

Partitioning Multiple Objective Optimal Solutions with Applications in Radiotherapy Design

Allen Holder[†]

September 13, 2001

Abstract

The optimal partition for linear programming is induced by any strictly complementary solution, and this partition is important because it characterizes the optimal set. However, constructing a strictly complementary solution in the presence of degeneracy was not practical until interior point algorithms became viable alternatives to the simplex algorithm. We develop analogs of the optimal partition for linear programming in the case of multiple objectives and show that these new partitions provide insight into the optimal set (both pareto optimality and lexicographic ordering are considered). Techniques to produce these optimal partitions are provided, and examples from the design of radiotherapy plans show that these new partitions are useful.

Key words: Optimal Partition, Multiple Objective Programming, Interior Point Algorithms, Radiation Oncology

[†] School of Business Administration, The University of Mississippi, MS USA. Research conducted at Trinity University, TX, and The University of Mississippi, MS. This research was partially supported by ONR Grant N00014-01-1-0917.

1 Introduction

Interior point algorithms have “revolutionized” the study of linear programming [27], but there has been substantially less work done in the related field of multiple objective programming, exceptions being [1], [3] and [5]. The primary purpose of this work is to begin to develop results for multiple objective programming that rely on, and that are similar to, recent results in the linear programming literature. The focus of this paper is on the *optimal partition*, which is important in linear programming because it characterizes the optimal set. Because of this, the optimal partition is fundamental when analyzing algorithms and performing sensitivity analysis. We define, and show how to generate, two new partitions for multiple objective programming. We also demonstrate that these new sets are useful in the design of radiotherapy treatment plans.

Throughout we are interested in the multiple objective linear program,

$$(MOLP) \quad \min\{Cx : Ax = b, x \geq 0\},$$

where $C \in \mathbb{R}^{p \times n}$, $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$. Without loss of generality, we assume that the rank of A is m . The p linear objectives are represented by the p rows of C , which are denoted by c^i , $i = 1, 2, \dots, p$. The feasible region is denoted by \mathcal{P} and is referred to as *decision space*. The set, $\{Cx : x \in \mathcal{P}\}$, is called *objective space*. The optimal face for the single objective linear program $\min\{c^i x : x \in \mathcal{P}\}$ is \mathcal{P}_i^* . We use \mathbb{R}_+^n to denote the set of vectors in \mathbb{R}^n with each component being non-negative, and \mathbb{R}_{++}^n indicates the further restriction that the n components are positive. The Null and Column spaces of a matrix A are $\text{null}(A)$ and $\text{col}(A)$, respectively. The vector e is the all ones vector, where length is decided by the context in which it is used. Any further notation is standard, and we refer the reader to *The Mathematical Programming Glossary* [13].

The “sense” of minimization is not uniquely defined when there are several objectives (the problem being that there is not a readily available complete ordering for \mathbb{R}^n , $n \geq 2$). Two standard approaches are *pareto* optimization and *lexicographic* optimization. A *pareto* optimal solution is a feasible element x such that there does not exist another feasible element y such that $Cy \leq Cx$, with strict inequality in at least one component. The set of all pareto optimal solutions is called the *Efficient Frontier*, denoted by \mathcal{E} , and we use the phrases “ x is pareto optimal” and “ x is efficient” to mean that $x \in \mathcal{E}$. Practitioners often desire to generate, or sample, the entire efficient frontier. This is because a complete description of the efficient frontier shows a decision maker how the objective functions interact. Recent techniques for parameterizing the efficient frontier are described by Dennis and Das [21]. We make use of the well known result that a feasible element is efficient if, and only if, the element is optimal for a weighted sum of the objectives [7]. In other words, $x \in \mathcal{E}$ is the same as knowing there exists $w \in \mathbb{R}_{++}^p$ such that x solves $\min\{w^T Cx : x \in \mathcal{P}\}$.

Another standard manner to optimize when there are multiple objectives is to lexicographically order the objectives. Such an ordering scheme makes sense when the objectives are easily ranked. For example, a modeler may prefer to minimize the function f , but only after function g is minimized. This is exactly the situation when lexicographic optimization is appropriate. A typical sequential technique for solving the lexicographic MOLP is discussed in section 5.

The overriding goal of this paper is to show how to ascertain whether or not a decision variable is zero in **every** optimal solution. Such information is important because it supplies a decision maker with a collection of variables that do not appear in any optimal solution. The case of pareto optimization is considered in Section 3. We begin by defining the MOLP optimal partition, and in Theorem 1 we relate this partition to the optimal partitions of the single objective linear programs. An algorithm that constructs the MOLP optimal partition in the special case of two objectives is developed. An example from the radiotherapy literature is developed in Section 4, and we show how the MOLP optimal partition is useful. In Section 5 we define the lexicographic optimal partition, and again use the radiotherapy model as an example of its usefulness. Section 2 reviews the basic ideas of a path-following interior point algorithm and shows how these algorithms provide a solution that induces the linear programming (LP) optimal partition. Readers familiar with interior point algorithms will feel comfortable skipping all but the last paragraph of this section.

2 A Brief Review of a Path Following Interior Point Algorithm

The remaining sections rely on the fact that a path-following interior point algorithm terminates with a solution that induces the optimal partition. As such, a basic understanding of a path-following interior point algorithm is now presented (complete algorithm developments are found in [25], [28], and [29]). Consider the linear program (LP) $\min\{c^T x : Ax = b, x \geq 0\}$ and its associated dual (LD) $\max\{b^T y : A^T y + s = c, s \geq 0\}$ (as in (MOLP), we assume that $A \in \mathbb{R}^{m \times n}$ has full row rank). The necessary and sufficient conditions for optimality are

$$Ax = b, \quad A^T y + s = c, \quad x \geq 0, \quad s \geq 0, \quad \text{and} \quad x_i s_i = 0, \quad i = 1, 2, \dots, n.$$

A fundamental concept in the theory of interior point algorithms is that of the *central path*, which is an infinitely smooth curve contained within the *strict interior* of the feasible region. (The strict interior of the primal feasible region is $\{x : Ax = b, x > 0\}$, and the strict interior of the dual feasible region is $\{(y, s) : A^T y + s = c, s > 0\}$.) The central path is formed by replacing the complementarity constraint $x_i s_i = 0$ with $x_i s_i = \mu$, for $i = 1, 2, \dots, n$ and $\mu > 0$. Notice that μ being positive means that we require feasible elements such that $x > 0$ and $s > 0$. Hence, we make the assumption that the strict interiors of the primal and the dual are nonempty. For $\mu > 0$, the central path is defined by

$$Ax = b, \quad A^T y + s = c, \quad x > 0, \quad s > 0, \quad \text{and} \quad x_i s_i = \mu, \quad i = 1, 2, \dots, n.$$

The full row rank of A implies that this system has a unique solution for every $\mu > 0$, which we denote $(x(\mu), y(\mu), s(\mu))$. Notationally, the *primal central path* is the parameterization $x(\mu)$, and the *dual central path* is the parameterization $(y(\mu), s(\mu))$. As $\mu \downarrow 0$, the central path converges to a unique optimal solution known as the *analytic center solution*, denoted (x^*, y^*, s^*) . A path-following interior point algorithm follows the central path towards optimality, and hence, terminates at the analytic center solution.

Unless there is a unique solution to (LP) , (x^*, y^*, s^*) is not basic. Instead, the analytic center solution is *strictly complementary*, meaning that $x_i^* s_i^* = 0$ and $x_i^* + s_i^* > 0$, for $i = 1, 2, \dots, n$. Strictly complementary solutions have been known to exist since 1956 [11], but until path following interior point algorithms were implemented, these solutions were not easily calculated [32]. Any strictly complementary solution, and in particular the analytic center solution, induces the optimal partition, which is written $(B|N)$ and is defined by

$$B = \{i : x_i^* > 0\} \quad \text{and} \quad N = \{1, 2, \dots, n\} \setminus B.$$

These sets are important because they indicate the variables that are allowed to be positive at optimality (those in B) and the variables that are forced to zero at optimality (those in N). Making the convention that a set subscript on a vector (or a matrix) indicates the subvector (or the submatrix) that corresponds to the components (or the columns) contained within the set, we have that the optimal set for (LP) is $\{x : A_B x_B = b, x_B \geq 0, x_N = 0\}$ and the optimal set for (LD) is $\{(y, s) : A_B^T y = c_B, A_N^T y + s_N = c_N, s_N \geq 0, s_B = 0\}$. Hence, the optimal partition characterizes the optimal sets. The reason that (x^*, y^*, s^*) is called the *analytic center* solution is that x^* is the unique solution to

$$\max \left\{ \sum_{i \in B} \ln(x_i) : A_B x_B = b, x_B > 0, x_N = 0 \right\}$$

and (y^*, s^*) is the unique solution to

$$\max \left\{ \sum_{i \in N} \ln(s_i) : A_B^T y = c_B, A_N^T y + s_N = c_N, s_N > 0 \right\}.$$

In general, the analytic center of a bounded polyhedron, say $\{x : Ux = u, x \geq 0\}$, is the unique solution to

$$\max \left\{ \sum_{i \in I} \ln(x_i) : Ux = u, x_I > 0, x_{I^c} = 0 \right\},$$

where I indexes the components of x that are allowed to be positive, and I^c is the complement of I .

Having an algebraic description of the optimal face is important because it provides a vehicle from which to base a complete sensitivity analysis. Before interior point algorithms became viable alternatives to simplex based procedures, sensitivity analysis was conducted with a basic optimal solution. However, in the presence of degeneracy there are problems with a basic optimal approach to sensitivity analysis. This is because the analysis depends on the specific basic optimal solution provided by the algorithm. In the presence of degeneracy, the analysis provided depends on the basic optimal solution produced by the specific algorithm. However, the optimal partition is unique, and several researchers have investigated how to base post-optimal solution analysis on the optimal partition [2, 12, 14, 15, 18, 19, 20, 23, 24, 25, 30, 31].

Before continuing, we require some brief notes on notation. For $i = 1, 2, \dots, p$, the optimal partition of the single objective linear program $\min\{c^i x : x \in \mathcal{P}\}$ is $(B^i|N^i)$, and hence,

$$\mathcal{P}_i^* = \{x : A_{B^i} x_{B^i} = b, x_{B^i} \geq 0, x_{N^i} = 0\} = \{x \in \mathcal{P} : x_{N^i} = 0\}.$$

The *strict interior* of \mathcal{P}_i^* is $(\mathcal{P}_i^*)^\circ = \{x \in \mathcal{P}_i^* : x_{B^i} > 0\}$.

3 The Optimal Partition for Multiple Objective Linear Programming

In this Section we consider the case of pareto optimization. We begin by defining an analog of the LP optimal partition, which we succinctly refer to as the MOLP optimal partition. Moreover, using parametric programming results, we provide an algorithm that produces the MOLP optimal partition for bicriteria optimization. We desire that the MOLP optimal partition provides information about the efficient frontier that is similar to the information provided by the LP optimal partition. With this in mind, we make the following definition.

Definition 1 Let \mathcal{E} be the efficient frontier of $\min\{Cx : Ax = b, x \geq 0\}$. The MOLP optimal partition, denoted $(B | N)^{molp}$, is defined by

$$\begin{aligned} N^{molp} &= \{i : x_i = 0 \text{ for all } x \in \mathcal{E}\} \text{ and} \\ B^{molp} &= \{1, 2, 3, \dots, n\} \setminus N. \end{aligned}$$

The definition of the MOLP optimal partition retains the quality that an index being in N indicates that the component is zero in every pareto optimal solution. Likewise, an index in B demonstrates that the component is allowed to be positive on the efficient frontier. A property that is unfortunately relinquished is that the MOLP optimal partition is not capable of characterizing the efficient frontier —i.e. $\mathcal{E} \neq \{x \in \mathcal{P} : x_{N^{molp}} = 0\}$. However, we do have that $\mathcal{E} \subseteq \{x \in \mathcal{P} : x_{N^{molp}} = 0\}$, where equality holds in the case of a single objective (and hence the MOLP optimal partition reduces to the LP optimal partition).

We point out that B^{molp} is not the union of the the B^i 's, and that in general there is no simple relationship between B^{molp} and the B^i 's. This is because the efficient frontier may contain faces that are not optimal for any of the single objectives, and the single objective optimal faces need not be contained within the efficient frontier. With that said, the next theorem provides relationships between the MOLP optimal partition and the single objective optimal partitions.

Theorem 1 We have that

$$\mathcal{E} \supseteq \mathcal{P}_i^*, \text{ for } i = 1, 2, \dots, p, \Leftrightarrow B^{molp} \supseteq \bigcup_{i=1}^p B^i \text{ and } N^{molp} \subseteq \bigcap_{i=1}^p N^i,$$

and for any fixed j between 1 and p ,

$$\mathcal{E} \subseteq \mathcal{P}_j^* \Leftrightarrow B^{molp} \subseteq B^j \text{ and } N^{molp} \supseteq N^j.$$

Moreover,

$$\mathcal{E} = \bigcup_{i=1}^p \mathcal{P}_i^* \Rightarrow B^{molp} = \bigcup_{i=1}^p B^i \text{ and } N^{molp} = \bigcap_{i=1}^p N^i.$$

Proof: Suppose that $\mathcal{E} \supseteq \mathcal{P}_i^*$, for $i = 1, 2, \dots, p$. Let $j \in \bigcup_{i=1}^p B^i$. Then $j \in B_k$ for some k , and there exists $x \in \mathcal{P}_k^* \subseteq \mathcal{E}$ such that $x_j > 0$. Hence $j \in B^{molp}$, and the sufficiency of the first statement is established.

Now assume that there exists j such that $\mathcal{P}_j^* \not\subseteq \mathcal{E}$. We first show that $(\mathcal{P}_j^*)^o \cap \mathcal{E} = \emptyset$. Suppose for the sake of attaining a contradiction that $x \in (\mathcal{P}_j^*)^o \cap \mathcal{E}$. Since $x \in \mathcal{E}$, there exists $w \in \mathbb{R}_{++}^p$ such that x solves $\min\{w^T Cx : Ax = b, x \geq 0\}$. Let $(B' | N')$ be the optimal partition of this linear program. Then $\mathcal{O} = \{x \in \mathcal{P} : A_{B'}x_{B'} = b, x_{B'} \geq 0, x_{N'} = 0\}$ is contained within \mathcal{E} . Since $x \in (\mathcal{P}_j^*)^o$, we have that $x_{B_j} > 0$, and since $x \in \mathcal{O}$, the only components of x allowed to be positive are indexed by B' . So, $B^j \subseteq B'$. We now have the contradiction that

$$\begin{aligned} \mathcal{P}_j^* &= \{x \in \mathcal{P} : A_{B_j}x_{B_j} = b, x_{B_j} \geq 0, x_{N_j} = 0\} \\ &\subseteq \{x \in \mathcal{P} : A_{B'}x_{B'} = b, x_{B'} \geq 0, x_{N'} = 0\} \\ &\subseteq \mathcal{E}. \end{aligned}$$

So, $\mathcal{P}_j^* \not\subseteq \mathcal{E}$ implies $(\mathcal{P}_j^*)^o \cap \mathcal{E} = \emptyset$, and hence there exists $x \in (\mathcal{P}_j^*)^o \setminus \mathcal{E}$. Subsequently, this means that $B^j \not\subseteq B^{molp}$, and the necessity of the first statement is proven.

Let $\mathcal{E} \subseteq \mathcal{P}_j^*$ for some j between 1 and p . Then every efficient element is contained in

$$\mathcal{P}_j^* = \{x \in \mathcal{P} : A_{B^j} x_{B^j} = b, x_{B^j} \geq 0, x_{N^j} = 0\},$$

from which we have that $\overset{m.o.l.p.}{B} \subseteq B^j$. To establish the necessity of the second statement we assume that $\overset{m.o.l.p.}{B} \subseteq B^j$, for some j between 1 and p . Then,

$$\begin{aligned} \mathcal{E} &\subseteq \{x \in \mathcal{P} : A_{\overset{m.o.l.p.}{B}} x_{\overset{m.o.l.p.}{B}} = b, x_{\overset{m.o.l.p.}{B}} \geq 0, x_{N^j} = 0\} \\ &\subseteq \{x \in \mathcal{P} : A_{B^j} x_{B^j} = b, x_{B^j} \geq 0, x_{N^j} = 0\} \\ &= \mathcal{P}_j^*. \end{aligned}$$

To prove the third implication assume that $\mathcal{E} = \bigcup_{i=1}^p \mathcal{P}_i^*$. Let $k \in \overset{m.o.l.p.}{B}$. Then, there exists $x \in \mathcal{E}$ such that $x_k > 0$. Since $x \in \mathcal{P}_j^*$, for some j , and $x_k > 0$, $k \in B^j$. So, $\overset{m.o.l.p.}{B} \subseteq \bigcup_{i=1}^p B^i$. Now let $k \in \bigcup_{i=1}^p B^i$. Then, $k \in B^j$ for some j , which implies there exists x in \mathcal{P}_j^* such that $x_k > 0$. Since $\mathcal{P}_j^* \subseteq \mathcal{E}$, we subsequently have that $k \in \overset{m.o.l.p.}{B}$. So, $\bigcup_{i=1}^p B^i \subseteq \overset{m.o.l.p.}{B}$, and the proof is complete. \blacksquare

The last statement in Theorem 1 is not reversible as demonstrated by the following example.

Example 1 Consider the following multiple objective linear program,

$$\min \left\{ \begin{pmatrix} x_1 \\ x_2 \\ -x_1 - x_2 \end{pmatrix} : x_1 + x_2 + x_3 = 1, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \right\}.$$

The single objective optimal partitions are

$$(B^1|N^1) = (\{2, 3\}|\{1\}), \quad (B^2|N^2) = (\{1, 3\}|\{2\}), \quad \text{and} \quad (B^3|N^3) = (\{1, 2\}|\{3\}).$$

Notice that if we add the objectives we have the linear program

$$\min\{0x_1 + 0x_2 + 0x_3 : x_1 + x_2 + x_3 = 1, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0\},$$

which means that every feasible element is pareto optimal. So $\mathcal{E} = \mathcal{P}$, and because $\mathcal{P}^o \neq \emptyset$, we have that $\overset{m.o.l.p.}{B} = \{1, 2, 3\}$, which is equal to $\bigcup_{i=1}^3 B^i$. However,

$$\begin{aligned} \mathcal{E} &= \mathcal{P} \\ &\neq \mathcal{P}_1^* \cup \mathcal{P}_2^* \cup \mathcal{P}_3^* \\ &= \{x \in \mathcal{P} : x_1 = 0\} \cup \{x \in \mathcal{P} : x_2 = 0\} \cup \{x \in \mathcal{P} : x_3 = 0\}. \end{aligned}$$

The geometry of this problem is depicted in Figure 1.

Theorem 1 indicates that if the efficient frontier is comprised entirely of the single objective optimal faces, $(\overset{m.o.l.p.}{B} | \overset{m.o.l.p.}{N})$ may be calculated by solving the p single objective linear programs. However, the sufficient condition that $\mathcal{E} = \bigcup_{i=1}^p \mathcal{P}_i^*$ is not practical because the nature of the efficient frontier would have to be known apriori.

This leads to the question of when is it possible to generate $(\overset{m.o.l.p.}{B} | \overset{m.o.l.p.}{N})$. In the special case of $p = 2$ —i.e. bicriteria optimization, we develop an algorithm to produce the MOLP optimal partition. For the remainder of this section we assume that $p = 2$. As previously mentioned, we have that $x \in \mathcal{E}$ is the same as x solving

$$(LP_w) \quad \min\{(w_1 c^1 + w_2 c^2)x : x \in \mathcal{P}\},$$

for some $w_1 > 0$ and $w_2 > 0$. An observation first made by Geoffrion [10] is that the solutions to (LP_w) correspond to the solutions of

$$(LP_\theta) \quad \min\{(c^1 + \theta \delta c)x : x \in \mathcal{P}\},$$

where $\theta \in (0, 1)$ and $\delta c = c^2 - c^1$. So, x being optimal to (LP_w) , where $w_1 > 0$ and $w_2 > 0$, implies there is $\theta \in (0, 1)$ such that x is optimal to (LP_θ) . Similarly, if x is optimal to (LP_θ) for $\theta \in (0, 1)$, there exists $w_1 > 0$ and $w_2 > 0$ such that x is optimal to (LP_w) . This means that the efficient frontier may be parameterized for bicriteria optimization, and hence, $(\overset{m.o.l.p.}{B} | \overset{m.o.l.p.}{N})$ may be found by calculating the optimal partitions as θ traverses the open interval $(0, 1)$.

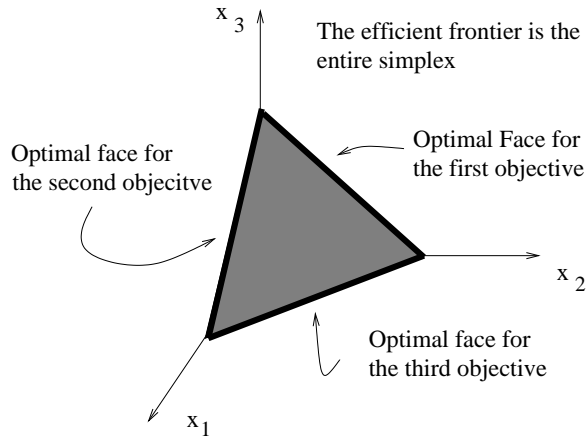


Figure 1: Every feasible element is pareto optimal, but the union of the single objective optimal faces is the boundary of the feasible region.

Parametric linear programming has been studied since the middle 1950s [9, 26], and the classical question asks how far the data describing the linear program can be perturbed so that the current optimal basis remains optimal. However, the popularity of interior point algorithms has lead math programmers to re-think the topic of parametric analysis, and we now understand that the analysis available from the optimal partition is substantially more robust than the analysis allowed by a basic optimal solution [20]. The work of Monteiro and Mehrotra [23] and Roos, Terlaky, and Vial [25] show how to generate the optimal partition under parameterization of either the objective coefficients or the right-hand side vector, and we use these algorithms to calculate $(B^{molp} | N^{molp})$. (also see [14] for simultaneous parameterization). We review how these algorithms work, but interested readers should see the original works for complete details.

The algorithms in [23] and [25] completely explain the classic result that $z^*(\theta) = \min\{(c^1 + \theta\delta c)^T x : x \in \mathcal{P}\}$ is a piecewise linear function ($z^*(\theta)$ is also concave) [8]. What the authors of [23] and [25] noticed was that the linearity intervals of $z^*(\theta)$ correspond with an optimal partition. For example, suppose that we want to parameterize between c^1 and c^2 , and that upon setting $\delta c = c^2 - c^1$, we find that $z^*(\theta)$ is the function depicted in Figure 2. The optimal objective value, z^* , traverses three linearity intervals as θ increases from 0 to θ_3 , and

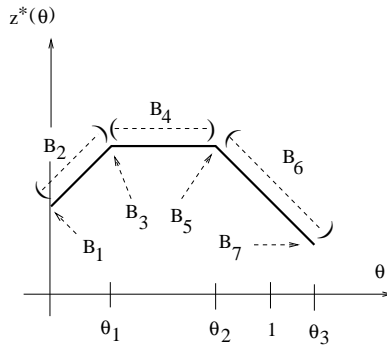


Figure 2: The optimal objective function as θ traverses the interval $[0, \theta_3]$

seven optimal partitions are observed. (Note: We use $(B_i | N_i)$ for the example because $(B^i | N^i)$ is the optimal partition for the i^{th} single objective linear program.) For $\theta = 0$ (the unperturbed problem), the components of an optimal solution that are allowed to be positive are indexed by B^1 . However, as θ increases from zero, the optimal partition changes immediately, and a new optimal partition is defined for $0 < \theta < \theta_1$. Similarly, as θ attains the value of θ_1 , the optimal partition changes to (B_3, N_3) , but only observes this partition instantaneously because (B_4, N_4) is the optimal partition for $\theta_1 < \theta < \theta_2$. This analysis continues until θ increases past 1, which in this case is where θ equals θ_3 and the optimal partition is $(B_7 | N_7)$. The linearity of z^* changes at 0, θ_1 , θ_2 ,

and θ_3 , and these values are called the *break points* of z^* .

The algorithm to produce z^* relies on two calculations. The first calculation figures out how much θ can increase from zero until the optimal partition of the unperturbed problem is forced to change. Let $(B|N)$ be the original optimal partition, and observe that this partition remains optimal so long as the following equations are consistent,

$$\begin{aligned} \text{Primal} &\rightarrow \left\{ \begin{array}{l} A_B x_B = b, x_B > 0 \\ \end{array} \right. && (\text{—i.e. } x_N = 0) \\ \text{Dual} &\rightarrow \left\{ \begin{array}{l} A_B^T y = c_B + \theta \hat{\alpha}_B, \\ A_N^T y + s_N = c_N + \theta \hat{\alpha}_N, s_N > 0. \end{array} \right. && (\text{—i.e. } s_B = 0) \end{aligned}$$

Changing θ does not alter the primal conditions. So, to calculate the range of θ over which the optimal partition remains optimal we solve

$$(LP_{Range}) \max \left\{ \theta : A_B^T y = c_B + \theta \hat{\alpha}_B, A_N^T y + s_N = c_N + \theta \hat{\alpha}_N, s_N \geq 0, \theta \geq 0 \right\}.$$

We assume for the moment that (LP_{Range}) has a finite optimal value, and we let (θ^*, y^*, s_N^*) be the analytic center solution. There are two cases to consider. If $\hat{\alpha}_B \notin \text{col}(A_B^T)$, the only value of θ for which $A_B^T y = c_B + \theta \hat{\alpha}_B$ is consistent is $\theta = 0$. So, in this case the optimal value is 0, and $s_N^* > 0$. Otherwise, $\hat{\alpha}_B \in \text{col}(A_B^T)$ and $A_B^T y = c_B + \theta \hat{\alpha}_B$ is consistent for every θ . In this situation, the optimal value is positive, and components of s_N^* are zero. Notice that the value of θ^* is a break point of z^* . The linear program (LP_{Range}) has its own optimal partition, which induces the new optimal partition at the break point. If $\theta^* = 0$, as it is in our example, s_N^* is positive and the optimal partition has not changed. If $\theta^* > 0$, there is a non-empty subcollection of N , say \tilde{N} , such that $s_{\tilde{N}} = 0$. The indices in \tilde{N} switch from N to B so that the optimal partition at the break point is $(B'|N') = (B \cup \tilde{N} | N \setminus \tilde{N})$.

The second calculation is to find the rate at which z^* changes over the next linearity interval. How to calculate the rate of change of z^* along $\hat{\alpha}$ is a question that was first answered by Mills [22] in 1956, where he showed that the directional derivative of $z^*(\theta^*)$ is

$$\min \left\{ \hat{\alpha}^T x : x \in \text{argmin}\{(c + \theta^* \hat{\alpha})^T x : x \in \mathcal{P}\} \right\}$$

Because $(B'|N')$ is the optimal partition for $\min\{(c + \theta^* \hat{\alpha})^T x : x \in \mathcal{P}\}$, we have that

$$\text{argmin}\{(c + \theta^* \hat{\alpha})^T x : x \in \mathcal{P}\} = \{x \in \mathcal{P} : A_{B'} x_{B'} = b, x_{B'}, x_{N'} = 0\}.$$

From this we see that the rate of change of z^* over the next linearity interval is

$$(LP_{Rate}) \min \left\{ \hat{\alpha}_{B'}^T x_{B'} : A_{B'} x_{B'} = b, x_{B'} \geq 0 \right\}.$$

As before, (LP_{Rate}) has its own optimal partition, which is used to form the optimal partition for the next linearity interval. If $x_{B'}^*$ is the analytic center solution, the indices of the positive components remain in B , while the indices of the zero components move to N . Let \tilde{B} be the subset of B' such that $x_{\tilde{B}}^* > 0$ and $x_{B' \setminus \tilde{B}} = 0$. Then, the optimal partition for the next linearity interval is $(B' \setminus \tilde{B} | N' \cup \tilde{B})$. The algorithm for the example illustrated in Figure 3 proceeds as follows.

Algorithm to Calculate the Optimal Partitions as θ traverses the interval $[0, \theta_3]$

1. Find the amount of movement along δc that $(B_1|N_1)$ can observe by solving

$$\begin{aligned} \max \theta : \quad & A_{B_1}^T y = c_{B_1} + \theta \delta c_{B_1} \\ & A_{N_1}^T y + s_{N_1} = c_{N_1} + \theta \delta c_{N_1} \\ & s_{N_1} \geq 0. \end{aligned}$$

For the example the solution is 0. If any components of s_{N_1} had been forced to zero at optimality, the indices for these components would have been added to B_1 to form the partition at the break point. However, $\theta = 0$ is a break point as indicated by the optimal value being zero, and hence none of the components of s_{N_1} are forced to zero at optimality.

2. Find the rate of change for $z^*(\theta)$ over the next linearity interval by solving

$$\begin{aligned} \min \delta c_{B_1}^T x_{B_1} : \quad & A_{B_1} x_{B_1} = b \\ & x_{B_1} \geq 0. \end{aligned}$$

The solution is the slope of the first linearity interval, and the partition corresponding to this interval is

$$B_2 = \{i : x_i^* > 0, i \in B_1\} \text{ and } N_2 = N_1 \cup \{i : x_i^* = 0, i \in B_1\},$$

where $x_{B_1}^$ is the analytic center solution.*

3. Find out how much movement $(B_2|N_2)$ can accommodate by solving

$$\begin{aligned} \max \theta : \quad & A_{B_2}^T y = c_{B_2} + \theta \delta c_{B_2} \\ & A_{N_2}^T y + s_{N_2} = c_{N_2} + \theta \delta c_{N_2} \\ & s_{N_2} \geq 0. \end{aligned}$$

The solution is θ_1 , and the optimal partition for this break point is

$$B_3 = B_2 \cup \{i : s_i^* = 0, i \in N_2\} \text{ and } N_3 = \{i : s_i^* > 0, i \in N_2\},$$

where $(\theta^, y^*, s_{N_2}^*)$ is the analytic center solution.*

4. Find the rate of change for $z^*(\theta)$ over the next linearity interval by solving

$$\begin{aligned} \min \delta c_{B_3}^T x_{B_3} : \quad & A_{B_3} x_{B_3} = b \\ & x_{B_3} \geq 0. \end{aligned}$$

The optimal value is the slope for the second linearity interval, and the optimal partition for this interval is

$$B_4 = \{i : x_i^* > 0, i \in B_3\} \text{ and } N_4 = N_3 \cup \{i : x_i^* = 0, i \in B_3\},$$

where $x_{B_3}^$ is the analytic center solution.*

5. Continue in a like fashion.

Once the optimal partitions are calculated, we use the result that the efficient frontier is parameterized as θ traverses the interval $(0, 1)$. Notice that $(B_1|N_1)$ is optimal for $\theta = 0$ only, which means that $B_1 \not\subseteq^{m.otp} B$. Similarly, $(B_7|N_7)$ is optimal for $\theta = \theta_3$, which is greater than 1. Hence, B_7 is not a subset of B . The remaining optimal partitions are optimal for some θ in $(0, 1)$, and hence,

$$B^{m.otp} = \bigcup_{i=2}^6 B_i \text{ and } N^{m.otp} = \{1, 2, \dots, n\} \setminus B^{m.otp}.$$

Notice that there is a subset relationship between the optimal partition at a break point, say $(B^3|N^3)$ at θ_1 , and the optimal partitions of the adjoining linearity intervals, $(B^2|N^2)$ and $(B^4|N^4)$. Specifically, we have $B^3 \supseteq B^2$ and $B^3 \supseteq B^4$. Such relationships hold in general, and we formalize this by defining the point-to-set map $(B(\theta)|N(\theta))$ to be the optimal partition for the objective coefficients $c^1 + \theta \delta c$. The following monotonicity property was first

proven by Adler and Monteiro [2], but is also found in [4] and [16], where the latter established the result for nonlinear perturbations.

Lemma 1 (Adler and Monteiro [2]) *For sufficiently small $\theta > 0$, we have*

$$N(0) \subseteq N(\theta) \text{ and } B(\theta) \subseteq B(0).$$

Lemma 1 is useful because it shows that we do not need every optimal partition to construct B^{mopt} . Instead, we only need the optimal partitions that correspond to the break points of z^* . For the example depicted in Figure 2, we have that $B^{mopt} = B_3 \cup B_5$.

The following algorithm uses Lemma 1 to construct B^{mopt} . The concept is the same as in the example, calculate a range of θ over which an optimal partition remains optimal, calculate the rate of change for the next linearity interval, and repeat. From Lemma 1 we have that indices need to be added to B^{mopt} only after the range calculation.

Algorithm to Calculate $(B^{mopt} | N^{mopt})$ for Bicriteria Optimization

Step 1 - Initial Solve: Initialize $s = 0$, $\Theta_0 = 0$, and $B_0 = \emptyset$. Set c to be c^1 and δc to be $c^2 - c^1$. Solve $\min\{cx : Ax = b, x \geq 0\}$ with a path-following interior point algorithm to get $(B^1 | N^1)$. Set $(B | N)$ to be $(B^1 | N^1)$.

Step 2 - Calculate Range: Calculate

$$\theta_s = \max\{\theta : A_B^T y_B = c_B + \theta \delta c_B, A_N^T y_N + s_N = c_N + \theta \delta c_N, s_N \geq 0, 0 \leq \theta \leq 2\},$$

Let (θ^*, y^*, s_N^*) be the analytic center solution (or any strictly complementary solution), and set

$$\begin{aligned} B' &= B \cup \{i : i \in N, s_i^* = 0\} \text{ and} \\ N' &= \{i : i \in N, s_i^* > 0\}. \end{aligned}$$

Set $\Theta_{s+1} = \Theta_s + \theta_s$ (notice the difference between θ and Θ).

Step 3 - Build B^{mopt} :

- a. If $s = 0$ and $\Theta_{s+1} = 0$, proceed to step 4.
- b. If $s = 0$ and $0 < \Theta_{s+1} < 1$, set $B_{s+1}^{mopt} = B_s^{mopt} \cup B'$.
- c. If $s = 0$ and $\Theta_{s+1} \geq 1$, set $B = B^1$ and go to step 6.
- d. If $s \geq 1$ and $0 < \Theta_{s+1} < 1$, set $B_{s+1}^{mopt} = B_s^{mopt} \cup B'$.
- e. If $s \geq 1$ and $\Theta \geq 1$, go to step 6.

Step 4 - Update Step Index: Index s up by one.

Step 5 - Calculate Rate of Change: Solve

$$\min\{\delta c_{B'} x_{B'} : A_{B'} x_{B'} = b, x_{B'} \geq 0\}$$

and set

$$\begin{aligned} B &= \{i : i \in B', x_i^* > 0\} \text{ and} \\ N &= N' \cup \{i : i \in B', x_i^* = 0\}, \end{aligned}$$

where $x_{B'}^*$ is the analytic center solution (or any strictly complementary solution). Go to step 2.

Step 6 - Stop: Set

$$B^{mopt} = B_{s+1}^{mopt} \text{ and } N^{mopt} = \{1, 2, 3, \dots, n\} \setminus B^{mopt}.$$

The optimization problem in Step 2 calculates the range of θ for which the current optimal partition remains optimal (and corresponds to a linearity interval of z^*). Notice that we bound θ above by 2, which guarantees that a linearity interval does not continue as $\theta \rightarrow \infty$.

The conditional statements in 3(a) through 3(c) handle the possible situations during the first iteration. If the first range calculation returns 0, then $(B^1|N^1)$ is optimal for $\theta = 0$ only, and hence $B^1 \not\subseteq^{m\text{olp}} B$. In this case we do not add any components to B , and we continue by finding the rate of change for the next linearity interval. If the first range calculation returns a value between 0 and 1, we have that $B^1 \subseteq^{m\text{olp}} B$ and that the new partition $(B'|N')$ is an optimal partition for a break point of z^* . Moreover, Lemma 1 implies $B^1 \subseteq B'$, and because the range calculation returned a value between 0 and 1, we have $B^1 \subseteq B' \subset^{m\text{olp}} B$. So, instead of adding the indices in B^1 to B , we add the indices of B' to B . Finally, if the first range calculation returns a value of at least 1, then $(B^1|N^1)$ is the unique optimal partition for every θ in $(0, 1)$. In this case we have that $(B|N)^{m\text{olp}} = (B^1|N^1)$. The last two conditional statements in Step 3 continue to add components to B so long as $\Theta < 1$. Again, from Lemma 1 we only add indices to B after we calculate an optimal partition corresponding to a break point of z^* .

4 An Example from Radiotherapy Treatment Design

The purpose of this section is to show that the MOLP optimal partition is useful in the design of radiotherapy treatment plans. We use a modification of the model in [17]. External beam radiotherapy design is concerned with choosing a collection of positions around a patient so that radiation from an external source is focused on a tumorous region and not on other critical structures. Energy is produced in a linear accelerator and is directed through a gantry, which is capable of being placed continuously around the patient. In modern treatment facilities, the head of the gantry contains a multi-leaf collimator, which is used to “shape” the beam of energy. A *treatment plan* consists of (1) a set of angles to place the gantry, (2) the amount of radiation to be administered along each angle, and (3) the shape of the beam at each angle. Since the gantry is capable of moving continuously around the patient, and the shape of the beams is controlled by over 50 leaves in the collimator, designing plans is a complex procedure requiring some sort of automated aid.

The fact that the accumulative dose to a cell is additive means that linear models are natural and appropriate. Linear models rely on the fact that there is a linear transformation that takes radiation at the gantry and deposits this energy into the anatomy. To understand how this transformation is formed, we set $x_{(a,i)}$ to be the dose along the i^{th} sub-beam of angle a , and $d_{(p,a,i)}$ to be the distance from where sub-beam $x_{(a,i)}$ enters the image to where it reaches pixel p . We further define $A_{(p,a,i)}$ to be the product of $e^{-\mu d_{(p,a,i)}}$ and the geometric area common to both sub-beam $x_{(a,i)}$ and pixel p . The factor $e^{-\mu d_{(p,a,i)}}$ describes how the beam of radiation *attenuates* as it passes through the body, and μ is called the *attenuation coefficient*. The value of μ depends on the energy of the beam. Gamma-rays from Cobalt-60 have a relatively low energy of 1.25 MeV and correspond to a μ value of approximately 0.06. Linear accelerators in modern facilities are capable of producing beams with intensities of up to 15 MV, and typical values of μ are 0.02 [6]. To calculate $A_{(p,a,i)}$ for the pixel, angle, and sub-beam depicted in Figure 3, we first calculate that sub-beam i intersects 40% of pixel p . Second, we calculate the distance along sub-beam i from the edge of the patient to pixel p . Third, we set $A_{(a,p,i)} = 0.40e^{-\mu d_{(p,a,i)}}$, where μ is decided by the energy of the beam.

The accumulative dose to pixel p is found by adding the energy deposited in pixel p from every sub-beam. We “discretize” the angle positions so that the continuous rotation of the gantry is approximated by a set number of evenly spaced angles. So, the accumulative dose to pixel p is

$$\sum_{(a,i)} A_{(p,a,i)} x_{(a,i)}.$$

The dose deposition matrix, denoted by A , is formed by letting $A_{(p,a,i)}$ to be the element in the p^{th} row and the $(a,i)^{\text{th}}$ column. Allowing the vector x to be comprised of the $x_{(a,i)}$'s, where the order corresponds to the columns of A , we have that the p^{th} component of Ax is the total dose received by pixel p under plan x .

What happens in a clinical setting is that a dosimetrist highlights regions in a patient image that correspond to the tumor and the surrounding organs. Once the image is segmented into regions, a tumoricidal dose for the tumor is decided, and the pixels flagged as cancerous are asked to receive a tumoricidal level of radiation. At the same time, pixels corresponding to other structures in the anatomy are given bounds on the amount of radiation that they should receive. For example, suppose that pixel \bar{p} is within the tumor and that the tumoricidal dose is 80Gy. The linear constraint

$$\sum_{(a,i)} A_{(\bar{p},a,i)} x_{(a,i)} \geq 80$$

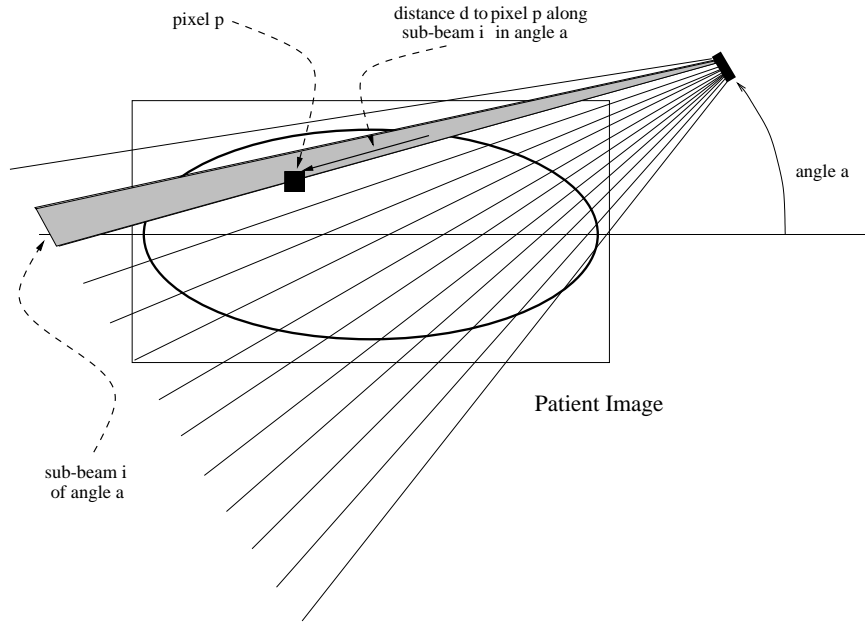


Figure 3: The geometry of treating a patient image along angle a .

guarantees that pixel \bar{p} receives at least 80Gy's. Similarly, if pixel \hat{p} is within the liver, and the liver's upper bound is 20Gy's, the constraint

$$\sum_{(a,i)} A_{(\hat{p},a,i)} x_{(a,i)} \leq 20$$

guarantees that pixel \hat{p} receives no more than 20Gy's.

The rows of the dose deposition matrix are partitioned and reordered into the rows that correspond to the tumor, the rows that correspond to the critical structures, and the rows that correspond to the remaining good tissue. This reordering is denoted with the sub-matrices A_T , A_C , and A_G , as indicated below,

$$A = \begin{bmatrix} A_T \\ A_C \\ A_G \end{bmatrix} \begin{array}{l} \leftarrow \text{Tumor} \\ \leftarrow \text{Critical Structures} \\ \leftarrow \text{Remaining Good Tissue.} \end{array}$$

Sub-beams that do not intersect the tumor are removed from consideration by eliminating the columns of A that have a corresponding zero column in A_T . For notational brevity, we keep the A notation for the sub-matrix with these columns removed. In what follows, $A \in \mathbb{R}^{m \times n}$, $A_T \in \mathbb{R}^{m_T \times n}$, $A_C \in \mathbb{R}^{m_C \times n}$, and $A_G \in \mathbb{R}^{m_G \times n}$.

A *prescription* is comprised of a physician's aspirations for the tumor and upper bounds for the non-tumorous tissue. Specifically, a prescription is the 4-tuple (TUB, TLB, CUB, GUB) , where

- TUB is a m_T vector of upper bounds for the tumor,
- TLB is a m_T vector of lower bounds for the tumor,
- CUB is a m_C vector of upper bounds for the the critical structures, and
- GUB is a m_G vector of upper bounds for the remaining good tissue.

Because a uniform, tumoricidal dose is to be delivered to the tumor, the lower and upper bounds for the tumor are taken to be a fixed percentage of the physician's goal for the tumor. Supposing that the physician's goal for a tumorous cell is TG , values for TUB_i and TLB_i are $(1 + \epsilon)TG$ and $(1 - \epsilon)TG$, respectively. Here, ϵ is a percentage of variation for the tumor dosage and is called the *tumor uniformity level*. Typical values of ϵ found in the literature range from 0.02 to 0.15. The vector GUB describes the highest amount of radiation that any single pixel is allowed, and in general no tissue should receive more than 10% of the tumor's desired dose. As such, we set $GUB_i = TG(1.10)$.

The model we investigate is

$$\min \begin{pmatrix} e^T \alpha \\ e^T \beta + e^T \gamma \end{pmatrix}$$

such that

$$\begin{aligned} TLB - \alpha &\leq A_T x \leq TUB \\ &A_C x \leq CUB + \beta \\ &A_G x \leq GUB + \gamma \\ 0 &\leq \alpha \leq TLB \\ -CUB &\leq \beta \\ 0 &\leq \gamma \\ 0 &\leq x. \end{aligned} \tag{1}$$

The first set of constraints attempts to restrict the tumor's dose to be within its prescribed lower and upper bounds. The discrepancy between the tumor's lower bounds and the amount of radiation deposited in the tumor is recorded in α . The reason that the tumor is not guaranteed to be within its limits is that α may be non-zero. If α equals its upper bound of TLB , no radiation is required to treat the tumor, and if α equals 0, the tumor is receiving its prescribed minimum dose. With this interpretation we see that the first objective is to minimize the accumulative amount the tumor is under its prescribed lower limit. The vectors β and γ have similar roles for the critical structures and the normal tissue. The difference between β and γ is their lower bounds. If β attains its lower bound of $-CUB$, we have a plan that deposits no radiation in the critical structures. The lower bound on γ is zero because plans are considered acceptable once the remaining areas receive less than 10% of the maximum tumor dose. The second objective minimizes the sum of the radiation on the critical structures and the discrepancy on the normal tissue.

The objectives "compete" for energy in the following manner. The desire to drive $e^T \alpha$ to zero removes low energy plans from consideration, but the competing goal of driving $e^T \beta + e^T \gamma$ to $-e^T CUB$ removes high energy plans from consideration. There may be plans which allow both $e^T \alpha = 0$ and $e^T \beta + e^T \gamma = -e^T CUB$. Such plans are called *unencumbered* because they deliver a uniform, tumoricidal dose without depositing radiation in the critical structures. There are advantages to having the analytic center solution when such plans are available [17].

Our goal is to show the value of constructing the MOLP optimal partition for this problem. We used Radiotherapy optimAl Design (*RAD*), a free software package built in Matlab[®] and available at <http://www.trinity.edu/aholder/research/oncology/rad.html>. The model used in *RAD* is identical to our model except that it uses the single objective $\omega \cdot e^T \alpha + e^T \beta + e^T \gamma$, where ω weights the importance of attaining a uniform, tumoricidal dose. The only alteration required was to change the objective to $\theta e^T \alpha + (1 - \theta)(e^T \beta + e^T \gamma)$ and implement the algorithm described in Section 3.

While Matlab[®] is a robust language, the path-following interior point algorithm implemented in the optimization tool-box had some numerical difficulties. We declared a component of an optimal solution to be zero if it was less than 10^{-3} , and the problem we encountered was detecting a change in the optimal partition after some range calculations (smaller tolerance levels, down to 10^{-16} , led to the same problem). We handled the problem in the following manner. If $\{i : i \in N, s_i^* = 0\} = \emptyset$ in Step 2, we set $\Theta_{s+1} = \Theta_s + 10^{-3}$ and calculated the optimal partition for $c^1 + \Theta_{s+1} \hat{c}$ —i.e. we advanced our progress through the interval by 10^{-3} . While we understand that we may miss some linearity intervals with this technique, digital computation and internal algorithm tolerances require that a threshold be used to decide if an optimal value is indeed zero. Experience shows that using 10^{-3} provides stable, reliable results.

All experiments were conducted on a 64×64 grid, with 72 equally spaced angles to choose from. Each angle was comprised of 32 sub-beams, and the tumor uniformity level was $\pm 4\%$.

Consider the geometry in Figure 4, where a tumor is surrounded by three critical structures. We note that it is impossible to achieve a uniform, tumoricidal dose without violating some of the other dose restrictions. The algorithm to calculate $(B | N)^{m.o.p.}$ found 194 break points. This corresponds to 399 different partitions and 195 linearity intervals, with the longest being (0.709, 0.999) and the shortest having a length of 0.001 (because of the numerical instability previously discussed). We observed that the optimal partition became increasingly stable as importance on treating the tumor increased. Figures 5, 6, and 7 depict how the radiation is deposited by the optimal plan for the first, middle, and last break points: $\theta = 0.001$, $\theta = 0.397$, and $\theta = 0.999$. In Figure 5 we see that the critical structures receive no radiation and that there are large portions of the tumor that are not treated. The entire tumor is treated in the deposition pattern in Figure 6, but at the expense of the critical structures receiving an increased dose. Also, the uniformity level is not within its prescribed deviation allowance of $\pm 4\%$. In Figure 7 we have that the tumor is completely treated within its uniformity level. However, the lower part of

the critical structure on the lower left side of the tumor is receiving levels of radiation above its prescribed upper bound of 60 Gy's. Similar discrepancies occur for the upper left region of the critical structure on the upper right side of the tumor. Interestingly, the critical structure restricted to 20 Gy's always adheres to its prescribed upper bound.

Figures 8 and 9 provide information about which angles are used, and not used, as θ traverses the interval $(0, 1)$. We saved the optimal partitions at each break point, and used this information to calculate how often an angle is used. We say that an angle is *used at level k* if there are k sub-beams from that angle with positive amounts of radiation in the optimal plan. Furthermore, we say that an angle is *on* provided its level of use is at least 1. In Figure 8 we calculated each angle's level of use, and then added these together for each of the 194 break points. These values are recorded above the circle around the image. The highest peak is at 90° and has a value of 1,424, which means that 1,424 sub-beams were used from this angle as θ traversed the interval $(0, 1)$. Figure 9 is similar, but instead of accumulating sub-beams from each angle, the percentage of times an angle is on is displayed over the circle. Angle 85° was on in 100% of the optimal partitions, and while 90° had the highest amount of sub-beam usage, it was not on in each optimal partition (it was used in 99.75% of the optimal partitions).

The graphics in Figures 8 and 9 indicates which sub-beams, and hence angles, are in B^{mop} and N^{mop} . In either graph, an angle having a value of zero implies that the sub-beams from that angle were not used in **any** optimal plan. Such information is important because this collection of angles can be excluded from consideration when a dosimetrist is uncertain about the relative importance of the objectives. So, no matter what importance is placed on attaining a uniform, tumoricidal dose, the angles in N^{mop} are never used. Moreover, Figures 8 and 9 indicate to a dosimetrist which angles are used in a large number of optimal plans, and hence, should be considered under most circumstances.

5 An Optimal Partition for Lexicographic Multiple Objective Linear Programs

In this section we develop an optimal partition when the objectives in $(MOLP)$ are lexicographically ordered. As is consistent with our notation, we denote the optimal set by \mathcal{P}_L^* . The standard manner to solve a lexicographic MOLP is the following sequential technique.

A traditional solution procedure for lexicographic multiple-objective linear programming

- calculate $z_1^* = \min\{c^1 x : x \in \mathcal{P}\}$,
- calculate $z_2^* = \min\{c^2 x : x \in \mathcal{P}, c^1 x = z_1^*\}$,
- calculate $z_3^* = \min\{c^3 x : x \in \mathcal{P}, c^1 x = z_1^*, c^2 x = z_2^*\}$,
- \vdots
- calculate $z_p^* = \min\{c^p x : x \in \mathcal{P}, c^1 x = z_1^*, c^2 x = z_2^*, \dots, c^{p-1} x = z_{p-1}^*\}$.

The value of z_p^* is the optimal value of the lexicographic MOLP, and the optimal set is $\mathcal{P}_L^* = \{x : x \in \mathcal{P}, c^1 x = z_1^*, c^2 x = z_2^*, \dots, c^p x = z_p^*\}$. To see how the optimal partition is useful we need only recognize that

$$\mathcal{P}_1^* = \{x : x \in \mathcal{P}, c^1 x = z_1^*\} = \{x : A_{B^1} x_{B^1} = b, x_{B^1} \geq 0, x_{N^1} = 0\},$$

Hence, to calculate z_2^* we may instead solve

$$(LP_2) \quad \min\{c_{B^1}^2 x_{B^1} : A_{B^1} x_{B^1} = b, x_{B^1} \geq 0\}.$$

The difference between solving (LP_2) and the traditional approach is that instead of adding a constraint to the original problem, we remove a collection of variables from consideration. This means that the problem size decreases instead of increases.

The linear program (LP_2) has its own optimal partition, which we denote by $(B_L^2 | N_L^2)$. We use B_L^2 to calculate z_3^* exactly as we used B^1 to calculate z_2^* . The procedure continues as follows.

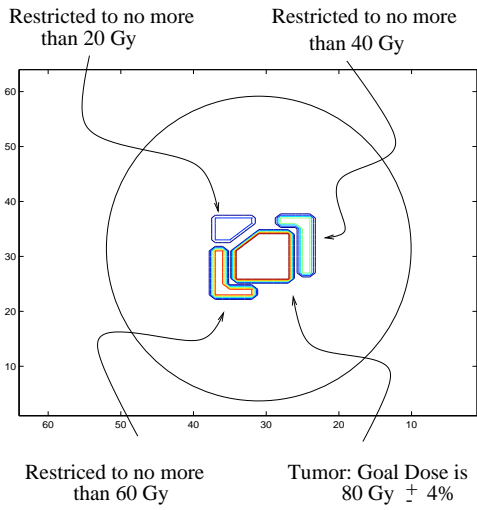


Figure 4: A tumor surrounded by 3 critical structures.

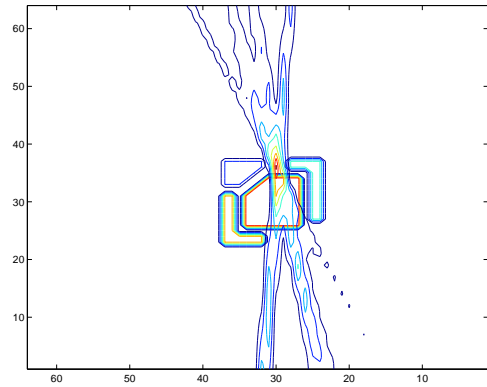


Figure 5: A contour plot for the break point $\theta = 0.001$. Neither the tumor or the critical structures receive much radiation.

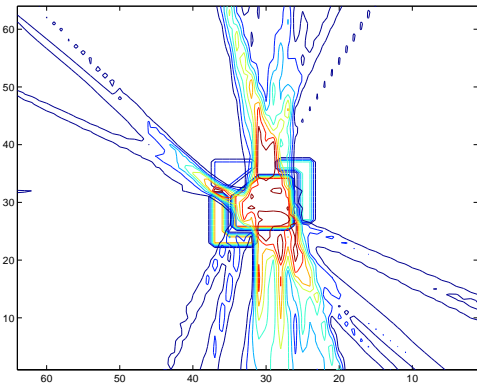


Figure 6: A contour plot for the break point $\theta = 0.397$. The tumor receives a more uniform dose, but at the expense of the critical structures receiving increased radiation.

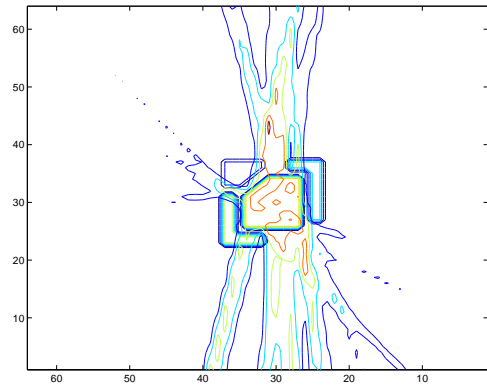


Figure 7: A contour plot for the break point $\theta = 0.999$. The tumor receives a uniform dose, but portions of the critical structures receive high levels of radiation.

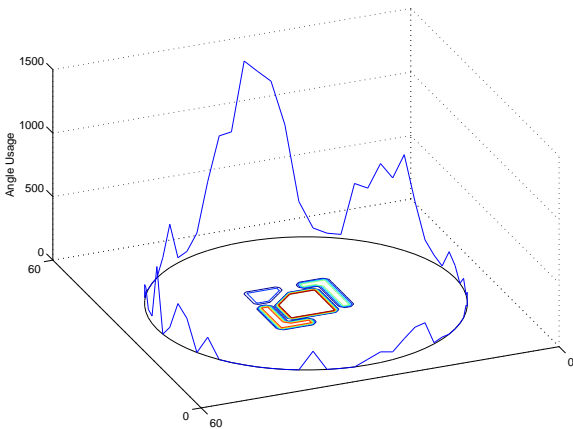


Figure 8: Accumulative totals of sub-beam usage along each angle.

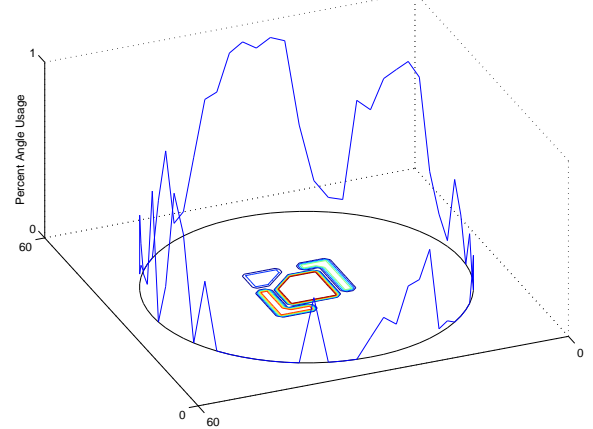


Figure 9: Percent totals for each angle.

- Calculate $z_1^* = \min\{c^1 x : x \in \mathcal{P}\}$, and set $(B^1|N^1) = (B_L^1|N_L^1)$.
- Calculate $z_2^* = \min\{c_{B_L^1}^2 x_{B_L^1} : A_{B_L^1} x_{B_L^1} = b, x_{B_L^1} \geq 0\}$, and let $(B_L^2|N_L^2)$ be the optimal partition.
- Calculate $z_3^* = \min\{c_{B_L^2}^3 x_{B_L^2} : A_{B_L^2} x_{B_L^2} = b, x_{B_L^2} \geq 0\}$, and let $(B_L^3|N_L^3)$ be the optimal partition.
- \vdots
- Calculate $z_p^* = \min\{c_{B_L^{p-1}}^p x_{B_L^{p-1}} : A_{B_L^{p-1}} x_{B_L^{p-1}} = b, x_{B_L^{p-1}} \geq 0\}$, and let $(B_L^p|N_L^p)$ be the optimal partition.

At each step of this procedure we have that $B_L^i \subseteq B_L^{i-1}$. Moreover, B_L^p indexes the collection of variables that are allowed to be positive at optimality. Setting $B_L = B_L^p$ and $N_L = \{1, 2, \dots, n\} \setminus B_L^p$, we define the *lexicographic Optimal Partition* to be $(B_L|N_L)$. We remark that since every lexicographic optimal solution is a pareto optimal solution, we have $B_L \subseteq \overset{mopt}{B}$. This relation is independent of the ordering on the objectives. Notice that as in linear programming, the lexicographic optimal partition characterizes the optimal set because $\mathcal{P}_L^* = \{x \in \mathcal{P} : x_{N_L} = 0\}$. Also notice that the lexicographic optimal partition is the LP optimal partition if there is a single objective function.

Similar to the MOLP optimal partition, the lexicographic MOLP optimal partition provides insights into the radiotherapy design problem. As an example we consider the situation depicted in Figure 10, where a tumor has grown around a critical structure, say the brain stem, and the optical sockets are restricted critical regions. The difference between the approach taken in this section and that of the last, is that we now assume that the dosimetrist has prioritized the three goals of 1) attaining a uniform, tumoricidal dose to the tumor [**Tumor**], 2) designing a plan that deposits as little radiation as possible to the critical structures [**Critical**], and 3) keeping the normal tissue under 110% of the tumoricidal dose [**Normal**]. For example, if the patient is declared terminal, the overriding treatment objective might be to do no harm to the critical structures. In this situation, the goal of attaining a uniform, tumoricidal dose might be last on the priority scale, and the hierarchy for the goals would be Critical/Normal/Tumor. For this priority listing, the objective function in (1) changes to

$$\begin{pmatrix} e^T \beta \\ e^T \gamma \\ e^T \alpha \end{pmatrix} \begin{array}{l} \leftarrow \text{Priority 1: Decrease Critical Structure Dose} \\ \leftarrow \text{Priority 2: Limit Dose to Normal Tissue} \\ \leftarrow \text{Priority 3: Obtain Uniform, Tumoricidal Dose.} \end{array}$$

Figures 11, 12, and 13 depict optimal plans for various priority listings. While each plan is different, the most noticeable difference is between the plans in Figure 12, where the tumor has the highest priority, and those in Figures 11 and 13. While all of these plans exhibit a high dose to the area between the optical cavities (the nasal cavity), the plan where the tumor is the first priority deposits high amounts of radiation in parts of the anatomy that the other two do not. Since the plan in Figure 13 attains tumor uniformity, the high levels found in Figure 12 do not mean that adhering to the normal tissue limits is impossible if the tumor is to be within its prescribed limits. Rather, the high doses in Figure 12 are consequences of the combined affect of the tumor having the highest priority and the normal tissue having the lowest priority —i.e. by the time that the tumor attains its prescription and the critical structure dose is minimized, the only tissue remaining to focus energy on is the normal tissue.

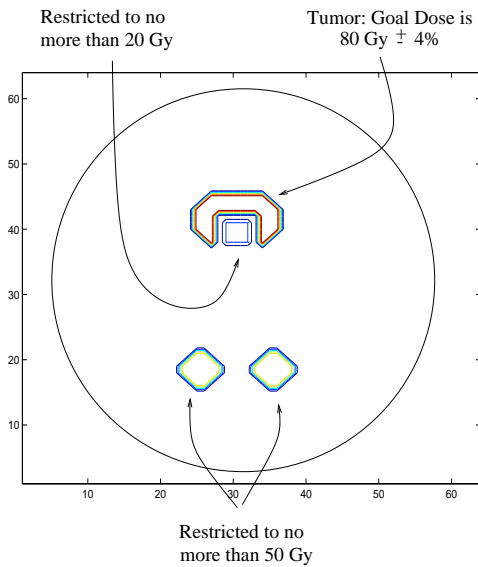


Figure 10: A tumor that has grown around a critical structure.

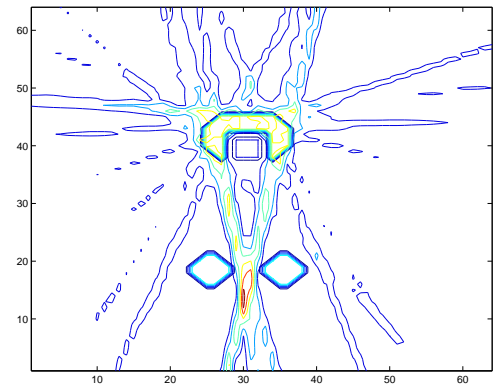


Figure 11: Optimal Plan for Priority List Critical/Normal/Tumor.

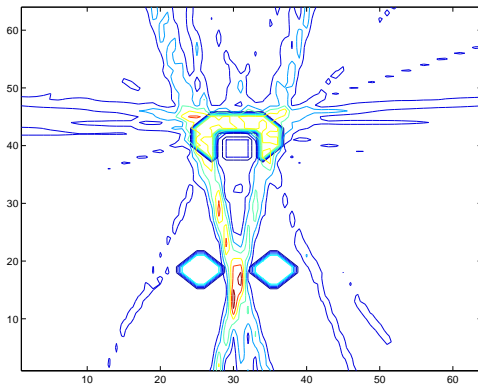


Figure 12: Optimal Plan for Priority List Tumor/Critical/Normal.

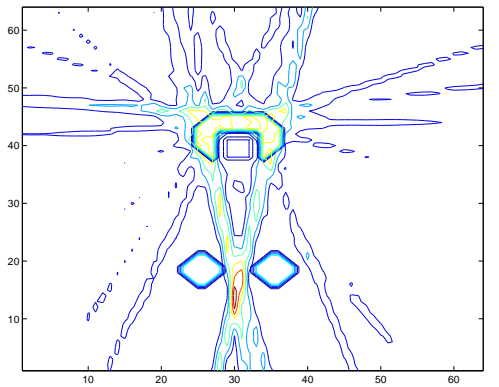


Figure 13: Optimal Plan for Priority List Normal/Tumor/Critical.

The optimal partition approach to solving a lexicographic MOLP is useful because such an approach indicates how the minimization of each objective function removes angles from consideration. As before, we adapted the *RAD* software package by altering the objective function to the appropriate priority list, and variables with a value less than 10^{-3} were declared zero after each sequential linear program was solved. The remaining positive variables were allowed in subsequent models. Table 1 provides each angles level of use after each sequential solve. For example, the first two rows correspond to the priority listing of Tumor/Critical/Normal, and the two pictures under **1st Priority** show each angles level of use after guaranteeing a uniform, tumoricidal dose (the top graph shows the angle use around the patient image and the lower graph is angle vs. angle use). From these two graphs we see that no angles are excluded after minimizing the first objective. The two graphs under **1st & 2nd Priority** show that angles 95° and 275° are removed from consideration upon minimizing the dose to the critical structures. The final two graphs under **1st, 2nd, & 3rd Priority** show that many angles are no longer used once the 3rd priority of minimizing the normal tissue dose is considered.

For all priority lists, the graphs in the last column indicate which angles are in B_L , those with a positive usage, and the angles in N_L , those with a zero usage. While it may appear that the angles in N_L are similar for each priority list, only 9 angles are zero under all priority choices: 55° , 75° , 120° , 125° , 205° , 225° , 250° , 310° , and 355° . These are the angles that can be excluded from consideration regardless of how the dosimetrist orders the priorities (which means we have removed 12.5% of the angles from consideration).

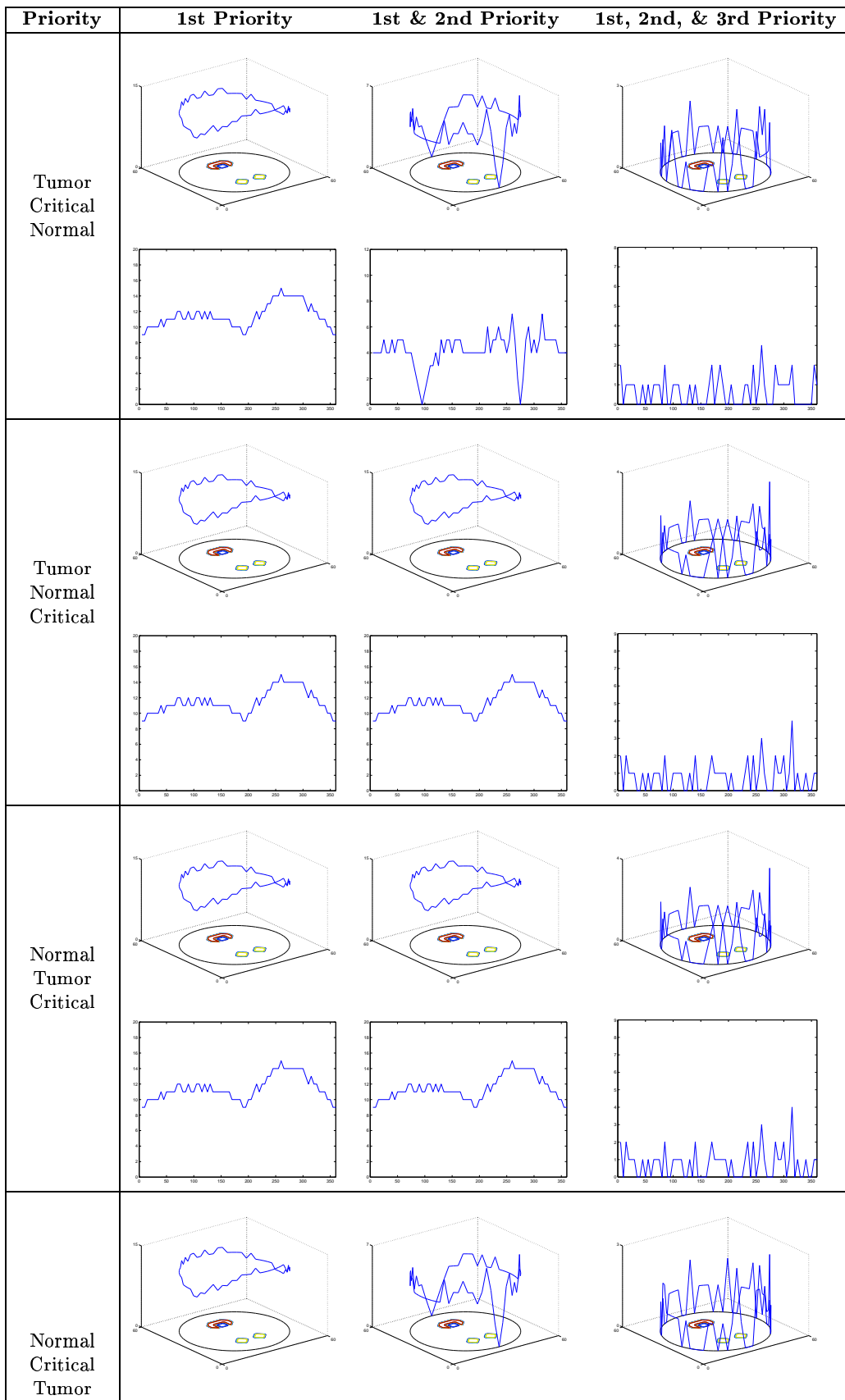


Table 1: How an Angles Candidacy is Affected Under Sequential Minimization

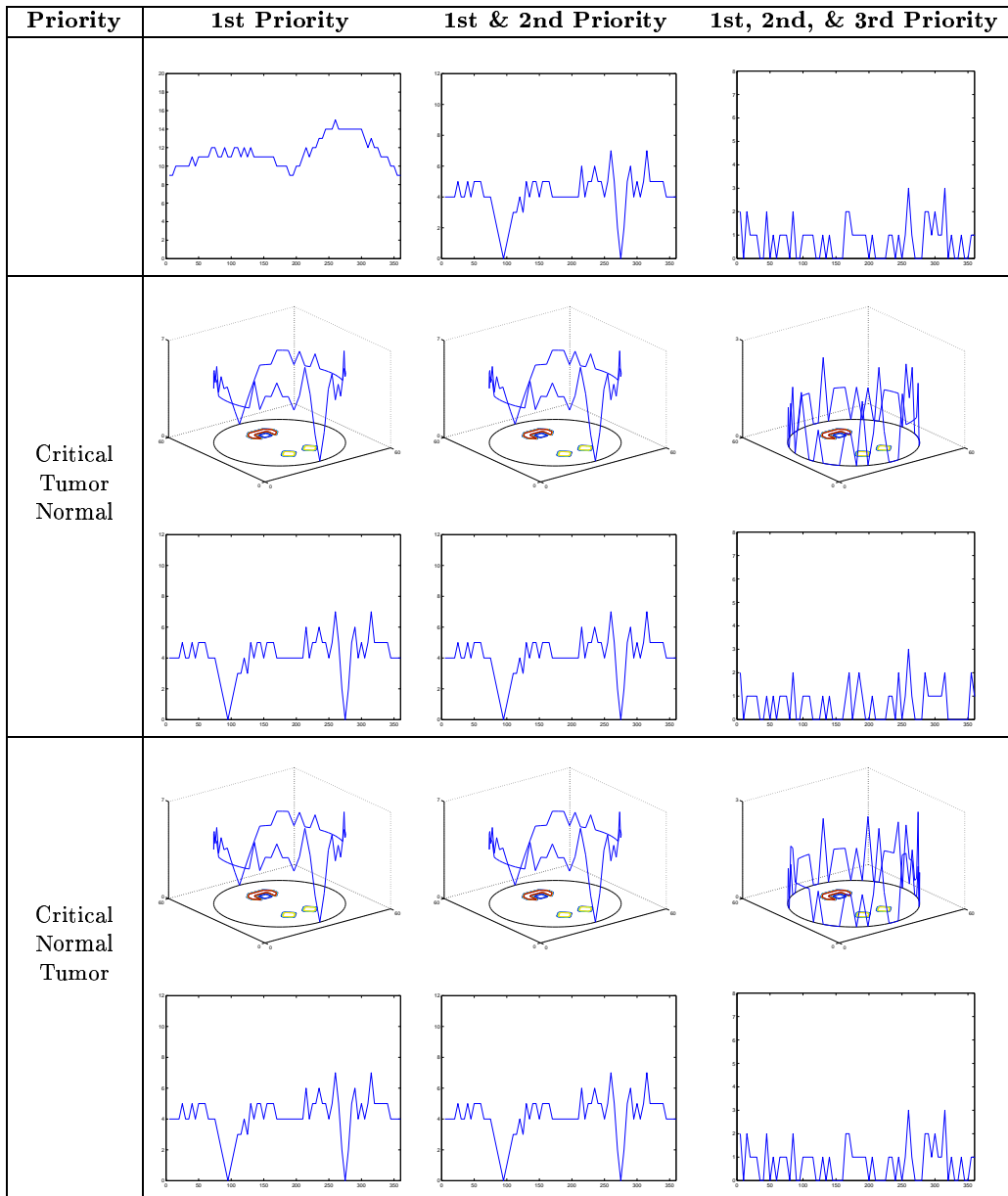


Table 1: How an Angles Candidacy is Affected Under Sequential Minimization

6 Conclusions

The aspirations for this paper were to develop optimal partitions for the Multiple Objective Linear Programming problem that were natural extensions of the LP optimal partition. The MOLP optimal partition and the lexicographic optimal partition achieve this goal because 1) the definitions are equivalent if there is a single objective, and 2) they classify the optimal set —i.e. $\mathcal{E} \subseteq \{x \in \mathcal{P} : x_{m_{olp}} = 0\}$ and $\mathcal{P}_L^* = \{x \in \mathcal{P} : x_{N_L} = 0\}$. Theorem 1 illustrates how the structure of the efficient frontier and the MOLP optimal partition are related. Furthermore, we show how to produce the MOLP optimal partition for bicriteria optimization and the lexicographic optimal partition for an arbitrary number of objectives. Both partitions are shown to be useful in the design of radiotherapy treatment plans.

There are several avenues for further research. First, it would be nice to know the relationship between $(B_L|N_L)$ and $(B^{\text{molp}}, N^{\text{molp}})$. For example, under what conditions is it true that the union of the B_L 's, taken over

all priority lists, equal to $B^{m.o.p}$. Second, the authors of [5] classify objective functions as opposing, repelling, and allying, and these concepts might provide alternative ways to calculate $(B^{m.o.p}|N)$ and $(B_L|N_L)$. We make the following conjectures:

- If all the objective functions ally with each other, $\bigcup_{i=1}^p B^i = B^{m.o.p}$, and
- If the objective functions oppose each other, $B^i \not\subseteq B^{m.o.p}$, $i = 1, 2, \dots, p$.

Third, the radiotherapy plans produced by \mathcal{RAD} are not practical because the large number of angles being used forces exorbitant treatment times. Hence, plans must be pruned to the capabilities of the treatment facility. Combining the partitions developed here might lead to procedures that intelligently achieve practical treatment plans.

References

- [1] S. S. Abhyankar, T. L. Morin, and T. Trafalis. Efficient faces of polytopes: Interior point algorithms, parameterization of algebraic varieties and multiple objective optimization. In J. Lagarias and M. Todd, editors, *Mathematical Developments Arrising From Linear Programming*, pages 319–341. American Mathematical Society, 1990.
- [2] I. Adler and R. Monteiro. A geometric view of parametric linear programming. *Algorithmica*, 8:161–176, 1992.
- [3] A. Arbel. A multiobjective interior primal-dual linear programming algorithm. *Computers & Operations Research*, 21(4):433–445, 1994.
- [4] A. Berkelaar, C. Roos, and T. Terlaky. The optimal set and optimal partition approach to linear and quadratic programming. In T. Gal and H. Greenberg, editors, *Recent advances in sensitivity analysis and parametric programming*, chapter 6. Kluwer Publishers, 1997.
- [5] R. Caron, H. Greenberg, and A. Holder. Analytic centers and repelling inequalities. Technical Report CCM 142, Center for Computational Mathematics, University of Colorado at Denver, 1999. To appear in *European Journal of Operations Research*.
- [6] A. Cormack and E. Quinto. The mathematics and physics of radiation dose planning using x-rays. *Contemporary Mathematics*, 113:41–55, 1990.
- [7] M. Ehrgott. *Multicriteria Optimization*, volume 491 of *Lecture Notes in Economics and Mathematical Systems*. Springer, New York, NY, 2000.
- [8] T. Gal. *Postoptimal analysis, parametric programming, and related topics: degeneracy, multicriteria decision making, redundancy*. Walter de Gruyter & Co., New York, NY, 2nd edition, 1995.
- [9] S. Gass and T. Saaty. Parametric objective function (part 2) - generalization. *Operations Research*, 3(4):395–401, 1955.
- [10] A. Geoffrion. Solving bicriterion mathematical programs. *Operations Research*, 15:39–54, 1967.
- [11] A. Goldman and A. Tucker. Theory of linear programming. In H. Kuhn and A. Tucker, editors, *Linear Inequalities and Related Systems*, volume 38, pages 53–97. Princeton University Press, Princeton, New Jersey, 1956.
- [12] H. Greenberg. The use of the optimal partition in a linear programming solution for postoptimal analysis. *Operations Research Letters*, 15(4):179–185, 1994.
- [13] H. Greenberg. *Mathematical Programming Glossary*. World Wide Web, <http://www-math.cudenver.edu/~hgreenbe/glossary/glossary.html>, 1996-2001.
- [14] H. Greenberg. Simultaneous primal-dual right-hand-side sensitivity analysis from a strictly complementary solution of a linear program. *SIAM Journal on Optimization*, 10(2):427–442, 2000.
- [15] H. Greenberg, A. Holder, C. Roos, and T. Terlaky. On the dimension of the set of rim perturbations for optimal partition invariance. Technical Report CCM No. 94, University of Colorado, at Denver, 1996. To appear in *SIAM Journal of Optimization*.
- [16] A. Holder. *Sensitivity analysis and the analytic central path*. PhD thesis, University of Colorado at Denver, Denver, CO, 1998. <http://www-math.cudenver.edu/graduate/thesis/phdindex.html>.

- [17] A. Holder. Desinging radiotherapy plans with elastic constraints and interior point methods. Technical Report No. 49, Trinity University, San Antonio, TX, 2000.
- [18] A. Holder, J. Sturm, and S. Zhang. Marginal and parametric analysis of the central optimal solution. Technical Report No. 48, Trinity University Mathematics, 1999. To appear in *Information Systems and Operational Research*.
- [19] B. Jansen. *Interior Point Techniques in Optimization*. PhD thesis, Delft University of Technology, Delft, Netherlands, 1995.
- [20] B. Jansen, J.J. de Jong, C. Roos, and T. Terlaky. Sensitivity analysis in linear programming: Just be careful! *European Journal of Operations Research*, 101:15–28, 1997.
- [21] J. Dennis and I. Das. Normal-boundary intersection: A new method for generating pareto optimal points in nonlinear multicriteria optimization problems. *SIAM Journal on Optimization*, 8(3):631–657, 1998.
- [22] H. Mills. Marginal values of matrix games and linear programs. In H. Kuhn and A. Tucker, editors, *Linear Inequalities and Related Systems*. Princeton University Press, Princeton, NJ, 1956.
- [23] R. Monteiro and S. Mehrotra. A general parametric analysis approach and its implication to sensitivity analysis in interior point methods. *Mathematical Programming*, 72:65–82, 1996.
- [24] C. Roos. Interior point approach to linear programming: Theory, algorithms & parametric analysis. In A. van der Burgh and J. Simonis, editors, *Topics in Engineering Mathematics*, pages 181–216. Kluwer Academic Publishers, Netherlands, 1992.
- [25] C. Roos, T. Terlaky, and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization: An Interior Point Approach*. John Wiley & Sons, New York, NY, 1997.
- [26] T. Saaty and S. Gass. Parametric objective function (part 1). *Operations Research*, 2(3):316–3–19, 1954.
- [27] M. Wright. The interior-point revolution in constrained optimization. Technical Report 98-4-09, Bell Laboratories, Murray Hill, New Jersey, 1998.
- [28] S. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, PA, 1997.
- [29] Y. Ye. *Interior Point Algorithms Theory and Analysis*. John Wiley & Sons, Inc., New York, NY, 1997.
- [30] A. Yildirim and M. Todd. An interior-point approach to sensitivity analysis in degenerate linear programs. Technical Report 1268, School of Operations Research and Industrial Engineering, Cornell University, 2000.
- [31] A. Yildirim and M. Todd. Sensitivity analysis in linear programming and semidefnate programming using interior-point methods. *Mathematical Programming*, 90(2):229–261, 2001.
- [32] S. Zhang. On the strictly complementary slackness relation in linear programming. In *Advances in Optimization and Approximation*, pages 347–361. Kluwer, 1994.