A hybrid framework for optimizing beam angles in radiation therapy planning

Gino J. Lim*and Laleh Kardar and Wenhua Cao

December 29, 2013

Abstract

The purpose of this paper is twofold: (1) to examine strengths and weaknesses of recently developed optimization methods for selecting radiation treatment beam angles and (2) to propose a simple and easy-to-use hybrid framework that overcomes some of the weaknesses observed with these methods. Six optimization methods –branch and bound (BB), simulated annealing (SA), genetic algorithms (GA), nested partitions (NP), branch and prune (BP), and local neighborhood search (LNS)– were evaluated. Our preliminary test results revealed that (1) one of the major drawbacks of the reported algorithms was the limited ability to find a good solution within a reasonable amount of time in a clinical setting, (2) all heuristic methods require selecting appropriate parameter values, which is a difficult chore, and (3) the LNS algorithm has the ability to identify good solutions only if provided with a good starting point. On the basis of these findings, we propose a unified beam angle selection framework that, through two sequential phases, consistently finds clinically relevant locally optimal solutions. Considering that different users may use different optimization approaches among those mentioned above, the first phase aims to quickly find a good feasible solution using SA, GA, NP, or BP. This solution is then used as a starting point for LNS to find a locally optimal solution. Experimental results using this unified method on five clinical cases show that it not only produces consistently good-quality treatment solutions but also alleviates the effort of selecting an initial set of appropriate parameter values that is required by all of the existing optimization methods.

Keywords: IMRT, IMPT, treatment planning, beam angle optimization

*Corresponding author: Department of Industrial Engineering, University of Houston, ginolim@uh.edu, fax: 713-743-4190

1 Introduction

Radiation therapy is the treatment of cancerous tissue with ionizing radiation that damages the DNA of cells by disabling the cell's ability to multiply. It aims to deliver a sufficient radiation dose to the tumor to destroy or control the growth of the cancerous cells, while minimizing the amount of radiation absorbed by healthy tissues. Intensity-modulated radiation therapy with high-energy photons (IMRT) and with charged protons (IMPT) are two of the most advanced delivery modalities for radiation therapy that allow a high degree of conformity of the radiation dose distribution to the desired target area. Before a treatment plan is designed, a medical imaging technique such as computed tomography (CT), magnetic resonance imaging (MRI), or positron emission tomography (PET) is used to localize tumors. These images allow for three-dimensional representation of the tumor and surrounding normal tissue for the purpose of treatment planning. Three types of volumes are considered: the planning target volume (PTV), which includes the cancerous area; the organs-at-risk (OARs) or critical structures that are located close to the PTV; and the remaining portions that are considered normal tissue.

Radiation is delivered to the target volume from a gantry that stops at a fixed number of selected beam angles around the patient. The selection of suitable beam angles is generally based upon the treatment planner's experience or intuition. Treatment angles for a specific cancer are often decided by the radiation planner in a clinical setting and not necessarily mathematically optimized. If a new set of treatment beam angles is needed, the selection process can be time consuming because it often requires lengthy trial-and-error attempts. Generally, when a relatively large number of beams are used, the final plan may be less sensitive to the choice of individual angles; however, it may result in exposing a larger volume of healthy tissues to radiation. Furthermore, changing beam directions during the treatment is time consuming and may increase treatment delivery time.

It has been argued that a smaller number of beams placed at the optimal angles can generate plans as good as, or even better than, those generated by a larger number of non-optimized beams. Depending on the type of cancer, four to nine beams are typically used in IMRT [1, 2, 3], whereas two to three beams are typically used in IMPT [4, 5]. Using optimization methods to identify beam angles can expedite the treatment planning process and eliminate the dependency on a planner's experience. Selecting an optimal set of treatment beam angles, called beam angle optimization (BAO), is complex and involves a very large-scale combinatorial optimization problem. Typically, the BAO problem also includes a sub-problem, fluence map

optimization (FMO), which optimizes the intensity profile for each of the selected beam angles to ensure the resulting treatment plan meets prescribed requirements.

To date, the BAO problem has been extensively studied and fruitful results have been achieved. Global optimization approaches, including mixed integer programming (MIP) approaches [6, 7, 8], simulated annealing (SA) [2, 9, 10, 11, 12], genetic algorithms (GA) [13, 14, 15, 16, 17, 18], nested partitions (NP) algorithms [19], particle swarm algorithms [20], the response surface method [21], and artificial neural network algorithms [22] have been proposed by many researchers. These approaches are designed to avoid entrapment in local minima and to converge to global optimal solutions. However, a large number of iterations, i.e., longer computation times, are required to achieve global optimal or good-quality solutions [23]. These longer times make global approaches less attractive to treatment planner in clinical practice. In addition, the performance of those approaches is largely dependent on the choice of algorithmic parameters. This requirement may also limit the application of the proposed methods in real-world practice because determining appropriate algorithmic parameters is usually a difficult task. In addition to global optimization approaches, successful attempts to optimize beam angles have also been made by applying local optimization approaches [6, 12, 24, 25, 26]. However, although local search algorithms can find local optimal solutions quickly, the resulting solution quality is often sensitive to the initial starting point [6].

Despite the rich literature on various BAO methods, studies comparing the performance of those methods have been rarely reported [27]. Recently, Lim et al. [23] conducted a preliminary comparison study of various BAO methods for IMRT, which motivated the current study. In this study, a collection of both deterministic methods (branch and bound (BB), branch and prune (BP), and local neighborhood search (LNS)) and stochastic methods (SA, GA, and NP) were implemented and tested on five patient cases, including IMRT and IMPT and four different cancer types. Having observed strengths and weaknesses of each method, we propose a hybrid scheme in which SA, GA, NP, or BP is used to provide a starting solution for LNS to find a local optimal solution, and we tested this approach in the same five cases. The primary purpose of the hybrid approach is to deliver consistently good-quality solutions within a clinically acceptable amount of computational effort. It is known that users often make a considerable effort to fine-tune the parameter values used in most stand-alone methods. Therefore, the hybrid approach is also designed to alleviate the impact of the parametric settings of the stand-alone methods; for example, the randomness of selecting an initial starting point for LNS is avoided. With a little initial effort, appropriate parameter values of SA, GA, NP, and BP can be found to deliver consistently good starting solutions for LNS.

The rest of this paper is organized as follows. Section 2 describes optimization models. Solution methods are discussed in Section 3. In Section 4, five clinical cases are used to compare the efficiency and quality of the presented solution methods, including the hybrid approach. Finally, conclusions are presented in Section 5.

2 Optimization Model

To formulate an optimization model, we assume that the structures are irradiated using a set of beams, each corresponding to a particular beam angle. For each beam angle, the aperture of this beam is decomposed into small beamlets. The primary decision variables of the optimization model are the beamlet weights $w_{a,b}$ representing the intensity of radiation delivered by the beam angle a and over the beamlet b, where $a \in A'$, $b \in B_a$, A' denotes the set of treatment beam angles, and B_a represents the set of beamlets at angle a. Let T denote the set of target structures; S, the set of OARs; and N, the set of normal structures. **In practice, each structure $s \in T \cup S \cup N$ is discretized into a finite number v_s of cubes, which are known as voxels. The set of all voxels will be denoted by $V = \cup_s v_s$.** The total dose D_j that a voxel j receives is

$$D_j = \sum_{a \in A'} \sum_{b \in B_a} w_{a,b} \cdot d_{j,a,b} \qquad j \in V, \tag{1}$$

where $d_{j,a,b}$ denotes the dose contribution to voxel j from beamlet (a,b) at unit weight. Table 1 shows all model parameters used in this paper. A hot spot is a portion of the organ that receives a radiation dose above the desired dose level, and a cold spot is a portion of the organ that receives a radiation dose below the required dose level. Using this notation, we explain two optimization models in the following subsections.

2.1 Linear Programming Model

The linear programming (LP) model for FMO described by Lim and Cao [6] is adopted for this study. Given a fixed set of beam angles as an input parameter, the model aims to find the optimal amount of radiation that each beamlet delivers. The LP model allows evaluation of the treatment plan quality for the given beam

Table 1: Optimization model parameters

Notation	Definition
θ_L	Cold spot control parameter on targets
θ_U	Hot spot control parameter on targets
ϕ	Hot spot control parameter on OAR
η	Maximum number of treatment beam angles
L	Lower reference bound on voxel
U	Upper reference bound on voxel
λ_t^+	Penalty coefficient for hot spots on targets
λ_t^-	Penalty coefficient for cold spots on targets
λ_s	Penalty coefficient for hot spots on OAR
λ_n	Penalty coefficient for normal structure

angles. The LP formulation for solving FMO is

$$\underset{w}{\text{Min}} \quad f(D) \tag{2}$$

$$\text{s.t.} \quad L_j \leq D_j \leq U_j \qquad j \in v_s \quad s \in T \tag{3}$$

$$D_j \leq U_j \qquad j \in v_s \quad s \in \bar{N} \subseteq N \tag{4}$$

$$D_j \leq U_j \qquad j \in v_s \quad s \in S \tag{5}$$

$$0 \leq w_{a,b} \leq M_{a,b} \qquad a \in A \quad b \in B_a. \tag{6}$$

Constraint (3) imposes upper and lower bounds on voxels to avoid an undesirable dose distribution to the targets. In constraint (4), a strict dose upper bound is imposed on normal tissue, excluding voxels that are close (within δ mm, where $\delta > 0$) to the tumor region, to avoid the streaking issue in IMRT treatment planning [28]. Constraint (5) imposes upper bounds on some of the OARs. Constraint (6) sets an upper bound on the beamlet weights. Note that except for the upper bound for the OAR and/or normal tissue and the lower and upper bound for the tumor, other dose-volume constraints are not included in the LP model; they are only checked after the optimization.

We use a penalty-based linear objective function that has four terms corresponding to the targets, the

OARs, and normal tissue, which is written as follows:

$$f(D) = \sum_{s\in T}\Big(\sum_{j\in v_s}(\lambda^{+}_{t_s}||(D_j-\theta_U)_+||_\infty + \lambda^{-}_{t_s}||(\theta_L-D_j)_+||_\infty)\Big) + \sum_{s\in S}\Big(\frac{1}{|v_s|}\sum_{j\in v_s}\lambda_{s_s}||(D_j-\phi)_+||_1\Big) + \sum_{s\in N}\frac{1}{|N_s|}\sum_{j\in v_s}\lambda_{n_s}||D_j||_1. \tag{7}$$

The notation $(\cdot)_+$ represents $\max\{\cdot,0\}$. The first two terms are based on the L_∞ - norm that penalizes maximum violations for both hot spots and cold spots on the target. The third and the fourth terms penalize the average overdose on OARs hot spots and normal tissue, respectively. The values of the optimization model parameters are selected on the basis of interaction with physicians.

2.2 Mixed Integer Programming

The MIP model is designed to select η beam angles out of a given set of candidate angles A, where ($\eta << |A|$), and to optimize the beamlet weights $w_{a,b}$ accordingly. The MIP model is constructed by adding binary variables ψ_a, where $a \in A$, to the LP model, in which ψ_a denotes whether angle a is selected. The corresponding MIP formulation is

$$\min_{w,\psi} \quad f(D) \tag{8}$$

$$\text{s.t.} \quad L_j \le D_j \le U_j \qquad j \in v_s \quad s \in T \tag{9}$$

$$D_j \le U_j \qquad j \in v_s \quad s \in \bar{N} \subseteq N \tag{10}$$

$$D_j \le U_j \qquad j \in v_s \quad s \in S \tag{11}$$

$$\sum_{a\in A}\psi_a \le \eta \tag{12}$$

$$0 \le w_{a,b} \le M_{a,b}\cdot\psi_a \qquad a \in A \quad b \in B_a \tag{13}$$

$$\psi_a \in \{0,1\} \quad a \in A.$$

Solving the MIP model optimally is computationally challenging, and solution approaches for selecting treatment angles within a clinically acceptable time limit are needed. In the next section, we describe implementations of different BAO solution methods found in the literature.

3 Solution Methods

The following solution methods are considered in this study:

- Simulated annealing (SA) [12]
- Genetic algorithm (GA) [17]
- Nested partitions (NP) [19]
- Branch and prune (BP) [6]
- Local neighborhood search (LNS) [6, 12].

In the following sections, we give a brief overview of each of these methods. Note that NP and BP require a score function to select a promising angle set. For this purpose, we use the score function developed by Lim *et al.* [28]. A beam angle score function is a function that is used to evaluate a particular angle's merit to its resulting treatment. After solving the LP problem of Eqs. (2-6) for a given beam angle set A' ($A' \subseteq A$), the total dose delivered to the targets, OARs, and normal structures by beam angle a can be calculated as

$$D(\Omega)_a = \sum_{j \in v_\Omega} \sum_{b \in B_a} w_{a,b} \cdot d_{j,a,b} \qquad a \in A',\ \Omega \in \{T,S,N\}. \tag{14}$$

Three subscores are calculated to capture the dose contribution to the targets (s_a^T) and the average dose contribution to the OARs (s_a^S) and normal structures (s_a^N). These scores can be defined as follows:

$$s_a^T = \frac{D(T)_a}{uD(T)_a}, \quad s_a^S = \frac{D(S)_a}{C_a^S}, \quad s_a^N = \frac{D(N)_a}{C_a^N}, \quad a \in A', \tag{15}$$

where $uD(T)_a$ is the total dose on the targets by a uniform weight of 1 from angle a. C_a^S and C_a^N are the numbers of voxels in the OARs and the normal structures, respectively, that receive positive dose from angle a. The merit score of an angle a (S_a for $a \in A'$) is defined using a linear combination of the three normalized scores given as follows:

$$S_a = \left(\frac{s_a^T}{\sum_{a \in A'} s_a^T} \right) - k_s \left(\frac{s_a^S}{\sum_{a \in A'} s_a^S} \right) - k_N \left(\frac{s_a^N}{\sum_{a \in A'} s_a^N} \right), \quad 0 \le k_S,\ k_N \le 1. \tag{16}$$

This score acts as a measure to determine the quality of angles. Angles with higher scores are considered to provide better quality. The weighting coefficients, k_S and k_N, are used as avoidance factors for OARs and normal tissue.

3.1 Simulated Annealing

SA is a probabilistic search algorithm that models the physical process of heating a metal and then slowly lowering the temperature into a minimum energy crystalline structure (the annealing process). Algorithm 1 is a pseudocode of SA.

The algorithm starts with an initial solution of the problem and its corresponding objective function value. At each iteration of the SA algorithm, a new solution of the problem is generated within the neighborhood of the current solution by perturbing one or more decision variables. In the SA implementation by Aleman *et al.* [12], a new candidate angle set is constructed by swapping each element of the current solution with one in its neighborhood. Although changing all angles in each iteration brings additional complexity to the problem, it allows the problem to escape from local minima. The neighborhood of an individual beam angle is defined by bounds of $\pm\sigma$ degrees. That is, the neighborhood of an angle a_i consists of angles that are within the interval of $[a_i-\sigma,\ a_i+\sigma]$. We use a geometric probability distribution [12] to select neighborhood angles for individual beam angles. This distribution gives a higher selection probability to angles that are close to angle a_i and also allows for angles further from the current angle to be selected with a certain probability, that potentially increases the likelihood of finding a global optimal solution.

After selecting a new solution, the associated objective value is computed using the LP model and compared with the current objective value. If the new solution improves the objective value, it replaces the old one. Otherwise, a random number r is generated uniformly in the range $[0,1]$ and compared with $y = e^{(-(z^{new}-z^{old})/T(i))}$, where z^{new} and z^{old} are the objective values of the new and current solutions, respectively, and T is the system temperature. If $r < y$, the new (worse) solution is accepted. Otherwise, it is rejected and the current solution remains as is. The system temperature is initiated by the user before running the algorithm. When the system temperature is high, it is more likely that a worse solution might be accepted. To prevent the algorithm from being trapped in a local minimum, it usually starts with a relatively high temperature. At each iteration, the temperature is decreased according to an exponential cooling schedule, $T(i+1) = \beta \cdot T(i)$, where β is the cooling rate, which is a constant and typically lies between 0.8 and 0.99 [29]. The algorithm stops when either of the following two conditions is met: (a) a certain number of

Algorithm 1 Simulated Annealing

Generate an initial set of angles $\bar{A}$;
Solve the LP model with $\bar{A}$ and store objective value $\bar{z}$;
$T(1) := T_0$; $i := 1$; $\bar{z}^* := \bar{z}$; $\bar{A}^* := \bar{A}$;
while Stopping criteria are not met **do**
    Choose a new angle set A^i from the neighborhood of $\bar{A}^*$ randomly;
    Solve the LP model with A^i and store objective value z_i ;
    if $z_i < \bar{z}^*$ **then**
        $\bar{z}^* := z_i$; $\bar{A}^* := A^i$;
    else
        Choose a random r uniformly in the range $[0,1]$;
        if $r < \exp[-(z_i - \bar{z}^*)/T(i)]$ **then**
            $\bar{z}^* := z_i$; $\bar{A}^* := A^i$;
        end if
    end if
    Decrease T according to temperature cooling schedule $T(i+1) = \beta \cdot T(i)$;
    $i \leftarrow i+1$;
end while
return ($\bar{A}^*$, $\bar{z}^*$) as optimal solution.

iterations is reached, or (b) there is no change in the objective value for a certain number of consecutive iterations.

3.2 Genetic Algorithm

Genetic algorithms are a class of evolutionary algorithms that imitate the process of natural evolution. These algorithms encode a potential solution to a specific problem on a simple chromosome-like data structure. In this structure, genes are represented by unknown variables in a solution, while a chromosome is represented by the solution comprised of such variables. Multiple chromosomes represent a population of individuals. The selection of individuals is based on their fitness values. Genetic operations such as selection, crossover, and mutation are carried out on individuals with high fitness values. This results in the next generation of individuals having relatively higher fitness values. The process repeats until the stopping criteria are met, and the fittest individual in the last generation offers the approximate optimal solution. Algorithm 2 shows a pseudocode for GA.

As applied to BAO, the algorithm begins by creating a population of individuals for the first-generation set. Each individual or chromosome represents a combination of beam angles as a solution for the BAO problem. The population size is set to be twice the number of treatment angles. For example, if the desired number of treatment angles is six, the initialization process creates a population of 12 beam angle sets. The

Algorithm 2 Genetic Algorithm

1: $g := 1$;
2: Generate initial population of solutions $P^{(g)}$;
3: Evaluate the solutions by calculating the objective value of the LP model;
4: **while** Stopping criteria are not met **do**
5: Generate child solutions via crossover and mutation;
6: Evaluate the new solutions by calculating the objective value of the LP model;
7: Update the population $P^{(g)}$;
8: $g \leftarrow g + 1$;
9: **end while**
10: **return** the fittest individual as optimal solution.

initial population is generated using two methods: (a) beam angle sets initialized with equi-spaced angles, and (b) beam angle sets that are randomly generated.

After generation of an initial population, all sets in the population are evaluated based on their fitness values. The fitness value of a set is calculated using the objective function of the LP model. A smaller objective value represents a better beam angle set. Fitter sets will go forward to form a "mating" pool for the next generation. Next, the crossover operation selects a pair of beam angle sets and chooses the crossover positions randomly. The two parts of the sets are exchanged. Any two randomly selected "parent" angle sets are considered for crossover according to a specific crossover probability, P_x, which normally lies between 0.5 and 0.95 [17]. If the fitness values of the two new "children" are better (e.g., have a lower objective value) than those of the parent sets, the new sets will be selected as members of the new generation; otherwise, their parents will be kept in the mating pool for the next generation. The newly formed sets from the crossover operation are randomly selected for the mutation operation with a probability of P_m. First, an angle with the lowest score is selected for removal. A replacement angle is generated randomly such that it is not in the current angle set. Once again, the fitness of the new beam angle set is evaluated. This completes one GA iteration. The GA stops when either of the following two conditions are met: (a) a predefined number of generations, or (b) no change in the solution for a certain number of consecutive generations, whichever comes first. The individual with the best fitness value in the last generation will be reported as the optimal angle set.

3.3 Nested Partitions

The NP method is primarily used for solving combinatorial optimization problems with a finite feasible region and has been proven to converge to a global optimal with probability 1 [30]. It is a

randomized optimization method that systematically partitions the feasible region and concentrates the search in regions that are the most promising. The most promising region in each iteration is selected based on the information gathered from random sampling of the entire feasible region. The promising region is partitioned into subregions, and the remaining region is aggregated into one region and it is termed "surrounding region". Solution sets are then sampled from each of these regions using a random sampling approach. The performance function (i.e., objective function) values of the randomly selected samples are used to calculate the promising index for each region. The region with the best promising index becomes the most promising region for the next iteration. If the surrounding region is found to be the best, the algorithm backtracks to a larger region that contains the old most promising region. The new most promising region is then partitioned and solutions are sampled in a similar fashion until the algorithm converges.

Algorithm 3 illustrates the general structure of NP implemented in this study. Given a set of candidate beam angles, an initial starting solution is generated (i.e., Section 4), and the corresponding LP model is solved. Based on the LP solution, the merit scores are calculated for these angles using Eq. (16). An angle with the highest merit score is selected for partitioning to define an initial most promising region. The solution space is divided into two regions based on this angle: the most promising region with all beam angle sets including this angle and the surrounding region with the remaining beam angle sets in the feasible solution space. For example, in our implementation, we start with an equi-spaced angle set and its corresponding objective value. Suppose a_1 is the angle with the highest merit score which will be selected to determine the regions for the second iteration. For the promising region, the first angle remains fixed as a_1 while the remaining angles can be any other angles in the candidate beam angle set. The surrounding region contains all the beam angle sets excluding a_1. The most promising region is further partitioned by fixing the second angle with the second highest merit score. Thus, one subregion includes both angles (called the left subregion), while the other only includes the first angle (called the right subregion). After defining the regions, randomly sampled angle sets are obtained from all regions.

There are different ways to obtain random samples. In this paper, we considered the biased sampling scheme proposed by Zhang et al. [19]. Using this scheme, a frequency index is defined for each angle to indicate how many times a particular angle has appeared in the best angle sets of the previous iterations. An angle with a high-frequency index is deemed to be high quality and is given higher selection probability.

Algorithm 3 Nested Partitions

Generate an initial set of angles $\bar{A}$;
Solve the LP model with $\bar{A}$ and find objective value $\bar{z}$;
$i := 1$; $\bar{A}^* := \bar{A}$; $\bar{z}^* := \bar{z}$;
while Stopping criterion is not met **do**
    Compute scores S_a, $\forall a \in \bar{A}^*$;
    Select an angle with the highest score for partitioning;
    Partition solution space into promising and surrounding regions;
    Sample each region and generate angle sets, $N(A^i) = \{A^i_1, ..., A^i_n\}$;
    for $j = 1 : n$ **do**
        Solve the LP model with A^i_j and store objective value z_j;
    end for
    $j^* \leftarrow \arg\min_j z_j$;
    if $z_{j^*} < \bar{z}^*$ **then**
        $\bar{A}^* := A^i_{j^*}$; $\bar{z}^* := z_{j^*}$
        if $A^i_{j^*}$ is from surrounding region **then**
            Backtrack;
        end if
    end if
    $i \leftarrow i + 1$
end while
return $(\bar{A}^*, \bar{z}^*)$ as optimal solution.

After the sampling, all samples in those regions are evaluated based on their objective function values. The sample with the lowest objective value is selected as the best angle set found so far, and the region to which the sample belongs becomes the next most promising region. If the best sample set resides in the surrounding region, backtracking is performed to reverse a sequence of incorrect moves.

Different versions of the NP algorithm can be developed using different backtracking rules. The algorithm can get back to either the root node (Rule I) or any other nodes along the path that leads to the current promising region (Rule II). In this paper, we used a backtracking scheme that utilizes the frequency index and Rule I [19]. Using this scheme, rather than getting back to the root node and starting with an empty angle set, we select angles with higher frequency indexes to define a most promising region. Then, the computational effort concentrates on the new most promising region, and it is partitioned and sampled in a similar manner. NP stops when a singleton region is reached in which all the required number of angles are specified.

3.4 Branch and Prune

BP is an iterative algorithm that utilizes the integer programming (IP) branch-and-bound tree node elimination approach to concentrate the search effort on the region that more likely contains high-quality feasible solutions. The general structure of the BP algorithm is illustrated in Algorithm 4 .

The algorithm begins by solving the LP relaxation of the IP model, including all candidate angles at the root node. Based on the score values of individual angles, in the second iteration a new angle set is constructed by eliminating the angle with the lowest score value. Similarly, another angle set can be constructed by removing an angle with the second lowest score value. In general, for a given value of c (typically 1, 2, or 3), c child nodes are formed from the parent node by eliminating each of the c worst angles from the current angle set one at a time. All new nodes are evaluated by solving the corresponding LP relaxation models. The node with the smallest objective value is then selected as the promising node for further branching in the next iteration, and the rest of the nodes are pruned accordingly. The algorithm stops when there are only $\eta + \alpha$ candidate angles left in the selected node, where η is the desired number of treatment angles. The value of α is recommended to be a small positive value [6]. The set with $\eta + \alpha$ angles is then fed to the MIP model to select the final η angles.

Algorithm 4 Branch and Prune

$i := 1$;
Initialize $(A^{(i)} = A)$;
while $|A^{(i)}| > \eta + \alpha$ **do**
    Solve the LP model with $A^{(i)}$;
    Compute scores $S_a^{(i)}$, $\forall a \in A^{(i)}$;
    Select c^i angles, $\{\bar{a}_1, ..., \bar{a}_{c^i}\} \subset A^{(i)}$, corresponding to the c^i lowest scores;
    Form c^i child nodes with angle set $B^j = \{A^{(i)} \backslash \bar{a}_j\}$, $j = 1, ..., c^i$;
    for $j = 1 : c^i$ **do**
        Solve the LP model with B^j and store objective value z_j;
    end for
    $j^* \leftarrow \arg\min_j z_j$;
    $i \leftarrow i + 1$;
    $A^{(i)} \leftarrow B^{j^*}$;
end while
Solve MIP with $A^{(i)}$ to find optimal solution.

3.5 Local Neighborhood Search

LNS is a solution search technique to find high-quality local optimal solutions by iteratively searching for a better solution within the neighborhood of the current solution. A pseudocode for the LNS is given in Algorithm 5. In order to design an LNS approach, a neighborhood structure must be defined. The neighborhood of the current solution may contain a large number of solutions that are close to the current one. Examining a larger neighborhood may lead to a better solution, but it may take longer to converge to a solution than with a smaller neighborhood. Hence, a neighborhood structure should be constructed considering both of these performance criteria. To tackle the trade-off between the solution time and quality, neighbors of an individual beam angle, including geometrically adjacent and the opposite angles, are used to construct the neighborhood of a beam angle set. A one-angle-exchange algorithm is performed where one angle of the current angle set is swapped with one of this angle's neighbors, one at a time, to search for a better solution. Once all candidate solution sets are constructed, the LP model is solved for each candidate solution, and the angle set with the minimum objective value is selected. If a new solution yields a better objective value than the current solution, both the solution and the objective value are updated accordingly, and the search algorithm continues. Otherwise, the algorithm stops, and the current solution is reported as optimal.

Algorithm 5 Local Neighborhood Search

Generate an initial set of angles $\bar{A}$;
Solve the LP model with $\bar{A}$ and find objective value $\bar{z}$;
$i := 1$; $\bar{A}_i^* := \bar{A}$; $\bar{z}_i^* := \bar{z}$;
repeat
    Generate new angle sets, $N(A^i) = \{A_1^i, ..., A_n^i\}$, from the neighborhood of $\bar{A}_i^*$;
    for $j = 1 : n$ **do**
        Solve the LP model with A_j^i and store objective value z_j;
    end for
    $j^* \leftarrow \arg\min_j z_j$;
    $\bar{A}_{i+1}^* := A_{j^*}^i$; $\bar{z}_{i+1}^* := z_{j^*}$;
    $i \leftarrow i + 1$;
until $z_i^* < z_{i-1}^*$
return $(\bar{A}_{i-1}^*, \bar{z}_{i-1}^*)$ as optimal solution.

3.6 Hybrid Approaches

Methods such as SA, GA, and NP are global optimization algorithms for solving non-convex problems in which many local optima may exist. These methods are designed to prevent a solution from being trapped in local optima, and to converge to the global optimal in a reasonable computation time. This convergence may require quickly examining a great number of interim solutions. However, for problems such as BAO in radiotherapy planning, evaluating a solution (i.e., solving an FMO model) is often very time consuming in itself. Hence, global algorithms may not guarantee good-quality solutions when the time to find such solutions is limited. In addition, the performance of these algorithms is known to be sensitive to the algorithmic parameters used. Setting appropriate parameters for these algorithms often requires tedious trial-and-error attempts.

On the other hand, local optimization algorithms such as LNS can converge to a local optimal solution rather quickly, but the solution quality depends on the starting points used [6]. Heuristic approaches such as BP may also be efficient in generating good-quality solutions but do not guarantee optimality. Clearly, to achieve a reasonable convergence speed and robustness in solution quality, a straightforward strategy is to combine the key elements of different methods and take advantage of their complementary strengths. In this study, we consider four hybrid optimization techniques that incorporate LNS into SA, GA, NP, and BP. The resulting hybrid approaches operate in two sequential phases. First, one of the SA, GA, NP, or BP methods is used to provide a good, feasible solution quickly. Second, using the first-phase solution as the starting solution, LNS converges to a local optimal solution.

4 Numerical Experiments and Results

The solution methods were tested on five clinical cases: two IMRT cases (one prostate cancer and one pancreas cancer) and three IMPT cases (prostate, lung, and -and-neck cancers). The volumes of interest (VOIs) and the number of voxels within each volume for each case are displayed in Table 2. **For the IMRT cases, a small fraction of sampled normal voxels are included in the optimization model as in [28] (23,433 normal voxels for the prostate case and 9,229 normal voxels for the pancreas case). In the IMPT cases, we did not include the normal voxels in the optimization model because it is a common practice to not include normal voxels in the IMPT literature [5, 31]. This claim was further confirmed on our**

own experiments that adding normal voxels to IMPT optimization did not help improve dosimetrical outcomes of the treatment plans, but it took significantly longer to obtain solutions. Table 3 lists the dose-volume requirements for all the VOIs.

Table 2: VOIs and number of voxels within each volume for different cases

Case	Volume	Number of voxels
Prostate (IMRT)	PTV	5,246
	OAR (rectum)	1,936
Pancreas (IMRT)	PTV	1,244
	OAR (liver)	50,391
	OAR (left kidney)	9,116
	OAR (right kidney)	9,116
	OAR (spinal cord)	489
Prostate (IMPT)	PTV	2,310
	OAR (rectum)	2,462
	OAR (bladder)	4,565
Head and Neck (IMPT)	PTV	436
	OAR (brain stem)	687
	OAR (optic chiasm)	34
Lung (IMPT)	PTV	796
	OAR (esophagus)	1,079
	OAR (heart)	10,441
	OAR (lung)	90,840
	OAR (spinal cord)	686

Scenarios with 12 and 36 candidate beam angles were considered to test the efficiency and solution quality of each method. In this study, we only considered solutions involving coplanar beams (beams that are obtained from the rotation of only the gantry). To generate the 12 and 36 beam angles, the 360° gantry angle was sampled with discrete angle steps of 30° and 10°, respectively. For the IMRT cases, our aim was to select six beam angles for the prostate case and four beam angles for the pancreas case. For the IMPT cases, three beam angles were selected for all cases. All solution methods were initiated with equally spaced angles starting from 0°. All experiments were performed on a Linux computer with a Xeon Quad processor 2.8 GHz and 16 GB RAM. Both the LP model and the MIP model were solved using CPLEX 12.1 with $\delta \approx 10$ mm. Our experiments illustrated that the interior point method and the dual-simplex method were more efficient than other methods in solving the LP model. Therefore, these two methods were used to solve the LP model. The MIP model was solved using the branch-and-bound (BB) method.

Parameter values of each algorithm are taken from the previous studies without tuning them for specific

Table 3: Dose volume requirements for the VOIs of different cases

Case	Volume	Constraints
Prostate	PTV	Prescription: 76 Gy
	PTV	Volume receiving at least the prescription dose $\geq$ 95%
	Rectum	Volume receiving doses higher than 60 Gy: $\leq$40%
	Rectum	Volume receiving doses higher than 70 Gy: $\leq$25%
	Bladder	Volume receiving doses higher than 70 Gy: $\leq$25%
Pancreas	PTV	Prescription: 54 Gy
	PTV	Volume receiving at least the prescription dose $\geq$ 95%
	Spinal cord	Max dose: 45 Gy
Head and neck	PTV	Prescription: 74 Gy
	PTV	Volume receiving at least the prescription dose $\geq$ 95%
	Brain stem	Max dose: 55 Gy
	Optic chiasm	Max dose: 54 Gy
Lung	PTV	Prescription: 70 Gy
	PTV	Volume receiving at least the prescription dose $\geq$ 95%
	Esophagus	Volume receiving doses higher than 55 Gy: $\leq$65%
	Heart	Volume receiving doses higher than 45 Gy: $\leq$65%
	Lung	Volume receiving doses higher than 20 Gy: $\leq$37%
	Spinal cord	Max dose: 45 Gy

cancer cases. These values are used for both stand-alone and hybrid approaches. For the SA algorithm, parameters β (the cooling rate) and T_0 (the initial temperature) were set to 0.95 and 75, respectively. The neighborhood with σ=180 was used so that the entire solution space was considered [12]. However, when generating a neighborhood angle set using geometric distribution with $p = 0.25$, a higher selection probability was given to angles that were closer to the current beam angle. Using this approach, a beam 60° away from the incident beam angle had less than 20% probability of being selected. For the GA method, the crossover operation was performed using a crossover probability (P_x) of 0.9 [17], and the mutation probability was set to 0.01. To implement the NP method, one needs to determine the number of angle sets that would be examined in each iteration. For our comparison study, in each iteration, 25 samples were evaluated: 20 samples were generated from the promising region (10 samples from each subregion), and five samples were generated from the surrounding region [19]. Parameters α and c in the BP algorithm were set to 2 and 3, respectively [6].

Each solution method had different parameters to control stopping. To implement the hybrid strategies, the stopping control parameters for the SA, GA, and NP algorithms were adjusted to quickly find starting points for the LNS. Table 4 lists the stopping criteria for stand-alone and hybrid approaches. Note that

for BP with 36 candidate beam angles, in order to avoid solving large LP problems to generate the initial solution for the LNS, we used the solution obtained from BP with 12 candidate angles as the initial solution for the LNS.

For each case, the deterministic methods, BB and BP, were run once, and the resulting objective values and solution times were reported. For stochastic methods, SA, GA, and NP, 10 random runs were conducted for each case, and means and standard deviations of the objective values and solution times were recorded.

Table 4: Stopping criteria

SA	Number of iterations exceeds 500; or There is no better solution found in 100 successive iterations.
SA-LNS	Number of iterations exceeds 100; or There is no better solution found in 20 successive iterations.
GA	Number of generations exceeds 50; or There is no better solution found in 20 successive generations.
GA-LNS	Number of generations exceeds 20; or There is no better solution found in 5 successive generations.
NP	All of the required number of angles are specified (25 Samples are evaluated in each iteration).
NP-LNS	All of the required number of angles are specified (10 Samples are evaluated in each iteration).

4.1 IMRT: Prostate Case

Table 5 shows the results for the IMRT prostate case with 12 and 36 candidate angles. The results show that for the 12 beam angles, all stand-alone methods provided solutions that were comparable to the MIP (BB) model. Note that BP found the optimal solution in this configuration. The stand-alone methods obtained such objective values up to 82% quicker than did MIP. BP had the shortest solution time, followed by GA, NP, and SA. We repeated all the experiments with 36 candidate beam angles to select six treatment angles. The MIP model failed to find an optimal solution within 1 day, and the process was terminated manually. The last recorded objective function value was 0.294 with a dual gap of 57.44%. BP yielded the best objective value among all solution methods. For stochastic methods, GA outperformed NP and SA in terms of both objective value and solution time. Note that the solution time of the BP algorithm was longer than that of other solution methods, mainly because there were larger LP models to solve in the first few iterations of BP.

Table 5: Comparison of results of stand-alone and hybrid solution methods based on 12- and 36-candidate-beam-angle configurations for the prostate cancer case (IMRT)

η=6		$\|A\| = 12$		$\|A\| = 36$	
Solution method		Obj.	Time (min)	Obj.	Time (min)
Deterministic	BB	0.271	120.8	NA	NA
	BP	0.271	20.8	0.264	149.7
Stochastic	SA	0.272 (0.001)	62.7 (5.3)	0.270 (0.006)	72.5 (12.9)
	GA	0.273 (0.002)	49.6 (10.4)	0.267 (0.006)	66.7 (17.3)
	NP	0.272 (0.001)	53.7 (10.7)	0.269 (0.003)	91.2 (8.5)
Hybrid	BP-LNS	0.271	28.3	0.262	55.4
	SA-LNS	0.272 (0.001)	23.8 (2.7)	0.255 (0.005)	69.6 (14.8)
	GA-LNS	0.271 (0.001)	21.2 (9.4)	0.256 (0.002)	62.8(3.9)
	NP-LNS	0.272 (0.001)	29.7 (3.8)	0.253 (0.003)	72.3 (10.1)

Obj., objective function value. For stochastic methods, a (b) denotes mean (standard deviation) of results.

Next, we compared the performance of the hybrid strategies with the results obtained from each algorithm applied alone. For the prostate case with 12 beam angles, the hybrid methods appeared to be more efficient than the stand-alone algorithms; we observed 62%, 57%, and 45% reductions in computation times for SA-LNS, GA-LNS, and NP-LNS, respectively, compared to SA, GA, and NP applied alone. For the 12-beam configuration, LNS improved the objective value when initialized by the solution given by GA. For the 36-beam configuration, the hybrid approaches improved objective values and reduced the computational effort when compared to the stand-alone algorithms (Table 5).

To further illustrate the performance of the hybrid approaches, Figure 1 compares the objective values from LNS solved with starting points obtained from 20 random runs of SA, GA, and NP and one run of BP on the IMRT prostate case with 36 candidate angles. In all cases, local optimal values achieved by LNS ranged from 0.248 to 0.263, while the starting solutions ranged from 0.265 to 0.349. It is interesting to note that GA-LNS has a much smaller range in the final objective values than their counterparts regardless of the starting solutions used. We have observed a similar result for the 12 candidate angles scenario (see Figure 5 in Appendix). Although the starting solutions were selected from a wider range of objective values, the LNS consistently converged to local optimal solutions within a narrow range (or a unique optimum for GA-LNS). Furthermore, the global optimal solution was reached in both GA and BP methods.

The quality of the plans can be evaluated using various dose-volume metrics. One metric often used for plan evaluation is the percentage of the PTV that receives the prescribed dose (typically, 95%

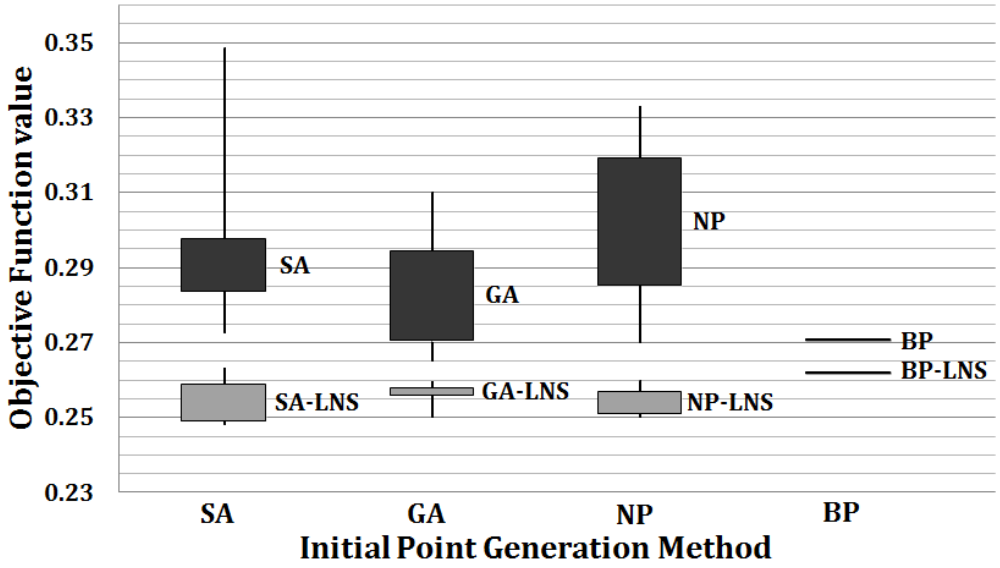

Figure 1: LNS objective value comparison for the IMRT prostate case with 36 candidate angles. The upper black boxes are the objective values of the starting points obtained from SA, GA, NP, and BP, and the lower gray boxes are the objective values obtained from LNS.

of the PTV is required to be covered by the prescription dose). In our analysis, in order to make a fair comparison among different algorithms and to be consistent with the common practice of treatment planning, all plans were normalized in such a way that 95% of the PTV receives the prescription dose. We then evaluated the volume of the PTV receiving at least 110% of the prescribed dose in order to assess the uniformity of dose distribution within the PTV. Table 6 summarizes the achieved dose−volume results for each 36-beam-angle solution method in the IMRT prostate case. We can see from Table 6 that better rectum sparing and PTV dose uniformity were achieved by using the hybrid approaches instead of the stand-alone methods.

4.2 IMRT: Pancreas Case

All solution methods with the same parameters were applied to the pancreas case. Complete numerical results are tabulated in Table 7 for 12- and 36-candidate-angle configurations. For the 12-angle configuration, the MIP model achieved the optimal objective function value of 0.110. In this case, BP found the optimal solution. Among the stochastic methods, SA performed very well with respect to the objective value. Similar average objective values were obtained from NP and GA, but GA showed a higher deviation than NP. Among all single-solution methods, BP attained the lowest computational time. All of the methods resulted

Table 6: Dose-volume results obtained from the BAO solution methods in the prostate case (IMRT) with 36 candidate angles

Solution method	% volume of PTV receiving at least 110% of the prescribed dose	% volume of rectum receiving at least 60 Gy
BB	NA	NA
BP	13.2	6.8
SA	15.4 (3.2)	5.7 (0.9)
GA	13.8 (1.3)	6.4 (0.3)
NP	15.4 (3.0)	6.7 (3.03)
BP-LNS	12.4	2.2
SA-LNS	11.1 (0.4)	2.6 (0.9)
GA-LNS	11.5 (1.1)	2.8 (1.8)
NP-LNS	11.7 (1.5)	3.1 (1.4)

For stochastic methods, a (b) denotes mean (standard deviation) of results.

in lower solution times compared to the MIP model, except for SA. SA had the worst time performance for

Table 7: Comparison of results of stand-alone and hybrid solution methods based on 12- and 36-candidate-beam-angle configurations for the pancreas cancer case (IMRT)

η=4		$\|A\| = 12$		$\|A\| = 36$	
Solution method		Obj.	Time (min)	Obj.	Time (min)
Deterministic	BB	0.110	50.9	NA	NA
	BP	0.110	17.1	0.072	273.3
Stochastic	SA	0.112 (0.004)	102.3 (40.6)	0.083 (0.012)	112.8 (6.4)
	GA	0.118 (0.013)	35.9 (11.3)	0.115 (0.020)	46.6 (2.1)
	NP	0.117 (0.005)	35.9 (4.1)	0.079 (0.008)	54.6 (1.8)
Hybrid	BP-LNS	0.110	30.3	0.071	45.3
	SA-LNS	0.110 (0.000)	43.6 (12.6)	0.069 (0.001)	91.5 (19.7)
	GA-LNS	0.110 (0.000)	34.6 (4.3)	0.071 (0.003)	78.9 (9.7)
	NP-LNS	0.112 (0.004)	41.3 (6.7)	0.073 (0.005)	60.1 (14.7)

Obj., objective function value. For stochastic methods, a (b) denotes mean (standard deviation) of results.

this case: almost 100% worse than the MIP model and at least 184% worse than the other methods.

In the 36-angle setup, the MIP model was run for 1 day. We recorded the objective value of 0.075 when the dual gap was 37.55%. Among the stand-alone methods, the best objective value (0.072) was obtained by BP, but the CPU time for this method was rather high. GA had the lowest computational time, but its objective value was worse than those of BP, NP, and SA. Among the stochastic methods, SA again consumed more CPU time than the other methods. Considering the hybrid approaches, for the 12-angle set-up, the hybrid methods consistently produced better quality solutions with smaller deviations than the

stand-alone methods. The solution time for SA-LNS was significantly lower than that of the single SA method. With 36 beam angles, every hybrid approach outperformed the respective stand-alone methods and provided relatively lower objective function values. Furthermore, the objective value deviations for the hybrid methods were rather small, especially when compared with the results from the stand-alone methods.

Figure 2 compares objective values obtained from LNS using the starting solutions provided by 20 random runs of SA, GA, and NP and one run of BP on the pancreas case with 36 candidate angles. With 36 candidate angles, in all cases local optimal values achieved by LNS ranged from 0.068 to 0.075, while the starting solutions ranged from 0.072 to 0.206. For this case, the LNS converged to solutions with a very narrow objective value range although some starting solutions were quite off from the final solutions. For the 12 candidate angles scenario, LNS was able to find the same local optimal solution in all runs, except for one run in NP, with the objective value of 0.120 (see Figure 6 in Appendix). In fact, the local optimal solution turned out to be the global optimum.

Table 8 shows the dose-volume results obtained from the solution methods in the pancreas case for the 36-angle configuration. As can be seen from Table 8, the hybrid approaches had advantages over the stand-alone methods with respect to dose uniformity within the target. The maximum dose delivered to one of the OARs, the spinal cord, remained below its allowed limit of 45 Gy for all solution methods and improved up to 7.5% with hybrid approaches.

Table 8: Dose-volume results obtained from the BAO solution methods in the pancreas case (IMRT) with 36 candidate angles

Solution method	% volume of PTV receiving at least 105% of the prescribed dose	Maximum spinal cord dose (Gy)
BB	NA	NA
BP	4.2	11.9
SA	5.0 (0.9)	12.1 (1.6)
GA	5.4 (0.5)	12.0 (0.6)
NP	4.9 (0.9)	12.2 (0.8)
BP-LNS	4.3	11.4
SA-LNS	4.0 (1.2)	11.2 (0.1)
GA-LNS	4.3 (0.9)	11.6 (0.0)
NP-LNS	4.4 (1.1)	11.5 (0.1)

For stochastic methods, a (b) denotes mean (standard deviation) of results.

For this case, the LP model was solved using the dual-simplex method for all solution methods except

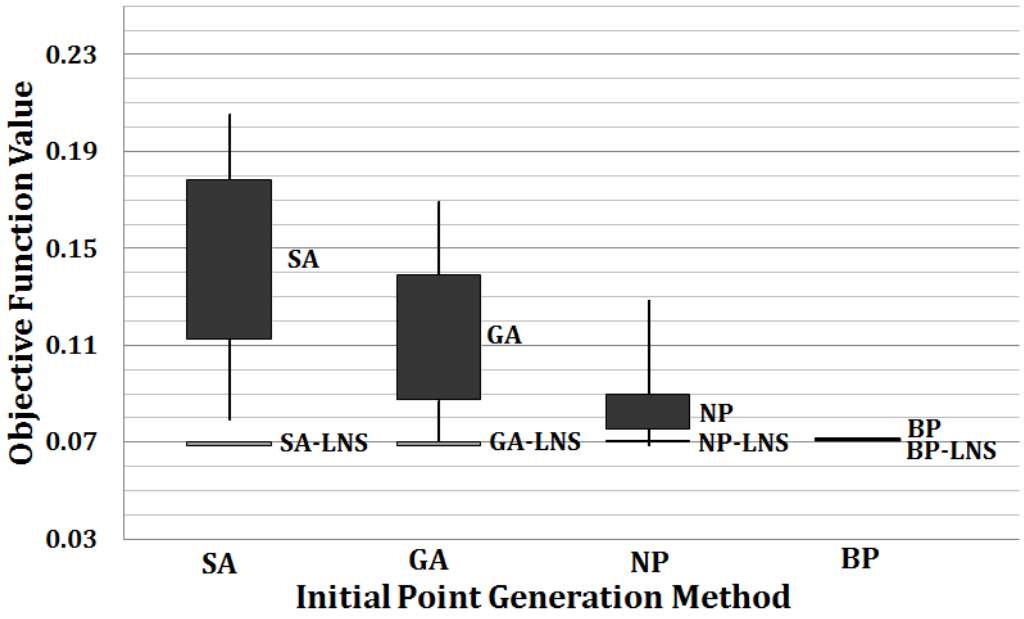

Figure 2: LNS Objective value comparison for the IMRT pancreas case with 36 candidate angles. The upper black boxes are the objective values of the starting points obtained from SA, GA, NP, and BP and the lower lines are the objective values obtained from LNS.

for BP, which was solved using the interior point method. The interior point method is known to be a powerful tool for solving large-scale linear programming models. However, we observed a drawback when the problem was infeasible, which occurred for some of the selected angle sets in the pancreas cancer case because of the size and relative position of the target and OARs. When the LP model was solved for the pancreas data using the interior point method, the solution time was unreasonably high for most of the methods used in this paper; it took an unexpectedly long time to conclude that the corresponding angle set was infeasible. On the other hand, the dual-simplex method quickly determined the infeasibility of the problem, which helped in terminating the algorithms faster (Table 9). BP, however, never encountered infeasible solutions in our preliminary experiments, and we noted much faster solution times with the interior point method. This observation could highlight the importance of using score function as a measure of angle quality in the BP algorithm.

4.3 IMPT: Prostate Case

To further evaluate the solution methods and the effectiveness and robustness of the hybrid approaches, they were applied to another prostate cancer case for IMPT treatment planning. Objective values and compu-

Table 9: Solution time comparison using the dual-simplex and interior point methods for the pancreas cancer case with 36 candidate beam angles (IMRT)

Solution method	CPU time (min)	
	Dual-simplex method	Interior point method
BP	650.7	273.3
NP	54.6	102.3
SA	112.8	680.2
GA	46.6	180.6

tation times for the different methods are presented in Table 10. For the 12-angle setup, the MIP (BB) model obtained the optimal objective value of 0.165 after almost 23 hours of computation time, which is too long to be practical. This long computation is mainly attributed to the large amount of data in IMPT treatment planning. Among the stand-alone methods, SA found the optimal solution, while GA produced a near-optimal solution. NP resulted in the shortest computation time, but it returned the highest objective value. All of the hybrid approaches returned the optimal solution at a slight cost of computation time when compared to their stand-alone counterparts. However, all of the hybrid approaches consumed significantly less computation time compared to that of the MIP.

Table 10: Comparison of results of stand-alone and hybrid solution methods based on 12- and 36-candidate-beam-angle configurations for the prostate cancer case (IMPT)

η=3		$\|A\| = 12$		$\|A\| = 36$	
Solution method		Obj.	Time (min)	Obj.	Time (min)
Deterministic	BB	0.165	1393.2	NA	NA
	BP	0.170	420.2	NA	NA
Stochastic	SA	0.165 (0.000)	163.9 (23.7)	0.161 (0.003)	214.5 (47.2)
	GA	0.167 (0.002)	58.5 (10.2)	0.158 (0.001)	111.1 (14.7)
	NP	0.174 (0.005)	30.9 (9.5)	0.171 (0.002)	44.5 (3.7)
Hybrid	BP-LNS	0.165	444.7	0.159	456.49
	SA-LNS	0.165 (0.000)	174.1 (11.4)	0.154 (0.004)	182.3 (22.8)
	GA-LNS	0.165 (0.000)	70.9 (8.6)	0.155 (0.001)	118.2 (13.5)
	NP-LNS	0.165 (0.001)	66.7 (6.4)	0.159 (0.005)	64.4 (22.5)

Obj., objective function value. For stochastic methods, a (b) denotes mean (standard deviation) of results.

For the 36-angle configuration, CPLEX was unable to find an integer solution for the MIP model within 1 day of running time. However, due to the small number of possible beam angle combinations, an exhaustive search method was used to find the global optimal solution by enumerating and comparing the objective functions for all 7,140 possible solutions. Using this method, the global optimal solution, with an objective

value of 0.152, was obtained in approximately 3 days. Note that in this study, the sole purpose of employing this exhaustive search method was to use its global optimal solution as the benchmark to examine the efficiency and measure the quality of the solutions obtained using other methods.

Again due to the large amount of data in one LP model, the BP algorithm was unable to find a feasible solution for the 36-angle configuration after 1 day of running time. For stochastic stand-alone approaches, GA achieved the lowest objective value (0.158), while NP had the fastest solution time. The objective values of all hybrid approaches were consistently better than those of the corresponding stand-alone methods, with a slight increase in the CPU run time (Table 10).

Table 11 shows the radiation dose-volume results from BAO solution methods for IMPT prostate case with 36 candidate angles. As can be seen from Table 11, the rectum and bladder sparing improved by up to 6.5% and 4.8%, respectively, through the use of hybrid plans.

Table 11: Dose-volume results obtained from the BAO solution methods in the prostate case (IMPT) with 36 candidate angles

Solution method	% volume of PTV receiving at least 105% of prescription dose	% volume of rectum receiving at least 60 Gy	% volume of bladder receiving at least 70 Gy
BB	NA	NA	NA
BP	NA	NA	NA
SA	0.0 (0.0)	9.0 (0.4)	8.2 (0.5)
GA	0.0 (0.0)	9.0 (0.1)	8.2 (0.2)
NP	0.0 (0.0)	9.3 (0.2)	8.3 (0.3)
BP-LNS	0.0 (0.0)	8.3	7.9
SA-LNS	0.0 (0.0)	8.7 (0.1)	7.9 (0.1)
GA-LNS	0.0 (0.0)	8.7 (0.2)	7.9 (0.2)
NP-LNS	0.0 (0.0)	8.7 (0.4)	7.9 (0.2)

For stochastic methods, a (b) denotes mean (standard deviation) of results.

4.4 IMPT: Head-and-Neck Case

The solution methods were compared for a head-and-neck cancer case (Table 12). For the 12-angle configuration, the optimal solution, with objective value of 0.235, was found for the MIP model. Among the stand-alone methods, GA produced the best solution, followed by SA and NP. The BP approach was deemed the most undesirable algorithm for this case since it produced the worst result in terms of objective function value. As shown in Table 12, all of the hybrid approaches performed well; the global optimal value was

captured in all cases, in a fraction of the solution time used by BB.

For the 36-angle configuration, we let CPLEX run the MIP (BB) model with the time limit of 1 day. Since the data for this instance was small, CPLEX was able to find the optimal solution to the MIP model in 920.3 minutes with an objective value of 0.206. Among the stand-alone methods, the best solution was generated by the SA method. For all cases, the hybrid approaches resulted in near-optimal solutions, while using a much smaller amount of computational resources when compared to BB. Up to 16% improvement in the objective function value was achieved compared to stand-alone methods. These results demonstrate the superiority of the hybrid approaches in comparison with the stand-alone methods.

Table 12: Comparison of results of stand-alone and hybrid solution methods based on 12- and 36-candidate-beam-angle configurations for the head-and-neck cancer case (IMPT)

η=3		$\|A\| = 12$		$\|A\| = 36$	
Solution method		Obj.	Time (min)	Obj.	Time (min)
Deterministic	BB	0.235	25.3	0.206	920.3
	BP	0.253	14.9	0.259	237.2
Stochastic	SA	0.239 (0.008)	26.6 (8.1)	0.219 (0.008)	30.8 (8.6)
	GA	0.237 (0.001)	14.9 (1.8)	0.231 (0.006)	22.2 (8.3)
	NP	0.240 (0.006)	10.8 (0.5)	0.225 (0.01)	20.4 (2.2)
Hybrid	BP-LNS	0.235	21.2	0.217	21.4
	SA-LNS	0.235 (0.00)	15.0 (1.4)	0.208 (0.003)	22.24 (6.6)
	GA-LNS	0.235 (0.00)	14.7 (1.1)	0.209 (0.004)	24.8 (4.7)
	NP-LNS	0.235 (0.00)	10.3 (2.9)	0.207 (0.003)	19.8 (5.1)

Obj., objective function value. For stochastic methods, a (b) denotes mean (standard deviation) of results.

An overview of the dose-volume metrics in the head-and-neck cancer case with 36 candidate angles is given in Table 13. The hybrid plans had better PTV dose uniformity than did the single-solution plans. In all plans, the brain stem and optic chiasm were within their tolerance limits of 55 and 54 Gy, respectively. The maximum doses to the brain stem obtained from the hybrid plans were comparable to those of the stand-alone methods. The maximum doses to the optic chiasm were reduced by up to 3.5 Gy through the use of hybrid methods.

4.5 IMPT: Lung Case

Table 14 shows how different solution methods performed in a lung cancer case for IMPT planning. For both the 12- and 36-angle configurations, we set a computational time limit of 1 day. For the 12-angle setup, the optimal solution to the MIP (BB) model, computed by CPLEX, was obtained in 127.8 minutes

Table 13: Dose-volume results obtained from the BAO solution methods in the head-and-neck case (IMPT) with 36 candidate angles

Solution method	% volume of PTV receiving at least 105% of prescription dose	Maximum brain stem dose (Gy)	Maximum optic chiasm dose (Gy)
BB	13.2	54.1	38.2
BP	17.2	54.3	45.1
SA	17.7 (2.4)	54.6 (0.2)	41.1 (2.8)
GA	17.5 (2.2)	53.9 (0.3)	43.7 (2.5)
NP	16.9 (2.0)	54.7 (0.3)	41.4 (2.7)
BP-LNS	16.4	54.0	43.4
SA-LNS	13.3 (0.5)	54.0 (0.0)	39.9 (3.1)
GA-LNS	13.5 (0.8)	53.9 (0.1)	40.2 (3.0)
NP-LNS	13.4 (0.5)	54.0 (0.0)	39.9 (2.9)

For stochastic methods, a (b) denotes mean (standard deviation) of results.

with an objective value of 0.138. Among the stand-alone methods, NP outperformed the others in terms of objective function value and computation time. The longest running time was observed with the BP method. However, the time needed by BP did not yield a better solution. The SA algorithm did as well as BP in about half the time used by BP.

Table 14: Comparison of results of stand-alone and hybrid solution methods based on 12- and 36-candidate-beam-angle configurations for the lung cancer case (IMPT)

η=3		$\|A\| = 12$		$\|A\| = 36$	
Solution method		Obj.	Time (min)	Obj.	Time (min)
Deterministic	BB	0.138	127.8	NA	NA
	BP	0.142	230.7	NA	NA
Stochastic	SA	0.142 (0.005)	118.6 (12.2)	0.137 (0.009)	112.2 (15.1)
	GA	0.147 (0.004)	71.4 (3.2)	0.135 (0.007)	157.2 (8.6)
	NP	0.139 (0.002)	64.3 (2.4)	0.133 (0.005)	67.5 (3.4)
Hybrid	BP-LNS	0.142	235.2	0.142	235.5
	SA-LNS	0.139 (0.002)	71.4 (7.3)	0.128 (0.002)	75.6 (9.12)
	GA-LNS	0.139 (0.001)	65.3 (4.2)	0.128 (0.002)	139.8 (8.2)
	NP-LNS	0.139 (0.001)	62.7 (7.2)	0.128 (0.001)	61.2 (5.63)

Obj., objective function value. For stochastic methods, a (b) denotes mean (standard deviation) of results.

For the 36-angle setup, the time limit was reached, and no integer solution was reported by MIP. The global optimal solution, with an objective value of 0.126, was obtained by an exhaustive search; it took approximately 2 days to solve all 7,140 LP models. For this case, as for the 12-angle configuration, NP turned out to be the best of the single methods. As shown in Table 14, the proposed hybrid approaches

achieved better results in terms of objective values and computation times than their corresponding stand-alone methods for both the 12- and 36-angle configurations.

The dose-volume results in the lung case with 36 candidate angles are presented in Table 15. Data for heart and esophagus are not shown since dose-volume metrics indicate doses well below the acceptable limits. Each of the treatment plans obtained using the BAO solution methods met clinical dose-volume constraints. However, the hybrid plans allowed slight improvement in PTV dose uniformity and OAR sparing.

Table 15: Dose-volume results obtained from the BAO solution methods in the lung case (IMPT) with 36 candidate angles

Solution method	% volume of PTV receiving at least 110% of prescription dose	% volume of lung receiving at least 20 Gy	Maximum cord dose (Gy)
BB	NA	NA	NA
BP	NA	NA	NA
SA	17.3 (3.1)	8.6 (0.9)	21.5 (1.2)
GA	17.1 (4.1)	8.7 (1.0)	20.3 (1.1)
NP	14.7 (3.4)	8.2 (0.9)	21.5 (1.7)
BP-LNS	18.9	8.6	21.5
SA-LNS	15.2 (3.8)	8.2 (0.6)	20.4 (1.8)
GA-LNS	16.7 (0.2)	7.9 (0.7)	19.2 (2.3)
NP-LNS	14.0 (3.2)	7.9 (0.2)	20.1 (1.8)

For stochastic methods, a (b) denotes mean (standard deviation) of results.

4.6 Distributions of Optimized Beam Angles

Figures 3 and 4 compare the equally spaced beam angles and beam angles selected under different solution methods for the IMRT pancreas case and the IMPT prostate case with 36-angle configuration, respectively. These beam angles were selected based on the best results among multiple runs of each algorithm. In Figure 3, all optimized beam angles are significantly different from the uniformly spaced angles. Moreover, the beam angle methods associated with better objective values (i.e., NP, SA-LNS, GA-LNS, and NP-LNS) shared a similar pattern (i.e., in the range between 90° and 210°), while the methods with worse objective values (SA and GA) exhibited a different pattern. In Figure 4, the optimized beam angles showed evident characteristics. For example, angles from the lateral directions toward the patient body (angles in 90° and 270°) were preferred with regard to better objective values, over angles from the anterior direction (angles

adjacent to 0°). Moreover, the solutions with the two lowest objective values (SA-LNS, GA-LNS, and NP-LNS) all consisted of one lateral angle (90° or 270°) from one side (left or right) and two oblique angles from the other side.

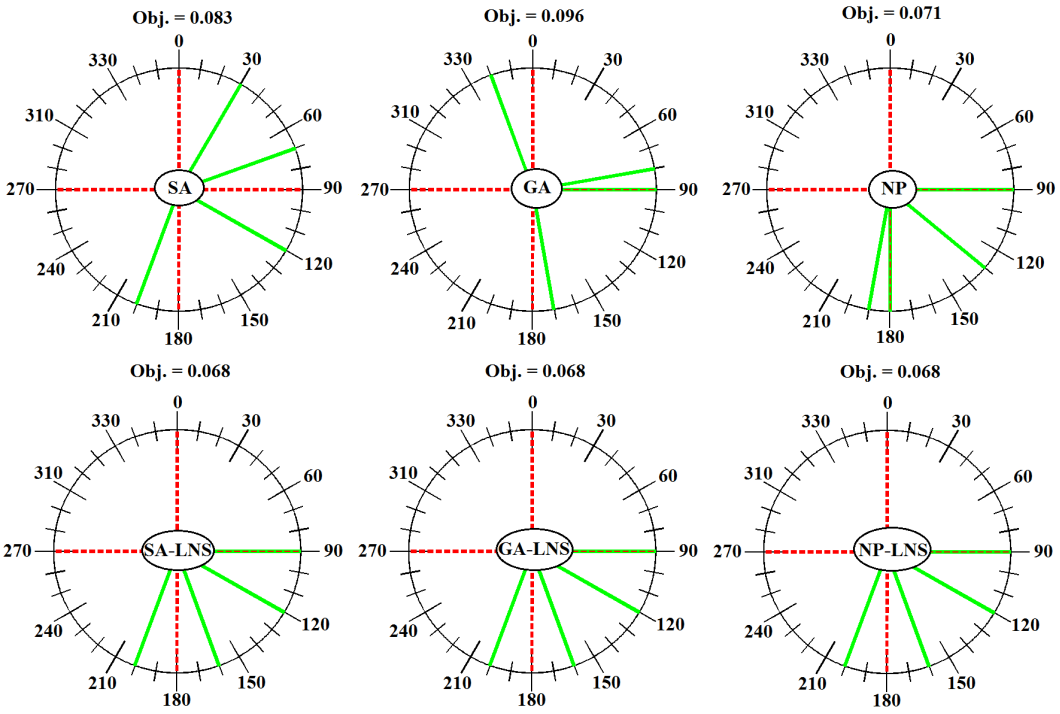

Figure 3: Beam angles selected for the IMRT pancreas case with a 36-angle configuration. Equi-spaced beam angles are denoted by dashed lines, and final beam angels by solid lines.

5 Conclusion

In this paper, a collection of previously reported solution methods for beam angle and fluence map optimization in radiation therapy treatment planning, specifically, BB, SA, GA, NP, BP, and LNS, selected from literature, were implemented and thoroughly compared on five clinical cancer cases. We found that global approaches require a considerable number of iterations to generate high-quality solutions, while local approaches yield converging solutions more quickly but were sensitive to starting points. We therefore developed a hybrid BAO framework to enhance the performance of these stand-alone methods by combining LNS with the SA, GA, NP, or BP method to take advantage of both global and local searching.

Comparing the hybrid approach results with ones obtained from the corresponding stand-alone methods on five clinical cancer cases demonstrated that the hybrid approaches are effective and robust as they

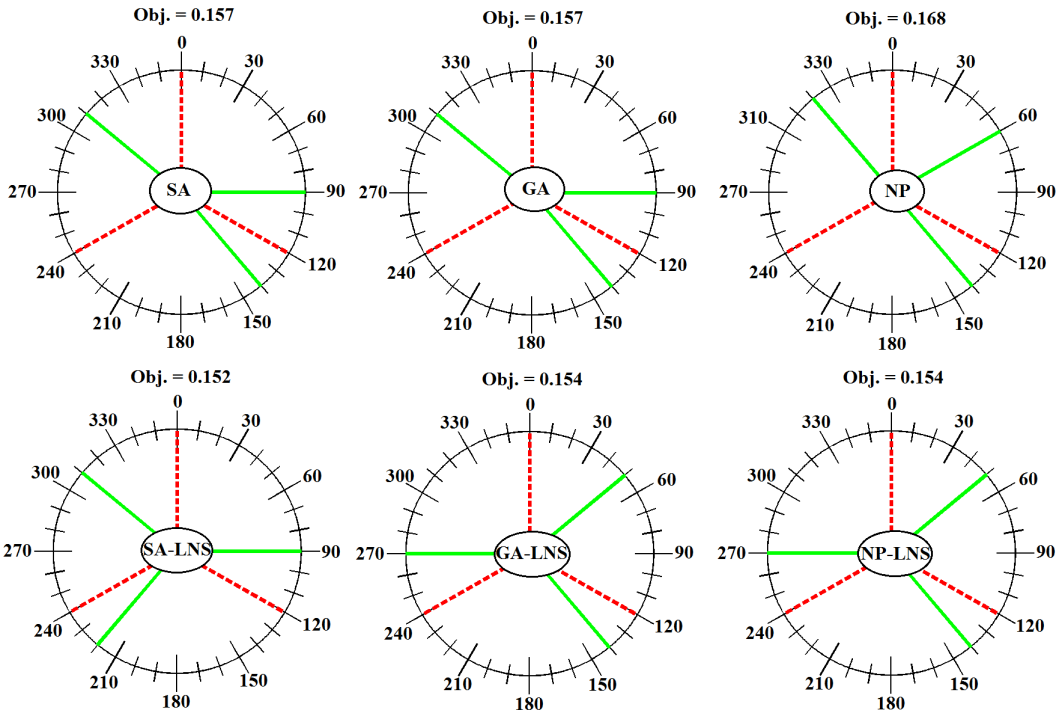

Figure 4: Beam angles selected for the IMPT prostate case with a 36-angle configuration. Equi-spaced beam angles are denoted by dashed lines, and final beam angels by solid lines.

quickly and consistently delivered good local optimal solutions. For smaller problems, e.g., problems with 12 candidate beam angles, the stand-alone solution methods tend to find reasonably good results, leaving little room for improvement by hybrid approaches. However, for larger problems, e.g., problems with 36 candidate beam angles, the hybrid methods show meaningful improvements in terms of solution quality. An advantage of the hybrid approaches was that we do not need to select the parameters of the solution methods carefully; which means hybrid strategies can relax the parameter dependence to some extent.

References

[1] S. Söderström and A. Brahme, "Which is the most suitable number of photon beam portals in coplanar radiation therapy?," *International Journal of Radiation Oncology, Biology, Physics*, vol. 33, no. 1, pp. 151–159, 1995.

[2] T. Bortfeld and W. Schlegel, "Optimization of beam orientations in radiation therapy: some theoretical considerations," *Physics in Medicine and Biology*, vol. 38, no. 2, pp. 291–304, 1993.

[3] J. Stein, R. Mohan, X. Wang, T. Bortfeld, Q. Wu, K. Preiser, C. Ling, and W. Schlegel, "Number and orientations of beams in intensity-modulated radiation treatments," *Medical Physics*, vol. 24, no. 2, pp. 149–160, 1997.

[4] A. Trofimov, P. Nguyen, J. Coen, K. Doppke, R. Schneider, J. Adams, T. Bortfeld, A. Zietman, T. DeLaney, and W. Shipley, "Radiotherapy treatment of early-stage prostate cancer with IMRT and protons: a treatment planning comparison," *International Journal of Radiation Oncology, Biology, Physics*, vol. 69, no. 2, pp. 444–453, 2007.

[5] W. Cao, G. J. Lim, A. Lee, Y. Li, W. Liu, X. R. Zhu, and X. Zhang, "Uncertainty incorporated beam angle optimization for IMPT treatment planning," *Medical Physics*, vol. 39, no. 8, pp. 5248–5256, 2012.

[6] G. Lim and W. Cao, "A two-phase method for selecting IMRT treatment beam angles: Branch-and-prune and local neighborhood search," *European Journal of Operational Research*, vol. 217, no. 3, pp. 609–618, 2012.

[7] C. Wang, J. Dai, and Y. Hu, "Optimization of beam orientations and beam weights for conformal radiotherapy using mixed integer programming," *Physics in Medicine and Biology*, vol. 48, no. 2, pp. 4065–4076, 2003.

[8] E. Lee, T. Fox, and I. Crocker, "Integer programming applied to intensity-modulated radiation therapy treatment planning," *Annals of Operations Research*, vol. 119, no. 1, pp. 165–181, 2003.

[9] A. Pugachev, A. Boyer, and L. Xing, "Beam orientation optimization in intensity-modulated radiation treatment planning," *Medical Physics*, vol. 27, no. 5, pp. 1238–1245, 2000.

[10] A. Pugachev and L. Xing, "Incorporating prior knowledge into beam orientation optimization in IMRT," *International Journal of Radiation Oncology, Biology, Physics*, vol. 54, no. 5, pp. 1565–1574, 2002.

[11] D. Djajaputra, Q. Wu, Y. Wu, and R. Mohan, "Algorithm and performance of a clinical IMRT beam-angle optimization system," *Physics in Medicine and Biology*, vol. 48, no. 19, pp. 3191–3212, 2003.

[12] D. Aleman, A. Kumar, R. Ahuja, H. Romeijn, and J. Dempsey, "Neighborhood search approaches to beam orientation optimization in intensity modulated radiation therapy treatment planning," *Journal of Global Optimization*, vol. 42, no. 4, pp. 587–607, 2008.

[13] G. Ezzell, "Genetic and geometric optimization of three-dimensional radiation therapy treatment planning," *Medical Physics*, vol. 23, no. 3, pp. 293–305, 1996.

[14] Q. Hou, J. Wang, Y. Chen, and J. Galvin, "Beam orientation optimization for IMRT by a hybrid method of the genetic algorithm and the simulated dynamics," *Medical Physics*, vol. 30, no. 9, pp. 2360–2367, 2003.

[15] X. Wu and Y. Zhu, "A mixed-encoding genetic algorithm with beam constraint for conformal radiotherapy treatment planning," *Medical Physics*, vol. 27, no. 11, pp. 2508–2506, 2000.

[16] E. Schreibmann, M. Lahanas, L. Xing, and D. Baltas, "Multiobjective evolutionary optimization of the number of beams, their orientations and weights for intensity-modulated radiation therapy," *Physics in Medicine and Biology*, vol. 49, no. 5, pp. 747–770, 2004.

[17] Y. Li, J. Yao, and D. Yao, "Automatic beam angle selection in IMRT planning using genetic algorithm," *Physics in Medicine and Biology*, vol. 49, no. 10, pp. 1915–1932, 2004.

[18] J. Lei and Y. Li, "An approaching genetic algorithm for automatic beam angle selection in IMRT planning," *Computer Methods and Programs in Biomedicine*, vol. 93, no. 3, pp. 257–265, 2009.

[19] H. Zhang, L. Shi, R. Meyer, D. Nazareth, and W. D'Souza, "Solving beam-angle selection and dose optimization simultaneously via high-throughput computing," *INFORMS Journal on Computing*, vol. 21, no. 3, pp. 427–444, 2009.

[20] Y. Li, D. Yao, J. Yao, and W. Chen, "A particle swarm optimization algorithm for beam angle selection in intensity-modulated radiotherapy planning," *Physics in Medicine and Biology*, vol. 50, no. 15, pp. 3491–3514, 2005.

[21] D. M. Aleman, H. E. Romeijn, and J. F. Dempsey, "A response surface approach to beam orientation optimization in intensity-modulated radiation therapy treatment planning," *INFORMS Journal on Computing*, vol. 21, no. 1, pp. 62–76, 2009.

[22] C. Rowbottom, S. Webb, and M. Oldham, "Beam-orientation customization using an artificial neural network," *Physics in Medicine and Biology*, vol. 44, no. 9, pp. 2251–2262, 1999.

[23] G. J. Lim, L. Kardar, W. Cao, and A. Kelkar, "Radiation therapy beam angle optimization methods: A comparison study," *Proceedings of the Industrial and Systems Engineering Research Conference (ISERC)*, G. Lim and J.W. Herrmann, eds., ID 122, 2012.

[24] V. Misic, D. Aleman, and M. Sharpe, "Neighborhood search approaches to non-coplanar beam orientation optimization for total marrow irradiation using IMRT," *European Journal of Operational Research*, vol. 205, no. 3, pp. 522–527, 2010.

[25] C.-H. J. Lee, D. M. Aleman, and M. B. Sharpe, "A set cover approach to fast beam orientation optimization in intensity modulated radiation therapy for total marrow irradiation," *Physics in Medicine and Biology*, vol. 56, no. 17, pp. 5679–5695, 2011.

[26] M. Bangert and U. Oelfke, "Spherical cluster analysis for beam angle optimization in intensity-modulated radiation therapy treatment planning," *Physics in Medicine and Biology*, vol. 55, no. 19, pp. 6023–6037, 2010.

[27] M. Ehrgott, A. Holder, and J. Reese, "Beam selection in radiotherapy design," *Linear Algebra and its Applications*, vol. 428, no. 5, pp. 1272–1312, 2008.

[28] G. Lim, J. Choi, and R. Mohan, "Iterative solution methods for beam angle and fluence map optimization in intensity modulated radiation therapy planning," *OR Spectrum*, vol. 30, no. 2, pp. 289–309, 2008.

[29] R. Eglese, "Simulated annealing: a tool for operational research," *European Journal of Operational Research*, vol. 46, no. 3, pp. 271–281, 1990.

[30] L. Shi and S. Ólafsson, "Nested partitions method for global optimization," *Operations Research*, vol. 48, no. 3, pp. 390–407, 2000.

[31] W. Cao, G. Lim, Y. Li, and X. Z. Zhang, "Using beam angle optimization to improve treatment plan quality of intensity modulated proton therapy (IMPT) for prostate cancer," *Medical Physics*, vol. 38, no. 6, p. 3741, 2011.

Appendix

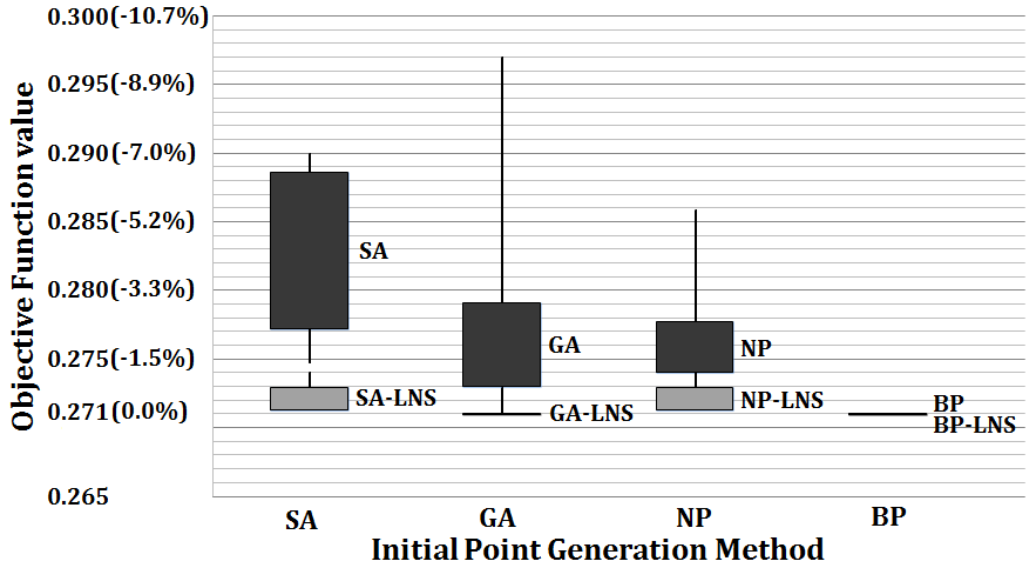

Figure 5: LNS objective value comparison for the IMRT prostate case with 12 candidate angles. The upper black boxes are the objective values of the starting points obtained from SA, GA, NP, and BP, and the lower gray boxes are the objective values obtained from LNS. Numbers in parentheses are the percentages of objective function deterioration with respect to optimal value.

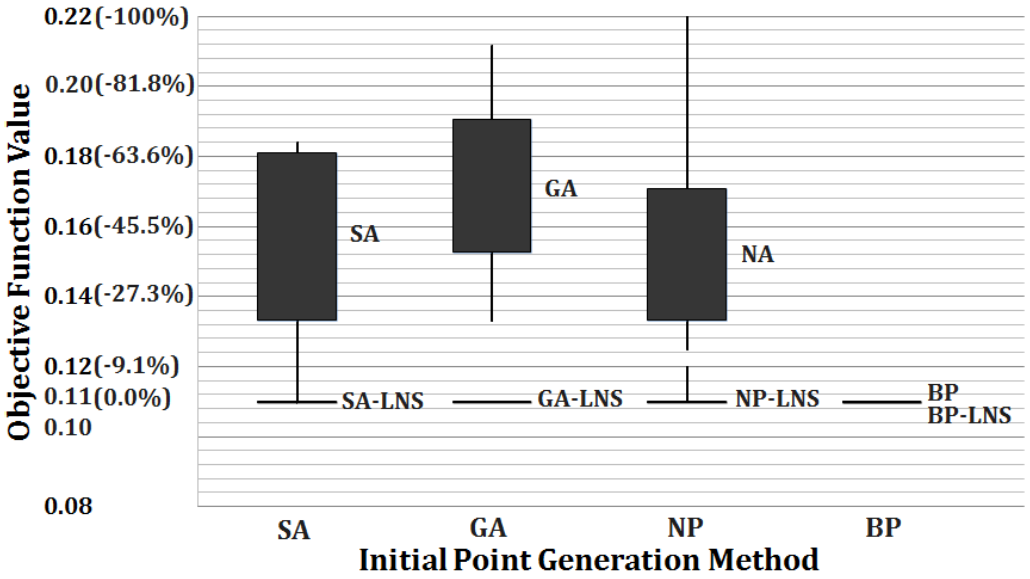

Figure 6: LNS Objective value comparison for the IMRT pancreas case with 12 candidate angles. The upper black boxes are the objective values of the starting points obtained from SA, GA, NP, and BP and the lower lines are the objective values obtained from LNS. Numbers in parentheses show the percentage of objective function deterioration with respect to optimal value.