A theoretical stochastic control framework for adapting radiotherapy to hypoxia

Fatemeh Saberian, Archis Ghate, Minsun Kim

May 26, 2015

Abstract

Hypoxia, that is, insufficient oxygen partial pressure, is a known cause of reduced radiosensitivity in solid tumors. It is thus believed to adversely affect the outcome of fractionated radiotherapy. Oxygen partial pressure varies spatially and temporally over the treatment course and exhibits inter-patient and intra-tumor variation. Emerging advances in non-invasive functional imaging offer the future possibility of adapting radiotherapy plans to this uncertain spatiotemporal evolution of hypoxia over the treatment course.

We study the potential benefits of such adaptive planning via a theoretical stochastic control framework using computer-simulated evolution of hypoxia on computer-generated test-cases in head-and-neck cancer. Exact solution of the resulting control problem is computationally intractable. We develop an approximation algorithm, called certainty equivalent control, that calls for the solution of a sequence of convex programs over the treatment course; dose-volume constraints are handled using a simple constraint generation method. The convex programs are solved using an interior point algorithm with a logarithmic barrier via Newton’s method and backtracking line search.

We perform numerical experiments on four test cases by using a first-order vector autoregressive process with exponential and rational-quadratic covariance functions from the spatiotemporal statistics literature to simulate hypoxia evolution. Our results suggest that dynamic planning could lead to a considerable improvement in the number of tumor cells remaining at the end of the treatment course. Through these simulations, we also gain insights into when and why dynamic planning is likely to yield the largest benefits.

1 Introduction

Hypoxia, that is, an inadequate supply of oxygen in living tissue, has long been known to adversely affect the outcome of radiotherapy in solid tumors and especially in head-and-neck tumors [19, 20, 29, 32, 39, 40, 51, 52, 53, 65]. Specifically, hypoxia induces radioresistance in these tumor cells and hence a higher radiation dose may be needed to achieve the same degree of cell-kill as compared to that for well-oxygenated tumor cells.

Oxygen partial pressure within tumors varies both spatially and temporally over the treatment course [30, 38, 51, 52, 53]. Moreover, there could be a considerable inter-patient variation in intra-tumor oxygen partial pressure [1, 51]. Similarly, there could be a variation in oxygen partial pressure among tumors that are otherwise histologically identical [44]. Intensity Modulated Radiation Therapy (IMRT) can in principle be used to selectively boost radiation dose to tumor subregions that are known to be hypoxic, with the hope of improving tumor control in individual patients [2, 51, 52, 53, 54, 67]. However, traditionally, invasive techniques were needed to measure spatiotemporal distribution of oxygen partial pressure in solid tumors [1, 11, 28, 30, 38, 52, 62]. This curtailed the prevalence of such dose boosting in practice [11].
Non-invasive functional imaging techniques are now becoming available for directly or indirectly assessing intra-tumor oxygen partial pressure and other relevant biological information in vivo [10, 11, 22, 27, 36, 43, 55, 56]. Examples include magnetic resonance spectroscopic imaging (MRSI), single photon emission computed tomography (SPECT), and positron emission tomography (PET). For instance, $^{18}$F-fluoromisonidazole (FMISO) is an extensively studied PET agent whose intracellular retention depends on oxygen concentration. FMISO was the first radiopharmaceutical employed to quantitatively image hypoxia in human tumors with PET [11, 26, 44]. Similarly, PET imaging with the $^{18}$F-2-deoxyglucose (FDG) tracer, and MRSI with choline/citrate ratios can be used to assess spatial distribution of tumor cell density (and hence to indirectly locate radioresistant tumor regions of high cell density) [9, 54, 67, 69]. Research is currently ongoing to enhance the accuracy and precision of such functional imaging techniques. There seems to be hope in the medical community that this type of imaging will become clinically viable within a decade [51, 54, 67]. This creates the future possibility of adapting IMRT treatment plans to the uncertain, spatiotemporal evolution of oxygen partial pressure and tumor cell density as inferred from functional images acquired over the treatment course for individual patients [22, 51, 52, 53, 54, 57, 58, 59, 67].

Stochastic control is the appropriate mathematical framework for rigorously formulating the problem of optimally adapting IMRT plans. Thus, the aim of this study is to estimate any potential benefits of such adaptive planning using computer simulations of computer-generated head-and-neck cancer test cases within a general, theoretical stochastic control framework that we originally envisioned in [25].

2 A theoretical stochastic control formalism

The conceptual idea in the abstract stochastic control formalism that we outlined in [25] is to view the tumor and normal tissues as a dynamic system whose state evolution over the treatment course is affected stochastically by the control variables chosen by the treatment planner in different treatment sessions. The planner’s goal is to optimize some quantitative measure of treatment efficacy. Any concrete implementation of this abstract idea calls for precise mathematical definitions of the tumor state dynamics, the normal tissue state dynamics, the feasible control variables, and the treatment efficacy measure. These definitions are of course problem dependent. We formalize them in detail in the specific context of hypoxia in this section.

We consider a treatment course with $T$ sessions indexed by $t = 1, 2, \ldots, T$. The number of beamlets in the radiation field is $K$ and we use $u^t \in \mathbb{R}^K_+$ to denote the fluence-map, that is, the beamlet intensity vector, employed in the $t$th treatment session.

2.1 Tumor state dynamics

In our model, the tumor state includes the cell density and oxygen partial pressure in all tumor voxels as observed in functional images taken prior to a treatment session. The tumor is assumed to include $n$ equal-volume voxels (each with volume $\nu$) indexed by $i = 1, 2, \ldots, n$. Let $A$ denote the $n \times k$ tumor dose deposition coefficient matrix and let $A_i$ denote its $i$th row that corresponds to the $i$th tumor voxel. That is, according to the linear dose deposition model, $A_i u^t$ is the dose deposited in the $i$th tumor voxel in the $t$th session. Let $x^t_i$ denote the tumor cell density in tumor voxel $i$ as assessed from a functional image acquired at the beginning of the $t$th session.

We model the evolution of tumor cell density in each voxel using the standard linear-quadratic (LQ) dose-response model from radiobiology [18]. Let $\alpha^t_i$ and $\beta^t_i$ denote the radiosensitivity parameters of the LQ model in session $t$ for tumor voxel $i$. The dependence of these radiosensitivity
parameters on oxygen partial pressure is characterized using the standard notion of Oxygen Enhancement Ratio (OER) [51, 52, 65]. Specifically, let $y_i^t$ denote the oxygen partial pressure in tumor voxel $i$ as assessed from a functional image acquired at the beginning of session $t$. Then, we have, for $i = 1, 2, \ldots, n$,

$$
\alpha_i^t = \frac{\alpha}{\text{OER}_\alpha} \left( \frac{y_i^t \text{OER}_\alpha + \kappa}{y_i^t + \kappa} \right), \quad \text{and}
$$

$$
\beta_i^t = \frac{\beta}{(\text{OER}_\beta)^2} \left( \frac{y_i^t \text{OER}_\beta + \kappa}{y_i^t + \kappa} \right)^2,
$$

where $\alpha$ and $\beta$ are radiosensitivity parameters under well-oxygenated conditions, and OER$_\alpha$, OER$_\beta$, and $\kappa$ are other parameters of the OER model. The tumor cell dynamics are then given by the standard LQ formula

$$
x_{i}^{t+1} = x_{i}^{t} \exp(-\alpha_i^t(A_i u^t) - \beta_i^t(A_i u^t)^2), \quad i = 1, 2, \ldots, n.
$$

As in [31, 53], we ignore the effect of tumor proliferation for notational simplicity; it can, however, be incorporated without any changes to our methodology.

Currently, quantitative data on voxel-by-voxel spatiotemporal evolution of oxygen partial pressure in functional images taken over the treatment course are not available to us. In addition, as stated above, there is an inter-tumor and an inter-patient uncertainty in this evolution. We therefore model the stochastic, spatiotemporal evolution of oxygen partial pressure using a random walk as an example. Specifically, we assume that the oxygen partial pressure evolves according to

$$
\ln y_{t+1}^i = \ln y_t^i + \theta_t^i.
$$

Here, $\theta_t^i \in \mathbb{R}^n$ is stochastic noise that we assume to be independent across $t$ with zero mean and with a spatial covariance matrix $\Sigma$. Thus (4) can also be seen as a first order vector autoregressive process. This particular choice for modeling partial pressure evolution is mainly for concreteness. It seemed more realistic than the current approach in the theoretical literature of sampling hypoxia values as independent and identically distributed random variables over space and time [60]. This random walk has the added benefit that the resulting probability distribution of oxygen partial pressure well-approximates the lognormal distribution (existing clinical studies have employed the lognormal distribution to fit oxygen partial pressure data [1, 28, 60]). Our stochastic control formalism can be easily generalized to any Markovian model of oxygen partial pressure evolution. In fact, as we shall see in Remark 3.1, our treatment plan optimization algorithm is model-free in that it does not actually require an explicit model of hypoxia evolution. In this sense, our random walk model here serves as a means to simulate hypoxia evolution on a computer within our theoretical framework so that we could test our optimization algorithm without real hypoxia images.

In summary, the cell density vector $x^t \triangleq (x_1^t, \ldots, x_n^t) \in \mathbb{R}_+^n$ and the oxygen partial pressure vector $y^t \triangleq (y_1^t, \ldots, y_n^t) \in \mathbb{R}_+^n$ define the tumor states in our model with dynamics given by Equations (3) and (4).

### 2.2 Control variables and normal tissue dynamics

The control variables in our stochastic control formalism correspond to the fluence-maps $u^t$ employed in different treatment sessions. Feasible fluence-maps are subject to maximum dose, mean dose and dose-volume constraints on serial and parallel normal tissues. To explicitly model the
temporal component of the problem, these constraints are expressed in terms of biologically effective dose (BED) equivalents. We also include smoothness constraints on the fluence-maps. The details of these components of our formulation are described in this section.

Let $O_1, O_2, \ldots, O_M$ denote the $M$ different normal tissues under consideration. Let $\mathcal{M} = \{1, \ldots, M\}$ be the set of indices of these normal tissues. For $m \in \mathcal{M}$, let $n_m$ denote the number of voxels in $O_m$. These voxels are indexed by $j = 1, 2, \ldots, n_m$. Let $\mathcal{N}_m$ denote the set $\{1, 2, \ldots, n_m\}$ of these voxels. All normal tissue voxels are assumed to have equal volume. Let $A^m$ be the $n_m \times K$, non-negative dose deposition matrix for $O_m$. Let $A^m_j$ be the $j$th row of this matrix; this is the row that corresponds to the $j$th voxel in $O_m$. That is, $A^m_j u$ is the dose delivered to the $j$th voxel in $O_m$ in each session.

Let $\alpha_m$ and $\beta_m$ be the parameters of the LQ model for normal tissue $O_m$ and we define $\rho_m \triangleq 1/(\alpha_m/\beta_m)$. Then, the BED corresponding to the dose delivered to the $j$th voxel in $O_m$ in the $t$th session is given by

$$ D^m_j = (A^m_j u^t) + \rho_m (A^m_j u^t)^2. \quad (5) $$

The set of feasible fluence-maps in session $t$ is defined through BED and smoothness constraints as described next.

### 2.2.1 Maximum BED constraints for serial normal tissues

Let $\mathcal{M}_1 \subseteq \mathcal{M}$ be the set of indices of serial normal tissues for which we wish to include maximum dose constraints. These are the normal tissues whose function is hampered even when a small region is damaged by radiation. For example, maximum dose constraints are included on the spinal cord in head-and-neck cancer. Suppose for any $m \in \mathcal{M}_1$, that a total dose $D_{\text{max}}^m$ is known to be tolerated by each voxel in $O_m$ if administered in $N^m_{\text{conv}}$ equal-dose fractions. The BED corresponding to this schedule equals

$$ \text{BED}^m_{\text{max}} = D_{\text{max}}^m(1 + \rho_m(D_{\text{max}}^m/N^m_{\text{conv}})). \quad (6) $$

We use the standard approach of comparing normal tissue BED. In particular, let $z_j^{t,m}$ denote the BED of doses delivered to voxel $j$ in $O_m$ in the first $t - 1$ sessions. Then, a dose of $(A^m_j u^t)$ can be tolerated by the $j$th voxel in $O_m$ in session $t$ if

$$ z_j^{t,m} + (A^m_j u^t) + \rho_m (A^m_j u^t)^2 \leq \text{BED}^m_{\text{max}}, \quad j = 1, 2, \ldots, n_m. \quad (7) $$

Thus, for each $m \in \mathcal{M}_1$, we define the BED vectors $z^{t,m} \triangleq (z_1^{t,m}, \ldots, z_{n_m}^{t,m})$ as the state of $O_m$ at the beginning of session $t$.

### 2.2.2 Mean BED constraints for parallel normal tissues

Let $\mathcal{M}_2 \subseteq \mathcal{M}$ be the set of indices of parallel normal tissues for which we wish to include mean dose constraints. These are the normal tissues where a sufficiently small portion can be damaged without affecting the organ function. For example, mean dose constraints are included on the parotid glands in IMRT for head-and-neck cancer. Suppose for any $m \in \mathcal{M}_2$, that mean dose $D_{\text{mean}}^m$ is known to be tolerated by $O_m$ if administered in $N^m_{\text{conv}}$ equal-dose fractions. The BED corresponding to this mean dose is given by

$$ \text{BED}^m_{\text{mean}} = D_{\text{mean}}^m(1 + \rho_m(D_{\text{mean}}^m/N^m_{\text{conv}})). \quad (8) $$

Let $v^{t,m}$ denote the mean BED over all voxels in $O_m$ of doses delivered in the first $t - 1$ sessions. These will be included among our normal tissue state variables. Then, doses $(A^m_j u^t)$ delivered to
voxels $j \in \mathcal{N}_m$ in session $t$ can be tolerated by $O_m$ if
\[
\sum_{j=1}^{n_m} (A^m_ju^t_j) + \rho_m \sum_{j=1}^{n_m} (A^m_ju^t_j)^2 \leq \text{BED}_{\text{mean}}^m.
\] (9)

Thus, for each $m \in \mathcal{M}_2$, we define the mean BED $v^{t,m}$ as the state of $O_m$ at the beginning of session $t$.

### 2.2.3 Dose-volume constraints for parallel normal tissues

Let $\mathcal{M}_3 \subseteq \mathcal{M}$ be the set of indices of normal tissues with dose-volume constraints. Examples include rectum, bladder, and lung. In particular, suppose for any $m \in \mathcal{M}_3$, that no more than a volume fraction $\phi_m$ of normal tissue $O_m$ can receive a dose more than $D_{\text{dv}}^m$ if administered in $N_{\text{conv}}^m$ fractions. The BED of total dose $D_{\text{dv}}^m$ administered in $N_{\text{conv}}^m$ equal-dosage fractions is given by
\[
\text{BED}_{\text{dv}}^m = D_{\text{dv}}^m(1 + \rho_m(D_{\text{dv}}^m/N_{\text{conv}}^m)).
\] (10)

Let $w_{j}^{t,m}$ denote the BED of doses delivered to voxel $j$ in $O_m$ in the first $t-1$ sessions. Since all voxels in $O_m$ have equal volume, the volume fraction is the same as the voxel fraction. For each $m \in \mathcal{M}_3$ and for $j = 1, 2, \ldots, n_m$, we thus define binary-valued functions $f^m_j(u^t; w^{t,m})$ such that
\[
f^m_j(u^t; w^{t,m}) = \begin{cases} 
1 & \text{if } w^{t,m} + (A^m_ju^t_j) + \rho_m(A^m_ju^t_j)^2 > \text{BED}_{\text{dv}}^m, \\
0 & \text{if } w^{t,m} + (A^m_ju^t_j) + \rho_m(A^m_ju^t_j)^2 \leq \text{BED}_{\text{dv}}^m.
\end{cases}
\] (11)

In words, $f^m_j(u^t; w^{t,m})$ is one if the BED resulting after delivering dose by fluence-map $u^t$ to voxel $j$ exceeds the tolerance $\text{BED}_{\text{dv}}^m$; $f^m_j(u^t; w^{t,m})$ is zero otherwise. We use the integer $K_m$ to denote $\lfloor n_m \phi_m \rfloor$, that is, the largest integer that is at most $n_m \phi_m$. Then, the dose-volume constraints are expressed as
\[
\sum_{j=1}^{n_m} f^m_j(u^t; w^{t,m}) \leq K_m, \quad m \in \mathcal{M}_3.
\] (12)

These constraints ensure that there are at most $K_m$ voxels for which $w^{t,m} + (A^m_ju^t_j) + \rho_m(A^m_ju^t_j)^2 > \text{BED}_{\text{dv}}^m$; in other words, there are at least $n_m - K_m$ voxels for which $w^{t,m} + (A^m_ju^t_j) + \rho_m(A^m_ju^t_j)^2 \leq \text{BED}_{\text{dv}}^m$. Thus, for each $m \in \mathcal{M}_3$, we define the BED vectors $w^{t,m} \triangleq (w_1^{t,m}, \ldots, w_{n_m}^{t,m})$ as the state of $O_m$ at the beginning of session $t$.

### 2.2.4 Fluence-map smoothness constraints

Finally, to ensure that the intensity profile is deliverable in practice using a multi-leaf collimator, we put a smoothness constraint on each radiation field $[6, 47, 63]$. In particular, for each radiation field, we bound the absolute relative difference between intensities of each pair of nearest neighbor beamlets by a fraction $\xi$. Then the smoothness constraints can be written compactly in matrix format as
\[
S u \leq 0,
\] (13)
where $S$ is a block diagonal matrix with entries $-(1+\xi)$, $(1-\xi)$, $-1$, $0$, $+1$ at appropriate locations.
2.2.5 Set of feasible fluence-map policies

Recall that the states of various normal tissues at the beginning of treatment session $t$ are denoted by $z^t_{m}$ for $m \in \mathcal{M}_1$, $v^t_{m}$ for $m \in \mathcal{M}_2$, and $w^t_{m}$ for $m \in \mathcal{M}_3$, respectively. Let $|\mathcal{M}_1|$, $|\mathcal{M}_2|$, and $|\mathcal{M}_3|$ denote the cardinalities of sets $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$, respectively. We now define the more compact normal tissue state notation $\tilde{z}^t \triangleq (z^t_{1}, \ldots, z^t_{|\mathcal{M}_1|})$, $\tilde{v}^t \triangleq (v^t_{1}, \ldots, v^t_{|\mathcal{M}_2|})$, and $\tilde{w}^t \triangleq (w^t_{1}, \ldots, w^t_{|\mathcal{M}_3|})$. Thus, the combined state of all normal tissues is written as $[\tilde{z}^t; \tilde{v}^t; \tilde{w}^t]$. The set of feasible fluence-maps in session $t$ depends on this normal tissue state through constraints (7), (9), and (12). The set of fluence-maps $u^t \in \mathbb{R}^K_+$ that satisfy these constraints as well as the smoothness constraint (13) is denoted by $\mathcal{U}^t([\tilde{z}^t; \tilde{v}^t; \tilde{w}^t]) \subset \mathbb{R}^K_+$. Then, in the standard language of the stochastic control literature [4], a feasible policy assigns a fluence-map from the set $\mathcal{U}^t([\tilde{z}^t; \tilde{v}^t; \tilde{w}^t])$ to every possible combination of tumor and normal tissue states $[x^t; y^t; \tilde{z}^t; \tilde{v}^t; \tilde{w}^t]$ in each treatment session $t = 1, 2, \ldots, T$. We denote the set of such feasible policies by $\mathcal{P}$ and generic policies in this set by $\pi$.

2.3 Bellman’s equations of dynamic programming

As in [16, 25], we use the total number of tumor cells remaining (TNTCR) as our treatment efficacy measure. That is, our goal is to minimize the TNTCR, which is given by

$$\text{TNTCR} \triangleq \sum_{i=1}^{n} \nu x_i^{T+1}. \quad (14)$$

Note that this perhaps less familiar objective is equivalent to the more common objective of maximizing tumor control probability (TCP); this is because the negative natural logarithm of TCP equals TNTCR. We employ the TNTCR objective here because it turns out to be algebraically less cumbersome than TCP while writing our models and our convexity proof in Lemma 3.3. Moreover, when the initial cell density is equal in all tumor voxels, which is an assumption we will later make as in [31] (this assumption is made merely for simplicity and can be removed easily without any changes to our solution algorithm), the percentage improvement in TNTCR achieved by one solution method over another does not depend on the value of this initial cell density; this is not the case with TCP. This feature of TNTCR made it appealing for us as compared to the TCP.

Finally, numerical calculations with TNTCR on our computer platform also seemed more reliable and less ill-conditioned than those with TCP. Also note that minimizing TNTCR is consistent in principle with the idea of maximizing the equivalent uniform dose (EUD) [7, 37, 66].

Now recall that the initial state is given by $[x^1; y^1; z^1; v^1; w^1]$; in fact, note that all components of $z^1$, $v^1$, and $w^1$ are zero since no dose is delivered prior to the first session. Let $J^1_\pi([x^1; y^1; z^1; v^1; w^1])$ denote the TNTCR achieved by the end of the treatment course if policy $\pi \in \mathcal{P}$ is implemented through the treatment course starting in state $[x^1; y^1; z^1; v^1; w^1]$ at the beginning of the first session. The problem of minimizing the TNTCR can then be formulated as the stochastic control problem

$$J^1_\pi \triangleq \min_{\pi \in \mathcal{P}} J^1_{\pi}([x^1; y^1; z^1; v^1; w^1]). \quad (15)$$

The standard approach for solving such problems, at least in theory, is to employ Bellman’s backward recursive algorithm of dynamic programming [4]. In order to present this algorithm, we define the optimal TNTCR-to-go functions $J^t([x^t; y^t; z^t; v^t; w^t])$ for all possible states $[x^t; y^t; z^t; v^t; w^t]$ and for $t = 1, 2, \ldots, T + 1$. Specifically, $J^t([x^t; y^t; z^t; v^t; w^t])$ is interpreted as the minimum TNTCR reached at the end of the treatment course given that the state at the beginning of session $t$ is
These functions can in principle be obtained by solving, in the backward order $t = T, T - 1, \ldots, 1$, the non-linear stochastic optimization problems

$$J^t([x^t; y^t; z^t; v^t; w^t]) = \min_{u^t \in \mathcal{U}^t([z^t; v^t; w^t])} \mathbb{E}\left(J^{t+1}([x^{t+1}; y^{t+1}; z^{t+1}; v^{t+1}; w^{t+1}])\right),$$

subject to

$$x_{i}^{t+1} = x_{i}^{t} \exp(-\alpha_{i}^t(A_{i}u^{t}) - \beta_{i}^t(A_{i}u^{t})^2), \quad i = 1, 2, \ldots, n,$$

$$\ln y_{i}^{t+1} = \ln y_{i}^{t} + \theta^{t},$$

$$z_{j}^{t+1,m} = z_{j}^{t,m} + (A_{j}^m u^{t}) + \rho_{m}(A_{j}^m u^{t})^2, \quad j = 1, 2, \ldots, n_m, \quad m \in \mathcal{M}_1,$$

$$v_{i}^{t+1,m} = v_{i}^{t,m} + \sum_{j=1}^{n_m} (A_{j}^m u^{t}) + \rho_{m} \sum_{j=1}^{n_m} (A_{j}^m u^{t})^2, \quad m \in \mathcal{M}_2,$$

$$w_{j}^{t+1,m} = w_{j}^{t,m} + (A_{j}^m u^{t}) + \rho_{m}(A_{j}^m u^{t})^2, \quad j = 1, 2, \ldots, n_m, \quad m \in \mathcal{M}_3,$$

for all possible states $[x^t; y^t; z^t; v^t; w^t]$ starting with the boundary condition

$$J^{T+1}([x^{T+1}; y^{T+1}; z^{T+1}; v^{T+1}; w^{T+1}]) = \sum_{i=1}^{n} \nu x_{i}^{T+1}.$$ 

The fluence $u^t \in \mathcal{U}^t([z^t; v^t; w^t])$ that achieve the minimum in (16) then define an optimal policy for the stochastic control problem (15).

Unfortunately, for our problem, this backward algorithm is not implementable in practice because it requires that the minimization in (16) be performed for an uncountable number of states. A state discretization is computationally impractical due to the curse of state-space dimensionality [41]. Even if a state discretization were possible, each minimization would require the solution of a non-linear stochastic programming problem; this is again impractical due to the curse of action-space dimensionality [41]. Thus, as in [25], we instead consider an approximate dynamic programming technique called Certainty Equivalent Control [4] that exploits the structure of our stochastic control problem and does not require discretization.

### 3 Solution method: Certainty Equivalent Control

Certainty Equivalent Control (CEC) [4] is an easy to implement approximation technique for stochastic control problems. It begins by assuming that all future noise in the problem is replaced by its nominal values, for example, the expected values, thus leading to a deterministic control problem. This control problem is then solved to obtain an optimal sequence of controls for all future periods. However, only the control obtained for the first period is implemented while the others are discarded. The system then evolves stochastically to the next state and the process is repeated until the terminal course is completed. Thus, in one complete run of CEC, a total of $T$ deterministic control problems are solved; the first one includes $T$ sessions, the second includes $T - 1$ sessions, and ultimately, the last one involves only one session.

**Certainty Equivalent Control**

**INITIALIZE:** Set $t = 1$ and begin with a given initial state $[x^{T}; y^{T}; z^{T}; v^{T}; w^{T}]$.

**DO WHILE** $t \leq T$,

1. let the state at the beginning of session $t$ be $[x^{t}; y^{t}; z^{t}; v^{t}; w^{t}]$;
2. fix the oxygen partial pressure vectors $y^t, y^{t+1}, \ldots, y^T$ at some “nominal” values $\tilde{y}^t, \tilde{y}^{t+1}, \ldots, \tilde{y}^T$. Substitute these in expressions (1) and (2) to obtain nominal radiosensitivity parameters $\hat{\alpha}_i^t, \hat{\alpha}_i^{t+1}, \ldots, \hat{\alpha}_i^T$ and $\hat{\beta}_i^t, \hat{\beta}_i^{t+1}, \ldots, \hat{\beta}_i^T$ for tumor voxels $i = 1, 2, \ldots, n$;

3. for each $m \in M_3$ and for $j = 1, 2, \ldots, n_m$, define binary-valued functions

$$g_j^m(u^t, u^{t+1}, \ldots, u^T; w^{t,m}) = \begin{cases} 
1 & \text{if } w^{t,m} + \sum_{l=t}^T [(A_j^m u^l)^2 + \rho_m (A_j^m u^l)^2] > \text{BED}_d^{m}, \\
0 & \text{if } w^{t,m} + \sum_{l=t}^T [(A_j^m u^l)^2 + \rho_m (A_j^m u^l)^2] \leq \text{BED}_d^{m};
\end{cases}$$

(23)

4. solve the deterministic optimization problem

$$(P_l) \min \sum_{i=1}^n \nu x_i^t \prod_{l=t}^T \exp(-\hat{\alpha}_i(A_i u^l) - \hat{\beta}_i(A_i u^l)^2)$$

subject to

$$\sum_{l=t}^T [(A_j^m u^l)^2 + \rho_m (A_j^m u^l)^2] \leq \text{BED}_\text{max}^m - z_j^{t,m}, \quad j = 1, 2, \ldots, n_m, \ m \in M_1,$$

(25)

$$\sum_{j=1}^{n_m} (A_j^m u^l)^2 + \rho_m \sum_{j=1}^{n_m} (A_j^m u^l)^2 \leq n_m \text{BED}_\text{mean}^m - n_m v^{t,m}, \ m \in M_2,$$

(26)

$$\sum_{j=1}^{n_m} g_j^m(u^t, u^{t+1}, \ldots, u^T; w^{t,m}) \leq K_m, \ m \in M_3,$$

(27)

$$Su^l \leq 0, \ l = t, t + 1, \ldots, T;$$

(28)

$$u^l \geq 0, \ l = t, t + 1, \ldots, T;$$

(29)

to obtain an optimal sequence of fluence-maps $u^t, u^{t+1}, \ldots, u^T$;

5. discard $u^{t+1}, u^{t+2}, \ldots, u^T$ and use fluence-map $u^t$ in session $t$;

6. calculate $\alpha_i^t$ and $\beta_i^t$ via expressions (1)-(2) using the imaged hypoxia values $y_i^t$, for $i = 1, 2, \ldots, n$; update tumor cell density and normal tissue states using $u^t$ as

$$x_i^{t+1} = x_i^t \exp(-\hat{\alpha}_i(A_i u_i^t) - \hat{\beta}_i(A_i u_i^t)^2), \ i = 1, 2, \ldots, n,$$

(30)

$$z_j^{t+1,m} = z_j^{t,m} + (A_j^m u_j^t)^2 + \rho_m (A_j^m u_j^t)^2, \ j = 1, 2, \ldots, n_m, \ m \in M_1,$$

(31)

$$u_j^{t+1,m} = u_j^{t,m} + \sum_{j=1}^{n_m} (A_j^m u_j^t)^2 + \rho_m \sum_{j=1}^{n_m} (A_j^m u_j^t)^2,$$

(32)

$$u_j^{t+1,m} = w_j^{t,m} + (A_j^m u_j^t)^2 + \rho_m (A_j^m u_j^t)^2, \ j = 1, 2, \ldots, n_m, \ m \in M_2;$$

(33)

7. sample $\theta^t$ from its probability distribution and set $\ln y^{t+1} = \ln y^t + \theta^t$;

8. update $t \leftarrow t + 1$.

END DO
Remark 3.1. If this CEC approach were implemented in practice, the fluence-map $u^t_i$ would be employed in the $t$th session and then the imaged oxygen partial pressure vector $y^t$ evolves stochastically to a new state $y^{t+1}$ that is again imaged at the beginning of the $(t+1)$st session. Consequently, an explicit model for oxygen partial pressure evolution, such as our random walk model, is not needed. Specifically, Step 7 in the above algorithm is then not needed; this step is needed in our computer simulation of this physical process based on our stochastic model of oxygen partial pressure evolution.

Remark 3.2. We emphasize that only one cell density image is needed in order to implement the CEC approach above in practice. This cell density image is acquired at the beginning of the treatment course wherein the planner can observe the initial cell density $x_1$. All subsequent cell densities are then calculated, when needed, via formula (30) using radiosensitivity parameters $\alpha^t_i$ and $\beta^t_i$ obtained by the OER expressions (1)-(2) using the imaged hypoxia values $y^t_i$, for $i = 1, 2, \ldots, n$. In fact, if one assumes that the initial tumor cell density is homogeneous across all tumor voxels, that is, if $x_1^i = c$, for $i = 1, 2, \ldots, n$, where $c$ is some fixed cell density value, then even an initial cell density image is not needed. This is because $c$ factors out from all our objective functions and thus has no effect on the fluence-maps obtained by our solution method. As such, the value of $c$ is then not needed anywhere by the above algorithm. It is for this reason that we call the above algorithm “CEC with hypoxia images” in Section 4 below.

It remains to choose nominal values of oxygen partial pressure vectors in Step 2 of the loop over sessions $t$ and then to devise an efficient procedure for solving problem $(P_t)$. We do this in the next section.

3.1 Efficient solution of problem $(P_t)$

Exact solution of $(P_t)$, as formulated in the above CEC algorithm, is computationally intractable. It is well-known in the IMRT literature that dose-volume constraints are some of the most difficult to handle [45]. Even without the dose-volume constraints (27), $(P_t)$ is a non-convex problem because although the constraints are convex, the objective function is in general not. In realistic instances of $(P_t)$, the number of beamlets $K$ is likely to equal a few thousand [45]. With $T = 35$ (this is the number of fractions typical in current treatment guidelines [33]), this implies that $(P_t)$ can include about a hundred thousand variables. Also, the number of constraints can be in the tens of thousands depending on the total number of normal tissue voxels [45]. We therefore propose an efficient procedure for approximate solution of $(P_t)$.

The first step is to choose nominal values for oxygen partial pressures. Since the partial pressure map $y^t$ is already available in the state as observed in the functional image acquired prior to treatment session $t$, we set $\hat{y}^t = y^t$. Moreover, if the treatment planner were not to re-plan in subsequent sessions, it is natural to plan the remaining sessions assuming that oxygen partial pressures in the subsequent sessions will be as seen in the image at the beginning of session $t$. As a result, we also set $\hat{y}^{t+1} = \hat{y}^{t+2} = \ldots = \hat{y}^T = y^t$. Now, given that the nominal values of oxygen partial pressures in all remaining sessions are chosen to be identical, and in fact equal to $y^t$, we denote them by $\hat{y}$ for simplicity. This also implies that the radiosensitivity parameters $\hat{\alpha}^t_i, \hat{\alpha}^{t+1}_i, \ldots, \hat{\alpha}^T_i$ and $\hat{\beta}^t_i, \hat{\beta}^{t+1}_i, \ldots, \hat{\beta}^T_i$ can be simply denoted by $\hat{\alpha}_i$ and $\hat{\beta}_i$, where, from formulas (1) and (2), we have,

$$\hat{\alpha}_i = \frac{\alpha}{\text{OER}_\alpha} \left( \frac{\hat{y}_i \text{OER}_\alpha + \kappa}{\hat{y}_i + \kappa} \right),$$

and

$$\hat{\beta}_i = \frac{\beta}{\text{OER}_\beta} \left( \frac{\hat{y}_i \text{OER}_\beta + \kappa}{\hat{y}_i + \kappa} \right).$$
\[ \hat{\beta}_i = \frac{\beta}{(\text{OER}_\beta)^2} \left( \frac{\hat{y}_i \text{OER}_\beta + \kappa}{\hat{y}_i + \kappa} \right)^2, \quad (35) \]

for tumor voxels \( i = 1, 2, \ldots, n \). Now, since the tumor’s dose-response parameters are time-invariant, it is natural, although not necessarily optimal (see [48, 61]), to search for equal-dosage treatment plans. That is, we set \( u^t = u^{t+1} = \ldots = u^T \) and simply call these fluence-maps \( u \). Note that this thought is also consistent with the current practice of using identical fluence-maps in all sessions. This simplifies problem \((P_t)\) to

\[
(Q_t) \min \sum_{i=1}^{n} \nu x_i^{t} \exp(-(T - t + 1)(\hat{\alpha}_i(A_i u) + \hat{\beta}_i(A_i u)^2)) \quad (36)
\]

subject to

\[
\begin{align*}
(T - t + 1) \left[ (A_j^m u) + \rho_m (A_j^m u)^2 \right] & \leq \text{BED}_{\text{max}}^m - z_j^{t,m}, \quad j = 1, 2, \ldots, n_m, \ m \in M_1, \quad (37) \\
(T - t + 1) \left[ \sum_{j=1}^{n_m} (A_j^m u) + \rho_m \sum_{j=1}^{n_m} (A_j^m u)^2 \right] & \leq n_m \text{BED}_{\text{mean}}^m - n_m v^{t,m}, \ m \in M_2, \quad (38) \\
\sum_{j=1}^{n_m} g_j^m(u, u, \ldots, u; w^{t,m}) & \leq K_m, \ m \in M_3, \quad (39) \\
S u & \leq 0, \quad (40) \\
u & \geq 0. \quad (41)
\end{align*}
\]

Moreover, since \( A_j^m u \geq 0 \), constraints \((37)\) can be equivalently rewritten as linear constraints wherein the right hand side is obtained by solving a quadratic equation. This yields the equivalent problem

\[
(Q_t) \min \sum_{i=1}^{n} \nu x_i^{t} \exp(-(T - t + 1)(\hat{\alpha}_i(A_i u) + \hat{\beta}_i(A_i u)^2)) \quad (42)
\]

subject to

\[
\begin{align*}
A_j^m u & \leq -1 + \frac{\sqrt{1 + 4 \rho_m (\text{BED}_{\text{max}}^m - z_j^{t,m})/(T - t + 1)}}{2 \rho_m}, \quad j = 1, 2, \ldots, n_m, \ m \in M_1, \quad (43) \\
(T - t + 1) \left[ \sum_{j=1}^{n_m} (A_j^m u) + \rho_m \sum_{j=1}^{n_m} (A_j^m u)^2 \right] & \leq n_m \text{BED}_{\text{mean}}^m - n_m v^{t,m}, \ m \in M_2, \quad (44) \\
\sum_{j=1}^{n_m} g_j^m(u, u, \ldots, u; w^{t,m}) & \leq K_m, \ m \in M_3, \quad (45) \\
S u & \leq 0, \quad (46) \\
u & \geq 0. \quad (47)
\end{align*}
\]

The dose-volume constraints \((45)\) in problem \((Q_t)\) are still difficult to tackle. We therefore propose a constraint generation approach to surmount this difficulty.
3.1.1 Constraint generation method

In particular, we first solve \((Q_t)\) but without constraints \((45)\). This problem is given by

\[
(Q_t) \quad \min \sum_{i=1}^{n} \nu x_i^t \exp(-(T - t + 1)(\hat{\alpha}_i(A_t u) + \hat{\beta}_i(A_t u)^2))
\]

subject to

\[
A_j^m u \leq -1 + \frac{\sqrt{1 + 4\rho_m(B_{\text{max}}^m - z_j^m)/(T - t + 1)}}{2\rho_m}, \quad j = 1, 2, \ldots, n_m, \ m \in \mathcal{M}_1,
\]

\[
(T - t + 1)\sum_{j=1}^{n_m} (A_j^m u) + \rho_m \sum_{j=1}^{n_m} (A_j^m u)^2 \leq n_m B_{\text{mean}}^m - n_m v_{t,m}, \ m \in \mathcal{M}_2,
\]

\[
Su \leq 0,
\]

\[
u \geq 0.
\]

Suppose \(\hat{u}\) is an optimal solution to problem \((Q_t)\). Then, for each \(m \in \mathcal{M}_3\), we find \(n_m - K_m\) voxels that receive the smallest doses among the \(n_m\) voxels in \(O_m\) under fluence-map \(\hat{u}\). Let subset \(\mathcal{N}_m(\hat{u}) \subseteq \mathcal{N}_m\) denote this group of voxels. We then re-solve \((Q_t)\) but this time by replacing the dose-volume constraints with tolerance limits on all voxels in the set \(\mathcal{N}_m^\prime(\hat{u})\). Finally, we note that these tolerance limits can be equivalently re-written as linear constraints whose right hand side is obtained by solving a quadratic equation. That is, this problem is given by

\[
(\tilde{Q}_t) \quad \min \sum_{i=1}^{n} \nu x_i^t \exp(-(T - t + 1)(\hat{\alpha}_i(A_t u) + \hat{\beta}_i(A_t u)^2))
\]

subject to

\[
A_j^m u \leq -1 + \frac{\sqrt{1 + 4\rho_m(B_{\text{max}}^m - z_j^m)/(T - t + 1)}}{2\rho_m}, \quad j = 1, 2, \ldots, n_m, \ m \in \mathcal{M}_1,
\]

\[
(T - t + 1)\sum_{j=1}^{n_m} (A_j^m u) + \rho_m \sum_{j=1}^{n_m} (A_j^m u)^2 \leq n_m B_{\text{mean}}^m - n_m v_{t,m}, \ m \in \mathcal{M}_2,
\]

\[
A_j^m u \leq -1 + \frac{\sqrt{1 + 4\rho_m(B_{\text{mean}}^m - w_j^m)/(T - t + 1)}}{2\rho_m}, \quad j \in \mathcal{N}_m^\prime(\hat{u}), \ m \in \mathcal{M}_3,
\]

\[
u \geq 0.
\]

In the next section, we describe a log-barrier interior point algorithm for solving \((\tilde{Q}_t)\) and \((\overline{Q}_t)\).

3.1.2 Log-barrier interior point algorithm for solving \((\tilde{Q}_t)\) and \((\overline{Q}_t)\)

Since problems \((\tilde{Q}_t)\) and \((\overline{Q}_t)\) are identical in structure, we present the interior point algorithm only for \((\tilde{Q}_t)\). Before proceeding, we equivalently re-write \((\tilde{Q}_t)\) in a more compact and convenient form. In particular, we use numbers \(b_j^m\), for \(j = 1, 2, \ldots, n_m\) and \(m \in \mathcal{M}_1\) to denote the right hand sides in constraint \((49)\). Similarly, we use positive definite matrices \(Q^m\), matrices \(P^m\), and numbers \(r^m\) to compactly write the quadratic constraints \((50)\). That is, we have,

\[
(\tilde{Q}_t) \quad \min \sum_{i=1}^{n} \nu x_i^t \exp(-(T - t + 1)(\hat{\alpha}_i(A_t u) + \hat{\beta}_i(A_t u)^2))
\]
subject to
\[ A^m_j u - b^m_j \leq 0, \quad j = 1, 2, \ldots, n_m, \quad m \in M_1, \] (60)
\[ u'Q^m u + P^m u - r^m \leq 0, \quad m \in M_2, \] (61)
\[ Su \leq 0, \] (62)
\[ u \geq 0. \] (63)

**Lemma 3.3.** Problem \((\hat{Q}_t)\) as formulated in (59)-(63) is convex if \(\alpha^2 \geq 2\beta\) and \(\text{OER}_\alpha \leq \text{OER}_\beta\). This convexity property also holds for \((Q_t)\).

**Proof.** The constraints in \((\hat{Q}_t)\) are linear and convex quadratic. So it only remains to show that the objective function is convex if \(\alpha^2 \geq 2\beta\) and \(\text{OER}_\alpha \leq \text{OER}_\beta\). That is, we need to show that
\[ \sum_{i=1}^n \nu x_i \exp(-(T-t+1)(\hat{\alpha}_i(A_iu) + \hat{\beta}_i(A_iu)^2)) \] is convex. Since composition of a convex function with an affine mapping is convex (see Section in [5]), and since \(A_i u \geq 0\) for all \(u \geq 0\), we have that \(\exp(-(T-t+1)(\hat{\alpha}_i(A_iu) + \hat{\beta}_i(A_iu)^2))\) is convex over \(u \geq 0\) if \(\alpha_i^2 \geq 2\beta_i\). From formulas (34) and (35), we see that \(\hat{\alpha}_i^2 \geq 2\beta_i\). Therefore, because \(a \geq 0\) and \(b \geq 0\), \(\frac{dh}{dz}\) is non-negative (and hence \(h(\cdot)\) is convex) over \(z \geq 0\) when \(a^2 \geq 2b\). Since composition of a convex function with an affine mapping is convex (see Section in [5]), and since \(A_i u \geq 0\) for all \(u \geq 0\), we have that \(\exp(-(T-t+1)(\hat{\alpha}_i(A_iu) + \hat{\beta}_i(A_iu)^2))\) is convex over \(u \geq 0\) for each \(i\) if \(\hat{\alpha}_i^2 \geq 2\hat{\beta}_i\). From formulas (34) and (35), we see that \(\hat{\alpha}_i^2 \geq 2\beta_i\) if
\[ \frac{\alpha^2(\hat{y}_i\text{OER}_\alpha + \kappa)^2}{\text{OER}_\alpha^2} \geq \frac{2\beta(\hat{y}_i\text{OER}_\beta + \kappa)^2}{\text{OER}_\beta^2}. \] (66)

After algebraic simplification, this yields that \(\hat{\alpha}_i^2 \geq 2\beta_i\) if \(\alpha^2 \geq 2\beta\) and \(\text{OER}_\alpha \leq \text{OER}_\beta\). \(\square\)

Fortunately, this sufficient condition for convexity is met for the values of \(\alpha, \beta, \text{OER}_\alpha, \text{OER}_\beta\) currently available in the clinical literature for head-and-neck cancer. Consequently, as discussed in [5], the log-barrier interior point algorithm with Newton’s method with backtracking line search is particularly suitable for solving \((\hat{Q}_t)\) and \((Q_t)\). Different components of this algorithm are described in various sections of [5] and we refer the interested reader to that book for a rigorous convergence analysis of such algorithms. We present this algorithm here using our notation for completeness.

For any log-barrier parameter \(\mu > 0\), the log-barrier function for problem (59)-(63) is given by
\[
\phi(u; \mu) \triangleq \sum_{i=1}^n \nu x_i \exp(-(T-t+1)(\hat{\alpha}_i(A_iu) + \hat{\beta}_i(A_iu)^2)) - (1/\mu) \left( \sum_{m \in M_1} \sum_{j=1}^{n_m} \ln( b^m_j - A^m_j u) + \sum_{m \in M_2} \ln(r^m - P^m u - u'Q^m u) + \sum_{k=1}^K \ln u_k + \sum_{l=1}^L \ln(-S_l u) \right).
\] (67)
The idea (see Algorithm 11.1 on page 569 of [5]) then is to use Newton’s procedure (Algorithm 9.5 on page 487 of [5]) to solve a sequence of unconstrained problems \( \min \phi(u; \mu_\eta) \) for a strictly increasing and divergent sequence of parameters \( \{\mu_\eta\}_{\eta=1}^\infty \). We use \( \nabla \phi(u; \mu) \) and \( \nabla^2 \phi(u; \mu) \) to denote the gradient and Hessian of \( \phi(u; \mu) \) with respect to the decision variables \( u \in \mathbb{R}^K \). The step-size \( \delta \) in Newton’s method is chosen by performing a backtracking line search (Algorithm 9.2 on page 464 of [5]) with parameters \( 0 < p < 0.5 \) and \( 0 < q < 1 \). Finally, note that the total number of constraints in \((\hat{Q}_t)\) is \( K + L + |M_2| + \sum_{m \in M_1} n_m \) and we use \( C \) to denote this integer in our stopping condition below.

**Log-barrier interior point algorithm**

**INITIALIZE:** Set \( \eta = 1 \) and begin with an initial guess fluence-map \( u \) in the interior of the feasible region of \((\hat{Q}_t)\). Set tolerance to \( \epsilon \) and pre-select a strictly increasing sequence of log-barrier parameters \( \mu_\eta \to \infty \).

**DO WHILE** \( C/\mu_\eta \geq \epsilon \)

\( \lambda^2 \triangleq \infty \).

**DO WHILE** \( \lambda^2 / 2 > \epsilon \)

\( \Delta u \triangleq -\nabla^2 \phi(u; \mu_\eta) \nabla \phi(u; \mu_\eta) \), and \( \lambda^2 \triangleq -(\nabla \phi(u; \mu_\eta))^\prime \Delta u \).

\( \delta = 1 \).

**DO WHILE** \( \phi(u + \delta \Delta u; \mu_\eta) > \phi(u) + p\delta \nabla \phi(u; \mu_\eta) \Delta u \)

\( \delta = q\delta \).

**END DO**

\( u \triangleq u + \delta \Delta u \).

**END DO**

\( \eta \leftarrow \eta + 1 \).

**END DO**

**TERMINATE:** Report fluence-map \( u \) as an optimal solution to \((\hat{Q}_t)\).

### 3.2 A streamlined version of CEC

Based on the above discussion, we propose the following simplified version of CEC for approximate solution of our stochastic control problem.

**Streamlined CEC**

**INITIALIZE:** Set \( t = 1 \) and begin with a given initial state \([x^1; y^1; z^1; v^1; w^1]\).

**DO WHILE** \( t \leq T \),

1. let the state at the beginning of session \( t \) be \([x^t; y^t; z^t; v^t; w^t]\);

2. fix the oxygen partial pressure vectors \( y^t, y^{t+1}, \ldots, y^T \) all at \( y^t \) and call this \( \tilde{y} \). Substitute \( \tilde{y} \) in Equations (34) and (35) to obtain nominal radiosensitivity parameters \( \tilde{\alpha}_i \) and \( \tilde{\beta}_i \) for tumor voxels \( i = 1, 2, \ldots, n \);

3. solve problem \((\hat{Q}_t)\) using the log-barrier interior point algorithm that uses Newton’s method along with a line search procedure (see Section 3.1.2 above) to obtain an optimal fluence-map \( \tilde{u} \) and, for each \( m \in M_3 \), let \( N^*_{m}(\tilde{u}) \subseteq N_m \) denote the subset of \( n_m - K_m \) voxels that receive the smallest doses among the \( n_m \) voxels in \( O_m \) under fluence-map \( \tilde{u} \);

4. solve problem \((\bar{Q}_t)\) using the log-barrier interior point algorithm that uses Newton’s method along with a line search procedure (see Section 3.1.2 above) to obtain an optimal fluence-map \( u_\ast \) and use it in session \( t \);
5. calculate $\alpha^t_i$ and $\beta^t_i$ via expressions (1)-(2) using the imaged hypoxia values $y^t_i$, for $i = 1, 2, \ldots, n$; update tumor cell density and normal tissue states using $u^*$ as

$$x^{t+1}_i = x^t_i \exp(-\alpha^t(A^t_i u^*) - \beta^t(A^t_i u^*)^2), \ i = 1, 2, \ldots, n,$$

$$z^{t+1,m}_j = z^{t,m}_j + (A^m_j u^*) + \rho^m(A^m_j u^*)^2, \ j = 1, 2, \ldots, n, m \in M_1,$$

$$v^{t+1,m}_j = v^{t,m}_j + \frac{\sum_{j=1}^{n_m} (A^m_j u^*) + \rho^m \sum_{j=1}^{n_m} (A^m_j u^*)^2}{n_m}, \ j = 1, 2, \ldots, n, m \in M_2,$$

$$w^{t+1,m}_j = w^{t,m}_j + (A^m_j u^*) + \rho^m(A^m_j u^*)^2, \ j = 1, 2, \ldots, n, m \in M_3;$$

6. sample $\theta^t$ from its probability distribution and set $\ln y^{t+1} = \ln y^t + \theta^t$;

7. update $t \leftarrow t + 1.$

END DO

This CEC method assumes that a hypoxia image is acquired at the beginning of every treatment session. Another alternative is to not use hypoxia images but instead acquire cell density images at the beginning of every session. Note that, after a few treatment sessions, hypoxia information is at least indirectly included in the cell density image because hypoxic regions are radioresistant, and hence, ceteris paribus, are likely to have a higher cell density. One change that is needed to the above CEC algorithm in this case is in Step 2 of the DO WHILE loop over $t \leq T$ — since a hypoxia image is not acquired, we do not know the oxygen partial pressure values $y^t$ and hence we cannot employ Equations (1) and (2) to calculate the values of $\hat{\alpha}_i$ and $\hat{\beta}_i$; we thus instead simply use the constant values $\alpha$ and $\beta$ under well-oxygenated conditions. Another change is in Step 5 — since the values of $\alpha^t_i$ and $\beta^t_i$ cannot be calculated, we cannot use the update (68) to calculate the next cell density vector $x^{t+1}$. Instead, this cell density is directly observed in an image acquired at the beginning of the $t + 1$st session. Consequently, we call this alternative approach “CEC with cell density images” below.

In the next section, we apply the CEC methodology developed here to four head-and-neck cancer test cases.

4 Results

We first describe the four test cases that were used in our simulation experiments. All test cases were generated using our in-house software Phantom Creator (PhanC) written in MATLAB [49]. These test cases were three-dimensional and were carefully developed to be representative of clinical scenarios in terms of geometry. Specifically, our test cases were similar to those we used in [47], which were in turn similar in structure to those in [17, 35, 45, 46, 50]. All cases used seven equally spaced coplanar beams and the beamlet resolution was $5 \times 5$ mm$^2$. All voxels were $5 \times 5 \times 5$ mm$^3$ leading to $\nu = 125$ mm$^3$. We acknowledge that this voxel-size is somewhat larger than what is used in the clinic today. This size helped us efficiently run hundreds of computer simulations to derive qualitative insights that we believe should hold independently of the voxel-size. Another practical issue in any future clinical implementation is that the voxel-resolution used in IMRT planning is likely to be finer than what is detectable with reasonable accuracy in a functional image. Since our tumor states correspond to voxel-by-voxel information acquired from functional images, we preferred to employ a coarser voxel resolution so as to avoid using a tedious notation that would account for two different resolutions in our problem formulation.
All cases included spinal cord, brainstem, left and right parotids and unspecified normal tissue between these critical organs. The total number of voxels in the head-and-neck target and in the normal tissues, and the total number of beamlets is shown in Table 1 below.

<table>
<thead>
<tr>
<th>case #</th>
<th># of beamlets (k)</th>
<th># of tumor voxels (n)</th>
<th># of normal tissues voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1680</td>
<td>9096</td>
<td>17040</td>
</tr>
<tr>
<td>2</td>
<td>1485</td>
<td>6384</td>
<td>15360</td>
</tr>
<tr>
<td>3</td>
<td>1500</td>
<td>7248</td>
<td>16536</td>
</tr>
<tr>
<td>4</td>
<td>1584</td>
<td>7350</td>
<td>13850</td>
</tr>
</tbody>
</table>

Table 1: Description of the geometry used in head-and-neck cancer cases.

The conventional fractionation schedule was assumed to include $N_{\text{conv}} = 35$ fractions. We included maximum dose constrains for spinal cord, brainstem and unspecified normal tissue. A dose-volume constraint for unspecified normal tissue was also added. Mean dose constrains were used for left and right parotids. The tolerance dose values for various normal tissues were similar to [3, 23, 33, 34] and are listed in Table 2 below.

<table>
<thead>
<tr>
<th>Normal tissue</th>
<th>$D_{\text{max}}$ (Gy)</th>
<th>$D_{\text{mean}}$ (Gy)</th>
<th>$D_{dv}$ (Gy), $\phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinal cord</td>
<td>45</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Brainstem</td>
<td>50</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Left and right parotids</td>
<td>N/A</td>
<td>28</td>
<td>N/A</td>
</tr>
<tr>
<td>Unspecified normal tissue</td>
<td>77</td>
<td>N/A</td>
<td>70, 0.05</td>
</tr>
</tbody>
</table>

Table 2: Tolerance doses for various normal tissues in our head-and-neck test cases where the dose is administered in $N_{\text{conv}} = 35$ equal-dose fractions. Recall that for dose-volume type constraints no more than a volume fraction $\phi$ of the normal tissue can receive dose more than $D_{dv}$.

Dose-response parameter estimates were chosen from the existing clinical literature [12, 13, 14, 15, 21, 42, 64, 68]. Head-and-neck tumor was assumed to exhibit $\alpha = 0.35 \text{ Gy}^{-1}$ and $\beta = 0.035 \text{ Gy}^{-2}$ under well-oxygenated conditions. The $\alpha/\beta$ ratios for all normal tissues were fixed at 3 Gy. Parameters OER$_{\alpha}$, OER$_{\beta}$, and $\kappa$ were fixed at 2.5, 3, and 3.28 mm Hg, respectively, as in [51, 52, 65]. Note that the two sufficient conditions for convexity in Lemma 3.3 hold for this choice of parameters. All computer simulations were performed on a 3.1 GHz iMac desktop with 16 GB RAM.

The stochastic vectors $\theta_t$ were assumed to be independent across treatment sessions and to have a multivariate normal distribution with zero mean. Intuitively, tumor voxels that are close to each other should have similar oxygen partial pressures. This was captured by using a distance-based covariance matrix $\Sigma$ for each $\theta_t$. Suppose $D_{ij}$ denotes the distance between tumor voxels $i$ and $j$. We used two standard covariance functions from spatiotemporal statistics [8] to define matrix $\Sigma$: the exponential function and the rational-quadratic function. That is, the entries $\sigma_{ij}$ of matrix $\Sigma$ were obtained by

$$\sigma_{ij} = \exp\left(-\frac{D_{ij}}{\sigma}\right)$$

(72)

for the exponential case, and by

$$\sigma_{ij} = \left(1 + \frac{D_{ij}^2}{2\sigma}\right)^{-\sigma}$$

(73)

for the rational-quadratic case, where $\sigma > 0$ is a parameter of these functions. Note that other researchers have also used different distance-based (deterministic) models of hypoxia evolution.
[51, 59], and our thought process here is similar to theirs. Finally, the initial cell density was fixed at the normalized value of 1 in all tumor voxels since our results below do not depend on this value as long as it is invariant across voxels. This assumption of homogeneous initial cell density is also made in [31], and in fact can be easily removed without any change to our solution procedure here (see [16, 24, 25]). As such, the assumption is made purely for notational simplicity and so that we can study the effect of spatiotemporal variations in hypoxia without the confounding effect of initial cell density within our theoretical framework.

We first compare the TNTCR values attained by three different methods for all four test cases. The first method, which we see as the base case or the conventional approach, is static; that is, it does not acquire any images, and hence does not adapt to hypoxia either directly or indirectly. In particular, we simply solved problem \( Q_1 \) once at the beginning of the treatment session with tumor radiosensitivity values fixed at \( \alpha = 0.35 \text{ Gy}^{-1} \) and \( \beta = 0.035 \text{ Gy}^{-2} \), and used the resulting fluence-map vector in all sessions. The second method employed CEC with hypoxia images and the third method employed CEC with cell density images as described above. For each test case, we performed two sets of 30 simulations each. For one set, we used the exponential covariance function and for the second set, we used the rational-quadratic function. The \( \sigma \) value was fixed at 0.1 for all simulations (we also tried a few other values of \( \sigma \), but our qualitative conclusions were identical in all cases, so we report results for only one value here). We compared the average TNTCR over 30 simulations for these three methods. The percentage improvements of the two CEC methods over the static method were computed as follows:

\[
\text{Percentage improvement} = \frac{\sum_{\text{sim}=1}^{30} \text{TNTCR}^\text{static}_{\text{sim}} - \sum_{\text{sim}=1}^{30} \text{TNTCR}^\text{CEC}_{\text{sim}}}{\sum_{i=1}^{30} \text{TNTCR}^\text{static}_{\text{sim}}} \times 100. \tag{74}
\]

<table>
<thead>
<tr>
<th>case #</th>
<th>hypoxia images</th>
<th>cell density images</th>
<th>hypoxia images</th>
<th>cell density images</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14%</td>
<td>13%</td>
<td>24%</td>
<td>22%</td>
</tr>
<tr>
<td>2</td>
<td>63%</td>
<td>62%</td>
<td>66%</td>
<td>63%</td>
</tr>
<tr>
<td>3</td>
<td>60%</td>
<td>59%</td>
<td>59%</td>
<td>58%</td>
</tr>
<tr>
<td>4</td>
<td>13%</td>
<td>11%</td>
<td>26%</td>
<td>24%</td>
</tr>
</tbody>
</table>

Table 3: Improvement in average TNTCR obtained by CEC over the static approach for exponential and rational-quadratic covariance functions.

The improvement numbers are reported in Table 3. Figures 1(a) and (b) show the improvement in TNTCR for each simulation for all cases, for the exponential and rational-quadratic covariance functions, respectively, when only hypoxia images are used. Similarly, Figures 2(a) and (b) show the improvement in TNTCR for all cases, for the exponential and rational-quadratic covariance functions, respectively, when only cell density images are used.

Table 3, and Figures 1 and 2 show that adapting the treatment plan to the spatiotemporal evolution of hypoxia, either directly via hypoxia images or indirectly via cell density images, led to an improvement in average TNTCR for both covariance functions; in fact, adaptive treatment led to a better TNTCR in every simulation.
The improvements were essentially identical no matter whether hypoxia images were used or cell density images were used. This seems to corroborate our original belief that sufficient information about the spatiotemporal evolution of hypoxia might be included in cell density images. A potential clinical implication of this could be that the treatment planner might be able rely on other logistical and technological considerations in choosing which images to use in future practice.

There was a marked difference in the improvement achieved by CEC, with both covariance functions, for cases 1 and 4 as compared to cases 2 and 3. Specifically, the improvement for cases 1 and 4 was smaller than that for cases 2 and 3. We believe that this can be explained based on the difference in the geometry of these cases. The spinal cord and brainstem were much closer to
the tumor in cases 1 and 4 than in cases 2 and 3. This implies that cases 2 and 3 inherently had more leeway for improvement via better treatment planning; as such, cases 2 and 3 were “easier”.

A careful comparison of Figure 1(a) with Figure 1(b) shows that the variance in improvements across different simulations was higher with the rational-quadratic covariance function than the exponential one. Similarly for Figures 2(a) and (b). We believe that this is because the rational-quadratic covariance decreases with increasing distance at a slower rate than the exponential function. This means that the oxygen partial pressure in distant voxels tends to be more correlated in the rational-quadratic case thus leading to a higher variance in improvements.

We wanted to study the root-cause of the improvements achieved by CEC over the static method. Intuitively, it should be that CEC is better able to spatiotemporally redistribute dose based on observed images. This is verified by Figures 3, 4, 5, and 6. Specifically, Figures 3 and 4 show that the average dose (over all tumor voxels) delivered by CEC in various treatment sessions was different and in fact was higher in almost every simulation than that delivered by the static method. The CEC dose showed an upward trend over treatment sessions. This temporal dose boost can be explained, for example in the case of cell density images, as follows. Both the CEC and static methods start out with an identical dose distribution in the first treatment session. As treatment progresses, the CEC method, which gets the opportunity to re-plan, discovers that there is room for higher doses that will reduce the TNTCR while still protecting the normal tissues; it thus administers these higher doses in later sessions. Figures 5 and 6 further elaborate this point by contrasting the spatial distribution of the doses delivered by CEC in each simulation against that delivered by the static method, for cases 3 and 2 with the rational-quadratic and exponential covariance functions, respectively.

![Graphs showing average dose per session over treatment sessions](image)

Figure 3: The average (over all tumor voxels) dose (Gy) per session delivered by CEC with hypoxia images in each one of the 30 simulations over 35 treatment sessions in case 1: (a) exponential covariance function (b) rational-quadratic covariance function. The solid flat line shows the constant dose (Gy) delivered by the static method in every session.

Finally, we employed a random walk model to simulate the spatiotemporal evolution of oxygen partial pressure in this paper. In lieu of actual hypoxia images, it is therefore important to validate
Figure 4: The average (over all tumor voxels) dose (Gy) per session delivered by CEC with cell density images in each one of the 30 simulations over 35 treatment sessions in case 1: (a) exponential covariance function (b) rational-quadratic covariance function. The solid flat line shows the constant dose (Gy) delivered by the static method in every session.

that this simulation method generates a reasonable distribution of hypoxia. Toward this end, as an example, we show histograms of the hypoxia distribution generated over different treatment sessions in one simulation for case 2 with the exponential covariance function in Figure 7. The figure shows that the shape of the hypoxia distribution does indeed look lognormal, and in fact, as the treatment progresses, matches with the shapes shown in [1, 28, 59, 60] based on clinical literature as well as simulation studies. Histograms of the corresponding $\alpha_t^i$ and $\beta_t^i$ values are also shown in Figures 8 and 9. The spikes in the histograms near 100mm Hg are somewhat artificial because we truncated the partial pressure values at 100mm Hg. We believe that this truncation has essentially no effect on our results since the radiosensitivity parameters $\alpha_t^i$ and $\beta_t^i$ asymptotically and monotonically approach their well-oxygenated values of 0.35 Gy$^{-1}$ and 0.035 Gy$^{-2}$, respectively, as the oxygen partial pressure increases. In fact, these parameters are already close to their well-oxygenated values at 100mm Hg. The specific value of 100 was chosen from the histograms shown in [31, 60].

5 Conclusions and discussion

We applied the general stochastic control formalism for dynamic radiotherapy planning from [25] to hypoxia. We developed in detail a concrete implementation of an approximate control method called CEC by specifically tailoring it to our problem. This method involved the solution of a sequence of convex optimization problems; dose-volume constraints were tackled via a simple constraint generation procedure. We implemented two versions of this CEC method. One used (simulated) hypoxia images and hence adapted to the spatiotemporal evolution of hypoxia directly; the other used (simulated) cell density images and thus adapted indirectly. The evolution of oxygen partial pressure was simulated using a first order vector autoregressive process with a distance-based covariance matrix from the literature in statistics. This simulation approach led to partial pressures that followed a lognormal distribution as observed in the clinical literature. The effect of hypoxia on tumor-radiosensitivity was modeled using the well-known OER approach from radiobiology.

We performed computer simulations to quantify any potential benefits of dynamically redesign-
Figure 5: The dose (Gy) delivered by CEC with hypoxia images and by the static method in various treatment sessions for case 3 with the rational-quadratic covariance function.
Figure 6: The dose (Gy) delivered by CEC with cell density images and by the static method in various treatment sessions for case 2 with the exponential covariance function.
ing the fluence-maps based on hypoxia or cell density information acquired over the treatment course. Our simulations suggest that such dynamic planning could offer an improvement over static planning. This conclusion is similar in essence to those drawn in [31, 53] using dynamic planning to treat a canine patient and also to that in [51, 59] using more stylized (spherical) models of hypothetical tumors.

Significant challenges would need to be overcome, however, before any future clinical implementation of our theoretical approach. Potential benefits of adaptive planning are likely to depend crucially on the accuracy of the functional images used. There is hope that functional imaging technology will continue to improve over the next few decades; this could impact the value of adaptive planning. Ultimately, clinical trials will need to be run to assess the benefits of adapting to hypoxia. We hope that computer simulation studies, such as the one we presented here, would help guide the design of such trials.

6 Acknowledgments

Research funded in part by the National Science Foundation through grant #CMMI 1054026.
Figure 7: A histogram of oxygen partial pressure in different tumor voxels for various treatment sessions for case 2 with the exponential covariance function.
Figure 8: A histogram of tumor alpha values (Gy$^{-1}$) over different tumor voxels for various treatment sessions for case 2 with the exponential covariance function.
Figure 9: A histogram of tumor beta values (Gy$^{-2}$) over different tumor voxels for various treatment sessions for case 2 with the exponential covariance function.
References


