&\in\; \mathbb{Z}_{\geq 0}^{m \times w}, & k = 1, \ldots, d, \\ s_k \;&\in\; \{1, \ldots, n\}\,, & k = 1, \ldots, d. \end{aligned}$$

Due to Theorem 2.4, each subfield A_k can be presented by a pair of matrices L^k and R^k. Moreover, the beam-on time and interleaf collision constraints can be displayed using this pair of matrices. First we reformulate Theorem 2.4 in terms of cumulative sums of the elements of the matrices. Let us denote by $c^l_{i,j}$ and $c^r_{i,j}$ the row-wise cumulative sum of the entries of the matrices L and R, respectively, i.e.,

$$c^l_{i,j} = \sum_{q=1}^{j} l_{i,q}, \quad c^r_{i,j} = \sum_{q=1}^{j} r_{i,q}, \qquad i = 1, \dots, m, \; j = 1, \dots, n+1. \tag{6}$$

Theorem 3.4. A matrix $A \in \mathbb{Z}^{m \times n}_{\geq 0}$ has a C1 decomposition w.r.t. $\mathcal{C}'$ with total beam-on time β if and only if there exist $m \times (n+1)$ matrices $C^l = (c^l_{i,j})$ and $C^r = (c^r_{i,j})$ with non-negative entries such that

$$\begin{aligned}
c^l_{i,j} - c^r_{i,j} &= a_{i,j}, && i = 1, \dots, m, \; j = 1, \dots, n, && (7)\\
\beta &= c^l_{i,n+1} = c^r_{i,n+1}, && i = 1, \dots, m, && (8)\\
c^l_{i,j-1} &\leq c^l_{i,j}, && i = 1, \dots, m, \; j = 2, \dots, n+1, && (9)\\
c^r_{i,j-1} &\leq c^r_{i,j}, && i = 1, \dots, m, \; j = 2, \dots, n+1, && (10)\\
c^l_{i-1,j} &\leq c^r_{i,j}, && i = 2, \dots, m, \; j = 1, \dots, n+1, && (11)\\
c^l_{i,j} &\leq c^r_{i-1,j}, && i = 2, \dots, m, \; j = 1, \dots, n+1. && (12)
\end{aligned}$$

Constraints (9) and (10) ensure that the entries of the matrices C^l and C^r represent cumulative sums. The interleaf collision constraints are given by constraints (11) and (12). Theorem 3.4 is valid for the unconstrained problem as well, since we can just disregard the interleaf collision constraints in that case. The existence of matrices C^l and C^r represents the necessary and sufficient condition for a total beam-on time of β in a more compact form than matrices L and R. Due to (6), matrices L and R can be obtained easily from the matrices C^l and C^r.

The problem (FSBOT) can be represented in terms of the matrices C^l and C^r as the following integer program, (FSBOT'):

$$
\begin{aligned}
\text{(FSBOT')} \qquad & \min \sum_{k=1}^{d} \beta_k \\
\text{s.t.} \quad & \sum_{k=1}^{d} [\, C^{lk} - C^{rk} \,]_{s_k} \;=\; A && && (13) \\
& \beta_k \;=\; c^{lk}_{i,w+1} \;=\; c^{rk}_{i,w+1}, && k = 1,\dots,d,\ i = 1,\dots,m, && (14) \\
& c^{rk}_{i,j} \;\le\; c^{lk}_{i,j}, && k = 1,\dots,d,\ i = 1,\dots,m, \\
& && j = 1,\dots,w, && (15) \\
& c^{lk}_{i,j-1} \;\le\; c^{lk}_{i,j}, && k = 1,\dots,d,\ i = 1,\dots,m, \\
& && j = 2,\dots,w+1, && (16) \\
& c^{rk}_{i,j-1} \;\le\; c^{rk}_{i,j}, && k = 1,\dots,d,\ i = 1,\dots,m, \\
& && j = 2,\dots,w+1, && (17) \\
& c^{lk}_{i-1,j} \;\le\; c^{rk}_{i,j}, && k = 1,\dots,d,\ i = 2,\dots,m, \\
& && j = 1,\dots,w+1, && (18) \\
& c^{lk}_{i,j} \;\le\; c^{rk}_{i-1,j}, && k = 1,\dots,d,\ i = 2,\dots,m, \\
& && j = 1,\dots,w+1, && (19) \\
& c^{lk}_{i,j}, c^{rk}_{i,j}, \beta_k \;\in\; \mathbb{Z}_{\ge 0}, && k = 1,\dots,d,\ i = 1,\dots,m, \\
& && j = 1,\dots,w+1, \\
& s_k \;\in\; \{1,\dots,n\}, && k = 1,\dots,d.
\end{aligned}
$$

The integer program (FSBOT') can be used for both constrained and unconstrained versions of the problem. For the unconstrained case we have to remove constraints (18) and (19) which represent the interleaf collision constraints. Some of the constraints in the formulation are redundant and can be removed or reformulated to make the formulation compact and tighter. However, we keep the formulation as it is stated in order to avoid complicated notations and make it easier to follow the main ideas.

For any fixed positions $(s_1, \dots, s_d)$ of the subfields the corresponding integer program can be solved efficiently. Indeed, the feasible set is an integral polyhedron.

Theorem 3.5. For any fixed positions of the submatrices the problem (FSBOT') can be solved in polynomial time.

Proof. We provide a sketch of the proof. We show that the feasible set defined by constraints (13) to (19) is an integral polyhedron. The coefficient matrix provided by (13) to (19) can be represented by a block matrix $[\tilde{C}^l \ \tilde{C}^r]$ where $\tilde{C}^l$ and $\tilde{C}^r$ represent coefficients corresponding to the variables $c^{lk}_{i,j}$ and $c^{rk}_{i,j}$, respectively. Consider any subset J^l of columns of the matrix $\tilde{C}^l$. One can show that the set J^l can be partitioned into two subsets J^l_1 and J^l_2 such that the following inequality holds for any row i of the matrix $\tilde{C}^l$

$$0 \;\le\; \sum_{j \in J^l_1} \tilde{c}^l_{i,j} - \sum_{j \in J^l_2} \tilde{c}^l_{i,j} \;\le\; 1.$$

Note that each row of the matrix $\tilde{C}^l$ has at most two non-zero entries. The same statement is true for the block matrix $\tilde{C}^r$. Then the proof immediately follows from the well known Ghouila-Houri characterization of total unimodularity [Ghouila-Houri, 1962].

□

One might develop a constructive algorithm to find an optimal splitting of the matrix A, in the sense of beam-on time, for fixed subfield positions $s_1, \ldots, s_d$. This can be done in the same manner as the constructive algorithm developed for the single field realization problem in Baatar et al. [2005]. However, in this paper we use integer programming to find the best splitting with the smallest cardinality. An intensity matrix A can be split in many different ways with the same total beam-on time and using the same positions for the subfields. Although the total beam-on time is the same, the cardinalities of the decompositions might differ.

Example 3.6. Consider a field splitting problem with $w = 5$ and a single row intensity matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 \end{pmatrix}.$$

Obviously, there is only one possible position for the first columns of the matrices $s_1 = 1$ and $s_2 = 4$. Moreover, one can easily see that the minimum beam-on time is 3 for the field splitting with $w = 5$. The matrix A can be split in two different ways such that minimum beam-on time is achieved:

$$\begin{aligned} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 \end{pmatrix} &= [\, 1 \;\; 1 \;\; 1 \;\; 0 \;\; 0 \,]_1 + [\, 1 \;\; 1 \;\; 2 \;\; 2 \;\; 2 \,]_4, \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 \end{pmatrix} &= [\, 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \,]_1 + [\, 0 \;\; 0 \;\; 2 \;\; 2 \;\; 2 \,]_4. \end{aligned}$$

The total minimum beam-on time for both cases is 3. However, we need three and two shape matrices, respectively, to achieve the minimum beam-on time.

The state-of-the-art exact algorithms proposed by Kamath et al. [2004] and Chen et al. [2011], for unconstrained beam-on time minimization, consider all possible positions of the subfields and for any fixed positions an optimal split is obtained using constructive algorithms. In this paper we follow the same exhaustive approach to identify the best positions $s_1^*, \ldots, s_d^*$.

3.2 Decomposition cardinality and lexicographic optimization

We use a lexicographic approach to find a Pareto optimal solution of (FS), i.e., first we minimize the total beam-on time and then the total number of shape matrices with respect to the minimum beam-on time. There are cases such that for any fixed positions, different sets of the matrices A_k can be constructed with the same total beam-on time. For example, see the instance considered in Example 3.6. Thus, to minimize the number of shape matrices we use only the positions of the matrices $s_1^*, \dots, s_d^*$ at which the minimum beam-on time can be achieved.

In this section, we focus on the constrained problem and formulate it as a mixed integer program which can be used for the unconstrained problem by removing the interleaf collision constraints. Suppose that $s_1^*, \dots, s_d^*$ are positions of the subfields at which the minimum beam-on time can be achived. Then the decomposition cardinality problem in the lexicographic approach can be formally written as an integer program, (FSDC):

$$
\begin{array}{llrcll}
\text{(FSDC)} & & \min \displaystyle\sum_{k=1}^{d}\sum_{t=1}^{|\mathcal{C}'|}\gamma_{kt} & & & \\
& & \text{s.t. } \displaystyle\sum_{k=1}^{d}[\ A_k\ ]_{s_k^*} & = & A, & \\
& & A_k & = & \displaystyle\sum_{t=1}^{|\mathcal{C}'|}\alpha_{kt}Y_t, & k=1,\dots,d, \\
& & \alpha_{kt} & \leq & M\gamma_{kt}, & k=1,\dots,d,\ t=1,\dots,|\mathcal{C}'|, \\
& & \beta^* & = & \displaystyle\sum_{k=1}^{d}\sum_{t=1}^{|\mathcal{C}'|}\alpha_{kt}, & \\
& & A_k & \in & \mathbb{Z}_{\geq 0}^{m\times w}, & k=1,\dots,d, \\
& & \gamma_{kt} & \in & \mathbb{B}, & k=1,\dots,d,\ t=1,\dots,|\mathcal{C}'|, \\
& & \alpha_{kt} & \in & \mathbb{Z}_{\geq 0}, & k=1,\dots,d,\ t=1,\dots,|\mathcal{C}'|, \\
& & Y_t & \in & \mathcal{C}', & t=1,\dots,|\mathcal{C}'|,
\end{array}
$$

where β^* is the minimum beam-on time.

For each single field the decomposition cardinality and the beam-on time have the relationship presented in the following. Let us therefore consider a single field realization of an intensity matrix B.

Theorem 3.7. An intensity matrix $B \in \mathbb{Z}_{\geq 0}^{m\times n}$ can be realized using p shape matrices if and only if for some q there exists a decomposition

$$B = \sum_{k=1}^{q}\alpha_k B_k \tag{20}$$

with $\alpha_k \in \mathbb{Z}_{\geq 0}$, $B_k \in \mathbb{Z}_{\geq 0}^{m\times n}$, $k = 1, \dots, q$, such that

$$p = \sum_{k=1}^{q} BOT(B_k) \tag{21}$$

Proof. The proof is straightforward. If B can be realized using p shape matrices, i.e.,

$$B \ = \ \sum_{k=1}^{p}\alpha_k S_k$$

then by choosing $B_k = S_k$, $k = 1, \dots, p$ we get the decomposition.

Suppose, for some q, there is a decomposition of B

$$B = \sum_{k=1}^{q} \alpha_k B_k$$

with

$$p = \sum_{k=1}^{q} BOT(B_k).$$

For each matrix B_k, $k = 1, \dots, q$, consider a realization

$$B_k = \sum_{j=1}^{BOT(B_k)} S_{kj}.$$

Then the matrix B can be represented as an integer linear combination of p shape matrices as

$$B = \sum_{k=1}^{q} \alpha_k \sum_{j=1}^{BOT(B_k)} S_{kj}.$$

□

Note that in Theorem 3.7 the number of matrices q is not fixed. Moreover, some of the shape matrices might be used several times. From Theorem 3.7 the following characterizations of the decompositions with smallest cardinality can immediately be deduced.

Corollary 3.8. Let p be the minimum decomposition cardinality of B. Then

1. The following statements are true for any decomposition $B = \sum_{k=1}^{q} \alpha_k B_k$ with $p = \sum_{k=1}^{q} BOT(B_k)$ where $\alpha_k \in \mathbb{Z}_{\geq 0}$, $B_k \in \mathbb{Z}_{\geq 0}^{m \times n}$, $k = 1, \dots, q$.
   (a) $B_k \neq B_h$ for all $k \neq h$, $k, h = 1, \dots, q$.
   (b) For any realizations of the matrices
   $$B_k = \sum_{j=1}^{q_k} \gamma_{kj} S_{kj} \text{ with } BOT(B_k) = \sum_{j=1}^{q_k} \gamma_{kj}, \; k = 1, \dots, q,$$
   - $\gamma_{kj} = 1$ for all $k = 1, \dots, q$, $j = 1, \dots, q_k$;
   - $S_{kj} \neq S_{ht}$ for all $k \neq h, k, h = 1, \dots, q$ and $j = 1, \dots, q_k$, $t = 1, \dots, q_t$.
2. There always exists a decomposition of B which satisfies the conditions in 1. and
   $$\alpha_k \neq \alpha_h$$
   for all $k \neq h$, $k, h = 1, \dots, q$.

Corollary 3.8 characterizes well the decompositions of a matrix B with the smallest cardinality. Moreover, these provide the opportunity to express the decomposition cardinality of a matrix by the sum of minimum beam-on times of the matrices used in the decomposition. In other words, the decomposition cardinality problem is equivalent to the decomposition of the intensity matrix into a positive linear combination of integer matrices such that the sum of total beam-on times of the integer matrices are minimized. Thus, the problem (FSDC) can be reformulated as

$$\text{(FSDC')} \qquad \begin{aligned} \min \sum_{k=1}^{d}\sum_{z=1}^{q_k} BOT(B_{kz}) & \\ \text{s.t. } \sum_{k=1}^{d}[\,\sum_{z=1}^{q_k} zB_{kz}\,]_{s_k^*} &= A, \\ \beta^* &= \sum_{k=1}^{d}\sum_{z=1}^{q_k} zBOT(B_{kz}), \\ B_{kz} &\in \mathbb{Z}_{\geq 0}^{m\times w}, \qquad z=1,\dots,q_k,\ k=1,\dots,d, \end{aligned}$$

where q_k is the number of different values of the coefficients of the matrices B_{kz} in the integer decomposition of the matrix A_k. The number of different values q_k can be determined by the largest possible entry of the matrix A_k and the coefficients are defined in this range.

Due to Theorem 3.4, we can represent the matrices B_{kz} by a pair of matrices C^{kzl} and

C^{kzr} which leads us to the following integer program, (FSDC"):

$$
\begin{array}{lrcll}
\text{(FSDC")} & \min \sum_{k=1}^{d} \sum_{z=1}^{q_k} \beta_{kz} & & & \\
& \text{s.t. } \sum_{k=1}^{d} [\sum_{z=1}^{q_k} z(C^{kzl} - C^{kzr}) \;]_{s_k^*} & = & A, & \\
& \beta^* & = & \sum_{k=1}^{d} \sum_{z=1}^{q_k} z\beta_{kz}, & \\
& \beta_{kz} & = & c^{kzl}_{i,w+1} \;=\; c^{kzr}_{i,w+1}, & i = 1, \ldots, m,\ z = 1, \ldots, q_k, \\
& & & & k = 1, \ldots, d, \\
& c^{kzr}_{i,j} & \leq & c^{kzl}_{i,j}, & i = 1, \ldots, m,\ j = 1, \ldots, w, \\
& & & & z = 1, \ldots, q_k,\ k = 1, \ldots, d, \\
& c^{kzl}_{i,j-1} & \leq & c^{kzl}_{i,j}, & i = 1, \ldots, m,\ j = 2, \ldots, w+1, \\
& & & & z = 1, \ldots, q_k,\ k = 1, \ldots, d, \\
& c^{kzr}_{i,j-1} & \leq & c^{kzr}_{i,j}, & i = 1, \ldots, m,\ j = 2, \ldots, w+1, \\
& & & & z = 1, \ldots, q_k,\ k = 1, \ldots, d, \\
& c^{kzl}_{i-1,j} & \leq & c^{kzr}_{i,j}, & i = 2, \ldots, m,\ j = 1, \ldots, w+1, \\
& & & & z = 1, \ldots, q_k,\ k = 1, \ldots, d, \\
& c^{kzl}_{i,j} & \leq & c^{kzr}_{i-1,j}, & i = 2, \ldots, m,\ j = 1, \ldots, w+1, \\
& & & & z = 1, \ldots, q_k,\ k = 1, \ldots, d, \\
& c^{kzl}_{i,j}, c^{kzr}_{i,j}, \beta_{kz} & \in & \mathbb{Z}_{\geq 0}, & i = 1, \ldots, m,\ j = 1, \ldots, w+1, \\
& & & & z = 1, \ldots, q_k,\ k = 1, \ldots, d.
\end{array}
$$

#	m	n	a_{max}	w	d	#pos	#	m	n	a_{max}	w	d	#pos
1	5	6	27	5	2	1	32	10	10	14	5	2	1
2	5	7	27	5	2	1	33	14	10	10	5	2	1
3	5	7	30	5	2	1	34	14	10	10	5	2	1
4	5	8	18	5	2	1	35	14	10	10	5	2	1
5	5	8	25	5	2	1	36	14	10	10	5	2	1
6	11	8	21	5	2	1	37	14	10	10	5	2	1
7	9	9	10	5	2	1	38	15	10	10	5	2	1
8	9	9	10	5	2	1	39	11	11	22	5	3	5
9	10	9	10	5	2	1	40	9	12	29	5	3	4
10	10	9	10	5	2	1	41	9	12	31	5	3	4
11	10	9	10	5	2	1	42	11	12	16	5	3	4
12	11	9	14	5	2	1	43	11	12	19	5	3	4
13	11	9	16	5	2	1	44	11	12	26	5	3	4
14	9	10	10	5	2	1	45	9	13	29	5	3	3
15	9	10	35	5	2	1	46	11	14	22	5	3	2
16	9	10	40	5	2	1	47	10	15	26	14	2	1
17	10	10	10	5	2	1	48	22	15	26	14	2	1
18	10	10	14	5	2	1	49	23	16	33	14	2	1
19	10	10	14	5	2	1	50	23	17	27	14	2	1
20	10	10	14	5	2	1	51	22	18	31	14	2	1
21	10	10	14	5	2	1	52	22	21	31	14	2	1
22	10	10	14	5	2	1	53	22	22	22	14	2	1
23	10	10	14	5	2	1	54	20	23	10	14	2	1
24	10	10	14	5	2	1	55	22	23	24	14	2	1
25	10	10	14	5	2	1	56	20	25	9	14	2	1
26	10	10	14	5	2	1	57	16	27	10	14	2	1
27	10	10	14	5	2	1	58	15	28	9	14	2	1
28	10	10	14	5	2	1	59	16	28	10	14	2	1
29	10	10	14	5	2	1	60	16	28	10	14	2	1
30	10	10	14	5	2	1	61	16	29	10	14	3	14
31	10	10	14	5	2	1	62	16	30	10	14	3	13

Table 1: Description of the 62 instances numbered by #: m number of rows, n number of columns, a_{max} maximum intensity level, w maximum separation in terms of columns, d number of subfields, #pos number of possible splitting positions.

4 Numerical results

We tested our approach using CPLEX 12.6 embedded in C++ on a Linux machine with 32Gb RAM, Intel Xeon 6 core, 3.5 GHz. We used 47 clinical examples varying in size from 5 to 23 rows and 6 to 30 columns, with a_{max} varying between 9 and 40. In addition, we used 15 instances of size 10×10 with entries randomly generated between 1 and 14. In Table 1 we show the dimensions and maximum intensity levels of the intensity matrices as well as the number of subfields and possible splitting positions for the subfields. The number of columns of the intensity matrices ranges from 6 – 30. A split width of 5 columns was used for matrices with less than 15 columns and a split width of 14 columns for the remaining 16 instances.

First we tested our proposed LP based approaches for constrained and unconstrained versions of the (FSBOT') problems. The computational results are shown in Table 2 and Table 3 for constrained and unconstrained (FSBOT') problems, respectively. For most of the instances the minimum beam-on time was attained at a single set of positions of the subfields. For several instances the minimum beam-on time was achieved at several different positions of the subfields, for the unconstrained (FSBOT') problem in 5 instances, 39, 42, 44, 45 and 62, and for the constrained problem in 4 instances 39, 42, 45 and 62. The C1 decomposition cardinalities of the solutions are shown in column DC of the tables where for multiple subfield positions the minimum, mean and maximum cardinalities are presented. All problems were solved in less than 1 second.

Next we tested the lexicographic optimization approach to the field splitting problem (FS). The decomposition cardinality was minimized using the optimal subfield positions obtained from solving the (FSBOT') problems. In this way we obtained the decomposition with the smallest number of shape matrices among those having the minimum beam-on time. For instances with multiple subfield positions for the (FSDC") problem we used the best decomposition cardinality as an upper bound for the next (FSDC") problems to reduce the computational effort. In our tests we set a time limit of 600 seconds for each (FSDC") problem. Table 4 and Table 5 present the results for unconstrained and constrained problems, respectively. The subcolumns "s", "f" and "a" of the column "status" represent the number of (FSDC") problems solved exactly, for which feasible solutions were found, or aborted reaching the time limit without finding a better solution than the current incumbent solution. Column ΔDC shows the percentage of improvement in the number of shape matrices. In case of multiple positions for subfields the minimum, mean and maximum improvements are presented.

For the unconstrained case, CPLEX was able to solve 54 instances exactly, for 6 instances it found a feasible solution and for 2 instances it reached the time limit and could not find a better solution than the decomposition that is obtained by solving the (FSBOT') problem. For the constrained (FSDC") problem, 50 instances were solved exactly and for 5 instances a feasible solution was obtained whereas 7 instances failed to produce a feasible solution within the time limit. The lexicographic approach reduced the number of shape matrices on average by 55% for unconstrained (FS) and by 49.4% for constrained (FS) in comparison to the optimal solutions of (FSBOT') that were computed without consideration of the number of shape matrices.

#	β^*	#pos	DC min	DC μ	DC max	#	β^*	#pos	DC min	DC μ	DC max
1	37	1		20		32	46	1		37	
2	34	1		21		33	26	1		24	
3	40	1		16		34	34	1		31	
4	36	1		20		35	32	1		32	
5	46	1		22		36	27	1		27	
6	36	1		30		37	32	1		28	
7	19	1		19		38	29	1		27	
8	21	1		20		39	50	2	39	40.5	42
9	22	1		21		40	69	1		45	
10	21	1		19		41	74	1		42	
11	25	1		22		42	29	2	25	25.5	26
12	27	1		23		43	53	1		35	
13	26	1		23		44	47	2	38	38	38
14	26	1		22		45	74	2	43	45	47
15	60	1		40		46	50	1		39	
16	100	1		37		47	49	1		32	
17	22	1		22		48	33	1		28	
18	47	1		41		49	35	1		34	
19	46	1		41		50	46	1		43	
20	48	1		41		51	41	1		41	
21	42	1		37		52	55	1		46	
22	45	1		40		53	47	1		46	
23	45	1		36		54	14	1		14	
24	48	1		38		55	44	1		40	
25	46	1		34		56	19	1		19	
26	49	1		36		57	17	1		17	
27	45	1		41		58	20	1		17	
28	53	1		46		59	17	1		17	
29	49	1		42		60	17	1		17	
30	50	1		39		61	19	1		18	
31	50	1		40		62	25	4	24	24.75	25

Table 2: Unconstrained (FSBOT'): β^* – minimum beam-on time, #pos – number of subfield positions where the minimum is achieved; min, μ and max – minimum, mean and maximum number of shape matrices, respectively.

			DC						DC		
#	β^*	#pos	min	μ	max	#	β^*	#pos	min	μ	max
1	37	1		20		32	46	1		38	
2	34	1		20		33	31	1		26	
3	40	1		17		34	40	1		35	
4	36	1		20		35	32	1		31	
5	46	1		23		36	27	1		26	
6	36	1		30		37	32	1		28	
7	19	1		19		38	29	1		28	
8	21	1		20		39	50	2	38	40	42
9	22	1		19		40	69	1		46	
10	21	1		19		41	76	1		46	
11	25	1		24		42	29	2	25	25	25
12	29	1		25		43	53	1		38	
13	26	1		22		44	51	1		39	
14	26	1		22		45	74	2	44	44.5	45
15	60	1		39		46	52	1		43	
16	100	1		36		47	54	1		35	
17	22	1		22		48	42	1		35	
18	47	1		42		49	48	1		47	
19	47	1		39		50	46	1		44	
20	48	1		42		51	41	1		41	
21	42	1		39		52	58	1		50	
22	45	1		40		53	58	1		56	
23	45	1		37		54	14	1		14	
24	54	1		43		55	44	1		40	
25	46	1		36		56	19	1		19	
26	49	1		38		57	18	1		18	
27	45	1		41		58	20	1		17	
28	53	1		45		59	18	1		18	
29	49	1		43		60	21	1		19	
30	54	1		42		61	20	1		19	
31	50	1		41		62	26	10	24	25.2	26

Table 3: Constrained (FSBOT'): β^* – minimum beam-on time, #pos – number of subfield positions where the minimum is achieved; min, μ and max – minimum, mean and maximum number of shape matrices, respectively.

		Status				$\Delta DC(\%)$					Status				$\Delta DC(\%)$		
#	DC	s	f	a	t (sec.)	min	μ	max	#	DC	s	f	a	t (sec.)	min	μ	max
1	8	1			5		60.0		32	13	1			14		64.9	
2	8	1			0		61.9		33	10	1			2		58.3	
3	7	1			2		56.3		34	12	1			1		61.3	
4	9	1			2		55.0		35	12	1			1		62.5	
5	10	1			4		54.5		36	12	1			2		55.6	
6	9	1			1		70.0		37	12	1			1		57.1	
7	8	1			0		57.9		38	12	1			1		55.6	
8	10	1			1		50.0		39	14	2	0	0	9	35.9	51.3	66.7
9	10	1			0		52.4		40	14	1			378		68.9	
10	10	1			0		47.4		41	16	1			402		61.9	
11	10	1			1		54.5		42	11	2	0	0	2	56.0	56.8	57.7
12	10	1			1		56.5		43	14	1			4		60.0	
13	9	1			0		60.9		44	14	2	0	0	8	63.2	63.2	63.2
14	10	1			1		54.5		45	16	2	0	0	338	60.5	63.2	66.0
15	13	1			60		67.5		46	16	1			2		59.0	
16	16		1		600		56.8		47	12	1			342		62.5	
17	11	1			1		50.0		48	28			1	600		0.0	
18	13	1			7		68.3		49	16		1		600		52.9	
19	14	1			8		65.9		50	43			1	600		0.0	
20	13	1			19		68.3		51	15		1		600		63.4	
21	13	1			7		64.9		52	21		1		600		54.3	
22	13	1			8		67.5		53	21		1		600		54.3	
23	13	1			6		63.9		54	8	1			1		42.9	
24	14	1			18		63.2		55	19		1		600		52.5	
25	14	1			31		58.8		56	13	1			3		31.6	
26	14	1			12		61.1		57	10	1			1		41.2	
27	14	1			10		65.9		58	13	1			2		23.5	
28	14	1			14		69.6		59	12	1			3		29.4	
29	13	1			13		69.0		60	10	1			1		41.2	
30	13	1			19		66.7		61	12	1			3		33.3	
31	13	1			18		67.5		62	14	4	0	0	36	41.7	42.8	44.0

Table 4: Unconstrained (FS): DC – minimum cardinality; Status (s, f, a) – number of (FSDC") problems solved exactly, with feasible solution found or aborted; ΔDC – decomposition cardinality improvement in percent; t – total time in seconds.

		Status				$\Delta DC(\%)$					Status				$\Delta DC(\%)$		
#	DC	s	f	a	t (sec.)	min	μ	max	#	DC	s	f	a	t (sec.)	min	μ	max
1	8	1			8		60.0		32	14	1			36		63.2	
2	8	1			1		60.0		33	12	1			2		53.8	
3	8	1			7		52.9		34	13	1			7		62.9	
4	9	1			10		55.0		35	12	1			5		61.3	
5	12	1			52		47.8		36	12	1			2		53.8	
6	9	1			3		70.0		37	12	1			3		57.1	
7	9	1			0		52.6		38	12	1			18		57.1	
8	10	1			1		50.0		39	15	2	0	0	33	60.5	62.4	64.3
9	10	1			1		47.4		40	16		1		600		65.2	
10	10	1			0		47.4		41	19		1		600		58.7	
11	10	1			1		58.3		42	12	2	0	0	3	52.0	52.0	52.0
12	11	1			1		56.0		43	15	1			14		60.5	
13	9	1			1		59.1		44	15	1			51		61.5	
14	11	1			1		50.0		45	18	2	0	0	1149	59.1	59.5	60.0
15	13	1			83		66.7		46	18	1			11		58.1	
16	16		1		600		55.6		47	15		1		600		57.1	
17	11	1			0		50.0		48	35			1	600		0.0	
18	14	1			26		66.7		49	47			1	600		0.0	
19	14	1			15		64.1		50	44			1	600		0.0	
20	14	1			51		66.7		51	41			1	600		0.0	
21	14	1			22		64.1		52	50			1	600		0.0	
22	14	1			12		65.0		53	56			1	600		0.0	
23	14	1			10		62.2		54	8	1			0		42.9	
24	16	1			61		62.8		55	40			1	600		0.0	
25	14	1			21		61.1		56	13	1			17		31.6	
26	15		1		600		60.5		57	10	1			5		44.4	
27	14	1			19		65.9		58	13	1			12		23.5	
28	15	1			32		66.7		59	12	1			6		33.3	
29	15	1			23		65.1		60	14	1			3		26.3	
30	17	1			24		59.5		61	12	1			17		36.8	
31	15	1			53		63.4		62	13	10	0	0	648	45.8	47.9	50.0

Table 5: Constrained (FS): DC – minimum cardinality; Status (s, f, a) – number of (FSDC") problems solved exactly, with feasible solution found or aborted; ΔDC – decomposition cardinality improvement in percent; t – total time in seconds.

5 Conclusion

In this paper we discussed the realization problem in IMRT with objective functions total beam-on time and total number of shape matrices. In particular, we focussed on the usage of linear accelerators and multileaf collimators with limited width (*maximum leaf spread constraint*) which led us to the investigation of field splitting with feathering. We addressed unconstrained and constrained (*interleaf collision constraint*) versions of the problem and developed a new approach to determine the minimum beam-on time for both these cases. Furthermore, we proved the decomposition cardinality problem with field splitting to be $\mathcal{NP}$-hard even for a single row intensity matrix and without feathering. We then introduced a lexicographic approach that minimizes the decomposition cardinality subject to minimum beam-on time. The approaches presented in this article use integer programming formulations that we implemented to obtain numerical results for clinical as well as randomly generated examples. In future work we intend to address the problems discussed in this paper by means of heuristics. This alternative approach will help to produce at least feasible solutions for those instances of (FSDC) for which the exact methods presented here failed to produce such solutions within the fixed time limit.

Acknowledgments

This work is partially supported by the Australian Government, Department of Innovation, Industry, Science and Research, Collaborative Research Networks Program (grant CRN2011:07). Further, we would like to acknowledge the support by the European Union Seventh Framework Programme (FP7-PEOPLE-2009-IRSES) under Grant agreement 246647 and by the New Zealand Government (Grant no. 649378v2) as part of the Optimization and its Applications in Learning and Industry project.

References

R. Ahuja and H. Hamacher. A Network Flow Algorithm to Minimize Beam-On Time for Unconstrained Multileaf Collimator Problems in Cancer Radiation Therapy. *Networks*, 45: 36–41, 2004.

D. Baatar, H. Hamacher, M.Ehrgott, and G. Woeginger. Decomposition of integer matrices and multileaf collimator sequencing. *Discrete Applied Mathematics*, 152:6–34, 2005.

D. Chen, K. Engel, and C. Wang. A new algorithm for a field splitting problem in intensity modulated radiation therapy. *Algorithmica*, 61:656–673, 2011.

M. Collins, D. Kempe, J. Saia, and M. Young. Nonnegative integral subset representation of integer sets. *Information Processing Letters*, 101:129–133, 2007.

M. Ehrgott, C. Güler, H. Hamacher, and L. Shao. Mathematical optimization in intensity modulated radiation therapy. *4OR*, 6:199–262, 2008.

K. Engel. A new algorithm for optimal multileaf collimator field segmentation. *Discrete Applied Mathematics*, 152:35–51, 2005.

A. Ghouila-Houri. Caractérisation des Matrices Totalement Unimodulaires. *Comptes Rendues de l'Académie des Sciences Paris*, 254:1192–1194, 1962.

T. Kalinowski. A duality based algorithm for multileaf collimator field segmentation with interleaf collision constraint. *Discrete Applied Mathematics*, 152:52–88, 2005.

S. K. Kamath, S. Sahni, S. Ranka, J. Li, and J. Palta. Optimal field splitting for large intensity modulated fields. *Medical Physics*, 31:3314–3323, 2004.

G. J. Lim and E. K. Lee, editors. *Optimization in Medicine and Biology.* CRC Press, London, 2008.

Y. Liu and X. Wu. Minimizing Total Variation for Field Splitting with Feathering in Intensity-Modulated Radiation Therapy. In D. Lee, D. Chen, and S. Ying, editors, *Frontiers in Algorithmics*, volume 6213 of *Lecture Notes in Computer Science*, pages 65–76. Springer Verlag, Berlin, 2010.

N. Neumann. Approximative Netzwerkmodelle in der Bestrahlungstherapie. Diploma thesis, University of Kaiserslautern, 2009.

P. Pardalos and H. Romeijn, editors. *Handbook of Optimization in Medicine.* Springer Verlag, Berlin, 2009.

I. Raschendorfer. Field Splitting in IMRT: Decomposition Time and Cardinality Objectives. Diploma thesis, University of Kaiserslautern, 2011.

W. Schlegel and A. Mahr. *3D-Conformal Radiation Therapy: A Multimedia Introduction to Methods and Techniques*. Springer Verlag, Berlin, 2002.

Q. Wu, M. Arnfield, S. Tong, Y. Wu, and R. Mohan. Dynamic splitting of large intensity-modulated fields. *Physics in Medicine and Biology*, 45:1731–1740, 2000.