Supervisor
Professor Ville Mustonen, University of Helsinki

I would also like to acknowledge Professor Dr. Niko Beerenwinkel and Dr. Robert Noble for their helpful comments.
“Nothing in biology makes sense except in the light of evolution.”
Theodosius Dobzhansky
Abstract

Cancer is a dynamic and complex microevolutionary process. All attempts of curing cancer thus rely on successfully controlling also the evolving future cancer cell population. Since the emergence of drug resistance severely limits the success of many anti-cancer therapies, especially in the case of the promising targeted therapies, we need urgently better ways of controlling cancer evolution with our treatments to avoid resistance.

This thesis characterizes acquired drug resistance as an evolutionary rescue and uses optimal control theory to critically investigate the rationale of aggressive maximum tolerated dose (MTD) therapies that represent the standard of care for first line treatment. Unlike the previous models of drug resistance, which mainly concentrate on minimizing the tumor volume, herein the optimal control problem is reformulated to explicitly minimize the probability of evolutionary rescue, or equivalently, maximizing the extinction probability of the cancer cells.

Furthermore, I investigate the effects of drug-induced resistance, where the rate of gaining new resistant cells increases with the dose due to increased genome-wide mutation rate and non-genetic adaptations (such as epigenetic regulation and phenotypic plasticity). This approach not only reflects the biological realism, but also allows to model the cost of control in a quantifiable manner instead of using some ambiguous and incomparable penalty parameter for the cost of treatment.

The major finding presented in this thesis is that MTD-style therapies may actually increase the likelihood of an evolutionary rescue even when only modest drug-induced effects are present. This suggests that significant improvements to treatment outcomes may be accomplished at least in some cases by treatment optimization. The resistance promoting properties of different anti-cancer therapies should therefore be properly investigated in experimental and clinical settings.

Keywords: drug-induced resistance, optimal control, treatment optimization, evolutionary rescue, Pontryagin’s minimum principle, Hamilton-Jacobi-Bellman equation, maximum tolerated dose (MTD)
Contents

1 Introduction 1

2 Theory of Optimal Control 6
  2.1 Pontryagin's Minimum Principle . . . . . . . . . . . . . . . . . . . . . . . . 7
  2.2 Simple Example with Quadratic Cost of Control . . . . . . . . . . . . 10
  2.3 Linear Cost of Control . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
  2.4 Isoperimetric Constraints . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
  2.5 Alternative Approach Using the Hamilton-Jacobi-Bellman Equa-
    tion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14

3 Drug Resistance as an Optimal Control Problem 16
  3.1 Evolutionary Rescue . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
  3.2 Drug-Induced Resistance as Cost of Control . . . . . . . . . . . . . . 19
  3.3 Problem 1: Minimize the Probability of an Evolutionary Rescue 20
  3.4 Time-Independent Control Law . . . . . . . . . . . . . . . . . . . . . . . . 28
  3.5 Problem 2: Discount for the Growth . . . . . . . . . . . . . . . . . . . . . 30
  3.6 Towards Evolutionary Rescue Experiments . . . . . . . . . . . . . . . 31

4 Discussion 35
1 Introduction

Cancer is a dynamic and complex microevolutionary process in which cells within the healthy tissue start to proliferate abnormally, escape from control and step by step acquire further neoplastic traits in their evolution towards a metastatic cancer ecosystem. During this progression the neoplastic cells must successfully pass through various selection barriers that have evolved as countermeasures against the threat that cancer imposes to the multicellular life. The first driver mutation, which initiates a benign neoplasm, essentially increases the cells’ birth rate above that of the healthy cells in equilibrium and thus grants them a selective advantage for an example by removing the need of external proliferative signalling. However, this first growth epoch will soon reach its end as the cells’ death rate will also consequently increase by the activation of apoptosis, susceptibility to higher immune predation and cellular senescence. Thus, in order to grow further, the neoplastic cells must gain new driver mutations that will deregulate these various countermeasures.

Even if the cells manage to sustain growth, they will soon reach the limits of their proliferative unit and other natural boundaries created by the tissue hierarchy. The ability to further invade the local tissue is a key malignant feature that initiates the series of events that lead to the development of a highly lethal disease, cancer [1]. Consequently, the development of cancer is a multi-step process where benign neoplasms acquire more and more malignant features as they pass through the various hallmarks of cancer [2].

Cancer is conventionally seen as a genetic disease and characterized by the accumulation of genetic and epigenetic alterations. The major focus of cancer research has thus been in the systematic study of mutations and genes that are responsible for the tumorigenesis and shed light on the pathogenesis of cancer. In the era of personalized medicine, there are exceedingly high hopes of utilizing the patient’s precise genomic data for better patient stratification, individualized treatment and the development of new targeted therapies that function by specifically targeting certain actionable molecular aberrations. Classical example of such targeted therapies include imatinib and other tyrosine-kinase inhibitors that revolutionized the treatment of chronic myeloid leukemia patients with Bcr-Abl mutations. Since then many more similar targeted therapies have been developed and put to clinical practice. PARP-inhibitors provide yet another promising example of how genetic biomarkers can improve treatment outcomes.

Despite these remarkable advances, the overall survival rates for many cancers have not significantly improved. The major obstacle that continues to halt the progress is drug resistance, a phenomenon where the drug loses its desired pharmacodynamical effects. While majority of the patients respond
initially well to the anti-cancer therapies and the tumor burden is efficiently reduced, many experience a relapse where drug resistant cells reoccupy the tumor niche. Even though the precise molecular mechanisms responsible for drug resistance are manifold and vary depending on the drugs used, all anti-cancer therapies are susceptible of developing resistance.

Traditional chemotherapies, such as alkylating and platinum-based agents, target unselectively all rapidly dividing cells (including healthy cells) and function by inducing severe DNA-damage, such as crosslinking, which then triggers apoptosis. Cancer cells can get around it by enhancing DNA repair mechanisms and impairing the cell cycle arrest and apoptosis mechanisms. TP53 is a crucial tumor suppressor gene that is found to be impaired in majority of cancers, highlighting the importance apoptotic mechanisms in preventing cancer. Antimetabolites which inhibit DNA synthesis can be bypassed by overexpressing genes responsible for DNA-synthesis. More targeted drugs, such as receptor and tyrosine kinase inhibitors, which halt cell proliferation can be circumvented by upregulating another pathways or by simply altering the target molecules. This makes the highly rejoiced targeted therapies usually even more susceptible to resistance. In the well-known example of imatinib a single point mutation is sufficient to prevent the competitive binding to the kinase. More general mechanisms which contribute to multidrug resistance include enhanced membrane transport using efflux pumps, autophagy and enzymatic deactivation.

Even though we have adequate biological understanding on the various different mechanisms that contribute to drug resistance, the fundamental questions of how and when, exactly, does drug resistance emerge remain largely unclear. For an example, it is very difficult or even impossible to distinguish whether the relapse resulted from the growth of resistant cells that were present already prior the treatment (intrinsic resistance) or from resistant cells that emerged only after the treatment was initiated (acquired resistance). In both cases cells with reduced sensitivity to the drug are strongly favoured by natural selection and thus can easily invade the population.

Classically, the emergence of drug resistance is explained by randomly occurring mutations that take place during the treatment period [3]. This somatic mutation theory of drug resistance dates back to the famous Luria-Delbrück experiments [4], where resistance to phages were studied in bacteria. The authors hypothesized that if resistance was induced by the media, each tested plate should contain roughly the same number of resistant mutants. However, since there were highly variable number of mutants at each tested plate, the authors concluded that the resistance to the virus must have a genetic origin.

The recent advances and experimental findings have, however, compli-
cated the study of drug resistance. Firstly, the possibility of intrinsic resistance is starting to look highly likely given the tremendous heterogeneity observed in many advanced cancers [5]. Secondly, cancer cells can activate certain resistance mechanisms even without a need for a mutation using epigenetic regulation in a post-adaptive manner. Cancer cells, especially stem-like cancer cells, can exhibit diverse phenotypic plasticity, where they can dynamically respond to the changes in their environment [6]. Essentially, cancer cells can enhance their survival by allocating more resources normally spent for proliferation as a form of life-history trade-off. Finally, there is ample experimental evidence suggesting that many anti-cancer therapies themselves may in fact induce drug resistance by increasing the genome-wide mutation rate and non-genetic adaptations in some dose-dependent fashion [7–9]. Studies in bacteria demonstrate that stress alone can be sufficient to activate a higher mutation rate as a form of adaptation to changing environment [10]. Recently, similar findings were found also in cancer [7].

Figure 1: Relapse can be explained by three distinct models of drug resistance all of which can occur simultaneously in a single clinical manifestation. In all cases, resistance is the consequence of strong selection pressure created by the harsh treatment. In intrinsic resistance model (1) resistant cells coexisted in the cancer population prior the treatment. In acquired resistance models (2) and (3) de novo resistance emerged during the treatment period either spontaneously by random mutations (2) and/or by drug-induced mutations (3). The crucial difference between the models (2) and (3) is that in the case of drug-induced resistance (3) the probability of gaining new resistance mutations follows some dose-dependency while in (2) this probability does not depend on the dosage used.
All attempts to cure cancer rely on controlling the evolution of cancer population. Unlike other genetic diseases, cancer is not a static abnormality which can be exhaustively classified and understood solely by the means of genetics. The crucial roles of the tumor microenvironment and cancer ecology have only very recently gained more attention (see e.g. [11,12]). For example, metastatic cancers have similar genetic characteristics as non-metastatic cancers [13] suggesting that metastasis is fundamentally an ecological rather than genetic phenomenon (as dispersal is in general).

By viewing cancer as an evolutionary process, we are forced to think not only the present but also the future cancer population. At the same time we understand that the evolution of cancer is not a completely random and independent process (like mutations), but we instead have the power to influence and control it by inducing different selection pressures with our treatments. These considerations then raise the intriguing questions of whether the evolutionary outcomes of individual tumors are predictable and what would be the best way of controlling them [14]. After all, all treatments are attempts to control this evolutionary process and so it may prove to be highly useful to approach these questions from the optimal control point of view. This gives rise to the mathematical study of cancer evolution and treatment optimization avenues.

When a new drug has been developed, it is thoroughly tested in clinical trials and the dosage is increased until the side-effects become intolerable and the so-called maximum tolerated dose (MTD) is reached. Virtually all approved cancer therapies are then to be administered either at a higher doses, with drug holidays in between, or more densely using lower doses (metronomic therapy) [15]. Both regimens lead to approximately the same cumulative dosage which is close to the MTD. The usage of aggressive MTD-style therapies have rooted deeply to oncology because the chances of a cure are thought to be maximized with such a regimen.

There exists a wealth of literature on using mathematical modelling in the context of cancer and drug resistance (for reviews see [16,17]). Majority of the previous optimal control models concentrate on minimizing the tumor volume while penalizing the use of control to account for the side-effects of treatment [18,19]. Typically, such analyses lead to the MTD-paradigm with metronomic therapies usually performing better. Other theoretical studies have criticized the MTD-paradigm directly often based on evolutionary theory: while effective at eradicating the sensitive cells from the tumor population, such aggressive therapies clear the space for the emerged resistant cells which in turn thrive in the new “pristine” environment. The elimination of the competing sensitive cells, which often have a higher fitness in the absence of treatment, only help the resistant cells grow faster leading
to a phenomenon termed competitive release [20]. Indeed, it is well-known from cancer biology that cancer cells actively engage in competition for resources, such as oxygen, nutrients, growth factors and space, in the tumor microenvironment.

This has then motivated various authors to suggest that containment strategies, which aim to use the minimal amount of control to keep the tumor burden in check, could be used to competitively suppress the emerged resistant cells [21–23]. Most notably, Gatenby et al. have proposed in [21] the so-called adaptive therapy which represents the most radical departure from the MTD-paradigm.

This thesis contributes to the field of mathematical oncology by characterizing acquired drug resistance as an evolutionary rescue and using optimal control theory to investigate the rationale of maximum tolerated dose (MTD) therapies that represent the standard of care for first line treatment. First, some general theory of optimal control is introduced in Chapter 2. Then Chapter 3 presents the concept of an evolutionary rescue and the problem of drug resistance is reformulated as an optimal control problem using a novel objective function: instead of focusing on minimizing the tumor volume we are looking for treatment strategies which explicitly minimize the probability of an evolutionary rescue, or equivalently, maximize the extinction probability of the cancer cells. Furthermore, the effects of drug-induced resistance are investigated while paying special attention on how the cost of control is modelled. Finally, Chapter 4 discusses the optimality of the MTD-style therapies, the relationship of elimination versus containment strategies and provides an outlook on how evolutionary processes in general can be predicted and controlled.
2 Theory of Optimal Control

Consider a system of well-defined and measurable state variables which evolve in time according to some transition rule, which describes a function from the current state to the future state. A characteristic property of such dynamical systems is that they can exhibit highly complex behavior even when the transition rules are very simple due to the coupled state variables which create non-linear feedback loops.

Optimal control theory deals with dynamical systems, where there exists at least one variable, the control variable, whose value can be chosen freely by the controller from the set of admissible controls. The controller must thus have some power over the evolution of the system and the question of interest is then to ask what is the best way to use this power. Each control strategy determines the corresponding evolution of the rest of the state variables and one such realization is referred to as trajectory (the term originates from the control of spacecrafts). Since we are interested in finding out the best control strategy, which yields the trajectory with the best outcome, it is essential to have a criterion by which to evaluate and compare the different trajectories. Well-defined optimal control problems thus usually satisfy the following general properties:

1.) the system’s evolution can be accurately modelled (the transition rule is known);

2.) there exists a control variable that can be chosen by the controller (from the admissible set of controls);

3.) the evolution of the state variables depend somehow on the control variable (controllability);

4.) each control strategy determines a single trajectory;

5.) each trajectory has a unique cost and these costs can be ordered;

Here we are interested in applying the theory of optimal control to biology. Thus, it needs some justification that these conditions are indeed satisfied. Suppose a harmful population, such as pathogen or pest, is spreading and we have a drug that we can use to control the growth of the population. How can we apply optimal control theory to manage the population most efficiently? Condition 1 requires that the population dynamics can be modelled. This means that we need to know how the population is growing both in the
absence of the drug and when the drug is used. Condition 2 states that we are free to choose the dose with respect to some constraints and condition 3 requires that the drug must be effective. Conditions 4 and 5 require that we can monitor the population size and quantify the costs related to the population size and secondly the potential costs that resulted from the usage of the drug. In the case of a pest population we can for an example assign an economic cost for the lost crop (which is proportional to the population size) and the pesticide used.

For an example, consider a harmful population that grows according to dynamics
\[
\dot{x}(t) := \frac{dx(t)}{dt} = f(x, u); \quad x(0) = x_0,
\]
where \(x\) and \(u\) are both functions of time and denote the population size and the control variable, respectively. Assume that we have a fixed time period \([0, T]\) when we can track the population size and apply control. The standard way to quantify the costs associated with the population size is using a very general cost functional of the form
\[
C := \Phi(x(T)) + \int_0^T L(t, x(t), u(t))dt,
\]
where \(\Phi\) is an increasing function and represents the end-cost of the population burden at the end of the control period. \(L\) is the Lagrangian term which takes into account the cumulative costs associated with population size and the possible side-effects of applying the control during the entire control period.

The problem discussed here concentrates on finding the optimal control strategy \(u^*(t)\) which minimizes the cost functional above. The minimization is subject to the constraint that the optimal control strategy \(u^*(t)\) and the associated optimally controlled trajectory \(x^*(t)\) are compatible with the dynamics and that the optimal control is admissible,
\[
u(t) \in [u_{\min}, u_{\max}] =: \mathcal{U} \text{ for all } t \in T.
\]

We next present a well-known Pontryagin’s Minimum Principle [24], which provides a necessary condition for the optimal control strategy [25].

2.1 Pontryagin’s Minimum Principle

Theorem 2.1. (Pontryagin) Consider the problem of minimizing the cost functional
\[
C(u) = \Phi(x(T)) + \int_0^T L(t, x(t), u(t))dt,
\]
(2.1)
where \( x = (x_1, ..., x_n) \) is the state variable, \( u = (u_1, ..., u_m) \) is the control variable, \( \Phi : \mathbb{R}^n \rightarrow \mathbb{R} \) and \( L : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R} \) are continuously differentiable functions and the terminal time \( T \) is fixed. The minimization must be done subject to dynamics
\[
\dot{x} = f(t, x(t), u(t)); \quad x(0) = x_0, \tag{2.2}
\]
where \( f = (f_1, ..., f_n) : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \) is also continuously differentiable with respect to all its component functions. Suppose that \( u^* \) is a (local) minimizer of the problem and let \( x^* \) denote the corresponding optimal trajectory. Then, there exists a costate (multiplier) variable \( \lambda^* = (\lambda_1, ..., \lambda_n) \) such that the triplet \( (u^*, x^*, \lambda^*) \) minimizes the Hamiltonian, which is defined as
\[
H(t, x, u, \lambda) := L(t, x, u) + \lambda^T f(t, x, u), \tag{2.3}
\]
so that
\[
H(t, x^*, u^*, \lambda^*) \leq H(t, x, u, \lambda) \tag{2.4}
\]
for all \( t \in [0, T] \) and all admissible controls \( u \in U \). Furthermore the triplet \( (u^*, x^*, \lambda^*) \) must satisfy the following system of equations:
\[
\begin{align*}
H_u &= L_u + \lambda^T f_u = 0 \tag{2.5} \\
-H_x &= -(L_x + \lambda^T f_x) = \dot{\lambda} \tag{2.6} \\
H_\lambda &= f(x, u) = \dot{x} \tag{2.7}
\end{align*}
\]
Additionally, if the final state of the state vector is not fixed (for some component of \( x \)), then the corresponding multipliers must satisfy the following boundary condition:
\[
\Phi_{x_i}(x_i(T)) = \lambda_i(T) \tag{2.8}
\]
for every component \( i \) for which the corresponding state variable \( x_i \) has a free end-state \( x_i(T) \).

Pontryagin’s Minimum Principle is based on the variational approach and provides a necessary condition for the optimal control strategy. Essentially the first step is to transform the constrained optimization problem to ordinary unconstrained optimization problem by adjoining the constraint with the Lagrange multiplier function which leads to the definition of the Hamiltonian. Then, one computes the independent variations of the state, control and costate (multiplier) variables and sets the total variation of the cost functional to zero (as usual). The stationary solution is obtained when the coefficients of the individual variations are zero, which yields the equations (2.5)-(2.7).

The optimal control strategy \( u^* \) can in some be cases solved from eq. (2.5) as a function of the state and costate variables. The optimality for
the obtained candidate solution $u^*(t, x, \lambda)$ must then be verified using some second order condition. The Legendre-Clebsch condition [19] is often the most convenient way:

\[
\frac{\partial^2 H}{\partial u^2} > 0 \text{ for a minimum.} \tag{2.9}
\]

Once the optimality has been verified, the optimal control strategy must then be solved together with state and costate variables using the dynamics (2.6)-(2.7) and the boundary conditions (2.2) and (2.8). Solving these coupled ordinary differential equations is in principle simple, but the handling of the mixed boundary conditions requires somewhat sophisticated numerical methods. The following algorithm, known as **Forward-Backward-Sweep Method** [26,27], provides an intuitive and efficient way to solve the optimal control problem:

**Forward-Backward Sweep Method**

1.) Using an educated guess for $u^*$, solve $x$ forward in time using equation (2.7) and the initial condition (2.2).

2.) Then, using the guess for $u^*$ and the updated $x$, solve $\lambda$ backwards in time using equation (2.6) and boundary condition (2.8).

3.) Now update $u^*(x, \lambda)$ for the next iteration using the updated $x$ and $\lambda$ and equation (2.5).

4.) Iterate until convergence is achieved.

Finally, the admissibility of the control must be satisfied. The most simple type of constraint limits the rate of control admissible to some bounded interval

\[ u_{\text{min}} \leq u(t) \leq u_{\text{max}}, \forall t \in [0, T]. \]

These constraints can be implemented simply by defining

\[ u^*(t) = \min(u_{\text{max}}, \max(u_{\text{min}}, u(t))) \tag{2.10} \]

as if the optimal control would like to use more or less control than what is admissible then the best possible thing to do is to use as much or as little control as possible; in the interior these constraints have no effect. We refer to the extreme controls as *bang-bang* and the intermediate controls
as singular. More complicated constraints, such as isoperimetric constraints, which limit only the cumulative control used during the entire control period, are discussed in section 2.4.

2.2 Simple Example with Quadratic Cost of Control

Consider a harmful population which grows according to the logistic model with constant per capita growth rate $r > 0$ and scaled carrying capacity $K = 1$. Suppose we can increase the net death rate by applying some amount of control $u > 0$ which leads to population dynamics

$$\dot{x} = (r - u)x(1 - x); \quad x(0) = x_0. \quad (2.11)$$

Furthermore, assume that we are interested in minimizing the following cost functional:

$$\min_{u \in U} x(T) + \int_0^T (x(t) + Qu^2(t))dt, \quad (2.12)$$

where the admissible set of controls is defined as $U = [0, u_{\text{max}}]$. The parameter $Q > 0$ determines the relative cost of control compared to the cost related to the population. The Hamiltonian for this system is given by

$$H = x + Qu^2 + \lambda(r - u)x(1 - x).$$

Thanks to the quadratic cost of control in the Lagrangian term we can use equation (2.5) to obtain the candidate optimal control strategy:

$$u^*(t, x, \lambda) = \frac{\lambda(t)x(t)(1 - x(t))}{2Q}.$$ 

Note that since

$$\frac{\partial^2 H}{\partial u^2} = 2Q > 0$$

this is indeed a minimizing solution by the Legendre-Clebsch condition. The dynamics of $\lambda$ is governed by equation (2.6) which yields

$$\dot{\lambda} = -(1 + \lambda(r - u)(1 - 2x)) = \lambda(u - r)(1 - 2x) - 1.$$ 

Furthermore, since the end state $x(T)$ is not fixed, equation (2.8) provides a boundary condition

$$\lambda(T) = \Phi_x(x(T)) = 1,$$

which allows us to solve the multiplier function backwards in time.
Figure 2: Numerical solution using Forward-Backward Sweep Method to the optimal control problem (2.12) with parameter values $T = 1$, $r = 0.5$, $Q = 0.01$, $u_{max} = 5$ and $x_0 = 0.3$.

Figure 3: The optimal control strategy is highly sensitive to the cost of control parameter $Q$. As $Q \to 0$ the optimal control reduces to bang bang solution (with no switches).
2.3 Linear Cost of Control

In some cases the control variable $u$ vanishes from equation (2.5) upon differ-entiation, for an example when all the control dependencies are linear. This leads to a degenerate case, where the Forward-Backward Sweep Method cannot be used directly as we don’t have an expression for $u^*(t, x, \lambda)$ which to pass to the algorithm. In this case it is convenient to define a switching function $\Psi(t) := H_u$ whose sign determines the optimal bang-bang controls as follows:

$$u^*(t) = \begin{cases} 
0, & \text{if } \Psi(t) > 0 \\
u_{\text{max}}, & \text{if } \Psi(t) < 0 \\
\text{singular,} & \text{if } \Psi(t) = 0.
\end{cases} \quad (2.13)$$

The possibility of singular controls must, however, be excluded as it is possible that the switching function vanishes on some interval $I \subset [0, T]$. If this occurs, the optimal control must be synthesized from the singular and bang-bang controls, which in general can be very tricky.

Assume that we have such an interval $I = \{t \in [0, T] : \Psi(t) = 0\}$. Then, necessarily all its time-derivatives must also vanish and we get condition

$$\frac{d^k \Psi(t)}{dt^k} = 0 \text{ for } t \in I \text{ and } k = 1, 2, \ldots \quad (2.14)$$

By differentiating the switching function with respect to time, we can force the control variable to appear. Due to interesting Lie-algebraic properties it can be shown [19] that the control variable can only appear after even number of differentiations. After this, we can obtain the candidate optimal singular control $u^*(t, x, \lambda)$ which to pass to the Forward-Backward Sweep Method. However, the optimality of the singular control must first be verified as the singular control can also be the worst (maximizing) solution. The generalized Legendre-Clebsch condition takes the form

$$(-1)^k \frac{\partial}{\partial u} \frac{d^k \partial H}{dt^kd\partial u} \geq 0 \text{ for a minimum.} \quad (2.15)$$

2.4 Isoperimetric Constraints

It is often the case that the amount of control available to use during the control period is restricted, for an example due to limited supply, economic considerations or the intolerable side-effects of the control. Taking these types of constraints account in the optimal control problems lead to the following isoperimetric constraint:

$$\int_0^T u(t) dt = D_{\text{max}}. \quad (2.16)$$
Note that imposing such a constraint may change the optimal control profile qualitatively, unlike the rate limiting constraints discussed in the previous section, because the question of when is it best to use the control rises.

The isoperimetric constraint can be handled by introducing a second state variable \[ y(t) := \int_0^t u(t)dt, \] (2.17) which represents the cumulative control used at time \( t \). The dynamics of \( y \) is simply \( \dot{y} = u \) with boundary conditions \( y(0) = 0 \) and \( y(T) = U_{max} \). The Hamiltonian for the problem (2.12) reads

\[ H = x + Qu^2 + \lambda_1 (r - u)x(1 - x) + \lambda_2 u. \]

Differentiating the Hamiltonian with respect to \( u \) and setting it zero yields the optimal control solution

\[ u^* = \frac{\lambda_1 x(1 - x) - \lambda_2}{2Q}, \]

where the multipliers must satisfy the dynamics

\[ \dot{\lambda}_1 = -H_x = -(1 + \lambda_1 (r - u)(1 - 2x)) = \lambda_1 (u - r)(1 - 2x) - 1, \quad \lambda_1(T) = 1 \]

and

\[ \dot{\lambda}_2 = -H_y = 0. \]

Note that the end-state of the second state variable is not free but fixed \( y(T) = D_{max} \) as we impose the isoperimetric constraint. Thus, we have no boundary condition for the second multiplier. However, since the second state variable does not appear in the Hamiltonian, its derivative is zero and the second multiplier is just a constant unknown scalar.

Solving the optimal control problem numerically with the isoperimetric constraint requires the implementation of some sort of shooting method together with the Forward-Backward Sweep Method where the value of the unknown multiplier is guessed and adjusted until the total cumulative control used is approximately equal to the constraint \( D_{max} \). This gives equation

\[ \int_0^T u^*(t)dt = \int_0^T \frac{\lambda_1 x(1 - x)}{2Q}dt - \int_0^T \frac{\lambda_2}{2Q}dt = D_{opt} - \frac{\lambda_2 T}{2Q} = D_{max}, \]

where \( D_{opt} \) is the optimal amount of control to use under no constraints. From this we can easily derive an expression for \( \lambda_2^* \):

\[ \lambda_2^* = \frac{2Q}{T}(D_{opt} - D_{max}). \]
If the constraint stops us from using as much control as would be optimal, which is the case when $D_{opt} > D_{max}$, then the multiplier is a positive scalar which decreases the rate of control enforcing the constraint. In the case $D_{opt} < D_{max}$ the constraint forces us to use more control than what would be optimal and thus the multiplier will be a negative number which increases the rate of control until the constraint is met.

This result shows that the multipliers are not just mere dummy-variables, but their have their own interesting interpretation as a measure of the sensitivity of the cost functional to the perturbations in the respective state variables [25]. The second multiplier tells us how the cost of the optimally controlled trajectory would change if we would change the value of $D_{max}$, that is, loosen the isoperimetric constraint. Consequently, $\lambda_1^*$ measures the sensitivity of the cost functional to tiny perturbations in the first state variable, the population size $x$. In this case, the first multiplier is always positive as increasing the population size will always increase the cost and decreasing the population size will decrease the cost.

As a special case, consider the interpretation of $\lambda_1^*(T)$, which now tells how the optimal cost changes if the trajectory would be perturbed at the last time point. The perturbation will not affect the cumulative cost as the value of the integral will not change and the only change in cost results directly from the end-cost. Thus, the derivative of the end-cost function $\Phi(x)$ will give the first order information of how the optimal cost changes, exactly confirming the natural boundary conditions set by the Pontryagin’s minimum principle in equation (2.8).

2.5 Alternative Approach Using the Hamilton-Jacobi-Bellman Equation

Richard Bellman was the first to notice that the optimal control has the following property: “an optimal policy has the property that, whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision”. This property is known as Bellman’s principle of optimality [29] and it can be stated more precisely as follows: if we have an optimal solution $u^* : [0, T] \rightarrow \mathcal{U}$ for the entire control period, then the restriction $u^* : [t_0, T] \rightarrow \mathcal{U}$ is itself the optimal solution for the corresponding restricted optimal control problem with initial condition $x(t_0) = x^*(t_0)$ for any $t_0 \in [0, T]$. While sounding innocent at first, this actually provides a very efficient way to solve optimal control problems as the hard optimization through the sequence of controls can be reduced to subsequent simple scalar optimizations which together
constitute the optimal control strategy due to the Markovian nature of the problem [30].

This is accomplished by introducing the optimal cost-to-go $J(x,t)$, which gives the optimal cost of the optimally controlled trajectory which starts from $x$ at time $t$. The optimal cost $V(x_0)$ thus equals to $J(x_0,0)$. The evolution of the cost-to-go function can be solved using the Hamilton-Jacobi-Bellman equation

$$-\partial_t J(x,t) = \min_{u \in U} \left( L(t,x,u) + f(t,x,u)\partial_x J(x,t) \right)$$  \hspace{1cm} (2.18)$$

which must be solved with the boundary condition (end-cost)

$$J(x,T) = \Phi(x).$$  \hspace{1cm} (2.19)$$

By recording the minimizing control at each point, the HJB-approach gives a control map

$$u^*(x,t) = \arg \min_u \left( L(t,x,u) + f(t,x,u)\partial_x J(x,t) \right)$$  \hspace{1cm} (2.20)$$

from which the optimal control can be read for any conceivable state. Whereas the Pontryagin’s minimum principle can be used to calculate the optimal control and trajectory for any given initial condition, the HJB-approach essentially solves the problem for all possible states simultaneously. This of course comes at a higher computational cost. Visualizing the control map may nevertheless give interesting insights to the problem. For an example, the control map may pinpoint the existence of a stationary profile which can be used to obtain a closed loop control which acts as direct feedback of the state variable instead of being an explicit function of time.

Additional advantage of the HJB-approach is that it can incorporate stochastic population dynamics easily by extending the Hamilton-Jacobi-Bellman equation with a second-order term:

$$-\partial_t J(x,t) = \min_{u \in U} \left( L(t,x,u) + f(t,x,u)\partial_x J(x,t) + \frac{1}{2} \nu(t,x,u)\partial_x^2 J(x,t) \right)$$  \hspace{1cm} (2.21)$$

where $x$ is now a stochastic process satisfying $\langle dx \rangle = f(t,x,u)dt$ and $\langle dx^2 \rangle = \nu(t,x,u)dt$; the deterministic dynamics gives the expected mean-field model while the term $\nu(t,x,u)$ gives the variance of the fluctuating trajectories around the expected (deterministic) trajectory. These stochastic fluctuations may play an important role in the optimal control problem as the optimal control may anticipate and take advantage of the stochasticity. Taking $\nu \to 0$ leads back to the deterministic case.
3 Drug Resistance as an Optimal Control Problem

We have now the necessary tools to move on to the problem of drug resistance. Ledzewicz and Schättler among others have considered various optimal control problems on the subject using similar quadratic and linear cost functionals (see e.g. [18, 19]). These control structures specifically aim at minimizing the tumor volume as efficiently as possible while penalizing the use of control with some penalty parameter, which then trades-off the benefits of a lower tumor burden to the side-effects of the treatment.

The fundamental problem with such formulations, however, lies precisely in the discrepancy of the assumed benefits of the lower tumor burden and the ambiguously defined cost of control. It is evident that anti-cancer therapies, chemotherapies especially, have severe side-effects that greatly impact the patient’s quality of life. Younger patients who have undergone these treatments have also a considerable risk of developing a second cancer partly due to genotoxic side-effects which leave distinct mutational signatures [9]. Many cancer therapies also have a significant economic cost. Neglecting these costs altogether would of course not depict reality, but how can these costs be compared to the cost of the tumor burden in any meaningful way?

Firstly, what good does a temporary reduction in the tumor burden achieve if the tumor returns and is even more difficult to treat? Is this indeed a suitable and desirable objective for the treatment? Secondly, even if we agree that shrinking the tumor is at all situations a desirable outcome of the treatment, how is this benefit comparable to the assumed cost? In which units are the costs and benefits measured and, more specifically, can they even have a same unit? Besides the economic costs of the treatment, how can we actually even quantify the side-effects and toxicity experienced by the patient? From this point of view, all such formulations seem rather ambiguous and are based on desirable mathematical properties (such as convexity) rather than actually measurable and quantifiable parameters and costs. This is especially troubling given the high sensitivity of the optimal control to the penalty parameter (see figure 3).

Given the widespread possibility of drug resistance, it is perhaps time to rethink the objectives of the treatment and shift the focus from simply minimizing the tumor size. More specifically, we need to ask, why do we want to reduce the tumor size as quickly as possible in the first place? In fact, it is rarely the case that the massive tumor burden on its own kills the patient. Rather, it is the ability to metastasize and develop resistance and the consequent future growth potential which make cancer such a lethal
disease. Thus, we should ideally be designing therapies that take into account the evolvability of the cancer and then, given these evolutionary potentials, do the best that can be done.

The approach taken in this thesis utilizes the concept of evolutionary rescue to characterize acquired drug resistance and then reformulates the objective function to explicitly maximize the extinction probability of the cancer cells. Since all the treatments are essentially designed to be extinction events, such an objective function incorporates the true objectives of treatment while also specifically acknowledging that emergence of resistance denotes treatment failure.

3.1 Evolutionary Rescue

Consider a population that is experiencing a dramatic change in its environment and is heading towards extinction (for an example due to climate change). The key question in such situations is whether adaptive evolution can happen fast enough to save the population from going extinct. If the population is saved, we say that an evolutionary rescue has occurred [31]. Introduced first in the field of conservation biology, where the objective has been to design the most efficient intervention programs to save endangered species from extinction, the concept of evolutionary rescue can be readily applied also to the study of drug resistance [32] with the opposite goal in mind.

Evolutionary rescue can occur either by standing variation or by de novo mutation. Rescue by standing variation corresponds to the intrinsic resistance model where the population is readily diverse enough to contain individuals who can survive also in the changed environment. This is one of the major reasons why mechanisms which promote diversity are preserved and explains why heterogeneous populations carry a lower extinction risk. This is also the reason why higher degrees of intratumor heterogeneity are usually associated with a poorer prognosis. Rescue by de novo mutation, on the other hand, corresponds to the acquired resistance model where new, better adapted individuals are created by the mutational process after the application of therapy.
Figure 4: Illustration of an evolutionary rescue by de novo mutation.

Figure 4 provides an illustrative example of an evolutionary rescue: first, the population is rapidly declining as the treatment eradicates the sensitive cells. The sensitive cells can however acquire mutations which reduce their sensitivity to the drug. The emerged resistant cells will be strongly favoured by natural selection but can nevertheless be lost due to stochastic extinctions. If the resistant cells manage to establish, they will soon reoccupy the tumor niche and the population is rescued from extinction.

We use the term rescue window to the time period when the sensitive cells can acquire mutations and the population can still be rescued. Since the probability of gaining mutations is proportional to the population size, a characteristic feature of the rescue window is that the probability of an evolutionary rescue sharply decreases with the population size. This corresponds to the classical somatic mutation theory of drug resistance and gives rise to the MTD-therapies which aim to minimize the probability of an evolutionary rescue by making the rescue window as short as possible. Indeed, it is easy to show that this is the case when the mutations are assumed to occur spontaneously with a constant rate irrespective of the treatment. The focus of this thesis, however, lies precisely in the fact that such assumption
may in many cases be invalid due to drug-induced genetic and epigenetic changes that promote the emergence of resistance.

3.2 Drug-Induced Resistance as Cost of Control

The traditional cytotoxic chemotherapies may indeed have a quantifiable “biological cost”, namely drug-induced resistance. Drug-induced resistance can occur either if the genome-wide mutation rate increases during the treatment or if other forms of adaptive evolution, such as epigenetic regulation and phenotypic plasticity, take place. Essentially, drug-induced resistance means that the application of the drug somehow speeds up the evolution towards the resistant phenotype. In practice, however, the presence and the strength of the drug-induced effects are difficult to assess because of the confounding alternative models of resistance (see figure 1) and the intrinsic stochasticity which is necessarily present at different levels.

Even though we presently lack a quantitative understanding of the extent of drug-induced phenomena, there is nevertheless plenty of evidence of drug-induced mutagenesis and phenotypic changes not only in bacteria [10] but also in cancer [7–9]. The concept of drug-induced resistance is still a fairly novel one and thus the effects of drug-induced resistance have not received much attention in the previous theoretical and optimal control models of drug resistance. Greene et al. take one of the first steps in their recent paper [33] where they show that drug-induced effects may have significant impact on the optimal treatment strategies. They use a linear dose-dependent mutation rate and show that the slope of the dose-dependent mutation rate is theoretically identifiable. In [34] the same authors consider a time optimal control problem and find that a low-dose therapy maximizes the time until treatment failure. Essentially their solution represents the optimal containment strategy where the tumor burden is held as high as possible throughout the treatment period which then maximally prolongs the time until treatment failure by competitive release of the resistant cells.

Herein the optimal control problem is reformulated such that we seek to find the optimal elimination strategy, which minimizes the probability of an evolutionary rescue, in the presence of drug-induced effects. Another key insight introduced here is that the cost of control is realized indirectly where using control now reduces the possibility to control the system in the future. Naturally, if one necessarily wants, the conventional penalty terms for the cost of control can be easily carried also to our formulation of the problem. Consequently, it is not our intention to replace the evident other costs of the treatment by focusing on drug-induced resistance, but instead highlight that this is yet another, often neglected, cost of the treatment.
We concentrate our analysis to the linear dose-dependent mutation rate. This decision can be justified by the experimental evidence and previous literature. We are not interested on what is the precise cause (genetic or epigenetic) of the change in phenotype as our model merely distinguishes the phenotypes and the transitions between those phenotypes. This way the linear dose-dependency can also be seen as a conservative lower-bound to the true dependency, which very well may be superlinear due to the adaptive changes in phenotype.

3.3 Problem 1: Minimize the Probability of an Evolutionary Rescue

Let’s consider the problem of minimizing the probability of an evolutionary rescue by de novo mutation. First consider the following minimal dynamical model for drug resistance, where we assume that resistance is strictly a binary phenotype:

\[
\begin{\array}{l}
\dot{S} = r_S(N)S - d(u)S - \mu(u)S; \quad S(0) = S_0 \\
\dot{R} = r_R(N)R + \mu(u)S; \quad R(0) = R_0.
\end{array}
\] (3.1)

Here, only the phenotype of the cell is distinguished: sensitive cells $S$ may be controlled by administering cytotoxic treatment, which we assume has no effect on the resistant cells $R$. The control variable $u$ gives the concentration of the drug while the function $d(u)$ gives the pharmacodynamical effect, i.e. translates how the obtained drug concentration leads to the death of the sensitive cells. Sensitive cells can become resistant via the mutational process where $\mu(u)$ gives the mutation rate, which we allow to explicitly depend on the control variable due to the possible drug-induced effects. The model does not distinguish the precise cause (genetic or epigenetic) of the change in phenotype, but merely the transition from the sensitive compartment to the resistant compartment. For simplicity, however, we use the term “mutation” to refer to both genetic and epigenetic events that confer resistance to the drug, which is standard in the context of mathematical models.

The model allows the sensitive and resistant cells to have separate growth dynamics to include the effects of competition and the possible fitness costs related to maintaining the resistance mechanisms. Thus, the intrinsic per capita growth rates generally satisfy $0 < r_R \leq r_S$, while the effective per capita rates may depend on the total population size $N(t) := S(t) + R(t)$ as a result of competition. Constant growth rates lead to the exponential growth model (which is suitable only for small populations) while common density dependent choices include logistic $r_i(N) = r_i(1 - \frac{N}{K})$ and Gompertz
\( r_i(N) = r_i \log(\frac{K}{N}) \) growth models, where \( K \) is the assumed common carrying capacity, \( r_i > 0 \) and \( i \in \{S, R\} \).

Let us first consider the case where there are no resistant cells present prior to the treatment and where there is no competition between the sensitive and resistant cells. The emergence of new resistant mutants is a time-inhomogeneous Poisson process depicted in the rescue window (see figure 4), where the rate of gaining a new mutant at time \( t \) is given by the product of the population size \( S(t, u(t)) \) and the mutation rate \( \mu(u(t)) \). Note that both of these factors now depend on the control variable, since we explicitly take into account the drug-induced effects.

Since there are no or only few resistant cells present, we need to use stochastic population dynamics to model the growth of small populations that have a considerable stochastic extinction risk due to inherent demographic fluctuations. The stochastic extinction risk for a simple birth-death process founded by a single cell is given by \( q = \frac{1}{R_0} = \frac{\delta_0}{\beta_0} \), where \( \delta_0 \) and \( \beta_0 \) are the intrinsic death and birth rates respectively, leading to the net growth rate \( r_i = \beta_0 - \delta_0 \). For simplicity, we assume that this stochastic extinction probability is a constant given property of the resistant cell and specifically does not depend on the total population size (which would of course indirectly depend on the control variable).

With these assumptions in mind, the intensity, or the total cumulative rate of generating rescue mutations, is given by

\[
\int_0^T S(t, u(t))\mu(u(t))\pi_f dt, \quad (3.2)
\]

where \( \pi_f = 1 - q \) is the probability of establishment. This quantity also corresponds to the expected number of rescue mutants generated during the treatment period. The probability of an evolutionary rescue by \textit{de novo} mutation is then \cite{31}

\[
\Pr(\text{ER}) = 1 - \exp \left( - \int_0^T S(t, u(t))\mu(u(t))\pi_f dt \right). \quad (3.3)
\]

The exponential term is just the zero class of Poisson distribution and hence the complement of this gives the probability that there is at least one cell which survived the rescue window. A suitable objective of the treatment is then to minimize this quantity, which is equivalent with maximizing the extinction probability of the cancer cells. Since the probability of establishment does not depend on the control variable, the optimal control problem reduces to

\[
\min_{u \in \mathcal{U}} \int_0^T S(t, u(t))\mu(u(t))dt, \quad \mathcal{U} = [0, u_{\text{max}}]. \quad (3.4)
\]
Assuming the logistic growth model with a scaled carrying capacity $K = 1$, the dynamics of the sensitive cells becomes

$$
\dot{S} = r_s S (1 - S) - d(u) S - \mu(u) S; \quad S(0) = S_0. \tag{3.5}
$$

Finally, we assume a Hill-type pharmacodynamics

$$
d(u) = d_{max} \left(1 - \frac{1}{1 + \left(\frac{u}{h}\right)^k}\right) \tag{3.6}
$$

and linear dose-dependent mutation rate

$$
\mu(u) = \mu_0 + \alpha u. \tag{3.7}
$$

The Hamiltonian for the problem reads

$$
H = S \mu(u) + \lambda_1 S (r_s (1 - S) - d(u) - \mu(u)). \tag{3.8}
$$

Figure 5: Illustration of typical sigmoidal pharmacodynamical responses observed on logarithmic concentration scales. The parameter $k$ gives the steepness of the curve, parameter $h$ is the dose which yields 50% of the maximal death rate $d_{max}$. 

22
The optimal control is found by studying the function
\[ H_u(t) = S(t)(\alpha - \lambda_1(d'(u(t)) + \alpha)) =: \Psi(t), \]
whose roots, should they exist, determine the singular controls. In case no roots exists for some time point, then the sign of this function determines the bang-bang controls as given by equation (2.13). The dynamics of the multiplier is
\[ \dot{\lambda}_1 = -\mu(u) - \lambda_1(r_S(1 - 2S) - d(u) - \mu(u)); \quad \lambda_1(T) = 0. \] Notice that since the cost functional (3.4) does not include the resistant cells and we assume that competitive interactions with the resistant cells are negligible during the treatment period, we need not explicitly model and track the resistant cells.

Consider the case where no dose-dependency is assumed ($\alpha = 0$). Then we have that
\[ \Psi = -S\lambda_1d'(u) < 0 \]
for all $t \in [0, T]$ since the population size and the associated multiplier must be positive over the entire control period. This is because there is no explicit cost on control and the sensitive cells have the sole function of producing de novo mutations. Thus, the optimal control reduces to MTD-solution as $\alpha \to 0$. On the other hand, the optimal control should also switch to use no control at all if $\alpha$ is sufficiently large. This is because the cost for $u \equiv 0$ is bounded from above. These latter cases are most likely not biologically relevant, but demonstrate that our formulation reduces to the bang-bang controls as a special case. Now, the real interest lies in the intermediate but still biologically relevant values of $\alpha$ where singular solutions may arise. Bare in mind that we have currently no good a priori knowledge of the true values of $\alpha$ and this needs to be estimated using dose-response studies on a case-by-case basis.

Forward-Backward-Sweep Method can be used to solve the optimal control as before with the slight complication that we now have no explicit expression for $u^*(t, x, \lambda_1)$ which to pass on to the algorithm. Thus, we need to numerically search the roots of $\Psi(t)$ at each time point separately using the updated value of $\lambda_1$ (notice that neither the root or the sign depends on $S$). Figure 6 illustrates the numerical root finding which determines the iterable optimal control. There in fact appear two distinct roots, but the first one (leading to very low-doses) can be shown to be always non-optimal using the Legendre-Clebsch condition (requires analysis of the second derivative of the dose-response). Thus, a log-spaced discretization for the root searching is preferable as it neglects to first root and finds the actual root more accurately.
Figure 6: Illustration of finding the roots, which determine the optimal control.

Figure 7: Solution of the problem (3.4) using parameter values $\alpha = 10^{-8}$, $r_S = 0.5$, $d_{max} = 0.8$, $u_{max} = 1000$, $h = 40$, $k = 2.5$, $\mu_0 = 10^{-6}$, $T = 30$ and $S(0) = 0.99$. 
Figure 7 shows the solution of the problem for specific parameter values: figure 7A plots the optimal treatment strategy $u(t)$ and figure 7B gives the optimally controlled trajectories, where the resistant cells followed a deterministic exponential growth (with establishment probability of 1). Deterministic dynamics always leads to an evolutionary rescue so the quantity of interest is the expected number of rescue mutants generated during the treatment period, which is just the area under the intensity curves depicted in figure 7C. In reality, of course, we sometimes succeed in driving the population extinct; the probability of a cure can be calculated from equation (3.3) if all the parameter values, absolute population sizes and the probability of establishment are known. The probability of establishment can also be estimated using stochastic simulations.

Figure 7C demonstrates the fundamental trade-off between population size and mutation rate. The solid blue line is the rate at which the optimally controlled population generates rescue mutants plotted as a function of time. The dashed lines are examples of intensities which result from higher and lower constant doses. Even though the higher dose eliminates the sensitive cells faster than the optimal dose, it suffers from a higher probability peak at the very beginning as it uses lot of control to a large population. On the other hand, while lower than optimal doses have a smaller probability peak at the beginning, it takes more time to eliminate the sensitive cells later thus leading to higher probabilities at the end. The optimal control is then such that it trades-off with the length of the rescue window and the early probability peak and balances these two opposite forces of action to constitute the optimal treatment strategy.

Figure 8 compares the optimal dose to various other constant doses. For an example, a constant dose of 200, which is still far from the region where the pharmacodynamical effect is plateauing, produces on average 49% more rescue mutants then the optimal strategy and leads to 76% higher residual resistant population. These numbers are of course purely speculative as they crucially depend on many (generally biologically plausible) parameter values, but underline that fairly modest drug-induced effects can, at least in some cases, lead to clinically significant differences in favour of the optimal strategy.

We also notice that a constant time average of the optimal dose performs almost as well as the optimal dose itself. Thus, the precise time implementation is not particularly significant, but rather the absolute dose administered. Finally, we notice that while the optimal strategy minimizes the intended cost functional (that is, the expected number of generated rescue mutants), the compared lower doses have a lower residual resistant population. We shall return to discuss the problem of finding the strategy which minimizes the
rescue fraction in section 3.5.

Figure 8: Comparing the optimal strategy $u_{opt}$ displayed in figure 7 to other constant strategies using deterministic dynamics. The optimal control minimizes the total expected number of rescue mutations, but if evolutionary rescue does occur, it leads on average to a slightly higher residual resistant population than the compared lower doses.

Notice that in the comparison above we have no way of determining where the true toxicity constraint is, but instead this is a specific property of the drug used and the MTD is to be determined in the clinical trials. Thus, we regard the MTD to be a given parameter, which together with the parameter $\alpha$ determine how strong the drug-induced effects (and the consequent benefits of the optimal treatment) are. This is demonstrated in figure 9, where the costs of the MTD-treatments are compared to the optimal treatment strategy in the relevant region of parameter $\alpha$. The heat map gives the fold change of the baseline mutation rate $\mu_0$ that is assumed in each scenario. The plotted contour lines give the relative cost of the respective MTD-treatment compared to the optimal treatment strategy: for an example, the 1.5 line gives all the different scenarios where MTD-treatment produces on average 50% more rescue mutants than the optimal strategy.
Figure 9: The heat map demonstrates what fold-changes to the baseline mutation rate $\mu_0$ are assumed for different MTD-treatments in the relevant parameter range of $\alpha$. The contour lines plotted on top show the relative cost of the respective MTD-dose compared to the optimal strategy. For an example, the 1.5 line gives the different scenarios where the MTD-dose produces on average 50% more rescue mutants than the optimal strategy.

Figure 9 reveals that highly significant differences in favour of the optimal strategy can be obtained already for very modest drug-induced effects (of order 2- to 3-fold increases in mutation rate) if the drug is well tolerated. On the other hand, even if the MTD-dose is very low, there is room for clear improvement if the drug is simultaneously highly mutagenic. Of course, it is possible that the MTD-dose is even lower than the optimal dose, in which case the MTD is optimal.

Another point to remember is that even if the optimal strategy performs significantly better in terms of the expected number of mutant establishment attempts, these differences may not be significant in the probability level if the baseline mutation rate $\mu_0$ is too large. In these cases the probability of an evolutionary rescue is in any case very high and cannot be significantly improved by treatment optimization. The entire point of eliminating the sensitive cells may then be questioned and containment strategies should be considered instead. The challenge lies in quantifying this probability reliably, so that it could guide treatment decisions.
3.4 Time-Independent Control Law

Would it be possible to obtain a closed-loop control law, which would not give the optimal control strategy as a function of time but rather as a function of the population size? Such solution would generalize the result as it no longer would depend on the initial condition and secondly it would eliminate the artificial boundary effects that result from the arbitrarily fixed treatment period. After all, the total tumor burden can be easily clinically evaluated and adjusting the dose dynamically with respect to that is easier than along some strict time schedule. The question of when such control law can be obtained in general is an interesting mathematical question.

The control law can be obtained analytically using the Hamilton-Jacobi-Bellman equation. The control law we are looking for is such that it is independent of time. This means that the cost-to-go $J(x, t)$ should only depend on the population size $x$. In other words, we are looking for a stationary solution where $J(x, t') = J(x, t' + dt)$ and thus $\partial_t J(x, t') = 0$. (In general, also cases where $\partial_t J(x, t') = \text{constant}$ are called stationary solutions.) Using the HJB-equation at time $t'$ we get

$$
\partial_t J(x, t') = \min_{u \in U} \left( L(t', x, u) + f(t', x, u) \partial_x J(x, t') \right) = 0.
$$

(3.11)

Carrying out the minimization by formally differentiating with respect to $u$ leads to an implicit condition

$$
L_u(t', x, u) + f_u(t', x, u) \partial_x J(x, t') = 0,
$$

which the optimal control must satisfy, if an interior optimum exists. Then, substituting

$$
\partial_x J(x, t') = -\frac{L_u(t', x, u)}{f_u(t', x, u)}
$$

back to the HJB-equation, we obtain

$$
L(t', x, u) - f(t', x, u) \frac{L_u(t', x, u)}{f_u(t', x, u)} = 0
$$

(3.12)

from which the stationary profile $u(x)$ can be solved. In our case, equation (3.11) reduces to

$$
S\mu(u) + S(r_S(1 - S) - d(u) - \mu(u))\left(\frac{\alpha}{d'(u)} + \alpha\right) = 0,
$$

which can be further simplified by cancelling $S$ from the both terms and neglecting the mutational term $\mu(u)$ from the dynamics $f$. This yields equation

$$
\mu_0 + \alpha u + (r_S(1 - S) - d(u))\left(\frac{\alpha}{d'(u)}\right) = 0,
$$

(3.13)
which can be solved numerically for $u$ for every fixed population size $S$. Note that the boundary optima must be excluded separately. Result for the same parameters as before is presented in figure 10.

Figure 10: Control law for the same problem as presented in figure 7.

Equation (3.12) can now be used to the sensitivity analysis for the parameter $\alpha$: we can simply solve the stationary profile ($u(S)$) separately for each $\alpha$ or any other parameter of interest. Perturbing the parameter $\alpha$ preserves the almost linear profile only adjusting the intercepts (the amount of control). Remember that this analysis does not exclude the boundary optima and they must always be compared to the singular solution.

The same stationary profile can be obtained numerically more simply by first solving the optimal control for initial condition $S(0) = 1$ and then taking the end-time $T$ sufficiently large so that the population size at the end is practically zero. Then, given that $S$ is a monotonically decreasing function (i.e. the optimal control does not allow the population to regrow in between ), there exists an inverse function $S^{-1} : [0, 1] \to [0, T]$ which gives the time at which the population was at any given size. Now, if indeed $u(S)$ does exist, it must be unique and thus the optimal control for some population size $S'$ must satisfy $u(S') = u(t')$ where $S(t') = S'$. Then the stationary profile can be obtained using the inverse function as $u(S') = u(S^{-1}(S'))$. Applying this to the optimal solution displayed in figure 7 gives the same control profile as
3.5 Problem 2: Discount for the Growth

Figure 8 shows that the mutation minimizing strategy may nevertheless lead to a higher resistant population. This is because the time when the mutations occur has a significant role: the resistant cells which emerge very early of course have more time to generate growth than the resistant cells that emergence only at the end of rescue window. If we wish to minimize the number of resistant cells, we need to assign a higher cost for the mutations that occur early. More precisely, we need to discount the mutations by the growth they can generate. If we assume that the resistant cells grow exponentially at a rate $r_R \leq r_S$, then a resistant cell that emerges at time $t$ can generate

$$e^{r_R(T-t)}$$

resistant cells during the treatment period. Then, if there are no resistant cells at the beginning of the treatment period, the number of resistant cells at the end can be written as

$$R(T) = \int_0^T S(t, u(t))\mu(u(t))e^{r_R(T-t)}dt, \quad \text{if } R(0) = 0. \quad (3.15)$$

We can take this as our cost functional and refer to this as the discounted problem. The solution procedure is almost identical with the exception that the exponential term (3.14) needs to be carried to the switching function $\Psi$ and to the multiplier dynamics.

We notice that the optimal control strategy, which minimizes the number of resistant cells at the end, uses even less control than the mutation minimizing strategy but still preserves the same linear profile of the solution. Essentially the optimal solution balances the early mutational peak even further to avoid generating too many resistant cells very early on. This of course comes at the expense that more mutations are generated later on and the total (expected) number of mutations is higher.
3.6 Towards Evolutionary Rescue Experiments

In figure 12 the costs of various constant therapies are plotted in the relevant region of parameter $\alpha$. The resulting surface displays an interesting non-monotonic dose response: when drug-induced effects are negligible, the residual tumor burden decreases with the dose thus leading to the MTD-paradigm. However, with drug-induced effects present, the residual tumor burden will first decrease with moderate doses and then start to increase with higher doses. The optimal constant dose is plotted as a function of $\alpha$ with the red solid line.
Figure 12: Rescue fractions $R(T)$ plotted for constant therapies in the relevant region of parameter $\alpha$. The red line displays the solution to the discounted problem and gives the dose which minimizes the expected number of resistant cells at the end of the treatment period. Note that if very low doses are used, then the residual tumor burden $N(T)$ will be higher as there will remain sensitive cells.

Since the cost function of the discounted problem is highly intuitive, it is just the population size after treatment, these results provide striking predictions that are immediately amenable to experimental investigation both in cancer and in the context of antibiotic resistance. Using antibiotics with reported drug-induced effects [10] to bacterial populations at a various (constant) doses ranging from low to high, we should be able to test whether the observed rescue fractions comply with the predicted non-monotonic relationship (see figure 13). This experimental setup can also be used to identify, whether drug-induced effects are present or not, with the crucial advantage that we do not need explicit knowledge of the birth and death rates which can confound the results obtained in fluctuation assays.
Figure 13: Consider an experiment where varying constant doses are used to control a population with $N(0) = 1$. If no drug-induced effects are present ($\alpha = 0$), then the expected rescue fractions $N(T)$ should be decreasing function of the dose used. However, if drug-induced effects were to be present ($\alpha > 0$), the rescue fractions should display a non-monotonic dose response.

The outcome of such evolutionary rescue experiments is a bimodal distribution of the rescue fractions (see figure 14). The first zero-peak corresponds to the expected proportion of successful therapies which drive the population extinct. The optimal strategy which maximizes the probability of a cure was calculated in Problem 1. The second peak corresponds to the expected final-size of non-extinct populations. Problem 2 discussed the problem of finding the treatment strategy which minimizes the mean of this distribution conditioned on non-extinction (a cure is never seen in deterministic dynamics).

The rationale of all elimination strategies lies in the fact that one has a reasonable chance to cure the disease. More specifically, this is the case only when the mutation is the rate-limiting step towards resistance and the risk of pre-existing resistance is not too large. The effective baseline mutation rate $N_0\mu_0$ is a key quantity in determining the probability of an evolutionary rescue. Figure 14 illustrates the role of effective baseline mutation rate on the probability of an evolutionary rescue.
Figure 14: 500 stochastic simulations of three different constant therapies were performed using the Gillespie algorithm. Each treatment strategy leads to a characteristic bimodal distribution where the zero-peak corresponds to the proportion of successful therapies. The red distribution displays the optimal (constant) treatment strategy, which maximizes the probability of a cure. However, if the effective baseline mutation rate is greater than 1, evolutionary rescue will occur with a very high probability.
4 Discussion

MTD-style therapies have rooted deeply to oncology because the chances of a complete cure are assumed to be maximized by such a regimen. This dates back to the classical somatic mutation theory of drug resistance where new resistant cells arise spontaneously at a constant rate irrespective of the treatment. In these cases the probability of an evolutionary rescue is indeed minimized by eradicating the sensitive cells as fast as possible and consequently the MTD-paradigm constitutes the optimal treatment strategy when resistance is conferred primarily by \textit{de novo} mutation.

The results obtained in this thesis demonstrate that the situation changes if there are even modest drug-induced effects present. This leads to a novel trade-off where the population size can be decreased only at the expense of also simultaneously increasing the mutation probability. In such cases, the MTD-strategy actually increases the likelihood of an evolutionary rescue and thus the treatment outcomes may, at least in some cases, be significantly improved by treatment optimization. Our results thus confirm the observations made by Greene \textit{et al.} in [33] that the drug-induced effects have a crucial impact to the optimal treatment strategy. Consequently, it will be of great importance to properly investigate the various mutagenic and resistance promoting properties of different anti-cancer therapies in experimental and clinical settings.

The main focus of this thesis was to solve the optimal elimination strategy, which minimizes the probability of an evolutionary rescue, in the case of dose-dependent mutation rate. Recently, however, the very founding objective of eliminating the tumor burden has been challenged and so-called containment strategies have been proposed to specifically avoid the competitive release of the resistant cells. Such paradigm shift in treatment may greatly improve treatment outcomes especially in those situations where there is high abundance of pre-existing resistant cells and a complete cure cannot be expected. First proofs of concept have already be made by Gatenby and colleagues in cancer [35] and recently also in the context of antibiotic resistance [36]. Hansen \textit{et al.} then argue from this basis that all viable treatment strategies must trade-off against minimizing mutations (to prevent the emergence of new resistant cells) and maximizing competition (to suppress the growth of the existing resistant cells).
Figure 15: Treatment optimization in the multi-objective sense: the Pareto-frontiers illustrates the trade-offs between the two alternative objectives of the treatment where mutations can be further reduced only at the expense of also reducing the competition. When no drug-induced effects are present, we lie on the blue curve and the mutation minimizing solution is given by the MTD, plotted as a red dot. However, if drug-induced effects are present, the optimization is done on a completely different, yellow Pareto-frontier. Using MTD in these cases leads to a particularly detrimental situation where the red dot lies below the trade-off curve and thus significant improvement could be achieved by treatment optimization by moving to the green dot on the yellow Pareto-frontier. The problem of determining the optimal treatment on the Pareto-frontier still remains and is patient-specific. The blue and yellow arrows demonstrate the direction to which the cumulative control used increases when moved on the frontier. Note that the shape of the Pareto-frontiers need not be necessarily concave, but also linear and convex frontiers are possible depending on the optimality of the intermediate solutions.

To put the results of this thesis to a wider context, consider figure 15 which illustrates this fundamental trade-off between the two alternative evolutionary pressures that can be induced by treatment. So far, the discussion of treatment optimization has exclusively concentrated on the blue trade-off curve (or Pareto-frontier in the language of multi-objective optimization) and
the common belief remains to be that MTD-style treatment strategies are
the ones which minimize the mutations. The key result of this thesis is that
this is not generally true when there are drug-induced effects present as the
optimization must then be done on a completely different trade-off curve.
Neglecting the effects of drug-induced resistance can lead to situations where
the MTD-strategy lies well below the Pareto-frontier and is thus a particu-
larly detrimental strategy as it fails on both aspects. Using the methods
presented here, one can identify the optimal mutation minimizing solution
and thus potentially gain significant improvements.

The approach taken in this thesis has many advantages. It proposes a
novel way of formulating the precise objectives of the treatment in evolu-
tionary terms and contains an interesting theoretical framework for further
treatment optimization avenues. Furthermore, it specifically considers the
effects of drug-induced resistance, an often neglected cost of treatment, and
highlights the importance of these effects. The results obtained provide, to
my best knowledge, a new qualitative insight that is potentially exploitable
by treatment optimization. The predictions made by this work are also im-
mediately amenable to experimental investigation where a lot of future study
is needed to gain a better quantitative knowledge of the scope and extent of
the drug-induced effects.

However, there are also potential weaknesses that may limit the use-
fulness of these results, especially when it comes to the more quantitative
predictions. Many of the underlying assumptions of the dynamical model
can be questioned. This is especially the case of the linear dose-dependent
mutation rate, where we currently lack a good understanding of the true
dose-dependency. This choice was however justified as a plausible and fairly
conservative motivated by some of the experimental evidence and also pre-
vious literature. Similar qualitative results are likely to be observed for any
monotonically increasing dose-dependency. Also the assumptions related to
the growth dynamics and other more implicit assumptions of the dynamical
model can be revisited when needed without excessively complicating the
solution of optimal controls. Indeed, the purpose of the minimal model pre-
sented was to setup a working model to study the control structures instead
of providing the most realistic and comprehensive description of the under-
lying processes. Finally, given the wide-spread use of MTD-style therapies,
these results may be significant and worth of further investigation even if
they only apply only to certain drugs.

There remain many outstanding questions for future research related to
the topics discussed. Firstly, to be able to design and implement the opti-
mal mutation minimizing treatment strategies, we need to have an adequate
quantitative understanding of the dose-dependency of the mutation rate. In
the case of containment strategies we need to be able to better characterize and quantify the intracellular competition to which all containment strategies rely on. Secondly, it is clear that we need to carefully assess and identify those situations where containment strategies would most likely perform better than elimination strategies. Could cancers be contained so successfully that in cases where a complete cure is not likely (for an example due to great number of pre-existing resistant cells) the cancer could be transformed into a chronic and manageable disease?

Finally, it would be crucial that also the metastatic potential of different cancers could be quantified, predicted and incorporated to the treatment optimization. For example, how does the tumor burden affect the probability of seeding a metastasis? Do containment strategies, which aim to deliberately maintain a high tumor burden, significantly increase the risk of metastases? These are important questions which among other factors determine whether the usefulness of containment strategies is limited only at prolonging the survival in highly advanced cancers or do containment strategies denote a more widespread paradigm change in treatment.

By understanding cancer as an evolutionary process, the emergence of resistance is no longer a mystery but rather a fairly direct consequence of the treatment. In fact, given the tremendous heterogeneity and complexity of all clinical cancers, the true mystery and question to be asked is how can the current treatments so often be successful. What is the role of the immune system in permanently driving the cancer cells extinct and how would this change if other treatment strategies than MTD were to be used? What distinguishes curable cancers from relapsing ones, the metastasizing cancers from local ones and so on? We need urgently tools that could be used to quantify and predict these different evolutionary potentials. Ultimately, the goal of these attempts is in the pursuit of a unified theory on how evolutionary processes can be predicted and controlled in cancer and beyond.
References


Recent paper which uses a linear dose-dependent mutation rate to show that drug-induced resistance can lead to qualitatively different treatment responses and that slope of the dose-dependent mutation rate is theoretically identifiable.

