An algorithm for optimal multileaf collimator field segmentation with interleaf collision constraint

Thomas Kalinowski
Universität Rostock
Fachbereich Mathematik
D-18051 Rostock
Germany

February, 2003

Abstract

Intensity maps are nonnegative matrices describing the intensity modulation of beams in radiotherapy. An important step in the planning process is to determine a segmentation, that is a representation of an intensity map as a positive combination of special matrices corresponding to fixed positions of the multileaf collimator, called segments. We consider the problem of constructing segmentations with small total numbers of monitor units and segments. Generalizing the approach of [5] so that it applies to the segmentation problem with interleaf collision constraint, we show that the minimal number of monitor units in this case can be interpreted as the length of a longest path in a layered digraph. In addition we derive an efficient algorithm that constructs a segmentation with this minimal number of monitor units.

Key words: leaf sequencing, radiation therapy optimization, intensity modulation, multileaf collimator, IMRT

1 Introduction

The objective in radiotherapy planning for cancer treatment is to irradiate the tumor as efficient as possible without damaging the organs near to it. The first step is to determine an intensity function which describes the distribution of radiation over a rectangular target area. After discretization an intensity function can be considered as an $m \times n$ matrix A with nonnegative entries. One way of realizing such an intensity map is the usage of a multileaf collimator (MLC). An MLC has a pair of leaves for each row of the matrix, which can be shifted in horizontal direction and so open certain regions of the rectangle. By irradiating successively with different leaf-positions (called segments) it can be achieved that every region receives the amount of radiation that is prescribed by the intensity map. Due to technological restrictions in some of the currently used MLCs there is an additional condition for the possible segments: The interleaf collision constraint (ICC) excludes positions in which the left leaf of row i and the right leaf of row $i \pm 1$ overlap.
In this paper we consider the problem of determining the segments for a given matrix A in a good way. Two important objectives in this step are to minimize the total number of monitor units (TNMU) and the number of segments (NS). For the case of an MLC without ICC there are several segmentation algorithms (see [2–4, 6, 8–10]), some of them providing the optimal TNMU but a large NS, others reducing the NS at the price of an increased TNMU. In principle both, TNMU and NS, can be optimized by integer programming and this method can be adapted to additional restrictions like ICC (see [7]). But this approach is applicable only for small problem sizes. Another approach is the reformulation of the segmentation problem in a network flow setting. In [1] this is done for MLC-segmentation with ICC. In [5] there is presented an efficient segmentation algorithm which yields the optimal TNMU and a very small NS for the segmentation problem without ICC. The algorithm is derived from an explicit formula for the smallest possible TNMU. Here we generalize this approach to characterize the smallest possible TNMU with ICC as the maximal length of a path in a layered digraph.

2 A Linear Programming formulation

Throughout we use the notation $[n] := \{1, 2, \ldots, n\}$ for positive integers n. Let $A = (a_{i,j})_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$ be an $m \times n$-matrix with nonnegative integer entries. In addition we put

$$\begin{aligned} a_{i,0} &= a_{i,n+1} = 0, && i \in [m], \\ d_{i,j} &= a_{i,j} - a_{i,j-1}, && i \in [m],\ j \in [n+1]. \end{aligned}$$

A segment is a matrix that corresponds to a position of an MLC with interleaf collision constraint. This is made precise in the following definition.

Definition. A *segment* is an $m \times n$-matrix $S = (s_{i,j})$, such that there exist integers $l_i,\ r_i \quad (i \in [m])$ with the following properties:

$$l_i \leq r_i + 1 \qquad (i \in [m]), \tag{1}$$

$$s_{i,j} = \begin{cases} 1 & \text{if } l_i \leq j \leq r_i \\ 0 & \text{otherwise} \end{cases} \qquad (i \in [m],\ j \in [n]), \tag{2}$$

$$\text{ICC:} \quad l_i \leq r_{i+1} + 1,\ r_i \geq l_{i+1} - 1 \qquad (i \in [m-1]). \tag{3}$$

A segmentation of A is a representation of A as a sum of segments, i.e.

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

with segments S_i $(i = 1, 2, \ldots, k)$ and positive integers u_i $(i = 1, 2, \ldots, k)$. The segmentation problem is to find a segmentation of A with minimal value of $\sum_{i=1}^{k} u_i$. By $\mathcal{F}$ we denote the subsets of $V := [m] \times [n]$ that correspond to segments, that is

$$\mathcal{F} = \{T \subseteq V :\ \text{There exists a segment } S \text{ with } (i,j) \in T \iff s_{i,j} = 1\}.$$

The segmentation problem can be formulated as a linear program:

$$(P) \qquad \begin{array}{rll} \text{minimize} & \sum\limits_{T \in \mathcal{F}} f(T) & \text{subject to} \\ f(T) & \geq 0 & \forall T \in \mathcal{F}, \\ \sum\limits_{T \in \mathcal{F} : (i,j) \in T} f(T) & = a_{i,j} & \forall (i,j) \in V. \end{array}$$

The dual program is

$$(D) \qquad \begin{array}{rll} \text{maximize} & \sum\limits_{(i,j) \in V} a_{i,j} g(i,j) & \text{subject to} \\ \sum\limits_{(i,j) \in T} g(i,j) & \leq 1 & \forall T \in \mathcal{F}. \end{array}$$

To solve the segmentation problem we proceed in two steps: first we construct a feasible solution for the program (D) which yields a lower bound for the minimal TNMU in a segmentation, and in the second step we construct a sequence of segments that realizes this lower bound. We define a directed acyclic graph $\overrightarrow{G} = (V \cup \{0,1\}, E)$. For E we take all possible edges of the forms $(0,(i,1))$ and $((i,n),1)$, as well as all the edges between the $j-$th and the $(j+1)-$th column $(j = 1,2,\dots,n-1)$, precisely $E = E_1 \cup E_2 \cup E_3$, where

$$\begin{aligned}
E_1 &= \{(0,(i,1)) \ : \ i \in [m]\},\\
E_2 &= \{((i,n),1) \ : \ i \in [m]\},\\
E_3 &= \{((i,j),(i',j+1)) \ : \ i,i' \in [m], j \in [n-1]\}.
\end{aligned}$$

We define a length function δ (associated with A) on E by

$$\begin{aligned}
\delta(0,(i,1)) &= a_{i,1} && (i \in [m]),\\
\delta((i,n),1) &= 0 && (i \in [m]),\\
\delta((i,j),(i',j+1)) &= \max\{0, d_{i',j+1}\} - \sum_{k=i}^{i'-1} a_{k,j} \\
& \qquad\qquad (i,i' \in [m], i \le i', j \in [n-1]),\\
\delta((i,j),(i',j+1)) &= \max\{0, d_{i',j+1}\} - \sum_{k=i'+1}^{i} a_{k,j} \\
& \qquad\qquad (i,i' \in [m], i \ge i', j \in [n-1]).
\end{aligned}$$

For a path $P = (v_0, v_1, \dots, v_l)$ in $\overrightarrow{G}$ its length is

$$\delta(P) = \sum_{i=1}^{l} \delta(v_{i-1}, v_i).$$

The distance $D(v,w)$ of two vertices $v, w \in V \cup \{0,1\}$ is defined by

$$D(v,w) = \max\left\{\delta(P) \ : \ P \ (v,w) - \text{path in} \overrightarrow{G}\right\}.$$

Finally, we define the *complexity* $c(A)$ of the matrix A to be the distance of 0 and 1 in $\overrightarrow{G}$,

$$c(A) = D(0,1).$$

Now we are prepared to formulate our main result.

Theorem 1. *For every segmentation* $A = \sum_{i=1}^{k} u_i S_i$*, we have*

$$\sum_{i=1}^{k} u_i \geq c(A).$$

In addition, there exists a segmentation with $\sum_{i=1}^{k} u_i = c(A)$.

Note that the minimal TNMU in a segmentation without ICC can be interpreted analogously. In this case the minimal TNMU equals (see [5])

$$\max_{1\leq i\leq m} \sum_{j=1}^{n} \max\{0, d_{i,j}\},$$

that is the maximal length of a $(0,1)-$path in the graph that is obtained from $\overrightarrow{G}$ by deleting all the edges $((i,j),(i',j+1))$ with $i \neq i'$.

3 The lower bound

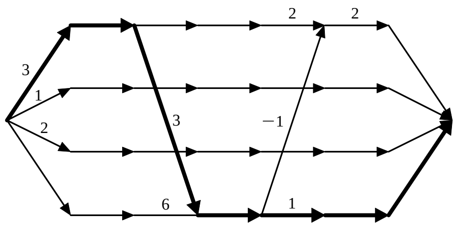
In this section we show how the $(0,1)-$paths in $\overrightarrow{G}$ correspond to certain feasible solutions for the program (D) and derive the lower bound of the theorem. For a path $P = (0,(i_1,1),(i_2,2),\ldots,(i_n,n),1)$ in $\overrightarrow{G}$ we put $i_{n+1} = i_n$ and define a function $g : V \to \{0,1,-1\}$ as follows:

$$g(i,j) = \begin{cases} 1 & \text{if } i = i_j = i_{j+1},\ d_{i,j} \geq 0,\ d_{i,j+1} < 0, \\ 1 & \text{if } i = i_j = i_{j+1} \neq i_{j+2},\ d_{i,j} \geq 0, \\ -1 & \text{if } i = i_j = i_{j+1} = i_{j+2},\ d_{i,j} < 0 \text{ and } d_{i,j+1} \geq 0, \\ -1 & \text{if } i_j \leq i < i_{j+1} \text{ or } i_{j+1} < i \leq i_j, \\ -1 & \text{if } i = i_{j+1} \neq i_j \text{ and } d_{i,j+1} \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Example 1. Fig. 1 shows a path of length 7 with respect to the matrix

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 & 2 & 4 \\ 1 & 1 & 1 & 2 & 3 & 3 \\ 2 & 2 & 2 & 1 & 1 & 1 \\ 0 & 0 & 6 & 0 & 1 & 1 \end{pmatrix}$$

and the corresponding function g.



1	−1				
	−1				
	−1				
	−1	1	−1		1

Figure 1: A path of length 7 with respect to A and the corresponding function g

We want to prove that for every $(0,1)-$path P the corresponding function g is a feasible solution for the program (D). Thus we have to show that, for every $T \in \mathcal{F}$,

$$\sum_{(i,j)\in T} g(i,j) \leq 1.$$

A $(0,1)-$path P is uniquely determined by the indices of the colums in which P changes the row and the indices of the rows in which P runs between the row changes. So let $x_1, x_2, \ldots, x_{k-1}$ with $0 < x_1 < x_2 < \cdots < x_{k-1} < n$ denote the indices of the columns where P changes the row, i.e.

$$(i, x_t), (i', x_t + 1) \in P \quad \text{with} \quad i \neq i' \quad (t \in [k-1]).$$

In addition let i_t^* be the row index with $(i_t^*, x_t) \in P$ $(t = 1, 2, \ldots, k-1)$ and i_k^* the index with $(i_k^*, n) \in P$. Finally, we put $x_0 = 0$, and $x_k = n + 1$. Thus

$$P = (0, (i_1^*, 1), (i_1^*, 2), \ldots, (i_1^*, x_1), (i_2^*, x_1 + 1), \ldots, (i_2^*, x_2), \ldots, \ldots, (i_k^*, n), 1),$$

and P is uniquely determined by its parameters $(i_1^*, x_1), \ldots, (i_k^*, x_k)$.

Lemma 2. *Let P be a $(0,1)-$path with parameters $(i_1^*, x_1), \ldots, (i_k^*, x_k)$, and let g be defined according to* (4)*. In addition let $1 \leq l \leq r+1 \leq n+1$. Then, for every $i \in [m]$,*

$$\sum_{j=l}^{r} g(i,j) \leq 1,$$

and equality implies $x_{t-1} < l \leq r < x_t$ for some $t \in [k]$ with $i_t^* = i$.

Proof. We choose an arbitrary $i \in [m]$ and put

$$J(i) := \{j \in [n] \ : \ g(i,j) \neq 0\}.$$

We denote the elements of $J(i)$ by $j_1, j_2, \ldots, j_p$ $(j_1 < j_2 < \cdots < j_p)$. Then the following claims follow immediately from the definition of g.

1. If $i = i_1^*$ and $k > 1$ then $x_1 \in J(i)$ and the sequence

$$g(i, j_1), g(i, j_2), \ldots, g(i, x_1)$$

is an alternating $(1, -1)-$sequence ending with -1.

2. If $i = i_k^*$, $k > 1$ and $J(i) \cap \{x_{k-1}, x_{k-1} + 1, \ldots, n\} \neq \emptyset$ then for

$$q = \min\{\tau \ : \ j_\tau \geq x_{k-1}\},$$

$g(i, j_q), g(i, j_{q+1}), \ldots, g(i, j_p)$ is an alternating $(1, -1)-$sequence starting with -1 and ending with 1.

3. If $i = i_1^*$ and $k = 1$ then the sequence $g(i, j_1), g(i, j_2), \ldots, g(i, j_p)$ is empty or an alternating $(1, -1)-$sequence starting and ending with 1.

4. If $i = i_t^*$ for $2 \leq t \leq k - 1$ then $x_t \in J(i)$ and for

$$q = \min\{\tau \ : \ j_\tau \geq x_{t-1}\},$$

$g(i, j_q), g(i, j_{q+1}), \ldots, g(i, x_t)$ is an alternating $(1, -1)-$sequence starting and ending with -1.

5. If $j \in J(i)$ and (i, j) does not correspond to a term in one of the sequences described in the first 4 cases then $j = x_t$ for some $t \in [k - 1]$ with $i \neq i_t^*$ and $i \neq i_{t+1}^*$ and $g(i, j) = -1$.

Consequently we obtain, for $1 \leq q \leq p - 1$,

$$g(i, j_q) = 1 \ \Rightarrow \ g(i, j_{q+1}) = -1.$$

Now the first part of the lemma follows from

$$\sum_{j=l}^{r} g(i, j) = \sum_{\tau=q}^{q'} g(i, j_\tau) \quad \text{for some } q, q' \in [p].$$

Suppose

$$\sum_{j=l}^{r} g(i, j) = \sum_{\tau=q}^{q'} g(i, j_\tau) = 1.$$

Then the sequence $g(i, j_q), g(i, j_{q+1}), \ldots, g(i, j_{q'})$ has to be an alternating $(1, -1)-$sequence starting and ending with 1. This implies

$$x_{t-1} < l < x_t \quad \text{and} \quad x_{t'-1} < r < x_{t'}$$

for some $t, t' \in [k]$ with $i = i_t^* = i_{t'}^*$. Assume $t \neq t'$ and put

$$t'' = \min\{\sigma > t \ : \ i_\sigma^* = i\} \quad \text{and} \quad q'' = \min\{\tau \ : \ j_\tau \geq x_{t''-1}\}.$$

Then

$$j_q < x_t < j_{q''} < j_{q'}, \quad g(i, x_t) = g(i, j_{q''}) = -1,$$

and $g(i,j) \leq 0$ for all j with $x_t < j < j_{q''}$. So $g(i, j_q), g(i, j_{q+1}), \dots, g(i, j_{q'})$ contains two consecutive $(-1)-$terms, which is a contradiction. Hence $t = t'$ and the second part of the lemma follows. □

The next lemma gives a condition that must hold if the sum of the $g(i,j)$ over a row of a segment vanishes. (By a row of a segment we mean the part of the row that is left open by the MLC in the corresponding leaf position.)

Lemma 3. *Let P be a $(0,1)-$path with parameters $(i_1^*, x_1), \dots, (i_k^*, x_k)$, and let g be defined according to* (4)*. Assume $i \in [m]$, $1 \leq l \leq r+1 \leq n+1$ and*

$$\sum_{j=l}^{r} g(i,j) = 0.$$

Suppose in addition that for $t \in [k-1]$ one of the following conditions holds

1. *$i_t^* < i < i_{t+1}^*$ and $l \leq x_t$*
2. *$t \geq 2$, $i_t^* = i < i_{t+1}^*$ and $l \leq x_{t-1}$*
3. *$i_t^* > i > i_{t+1}^*$ and $l \leq x_t$ for some $t \in [k-1]$*
4. *$t \geq 2$, $i_t^* = i > i_{t+1}^*$ and $l \leq x_{t-1}$*

Then $r < x_t$.

Proof. We consider only the first two cases that are illustrated in Fig. 2. The other two are treated analogously. Assume $r \geq x_t$. In order to derive a contradiction we use the following observation several times. If P leaves row i in (i,j) then $g(i,j) = -1$, and if P enters row i' in (i', j'), $j' > 1$, then either $g(i', j'-1) = -1$ or the first nonvanishing $g(i', j'')$ we meet on the subpath starting with (i', j') equals -1. We put

$$J = \{j \ : \ l \leq j \leq r, \ g(i,j) \neq 0\},$$

and denote the elements of J by $j_1, j_2, \dots, j_p$ $(j_1 < j_2 < \dots < j_p)$. In particular $j_q = x_t$ for some $q \in [p]$.

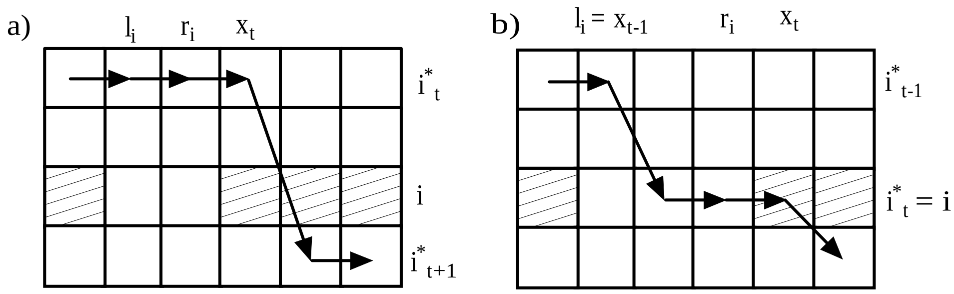


Figure 2: Illustration of Lemma 3. a) $i_t^* < i < i_{t+1}^*$ and $l \leq x_t$, b) $t \geq 2$, $i_t^* = i < i_{t+1}^*$ and $l \leq x_{t-1}$.

Case 1: $g(i, j_1) = -1$.

By assumption $g(i, j_1), \ldots, g(i, j_p)$ is an alternating $(1, -1)-$sequence starting with -1 and ending with 1. From $g(i, x_t) = -1$ follows $q < p$ and by construction of g the contradiction

$$g(i, j_q) = g(i, j_{q+1}) = -1.$$

Case 2: $g(i, j_1) = 1$.

This implies $l < x_{t'}$ for some $t' < t$ with $i_{t'}^* = i$, and consequently $j_{q'} = x_{t'}$ for some $q' \in [p-1]$, $q' < q$. Thus

$$g(i, j_{q'}) = g(i, j_{q'+1}) = -1,$$

and $g(i, j_1), g(i, j_2), \ldots, g(i, j_p)$ contains two consecutive $(-1)-$terms. By assumption this implies $g(i, j_p) = 1$, hence $p > q$, and by construction of g,

$$g(i, j_q) = g(i, j_{q+1}) = -1.$$

But now $g(i, j_1), g(i, j_2), \ldots, g(i, j_p)$ contains two pairs of consecutive $(-1)-$terms (if $q'+1 < q$) or three consecutive (-1)-terms (if $q'+1 = q$). Again this yields a contradiction.

□

The following lemma is the crucial step in the proof of the feasibility of g. We show that the ICC implies that in any segment, between two rows in which the values of g add up to 1 there is a row in which this sum is at most -1.

Lemma 4. *Let P be a $(0,1)-$path with parameters $(i_1^*, x_1), \ldots, (i_k^*, x_k)$, and let g be defined according to (4). Suppose S is a segment described by $l_1, l_2, \ldots, l_m, r_1, r_2, \ldots, r_m$ and there are row indices i_0, i_1 ($1 \leq i_0 < i_1 \leq m$) such that*

$$\sum_{j=l_{i_0}}^{r_{i_0}} g(i_0, j) = 1 \quad \textit{and} \quad \sum_{j=l_{i_1}}^{r_{i_1}} g(i_1, j) = 1.$$

Then there exists a row index i with $i_0 < i < i_1$ and $\sum\limits_{j=l_i}^{r_i} g(i,j) \leq -1$.

Proof. W.l.o.g. we may assume that there is no row i with $i_0 < i < i_1$ and

$$\sum_{j=l_i}^{r_i} g(i,j) = 1.$$

Suppose that for all i with $i_0 < i < i_1$, $\sum\limits_{j=l_i}^{r_i} g(i,j) = 0$. By Lemma 2 there are $t, t' \in [k]$ such that

$$\begin{aligned} x_{t-1} < l_{i_0} \leq r_{i_0} < x_t, \quad & i_t^* = i_0 \quad \text{and} \\ x_{t'-1} < l_{i_1} \leq r_{i_1} < x_{t'}, \quad & i_{t'}^* = i_1. \end{aligned}$$

W.l.o.g. we may assume $t < t'$. Now let $i_0 = z_0 < z_1 < \ldots < z_p = i_1$ be an increasing sequence of row indices such that there is a corresponding sequence $t = t_0 < t_1 < \ldots < t_p \leq t'$ with $i_{t_q}^* = z_q$ ($0 \leq q \leq p$) and in addition for $0 \leq q \leq p-1$ there is no τ with $t_q < \tau < t_{q+1}$ and $z_q < i_\tau^* \leq z_{q+1}$. Precisely, we put

$$t_0 = t \quad \text{and} \quad z_0 = i_0,$$

and for $q \geq 1$ and $z_{q-1} < i_1$,

$$t_q = \min\{\tau \ : \ i_\tau^* > z_{q-1}\} \quad \text{and} \quad z_q = i_{t_q}^*.$$

So for some q we obtain $z_q = i_1$, and then we put $p = q$ (see Fig. 3).

Claim 1: For $0 \leq q \leq p-1$,

$$r_{z_q} < x_{t_q} \ \Rightarrow \ r_i < x_{t_{q+1}-1} \text{ for all } i \text{ with } z_q \leq i < z_{q+1}.$$

Claim 2: For $0 \leq q \leq p-1$ we have $r_{z_q} < x_{t_q}$.

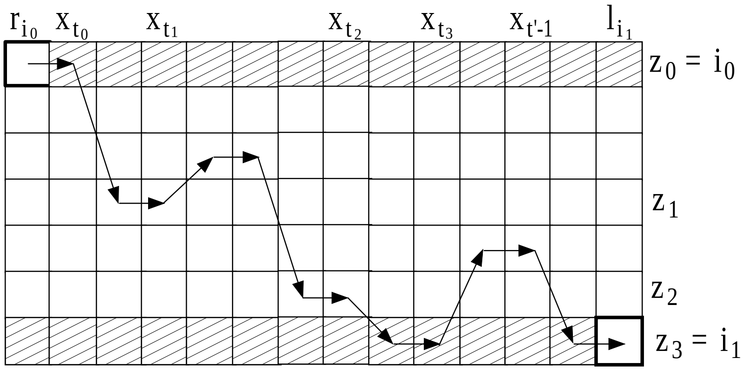


Figure 3: Situation in the proof of Lemma 4 with $p = 3$.

Proof of Claim 1. We proceed by induction on i. By assumption

$$r_{z_q} < x_{t_q} \leq x_{t_{q+1}-1}.$$

So let $z_q < i < z_{q+1}$ and assume $r_{i-1} < x_{t_{q+1}-1}$. The ICC implies $l_i \leq x_{t_{q+1}-1}$ and by Lemma 3 we obtain

$$r_i < x_{t_{q+1}-1}.$$

Proof of Claim 2. Here we use induction on q. Clearly,

$$r_{z_0} = r_{i_0} < x_t = x_{t_0}.$$

So let $q > 0$ and assume by induction $r_{z_{q-1}} < x_{t_{q-1}}$. Then by claim 1,

$$r_{z_q-1} < x_{t_q-1}.$$

Thus $l_{z_q} \leq x_{t_q-1}$, and hence, again by Lemma 3, $r_{z_q} < x_{t_q}$.
Combining claims 1 and 2 we obtain

$$r_{i_1-1} < x_{t_p-1} \leq x_{t'-1} < l_{i_1},$$

thus $r_{i_1-1} < l_{i_1} - 1$ in contradiction to the ICC. □

Lemma 5. *Let P be a $(0,1)-$path with parameters $(i_1^*, x_1), \ldots, (i_k^*, x_k)$, and let g be defined according to* (4). *Then g is feasible for* (D).

Proof. Let $T \in \mathcal{F}$ be arbitrary and let S be the corresponding segment with parameters l_i, r_i $(i \in [m])$. Then

$$\sum_{(i,j)\in T} g(i,j) = \sum_{i=1}^{m} \sum_{j=l_i}^{r_i} g(i,j).$$

By Lemma 2, for all $i \in [m]$, $\sum\limits_{j=l_i}^{r_i} g(i,j) \leq 1$, and by Lemma 4 between two rows i and i'' with $i < i''$ and

$$\sum_{j=l_i}^{r_i} g(i,j) = \sum_{j=l_{i''}}^{r_{i''}} g(i'',j) = 1$$

there is always a row i' with $i < i' < i''$ and $\sum\limits_{j=l_{i'}}^{r_{i'}} g(i',j) \leq -1$. Consequently,

$$\sum_{(i,j)\in T} g(i,j) \leq 1,$$

that is the feasibility of g. □

Lemma 6. *Let P be a $(0,1)-$path with parameters $(i_1^*, x_1), \dots, (i_k^*, x_k)$, and let g be defined according to* (4)*. Then*

$$\sum_{(i,j)\in V} g(i,j)a_{ij} = \sum_{t=1}^{k} \sum_{j=x_{t-1}+1}^{x_t-1} \max\{0, d_{i_t^*,j}\} - \sum_{t=1}^{k-1} \left(\sum_{i=i_t^*}^{i_{t+1}^*-1} a_{i,x_t} + \sum_{i=i_{t+1}^*+1}^{i_t^*} a_{i,x_t} \right).$$

Proof. In terms of the i_t^* and x_t the definition of g is

$$g(i,j) = \begin{cases} 1 & \text{if } x_{t-1} < j < x_t,\ i = i_t^*,\ d_{i,j} \geq 0,\ d_{i,j+1} < 0, \\ 1 & \text{if } x_{t-1} < j = x_t - 1,\ i = i_t^*,\ d_{i,j} \geq 0, \\ -1 & \text{if } x_{t-1} < j < x_t - 1,\ i = i_t^*,\ d_{i,j} < 0,\ d_{i,j+1} \geq 0, \\ -1 & \text{if } j = x_t,\ i_t^* \leq i < i_{t+1}^* \text{ or } i_{t+1}^* < i \leq i_t^*, \\ -1 & \text{if } j = x_t,\ i = i_{t+1}^*,\ d_{i,j+1} \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

This yields immediately

$$\sum_{j=x_{t-1}}^{x_t-1} g(i_t^*, j)a_{i_t^*,j} = \sum_{j=x_{t-1}+1}^{x_t-1} \max\{0, d_{i_t^*,j}\} \quad (t = 1, 2, \dots, k).$$

The remaining nonzero $g(i,j)$ correspond to the row changes of P, and we have to add for $t = 1, 2, \dots, k-1$,

$$\sum_{i=i_t^*}^{i_{t+1}^*-1} g(i, x_t)a_{i,x_t} = -\sum_{i=i_t^*}^{i_{t+1}^*-1} a_{i,x_t} \qquad \text{if } i_t^* < i_{t+1}^* \quad \text{and}$$

$$\sum_{i=i_{t+1}^*+1}^{i_t^*} g(i, x_t)a_{i,x_t} = -\sum_{i=i_{t+1}^*+1}^{i_t^*} a_{i,x_t} \qquad \text{if } i_t^* > i_{t+1}^*.$$

□

For the length of P to be equal to the value of the program (D) for the corresponding g we need an additional restriction on P. We call the $(0,1)-$path P with parameters $(i_1^*, x_1), \ldots, (i_k^*, x_k)$ feasible (with respect to A) if $d_{i_t^*, x_t} < 0$ for $t = 1, 2, \ldots, k-1$, which in particular implies that the last edges of the horizontal parts of P have length 0.

Lemma 7. *Let P be a feasible $(0,1)-$path and g defined according to (4). Then*

$$\sum_{(i,j)\in V} g(i,j) a_{i,j} = \delta(P).$$

Proof. Let P be given by parameters $(i_1^*, x_1), \ldots, (i_k^*, x_k)$. For $t \in [k]$ we denote by P_t the subpath from $(i_t^*, x_{t-1}+1)$ to (i_t^*, x_t). Thus

$$\delta(P) = \sum_{t=1}^{k} \delta(P_t) + \delta(0, (i_1^*, 1)) + \sum_{t=1}^{k-1} \delta((i_t^*, x_t), (i_{t+1}^*, x_t + 1)).$$

From the feasibility of P follows that the last edge of P_t has length 0 for all $t \in [k]$, and we obtain

$$\delta(P_t) = \sum_{j=x_{t-1}+2}^{x_t - 1} \max\{0, d_{i_t^*, j}\} \quad (t \in [k]).$$

In addition, $\delta(0, (i_1^*, 1)) = a_{i_1^*, 1} = \max\{0, d_{i_1^*, 1}\}$, and for $t \in [k-1]$,

$$\delta((i_t^*, x_t), (i_{t+1}^*, x_t + 1)) = \max\{0, d_{i_{t+1}^*, x_t + 1}\} - \sum_{i=i_t^*}^{i_{t+1}^* - 1} a_{i, x_t} - \sum_{i=i_{t+1}^*+1}^{i_t^*} a_{i, x_t}.$$

Thus

$$\delta(P) = \sum_{t=1}^{k} \sum_{j=x_{t-1}+1}^{x_t - 1} \max\{0, d_{i_t^*, j}\} - \sum_{t=1}^{k-1} \left(\sum_{i=i_t^*}^{i_{t+1}^* - 1} a_{i, x_t} + \sum_{i=i_{t+1}^*+1}^{i_t^*} a_{i, x_t} \right),$$

and the claim follows by Lemma 6. $\square$

Lemma 8. *There exists a feasible $(0,1)-$path P with $\delta(P) = c(A)$.*

Proof. For any $(0,1)-$path P with parameters $(i_1^*, x_1), \ldots, (i_k^*, x_k)$ denote by $R(P) \subseteq [k-1]$ the subset of indices that destroy the feasibility of P, that is

$$R(P) = \{t \in [k-1] :\ d_{i_t^*, x_t} \geq 0\}.$$

Then

$$\lambda(P) = \sum_{t \in R(P)} \left| i_t^* - i_{t+1}^* \right|$$

measures how far P is from being feasible. Let P_0 be a $(0,1)-$path with parameters $(i_1^*, x_1), \ldots, (i_k^*, x_k)$ and length $\delta(P_0) = c(A)$. If $\lambda(P_0) = 0$ then P_0 is feasible and there is nothing to do. So we assume that for $r \geq 1$ we have a $(0,1)-$path P_{r-1} with parameters

$$(i_1^*, x_1), \ldots, (i_k^*, x_k),$$

$\delta(P_{r-1}) = c(A)$ and $\lambda(P_{r-1}) > 0$. From this we construct a $(0,1)-$path P_r with $\delta(P_r) = c(A)$ and $\lambda(P_r) \leq \lambda(P_{r-1}) - 1$. This will prove the lemma, since after finitely many steps we obtain a path P with $\delta(P) = c(A)$ and $\lambda(P) = 0$. Let t be the smallest element of $R(P_{r-1})$.

Case 1: $d_{i_t^*,j} \geq 0$ for $x_{t-1} < j < x_t$.

We define P_r as follows.

1. If $i_t^* < i_{t+1}^* - 1$ and $i_{t-1}^* \neq i_t^* + 1$ the parameters of P_r are (see Fig. 4 and 6)

$$(i_1^*, x_1), \ldots, (i_{t-1}^*, x_{t-1}), (i_t^* + 1, x_t), (i_{t+1}^*, x_{t+1}), \ldots, (i_k^*, x_k).$$

2. If $i_t^* > i_{t+1}^* + 1$ and $i_{t-1}^* \neq i_t^* - 1$ the parameters of P_r are (see Fig. 9 and 11)

$$(i_1^*, x_1), \ldots, (i_{t-1}^*, x_{t-1}), (i_t^* - 1, x_t), (i_{t+1}^*, x_{t+1}), \ldots, (i_k^*, x_k).$$

3. If $i_{t-1}^* - 1 = i_t^* < i_{t+1}^* - 1$ or $i_{t-1}^* + 1 = i_t^* > i_{t+1}^* + 1$ the parameters of P_r are (see Fig. 7 and 12)

$$(i_1^*, x_1), \ldots, (i_{t-2}^*, x_{t-2}), (i_{t-1}^*, x_t), (i_{t+1}^*, x_{t+1}), \ldots, (i_k^*, x_k).$$

4. If $i_t^* + 1 = i_{t+1}^* \neq i_{t-1}^*$ or $i_t^* - 1 = i_{t+1}^* \neq i_{t-1}^*$ the parameters of P_r are (see Fig. 5 and 10)

$$(i_1^*, x_1), \ldots, (i_{t-1}^*, x_{t-1}), (i_{t+1}^*, x_{t+1}), \ldots, (i_k^*, x_k).$$

5. If $i_t^* + 1 = i_{t+1}^* = i_{t-1}^*$ or $i_t^* - 1 = i_{t+1}^* = i_{t-1}^*$ the parameters of P_r are (see Fig. 8 and 13)

$$(i_1^*, x_1), \ldots, (i_{t-2}^*, x_{t-2}), (i_{t+1}^*, x_{t+1}), \ldots, (i_k^*, x_k).$$

Case 2: $d_{i_t^*,j} < 0$ for some j with $x_{t-1} < j < x_t$.

We put

$$x := \max\{j \le x_t \ : \ d_{i_t^*,j} < 0,\ d_{i_t^*,j+1} \ge 0\},$$

and define P_r as follows.

1. If $i_t^* < i_{t+1}^* - 1$ the parameters of P_r are (see Fig. 14)

$$(i_1^*, x_1), \dots, (i_{t-1}^*, x_{t-1}), (i_t^*, x), (i_t^* + 1, x_t), (i_{t+1}^*, x_{t+1}), \dots, (i_k^*, x_k).$$

2. If $i_t^* > i_{t+1}^* + 1$ the parameters of P_r are (see Fig. 16)

$$(i_1^*, x_1), \dots, (i_{t-1}^*, x_{t-1}), (i_t^*, x), (i_t^* - 1, x_t), (i_{t+1}^*, x_{t+1}), \dots, (i_k^*, x_k).$$

3. If $i_t^* = i_{t+1}^* - 1$ or $i_t^* = i_{t+1}^* + 1$ the parameters of P_r are (see Fig. 15 and 17)

$$(i_1^*, x_1), \dots, (i_{t-1}^*, x_{t-1}), (i_t^*, x), (i_{t+1}^*, x_{t+1}), \dots, (i_k^*, x_k).$$

We have to show that $\delta(P_r) = c(A)$ and $\lambda(P_r) \le \lambda(P_{r-1}) - 1$. The last assertion follows from the fact that either

$$R(P_r) = R(P_{r-1}) \quad \text{or} \quad R(P_r) = R(P_{r-1}) \setminus \{x_t\},$$

and consequently,

$$\lambda(P_r) = \lambda(P_{r-1}) - 1 \quad \text{or} \quad \lambda(P_r) = \lambda(P_{r-1}) - \left| i_t^* - i_{t+1}^* \right|.$$

Now we check that in any case $\delta(P_r) \ge \delta(P_{r-1})$ and hence $\delta(P_r) = c(A)$. In the following let the vertices of $\overrightarrow{G}$ be denoted as in the corresponding figures. In addition for two vertices X and Y on a path P we denote by $D_P(X,Y)$ the length of the (X,Y)−subpath of P. Then in any case,

$$\begin{aligned} \delta(P_r) = \delta(P_{r-1}) - D_{P_{r-1}}(U,A) - D_{P_{r-1}}(A,B) - D_{P_{r-1}}(B,V) \\ + D_{P_r}(U,A') + D_{P_r}(A',B') + D_{P_r}(B',V). \end{aligned} \tag{5}$$

Cases 1.1(a), 1.4(a): (Fig. 4, 5)

Using $d_{i_t^*,j} \ge 0$ for $x_{t-1} < j \le x_t$ we obtain

$$\begin{aligned} D_{P_{r-1}}(A,B) &= a_{i_t^*,x_t} - a_{i_t^*,x_{t-1}+1}, \\ D_{P_r}(B',V) &= D_{P_{r-1}}(B,V) + a_{i_t^*,x_t}, \\ D_{P_r}(U,A') &= D_{P_{r-1}}(U,A) - \left(a_{i_t^*,x_{t-1}+1} - a_{i_t^*,x_{t-1}}\right) \\ &\qquad - a_{i_t^*,x_{t-1}} + \max\{0, d_{i_t^*+1,x_{t-1}+1}\} \\ &= D_{P_{r-1}}(U,A) - a_{i_t^*,x_{t-1}+1} + \max\{0, d_{i_t^*+1,x_{t-1}+1}\}. \end{aligned}$$

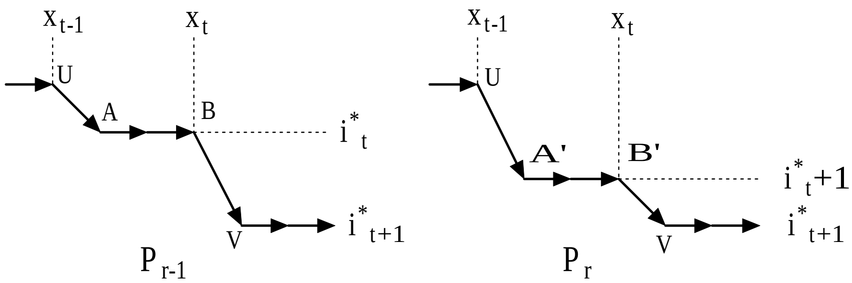


Figure 4: Transition from P_{r-1} to P_r in Case 1.1.a) $i^*_{t-1} < i^*_t$.

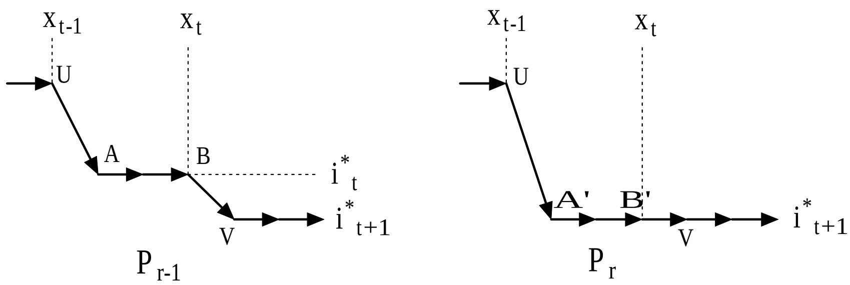


Figure 5: Transition from P_{r-1} to P_r in Case 1.4.a) $i^*_{t-1} < i^*_t$.

Substituting into (5) yields

$$\begin{aligned}\delta(P_r) = \delta(P_{r-1}) - D_{P_{r-1}}(U,A) - \big(a_{i^*_t,x_t} - a_{i^*_t,x_{t-1}+1}\big) - D_{P_{r-1}}(B,V) \\ + \big(D_{P_{r-1}}(U,A) - a_{i^*_t,x_{t-1}+1} + \max\{0, d_{i^*_t+1,x_{t-1}+1}\}\big) \\ + D_{P_r}(A',B') + \big(D_{P_{r-1}}(B,V) + a_{i^*_t,x_t}\big)\,,\end{aligned}$$

that is

$$\begin{aligned}\delta(P_r) &= \delta(P_{r-1}) + \max\{0, d_{i^*_t+1,x_{t-1}+1}\} + D_{P_r}(A',B') \\ &\geq \delta(P_{r-1}).\end{aligned}$$

Cases 1.1(b), 1.3(a), 1.5(a): (Fig. 6, 7, 8)

Again,

$$\begin{aligned}D_{P_{r-1}}(A,B) &= a_{i^*_t,x_t} - a_{i^*_t,x_{t-1}+1}, \\ D_{P_r}(B',V) &= D_{P_{r-1}}(B,V) + a_{i^*_t,x_t}.\end{aligned}$$

But in these cases

$$\begin{aligned}D_{P_r}(U,A') = D_{P_{r-1}}(U,A) - \big(a_{i^*_t,x_{t-1}+1} - a_{i^*_t,x_{t-1}}\big) \\ + a_{i^*_t+1,x_{t-1}} + \max\{0, d_{i^*_t+1,x_{t-1}+1}\}.\end{aligned}$$

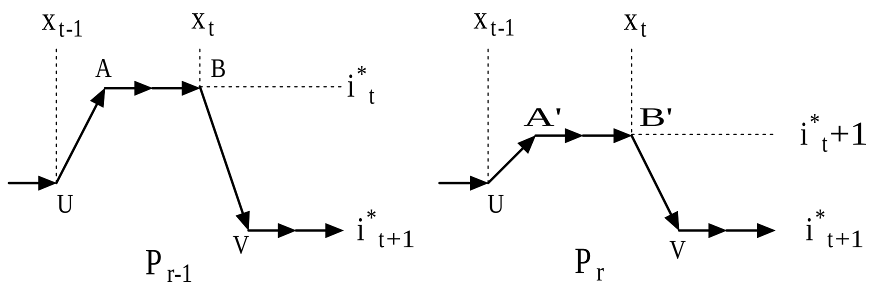


Figure 6: Transition from P_{r-1} to P_r in Case 1.1.b) $i^*_{t-1} > i^*_t$.

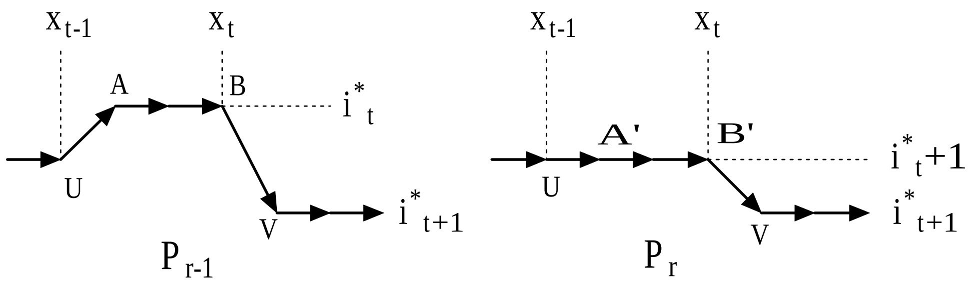


Figure 7: Transition from P_{r-1} to P_r in Case 1.3.a) $i^*_{t-1} > i^*_t$.

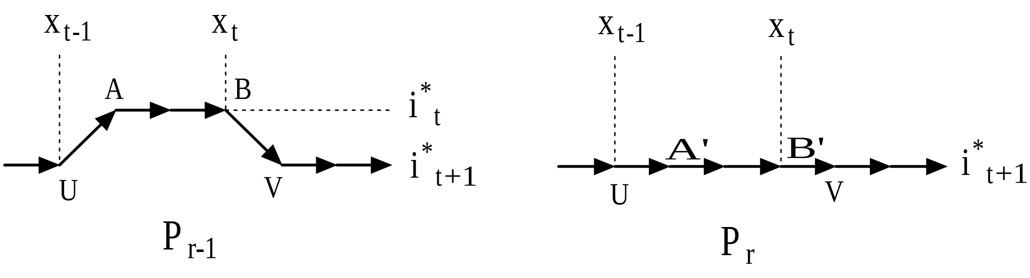


Figure 8: Transition from P_{r-1} to P_r in Case 1.5.a) $i^*_{t-1} > i^*_t$.

And substituting into (5) yields

$$\delta(P_r) = \delta(P_{r-1}) - D_{P_{r-1}}(U,A) - \left(a_{i_t^*,x_t} - a_{i_t^*,x_{t-1}+1}\right) - D_{P_{r-1}}(B,V) \\ + \left[D_{P_{r-1}}(U,A) - \left(a_{i_t^*,x_{t-1}+1} - a_{i_t^*,x_{t-1}}\right) + a_{i_t^*+1,x_{t-1}} + \right. \\ \left. \max\{0, d_{i_t^*+1,x_{t-1}+1}\}\right] + D_{P_r}(A',B') + \left(D_{P_{r-1}}(B,V) + a_{i_t^*,x_t}\right),$$

that is

$$\delta(P_r) = \delta(P_{r-1}) + a_{i_t^*,x_{t-1}} + a_{i_t^*+1,x_{t-1}} + \max\{0, d_{i_t^*+1,x_{t-1}+1}\} + D_{P_r}(A',B') \\ \geq \delta(P_{r-1}).$$

Cases 1.2(a), 1.4(b): (Fig. 9, 10)

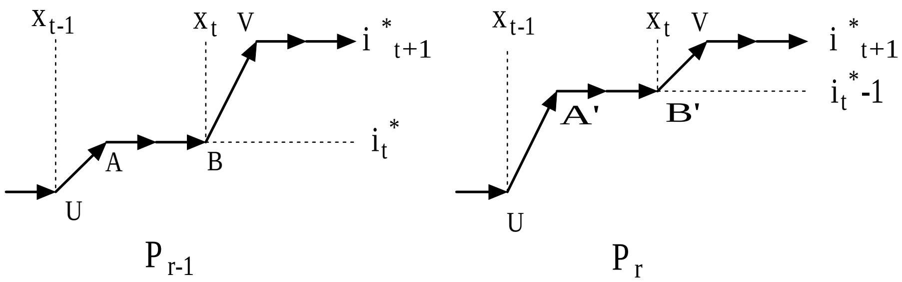


Figure 9: Transition from P_{r-1} to P_r in Case 1.2.a) $i_{t-1}^* > i_t^*$.

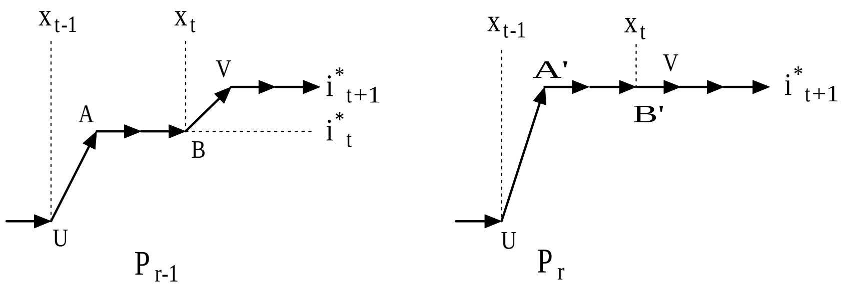


Figure 10: Transition from P_{r-1} to P_r in Case 1.4.b) $i_{t-1}^* > i_t^*$.

The computation is the same as in Case 1.1(a) but in the formula for $D_{P_r}(U,A')$ we have to replace $d_{i_t^*+1,x_{t-1}+1}$ by $d_{i_t^*-1,x_{t-1}+1}$.

Cases 1.2(b), 1.3(b), 1.5(b): (Fig. 11, 12, 13)

The computation is the same as in Case 1.1(b) but in the formula for $D_{P_r}(U,A')$ we have to replace $d_{i_t^*+1,x_{t-1}+1}$ by $d_{i_t^*-1,x_{t-1}+1}$.

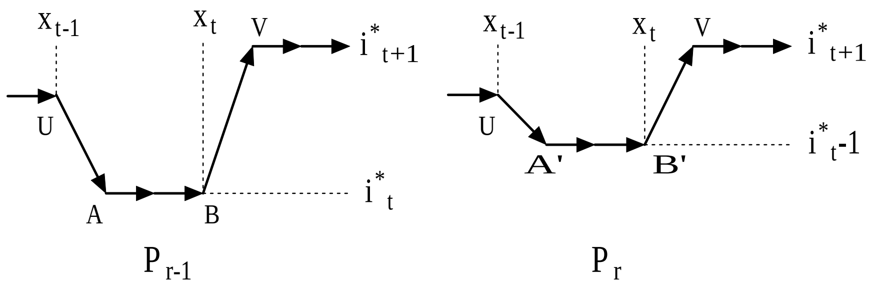


Figure 11: Transition from P_{r-1} to P_r in Case 1.2.b) $i^*_{t-1} < i^*_t$.

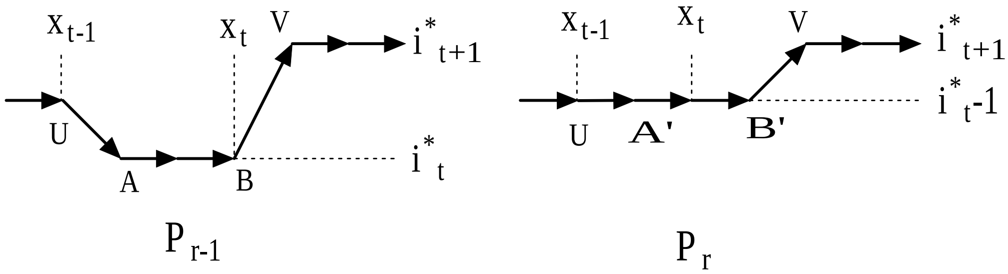


Figure 12: Transition from P_{r-1} to P_r in Case 1.3.b) $i^*_{t-1} < i^*_t$.

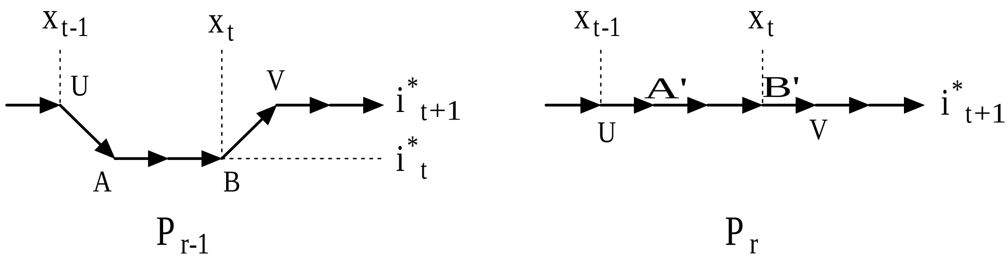


Figure 13: Transition from P_{r-1} to P_r in Case 1.5.b) $i^*_{t-1} < i^*_t$.

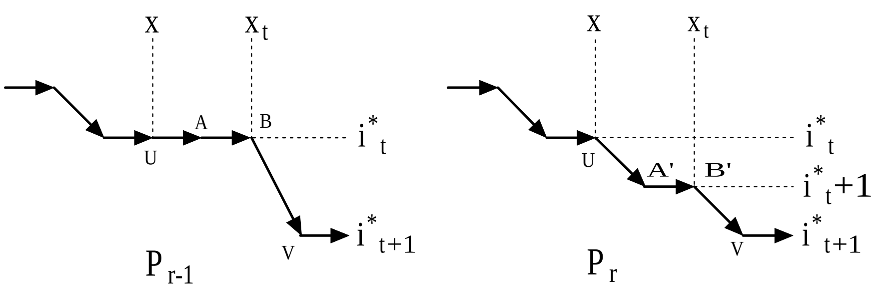


Figure 14: Transition from P_{r-1} to P_r in Case 2.1.

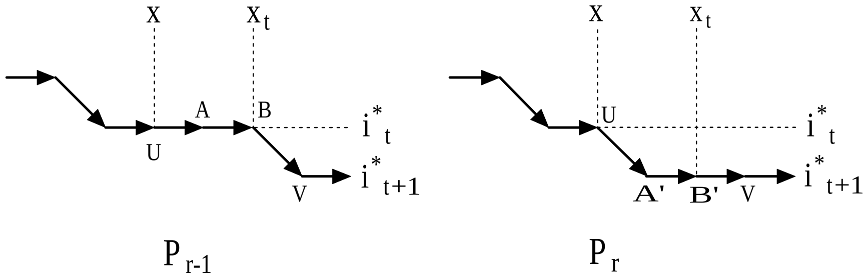


Figure 15: Transition from P_{r-1} to P_r in Case 2.3.a) $i^*_{t-1} < i^*_t$.

Cases 2.1, 2.3(a): (Fig. 14, 15)

Using $d_{i^*_t,j} \geq 0$ for $x < j < x_t$, and in particular $D_{P_{r-1}}(U,A) = a_{i^*_t,x+1} - a_{i^*_t,x}$, we obtain

$$\begin{aligned} D_{P_{r-1}}(A,B) &= a_{i^*_t,x_t} - a_{i^*_t,x+1}, \\ D_{P_r}(B',V) &= D_{P_{r-1}}(B,V) + a_{i^*_t,x_t}, \\ D_{P_r}(U,A') &= \max\{0, d_{i^*_t+1,x+1}\} - a_{i^*_t,x} \\ &= \max\{0, d_{i^*_t+1,x+1}\} + D_{P_{r-1}}(U,A) - a_{i^*_t,x+1}, \end{aligned}$$

and so with (5)

$$\begin{aligned} \delta(P_r) = \delta(P_{r-1}) - D_{P_{r-1}}(U,A) - \left(a_{i^*_t,x_t} - a_{i^*_t,x+1}\right) - D_{P_{r-1}}(B,V) \\ + \left(\max\{0, d_{i^*_t+1,x+1}\} + D_{P_{r-1}}(U,A) - a_{i^*_t,x+1}\right) \\ + D_{P_r}(A',B') + \left(D_{P_{r-1}}(B,V) + a_{i^*_t,x_t}\right), \end{aligned}$$

that is

$$\begin{aligned} \delta(P_r) &= \delta(P_{r-1}) + \max\{0, d_{i^*_t+1,x+1}\} + D_{P_r}(A',B') \\ 7 &\geq \delta(P_{r-1}). \end{aligned}$$

Cases 2.2, 2.3(b): (Fig. 16, 17)

The computation is the same as in Case 2.1 but in the formula for $D_{P_r}(U,A')$ we have to replace $d_{i^*_t+1,x+1}$ by $d_{i^*_t-1,x+1}$.

□

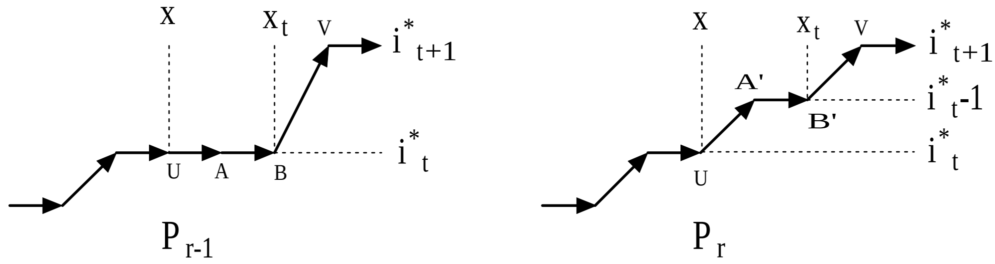


Figure 16: Transition from P_{r-1} to P_r in Case 2.2.

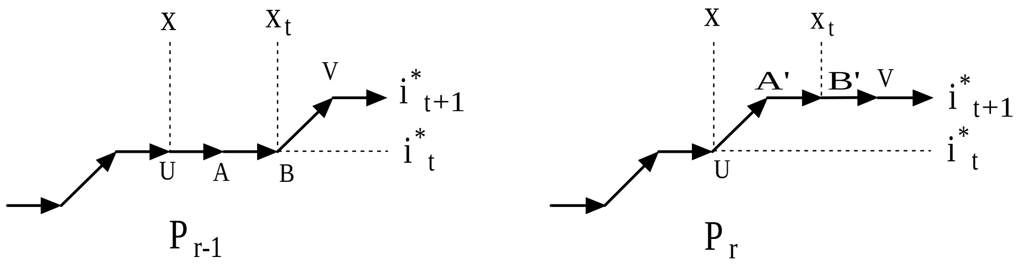


Figure 17: Transition from P_{r-1} to P_r in Case 2.3.b) $i^*_{t-1} > i^*_t$.

From Lemmas 5, 7 and 8 we deduce by duality that $c(A)$ is a lower bound for the sum of the coefficients of a segmentation of A and thus we have already proved the first half of the theorem.

4 The algorithm

In this section we assume $c(A) > 0$ and construct a segment S such that $A-S$ is still nonnegative and $c(A-S) \leq c(A)-1$. Iterating this construction we obtain a sequence of $c(A)$ segments whose sum is A. For $(i,j) \in V$ we denote by $\alpha_1(i,j)$ the maximal length of a $(0,(i,j))-$path, by $\alpha_2(i,j)$ the maximal length of an $((i,j),1)-$path and by $\alpha(i,j)$ the maximal length of a $(0,1)-$path through (i,j), that is

$$\begin{aligned}\alpha_1(i,j) &= D(0,(i,j)),\\ \alpha_2(i,j) &= D((i,j),1),\\ \alpha(i,j) &= \alpha_1(i,j)+\alpha_2(i,j).\end{aligned}$$

Now we define two subsets $V_1, V_2 \subseteq V$. In V_1 we collect the pairs (i,j) that determine local maxima or right ends of plateaus in the sequences

$a_{i,1}, a_{i,2}, \dots, a_{i,n}$ $(i = 1, 2, \dots, m)$, precisely

$$V_1 = \{(i,j) \in V \ : \ d_{i,j} \geq 0,\ d_{i,j+1} < 0\}.$$

The second subset V_2 is defined to be the set of pairs $(i,j) \in V_1$ with the following properties

1. There exists a $(0,1)-$path P of length $c(A)$ through (i,j).
2. The sequence $a_{i,1}, a_{i,2}, \dots, a_{i,j}$ is increasing, i.e. $a_{i,1} \leq a_{i,2} \leq \dots \leq a_{i,j}$.
3. The horizontal $(0,(i,j))-$path is a $(0,(i,j))-$path of maximal length.

In other words,

$$V_2 = \{(i,j) \in V_1 \ : \ \alpha(i,j) = c(A) \text{ and } \alpha_1(i,j) = a_{i,j}\}.$$

Observe that for $(i,j) \in V_1$, $\delta((i,j),(i,j+1)) = 0$ and thus, for $j'' > j$,

$$\begin{aligned} \delta((0,(i,1),(i,2),\dots,(i,j''))) &= \sum_{j'=1}^{j''} \max\{0, d_{i,j'}\} \\ &\geq \sum_{j'=1}^{j} d_{i,j'} + \sum_{j'=j+2}^{j''} d_{i,j'} = a_{i,j} + (a_{i,j''} - a_{i,j+1}) \\ &> a_{i,j''}, \end{aligned}$$

and hence $\alpha_1(i,j'') > a_{i,j''}$. In particular, for any fixed row i there is at most one column index j with $(i,j) \in V_2$. To see that $c(A) > 0$ implies $V_2 \neq \emptyset$ consider a feasible $(0,1)-$path P with $\delta(P) = c(A)$. If P is a horizontal path without any row change then $\delta(P) > 0$ implies that P contains an element of V_1. Otherwise let $((i,j),(i',j+1))$ be the first row change of P. Then by the feasibility of P, $d_{i,j} < 0$ und thus the subpath $0,(i,1),\dots,(i,j)$ contains an element of V_1. In both cases the first vertex on P which is in V_1 is in V_2 as well. We denote the elements of V_2 by

$$(i_1, j_1), (i_2, j_2), \dots, (i_t, j_t),$$

such that $i_1 < i_2 < \dots < i_t$. A segment S (given by the parameters $l_1, l_2, \dots, l_m, r_1, r_2, \dots, r_m$) is constructed according to the following strategy. In row i_k $(k \in [t])$ we choose the open part maximal under the condition that the right boundary is j_k, i.e. we put

$$r_{i_k} = j_k \quad \text{and} \quad l_{i_k} = \max\{j \leq j_k \ : \ a_{i_k,j} = 0\} + 1.$$

In the remaining rows we choose the open part in some sense minimal under the condition that the final result is a segment: The rows $i < i_1$ and $i > i_t$ remain closed. If $l_{i_k} > r_{i_{k+1}}+1$ we choose the open part in row i_k+1 maximal with $r_{i_k+1} = l_{i_k} - 1$. If necessary we repeat this step in the following rows, until finally $l_i \leq r_{i_{k+1}} + 1$ for some i with $i_k < i < i_{k+1}$. If $l_{i_{k+1}} > r_{i_k} + 1$ we proceed analogously, starting in row $i_{k+1} - 1$. For the details of the construction see Algorithm 1.

Example 2. Let

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 9 \\ 0 & 0 & 0 & 0 & 1 & 1 & 2 & 4 & 2 \\ 0 & 0 & 2 & 2 & 3 & 3 & 3 & 2 & 1 \\ 1 & 1 & 2 & 2 & 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 4 & 2 & 2 & 2 & 4 & 4 & 7 \\ 2 & 2 & 2 & 2 & 1 & 2 & 2 & 3 & 3 \\ 0 & 2 & 2 & 7 & 2 & 2 & 2 & 1 & 1 \end{pmatrix}.$$

Then $c(A) = 9$, $V_2 = \{(1,9),(5,3),(7,4)\}$ and the algorithm yields the segment

$$S = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where the bold 1's correspond to the elements of V_2. For the resulting matrix

$$A - S = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 2 \\ 0 & 0 & 1 & 1 & 3 & 3 & 3 & 2 & 1 \\ 1 & 1 & 2 & 2 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 2 & 2 & 2 & 4 & 4 & 7 \\ 2 & 2 & 2 & 2 & 1 & 2 & 2 & 3 & 3 \\ 0 & 1 & 1 & 6 & 2 & 2 & 2 & 1 & 1 \end{pmatrix}$$

we have $c(A - S) = 8$.

To prove the correctness of the algorithm we need an alternative description of paths in $\overrightarrow{G}$ that yields some insight into the relation between the constructed segment S and the path lengths. For this let $\overrightarrow{H}$ be a directed

Algorithm 1 Segment $S(A, V_2)$

for $(i, j) \in V_2$ **do**
  $l_i := \max\{j' \leq j \ : \ a_{i,j'} = 0\}{+}1$
  $r_i := j$
end for
for $i = 1$ to $i_1 - 1$ **do**
  $l_i := l_{i_1}$; $r_i := l_i - 1$
end for
for $i = i_t + 1$ to m **do**
  $l_i := l_{i_t}$; $r_i := l_i - 1$
end for
for $k = 1$ to $t - 1$ **do**
  if $j_k > j_{k+1}$ **then**
    $i := i_k$
    while $i < i_{k+1}$ and $l_i > r_{i_{k+1}} + 1$ **do**
      $i := i + 1$
      $r_i := l_{i-1} - 1$
      $l_i := \max\{j \leq r_i \ : \ a_{ij} = 0\} + 1$
    end while
    for $i' = i + 1$ to $i_{k+1} - 1$ **do**
      $r_{i'} := r_{i_{k+1}}$; $l_{i'} := r_{i'} + 1$
    end for
  else
    $i := i_{k+1}$
    while $i > i_k$ and $l_i > r_{i_k} + 1$ **do**
      $i := i - 1$
      $r_i := l_{i+1} - 1$
      $l_i := \max\{j \leq r_i \ : \ a_{ij} = 0\} + 1$
    end while
    for $i' = i_k + 1$ to $i - 1$ **do**
      $r_{i'} := r_{i_k}$; $l_{i'} := r_{i'} + 1$
    end for
  end if
end for

graph with vertex set $V \cup \{0,1\}$. As the edge set of H we take $E_0 = E_0^{(1)} \cup E_0^{(2)} \cup E_0^{(3)} \cup E_0^{(4)}$, where

$$\begin{aligned}
E_0^{(1)} &= \{(0,(i,1)) \ : \ i \in [m]\} \cup \{((i,n),1) \ : \ i \in [m]\},\\
E_0^{(2)} &= \{((i,j),(i,j+1) \ : \ i \in [m],\ j \in [n-1])\},\\
E_0^{(3)} &= \{((i,j),(i+1,j)) \ : \ i \in [m-1],\ j \in [n]\},\\
E_0^{(4)} &= \{((i,j),(i-1,j)) \ : \ 2 \le i \le m,\ j \in [n]\}.
\end{aligned}$$

Let the length function δ_0 on E_0 be defined by

$$\begin{aligned}
\delta_0(0,(i,1)) &= a_{i,1} && (i \in [m]),\\
\delta_0((i,n),1) &= 0 && (i \in [m]),\\
\delta_0((i,j),(i,j+1)) &= \max\{0, d_{i,j+1}\} && (i \in [m],\ j \in [n-1]),\\
\delta_0((i,j),(i+1,j)) &= -a_{i,j} && (i \in [m-1],\ j \in [n]),\\
\delta_0((i,j),(i-1,j)) &= -a_{i,j} && (2 \le i \le m,\ j \in [n]).
\end{aligned}$$

It is easy to see that there is a bijection between the paths in $\overrightarrow{G}$ and the paths in $\overrightarrow{H}$ with the additional restriction that the last edge is in $E_0^{(1)} \cup E_0^{(2)}$. In addition this bijection preserves the length, that is for a path P in $\overrightarrow{G}$ and the corresponding path Q in $\overrightarrow{H}$ we have

$$\delta(P) = \delta_0(Q).$$

In particular, there is a length–preserving bijection between the $(0,1)-$paths in $\overrightarrow{G}$ and $\overrightarrow{H}$. The advantage of Q compared to P is that possibly existing ,,long, skew" edges in P are replaced by a sequence of vertical edges and one horizontal edge, and the lengths of these edges are easier to control. For $v, w \in V \cup \{0,1\}$ we put

$$D_0(v,w) = \max\{\delta_0(Q) \ : \ Q \ (v,w)-\text{path in } \overrightarrow{H}\},$$

and analogous to α, α_1 and α_2 we define for $(i,j) \in V$,

$$\begin{aligned}
\beta_1(i,j) &= D_0(0,(i,j)),\\
\beta_2(i,j) &= D_0((i,j),1),\\
\beta(i,j) &= \beta_1(i,j) + \beta_2(i,j).
\end{aligned}$$

We need some information about the connection between the distances in $\overrightarrow{G}$ and $\overrightarrow{H}$. Obviously $\beta_2(i,j) = \alpha_2(i,j)$ for all $(i,j) \in V$. The next lemma is an analogous result about α_1 and β_1 for the vertices on $(0,1)-$paths of maximal length.

Lemma 9. *For all $(i,j) \in V$ with $\alpha(i,j) = c(A)$, we have $\beta_1(i,j) = \alpha_1(i,j)$.*

Proof. $\beta_1(i,j) \geq \alpha_1(i,j)$ is trivial, since for every $(0,(i,j))-$path in $\overrightarrow{G}$ there is the corresponding $(0,(i,j))-$path in $\overrightarrow{H}$ of the same length. Let Q_1 and Q_2 be a $(0,(i,j))-$path in $\overrightarrow{H}$ and a $((i,j),1)-$path in $\overrightarrow{H}$, respectively, with

$$\delta_0(Q_1) = \beta_1(i,j) \quad \text{and} \quad \delta_0(Q_2) = \beta_2(i,j) = \alpha_2(i,j).$$

By concatenating Q_1 and Q_2 we obtain a $(0,1)-$path Q in $\overrightarrow{H}$ with

$$\delta_0(Q) = \beta_1(i,j) + \alpha_2(i,j).$$

Since the last edge of Q is in $E_0^{(1)}$, this implies the existence of a $(0,1)-$path P in $\overrightarrow{G}$ with $\delta(P) = \beta_1(i,j) + \alpha_2(i,j)$. So

$$\beta_1(i,j) + \alpha_2(i,j) \leq c(A) = \alpha_1(i,j) + \alpha_2(i,j),$$

and thus $\beta_1(i,j) \leq \alpha_1(i,j)$. $\square$

Lemma 10. *Let $(i,j),(k,l) \in V$, $i > k$ and put $p = i - k$.*

1. *If $j < l$ and there are column indices $j'_1, j'_2, \ldots, j'_p$ such that*

$$j \leq j'_1 \leq j'_2 \leq \cdots \leq j'_p < l \quad \textit{and}$$

$$a_{i-q,j'_q} = 0 \qquad \textit{for } q = 1,2,\ldots,p,$$

then there exists a $((i,j),(k,l))-$path P in $\overrightarrow{G}$ with

$$\delta(P) \geq a_{k,l} - a_{i,j}.$$

2. *If $j > l$ and there are column indices $j'_1, j'_2, \ldots, j'_p$ such that*

$$l \leq j'_1 \leq j'_2 \leq \ldots \leq j'_p < j \quad \textit{and}$$

$$a_{k+q,j'_q} = 0 \qquad \textit{for } q = 1,2,\ldots,p,$$

then there exists a $((k,l),(i,j))-$path P in $\overrightarrow{G}$ with

$$\delta(P) \geq a_{i,j} - a_{k,l}.$$

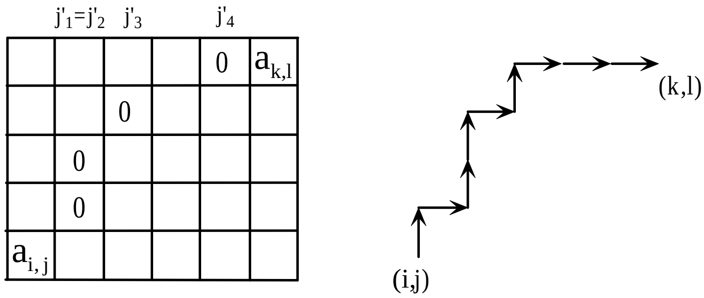


Figure 18: A part of a matrix A as described in the first case of Lemma 10 and the corresponding path Q.

Proof. We consider only the first case that is illustrated in Fig. 18. The second one is treated analogously. First we construct a $((i,j),(k,l))-$path Q in $\overrightarrow{H}$. We take $((i,j),(i-1,j))$ with length $-a_{i,j}$ as the first edge and complete this edge to a $((i,j),(k,l))-$path Q in such a way, that row changes occur only along the edges

$$((i-q, j'_q),(i-q-1, j'_q)) \qquad (1 \leq q \leq p-1).$$

This is possible by our assumption on the j'_q. Thus the vertical edges of Q, except for the first one, have length 0 and since the horizontal edges have nonnegative length in any case we conclude that the $((i,j),(k,j'_p))-$subpath of Q has length at least $-a_{i,j}$. Finally the length of the path

$$(k, i'_p),(k, i'_p+1),\dots,(k,l)$$

is at least $a_{k,l}$ and from $l > i'_p$ follows that the last edge of Q is in $E_0^{(2)}$ and thus there exists a $((i,j),(k,l))-$path P in $\overrightarrow{G}$ with

$$\delta(P) = \delta(Q) \geq a_{k,l} - a_{i,j}.$$

□

Lemma 11. *Algorithm 1 yields a segment S.*

Proof. Suppose the algorithm does not yield a segment. This is possible only if for some $k \in [t-1]$ the condition of the while–loop in line 14 (resp. line 24) holds for all $i \in \{i_k, i_k+1, \dots, i_{k+1}-1\}$ (resp. for all $i \in \{i_k+1, i_k+2, \dots, i_{k+1}\}$). If $j_k = j_{k+1}$ then

$$l_{i_k} \leq r_{i_{k+1}} \quad \text{and} \quad l_{i_{k+1}} \leq r_{i_k}.$$

So we may assume $j_k \neq j_{k+1}$. Let $j_k > j_{k+1}$. (The case $j_k < j_{k+1}$ can be treated analogously.) We put $p = i_{k+1} - i_k$ and

$$j'_q = l_{i_{k+1}-q} - 1 \qquad (q = 1, 2, \ldots, p).$$

The assumption that the while–condition is fulfilled for all $i_{k+1} - q$ ($q = 1, 2, \ldots, p$) implies

$$r_{i_{k+1}} + 1 \leq j'_1 \leq j'_2 \leq \ldots \leq j'_p < j_k \quad \text{and}$$

$$a_{i_{k+1}-q, j'_q} = 0 \qquad (q = 1, 2, \ldots, p).$$

Thus by Lemma 10 there is a $((i_{k+1}, j_{k+1} + 1), (i_k, j_k))-$path P_0 in $\overrightarrow{G}$ of length at least $a_{i_k,j_k} - a_{i_{k+1},j_{k+1}+1}$. Using $(i_{k+1}, j_{k+1}) \in V_2$ this yields

$$\delta_0(P_0) > a_{i_k,j_k} - a_{i_{k+1},j_{k+1}}.$$

Now we concatenate the path $0, (i_{k+1}, 1), (i_{k+1}, 2), \ldots, (i_{k+1}, j_{k+1} + 1)$ with P_0 to obtain a $(0, (i_k, j_k))-$path of length at least

$$a_{i_{k+1},j_{k+1}} + \delta(P_0) > a_{i_k,j_k},$$

in contradiction to $(i_k, j_k) \in V_2$. □

Let $S = (s_{i,j})$ be the result of algorithm 1. By construction $s_{i,j} = 1$ implies $a_{i,j} > 1$ and so the entries of $A - S$ ar nonnegative. We put

$$\begin{aligned} a'_{i,j} &= a_{i,j} - s_{i,j} && (i \in [m], j \in [n]), \\ a'_{i,0} &= a_{i,n+1} = 0 && (i \in [m]), \\ d'_{i,j} &= a'_{i,j} - a'_{i,j-1} && (i \in [m], j \in [n]). \end{aligned}$$

By δ' and δ'_0 we denote the length functions on $\overrightarrow{G}$ and $\overrightarrow{H}$, respectively, which correspond to $A' = (a'_{i,j})$, and by D' and D'_0 the corresponding distance functions. For $(i, j) \in V$ we put

$$\begin{aligned} \alpha'_1(i,j) &= D'(0,(i,j)), \\ \alpha'_2(i,j) &= D'((i,j),1), \\ \alpha'(i,j) &= \alpha'_1(i,j) + \alpha'_2(i,j), \\ \beta'_1(i,j) &= D'_0(0,(i,j)), \\ \beta'_2(i,j) &= D'_0((i,j),1), \\ \beta'(i,j) &= \beta'_1(i,j) + \beta'_2(i,j). \end{aligned}$$

By T we denote the subset of V which corresponds to the segment S, that is

$$T = \{(i,j) \in V \ : \ s_{i,j} = 1\}.$$

The next lemma asserts that for $(i,j) \in T$ the sequence $a_{i,1}, \ldots, a_{i,j}$ is increasing and the horizontal path from 0 to (i,j) has maximal length with respect to A in both of $\overrightarrow{G}$ and $\overrightarrow{H}$.

Lemma 12. *For $(i,j) \in T$ we have*

$$\beta_1(i,j) = \alpha_1(i,j) = a_{i,j} \quad \textit{and} \quad \alpha(i,j) = c(A).$$

Proof. Let $(i,j) \in T$. Clearly,

$$\beta_1(i,j) \geq \alpha_1(i,j) \geq a_{i,j}.$$

Assume P_0 is a $(0,(i,j))$−path in $\overrightarrow{G}$ with $\delta(P_0) > a_{i,j}$. We claim that for some $k \in [t]$ there is an $((i,j),(i_k,j_k))$−path P_1 in $\overrightarrow{G}$ of length at least $a_{i_k,j_k} - a_{i,j}$. To see this we distinguish 3 types of vertices in T:

1. $i = i_k$ and $j \leq j_k$ for some $k \in [t]$:

   The path $(i_k,j),(i_k,j+1),\ldots,(i_k,j_k)$ has length $a_{i_k,j_k} - a_{i_k,j}$.

2. $i_k < i < i_{k+1}$ for some $k \in [t-1]$ with $j_k > j_{k+1}$:

   By construction of S there are column indices $j'_1, j'_2, \ldots, j'_p$, where $p = i - i_k$, such that

   $$j \leq j'_1 \leq j'_2 \leq \cdots \leq j'_p < j_k \quad \text{and}$$

   $$a_{i-q,j'_q} = 0 \qquad (q = 1,2,\ldots,p).$$

   Thus the claim follows by Lemma 10.

3. $i_{k-1} < i < i_k$ for some $k \in \{2,3,\ldots,t\}$ with $j_{k-1} < j_k$:

   By construction of S there are column indices $j'_1, j'_2, \ldots, j'_p$, where $p = i_k - i$, such that

   $$j \leq j'_1 \leq j'_2 \leq \ldots \leq j'_p < j_k \quad \text{and}$$

   $$a_{i+q,j'_q} = 0 \qquad (q = 1,2,\ldots,p).$$

   Thus the claim follows by Lemma 10.

But now we can concatenate P_0 and P_1 to obtain a $(0,(i_k,j_k))-$path P in $\overrightarrow{G}$ with $\delta(P) > a_{i_k,j_k}$, in contradiction to $(i_k,j_k) \in V_2$. This proves $\alpha_1(i,j) = a_{i,j}$. In addition, concatenating the paths $(0,(i,1),(i,2),\ldots,(i,j))$, P_1 and a $((i_k,j_k),1)-$path of maximal length yields $\alpha(i,j) = c(A)$ and thus also $\beta_1(i,j) = \alpha_1(i,j)$ by Lemma 9. $\square$

Now we want to prove that for $(i,j) \in T$ the horizontal $(0,(i,j))-$path is still maximal with respect to A'. We need the following necessary condition for $\beta_1(i,j) > a_{i,j}$.

Lemma 13. *Suppose $\beta_1(i,j) > a_{i,j}$ and Q is a $(0,(i,j))-$path in $\overrightarrow{H}$ with $\delta_0(Q) = \beta_1(i,j)$. Then there exists a vertex $(i',j') \in V_1$ such that either*

- *$j' = 1$ and $((i',1),(i',2))$ is an edge of Q or*
- *$1 < j' < n$ and $((i',j'-1),(i',j'))$, $((i',j'),(i',j'+1)$ are edges of Q.*

If in addition $\beta(i,j) = c(A)$ then we can choose (i',j') even in V_2.

Proof. Let Q be a $(0,(i,j))-$path with $\delta_0(Q) = \beta_1(i,j)$ and assume there is no such vertex in V_1. We show $\delta_0(Q) = a_{i,j}$ which gives the desired contradiction. Clearly, $\delta_0(Q) \geq a_{i,j}$. The first edge of Q is of the form $(0,(i',1))$ and has length $a_{i',1}$. So we may assume that Q has more than one edge and proceed by induction on the number of edges of an initial subpath of Q.

Case 1: The last edge of Q is in $E_0^{(3)} \cup E_0^{(4)}$.

W.l.o.g. the last edge is $((i-1,j),(i,j))$ with length $-a_{i-1,j}$. Since by induction $\delta_0(Q \setminus \{(i,j)\}) = a_{i-1,j}$, we obtain $\delta_0(Q) = 0 \leq a_{i,j}$.

Case 2: The last edge of Q is in $E_0^{(2)}$, and the second last edge is in $E_0^{(3)} \cup E_0^{(4)}$. W.l.o.g. the last two edges of Q are $((i-1,j-1),(i,j-1))$ and $((i,j-1),(i,j))$. By induction the length of the $(0,(i-1,j-1))-$subpath of Q is $a_{i-1,j-1}$. Thus the length of the $(0,(i,j-1))-$subpath is 0 and by maximality of Q follows $a_{i,j-1} = 0$, hence $\delta_0(Q) = a_{i,j}$.

Case 3: The last two edges of Q are in $E_0^{(1)} \cup E_0^{(2)}$.

By induction the $(0,(i,j-1))-$subpath of Q has length $a_{i,j-1}$. By maximality of Q this implies $d_{i,j'} \geq 0$ for all j', $1 \leq j' \leq j-1$. Now $d_{i,j} \geq 0$, since otherwise $(i,j-1)$ is a vertex in V_1 that fulfills the conditions of the lemma. Thus $\delta_0(Q) = a_{i,j}$.

Now suppose $\beta(i,j) = c(A)$. Then we can complete Q to a $(0,1)-$path Q' of length $c(A)$. Let P be the corresponding $(0,1)-$path in $\overrightarrow{G}$, and let $(i',j') \in V_1$ be the first vertex on Q that has the claimed properties. Then $(i',j') \in P$ and the $(0,(i',j'))-$subpath of P has length $a_{i',j'}$, that is $(i',j') \in V_2$. $\square$

Lemma 14. *For $(i,j) \in T$ we have $\beta'_1(i,j) = \alpha'_1(i,j) = a'_{ij}$.*

Proof. Again trivially,

$$\beta'_1(i,j) \geq \alpha'_1(i,j) \geq a'_{i,j}.$$

Let $(i,j) \in T$ and assume $\beta'_1(i,j) > a'_{i,j}$. In particular, $j > 1$ since obviously $\beta'_1(i,1) = a'_{i,1}$ for all $i \in [m]$. There is a $(0,(i,j))-$path Q in $\overrightarrow{H}$ with

$$\delta'_0(Q) = \beta'_1(i,j) > a'_{i,j}.$$

W.l.o.g. we may assume that (i,j) is the first counterexample to the lemma on Q, i.e.

$$\beta'_1(i_0,j_0) = a'_{i_0,j_0} \quad \text{for all } (i_0,j_0) \in (Q \setminus \{(i,j)\}) \cap T.$$

Case 1: $(Q \setminus \{(i,j)\}) \cap T = \emptyset$.

Let e be the last edge of Q. Then $\delta_0(e_1) = \delta'_0(e_1)$ for all edges $e_1 \neq e$ of Q.

Case 1.1: $e \in E_0^{(2)}$.

Then $\delta_0(e) = \delta'_0(e)+1$, hence $\delta_0(Q) = \delta'_0(Q)+1$, and consequently (using Lemma 12),

$$\beta'_1(i,j) = \delta_0(Q) - 1 \leq \beta_1(i,j) - 1 = a'_{i,j}.$$

Case 1.2: $e \in E_0^{(3)} \cup E_0^{(4)}$.

W.l.o.g. $e = ((i-1,j),(i,j))$ and $\delta_0(e) = \delta'_0(e) = -a_{i-1,j}$, and thus

$$\delta_0(Q) = \delta'_0(Q) = \beta'_1(i,j).$$

Assume $\delta_0(Q) = \beta_1(i,j) = a_{i,j}$. Then $\delta_0(Q) > 0$, and thus

$$\delta_0(Q \setminus \{(i,j)\}) > a_{i-1,j}.$$

By Lemma 12, $\beta(i,j) = \alpha(i,j) = c(A)$ and consequently by Lemma 13, $Q \setminus \{(i,j)\}$ contains a vertex $(i_0,j_0) \in V_2 \subseteq T$. This is a contradiction and we conclude

$$\beta'_1(i,j) = \delta_0(Q) < \beta_1(i,j) = a_{i,j},$$

and thus $\beta'_1(i,j) = a'_{i,j}$.

Case 2: $(Q \setminus \{(i,j)\}) \cap T \neq \emptyset$.

Let (i_0, j_0) be the last vertex on $Q \setminus \{(i,j)\}$ that is in T and denote by Q_1 and Q_2 the $(0,(i_0,j_0))-$subpath and the $((i_0,j_0),(i,j))-$subpath of Q, respectively. By assumption $\delta_0'(Q_1) = a'_{i_0,j_0}$, and so w.l.o.g. we may assume $Q_1 = (0,(i_0,1),(i_0,2),\ldots,(i_0,j_0))$. We denote the edges of Q_2 by $e_1, e_2, \ldots, e_p$. For $p = 1$ we obtain

$$\begin{aligned} \delta_0'(Q) &= \delta_0'(Q_1) - a'_{i_0,j_0} = 0 && \text{if } e_1 \in E_0^{(3)} \cup E_0^{(4)} \quad \text{and} \\ \delta_0'(Q) &= \delta_0'(Q_1) + \max\{0, d'_{i,j}\} && \text{if } e_1 \in E_0^{(2)}. \end{aligned}$$

Since $e_1 \in E_0^{(2)}$ implies $(i,j),(i,j-1) \in T$ and thus $d'_{ij} = d_{i,j} \geq 0$ (Lemma 12), we obtain $\delta_0'(Q) \leq a'_{i,j}$ and consequently $\beta_1'(i,j) = a'_{i,j}$. So let $p > 1$. Then

$$\delta_0'(e_i) = \delta_0(e_i) \quad (2 \leq i \leq p-1),$$

$$\delta_0'(e_1) = \begin{cases} \delta_0(e_1) + 1 & \text{if } e_1 \in E_0^{(3)} \cup E_0^{(4)}, \\ \delta_0(e_1) + 1 & \text{if } e_1 \in E_0^{(2)} \text{ and } d_{i_0,j_0+1} \geq 0, \\ \delta_0(e_1) & \text{if } e_1 \in E_0^{(2)} \text{ and } d_{i_0,j_0+1} < 0 \quad \text{and} \end{cases}$$

$$\delta_0'(e_p) = \begin{cases} \delta_0(e_p) & \text{if } e_p \in E_0^{(3)} \cup E_0^{(4)}, \\ \delta_0(e_p) - 1 & \text{if } e_p \in E_0^{(2)}. \end{cases}$$

Case 2.1: $\delta_0'(Q_2) \leq \delta_0(Q_2)$.

$\delta_0'(Q) = \delta_0'(Q_1) + \delta_0'(Q_2) \leq a'_{i_0,j_0} + \delta_0(Q_2) < \delta_0(Q)$ implies

$$\beta_1'(i,j) < \beta_1(i,j) = a_{i,j},$$

and thus $\beta_1'(i,j) = a'_{i,j}$.

Case 2.2: $\delta_0'(Q_2) = \delta_0(Q_2) + 1$.

In this case $\delta_0'(Q) = \delta_0(Q)$ and $e_p \in E_0^{(3)} \cup E_0^{(4)}$, w.l.o.g. $e = ((i-1,j),(i,j))$ with length $-a_{i-1,j}$. Assume $\delta_0(Q) = \beta_1(i,j)$. Then $\delta_0(Q) > 0$ and thus

$$\beta_1(i-1,j) > a_{i-1,j}.$$

By Lemma 12, $\beta(i,j) = \alpha(i,j) = c(A)$, and by Lemma 13 there is a vertex $(i_1,j_1) \in V_2$ such that Q contains the edge $((i_1,j_1),(i_1,j_1+1))$. Now $\delta_0'(Q_2) = \delta_0(Q_2) + 1$ is possible only if

$$e_1 \in E_0^{(3)} \cup E_0^{(4)} \quad \text{or} \quad (e_1 \in E_0^{(2)} \text{ and } d_{i_0,j_0+1} \geq 0).$$

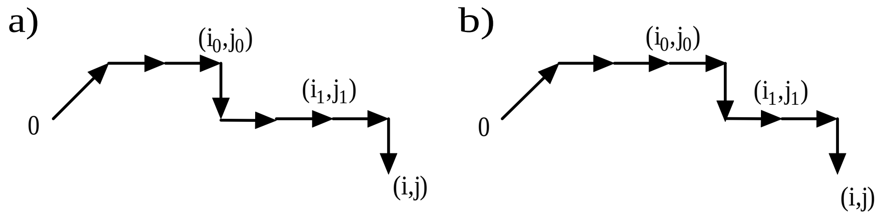


Figure 19: Paths Q as in Case 2.2. a) $e_1 \in E_0^{(3)} \cup E_0^{(4)}$, b) $e_1 \in E_0^{(2)}$ and $d_{i_0,j_0+1} \geq 0$.

Hence, using $d_{i_0,j'} \geq 0$ for $1 \leq j' \leq j_0$, $(i_1, j_1) \notin Q_1$ and we obtain the contradiction

$$(i_1, j_1) \in (Q \setminus \{(i,j),(i_0,j_0)\}) \cap V_2.$$

Thus $\delta_0'(Q) = \delta_0(Q) < \beta_1(i,j) = a_{i,j}$, and so $\beta_1'(i,j) = a_{i,j}'$.

□

Now we are prepared for the final step.

Lemma 15. $c(A') \leq c(A) - 1$.

Proof. Let Q be a $(0,1)-$path in $\overrightarrow{H}$ with $\delta_0(Q) = c(A')$ and let (i_0, j_0) be the last vertex on Q that is in T. We denote the $(0,(i_0,j_0))-$subpath and the $((i_0,j_0),1)-$subpath of Q by Q_1 and Q_2, respectively. By Lemmas 12 and 14,

$$\beta_1(i_0,j_0) = a_{i_0,j_0} = a'_{i_0,j_0} + 1 = \beta_1'(i_0,j_0) + 1,$$

and w.l.o.g. we may assume $Q_1 = (0,(i_0,1),(i_0,2),\ldots,(i_0,j_0))$. For the first edge e_0 of Q_2 we have $\delta_0(e_0) = \delta_0'(e_0)$ or $\delta_0(e_0) = \delta_0'(e_0) - 1$, and for all edges $e \neq e_0$ of Q_2, $\delta_0(e) = \delta_0'(e)$.

Case 1: $\delta_0(e_0) = \delta_0'(e_0)$.

$$\begin{aligned}\delta_0(Q) &= \delta_0(Q_1) + \delta_0(Q_2) = \delta_0'(Q_1) + 1 + \delta_0'(Q_2)\\ &= \delta_0'(Q) + 1 = c(A') + 1,\end{aligned}$$

and thus $c(A) \geq c(A') + 1$.

Case 2: $\delta_0(e_0) = \delta_0'(e_0) - 1$.

By the same argument as in the first case we only get

$$\delta_0(Q) = c(A').$$

Now assume $\delta_0(Q_2) = \alpha_2(i_0, j_0)$. From

$$\alpha(i_0, j_0) = c(A) \quad \text{and} \quad \alpha_1(i_0, j_0) = a_{i_0,j_0}$$

we deduce $\delta_0(Q) = c(A)$. By Lemma 13 either Q contains an edge $((i_1, j_1), (i_1, j_1 + 1))$ with $(i_1, j_1) \in V_2$ or the last edge of Q is $((i_1, j_1), 1)$ with $j_1 = n$ and $(i_1, j_1) \in V_2$. From $\delta_0(e_0) = \delta_0'(e_0) - 1$ follows that either

$$e_0 \in E_0^{(3)} \cup E_0^{(4)} \quad \text{or} \quad (e_0 \in E_0^{(2)} \text{ and } d_{i_0,j_0+1} \geq 0).$$

Hence, using $d_{i_0,j'} \geq 0$ for $1 \leq j' \leq j_0$, $(i_1, j_1) \notin Q_1$ and we obtain the contradiction

$$(Q_2 \setminus \{(i_0, j_0)\}) \cap V_2 \neq \emptyset.$$

Consequently, $\delta_0(Q_2) < \alpha_2(i_0, j_0)$ and there exists an $((i_0, j_0), 1)-$path Q_2^* with $\delta_0(Q_2^*) > \delta_0(Q_2)$. By concatenating Q_1 and Q_2^* we obtain a $(0, 1)-$path Q^* with $\delta_0(Q^*) > c(A')$, and thus

$$c(A) \geq c(A') + 1.$$

□

Now we collect the lemmas to prove the theorem.

Proof of the theorem. The lower bound is an immedate consequence of the Lemmas 5, 7 and 8 and duality. The existence of a segmentation with $\sum_{i=1}^{k} u_i = c(A)$ is proved by induction on $c(A)$. If $c(A) = 0$ then $A = 0$ and there is nothing to do. For $c(A) > 0$ we apply Algorithm 1 to construct a segment S with $c(A - S) \leq c(A) - 1$. By induction there are segments $S_2, S_3, \ldots, S_k$ and positive integers $u_2, u_3, \ldots, u_k$ such that

$$A - S = \sum_{i=2}^{k} u_i S_i \quad \text{and} \quad \sum_{i=2}^{k} u_i = c(A - S) \leq c(A) - 1,$$

and thus with $S_1 = S$ and $u_1 = 1$,

$$A = \sum_{i=1}^{k} u_i S_i \quad \text{and} \quad \sum_{i=1}^{k} u_i = c(A - S) + 1 \leq c(A).$$

□

L	TNMU (with ICC)	TNMU (without ICC)
3	15.4	14.0
4	19.5	17.9
5	23.6	21.7
6	27.7	25.6
7	31.8	29.4
8	35.7	33.2
9	39.7	37.0
10	43.7	40.9
11	47.7	44.7
12	51.7	48.5
13	55.8	52.3
14	59.7	56.2
15	63.7	59.8
16	67.7	63.3

Table 1: Test results for $m = n = 15$.

5 Results

Table 1 shows some test results of our algorithm in comparison with the optimal TNMU without ICC. The results of the third column are taken from [5]. For each row we computed segmentations for 10000 matrices with randomly chosen entries from $\{0, 1, \dots, L\}$ (uniformly distributed). On a 1.3 GHz PC the computation for the whole table took 206 seconds. In addition we mention that for $m = n = 15$ and $L = 10000$ the algorithm provides segmentations with an average of 39823.0 monitor units, compared to 37880.2 that are needed without ICC ([5]). With respect to time consumption the algorithm is completely practicable: The segmentation of one $100 \times 100-$matrix takes 14 seconds.

6 Conluding remarks

An important next step to make our method applicable to real world problems is to reduce the number of segments in the segmentation, that is the k in

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

A priori our algorithm yields a segmentation with $u_i = 1$ for all i and thus with $c(A)$ segments. A rather trivial approach is to determine how often successive calls of the algorithm yield the same segment and then merge these segments. But the reduction of the number of segments that is achieved by this method is not sufficient, and further research on this problem is necessary.
Acknowledgement. I would like to thank Prof. Konrad Engel for introducing the problem to me and making many helpful comments on the presentation of the results.

References

[1] N. Boland, H.W. Hamacher, and F. Lenzen. Minimizing beam-on time in cancer radiation treatment using multileaf collimators. under revision for NETWORKS, 2002.

[2] 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.

[3] A.L. Boyer and C.Y. Yu. Intensity–modulated radiation therapy with dynamic multileaf collimators. *Semin. Radiat. Oncol.*, 9:48–59, 1999.

[4] J. Dai and Y. Zhu. Minimizing the number of segments in a delivery sequence for intensity–modulated radiation therapy with a multileaf collimator. *Med. Phys.*, 28:2113–2120, 2001.

[5] K. Engel. A new algorithm for optimal multileaf collimator field segmentation. submitted, 2002.

[6] J.M. Galvin, X.G. Chen, and R.M. Smith. Combining multileaf fields to modulate fluence distributions. *Int. J. Radiat. Oncol. Biol. Phys.*, 27:697–705, 1993.

[7] M. Langer, V. Thai, and L. Papiez. Improved leaf sequencing reduces segments of monitor units needed to deliver IMRT using multileaf collimators. *Med. Phys.*, 28:2450–2458, 2001.

[8] W. Que. Comparison of algorithms for multileaf collimator field segmentation. *Med. Phys.*, 26:2390–2396, 1999.

[9] R.A.C. Siochi. Minimizing static intensity modulation delivery time using an intensity solid paradigm. *Int. J. Radiat. Oncol. Biol. Phys.*, 43:671–680, 1999.

[10] P. Xia and L. Verhey. Multileaf collimator leaf–sequencing algorithm for intensity modulated beams with multiple static segments. *Med. Phys.*, 25:1424–1434, 1998.