Genetic Instability in Cancer: An Optimal Control Problem

By Frederic Y. M. Wan, Alexander V. Sadovsky, and Natalia L. Komarova

Carcinogenesis (cancer generation) relies on a sequence of mutations that transforms cells, leading to their unchecked growth. An important phenomenon in cancer is genetic instability, or an increased rate of mutations, with a higher associated cell death rate. Given the trade-off between an increase in mutant cells by higher mutation rate and higher loss of mutants through a higher death rate, we ask the question: what mutation rate is most advantageous for cancer? To seek an answer to this question, we investigate an optimal control problem of normal and mutant cell growth where the abnormal mutation rate plays the role of a time-dependent control. The analysis of this problem shows that the best "strategy" for the fastest time to cancer is to start with a high level of genetic instability initially, and then to switch to a low level of genetic instability. The exact shape of the optimal mutation rate as a function of time depends on how genetic instability contributes to the death rate of cells. If the death rate is a linear or an increasing concave function of the mutation rate, the optimal mutation rate is bang-bang, which changes from its highest to its lowest value with a finite jump discontinuity. However, if the death rate is an increasing convex function of the mutation rate, then the optimal control is a continuous decreasing function of time. Two known mechanisms of cancer initiation have been considered, an activation of an oncogene and an inactivation of a tumor-suppressor gene. Mathematical justification of the results of the first, a one-step process, is reported herein.

Address for correspondence: Frederic Y. M. Wan, Department of Mathematics, Center for Mathematical and Computational Biology, Center for Complex Biological Systems, University of California, Irvine, Irvine, CA 92697-3875; e-mail: fwan@uci.edu

DOI: 10.1111/j.1467-9590.2009.00476.x

STUDIES IN APPLIED MATHEMATICS 125:1–38

1. Introduction

Homeostasis is a process for maintaining relative constancy in a biological system. In particular, homeostatic control of cells maintains a dynamic equilibrium (or acceptable variations) of cell populations in (organs of) multicellular organisms. Sometimes cells can break out of homeostatic control and enter a phase of abnormal expansion. Cancer is a manifestation of such uninhibited growth, which results when cell lineages lose the ability to maintain a sufficient rate of *apoptosis* [1], or programmed cell death, to counterbalance the rate of *mitosis* [1], or cell division. Mechanisms of homeostatic control can be disrupted by *genetic mutations* [1], to which every cell is prone to some extent. A highly elevated susceptibility to mutations is known in oncology as *genetic instability* [2–5, 37], and, being closely related to carcinogenesis, is of much interest to cancer researchers [6–11].

Genetic instability is known to involve the following two (and possibly other) competing effects on carcinogenesis: an increased frequency of cell deaths resulting from deleterious mutations, and an increased frequency of mutations producing cancerous cells. The former effect impedes cancer growth; the latter, accelerates it. This gives rise to the question, *What "amounts" of instability are most favorable to the onset of cancer?* This problem was first investigated by mathematical modeling in [12, 13]. The models were based on two conceptual premises, (a) the Darwinian view of a cell colony as a collection of phenotypes struggling for survival and predominance, and (b) the (hypothetical) ability to set the degree of genetic instability in the colony to any level that is constant in time. A colony of cells was regarded as a species undergoing a birth-and-death process, interpreted as a Darwinian microevolution, and subject to genetic instability. The degree of genetic instability was defined quantitatively as the probability of a genetic mutation of a specified type per cell, per mitosis. The "microevolutionary success" of the species was defined as the ability to produce a cancerous population of a given size M, and was measured as the inverse of the time required to reach the target population size. The above question was thus recast as a problem in optimization. The results in [12, 13] showed that "too much" instability impedes the growth of the colony by increasing the cell death rate (a result of too many deleterious mutations in cells), while "too little" instability impedes the growth by lowering the rate of acquisition of cancerous mutations. In the same paper, an optimal, time invariant level of genetic instability that maximizes the rate of progression was found. This level, defined as the probability of chromosomal loss per cell division, agreed well with available in vitro experimental measurements and turned out to be a robust mathematical result: it depended only logarithmically on the relevant parameter values, and its order of magnitude was consistent with the data [14].

The restriction in [13] that the level of genetic instability remain constant in time, however, served only to facilitate the first steps in the mathematical

modeling of the problem. There is growing biological evidence that the degree of genetic instability is high at relatively early stages of carcinogenesis, and decreases as cancer progresses [15, 16]. In [17], we removed the restriction of a fixed level of genetic instability and allowed it to be a function of time, $p(t)$, within a biologically admissible range. Following the general modeling framework of [13], we distinguished among several types of cells: normal cells, cancerous cells with homeostatic control impaired completely, and in some cases also cells containing intermediate mutations but still obeying homeostatic control. The entire cell colony would undergo birth-and-death process with mutations subject to certain magnitude constraints. Initially, homeostatic control would regulate the growth thus as to keep the colony at a constant size, near a "selection barrier" [18]. In this state of stalled growth, the cellular population would remain near its environment's "carrying capacity," defined by the available space, nutrients, and the cells' ability to divide and die. The colony, regarded as a Darwinian species "trying" to escape the regulation, may escape by a sequence of mutations until a *phenotype* [1] is produced with the homeostatic control mechanism lost. The resulting mutant cells would then spread in an unchecked manner. The evolution of these different cell populations in the colony was taken to depend on the mutation rate $p(t)$ due to genetic instability which acts as a control for the population growth. Of interest was an *admissible control p(t)* (or its normalized form $u(t)$ as defined in (6)) which steers the cancerous cells to a target population M in the shortest possible time.

This shortest time problem was discussed in [17] for two different mechanisms of homeostatic control loss. The control-theoretic analysis leading to the results reported there for the first mechanism, *activation of an oncogene*, is described herein. The dynamics of the one-step process in the context of genetic instability is more fully described and mathematically modeled in the next section. We note here only that our usage of optimal control theory for the relevant shortest time problems differs from that in many areas of biosciences [19, 20]. In most biomedical applications, optimal control theory is used to solve *design* problems, e.g., finding treatment strategies [21–24]. Here, we use it for the purpose of *analysis*, namely, to understand the experimentally observed behavior of developing cancer. This approach is similar in spirit to that taken by Iwasa and Levin [25], who analyzed the optimal timing of life-strategies of breeding and migrating organisms. In this ecological context, our analysis can be said to address the question whether the observed behavior of cancerous cell populations is essentially a consequence of an optimization process. The biological implications of the optimal control analysis of the cancer growth problem have already been reported in [17]. The present paper describes the theoretical development that enabled us to make the observations there, the principal one being the preference for a high level of genetic instability at the early stages of carcinogenesis to generate a large pool of cancerous

cells and for a relatively low level of mutation to allow the cancerous cells to multiply by their natural growth rate at the late stages.

The rest of this paper is organized as follows. In Section 2, we formulate mathematically a one-step genetic instability problem, where homeostatic control is lost by means of just one molecular event. In Section 3, we present the Hamiltonian formalism for solving the resulting optimal control problem. In Sections 4–6, we focus on a representative class of death rates given by (5) with a maximum normalized death rate $d_m = 1$. (The effects of a smaller value of d_m in the range of $(0, 1)$ have already been reported in [17].) The case of a general nonnegative, monotone increasing death rate $d(u)$ not investigated previously is examined in Section 7. Some results for a two-step genetic instability problem were presented in [17]. The mathematical analysis leading to these results are somewhat different from that for the one-step genetic instability mechanism and will be presented in a separate report.

2. Genetic instability by activation of an oncogene

2.1. A one-step system

Consider a cell colony undergoing a birth and death process with a time-varying mutation rate within a biologically admissible range. Cells go through a sequence of mutations until an advantageous, i.e., cancer-favorable, phenotype is achieved. Mutation rate can have two effects. On the one hand, an increased mutation rate can lead to a faster production of the advantageous mutants, thus accelerating the growth of the colony. On the other hand, a high mutation rate reduces the fitness of the mutated cells and thus leads to a higher death toll in the population. To delineate the net consequence of these two competing effects, we consider in this paper a colony of cells initially evolving at a constant population size near a selection barrier. This barrier can be overcome by the offspring of a mutant whose properties are different from those of normal cells. Here we consider the mutant cells to have an activated oncogene and show an increased division rate. We assume that such transformed cells are created by means of one molecular event, genetic or epigenetic. Denote by X_1 the population of cells that have not undergone cancerous mutations, and by X_2 the mutated type. The probability for a cell to acquire an inactivating mutation of a particular gene upon a cell division is denoted by $\bar{\mu}$; this quantity is called the "basic mutation rate". The probability p is an additional transformation rate resulting from genetic instability. This quantity measures the degree of genetic instability in cells. It is low in stable cells (cells without *chromosomal instability* or CIN), but can be highly elevated in chromosomally unstable cells. Effectively, if $p \ll \bar{\mu}$, then there is no genetic instability; $p \gg \bar{\mu}$ means a genetically unstable cell population. Both probabilities $\bar{\mu}$ and p are

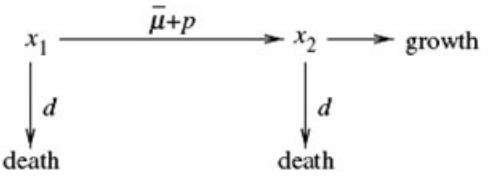

Figure 1. Growth and mutations diagram for the one-step process.

measured per gene per cell division. The new feature in our investigation first described in [17] is that p be allowed to vary with time. With these notations in mind, we can summarize the processes of growth and mutations above by means of a mutation diagram, Figure 1. The diagram is of the same type used in [26] and [27]. The goal is to find a strategy $p(t)$ which maximizes the growth of cancer; as such the mutation rate $p(t)$ due to genetic instability is the control of our optimal control problem.

The dynamics of the cell populations is modeled as follows. Cells reproduce and die, and the rate of renewal is normalized to be 1 for the type X_1. In the absence of dangerous mutants, X_2, the total number of cells, X_1, is taken to obey the well-known logistic growth law. The mutants X_2 expand at the rate $a > 1$. With $(.)' = d(.)/dt$, mutation diagram in Figure 1 translates into the following two ordinary differential equations (ODE) describing the rate of change of the two cell populations [36],

$$X_1' = (1 - \bar{\mu} - p - d(p)))X_1 - \phi X_1, \tag{1}$$

$$X_2' = (p + \bar{\mu})X_1 + a(1 - d(p))X_2 - \phi X_2, \tag{2}$$

where

$$\phi = (1 - d(p))X_1/N, \tag{3}$$

and

$$X_1(0) = N, \qquad X_2(0) = 0 \tag{4}$$

for a prescribed dependence of the death rate, $d(p)$, on the mutation rate p. The term ϕ provides the growth limiting factor leading to logistic growth in the absence of mutants, and accounts for the homeostatic control present in a system of X_1 cells. We note that choices of ϕ other than (3) are also possible (see Ref. [17]). With the fraction of normal cell population starting to mutate into cancerous mutants, X_2 cells break out of regulation and enter a phase of abnormal growth, which, depending on the growth rate, may or may not lead to carcinogenesis

2.2. Dependence of death rate on mutation rate

In our models, the death rate, d, is taken to be a function of the mutation rate p due to genetic instability. If p is small, then chromosome losses do not happen. If p is large, a cell often loses chromosomes, which results in an increased death rate. Therefore, in general, the function $d(p)$ is an increasing function of p. An example of this dependence of the death rate on p is

$$d(p) = \frac{d_m}{u_m^\alpha}\left\{u_m^\alpha - (u_m - (p - p_{\min}))^\alpha\right\} = d_m\left[1 - (1 - u)^\alpha\right], \quad (\alpha > 0), \tag{5}$$

where

$$u_m = p_{\max} - p_{\min}, \quad u = \frac{1}{u_m}(p - p_{\min}) \tag{6}$$

with $p_{\min} \le p \le p_{\max}$. The quantities $p_{\min}$ and $p_{\max}$ define a biologically relevant range of the mutation rate, p, and the nonnegative u is a normalized *gross chromosomal change (mutation) rate* which satisfies the inequality constraint $0 \le u \le 1$. The constant d_m defines the magnitude of the death rate, and is taken to be in the interval [0, 1]. The motivation for this particular form of dependency was given in [17]. We allow α, the exponent in Equation (5), to be a real positive number. In particular, we investigate the influence of the convexity of this function on the optimal solution (for the three cases $\alpha < 1$, $\alpha = 1$ and $\alpha > 1$). The special case $\alpha = 1$ yields a control problem which is linear in the control $u(t)$. This special case has been studied extensively in the optimal control literature (see [28] for example).

Every biologically acceptable normalized genetic instability rate $u(t)$, called an *admissible control*, determines a growth process for the cell colony. We seek the choice of an admissible control $u(t)$ in the range $0 \le u \le 1$ that allows the cancerous population to reach a given size, M, in the shortest possible time. The specified terminal population size M is called the *target*. An admissible control $\bar{u}(t)$ that steers the cancerous population to the target faster than any other control is an *optimal control*.

In the simplest case considered in [13], the class of admissible controls, $u(t)$, was restricted to constant functions. Then, the result of the optimization problem is a single normalized mutation rate value, $\bar{u}$ (or $\bar{p}$ in unnormalized form), which will depend on the parameters of the system. However, better growth times can be achieved if we allow p (and hence u) to be a function of time. It seems intuitive, and is evident from experimental results, that higher initial and lower subsequent values of $u(t)$ will better facilitate the growth. We will show in this paper that this is in fact the case and how the variation of $u(t)$ with time depends on the convexity of the death rate function $d(u)$.

2.3. The shortest time problem

To reduce the number of parameters in our problem, we let the population size of normal (noncancerous) cells at time t be normalized by its initial population size N and denote it by $x_1(t) = X_1(t)/N$. We let $x_2(t) = X_2(t)/M$ be the population size of dangerous mutant cells normalized by its final target population size M. In terms of $\{x_k(t)\}$, the two ODE (1)–(2) governing the time evolution of the two cell populations become:

$$x_1' = -(\mu + u_m u)x_1 + [1 - d(u)](1 - x_1)x_1 \equiv g_1(x_1, x_2, u), \tag{7}$$

$$x_2' = \frac{1}{\sigma}(\mu + u_m u)x_1 + [1 - d(u)](a - x_1)x_2 \equiv g_2(x_1, x_2, u), \tag{8}$$

where $(\,)' = d(\,)/dt$, $\sigma = M/N \gg 1$, $a \geq 2$, $10^{-1} \leq u_m \equiv p_{\max} - p_{\min} \leq 1$ and $0 < \mu \equiv \bar{\mu} + p_{\min} \ll 1$. (Typically, we have $\sigma \geq 10$, $a = 2$, and $\mu = 10^{-1} \ll u_m$.) For simplicity, we will first focus our discussion herein to the death rate of the form (5) with $d_m = 1$, i.e.,

$$d(u) = 1 - (1 - u)^{\alpha} \tag{9}$$

for some real parameter α. (Results for $0 \leq d_m < 1$ can be found in [17].) As we shall see later, the three cases $\alpha > 1$, $\alpha = 1$, and $\alpha < 1$ have to be treated separately. The case of a general nonnegative, monotone increasing $d(u)$ is treated in Section 7.

The two ODE (7) and (8) are subject to the following three auxiliary conditions:

$$x_1(0) = 1, \quad x_2(0) = 0, \quad x_2(T) = 1, \tag{10}$$

where T is the unknown time when the dangerous mutant cell population reaches the target size.

The problem now becomes one of choosing a control function $u(t)$, the normalized gross chromosomal change rate $u(t)$ to minimize the time T needed to reach the target population size of the dangerous mutants. The admissible controls are restricted to the class of piecewise continuous functions (with only finite jump discontinuities) on the interval $[0, T]$ satisfying the inequality constraint,

$$0 \leq u \leq 1, \tag{11}$$

denoted by Ω, and nonnegativity constraints on the cell populations,

$$x_1 \geq 0, \quad x_2 \geq 0. \tag{12}$$

This shortest time problem is conventionally recast in the standard form of choosing $u(t)$ from Ω to minimize the performance index

$$J = \int_0^T 1\, dt, \tag{13}$$

subject to the equations of state (7) and (8), the boundary conditions (10), and the inequality constraints (11) and (12) with $T > 0$ to be determined as a part of the solution.

3. The Hamiltonian and adjoint variables

3.1. The maximum principle

The key to the solution of our optimal control problem is the Maximum Principle [28–32] for the *Hamiltonian:*

$$H = 1 + \lambda_1(t)g_1 + \lambda_2(t)g_2, \tag{14}$$

where λ_1 and λ_2 are the two continuous and piecewise differentiable *adjoint* (or co-state) *variables* for the problem chosen to satisfy two *adjoint ODE*,

$$\lambda_1' = -\left(\lambda_1 \frac{\partial g_1}{\partial x_1} + \lambda_2 \frac{\partial g_2}{\partial x_1}\right), \quad \lambda_2' = -\left(\lambda_1 \frac{\partial g_1}{\partial x_2} + \lambda_2 \frac{\partial g_2}{\partial x_2}\right), \tag{15}$$

and (for the given auxiliary conditions on the state variable x_1 and x_2) one Euler boundary condition (also known as a *transversality condition*) for the problem:

$$\lambda_1(T) = 0. \tag{16}$$

Note that (7), (8), and (15) form a Hamiltonian system [28]. With the admissible controls restricted to be piecewise continuous, the state and adjoint variables are continuous throughout $(0, T)$.

The Maximum Principle requires every optimal control function $\bar{u}(t)$ of the shortest time problem to satisfy the following necessary conditions:

1. Four continuous functions $\{\bar{x}_1(t), \bar{x}_2(t), \bar{\lambda}_1(t), \bar{\lambda}_2(t)\}$ exist and satisfy the Hamiltonian system of four differential Equations (7), (8), (15), and four auxiliary conditions in (10) and (16) for the optimal control.
2. The terminal time T satisfies a free end condition $[H(T)]_{u=\bar{u}(T)} = 0$ which becomes

$$[H(T)]_{u=\bar{u}(T)} = [1 + \bar{\lambda}_2 \bar{g}_2]|_{t=T} = 0, \tag{17}$$

with

$$\begin{aligned}[\bar{g}_2]_{t=T} &= [g_2(\bar{x}_1(t), \bar{x}_2(t), \bar{\lambda}_1(t), \bar{\lambda}_2(t), \bar{u}(t))]_{t=T} \\ &= g_2(\bar{x}_1(T), 1, 0, \bar{\lambda}_2(T), \bar{u}(T))\end{aligned} \tag{18}$$

after simplification by the transversality condition (16) and the terminal condition in (10).

3. For all t in $[0, T]$, the Hamiltonian achieves its minimum with $u = \bar{u}(t)$, i.e.,

$$H(\bar{x}_1(t), \bar{x}_2(t), \bar{\lambda}_1(t), \bar{\lambda}_2(t), \bar{u}(t)) = \inf_{v \in \Omega} [H(\bar{x}_1(t), \bar{x}_2(t), \bar{\lambda}_1(t), \bar{\lambda}_2(t), v)]. \tag{19}$$

4. If there should be a finite jump discontinuity in the optimal control $\bar{u}(t)$ at the instant T_s, the Hamiltonian is required to be continuous at T_s:

$$[H]_{t=T_s-}^{t=T_s+} = 0. \tag{20}$$

Even with the help of the Maximum Principle, determination of the optimal control for our problem is far from straightforward (even for the case of $\alpha = 1$ so that $d(u) = u$). In principle, we simply have to solve a two-point boundary value problem (BVP) after $\bar{u}(t)$ is specified. The crux of the problem is finding the optimal control which may be either in terms of the state and adjoint variables (through the stationary condition (26) for example) or known completely (such as an upper or lower corner control), or a combination of the three, each for a different segment of $[0, T]$. An iterative scheme was developed in [17] for approximate solutions which converges to the optimal solution for the different combinations of parameter values used but was not proved to converge in general.

In the remaining subsections of this section, we will establish some preliminary results that depend only on the following properties of the death rate function:

$$d(0) = 0, \quad d(1) = 1, \quad 0 < d(u) < 1, \quad \text{and} \quad d'(u) > 0 \tag{21}$$

for $0 < u < 1$, and not the specific form (9). The optimal control $\bar{u}(t)$ will then be established for our one-step genetic instability problem in the next three sections for $\alpha = 1, \alpha > 1$, and $\alpha < 1$, separately. The final analytical characterization of the optimal control will in turn shed light on why the iterative scheme in [17] should converge.

To simplify the discussion, we take $d_m = 1$; discussion of $d_m < 1$ cases can be found in [17].

3.2. Adjoint variables adjacent to terminal time

For any admissible control $u(t)$, the two adjoint differential Equations (15) induced by the state Equations (7) and (8) take the form

$$\lambda_1' = -\lambda_1\{(1 - 2x_1)[1 - d(u)] - (u_m u + \mu)\} - \frac{\lambda_2}{\sigma}\{(u_m u + \mu) - \sigma x_2[1 - d(u)]\}, \tag{22}$$

$$\lambda_2' = -\lambda_2(a - x_1)[1 - d(u)]. \tag{23}$$

With the admissible controls restricted to be piecewise continuous (possibly with finite jump discontinuities only), the state and adjoint variables are continuous and piecewise continuously differentiable (with first derivatives continuous except possibly for finite jump discontinuities at finitely many points) throughout $(0, T)$. All equalities and inequalities involving derivatives of $\lambda_1(t)$ and $\lambda_2(t)$ are meant to hold everywhere in $[0, T]$ except at the finite number of jump discontinuities of the optimal control, if any.

The following lemma on the sign of $\lambda_2(t)$ (deduced from the ODE (23) and the end condition (17)) is important to the determination of the optimal control for the shortest time:

LEMMA 1. *$\lambda_2(t) < 0$ and $\lambda_2'(t) \geq 0$ for $0 \leq t \leq T$.*

Proof: From (17) and $0 \leq x_1 \leq 1$ (see (15)), we have

$$\lambda_2(T) = -\left[\frac{1}{g_2}\right]_{t=T}$$
$$= -\left[\frac{\sigma}{(\mu + u_m u)x_1 + \sigma[1 - d(u)](a - x_1)x_2}\right]_{t=T} < 0. \qquad (24)$$

given $0 \leq d(u) \leq 1$ by (11). With the same restrictions on the death rate, we have from (23)

$$[\lambda_2']_{t=T} = -[\lambda_2(a - x_1)\{1 - d(u)\}]_{t=T} \geq 0, \qquad (25)$$

(with $\lambda_2'(t) = 0$ for $u(t) = 1$ for example). It is understood that all derivatives of state and adjoint variables at the end points are one-sided derivatives. The two conditions (24) and (25) together implies $\lambda_2(t) < 0$ in some neighborhood of $t = T$. The lemma follows from this local result and the monotonicity of λ_2 as implied by the ODE (23) and (11). ■

From $\lambda_2(T) < 0$ and $\lambda_1(T) = 0$, we have $\lambda_2 - \sigma\lambda_1 < 0$ at the terminal time which in turn gives the following result useful in subsequent developments.

LEMMA 2. *In some interval (T_0, T) adjacent to the terminal time T, we have: (i) $(\lambda_2 - \sigma\lambda_1) < 0$ for either corner solution; (ii) $\lambda_1(t) > 0$ for a lower corner control, and (iii) $\lambda_1(t) < 0$ for an upper corner control.*

Proof: Property (i) follows immediately from $\lambda_2(T) - \sigma\lambda_1(T) = \lambda_2(T) < 0$ (see (24)) and the continuity of the adjoint variables. For the remaining two properties, we note that the first adjoint ODE requires

$$\lambda_1'(T) = -\frac{\lambda_2(T)}{\sigma}\{(u_m u + \mu) - \sigma[1 - d(u)]\}_{t=T}$$

which (with $d(1) = 1$) is positive for $u = 1$ and negative for $u = 0$ (given $d(0) = 0$ and $\sigma \gg (\mu + u_m) = O(1)$). By the continuity of the adjoint variables, λ_1 is increasing for $u = 1$ and decreasing for $u = 0$ in an interval adjacent to the terminal time T. Take (T_0, T) to be the smallest of the relevant intervals. ■

3.3. The interior control

Suppose $\bar{u}(t)$ is a control function that minimizes the time to cancer (and necessarily satisfies the optimality condition (19)). The prime candidates for such an *optimal control* are those well-defined $u_i(t)$ that render the Hamiltonian H stationary

$$\left[\frac{\partial H}{\partial u}\right]_{u=u_i} = 0, \tag{26}$$

and satisfy the strict-inequality version of the constraint (11), known as *interior controls*. To explore the consequences of (26), we rewrite the Hamiltonian as

$$H = 1 + \frac{1}{\sigma}(\mu + u_m u)R(t) + \{1 - d(u)\}D(t), \tag{27}$$

where

$$D(t) \equiv \lambda_1 x_1(1 - x_1) + \lambda_2 x_2(a - x_1), \quad R(t) = x_1(\lambda_2 - \sigma\lambda_1) \tag{28}$$

for brevity. Specializing it to the death rate (9) with $d^{\cdot} = \alpha(1 - u)^{\alpha - 1}$, the stationary condition (26) becomes

$$\begin{aligned} d^{\cdot} &= \alpha(1 - u_i)^{\alpha - 1} \\ &= \frac{u_m}{\sigma}\frac{R(t)}{D(t)} \equiv \frac{u_m}{\sigma} r(t), \quad (\,)^{\cdot} = \frac{d(\,)}{du}. \end{aligned} \tag{29}$$

For $\alpha \neq 1$, the condition (29) formally determines an *interior control* $u_i(t)$ in terms of $\{\lambda_i(t)\}$ and $\{x_j(t)\}$. A corresponding solution of the BVP for $\{x_k(t), \lambda_j(t)\}$ is called an "interior solution" for the shortest time problem.

As we see later, an "interior solution" may not be well defined or may violate the inequality constraints (11) in some part of the solution domain for some range of system parameter values. Even when it is well defined and satisfies (11), it may maximize the Hamiltonian (instead of minimizing it). Hence, some combination of the "*upper corner solution*" $u_1(t) = 1$, the "*lower corner solution*" $u_0(t) = 0$ and the *interior solution* $u_i(t)$ may have to be considered for an optimal solution.

To facilitate our discussion of the appropriate optimal control in subsequent sections, we note the following two exact relations upon observing the state and adjoint equations.

LEMMA 3. *With*

$$S(t) = x_1 [\lambda_2(a - 1 + \sigma x_2) + \lambda_1 \sigma x_1] \tag{30}$$

independent of the control variable $u(t)$, *we have*

$$R' = -[1 - d(u)]S(t), \quad D' = \frac{\mu + u_m u}{\sigma} S(t). \tag{31}$$

3.4. An upper corner control at the start but not at the end

Proposition 1 below is related to a well-known result for autonomous systems, and will be a key to our method for determining the optimal control configuration for different types of death rate functions.

PROPOSITION 1. *For an optimal control* $\bar{u}(t)$ *that leads to the target mutant cell population in the shortest time, the Hamiltonian (14) for the problem vanishes for all* t *in* $[0, T]$, *i.e.,* $\bar{H} \equiv [H]_{u=\bar{u}(t)} = 0$, *for all* t *in* $[0, T]$.

Proof: Except for points of a jump discontinuity of the control, we have

$$\frac{dH}{dt} = \frac{\partial H}{\partial u} u' + \sum_{i=1}^{z} \left\{ \frac{\partial H}{\partial x_i} x_i' + \frac{\partial H}{\partial \lambda_i} \lambda_i' \right\} = \frac{\partial H}{\partial u} \frac{du}{dt}. \tag{32}$$

For the optimal control $\bar{u}(t)$, the right hand side of (32) vanishes because either $\bar{u}(t)$ is an interior control so that $\partial H/\partial u = 0$ or it is a corner control in which case we have $d\bar{u}(t)/dt = 0$. Hence, $\bar{H}(t)$ is a constant in any interval where H is differentiable. With the free end condition (17), we have $\bar{H}(t) = 0$ in the interval $(T_s, T]$ if there should be a simple jump discontinuity in $\bar{u}$ at some earlier time $T_s < T$ in the interval $(0, T)$. The switching condition (20) required the constant $\bar{H}(T_s)$ on both sides of the jump in $\bar{u}(t)$ to be the same (and equal to zero by the free end condition). The observation allows us to extend $\bar{H}(t) = 0$ to the next switching point and finally for the entire interval $[0, T]$. ■

As an immediate consequence of Proposition 1, any optimal solution must start with an upper corner control.

PROPOSITION 2. *For any well-defined death rate* $0 \le d(u) \le 1$, *an optimal solution for our minimum terminal time problem requires (i)* $\bar{u}(0) = 1$, *and (ii)*

$$\lambda_2(0) - \sigma \lambda_1(0) = -\frac{\sigma}{\mu + u_m}. \tag{33}$$

Proof: By the two initial conditions in (10), we have from (14)

$$[H]_{t=0} = 1 + \frac{\mu + u_m u(0)}{\sigma} R(0) = 1 + \frac{1}{\sigma} [\mu + u_m u(0)] [\lambda_2(0) - \sigma \lambda_1(0)]. \tag{34}$$

By Proposition 1, we must have $H(t = 0) = 0$ which in turn requires

$$[\lambda_2 - \sigma \lambda_1]|_{t=0} < 0$$

(for we would have $H(0) > 0$ otherwise). By condition (19) of the Maximum Principle, we must have $\bar{u}(0) = 1$ to minimize $H(0)$ and (33) follows from (34). ■

The following expression for $D(t)$ near the initial time follows from Proposition 2 and will be needed in subsequent developments.

COROLLARY 1. *For sufficiently small t, the expression $D(t)$ is accurately given by*

$$D(t) = D'(0)t + o(t) = \frac{\mu + u_m}{\sigma} [\lambda_2(0)(a - 1) + \sigma \lambda_1(0)] t + o(t). \tag{35}$$

Proof: With $\bar{u}(0) = 1$, we have from the state and adjoint equations together with (33)

$$x_1'(0) = -(\mu + u_m), \quad x_2'(0) = \frac{\mu + u_m}{\sigma} \quad \lambda_1'(0) = 1 \quad \lambda_2'(0) = 0, \tag{36}$$

leading to

$$\begin{aligned} D'(0) = &\left[\lambda_1' x_1(1 - x_1) + \lambda_2' x_2(a - x_1) + \lambda_1 x_1'(1 - 2x_1)\right. \\ &\left. + \lambda_2 x_2'(a - x_1) - \lambda_2 x_1' x_2\right]_{t=0} \\ = &\frac{\mu + u_m}{\sigma} [\lambda_2(0)(a - 1) + \sigma \lambda_1(0)]. \end{aligned} \tag{37}$$

With $D(0) = 0$, let

$$D_2(t) = D(t) - D_1 t$$

where $D_1 = (\mu + u_m)[\lambda_2(0)(a - 1) + \sigma \lambda_1(0)]/\sigma$. Because $D(0) = 0$, we have

$$\lim_{t \to 0} \left[\frac{D_2(t)}{t} \right] = \lim_{t \to 0} \left[\frac{D(t) - D(0)}{t} - D_1 \right] = D'(0) - D_1 = 0$$

and hence (35). ■

At terminal time, we have $\lambda_1(T) = 0$ and therewith

$$H(t = T) = \begin{cases} 1 + \dfrac{1}{\sigma}\lambda_2(T)\{\mu x_1(T) + \sigma[a - x_1(T)]\} & (u(T) = 0) \\ 1 + \dfrac{1}{\sigma}\lambda_2(T)x_1(T)(\mu + u_m) & (u(T) = 1) \end{cases} . \quad (38)$$

Relevant biological parameter value ranges are $a \geq 2$, $\sigma \gg 1, u_m \leq 1$ and $\mu \ll 1$ so that $\sigma(a-1) > u_m$ and $x_1(1 + u_m/\sigma) < 2 \leq a$ for any feasible solution so that

$$[H(t = T)]_{u=0} < [H(t = T)]_{u=1} . \quad (39)$$

We have then the following result.

PROPOSITION 3. *An optimal control does not end with an upper corner control at terminal time, i.e.,* $\bar{u}(T) < 1$.

Note that the optimal control may or may not end with a lower corner control because an interior control, if feasible, would be superior.

3.5. Optimal control for the shortest time to cancer

The proofs of Lemmas 1–3, Propositions 1–3 and Corollary 1 only made use of the properties of $d(u)$ stipulated in (21) and not the specific functional form (9) for the death rate, particularly not the convexity of $d(u)$. In the next three sections, we establish the following optimal control for the shortest time to cancer for the oncogene activation problem for the class of death rate function (9) with $\alpha = 1, \alpha > 1$, and $\alpha < 1$, respectively:

- For $\alpha = 1$, the unique optimal control is bang-bang (see Proposition 6).
- For $\alpha > 1$, the unique optimal control is bang-bang (see Proposition 11).
- For $0 < \alpha < 1$, the unique optimal control is a monotone decreasing interior control, $\bar{u}(t) = u_i(t) \geq 0$ for $t \leq T_s^*$ with $u_i(0) = 1$ and $u_i(T_s^*) = 0$, and $\bar{u}(t) = 0$ for $T_s^* \leq t \leq T$ if $T_s^* < T$ as specified in Proposition 13.

We then extend the results for $\alpha \neq 1$ cases to general death rates with the properties (21) and the corresponding type of convexity.

4. Linear death rates ($\alpha = 1$)

4.1. Upper corner control for an initial time interval

In this section, we consider the special case $\alpha = 1$ for which the death rate (5) is linear in the control $u(t)$. For this case, the stationary condition (9) does

not involve the control u and therefore does not define an interior solution $u_i(t)$. The conventional method of solution is to examine the *singular solution*, possibly augmented by a most rapid approach (and departure) to accommodate the boundary conditions [28, 30]. With

$$H = 1 + \frac{\mu}{\sigma} R(t) + D(t) + u \left\{ \frac{u_m}{\sigma} R(t) - D(t) \right\}, \tag{40}$$

the singular solution corresponds to

$$\frac{u_m}{\sigma} R(t) = D(t), \tag{41}$$

if it exists. Note that (41) does not determine the optimal control. However with (41), the Hamiltonian (40) now becomes

$$H = 1 + \frac{\mu}{\sigma} R(t) + D(t) = 1 + \frac{\mu + u_m}{\sigma} R(t), \tag{42}$$

which corresponds to $[H]_{u=1}$. We state this observation in the form of the following basic Lemma 4.

LEMMA 4. *If the singular solution defined by* (41) *is feasible in some interval* (T_1, T_2), *then the corresponding control is an upper corner control,* $\bar{u}(t) = 1$, *in that interval.*

Given $D(0) = 0$ and $R(0) < 0$, a singular solution is feasible at least in some interval $[0, T_1^*)$ adjacent to the starting time. By the optimality condition (19) and the continuity of state and adjoint variables, the control $\bar{u}(t) = 1$ in $[0, T_1^*)$ is consistent with the following extension of Proposition 2 for a death rate linear in the control variable.

PROPOSITION 4. *For* $\alpha = 1$, *an optimal solution for the minimum terminal time problem requires*

$$(i)\ \bar{u}(t) = 1, \quad and\ (ii)\ R(t) = x_1(\lambda_2 - \sigma\lambda_1) = -\frac{\sigma}{\mu + u_m} < 0, \tag{43}$$

so that $\lambda_2 - \sigma\lambda_1 < 0$ *is monotone decreasing in t, all in some finite interval* $[0, T_1^*)$ *with* $T_1^* < T$.

Proof: Part (ii) of the Proposition is an immediate consequence of $H(t) = 0$ with $\bar{u}(t) = 1$ (see (42)). Proposition 3 requires $T_1^* < T$. ■

4.2. Lower corner control for a terminal time interval

We have $\bar{u}(T) < 1$ from Proposition 3. For $\alpha = 1$, a more specific statement can be made about what happens at and near the terminal time.

PROPOSITION 5. *For $\alpha = 1$, the optimal control must be a lower corner control $\bar{u}(t) = 0$ in some finite interval $(T_0^*, T]$ adjacent to the terminal time.*

Proof: With $\lambda_1(T) = 0$ and $\lambda_2(T) < 0$ (by Lemma 1), we have at the terminal time

$$\frac{u_m}{\sigma} R(T) = \frac{u_m}{\sigma} \lambda_2(T) > \lambda_2(T)(a-1) > \lambda_2(T)(a - x_1(T)) = D(T),$$

given $a \geq 2$, $\sigma \gg 1$, $u_m \leq 1$, and $\mu \ll 1$ so that $\sigma(a-1) > u_m a \geq 2$. Hence, a singular solution does not exist in some interval $(T_0^*, T]$ adjacent to the terminal time. For a choice between the upper and lower corner control for $(T_0^*, T]$, we recall from (39) that the lower corner control minimizes the Hamiltonian at the terminal time T. Given the continuity of the state and adjoint variables, the proposition follows. ■

Biologically, the population of mutant cells is too small initially for clonal expansion of mutants to offset the net gain of mutants associated with genetic instability. It is therefore more advantageous to maximize the mutation rate to produce more mutants. As the mutant population grows and the normal cell population declines with time, clonal expansion of mutants more than offsets the loss of mutation eventually (certainly at or before the break even point t_c in the exact solutions for the mutant population growth due to the two corner controls reported in the Appendix). The gross chromosomal change should therefore be (reduced and possibly) switched off beyond some switch point T_s to attain the maximum growth of mutant cells.

4.3. Optimal control is bang-bang

For $\alpha = 1$ so that the death rate is linear in the control $u(t)$, we know from the results of the previous subsections that an optimal control must start with $\bar{u}(t) = 1$ in an interval $[0, T_1^*)$ and must end with $\bar{u}(t) = 0$ in an interval $(T_0^*, T]$. Because a singular solution, when it exists, corresponds to an upper corner control (see Lemma 4), we expect the following existence theorem:

LEMMA 5. *For $\alpha = 1$, a bang-bang control $\bar{u}(t)$ that starts with an upper corner control for $[0, T_s)$ and switches to a lower corner control at $T_s(<T)$ is feasible. Moreover, T_s is a root of $G(T_s) = 0$ in $(0, T)$ where*

$$G(T_s) = D(T_s) - \frac{u_m}{\sigma} R(T_s). \tag{44}$$

Proof: At the switch point T_s, the switch condition (20) requires

$$G(T_s) \equiv [H]_{t=T_s-}^{t=T_s+} = [1 + \lambda_1 g_1 + \lambda_2 g_2]_{t=T_s-}^{t=T_s+} = 0. \tag{45}$$

After applications of the various continuity conditions on state and adjoint variables, $G(T_s)$ is simplified to (44). Note that while all state and adjoint

variables are continuous at T_s, the quantities g_1 and g_2 are generally discontinuous there because they depend on the control $u(t)$. Several efficient iterative numerical schemes have been developed for finding T_s as well as the unknown minimum terminal time T (see [17]). Here, we are concerned only with the existence of the switch point inside the interval $(0, T)$.

From $D(0) = 0$ (see Corollary 1) and (33), we have

$$G(0) = \frac{u_m}{\mu + u_m} > 0, \tag{46}$$

and from Lemma 1

$$G(T) = [\lambda_2\{a - x_1(1 + u_m/\sigma)\}]_{t=T} < 0, \tag{47}$$

because $x_1(T)(1 + u_m/\sigma) < a$ for the biological parameter values previously specified after Proposition A.1. These two limiting values of G ensure the existence of at least one location for the switch point in $(0, T)$. ■

The switch point is actually unique as proved in Lemma 6 by showing that $G(\tau)$ is a monotone decreasing function of its argument.

LEMMA 6. *There exists exactly one switching between the two corner controls in* $(0, T)$ *at some instant* T_s *which is the unique solution of* (44) *and* (45).

Proof: Upon observing the two exact relations (31), we get

$$G'(\tau) = \frac{\mu + u_m}{\sigma} S(\tau) \tag{48}$$

with

$$S(\tau) = x_1 [\lambda_2(a - 1 + \sigma x_2) + \lambda_1 \sigma x_1]$$

as found in (30). If $\lambda_1(\tau)$ is nonpositive, then $G(T_s)$ is monotone decreasing given $S(\tau) < 0$ in that case. We need only to consider the positive λ_1 case. For $\lambda_1 > 0$, let $S(\tau) \le x_1[\lambda_2(a - 1 + \sigma x_2) + \lambda_1^* \sigma x_1]$ where λ_1^* is the maximum (positive value) of $\lambda_1(\tau)$ assumed at τ^* with $\lambda_1'(\tau^*) \le 0$ and $u(t^*) = 0$ (because $\lambda_1'(\tau^*) > 0$ if $u(t^*) = 1$ by (22)). In that case, we have

$$\lambda_1^* \le \frac{-\lambda_2^*}{2x_1^* + \mu - 1} \left(x_2^* - \frac{\mu}{\sigma}\right)$$

with $1 - 2x_1^* - \mu > 0$ so that $\lambda_1^* > 0$ and therewith

$$S(t) \le S(\tau^*) \le x_1(\tau^*)\lambda_2(\tau^*) \left\{(a - 1) + \frac{(1 - \mu)(1 - x_1(\tau^*))}{1 - \mu - 2x_1(\tau^*)}\right\} < 0. \tag{49}$$

keeping in mind $\lambda_2 < 0$ by Lemma 1 as well as $a - 1 > 0$.

Then $G'(\tau) < 0$ follows from (49) and (48) and the lemma is proved. ■

Given the switch condition (20) does not depend on T_s and the solution of (45)–(44) is unique, it is not possible to have more than one switch between the two corner controls in $(0, T)$. Thus, the optimal control is unique and bang-bang with $T_1^* = T_0^* = T_s$; there is no subsequent upper-lower corner control combination beyond the first and only switch. This gives the following final result for the $\alpha = 1$ case.

PROPOSITION 6. *For $\alpha = 1$, the minimum time to the target mutated cell population, $x_2(T) = 1$, is attained by the bang-bang control*

$$\bar{u}(t) = u_{bb}(t) = \begin{cases} u_1(t) = 1 & (0 \le t < T_s) \\ u_0(t) = 0 & (T_s < t \le T) \end{cases} \tag{50}$$

with the only switch point T_s determined by (45)–(44).

For the switch point T_s, we note from the exact solutions for the two corner control obtained in the Appendix (see Proposition A.1) that the mutant cell population $x_2^{(1)}(t)$ generated by the upper corner control from the start becomes equal to that by the lower corner control from the start, $x_2^{(0)}(t)$, at some finite time $t_c > 0$, i.e., $x_2^{(1)}(t_c) = x_2^{(0)}(t_c)$. Given $x_2^{(1)}(t) - x_2^{(0)}(t) > 0$ for $t < t_c$ and $x_2^{(1)}(t) - x_2^{(0)}(t) < 0$ for $t > t_c$ (see (A.7)), the switch is expected to take place at some instant $T_s \le t_c$. This was confirmed by the numerical solutions obtained in [17].

5. Concave death rates ($\alpha > 1$)

5.1. Upper corner control at and near the starting time

We already know from Proposition 2 that $R(0) = \lambda_2(0) - \sigma\lambda_1(0) = -\sigma/(\mu + u_m) < 0$ and $\bar{u}(0) = 1$ for any α. By the continuity of state and adjoint variables, $R(t) = x_1(\lambda_2 - \sigma\lambda_1)$ remains negative for all t in some interval $[0, T_1^*)$ with $T_1^* > 0$. Given $D(t) = O(t)$ for t in $(0, T_2^*)$ for some $T_2^* > 0$ by (35), the stationary condition (29) leads to an interior control that is either not real-valued or larger than unity for sufficiently small t. In either case, the optimal control must be a corner control adjacent to starting time. With

$$[H(t)]_{u=1} = 1 + \frac{\mu + u_m}{\sigma} R(t), \quad [H(t)]_{u=0} = 1 + \frac{\mu}{\sigma} R(t) + D(t), \tag{51}$$

and with $D(t) < u_m R(t)/\sigma$ for sufficiently small positive t, say in the interval $(0, T_3^*)$, the optimality condition (19) leads to Proposition 7.

PROPOSITION 7. *For $\alpha > 1$, an optimal solution for the minimum terminal time problem requires (i) $\bar{u}(t) = 1$ in some finite interval $[0, T_1)$, and (ii) $x_1(\lambda_2 - \sigma\lambda_1) = -\sigma/(\mu + u_m) < 0$ in some interval $[0, T_1)$.*

Proof: Apply Proposition 1 to $[H]_{u=1}$ and take $T_1 = \min\{T_2^*, T_3^*\}$ with T_2^* and T_3^* as defined in the previous paragraph. ■

5.2. Lower corner control adjacent to terminal time

To see what happens beyond the interval $[0, T_1)$, we begin by proving that an interior control is not optimal at, and near the terminal time T for $\alpha > 1$.

PROPOSITION 8. *For $\alpha > 1$, the interior solution is maximizing for some interval adjacent to T, say $(T_0, T]$ with $0 < T_0 < T$.*

Proof: Differentiating the Hamiltonian twice partially with respect to the control u gives

$$\frac{\partial^2 H}{\partial u^2} = -D(t)d^{\cdot\cdot}(u). \tag{52}$$

Because $\lambda_1(T) = 0$ and $\lambda_2(T) < 0$ (from Lemma 1), we have

$$D(T) = [\lambda_1 x_1(1 - x_1) + \lambda_2 x_2(a - x_1)]_{t=T} = \lambda_2(T)[(a - x_1(T)] < 0. \tag{53}$$

By continuity, we have

$$-D(t) = -[\lambda_1 x_1(1 - x_1) + \lambda_2 x_2(a - x_1)] > 0$$

in $(T_0, T]$ for some $T_0 < T$. It follows that for $0 < u_i(t) < 1$

$$\frac{\partial^2 H}{\partial u^2} < 0 \tag{54}$$

in $(T_0, T]$, given $d^{\cdot\cdot}(u_i) < 0$ for $\alpha > 1$. In other words, the interior solution is maximizing in the interval $(T_0, T]$ and therefore is not optimal for a minimum terminal time. ■

As a consequence of Proposition 8, the optimal control adjacent to the terminal time must be a corner control.

PROPOSITION 9. *For $\alpha > 1$, any optimal solution for a minimum terminal time ends with a lower corner control in some finite interval $(T_0, T]$, i.e., $\bar{u}(t) = 0$ for t in $(T_0, T]$ for some $T_0 < T$.*

Proof: Because we can only choose between the two corner controls, the inequality (39) requires $\bar{u}(T) = 0$. Continuity of the state and adjoint variables

and the optimality condition (19) extend the lower corner control to a finite interval $(T_0, T]$ adjacent to the terminal time. ■

5.3. Switching between upper corner and interior control not feasible

Our findings so far require an optimal solution for the $\alpha > 1$ case to start with an upper corner control in some finite interval $[0, T_1)$ and end in a lower corner control in another finite interval $(T_0, T]$ with $T_1 \leq T_0$. One scenario would have $T_1 = T_0 (= T_s)$ in which case the optimal control is bang-bang as given by (50). However, unlike the $\alpha = 1$ case, we have not ruled out an interior control in an interior segment of the interval $[0, T]$, the interval $T_1 < t < T_0$ for instance. We show in this subsection that an interior control following the upper corner control at the start is maximizing. Hence, it is not possible to have a *two-switch solution* with an interior control in the interval $T_1 < t < T_0$ connecting the upper corner control at the start to the lower corner control at the terminal time. Lemma 7 is the first step of this effort.

LEMMA 7. *If the upper corner control at the start is followed by an interior control at some $T_s > 0$ for $\alpha > 1$, then $\bar{u}(t)$ is continuous at T_s (with $u_i(T_s) = 1$) and $\lambda_2(T_s) - \sigma\lambda_1(T_s) < 0$.*

Proof: At the switch point T_s, the switch condition (20) may be re-arranged to read

$$\left[\frac{1 - d(u_i)}{1 - u_i}\right]_{t=T_s} = \frac{u_m}{\sigma}\frac{R(T_s)}{D(T_s)} = d^{\cdot}(u_i(T_s)) \tag{55}$$

upon observing (29). However for a death rate of the form (9), we have $d^{\cdot}(u_i(T_s)) = \alpha(1 - u_i(T_s))^{\alpha-1}$ while the left hand side is equal to $(1 - u_i(T_s))^{\alpha-1}$. Hence, the switch condition requires

$$\left[(1 - u_i)^{\alpha-1}\right]_{t=T_s} = \left[\alpha(1 - u_i)^{\alpha-1}\right]_{t=T_s}$$

which cannot be met (given $\alpha > 1$) unless $u_i(T_s) = 1$ so that $\bar{u}(t)$ is continuous at T_s. The condition $u_i(T_s) = 1$ requires

$$H(t = T_s) = \left[1 + \frac{x_1}{\sigma}(\mu + u_m)(\lambda_2 - \sigma\lambda_1)\right]_{t=T_s}$$

and therewith $x_1(\lambda_2 - \sigma\lambda_1) = -\sigma/(\mu + u_m) < 0$ at $t = T_s$ by Proposition 1. ■

We can now show that it is not feasible for the upper corner control in $[0, T_s)$ to be followed by an interior control.

PROPOSITION 10. *An optimal control consisting of an upper corner control with $\bar{u}(t) = 1$ for $[0, T_s)$ at the start of the growth process followed by an*

interior control $u_i(t)$ *defined by* (26) (*or*(29)) *in* (T_s, T_1^*) *is maximizing in* $T_s < t < T_1^*$ *for any* $\alpha > 1$.

Proof: Suppose an upper corner control should be followed by an interior control at T_s. Consider first the case $T_s < T_1$ (with T_1 as defined in Proposition 7). In that case, we have from the previous lemma $\lambda_2 - \sigma\lambda_1 < 0$ at the switch point. Continuity of the adjoint variables extends $\lambda_2 - \sigma\lambda_1 < 0$ to a larger interval $[0, T_1^*)$. For t in (T_s, T_1^*), the stationarity condition (26) becomes

$$d^{\cdot} = \alpha(1-u)^{\alpha-1} = \frac{u_m}{\sigma}\frac{R(t)}{D(t)} \quad \left(T_s < t < T_1^*\right). \tag{56}$$

Given $d^{\cdot} > 0$ for $0 < u_i < 1$, we must have $D(t) = \lambda_1 x_1(1-x_1) + \lambda_2 x_2(a - x_1) < 0$ in the interval (T_s, T_1^*) in order for the stationary condition to specify a well-defined *unique* interior control. The expression (52) for $\partial^2 H/\partial u^2$ is then negative and shows that u_i maximizes $H(u)$ in that interval. The upper-interior control configuration in $[0, T_1^*)$ is therefore not optimal for our minimum time problem.

If $T_s = T_1$, then the same argument leads to the same conclusion that the interior control is not optimal in (T_s, T_1^*) with $T_1^* > T_1$. An upper corner-to-interior control configuration is therefore not optimal. adjacent to the starting time. ■

5.4. *Optimal control for* $\alpha > 1$ *is bang-bang*

It follows from Proposition 10 that the first switch in $\bar{u}(t)$ should be from the initial upper corner control to a lower corner control. At the switch point, the switch condition (20) again takes the form $G(T_s) = 0$ with $G(\cdot)$ given by (44). By the same proof as the one for Lemma 5, the existence of such a switch point T_s in $(0, T)$ is assured. There remains the possibility of subsequent combinations of interior-lower corner control starting at some $T_2 > T_s$ connecting the first lower corner control to last lower control adjacent to the terminal time (as required by Proposition 9), switching to the last lower corner control at T_3. We show presently that such a multiswitch control is also maximizing.

LEMMA 8. *For* $\alpha > 1$, *any multiswitch control with an interior control in* (T_2, T_3) *between two lower corner controls is maximizing in* (T_2, T_3) *and hence not optimal.*

Proof: The switch condition (20) at either switch point $\tau(=T_2$ or $T_3)$ may be simplified to read

$$d(u_i(\tau))D(\tau) = \frac{u_m}{\sigma}u_i(\tau)R(\tau)$$

or, upon observing (29) and setting $\xi = u_i(\tau)$,

$$g(\xi) \equiv d(\xi) - \xi d^{\cdot}(\xi)$$
$$= 1 - (1-\xi)^{\alpha} - \alpha\xi(1-\xi)^{\alpha-1} = 0 \tag{57}$$

Now $g(0) = 0$ and $dg/d\xi > 0$ for $0 < \xi < 1$; the condition (57) requires $u_i(T_s) = 0$. In that case, we have at either switch point, T_2 or T_3,

$$[H]_{t=\tau} = 1 + D(\tau) + \frac{\mu}{\sigma} R(\tau)$$
$$= \left[1 + \lambda_1 x_1(1-x_1) + \lambda_2 x_2(a - x_1) + \frac{\mu}{\sigma} x_1(\lambda_2 - \sigma\lambda_1)\right]_{t=\tau}.$$

With $\sigma \gg 1$, $\mu \ll 1$ and $a \geq 2$, Proposition 1 requires $D(t) < \mu R(t)/\sigma < 0$ in a neighborhood of each switch point. Because $d^{\cdot\cdot}(u_i) < 0$ for $\alpha > 1$, we have from the expression (52) for $\partial^2 H/\partial u^2$ that u_i is maximizing in some intervals (T_2, T_2^*) and (T_3^*, T_3) and therefore not optimal. ■

With u_i maximizing in (T_2, T_3), the optimal control for the shortest time must be the lower corner solution for $(T_s, T]$. Hence, the solution for our minimum time problem for the $\alpha > 1$ case is a one-switch bang-bang control.

PROPOSITION 11. *For $\alpha > 1$, the minimum time to target mutated cell population, $x_2(T) = 1$, is attained by the unique optimal (bang-bang) control (50) with the only switch point T_s determined by $G(T_s) = 0$ where $G(\cdot)$ is given by (44).*

Proof: The proposition is a consequence of Proposition 10, Lemma 8 together with Proposition 6. ■

6. Convex death rates ($\alpha < 1$)

6.1. Monotone decreasing interior control adjacent to initial and terminal time

From the numerical solutions reported in [17], we know the optimal mutation rate for the shortest time to cancer is not bang-bang for $\alpha < 1$, at least for the cases computed. In this section, we provide a more definitive description of the optimal mutation rate for the class of death rate functions (5) with $\alpha < 1$.

Again we have $\bar{u}(0) = 1$ from Proposition 2 and therewith $R(0) = \lambda_2(0) - \sigma\lambda_1(0) < 0$ for all α. However unlike the situation for $\alpha \geq 1$, the stationary condition $\partial H/\partial u = 0$ now defines a *unique* real-valued interior control for some interval $(0, T_i)$ after the initial time. To see this, we note the

relation $d^{\cdot}(u) = \alpha/(1-u)^{1-\alpha}$ enables us to re-write (29) as

$$u_i(t) = 1 - \left\{ \frac{\alpha\sigma}{u_m} \frac{D(t)}{R(t)} \right\}^{1/(1-\alpha)}. \tag{58}$$

With $\bar{u}(0) = 1$, the expression (35) in Corollary 1 and (33) of Proposition 2 may be used to give the asymptotic behavior of $u_i(t)$ for $0 \leq t \ll 1$:

$$\begin{aligned} u_i(t) &= 1 - \left\{ \frac{\alpha\sigma}{u_m} \left[\frac{D'(0)t + o(t)}{-[\sigma/(\mu + u_m)] + o(t)} \right] \right\}^{1/(1-\alpha)} \\ &\sim 1 - \left\{ -\frac{\alpha}{\sigma u_m} (\mu + u_m)^2 \left[\lambda_2(0)(a-1) + \sigma\lambda_1(0) \right] t + o(t) \right\}^{1/(1-\alpha)}, \end{aligned} \tag{59}$$

with $\sigma D'(0)/(\mu + u_m) = S(0) = \lambda_2(0)(a-1) + \sigma\lambda_1(0) < 0$ by the optimality condition (19) (see also Lemma 3) in order for $u_i(t)$ to be real-valued in some finite interval $(0, T_i)$ adjacent to the initial time with $0 < u_i(t) < 1$. The resulting interior control is minimizing.

LEMMA 9. *For $\alpha < 1$, the interior control is minimizing in some interval $(0, T_i)$ with $0 < T_i \leq T$ and therefore optimal for that interval.*

Proof: The lemma follows from

$$\frac{\partial^2 H}{\partial u^2} = -Dd^{\cdot\cdot}(u)$$

given $d^{\cdot\cdot}(u_i) > 0$ for $0 < \alpha < 1$ and

$$D(t) = D'(0)t + o(t) = \frac{\mu + u_m}{\sigma} \left[\lambda_2(0)(a-1) + \sigma\lambda_1(0) \right] t + o(t) < 0$$

for t in $(0, T_i)$. ■

Next, we show that the interior control is a monotone decreasing function adjacent to $t = 0$ and $t = T$.

LEMMA 10. *For $\alpha < 1$, the interior control $u_i(t)$ is well defined and monotone decreasing in some interval $(0, T_i)$.*

Proof: Recall that $D(t) < 0$ and $D'(t) < 0$ at least in some interval adjacent to $t = 0$. This implies $R(t) < 0$ and $R'(t) > 0$, the latter by Lemma 3. It follows that the ratio $D(t)/R(t)$ is positive and an increasing function of t in some interval $(0, T_i)$ and, by (58), $u_i(t)$ is a monotone decreasing function at least in $(0, T_i)$. ■

LEMMA 11. *For $\alpha < 1$, the interior control $u_i(t)$ is well defined and monotone decreasing in some interval $(T_i^*, T]$.*

Proof: At the terminal time, we have $\lambda_1(T) = 0$ so that

$$1 - u_i(T) = \left\{ \frac{\alpha\sigma}{u_m} \left[\frac{a - x_1(T)}{x_1(T)} \right] \right\}^{1/(1-\alpha)} > 0, \tag{60}$$

which defines a *unique* real-valued interior control $u_i(t)$ in some interval $(T_i^*, T]$. We show presently that $u_i'(t) < 0$ adjacent to the terminal time.

With $\lambda_1(T) = 0$, Lemma 1 gives

$$D(T) = \lambda_2(T)\,[a - x_1(T)] < 0, \quad R(T) = x_1(T)\lambda_2(T) < 0 \tag{61}$$

$$D'(T) = \frac{\mu + u_m u_i(T)}{\sigma} x_1(T)\lambda_2(T)(a - 1 + \sigma) < 0, \tag{62}$$

and

$$R'(T) = -[1 - d(u_i(T))]x_1(T)\lambda_2(T)(a - 1 + \sigma) > 0. \tag{63}$$

Continuity of the state, adjoint and interior control variables ensures also

$$D(t) < 0, \quad R(t) < 0, \tag{64}$$

$$D'(t) < 0, \quad R'(t) > 0, \tag{65}$$

in some interval (T_i^*, T). These inequality require the ratio $D(t)/R(t)$ in (58) to be positive and increasing with time; hence, the interior control $u_i(t)$is monotone decreasing in (T_i^*, T). ■

6.2. Lower corner control prior to terminal time?

Whether or not the two intervals $(0, T_i)$ and (T_i^*, T) overlap, a switch from the interior control to a corner control may take place for a number of reasons, including $u_i(t) < 0$. The following two results limit the switching possibilities.

LEMMA 12. *For $\alpha < 1$, switching between an interior control and an upper corner control is not admissible.*

Proof: At the switch point T_s (where $u_i(t)$ is well defined), the switch condition (20) requires

$$\begin{aligned} 1 - d(u_i(T_s)) &= \frac{u_m}{\sigma} \frac{R(T_s)}{D(T_s)} [1 - u_i(T_s)] \\ &= [(1 - u_i)\, d^{\cdot}(u_i)]_{t=T_s} \end{aligned}$$

or

$$[1 - u_i(T_s)]^{\alpha} = \alpha\,[1 - u_i(T_s)]^{\alpha}\,. \tag{66}$$

Given $0 < \alpha < 1$, the condition (66) requires $u_i(T_s) = 1$. However this requirement cannot be met because u_i is monotone decreasing from unity with $u_i(T_s) < 1$. ■

LEMMA 13. *For $\alpha < 1$, if there is a switch from an interior control to a lower corner control at T_0, then $u_i(T_0)$ must vanish.*

Proof: Suppose the first such switch takes place at T_0. For $\alpha < 1$, the switch condition (20) takes the form

$$d(u_i(T_0)) = \left[\frac{u_m u_i(t)}{\sigma}\frac{R(t)}{D(t)}\right]_{t=T_0} = u_i(T_0)d^{\cdot}(u_i(T_0))$$

which can be met only if $u_i(T_0) = 0$ given $0 < \alpha \leq d^{\cdot}(u_i) < \infty$, $d(0) = 0$, and $0 < d(u) < 1$ for $0 < u < 1$. ■

LEMMA 14. *For $\alpha < 1$, switching from a lower corner control to an upper corner control is not feasible.*

Proof: Suppose there is a switch from lower to upper corner control at some point T_s in $(0, T)$. By Lemma 11, the optimal control must end in an interior control or a lower corner control. However, a switch from upper corner control to an interior control at some $T_\ell > T_s$ is not feasible by Lemma 12. A switch from upper to lower corner control is also not feasible because there can only be one switch between the two corner controls by Lemma 6. ■

Remark 1. A less formal argument leading to the same conclusion is to note that at $t = T_s$, the quantity $D(t)$ must be more negative than $R(t)$; otherwise. the Hamiltonian (27),

$$H = 1 + \frac{1}{\sigma}(\mu + u_m u)R(t) + \{1 - d(u)\}D(t);$$

would not be minimized by the lower corner control. However with $x_2' > 0$, $x_1' < 0$ and $\lambda_1' < 0$ and $\lambda_2 < 0$, we see from the expressions for $D(t)$ and $R(t)$ in (28) that the inequality $D(t) < R(t)$ holds for $t > T_s$. In that case, a switch to the upper corner control would not minimize H and hence not optimal.

PROPOSITION 12. *If $\bar{u}(T_0^*) = 0$ for some T_0^* in $(0, T)$, then $\bar{u}(t) = 0$ for all t in $[T_0^*, T]$.*

Proof: Suppose T_0^* is the first zero of $\bar{u}(t)$ with $0 < T_0^* < T$. Then $\bar{u}(T_0^*) = 0$ cannot be preceded by an upper corner control because the latter cannot follow an interior control by Lemma 12. By Lemma 13, $\bar{u}(t)$ must be continuous at T_0^* with the interior control $u_i(t)$ positive and continuous for $(0, T_0^*)$ vanishing at T_0^*. Now, $\bar{u}(t)$ cannot switch from a lower corner to an upper corner control at some $T_\ell \geq T_0^*$ by Lemma 14. It also cannot switch to an interior control that is increasing from zero as such an interior control would not be optimal given the observation in Remark 1 (which can be made rigorous by showing $[D(t) - R(t)]' < 0$). Hence, we must have $\bar{u}(t) = 0$ for all t in $[T_0^*, T]$. ■

6.3. Corner control prior to terminal time?

Any optimal control should continue to be the positive monotone decreasing interior control $u_i(t)$ as long as it is well defined and does not violate the inequality constraint (11) because it is minimizing and superior to both corner controls. However, we saw in [17] that $u_i(t)$ may either become negative for $t > T_s^*$ (and hence necessitate a switch to the lower corner control) or remain positive for the entire interval $(0, T]$ depending on the system parameter values. The Lemma 15 confirm the latter possibility.

LEMMA 15. *If $\alpha\sigma(a-1)/u_m > 1$ for $\alpha < 1$, we have $u_i(t) < 0$ for $t > T_s^*$ for some $T_s^* < T$ (and $u_i(t)$ is therefore inadmissible in $(T^*, T]$ for some $T^* \geq T_s^* > 0$).*

Proof: We have from the stationary condition (58) and the end conditions at terminal time in (10) and (16):

$$u_i(T) = 1 - \left(\frac{\alpha\sigma\,[a - x_1(T)]}{u_m x_1(T)} \right)^{1/(1-\alpha)}. \tag{67}$$

With $0 < x_1(T) < 1$, the (positive) ratio inside the parentheses exceeds unity if $\alpha\sigma(a-1)/u_m > 1$. ■

Note that the condition $u_m/\sigma(a-1) < \alpha$ is only a sufficient condition for a switch to a lower corner control. At the other extreme, an obvious sufficient condition for $u_i(t)$ to be well defined and meets the inequality constraints (11) for $0 \leq t \leq T$ is:

$$\frac{\alpha\sigma\,[a - x_1(T)]}{u_m x_1(T)} \leq 1. \tag{68}$$

However, the condition (68) requires a knowledge of the unknown $x_1(T)$ and hence the solution of the problem. An explicit sufficient condition without any reference to the solution of the problem is possible (making use of the terminal time for the solution with the lower corner control throughout) but not useful.

We note only that the necessary condition $\alpha\sigma(a-1)/u_m < 1$ for the interior solution to hold for $0 < t \leq T$ is an immediate consequence of Lemma (15).

We are now ready for the characterization of the optimal control for our shortest time problem for the $\alpha < 1$ case

PROPOSITION 13. *For $\alpha < 1$, the optimal control $\bar{u}(t)$ starts with $\bar{u}(0) = 1$, follows continuously by a monotone decreasing interior solution, $\bar{u}(t) = u_i(t)$, for a period $(0, T_s^*)$ with (i) $\bar{u}(t) < 1$ for $t > 0$; (ii) $T_s^* < T$ if $\alpha\sigma(a-1) > u_m$, with $\bar{u}(t) = 0$ for all t in $[T_s^*, T]$, and (iii) $\alpha\sigma(a-1) < u_m$ if $T_s^* > T$.*

7. Genetic instability with a general death rate

In the previous three sections, we determined the optimal control for the shortest time to cancer for the three types of death rate function of the form (9) as described in Subsection 3.5. In this section, we consider more general form of the (normalized) death rate d as a function of the (normalized) mutation rate u. With the death rate associated with basic mutation already accounted for in the basic growth rates of cells, we continue to stipulate $d(0) = 0$ and will take $d(1) = 1$ and $d^{\cdot}(0) = \alpha > 0$ for simplicity but otherwise with no restriction on the form of $d(u)$ other than those stipulated in (21). For these general death rate functions, we have the following general results without any convexity/concavity assumption because they were proved in Section 3 without invoking the specific form (5) of the death rate.

PROPOSITION 14. *Lemmas* 1–3, *Propositions* 1–3, *and Corollary* 1 *all hold for general death rates with properties stipulated in* (21).

With the case of a linear $d(u)$ already treated in Section 4, we will discuss separately death rates which are *strictly concave* and those which are *strictly convex* in the next two sections. Note that the death rates considered are restricted to those with properties listed in (21) and $d^{\cdot}(0) = \alpha > 0$.

7.1. Strictly concave death rates

In this section, we consider (nonnegative, monotone increasing,) *strictly concave* death rates defined by

$$d^{\cdot}(v) < \frac{d(v) - d(u)}{v - u} < d^{\cdot}(u), \quad (0 < u < v < 1), \tag{69}$$

and

$$\frac{1 - d(u)}{1 - u} < d^{\cdot}(u), \quad \frac{d(u) - d(0)}{u - 0} = \frac{d(u)}{u} > d^{\cdot}(u). \tag{70}$$

for $0 < u < 1$. The class of death rates in (9) with $\alpha > 1$ is strictly concave.

Note: For a twice continuously differentiable, strictly concave death rate, we have

$$d^{\cdot\cdot}(u) < 0, \quad (0 < u < 1). \tag{71}$$

The stationary condition (26), taken in the form

$$D(t)d^{\cdot}(u_i) = \frac{u_m}{\sigma} R(t),$$

again does not define an interior control at the initial time given $D(t = 0) = 0$ and $R(0) = -\sigma/(\mu + u_m) < 0$. Then the proof of Proposition 7 applies to give the same result for the more general class of strictly concave death rates.

LEMMA 16. *For a strictly concave* $d(u)$ *with properties stipulated in* (21), *the optimal control starts with an upper corner control* $\bar{u}(t) = 1$ *with* $R(t) = x_1(\lambda_2 - \sigma\lambda_1) = -\sigma/(\mu + u_m) < 0$ *for a finite interval* $[0, T_1)$.

Next, we show that the optimal control must end in a lower corner control in two steps.

LEMMA 17. *Any interior control* $u_i(t)$ *is maximizing in an interval* $(T_0, T]$.

Proof: The proof is the same as that for Proposition 8 because only strict concavity of $d(u)$ and Lemma 1 (see also Proposition 14) were needed in that proof. ■

LEMMA 18. *Any optimal control ends with a lower corner control* $\bar{u}(t) = 0$ *over an interval* $(T_0, T]$.

Proof: The proof is identical to that Proposition 9. ■

LEMMA 19. *For the optimal control* $\bar{u}(t)$, *an upper corner control at the start cannot be followed by an interior control.*

Proof: As in Proposition 10, the continuity of the Hamiltonian (as a function of time) across any switch point T_s requires

$$\frac{1 - d(u_i)}{1 - u_i} = d^{\cdot}(u_i) \tag{72}$$

where $0 < u_i < 1$. The condition (72) violates the strict concavity of $d(u)$ (see (70) and cannot be met. ■

PROPOSITION 15. *For* $d(u)$ strictly concave, *the minimum time to the (normalized) target mutated cell population,* $x_2(T) = 1$, *is attained by the bang-bang control* (50) *with the only switch point* T_s *determined by* $G(T_s) = 0$ *where* $G(\cdot)$ *is given by* (44).

Proof: With the upper corner control at the start followed by a lower corner control, we only need to eliminate the possibility of interior-lower corner control or additional upper-lower corner control combinations following the admissible upper-lower corner control combination. This is accomplished as in Proposition 11 because only $d^{\cdot}(u_i) > 0$ and $d^{\cdot\cdot}(u_i) < 0$ were invoked in the proofs of leading up to that proposition. ■

7.2. Strictly convex death rates

In this section, we consider (nonnegative, monotone increasing,) strictly convex death rates $d(u)$ defined by

$$d^{\cdot}(v) > \frac{d(v) - d(u)}{v - u} > d^{\cdot}(u), \quad (0 < u < v < 1) \tag{73}$$

and

$$\frac{1 - d(u)}{1 - u} > d^{\cdot}(u) \quad \text{and} \quad \frac{d(u) - d(0)}{u - 0} = \frac{d(u)}{u} < d^{\cdot}(u). \tag{74}$$

for $0 < u < 1$. The death rate in (9) with $\alpha < 1$ is strictly convex.

Note: For a twice continuously differentiable, strictly convex death rate, we have

$$d^{\cdot\cdot}(u) > 0 \qquad (0 < u < 1). \tag{75}$$

We can now establish the analogues of Lemmas 9 and 10 (as well as Lemma 11) allowing us to state:

PROPOSITION 16. *For a strictly convex death rate with the properties* (21), *the optimal control* $\bar{u}(t)$ *starts with* $\bar{u}(0) = 1$, *follows continuously by a monotone decreasing interior solution,* $\bar{u}(t) = u_i(t)$, *for a period* $(0, T_s^*)$ *with* $\bar{u}(t) < 1$ *for* $0 < t < T_s^*$

The following analogue of Lemma 12 will have to be proved without the specific form (9) of the death rate:

LEMMA 20. *For a strictly convex death rate, a switch from an interior control to an upper corner control is not admissible.*

Proof: At the switch point T_s (where $u_i(t)$ is well defined), the switch condition (20) requires

$$1 - d(u_i(T_s)) = \frac{u_m}{\sigma} \frac{R(T_s)}{D(T_s)} [1 - u_i(T_s)]$$
$$= [(1 - u_i)\, d^{\cdot}(u_i)]_{t=T_s}$$

Because the death rate is strictly convex (see (74)), we must have $u_i(T_s) = 1$. However this requirement cannot be met because u_i is monotone decreasing from unity. ■

LEMMA 21. *For a strictly convex death rate, if there is a switch in $\bar{u}(t)$ from an interior control to a lower corner control at T_0, then $\bar{u}(T_0)$ must vanish.*

Proof: Suppose the first such switch takes place at T_0. The switch condition (20) and (26) requires

$$d(u_i(T_0)) = \left[\frac{u_m u_i(t)}{\sigma} \frac{R(t)}{D(t)}\right]_{t=T_0} = u_i(T_0) d^{\cdot}(u_i(T_0))$$

which because of strict convexity of $d(u)$ can only be met by $u_i(T_0) = 0$ given $0 < d^{\cdot}(u_i) < \infty$ and $d(0) = 0$ (see (74)). ■

PROPOSITION 17. *If $\bar{u}(T_0^*) = 0$ for some T_0^* in* $(0, T)$where$\alpha = \alpha^*(0)$, *then $\bar{u}(t) = 0$ for all t in $[T_0^*, T]$.*

Proof: The proof is the same as that for Proposition 12 except we need to invoke Lemmas 20 and 21 instead of Lemmas 12 and 13. ■

The sufficient condition for a switch to a lower corner control at $T_s^* < T$ is proved the same way as in Lemma 15. We have then the following second half of the optimal control characterization for strictly convex death rates:

PROPOSITION 18. *The optimal control for a strictly convex death rate with properties (21) starts with $\bar{u}(0) = 1$ follows continuously with an interior control $\bar{u}(t) = u_i(t)$ for t in $(0, T_s^*)$. If $\alpha\sigma(a - 1) > u_m$ where $\alpha = d^{\cdot}(0)$, then $T_s^* < T$ and $\bar{u}(t) = 0$ for all t in $[T_s^*, T]$. If on the other hand, $\bar{u}(t) = u_i(t)$ for $0 < t \leq T$, then we must have $\alpha\sigma(a - 1) < u_m$.*

Proof: At terminal time, we have by strict convexity and stationarity (see (26) and (27))

$$(0 <)\ d^{\cdot}(0) < d^{\cdot}(u_i(T)) = \frac{u_m}{\sigma} \frac{x_1(T)}{[a - x_1(T)]}.$$

so that $u_i(T) < 0$ if

$$\frac{u_m}{\sigma}\frac{x_1(T)}{[a - x_1(T)]} < \frac{u_m}{\sigma}\frac{1}{[a-1]} < d^{\cdot}(0) \equiv \alpha.$$

Thus, if $\alpha\sigma(a-1)/u_m > 1$, we have $u_i(t) < 0$ for $t > T^*$ for some $T^* < T$ so that $\bar{u}(t) = 0$ for all t in $(T^*, T]$ for some $T^* \geq T_s^* > 0$. We must have $T^* = T_s^*$, as a switch in control at T^* cannot be to an upper corner control because that cannot follow an interior control (which is monotone decreasing). ■

8. Conclusion

With the normalized mutation rate $u(t)$ of cells in our ODE models allowed to vary with time, the time path of $u(t)$ can be adjusted to give the shortest time to cancer. In this way, we have an optimal control problem with the minimization of the time it takes the cancerous colony to grow to a given size as the objective. The optimal control obtained as the solution of the mathematical problem gives the "best strategy" for the fastest growth of the cancerous colony. While biological implications of this study have been reported in [17], the present paper provides the mathematical analysis that led to the results there.

The control function, $u(t)$, which is the normalized mutation rate of cells, enters the problem in two different ways. On the one hand, it is responsible for generating cancerous cells out of normal cells; on the other hand, it contributes to cell death by means of a relationship, generally *nonlinear,* between the death rate and the mutation rate, $u(t)$. Another source of nonlinearity in our models is a logistic-type growth law for precancerous cell populations which are subject to homeostatic control. Both nonlinearities make the problem challenging from a control-theoretical prospective.

In general, there is no single "recipe" for solving such nonlinear optimal control problems. Our main strategy is to exploit the structure of the Hamiltonian for our model problems with the help of the Pontryagin Maximum Principle. The optimal solution is found to be generally a combination of an "interior" solution and the two "corner" solutions, the latter corresponding to constant control functions of the minimum and the maximum allowed value of the mutation rate. The precise nature of the optimal control is completely defined by the convexity of the death rate dependency on the mutation rate. If our death rate function $d(u)$ is a linear or a strictly concave monotonically increasing function of the mutation rate u, the optimal control is a one-switch "bang-bang" control. More specifically, u is piecewise constant, jumping from the "upper corner control" (the maximum allowed level of mutation rate) to the "lower corner control" (the minimum allowed level of mutation rate). On the other hand, if $d(u)$ is a strictly convex function of the mutation rate, then the optimal control $\bar{u}(t)$ starts at the maximum rate with $\bar{u}(0) = 1$ and follows continuously

with a continuous decreasing function of time, possibly switching to the lower corner control when the decreasing interior control reaches the lower bound. These characterizations of the optimal control were established first for death rate functions of the form (9) and then extended to strictly concave and convex without specifying the form of $d(u)$. The results obtained are in qualitative agreement with experimental observations about the decreased level of genetic instability at later stages of cancer development in several different cancers (see [15] for breast cancer, [16] for intestinal carcinoma in mice and humans, and [33–35] for proposed mechanisms of telomere-related reduction in the level of chromosomal instability).

There are several ways in which our theoretical results can be generalized and improved. For example, some preliminary results on a two-step model for carcinogenesis (e.g., *inactivation of a tumor suppressor gene*) involving a two-parameter family of functions $d(u)$ were reported in [17]. A more complete theoretical investigation of that model to document the similarities and differences in the optimal mutation rate would be of interest.

For the one-parameter family of death rate functions (9) investigated in the present paper, it was found that the optimal control strongly depends on the convexity of these functions. It would be interesting to generalize the results to other classes of death rates and investigate what happens, for instance, if the death rate varies in convexity type throughout the domain of $d(u)$.

Another possibility is to modify the model to include more stages of cancer progression. At this time we only consider one transformation (an activation of an oncogene or an inactivation of a tumor-suppressor gene). According to the theory of multistage carcinogenesis, the natural history of a tumor may consist of many such events. A more complicated, n-step model would have to be implemented to analyze multistage cancer developments [5].

In the present theory we take genetic instability as something imposed exogenously, that is, a control knob that can be turned to raise (or up-regulate) or lower (or down-regulate) the mutation rate. In reality, this is not as simple and one or more molecular steps may be required to increase or reduce the level of genetic instability. A step generating genetic instability has been included in the investigation of [13]. A similar mechanism needs to be included in more sophisticated versions of the current, time-dependent control problem.

Finally, it would be desirable to see how the theoretical findings fits with the actual biological reality. For this, we need more experimental data from detailed measurements which would quantify the level of genetic instability throughout stages of cancer progression. For example, an experiment correlating the rate of chromosome loss (in vitro) from cancer cells at different stages of colon cancer would provide very valuable information which is directly comparable with the outcome of our modeling technique. Also, quantifying the dependence of the cell death rate on the mutation rate would help design a more realistic shape for the function $d(u)$, which, as the present study shows, is an essential component of the optimal control problem.

Appendix: Exact solutions for corner controls

To gain further insight to the issue of switching from one type of control to another, we note the state equations admit an exact solution for the two corner controls. For the upper corner control $u_1(t) = 1$, the two state equations simplify to

$$\frac{dx_1^{(1)}}{dt} = -(u_m + \mu)x_1^{(1)}, \quad \frac{dx_2^{(1)}}{dt} = \frac{u_m + \mu}{\sigma}x_1^{(1)}. \tag{A.1}$$

The following exact solutions for these uncoupled first-order separable state equations are immediate.

LEMMA A.1. *For $u(t) = u_1(t) \equiv 1$ and a general set of initial conditions $x_k^{(1)}(t_s) = x_{ks}^{(1)}$, $k = 1, 2$, the exact solution for the two uncoupled first-order separable state Equation (A.1) is*

$$x_1^{(1)} = x_{1s}^{(1)}e^{-(u_m+\mu)\tau}, \quad x_2^{(1)} = x_{2s}^{(1)} + \frac{x_{1s}^{(1)}}{\sigma}\{1 - e^{-(u_m+\mu)\tau}\}, \tag{A.2}$$

with $\tau = t - t_s$ and a superscript (1) for upper corner control. The normalized mutated cell population $x_2^{(1)}$ is a positive, monotone increasing and strictly concave function of time for $t > 0$.

Remark A.1. The concavity of $x_2^{(1)}(t)$ also follows from

$$\frac{d^2x_2^{(1)}}{dt^2} = \frac{u_m + \mu}{\sigma}\frac{dx_1^{(1)}}{dt} = -\frac{(u_m + \mu)^2}{\sigma}x_1^{(1)} < 0.$$

For the lower corner control $u_0(t) = 0$, the two state equations simplify to

$$\frac{dx_1^{(0)}}{dt} = \left(1 - \mu - x_1^{(0)}\right)x_1^{(0)}, \quad \frac{dx_2^{(0)}}{dt} = \left(a - x_1^{(0)}\right)x_2^{(0)} + \frac{\mu}{\sigma}x_1^{(0)}. \tag{A.3}$$

LEMMA A.2. *For $u(t) = u_0(t) \equiv 0$ and the general initial conditions $x_k^{(0)}(t_s) = x_{ks}^{(0)}$, $k = 1, 2$, the exact solution for the two uncoupled first-order separable state Equation* (A.3) *is*

$$x_1^{(0)} = \frac{x_{1s}^{(0)}(1 - \mu)}{x_{1s}^{(0)} + \left(1 - \mu - x_{1s}^{(0)}\right)e^{-(1-\mu)\tau}},$$

$$x_2^{(0)} = \frac{(1 - \mu)e^{(a+\mu-1)\tau}}{x_{1s}^{(0)} + \left(1 - \mu - x_{1s}^{(0)}\right)e^{-(1-\mu)\tau}}\left\{x_{2s}^{(0)} + x_{1s}^{(0)}\frac{\mu}{\sigma}\frac{1 - e^{-(a+\mu-1)\tau}}{a - 1 + \mu}\right\}, \tag{A.4}$$

with $\tau = t - t_s$ and a superscript (0) for lower corner control. The normalized mutated cell population $x_2^{(0)}$ is a positive, monotone increasing and strictly convex function of time for $t > 0$.

Remark A.2. While the convexity of $x_2^{(0)}$ can be verified by computing $d^2x_2^{(0)}/dt^2$ directly, it follows more simply from

$$\begin{aligned}\frac{d^2x_2^{(0)}}{dt^2} &= \left(a - x_1^{(0)}\right)\frac{dx_2^{(0)}}{dt} + \left(\frac{\mu}{\sigma} - x_2^{(0)}\right)\frac{dx_1^{(0)}}{dt} \\ &= x_2^{(0)}\left[\left(a - x_1^{(0)}\right)^2 - x_1^{(0)}\left(1 - \mu - x_1^{(0)}\right)\right] + \frac{\mu}{\sigma}x_1^{(0)}\left(a + 1 - \mu - x_1^{(0)}\right) \\ &> 0,\end{aligned}$$

given $(a - x_1^{(0)})^2 > 1$ and $x_1^{(0)}(1 - \mu - x_1^{(0)}) < 1$ (because we have $a \geq 2$ and $\mu = O(10^{-1})$.

The exact cell populations corresponding to the two corner controls offer another perspective on why any optimal solution should start with an upper corner control and end on a lower corner control, at least for the $\alpha = 1$ case. Let us compare the corner solutions with the two corner controls applied from the start. In that case, we have $\tau = t$, $x_{1s}^{(1)} = x_{1s}^{(0)} = 1$, and $x_{2s}^{(1)} = x_{2s}^{(0)} = 0$ so that

$$x_1^{(1)} = e^{-(u_m+\mu)\tau}, \qquad x_2^{(1)} = \frac{1}{\sigma}\{1 - e^{-(u_m+\mu)\tau}\}, \tag{A.5}$$

$$x_1^{(0)} = \frac{(1-\mu)}{1 - \mu e^{-(1-\mu)\tau}}, \qquad x_2^{(0)} = \frac{\mu}{\sigma}\frac{1-\mu}{1 - \mu e^{-(1-\mu)\tau}}\frac{e^{(a+\mu-1)\tau} - 1}{a - 1 + \mu}. \tag{A.6}$$

PROPOSITION A.1. *$x_2^{(0)} - x_2^{(1)}$ changes sign exactly once at $t = t_c$ and $dx_2^{(1)}/dt - dx_2^{(0)}/dt$ also changes sign exactly once at $t = t_r < t_c$.*

Proof: We know from the ODE for x_2 in (A.1) and (A.3) that $x_2^{(1)}(0) = x_2^{(0)}(0)$ and $dx_2^{(1)}/dt > dx_2^{(0)}/dt$ (for $t \geq 0$) so that $x_2^{(1)}(t) > x_2^{(0)}(t)$ for a finite time interval $(0, t_c)$. At the same time, because of $d^2x_2^{(1)}/dt^2 < 0$ and $d^2x_2^{(0)}/dt^2 > 0$, the difference between the two first derivatives decreases with time and eventually they become equal at some time $t = t_r$ with $dx_2^{(1)}/dt < dx_2^{(0)}/dt$ for $t \geq t_r$. The reversal of inequality on the two first derivatives implies an intersection of $x_2^{(1)}(t)$ and $x_2^{(0)}(t)$ at a later time t_c. In other words, we have:

$$x_2^{(0)} - x_2^{(1)} \begin{cases} <0 & (t < t_c) \\ =0 & (t = t_c)\,, \\ >0 & (t > t_c) \end{cases} \qquad \frac{dx_2^{(0)}}{dt} - \frac{dx_2^{(1)}}{dt} \begin{cases} <0 & (t < t_r) \\ =0 & (t = t_r)\,. \\ >0 & (t > t_r) \end{cases} \tag{A.7}$$

■

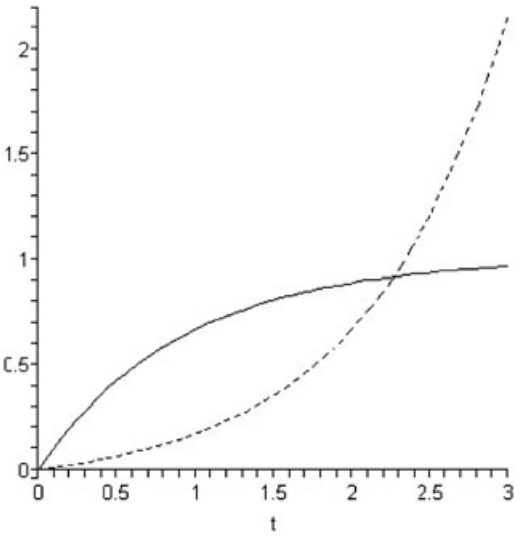

Figure A1. Graphs of $x_2(t)$ for upper (solid) and lower (dash) corner control ($\sigma = 10$, $a = 2$, $\mu = 0.1$, $u_m = 1$).

Both t_c and t_r ($<t_c$) can be determined from the exact solutions. With $\mu \ll u_m \leq 1$, $\sigma \gg 1$, $a \geq 2$, it not difficult to see from the expressions in (A.5) and (A.6)

$$e^{(a-1)t_c} \simeq \frac{a-1}{\mu}, \quad x_1^{(1)}(t_c) \simeq \left(\frac{\mu}{a-1}\right)^{\frac{u_m}{a-1}}, \quad x_2^{(1)}(t_c) \simeq \frac{1}{\sigma}\left[1 - x_1^{(1)}(t_c)\right] \tag{A.8}$$

and

$$e^{(a-1+u_m)t_r} \simeq \frac{u_m}{\mu}, \quad x_1^{(1)}(t_r) \simeq \left(\frac{\mu}{u_m}\right)^{\frac{u_m}{a-1+u_m}}, \quad x_2^{(1)}(t_r) \simeq \frac{1}{\sigma}\left[1 - x_1^{(1)}(t_r)\right]. \tag{A.9}$$

For $a = 2$ and $u_m = 1$, these expressions simplify further to

$$e^{t_c} \simeq \frac{1}{\mu}, \quad x_1^{(1)}(t_c) \simeq \mu, \quad x_2^{(1)}(t_c) \simeq \frac{1}{\sigma}\left[1 - \mu\right] \tag{A.10}$$

and

$$e^{t_r} \simeq \frac{1}{\sqrt{\mu}}, \quad x_1^{(1)}(t_r) \simeq \sqrt{\mu}, \quad x_2^{(1)}(t_r) \simeq \frac{1}{\sigma}\left(1 - \sqrt{\mu}\right). \tag{A.11}$$

For $\mu = 0.1$, we get from (A.10) to (A.11)

$$t_c \simeq 2.302585\ldots, \quad t_r \simeq 1.151292\ldots.$$

Accurate numerical solutions for t_c and t_r using the relevant exact relations in (A.7) give

$$t_c = 2.263757\ldots, \quad t_r = 1.131150\ldots,$$

showing the adequacy of the asymptotic results in (A.10)–(A.11). Graphs of $x_2^{(1)}(t)$ and $x_2^{(0)}(t)$ are given as the solid and dashed curve in Figure A1 for $\sigma = 10$, $a = 2$, $\mu = 0.1$ and $u_m = 1$ to illustrate a typical cross over of these two corner solutions, suggesting that a switch from upper to lower corner control occurring at or prior to t_c.

Acknowledgments

The research of A. Sadovsky (now at NASA Ames Research Center, Moffett Field) was supported by NIH Grants R01GM067247 and NIH National Research Service Award 5 T15 LM007443 from the National Library of Medicine. The research of F.Y.M. Wan was supported by R01GM067247, R01GM075309 and P50-GM076516. The two NIH R01 grants were awarded through the Joint NSF/NIGMS Initiative to Support Research in the Area of Mathematical Biology. The research of N.L. Komarova was supported by the Sloan Foundation Fellowship and NIH Grants R01AI058153, R01CA118545 and R01GN075309.

References

1. B. Alberts, D. Bray, J. Lewis, M. Raff, and K. Roberts, *Molecular Biology of the Cell*, Garland Publishing, Inc., New York and London, 1989.
2. C. Lengauer, K. W. Kinzler, and B. Vogelstein, Genetic instabilities in human cancers, *Nature* 396(6712):643–649 (1998).
3. S. M. Bailey and J. P. Murnane, Telomeres, chromosome instability and cancer, *Nucl. Acids Res.* 34(8):2408–2417 (2006).
4. A. Bardelli, et al., Carcinogen-specific induction of genetic instability, *Proc. Natl. Acad. Sci. U.S.A.* 98(10):5770–5775 (2001).
5. D. Wodarz and N. L. Komarova, Computational biology of cancer: Lecture notes and mathematical modeling, World Scientific, Singapore and Hackensack, NJ, 2005.
6. L. A. Loeb, C. F. Springgate, and N. Battula, Errors in DNA replication as a basis of malignant changes, *Cancer Res.* 34(9):2311–2321 (1974).
7. I. Tomlinson and W. Bodmer, Selection, the mutation rate and cancer: Ensuring that the tail does not wag the dog, *Nat. Med.* 5(1):11–12 (1999).
8. J. Breivik, Don't stop for repairs in a war zone: Darwinian evolution unites genes and environment in cancer development, *Proc. Natl. Acad. Sci. U.S.A.* S98(10):5379–5381 (2001).
9. R. Li, et al., Aneuploidy vs. gene mutation hypothesis of cancer: Recent study claims mutation but is found to support aneuploidy, *Proc. Natl. Acad. Sci. U.S.A.* 97(7):3236–3241 (2000).

10. I. M. Shih, et al., Evidence that genetic instability occurs at an early stage of colorectal tumorigenesis, *Cancer Res.* 61(3):818–822 (2001).
11. J. Marx, Debate surges over the origins of genomic defects in cancer, *Science* 297(5581):544–546 (2002).
12. N. Komarova, Does cancer solve an optimization problem? *Cell Cycle* 3(7):840–844 (2004).
13. N. L. Komarova and D. Wodarz, The optimal rate of chromosome lossfor the inactivation of tumor suppressor genes in cancer, *Proc. Natl. Acad. Sci. U.S.A.* 101(18):7017–7021 (2004).
14. C. Lengauer, K. W. Kinzler, and B. Vogelstein, Genetic instability in colorectal cancers, *Nature* 386(6625):623–627 (1997).
15. K. Chin, et al., In situ analyses of genome instability in breast cancer, *Nat. Genet.* 36(9):984–988 (2004).
16. K. L. Rudolph, et al., Telomere dysfunction and evolution of intestinal carcinoma in mice and humans, *Nat. Genet.* 28(2):155–159 (2001).
17. N. L. Komarova, A. Sadovsky, and F. Y. M. Wan, Selective pressures for and against genetic instability in cancer: A time-dependent problem, *J. Roy. Soc., Inter.* 5:105–121 (2007, 2008) (Online: June 19, 2007, doi:10.1098/rsif.2007.1054).
18. D. P. Cahill, et al., Genetic instability and darwinian selection in tumours, *Trends Cell Biol.* 9(12):57–60 (1999).
19. E. D. Sontag, Some new directions in control theory inspired by systems biology, *Syst Biol (Stevenage)* 1(1):9–18 (2004).
20. S. Lenhart and J. T. Workman, *Optimal Control Applied to Biological Models*, Chapman and Hall/CRC, New York, 2007.
21. G. W. Swan, Role of optimal control theory in cancer chemotherapy, *Math. Biosci.* 101(2):237–284 (1990).
22. D. Kirschner, S. Lenhart, and S. Serbin, Optimal control of the chemotherapy of HIV, *J. Math. Biol.* 35(7):775–792 (1997).
23. L. G. de Pillis, et al., Chemotherapy for tumors: An analysis of the dynamics and a study of quadratic and linear optimal controls, *Math. Biosci*. 209(1):292–315 (2007).
24. N. L. Komarova, et al., Dynamics of genetic instability in sporadic and familial colorectal cancer, *Cancer Biol. Ther.* 1(6):685–692 (2002).
25. Y. Iwasa and S. A. Levin, The timing of life history events, *J. Theor. Biol.* 172:33–42 (1995).
26. M. A. Nowak, et al., The role of chromosomal instability in tumor initiation, *Proc. Natl. Acad. Sci. U.S.A.* 99(25):16226–16231 (2002).
27. N. L. Komarova, A. Sengupta, and M. A. Nowak, Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability, *J Theor. Biol.* 223(4):433–450 (2003).
28. F. Y. M. Wan, *Inroduction to the Calculus of Variations and Its Applications*, Chapman and Hall, New York, 1995.
29. V. Boltyanskii, et al., *The Mathematical Theory of Optimal Processes*, John Wiley and Sons, New York, 1962.
30. A. Bryson and Y. C. Ho, *Applied Optimal Control*, Ginn and Company, Waltham, MA, 1969.
31. I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Prentice Hall, Englewood Cliffs, 1963.
32. L. S. Pontryagin, et al., *The Mathematical Theory of Optimal Control Processes*, Interscience Publishers, New York, 1962.

33. E. Samper, J. M. Flores, and M. A. Blasco, Restoration of telomerase activity rescues chromosomal instability and premature aging in Terc$^{-/-}$mice with short telomeres, *EMBO Rep.* 2(9):800–807 (2001).
34. S. E. Artandi and R. A. DePinho, A critical role for telomeres in suppressing and facilitating carcinogenesis, *Curr. Opin. Genet. Dev.* 10(1):39–46 (2000).
35. E. Gonzalez-Suarez, E. Samper, J. M. Flores, and M. A. Blasco, Telomerase-deficient mice with short telomeres are resistant to skin tumorigenesis, *Nat. Genet.* 26(1):114–117 (2000).
36. M. Eigen and P. Schuster, *The Hypercycle: A Principle of Natural Self-Organization*, Springer, Berlin, 1979.
37. K. Kinzler and B. Vogelstein, *The Genetic Basis of Human Cancer* (2nd ed.), The McGraw-Hill Companies, Inc., New York, 2002.

University of California

(Received October 8, 2009)