

ABSTRACT

Title of dissertation: MATHEMATICAL MODELS OF
TUMOR HETEROGENEITY
AND DRUG RESISTANCE

James Greene, Doctor of Philosophy, 2015

Dissertation directed by: Professor Doron Levy
Department of Mathematics

In this dissertation we develop mathematical models of tumor heterogeneity and drug resistance in cancer chemotherapy. Resistance to chemotherapy is one of the major causes of the failure of cancer treatment. Furthermore, recent experimental evidence suggests that drug resistance is a complex biological phenomena, with many influences that interact nonlinearly. Here we study the influence of such heterogeneity on treatment outcomes, both in general frameworks and under specific mechanisms.

We begin by developing a mathematical framework for describing multi-drug resistance to cancer. Heterogeneity is reflected by a continuous parameter, which can either describe a single resistance mechanism (such as the expression of P-gp in the cellular membrane) or can account for the cumulative effect of several mechanisms and factors. The model is written as a system of integro-differential equations, structured by the continuous “trait,” and includes density effects as well as mutations. We study the limiting behavior of the model, both analytically and numerically, and apply it to study treatment protocols.

We next study a specific mechanism of tumor heterogeneity and its influence on cell growth: the cell-cycle. We derive two novel mathematical models, a stochastic agent-based model and an integro-differential equation model, each of which describes the growth of cancer cells as a dynamic transition between proliferative and quiescent states. By examining the role all parameters play in the evolution of intrinsic tumor heterogeneity, and the sensitivity of the population growth to parameter values, we show that the cell-cycle length has the most significant effect on the growth dynamics. In addition, we demonstrate that the agent-based model can be approximated well by the more computationally efficient integro-differential equations, when the number of cells is large. The model is closely tied to experimental data of cell growth, and includes a novel implementation of transition rates as a function of global density.

Finally, we extend the model of cell-cycle heterogeneity to include spatial variables. Cells are modeled as soft spheres and exhibit attraction/repulsion/random forces. A fundamental hypothesis is that cell-cycle length increases with local density, thus producing a distribution of observed division lengths. Apoptosis occurs primarily through an extended period of unsuccessful proliferation, and the explicit mechanism of the drug (Paclitaxel) is modeled as an increase in cell-cycle duration. We show that the distribution of cell-cycle lengths is highly time-dependent, with close time-averaged agreement with the distribution used in the previous work. Furthermore, survival curves are calculated and shown to qualitatively agree with experimental data in different densities and geometries, thus relating the cellular microenvironment to drug resistance.

MATHEMATICAL MODELS OF TUMOR HETEROGENEITY
AND DRUG RESISTANCE

by

James Greene

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:

Professor Doron Levy, Chair/Advisor

Professor Pierre-Emmanuel Jabin

Professor Konstantina Trivisa

Professor Matei Machedon

Professor Wolfgang Losert, Dean's Representative

© Copyright by
James Greene
2015

Dedication

This work is dedicated to my wife, Jennifer, without whom this would be neither possible nor meaningful.

Acknowledgments

During my tenure as a graduate student at the University of Maryland, I have been fortunate to have been thoroughly supported academically, emotionally and financially by many people.

I would first like to thank my advisor, Doron Levy. Professor Levy has been very supportive of me over the last few years, and has provided me with many tools to make me a better applied mathematician and scientist. From assistance in writing, to networking/presentation opportunities, to daily research, Professor Levy has provided me hours of counsel and aid, for all of which I am acutely grateful. I have been lucky to have such an kind, encouraging, and intelligent advisor.

I must also thank Professor Levy and the other members of my committee, Professors Pierre-Emmanuel Jabin, Konstantina Trivisa, Matei Machedon, and Wolfgang Losert, for carefully reading through this dissertation.

There are a number of other professors, both at the University of Maryland and at the Pennsylvania State University, who have helped shape my academic career. I would especially like to thank Professor Eitan Tadmor for allowing me to be a part of the stimulating and collaborative CSCAMM environment, as well as Professor Joseph Previte and Professor Diane Henderson, for both introducing me to mathematical research as an undergraduate and encouraging me to pursue it further. Lastly, I would like to thank two of my high school mathematics teachers, Mrs. Kristin Meeker and Mrs. Renee Gerber, for first sparking my interest in math.

The Laboratory of Cell Biology at the National Institutes of Health, headed by

Dr. Michael Gottesman, has also been instrumental in my education. In particular I must thank Dr. Orit Lavi who, despite being a trained mathematician, always kept my work both biologically motivated and biologically accurate. Although our collaborations have not always run “smoothly”, my work has been better for them.

Throughout my time at UMD, I have been financially supported through a number of sources, including teaching assistantships with the Department of Mathematics, Professor Levy’s NSF research grants, a two-year NCI/UMD Partnership in Cancer Technology Fellowship, and directly from Dr. Gottesman’s laboratory. I am also thankful for receiving numerous travel grants from the department that allowed me to attend numerous conferences, workshops, and summer schools.

I appreciate the friendship and the mathematical conversations I have had with my colleagues, who made these graduate years productive and enjoyable, especially Dan Weinberg, Paul Koprowski, Ming Zhong, Amanda Galante, Shelby Wilson, and Dana Botesteanu.

I am also indebted to the many staff members from both the Department of Mathematics and CSCAMM, namely Celeste Regalado, William Schildknecht, Agi Alipio, Jeff Henrikson, and Scott Boutaugh.

The support and love of my family has been invaluable to any success that I have had. I owe my deepest thanks to my wife, Jennifer, for her unconditional love, never-ending faith, and immeasurable patience (the last of which every drop was necessary). My mother Susan Greene has always believed in me, and has always found a way be strong, even when it was most dark; I admire her most of all. And my sister Melinda Greene, who always gave me support whenever I needed it.

I would also like to thank the countless unnamed others who have helped me over the years, and apologize for any names I have inadvertently omitted from this list.

Contents

List of Tables	ix
List of Figures	x
List of Abbreviations	xix
1 Introduction	1
1.1 A brief introduction to cancer and chemotherapy	1
1.2 Tumor heterogeneity	5
1.3 The biology of drug resistance	7
1.3.1 Developmental origins	7
1.3.2 Mechanisms	9
1.4 Structure of the dissertation	11
2 The mathematics of drug resistance - a brief overview	13
2.1 Probability of resistant cells existing at diagnosis	13
2.2 Combination versus sequential therapy	15
2.3 Optimal treatment strategies	19
2.4 Modeling mechanisms: the P-glycoprotein pump	23
3 Cell density, mutations, and drug resistance: a structured population approach	26
3.1 Introduction	26
3.2 A mathematical model for cancer dynamics	30
3.3 Analysis and simulations	36
3.3.1 Trait-based growth	36
3.3.2 Density-dependent model	44
3.3.3 Selection/mutation model	54
3.4 Discussion	63
4 Clinical implications of drug resistance as a continuous variable	65
4.1 The role of cell density and intratumoral heterogeneity in multidrug resistance	65
4.1.1 Introduction	66
4.1.2 Quick guide to equations and assumptions	70
4.1.3 Results	71

4.1.3.1	Heterogeneity arising from treatment	71
4.1.4	Heterogeneity arising from both epigenetic and genetic changes	79
4.1.5	Discussion	84
4.2	Simplifying the complexity of resistance heterogeneity in metastasis .	87
4.2.1	Heterogeneity in primary tumors and in metastasis	88
4.2.2	Resistance level as a discrete variable	90
4.2.3	Resistance level as a continuous variable	92
4.2.4	Concluding remarks and future perspectives	102
5	Modeling intrinsic heterogeneity and growth of cancer cells	105
5.1	Introduction	106
5.2	Agent-based model	107
5.2.1	Model construction	108
5.2.2	Rate derivations	113
5.2.3	ABM algorithm	116
5.3	Integro-differential equation model	117
5.3.1	Approximation of expected values	118
5.3.2	The derivation of the IDE	120
5.3.3	Numerical methods	128
5.4	Results	130
5.4.1	Distribution of cell-cycle and apoptotic lengths	131
5.4.2	ABM simulation	133
5.4.3	Stochastic and deterministic comparison	136
5.4.4	Parameter estimation and variation	139
5.5	Conclusion	143
6	The cell-cycle and drug resistance - a spatial mechanism	148
6.1	Introduction	148
6.2	A mathematical model of cancer cell dynamics	150
6.2.1	Motility	150
6.2.2	Chemotherapy	153
6.2.3	Compartments in the ABM model	154
6.2.3.1	Quiescence (Q)	154
6.2.3.2	Proliferation (P)	157
6.2.3.3	Apoptosis (A)	160
6.2.4	The ABM algorithm	161
6.3	Results	162
6.3.1	Control parameter estimation	162
6.3.2	Cell-cycle distribution	167
6.3.3	Density-dependent drug resistance	171
6.4	Discussion	175
7	Conclusion	176
A	Chapter 3	179

B	Chapter 4	183
B.1	Supplemental material for Section 4.1	183
B.2	Supplemental material for Section 4.2	184
C	Chapter 5	186
C.1	Experiments	186
C.1.1	Experimental Design	186
C.1.1.1	Analysis of Proliferation Percentage	186
C.1.1.2	Evaluation of Cell Growth and Density	188
C.1.1.3	Analysis of Apoptosis Percentage	188
C.2	Variable table	190
	Bibliography	191

List of Tables

3.1	List of variables	31
3.2	Parameter values used in simulations	40
5.1	List of minimization parameters and ranges	142
6.1	List of variables for spatial ABM algorithm	163
C.1	List of parameters and variables	190

List of Figures

3.1	Numerical results of (3.8). (a) shows the rate profiles used; (b) a semi-log plot of the evolution of the density ρ ; (c) $n(x, t)/\rho(t)$ as a function of x at three representative times t . Note that in the simulations we have made the time scale change $t \rightarrow t\epsilon^{-1}$, where $\epsilon = 10^{-2}$, since we are primarily interested in studying the long-time behavior of the solution.	41
3.2	Numerical results of (3.8) with $c(x) \equiv 0$, and the same parameters appearing in Fig. 3.1. Shown is $n(x, t)/\rho(t)$ as a function of x at three representative times t	43
3.3	Simulations of (3.16) using trait parameters and time scale as in Fig. 3.1. G is given by (3.19). (a) the profiles of $n(x, t)$ at four times t ; (b) the variation of the net-growth rates as the density changes in time, at the same four times used in (a); (c) evolution of the density $\rho(t)$	46
3.4	Simulations of (3.16) using trait parameters and time scale as in Fig. 3.1 and G is given by (3.19). (a) the net-growth rates as a surface plot over t and x ; (b) is a contour plot of $n(x, t)$	46
3.5	$r(x) - c(x)$ and $d(x)$ used for Figs. 3.6 and 3.7. Only the difference between $r(x)$ and $c(x)$ is plotted since that is the term that controls the dynamics in all models. Note that $r(x) - c(x) < d(x) \forall x \in [0, 1]$	49
3.6	Simulations of (3.16) using the trait parameters shown in Fig. 3.5. G is given by (3.19). The time scale as in Fig. 3.1. (a) the profiles of $n(x, t)$ at four representative times; (b) the evolution of the density; (c) the variation of the net-growth rates as the density changes in time, at the same four times used in (a).	50
3.7	Simulations of (3.16) with the trait parameters shown in Fig. 3.5. G is given by (3.19). The time scale are identical to those in Fig. 3.1. (a) a surface plot of the net-growth rates; (b) a contour plot of $n(x, t)$	50
3.8	$r(x) - c(x)$ and $d(x)$ used for Figs. 3.9–3.11. Note that $\operatorname{argmax}_{0 \leq x \leq 1} (r(x) - c(x) - d(x)) = 0.5$ and that $r(0.5) - c(0.5) - d(0.5) > 0$	51
3.9	Simulations of (3.8) using rates shown in Fig. 3.8. (a) a semi-log plot of the evolution of the density ρ ; (b) the profiles of $n(x, t)$ at three representative times. Note that in the simulations we have made the time scale change $t \rightarrow t\epsilon^{-1}$, where $\epsilon = 10^{-2}$, since we primarily are interested in the long-time behavior.	52

3.10	Simulations of (3.16) using the trait parameters shown in Fig. 3.8. G is given by (3.19). (a) the profiles of $n(x, t)$ at four representative times; (b) the evolution of the density; (c) the variation of the net-growth rates as the density changes in time, at the same four times used in (a). Note that in the simulations we have made the time scale change $t \rightarrow t\epsilon^{-1}$, where $\epsilon = 10^{-2}$, since we primarily are interested in the long-time behavior.	53
3.11	Simulations of (3.16) with trait parameters in Fig. 3.8. G is given by (3.19). (a) the net-growth rates as a surface plot over t and x ; (b) a contour plot of $n(x, t)$. Note that in the simulations we have made the time scale change $t \rightarrow t\epsilon^{-1}$, where $\epsilon = 10^{-2}$, since we primarily are interested in the long-time behavior.	54
3.12	Simulations of (3.31), using the rate parameters shown in Fig. 3.1(a). G is given by (3.19). $\epsilon = 0.01$, and $\theta = 0.1$. (a) the profiles of $n(x, t)$ at four representative times; (b) the evolution of the density; (c) the variation of the net-growth rates as the density changes in time, at the same four times used in (a).	58
3.13	Simulations of (3.31), with rate parameters shown in Fig. 3.1(a). G is given by (3.19). $\epsilon = 0.01$, and $\theta = 0.1$. (a) a surface plot of the net-growth rates; (b) a contour plot of $n(x, t)$	58
3.14	Simulations of (3.31), at time $t = 10$. Plotted are $n(x, 10)$ obtained using rate parameters shown in Fig. 3.1(a). G is given by (3.19). $\theta = 0.1, 0.4, 0.7, 1$. For all plots, $\epsilon = 0.01$ is fixed.	61
3.15	Simulations of (3.31) at time $t = 100$. All other parameters are identical to those used in Fig. 3.14.	62
4.1	Heterogeneity arising from the administration of chemotherapeutic drugs, without genetic/epigenetic changes. (A) rate of cell division $r(x)$; natural death $d(x)$; and drug-induced death for a given dose $c(x)$. All rates are functions of the trait x . The two types of treatments are administered separately, in which drug 1 is a more effective treatment than drug 2. (B) numerical predictions of the net growth ($n(x, t)$) after a certain time ($t = 12.5$), when no alterations (e.g., mutations) occur, with different initial conditions (solid and dashed lines, IC_1 vs. IC_2). In addition to the two treatments, the control (growth without treatment) is also plotted and labeled as C_0 . This plot shows the selection process resulting from treatment with respect to the trait, and demonstrates the time scale it would take to develop a fast-growing tumor. (C) initial conditions, IC_1 : $n(x, t = 0) = 1$ for all x (solid curves); IC_2 : $n(x \leq 0.25, t = 0) = 0$, or $n(x > 0.25, t = 0) > 0$ (dashed curves). (D) and (E), cell density, $\rho(t)$, for the two types of treatments: drugs 1-2, with the initial condition of IC_1	72

4.2	Death rate as a function of dose and resistance level. (A) and (B) two common types of treatment (dashed and solid curves) for cells with different resistance levels (red curves for sensitive cells and blue for resistant cells). The cells may react differently to certain drugs. For example, a very toxic drug produces a lower IC_{50} value (solid vs. dashed lines), with a smaller distance between the two curves representing these cells (Δ_1 vs. Δ_2). Also, the slope can demonstrate the response to a treatment, in which the steeper slope is the desired one ($slope_1$ vs. $slope_2$). (C) and (D) two examples of death rates ($c_1(x, \alpha)$ vs. $c_2(x, \alpha)$), in which the heights and gradients of the functions vary according to the type of treatment. (E) and (F) the cell density as a function of the dose and duration of a treatment for the cases described in C and D.	74
4.3	Drug efficacy as a function of the dose and resistance level. Panels A and C represent the two common types of drug-induced death rates, $c_1(x, \alpha)$ and $c_2(x, \alpha)$ that are plotted in Fig. 4.2C and D, respectively. Given the exact function of the drug-induced death rate, in each panel (A and C), two different treatments were considered (drug 1 and drug 2), in a range of drug concentrations, for cells with different resistance levels. Panels B and D illustrate the population response during the treatments, and the density is plotted as a function of the drug dose and time. The shapes of the density curves vary according to the type of treatment and type of drug-induced death rate that were assumed. In all cases, a better treatment would give lower cell density at a lower drug dose (α^*), with a wide range of effective concentrations.	77
4.4	Heterogeneity arising from genetic/epigenetic changes. (A) numerical predictions of the net growth ($n(x, t)$) after a certain time ($t = 12.5$), when no chemotherapeutic drugs were administered. The effects of genetic/epigenetic changes act as a diffusion process, where $\epsilon_0 = 0, \epsilon_1 = 0.01, \epsilon_2 = 1$ with two initial conditions (solid and dashed lines, IC_1 vs. IC_2). The initial conditions are shown in Fig. 1C. This graph shows the cases in which the diffusion potentially controls the tumor dynamic and its heterogeneity. (B) and (C) net growth ($n(x, t)$) plotted at three time points: $t_1 = 2.5; t_2 = 5; t_3 = 12.5$, with initial condition IC_1 . The administration of the drug takes place over a certain period of time ($t_1 < t < t_2$), for a given dose. These graphs demonstrate the effect of varying the alteration rates on the dynamic in the presence of anticancer drugs ($\epsilon_1 = 0.01$ in (B), $\epsilon_2 = 1$ in (C)).	81
4.5	Dynamics with variations in the mutation rates over time with/without drugs. In all of these cases, there is an initial alteration (ϵ_1) fixed over time, and an additional alteration rate (ϵ_2) applied only when the drug (C_1) was applied for a certain period of time ($t_1 < t < t_2$, where $t_2 < t_3$). (A) $\epsilon_1 = 0.01, \epsilon_2 = 1$. (B) $\epsilon_1 = 0.01, \epsilon_2 = 0.01$. (C) $\epsilon_1 = 1, \epsilon_2 = 1$. (D) $\epsilon_1 = 1, \epsilon_2 = 0.01$	83

- 4.6 Comparison between the Stein and Lorz models, with resistance-level variables that are discrete vs. continuous. (A) Patient response data and two mathematical predictions of the total number of cancer cells during treatment are plotted. The data set (boxes) illustrates a representative relapse response (case 12, Fig. S1 in [1]). The Stein model, given in (4.2), describes the discrete approach of grouping the cells into two categories: resistant and sensitive (dashed line). The Lorz model, given in (4.3), describes the continuous approach to resistance level (solid line). (B) The cellular rates of cell division ($r(x)$, black solid line) and drug-induced death ($c(x)$, black dashed line) are assumed to be a sigmoid curves and the natural death rate is assumed to be a low constant value. (C) The trait distributions, $n(x, t)$, at the beginning ($t = 0$) and at the end ($t = 691$ days) of treatment. This plot is the solution of the exponential growth model described in equation (4.3), where $\theta = 0$ and $\epsilon = 0$, using initial condition IC₁ (see Figure 4.7(B)) and cellular rates shown in (B). 93
- 4.7 The Lorz model of exponential growth with four different initial trait distributions. (A) Four mathematical predictions of the total number of cancer cells over time, ($\rho(t)$ where $t \leq 691$ days), based on four different initial conditions (IC₁-IC₄). These curves are the solutions of (4.3), where $\theta = 0, \epsilon = 0$ along with the rates in Fig. 4.6B. (B) The four different initial conditions, labeled as IC₁-IC₄. (C) The trait distributions at the end of the treatment, $n(x, t = 691 \text{ days})$, using the same conditions as in panel A. (D-E) These two panels describe the effect of variations in the net-growth rates on the Lorz model. Using the same conditions as in Figure 4.6 (continuous-resistance model), we illustrate the outcome of the perturbed net-growth rates of equation (4.3) using 1000 simulations. The additional perturbation distribution assumes a normal distribution with mean 0, standard deviation 25% of the net-growth rate in Figure 4.6A applied uniformly over all traits. (D) shows the variation in the growth of cancer cells over time, while (E) shows the variation in the trait profile at the final time (691 days). The different colors are the result of different perturbations corresponding to the aforementioned normal distribution, while the black line corresponds to the simulation in Figure 4.6 (that is, no perturbation). The arrows stress the size of variations. 97

4.8	<p>Modeling cell growth with density dependence. As in Figure 4.7, the effects of four initial conditions on the density and final trait distribution of the third model are shown. (A) Four mathematical predictions of the total number of cancer cells over time, based on IC₁-IC₄. These curves are the solutions of equation (4.5), with $\theta = 0$ and $\epsilon = 0$ and rates as in Figure 4.6B. (B) The trait distributions at the end of the treatment ($t = 691$), using the same conditions as in panel A. (C) Lower plot: numerical predictions of the final cell density, $\rho(t = 691)$, using (4.5), where ϵ and θ vary over a range of parameters $[0, 1]$. All other parameters and functions are the same as in panels A and B. (D) Upper plot: as in panel C, ϵ and θ vary over a limited range of parameters, now in $[0, 0.1]$.</p>	101
5.1	<p>Model dynamics. Diagram of transitions between the three cellular compartments in the ABM. Q denotes the quiescent compartment with $N_q(t)$ cells at time t, P denotes the proliferation compartment with $N_p(t)$ cells at time t, and A denotes the apoptosis compartment with $N_a(t)$ cells at time t. Note that $N_q(t)$, $N_p(t)$, and $N_a(t)$ are all stochastic processes. The explicit transition rates between the compartments are shown in solid lines, and are labeled as $\alpha_p(t)$, $\alpha_{a_p}(t)$, and $\alpha_{a_q}(t)$. The implicit transition rates, due to the completion of cellular cycles, are shown in dotted lines, and have no closed-form expression. For example, $\Sigma_{P \rightarrow Q}$ corresponds to the rate of cell-cycle completion. The line originating from compartment A indicates cells that are removed from the simulation.</p>	109
5.2	<p>Experimental results with functional forms for cell-cycle and apoptosis fractions. Experimentally (blue and black), OVCAR-8 cells are allowed to attach to the plate for twenty-four hours, and then are incubated for a subsequent twenty-four hours. After this incubation, the cells were stained to determine which percentage of cells reside in the proliferation and apoptosis stages. The experiment was performed using two different techniques for cell synchronization: Data 1 and Data 2 (see C.1.1.1 for details). The functional forms (red) correspond to equations (5.4) and (5.3), with free parameters taking values $\beta_m = 0.75$, $\rho_m = 0.15$, $\epsilon = 1$, and $d = 0.03$. (a) Fraction of cells in division stage (P) as a function of the population density on the plate; (b) Fraction of cells in apoptosis stage (A) as a function of the population density on the plate. Note that we allow $\rho > 1$.</p>	111

5.3	(a) The mean and standard deviation of $\alpha_p(t)$ over all times $t \in [0, 96]$ hours, for 1000 Monte Carlo simulations. β and d appear as the red curves in Figures 5.2(a) and 5.2(b) respectively, and $\sigma = 3$ hours, $c = 1$ cell per hour. See equations (5.51) and (5.52) for the two sets of initial conditions. Note that the standard deviations are roughly an order of magnitude smaller than the mean values, for most times. (b) Error obtained in using the first term on the RHS in equation (5.27) as an approximation for $\alpha_p(t)$	121
5.4	Growth dynamics of OVCAR-8 human ovarian carcinoma cell line. Global density is measured on the vertical axis as a function of time. Both initial densities ($\rho(0) = 0.1$ and $\rho(0) = 0.8$) were generated by random seeding on a 24-well plate and cultured 24 hours prior to the first image. Data was collected every 24 hours, as indicated by stars. No data exists between the points, and the connecting line segments are added merely for viewing elucidation. The estimated density percentage is based on the average of two complete wells. The red curve represents a least-squares fit to the region of exponential growth, with doubling time $\mu = 24.4416$ hours.	131
5.5	Frequency distributions for the apoptosis process. The x -axis represents the time in the respective phase, while the y -axis labels a fraction of cells. For the experimental results (histogram in (a)-(c)), the x value indicates a range of times, so that for $x = i$ hours, the corresponding y -value (height of the bar) counts the relative frequency of cells with phase length between $i - 1$ and $i + 1$ hours. For the theoretical distribution (black curve), we allowed the phase length to be a continuous random variable, and fitted a gamma distribution with probability density function (PDF) using MATLAB's "fitdist.m" function. In (d), we used the PDFs obtained in (a)-(c) to obtain a single gamma distribution for the entire apoptotic process. (a) Phase 1, with $n = 56$ total cells measured; (b) Phase 2, with $n = 47$ total cells measured; (c) Phase 3, with $n = 49$ total cells measured; (d) Total length of time spent in apoptosis, using $n = 10^5$ simulations of (5.50).	134
5.6	First and second moments of the ABM. Shown are the temporal dynamics of the mean and variance of 1000 Monte Carlo simulations. β and d appear as the red curves in Figures 5.2(a) and 5.2(b) respectively, and $\sigma = 3$ hours, $c = 1$ cell per hour, $\gamma = 0.70$. Note that these parameters do not yield an optimal fit. See equations (5.51) and (5.52) for the two sets of initial conditions. In (b), the values of N_a for both sets of initial conditions essentially overlap, making them difficult to distinguish. The same is true for the variances of N_p and N_q in (c) and (d). (a) Sample means for the population density; (b) Sample means for the three cellular compartments; (c) Variances for simulations that satisfy $\rho(0) = 0.1$; (d) Variances for simulations that satisfy $\rho(0) = 0.8$	136

5.7	Comparison of ABM and IDE simulations. 1000 Monte Carlo realizations are plotted together with the corresponding IDE. β and d appear as the red curves in Figures 5.2(a) and 5.2(b) respectively, and $\sigma = 3$ hours, $c = 1$ cell per hour, $\gamma = 0.70$. (a) Cellular density for both sets of initial conditions ($\rho(0) = 0.1$ and $\rho(0) = 0.8$); (b) Compartment fractions for $\rho(0) = 0.1$; (c) Compartment fractions for $\rho(0) = 0.8$	139
5.8	Cellular growth dynamics for varying parameter sets. (a) A nonlinear least-squares algorithm was used to locate parameter values which minimize the difference between the experimental data and the IDE (equations (5.18)-(5.20), together with (5.9)-(5.10)). Due to the sparsity of the data (red), multiple local minimums were found. We note that all curves plotted have residuals with ℓ^2 norm $\in [0.0583, 0.0980]$, and are displayed in descending order by this measure. Similar plot styles in (a) and (b) correspond to the same parameter set. (a) Growth of cellular density in time; (b) Division (β) and death fractions ($d = \beta(1)$) for the fitted parameter sets. (c)-(d) Dependence of IDE solutions on cell-cycle standard deviation σ . σ is varied, while all other parameters are fixed from a local minimizing set located in (a) and (b). We plot the global density growth and cell-cycle fraction for each case. (c) Cellular density ρ for σ variation; (d) Proliferative compartment for σ variation.	141
5.9	Dependence of IDE solution on c . All parameters, excluding c , are fixed and obtained from an optimization appearing in Figure 5.8 (green curves). The parameter values are $\beta_m = 0.8, \rho_m = 0.2, \epsilon = 0.33, d = 0.01, \gamma = 0.5$ and $\sigma = 8.92$ hours. The black curve represents the local minimizer. (a) Cellular density ρ ; (b) Proliferative compartment; (c) Magnification of $t = 0$ dynamics of proliferative compartment; (d) Apoptotic compartment; (e) Quiescent compartment.	146
5.10	Dependence of IDE solution on ϵ . All parameters, excluding ϵ , are fixed and obtained from an optimization appearing in Figure 5.8 (green curves). The parameter values are $c = 12.19$ cells per hour, $\beta_m = 0.8, \rho_m = 0.2, d = 0.01, \gamma = 0.5$ and $\sigma = 8.92$ hours. The black curve represents the local minimizer. (a) Cellular density ρ ; (b) Proliferative compartment; (c) Apoptotic compartment; (d) Quiescent compartment.	147
6.1	Attraction and repulsion potentials used in the SDE model (6.1), via equations (6.4) and (6.5). (a) Attraction force potential, with local detection radius $\epsilon_a = 5$. (b) The repulsion potential used throughout this chapter.	153

6.2	Schematic representation of division/apoptotic/quiescent process. The model includes three compartments: proliferating cells (P), quiescent cells (Q), and apoptotic cells (A). Cells move stochastically through the three phases, with explicit rates represented by solid lines, and implicit rates given by dashed lines.	155
6.3	Transition rates α_p and α_{a_q} , with parameters values $\alpha_{\max} = 2.8, \rho_m = 0.2, \rho_r = 0.85$, and $d = 0.03$	157
6.4	The cell-cycle rate (equation (6.11)) for parameter values $p = 0.5, c_0 = 2.29$, and $\omega_c = 0.7931$	159
6.5	Sample initial and final configurations of the ABM used for comparison with 24 hour image control data. (a) and (c) are derived from the experimental data at $t = 0$ hours (Figures 6.6(a) and 6.6(c)), while (b) and (d) are calculated from the ABM after a period of 24 hours, and are compared to corresponding experimental data (Figures 6.6(b) and 6.6(d)). (b) corresponds to the initial conditions (a), while (d) corresponds to the initial conditions (c). The color scheme represents the three cellular compartments: yellow denotes quiescence, red denotes proliferation, and blue denotes apoptosis. $\epsilon_a = 5, \theta = 0.2, p = 0.5$, and all other parameters appear as in Figure 6.3.	164
6.6	Initial and final configurations of the experimental data used for model comparison. Note that all cells are marked with red fluorescence, independently of their position in division or apoptosis. (a) Data set 1 initial conditions. (b) Data set 1 after 24 hours of cell growth. (c) Data set 2 initial conditions. (d) Data set 2 after 24 hours of cell growth.	165
6.7	Spatial deviation parameter search. We plot the ℓ^1 spatial pixel distance between the experimental data and the simulations for a range of parameters ($\epsilon_a \in [2, 9], \theta \in [0, 0.4]$) for each data set. Also plotted are the minimums for each value of θ producing a small error (dots) and the least-squares line $\tilde{\epsilon}_a$ connecting these minimums (line). (a) Data set 1: $\tilde{\epsilon}_a = -17.714\sigma + 7.381$; (b) Data set 2: $\tilde{\epsilon}_a = -8.571\sigma + 9.24$.	166
6.8	Comparison of growth data of OVCAR-8 human ovarian carcinoma cell line to model output. Experimental data (red) is compared to 10 Monte Carlo simulations with growth parameter values $\beta_m = 2.8, \rho_m = 0.2, \rho_r = 0.85$, and $p = 0.5$	167
6.9	10 Monte Carlo long-time simulations of the ABM. (a) Cellular density growth analogous to Figure 6.8, where the mathematical model satisfies $\rho(0) = 0.1$ and $\rho(t = 145 \text{ hours}) = 1$. (b) The dynamics of the different compartments. Note that the apoptotic fraction (N_a/N) remains relatively constant, except for large time, when the mitotic checkpoint begins to take effect. (c) Fraction of apoptotic cells coming from Q (rate α_{a_q}) and P (rate α_{a_p}). Initially all apoptotic transitions originate from compartment Q, while as time and density increase, the percentage of transitions from P increases as well.	169

6.10	Cell-cycle distributions generated by local density. (a) Histogram for cell-cycle lengths based on the time interval of division. Division intervals are 10 hours in length. The red 28.5 denotes mitotic threshold length ($0.95\mu/p$) after which cells must transfer from P to A. (b) Mean of division lengths as a function of the occurrence interval. (c) Standard deviation of division lengths as a function of the occurrence interval.	170
6.11	Initial <i>in vitro</i> spatial configuration of cells for the killing curve experiment. The figures show the distribution of cells in a two-dimensional environment initialized in the ABM, which qualitatively agrees with the experimental seeding. All cells initially reside in compartment Q (yellow). (a) Configuration 1: $\rho(0) = 0.1$, uniform random on entire plate. (b) Configuration 2: $\rho(0) = 0.8$, uniform random on entire plate. (c) $\rho(0) = 0.1$, uniform random on compact ball.	171
6.12	Experimental killing curves for three geometrical configurations 1, 2, and 3. Relative cell survival (compared to the control case) is measured after three days of treatment as a function of dosage (x -axis). Note that the dense geometries (both on the entire plate 2 and locally in 3, see Figure 6.11) experience on overall greater survival rate, and thus are drug resistant.	172
6.13	Killing curves for three geometrical configurations 1, 2, and 3 computed from the ABM. Relative cell survival (compared to the control case) is computed after three days of treatment as a function of dosage (x -axis). Note that the dense geometries (2 and 3) experience on overall greater survival rate, and thus are drug resistant. The qualitative features of the model output presented here should be compared to the corresponding curves in Figure 6.12.	173

List of Abbreviations

A	Apoptotic compartment
ABM	Agent-based model
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
ABC	ATP-binding cassette
ECM	Extra-cellular matrix
IC	Initial condition
IC ₅₀	50% inhibitory concentration
IDE	Integro-differential equation
MDR	Multi-drug resistance
NIH	National Institutes of Health
ODE	Ordinary differential equation
OS	Overall survival
P	Proliferation compartment
Q	Quiescent compartment
PDE	Partial differential equation
PDF	Probability density function
P-gp	P-glycoprotein
SDE	Stochastic differential equation

Chapter 1: Introduction

The goal of this dissertation is to combine mathematical modeling with experimental data to study the emergence of drug resistance to chemotherapeutic agents. In this chapter, we overview the characteristics of cancer as a disease, and briefly discuss the related phenomena of tumor heterogeneity and multi-drug resistance (MDR). This introductory chapter will serve as the biological basis for the mathematical models developed later in this work.

1.1 A brief introduction to cancer and chemotherapy

In the United States, and throughout the world, cancer is a major public health concern. It accounts for one in four deaths in the U.S., where, as of 2015, it is the second leading cause of death [2]. Cancer is a collection of diseases, each of which is characterized by defects in the cellular pathways which govern proliferation [3, 4] and apoptosis [5]. It is thought that healthy cells undergo a series of genetic mutations [6–9] as well as epigenetic events which transform the cellular genotype/phenotype, allowing them to bypass homeostatic conditions, thus leading to uncontrolled growth. This growth can be harmful in a number of ways, and potentially leads to patient death via disruption of vital organ function [10], fatal

metastasis [11], and/or cachexia [12].

Although there are more than one hundred different types of cancer [13], affected cells appear to exhibit a number of common traits. We briefly discuss these traits here, which are summarized in the seminal works of Hanahan and Weinberg [14,15]. Indeed, an understanding of the basic characteristics of cancerous cells will be valuable in formulating assumptions for the mathematical models that will be presented in later chapters. The most well-known trait which cancer cells possess is the deregulation of cellular proliferation. Normal cells typically require signals from growth factors, neighboring cells, diffusing substrates, and the extra-cellular matrix (ECM) to enter a division cycle; however, many types of cancer cells require no such external signalling, and instead receive proliferation signals from mutated intracellular oncogenes. Analogously, resistance to anti-growth signals is also observed in most malignant cells. Furthermore, cancer cells may develop the capability to complete an infinite number of replications, making them essentially immortal. Normal cells experience telomere (i.e. chromosomal end) degradation after each successful division, which limits the total number of daughter cells they are able to produce in their lifetime. However, malignant cells possess a chromosomal recombination repair mechanism which allows them to circumvent this limit, and divide indefinitely. Besides the induction of excessive and unlimited growth, cancer cells also experience an overall reduction in cell death. For example, mutations in the p53 apoptotic pathway, possibly the most well understood tumor suppressor gene, are observed in over 50% of human cancers, and cause affected cells to evade apoptosis signaling in cases of DNA damage and hypoxic microenvironments. Cancer cells also exhibit

the ability to promote angiogenesis, i.e., the growth of new blood vessels, which is often necessary to sustain the large tumor populations produced when proliferation and apoptosis signaling is ignored. Lastly, most cancer cells also invade tissue and metastasize to other sites in the body, which results in the vast majority of cancer fatalities [11].

Today there exist four main cancer treatment strategies: surgery, chemotherapy, radiotherapy, and immunotherapy. Although surgical removal of the tumor is the most common method and first line of treatment for many cancers, it is often combined with chemotherapy and/or radiotherapy to kill remaining cancerous cells. Furthermore, many types of cancers cannot be removed via surgical procedures, which renders chemotherapy as an invaluable treatment option. Chemotherapy fails for a number of reasons; the most important is likely due to the development of drug resistance [16]. To understand this phenomenon, we next discuss what chemotherapy is, as well as its main mechanisms of action.

A chemotherapeutic agent is any chemical compound used to decrease a tumor load. Modern chemotherapy dates back to the 1940s, where an accidental spill of mustard gas in Bari Harbor, Italy, led to observations of depleted bone marrow and lymph nodes among the exposed [17]. Today, over 100 different chemotherapeutic agents exist, and are often applied together in combination therapies. Such treatments can be administered simultaneously or sequentially, an area to which mathematics has already contributed (see Chapter 2).

Chemotherapeutic agents can be classified as either cell-cycle nonspecific or cell-cycle specific [18]. Cell-cycle nonspecific drugs typically respond linearly as

a function of dosage, while cell-cycle specific agents exhibit a bounded response.

Furthermore, chemotherapies can be broadly classified via mechanism as below:

1. *alkylating agents*: impair cell function by bonding with amino, carboxyl, sulfhydryl, and phosphate groups in biologically significant molecules. Damage is typically done to the DNA molecules, and there is a high risk of developing leukemia at large doses. Examples include nitrogen mustards, nitrosoureas, and platinum complexes.
2. *antimetabolites*: interfere with the building blocks of DNA synthesis as structural analogs of natural metabolites. As they interfere with genetic synthesis, they are primarily active during the S phase of the cell-cycle. Examples include methotrexate, fluorouracil, and pentostatin.
3. *antitumor antibiotics*: interfere with enzymes involved in DNA replication, for instance by producing free oxygen radicals which cause strand breakage. Large doses can cause pulmonary disease. Examples include anthracyclines and bleomycin.
4. *topoisomerase inhibitors*: inhibit topoisomerase activity, resulting in the inability of the cell to synthesize DNA. Examples include topotecan and etoposide.
5. *mitotic inhibitors*: inhibit or stabilize tubulin assembly via rapid binding, thus blocking the polymerization of microtubules. Examples include the taxanes, epothilones, vinca alkaloids, and estramustine.

Other miscellaneous chemotherapeutic drugs exist, such as targeted (which attack cells based on genetic information such as cellular mutations [19–21]) and hormonal (which prevent cancer cells from obtaining hormones necessary for growth [22]) therapies, but the above list classifies the vast majority of chemotherapeutic agents.

Note lastly that the drugs in the list above act exclusively on cells in the cell-cycle. In fact, they do not distinguish between normal and cancerous cells, and rely on the accelerated division process of the disease. Thus, a fundamental principal guiding most chemotherapies is that the treatment is anticipated to affect cancer cells to a greater extent than normal cells. Of course, healthy cell death can have adverse side effects [23], which in combination with chemotherapeutic resistance often limits treatment efficacy.

1.2 Tumor heterogeneity

Much of the work in this dissertation is related to the observation that cells from a single tumor exhibit a range of responses to the same treatment. Thus, to understand the phenomena of drug resistance, it is necessary to first understand the different cellular genotypes and phenotypes that compose a cancer cell population. This section briefly reviews the experimental findings related to tumor heterogeneity, and how they relate to chemotherapeutic resistance.

The development of a tumor is a complex evolutionary process that involves perturbations in many essential cellular mechanisms. Furthermore, such perturbations are acquired through purely stochastic events and are also induced by en-

vironmental signals, including metabolic stress, inflammatory micro-environments, immune responses, and/or therapy [24]. Combinations of such signals produce a tumor that is heterogeneous, on both the genetic [25,26] and epigenetic [27–30] level. For example, many primary human tumors have been discovered to contain genetically and phenotypically distinct cellular sub-populations with different growth rates [31] and different metastatic potentials [32]. These, together with other manifestations of tumor heterogeneity, have been found to be a major contributor of drug resistance in chemotherapy [33,34].

In addition to cancer-related irreversible processes, heterogeneity can also arise via typical reversible biological processes that are stochastic, yet tightly regulated, in nature. Such mechanisms are intrinsic; they are not related to mutations or other cancer-specific phenomena, but instead are present in normal cell mechanics. Nevertheless, such intrinsic mechanisms add another layer of complexity to a cell's capacity to integrate information. Intra-cellular dynamics are mediated by many different molecular components (e.g. transcription factors, proteins, metabolites, RNA, etc.). Each such component operates at a different rate, often under different conditions, and responds to many dynamic inter- and intra-cellular signals, such as pH, temperature, and cellular density in the local environment. In order to maintain homeostasis despite a routinely noisy micro-environment, variations in gene expression [35], cell-cycle period [36], cell size and age [37], and cellular death period [28,38] of cells from the same clone must exist.

In this dissertation, the example of intrinsic heterogeneity we will be primarily interested in is the cell-cycle. The cell-cycle is one of the most studied biological

processes, and has obvious effects on cancer development, growth, and therapeutic resistance. Indeed, as most chemotherapies act on dividing cells (see Section 1.1), it is important to understand this dynamic from a therapeutic standpoint. For instance, being able to predict when the fraction of proliferating cells in the S phase is highest would provide clinicians with important information on the treatment protocol for an S-phase specific cytotoxic agent. Furthermore, if all stages of the cell-cycle can be calculated likewise, optimal combination treatment strategies can be formulated.

Tumors have many sources of heterogeneity, originating from induced mechanisms relating to the evolution of the disease, as well as from intrinsic mechanisms present in healthy cells. Understanding this heterogeneity and the role it plays in controlling the outcome of chemotherapy is of vital importance to overcoming drug resistance.

1.3 The biology of drug resistance

In this section, we briefly review the genetic and epigenetic origins of resistance to chemotherapy, as well as a partial list of the most well-studied mechanisms by which resistance presents.

1.3.1 Developmental origins

Drug resistance can arise from a variety of sources. As discussed in Section 1.2, cells can become resistant based on intrinsic factors, such as their location in the

body relative to the blood supply or their position in the cell-cycle relative to the chemotherapeutic schedule. Such mechanisms are not due to genetic variation, but rather to physiological conditions affecting the cells and their surrounding micro-environments. However, it is also important to understand the genesis of the explicit variations observed in differentiating wild-type sensitive and multi-drug resistant cancer cells.

Resistance due to explicit variations may develop via two main genetic events: point mutations and gene amplification. A point mutation occurs during cell division, and results in the replacement of a single base nucleotide with another nucleotide in either the DNA or RNA. Such an event is random, occurring with a small, nonzero probability, and is usually permanent and furthermore inherited by all future daughter cells. Point mutations alter the cellular genotype, and thus the phenotype, and can confer resistance through many different cellular mechanisms [39–42] (see Section 1.3.2 below). This pathway is the most well understood, and is the most commonly utilized mechanism for heterogeneity in mathematical models. Gene amplification, on the other hand, is a cellular process characterized by the overexpression of a region of DNA which contains one or more genes. Such an event provides the cell with additional copies of genetic material and thus alters the phenotype of the cell, which, in turn, can confer drug resistance through a variety of mechanisms [43–46]. Gene amplification, unlike point mutations, are typically non-hereditary, and may not even be permanent for the affected cell.

It is also important to note that drug resistance can either be a spontaneous or an induced event, where in the latter the applied cytotoxic agent causes resistance to

develop. Indeed, there is experimental evidence that such mutagenic drugs exist [47, 48]. Furthermore, drug resistance can be permanent or temporary, as well as relative or absolute. Here relative and absolute apply only to mutagenic drugs, and signifies that resistance is dependent or independent, respectively, of the applied dosage.

1.3.2 Mechanisms

Once a genetic or an epigenetic event occurs, a cell can become resistant through one or more specific mechanisms. From a mathematical perspective, it is often important to consider the precise characteristic (or combinations thereof) through which a cell expresses drug resistance. Indeed, understanding the mechanism and the correct method of mathematical implementation is a key step in the modeling process. Below, we present a survey of some of the most well-studied phenomena leading to drug resistance. For a more comprehensive list, we refer to Gillet and Gottesman [49], and the references therein. Note also that the mechanisms listed below can either act individually or synergistically, with the latter often resulting in MDR.

For a chemotherapeutic agent to be effective, it must first penetrate the cell membrane. Thus, the alteration of uptake transporters resulting in a decreased intracellular drug concentration can mediate resistance, as has been demonstrated clinically for SLC19A1/hRFC1, a member of the solute carrier genetic family [50]. As the uptake of chemotherapeutics is still not well understood, little can be concluded with certainty in regards to the entire family; although there is mounting evi-

dence to suggest that solute carriers play an important role in drug resistance [51,52]. Related to drug uptake is drug efflux, the transport of chemotherapeutics to the exterior of the cell. An increase in this transport rate will decrease the total concentration of drug that can reach the nucleus, hence reducing the drug's overall cytotoxic effect. This phenomenon is clinically well-established as a drug resistance mechanism, specifically relating to alterations involving the family of ATP-binding cassette (ABC) transporter proteins [49,53–55]. It is worth noting that other ATP-dependent drug efflux transporters have been shown to mediate resistance as well [56,57].

Additional mechanisms that can reduce the efficacy of treatment exist. Recent evidence suggests that drug sequestration, the process by which drug molecules are prevented from reaching target-specific regions via storage in lysosome, melanosome, golgi, and secretory intracellular compartments, is a major contributor to drug resistance in cancer treatment [58–61]. Drugs may also be metabolized before they enter the nucleus and induce cell death. For example, Gluatathione-S-transferases, a family Phase II detoxification enzymes, have been implicated in development of resistance via direct detoxification, as well as MAP kinase pathway inhibition [62]. Standard cancer “hallmarks” can also decrease the efficacy of some anti-cancer agents; it has been clinically observed that the response to platinum and alkylating agents can be substantially reduced in cancer cells with heightened DNA-damage repair mechanisms [63,64].

1.4 Structure of the dissertation

In this dissertation we develop mathematical models of tumor heterogeneity and drug resistance in cancer chemotherapy. In Chapter 2, we review some existing mathematical models related to drug resistance.

In Chapter 3, we develop a basic framework to accurately describe MDR via mathematical models. Our fundamental assumption that resistance should be viewed as a continuous (as opposed to discrete) variable arises from the complexity of drug resistance and the tumor micro-environment, and leads us to define a structured population model taking the form of a system of integro-differential equations (IDEs). The model includes mutations, as well as arbitrary density dependencies to account for a variety of growth dynamics. We analyze this model and its limiting distribution as a selection process, and identify the formulation which best describes tumor heterogeneity.

In Chapter 4, we apply the theoretical techniques developed in Chapter 3 to clinical scenarios. In Section 4.1 we generalize our IDE structured population model of Chapter 3 to include epigenetic effects, and discuss possible treatment strategies which minimize both intratumoral heterogeneity and overall tumor load. In Section 4.2, we apply discrete and continuum models to clinical response and relapse data. Here, we use patient data sets to argue in favor of modeling resistance as a continuum variable, and demonstrate how integrating all basic growth elements improves predictions concerning the resistance heterogeneity of metastases.

Understanding the origins and dynamics of tumor heterogeneity is a necessary

step to overcome drug resistance. In Chapter 5, we develop a mathematical model to describe cancer cell growth in the presence of intrinsic heterogeneity as discussed in Section 1.2. Here, heterogeneity arises from the stochastic nature of the cell-cycle and apoptosis process. We derive two novel mathematical models, an agent-based model (ABM) and an IDE, to describe the dynamics introduced by incorporating the cell-cycle into a mathematical model. We show that when the number of cells is large, the ABM is effectively approximated by the computationally efficient IDE. By analyzing the role all parameters play in the evolution of heterogeneity, we discover that the cell-cycle length has the most significant effect on the growth dynamics.

In Chapter 6, we extend the work of Chapter 5 to explain the variability of the cell-cycle via spatial mechanisms. Our main assumption is that the cell-cycle length increases as a function of local density measurements, so that spatial heterogeneity produces a distribution of cell-cycles in the observed 2D cultured cell environment. This effect, together with the principle cytotoxic mechanism of the chemotherapeutic (Paclitaxel) studied by our collaborators at the National Institutes of Health (NIH), are incorporated into the ABM discussed in Chapter 5. We show that such an assumption generates a traveling wave of approximately normal distributions describing the population cell-cycle distribution. We also show that such a model qualitatively predicts the structure of the experimental killing curves.

Concluding remarks are provided in Chapter 7.

Chapter 2: The mathematics of drug resistance - a brief overview

In this chapter, we review a subset of the existing mathematical literature relating to drug resistance and cancer chemotherapy. We emphasize that the discussion is limited to modeling techniques, and does not include “big data” approaches such as machine learning and other statistical analysis methods, as this dissertation deals only with the mathematical modeling aspects of drug resistance. For an excellent example of such techniques related to cancer classification, see [65].

This chapter features a series of subsections, each addressing a specific issue related to drug resistance and mathematical modeling. It closely follows the work of Lavi *et al.* [66], with a greater exposition of the mathematical details.

2.1 Probability of resistant cells existing at diagnosis

One of the fundamental questions related to drug resistance is the following: how likely is it that at the time of detection, a tumor contains resistant cancer cells? This question is closely related to another question: is drug resistance primarily a treatment-induced phenomenon, or rather a spontaneous genetic event dictated by purely random events? Such questions arise not only in chemotherapy and cancer, but in the related field of bacterial resistance to antibiotics. Most famously, this

issue was studied by Luria and Delbrück [67] in 1943, a work for which they were later awarded the 1969 Nobel Prize in Physiology or Medicine. In [67], a number of *Escherichia coli* bacteria were plated with T1 phage, and after a period of time, the number of virus resistant colonies were counted. This experiment was performed on many plates, and the distribution of the resistant colonies was studied. In the hypothesis of induced resistance, i.e. if the selective pressure of the virus directs bacterial evolution, the number of resistant colonies should be Poisson distributed, with *equal mean and variance*. If, on the other hand, mutations occurred randomly in each generation at a constant rate independently of the action of the virus, the resistant colonies are distributed according to the so-called Luria-Delbrück distribution, with mean μ and variance σ^2 given by

$$\sigma^2 = \mu \frac{aCN}{\log aCN}. \quad (2.1)$$

Here N is the number of observed bacteria, a is the mutation rate, and C is the number of sampled cultures, so that aCN is the total number of mutations in all experiments. Note that as long as sufficiently many samples are considered, the *variance should be appreciably larger than the mean*. This is precisely what Luria and Delbrück observed, thus concluding that mutations are random, rather than directed.

Similar questions have been raised in cancer research. Iwasa *et al.* [68, 69] mathematically studied the probability that resistant tumor cells exist at clinical detection as a function of tumor size M . Using branching processes theory, they

derived the generating function for the number of resistant cells at detection, $G(\xi)$, and thus the probability P that at least one such cell exists at detection:

$$P = 1 - G(0) = 1 - \exp\left(-Ma\frac{r}{d}\log\frac{r}{r-d}\right). \quad (2.2)$$

The derivation of (2.2) assumes that sensitive and resistant cells are equally fit. Here r and d are the birth and death rates, respectively, of the resistant cells, and a is the mutation rate as in (2.1). In general, these results are pessimistic, suggesting that resistance is almost always certain at detection (e.g., a typical M at detection is on the order of 10^9 , and a conservative estimate of $a \approx 10^{-8}$ yields $P \approx 1$). Similarly, in their original paper on chemotherapeutic drug resistance, Goldie and Coldman [70] reach an analogous conclusion: $P \propto 1 - \frac{1}{M}$. The results of [68,69] were extended by Tomasetti *et al.* [71] to populations of stem cells, accounting for different differentiation paths. A related approach was taken by Komarova *et al.* [72,73], combining probability theory with hyperbolic partial differential equations (PDEs) to address the question of drug resistance at detection.

2.2 Combination versus sequential therapy

It is often the case that multiple drugs are used for a specific disease. A natural question then arises: how should these therapies be combined to construct a complete treatment protocol? For instance, should the cytotoxic agents be applied sequentially or in combination? In general, the goal is to prevent the emergence of a cross-resistant cell line, i.e., cells that are resistant to all treatment options.

Often, combination therapies are not feasible due to toxicity constraints, so that physicians must use sequential scheduling. However, quantitative questions remain, such as the rate of alternating the individual drugs, as well the administration of rest periods. Such issues are well suited for mathematical inquiry, and have been studied extensively. In this section, we present a sample of such works.

Goldie and Coldman [74] first studied the problem of combination and sequential therapy administration in 1982. They assumed two non-cross resistant drugs, hence giving rise to three types of cancer cells: sensitive (S), resistant to drug #1 (R_1) and resistant to drug #2 (R_2). Treatment is presented in cycles, where there are alternating periods of cell growth (no treatment) followed by cell death (treatment). Cell type S is affected by either drug while, for instance, type R_1 is only affected by drug #2, and there is no cross resistance between R_1 and R_2 . Both drugs are assumed to follow the log-kill law, i.e. the probability of survival, q , as a function of dose D satisfies

$$\log q = -\beta D, \tag{2.3}$$

where $\beta > 0$ is a constant. All cell types grow and die stochastically with the same rate parameters. Mutations occur during cell division with constant rates: α_1 for drug #1 and α_2 for drug #2. Thus, a cell is doubly resistant and hence unaffected by treatment if it has undergone two successive mutations, via pathways $S \rightarrow R_1 \rightarrow R_2$ or $S \rightarrow R_2 \rightarrow R_1$. Using basic techniques from stochastic processes, Goldie and Coldman derived expressions for the probability of having no cross resistant cells

in a tumor of size N , $P_{1,2}(N|N_0)$, assuming no resistant cells existed at a smaller population N_0 :

$$P_{1,2}(N|N_0) = \exp \left(-\alpha_2[R_1(N) - R_1(N_0)] - \alpha_1[R_2(N) - R_2(N_0)] + 2\alpha_1\alpha_2(N - N_0) \right). \quad (2.4)$$

Here $R_1(\tilde{N})$ and $R_2(\tilde{N})$ are analytic formulas for the the expected resistant population sizes as a function of the total tumor mass ($\tilde{N} = S + R_1 + R_2$). Using simulations of drug cycling procedures while minimizing (2.4), Goldie and Coldman concluded that if possible, combination therapy was optimal. More interestingly, they also discovered that over all non-combination protocols, strictly alternated chemotherapy provides the smallest value of $P_{1,2}(N|1)$. Realizing the effect of mutations as a filtered Poisson process, more general analytical results were later proven by the same authors in [75], with recent extensions by Chen *et al.* [76].

Combination and sequential therapies in induced drug resistance were also studied by Panetta [77]. As in [74], therapy is cycled over two drugs, labeled A and B . Cells are modeled deterministically using both continuous and discrete dynamics. Between treatments, cells grow exponentially, whereas therapy instantaneously reduces the number of cells. Explicitly, denoting the number of cells sensitive to treatment A by x_A and those to treatment B by x_B , the overall effect of the simul-

taneous combination therapy $\{A, B\}$ is given by

$$x_A^+ = [f(D)(1 - R(D))]x_A^-, \quad (2.5)$$

$$x_B^+ = \tilde{f}(D)x_B^- + f^\alpha(D)\tilde{f}^{1-\alpha}(D)R(D)x_A^-, \quad (2.6)$$

where $+(-)$ superscripts denote cell masses just after (before) the chemotherapeutic dosage. D denotes the (equal) dosage of A and B , and f represents the survival fraction of cells sensitive to drug A . \tilde{f} is the survival fraction of cells resistant to drug A , but sensitive to drug B . Mutations occur to infer resistance to treatment A ; this is captured via $R(D)$ together with a corresponding growth term $f^\alpha(D)\tilde{f}^{1-\alpha}(D)R(D)x_A^-$.

Using the model (2.5)-(2.6), Panetta was able to investigate a variety of treatment protocols under different conditions. For example, in all treatment regimes, minimizing the rest period was shown to produce a better outcome in terms of overall tumor mass. Combination ($AB \rightarrow AB \rightarrow \dots$) and sequential ($A \rightarrow B \rightarrow A \rightarrow B \dots$) therapies were also compared. In the case of an initially sensitive tumor ($x_B(0) = 0$), both regimens are shown to have the same necessary and sufficient conditions for tumor eradication. However, for all $\alpha > 0$ (i.e., as long as drug A has some effect on the induced resistant cells), the sequential therapy eradicates the cancer faster. Different sequential pulsing therapies were also compared, such as repeated dosing strategies ($A \rightarrow A$) versus consistent switching ($A \rightarrow B$). Here, optimal strategies were dependent on tumor composition, i.e., the fraction $u := x_B/x_A$. Qualitatively, $A \rightarrow A$ is superior in cases when (i) the tumor is primarily sensitive, (ii) drug B is

less effective than drug A , or (iii) the mutation rate $R(D)$ is small.

Another related work is the study by Day *et al.* [78], where the authors construct computational models extending Goldie and Coldman's [74] work to asymmetrical treatments, and are able to design individualized treatments given detailed knowledge of the kinetic parameters of the tumor. Roe-Dale *et al.* [79] have also applied an ODE model combining the effects of drug resistance and cell-cycle dynamics to breast cancer dynamics. Clinically utilized treatments strategies (alternating cycles of CMF and doxorubicin) are analyzed, and it is concluded that drug resistance is paramount to understanding the superiority of the sequential regiment.

2.3 Optimal treatment strategies

Section 2.2 discusses the specific question of combination versus sequential administration when a number of therapies are available. More general issues can be studied, such as optimizing the treatment regimen to maximize the damage to the tumor while simultaneously minimizing the overall toxicity. In this section, we review a selection of such optimal control questions that are related to drug resistance in cancer therapy.

In [80], Ledzewicz and Schattler construct a two compartment ordinary differential equation (ODE) model to analyze properties of optimal controls relative to an objective functional J :

$$J = kN(T) + \int_0^T (\ell N(t) + m \cdot u(t)) dt. \quad (2.7)$$

Both sensitive and resistant cells are considered, and the authors study the one- and two-drug cases. T denotes the (fixed) treatment length, $N(t)$ the total number of cells at time t , $u(t)$ is the therapy dosage, and k, m, ℓ scaling parameters governing the relative importance of cancer cell kill versus toxicity constraints. In the case of one (two) drugs, $u \in [0, 1]$, $([0, 1]^2)$, and $m \in \mathbb{R}_+$ (\mathbb{R}_+^2).

For both the single and multiple drug scenarios, minimization of J as a function of u leads to Pontryagin's Minimum Principle [81], yielding necessary condition for the optimal control $u_*(t)$. In the case when only one drug is available, bang-bang type controls are shown to be optimal as long as the resistant population is not too large. More precisely, singular controls can be excluded in the region of state space where

$$qS > (2 - q)R. \tag{2.8}$$

Here q is the rate in which sensitive cells develop resistance. This result relies on the Legendre-Clebsh condition, a second order moment in the minimization procedure, and intuitively seems reasonable: as long as most cells are sensitive, one should apply the maximal drug dosage; however as resistant populations begin to dominate, toxicity effects will outweigh reduced tumor burden. For two drugs, a different procedure must be used, as the associated Hamiltonian is a quadratic function of two variables (i.e. the treatment functions), and thus must obtain its minimum at a vertex $(0, 0)$, $(1, 0)$, $(0, 1)$, or $(1, 1)$. Hence Ledzewicz and Schattler are able to dynamically calculate the dosages as a function of position in phase space, and

conclude that it is never optimal to initiate (or withdraw) both drugs simultaneously.

Swierniak, Kimmel, and Polanski have also extensively studied optimal control techniques applied to drug resistance models (see the review paper [82]). For example, in [83, 84], the authors study an infinite-dimensional system of coupled linear ODEs representing populations of increasingly resistant cancer cells. They assume that resistance is due to gene amplification, and that the process is subcritical, i.e., that deamplification is more probable than amplification. Assuming that any additional genes produce completely resistant cells and that all mutated cells experience the same growth dynamics, this (simplified) model becomes

$$\begin{aligned}
 \dot{N}_0(t) &= [1 - 2u(t)]\lambda N_0(t) - \alpha N_0(t) + dN_1(t), \\
 \dot{N}_1(t) &= \lambda N_1(t) - (b + d)N_1(t) + dN_2(t) + \alpha N_0(t), \\
 &\vdots \\
 \dot{N}_i(t) &= \lambda N_i(t) - (b + d)N_i(t) + dN_{i+1}(t) + bN_{i-1}(t), \\
 &\vdots
 \end{aligned} \tag{2.9}$$

with an objective functional analogous to (2.7). Here, N_0 denotes the number of sensitive cells, and N_i the number cells with i additional genetic material. $u(t)$ is the dosage of treatment at time t , λ the net growth rate, and a and b the probabilities of gaining and losing a gene during mutation, respectively. Using Laplace methods for integrals and Nyquist theory, sufficient conditions were derived which assured the asymptotic eradication of the total tumor cell population $N(t) = \sum_{i=1}^{\infty} N_i(t)$ during

constant treatment regimens ($u(t) \equiv u$):

$$\sqrt{d} - \sqrt{b} > \sqrt{\lambda}, \quad (2.10)$$

$$u > \frac{\alpha}{d \left(-b + d - \lambda + \sqrt{(b + d - \lambda)^2 - 4bd} \right)} + \frac{1}{2}. \quad (2.11)$$

A study of the general optimal control problem for arbitrary treatments was not fully undertaken, but gradient methods were used to find optimal bang-bang controls in [85]. For related works from this group see [86, 87].

Many cancer models focus on eradicating the disease completely. Realistically however, this is not always possible. Monro and Gaffney [88] focus on palliative or failed cure cancer treatment; that is, treatment based not on complete disease removal, but instead on prolonging survival and improving patient quality of life. Their goal is to derive optimal therapy protocols, but with objective functionals that are based on maximizing survival times, as opposed to minimizing the tumor cell load. The model of [88] consists of a coupled system of ODEs, representing the number of sensitive, $N_S(t)$, and resistant, $N_R(t)$, cells at time t :

$$\frac{dN_S}{dt} = -\beta \log \left(\frac{N(t)}{N_\infty} \right) (N_S - \lambda C(t) N_S + \tau_2 N_R - \tau_1 N_S), \quad (2.12)$$

$$\frac{dN_R}{dt} = -\beta \log \left(\frac{N(t)}{N_\infty} \right) (N_S - \tau_2 N_R + \tau_1 N_S). \quad (2.13)$$

Here $N(t) = N_S(t) + N_R(t)$, τ_1 and τ_2 are forward and backward mutation rates, respectively, and N_∞ is the carrying capacity for the Gompertzian tumor growth, characterized by a rate β . Note that treatment has no effect on resistant cells, and that

the chemotherapeutic effect ($\lambda C(t)N_S$) is proportional to the tumor growth rate, in accordance with the Norton-Simon hypothesis [89, 90]. Assuming that death occurs once the total tumor population reaches a critical value, N_{crit} , Monro and Gaffney studied the effect of dosage $C(t)$ on survival time T_{crit} , where $N(T_{\text{crit}}) = N_{\text{crit}}$. Numerical simulations provide a number of interesting conclusions. For example, in the case of continuous infusion, it is deduced that an intermediate level of chemotherapy C_* is optimal in maximizing T_{crit} . This has a simple biological explanation: the dosage C_* achieves a balance between sensitive cell kill and the growth inhibition of the resistant population due to the presence of sensitive cells.

2.4 Modeling mechanisms: the P-glycoprotein pump

Aside from studying the phenomenological aspects of drug resistance, mathematical models of specific resistance mechanisms have been formulated and analyzed. As discussed in Section 1.3.2, one of the best understood pathways of MDR is an increase of ABC transporter proteins on the cell membrane, which act to decrease the overall drug concentration interior to the cell. In this section, we overview a model introduced by Michelson and Slate [91] that incorporates this phenomenon.

P-glycoprotein (P-gp) is an energy dependent pump that acts by first binding with adenosine triphosphate (ATP) and subsequently to a cytotoxic agent inside of the cell. As ATP is synthesized to produce adenosine diphosphate (ADP), the drug is expelled through the cell membrane, thus acting as a pump for cytotoxic agents. Thus one method of reversing the action of the P-gp is to lower the overall cellular

concentration of ATP. However, Michelson and Slate argue against this methodology, both because the required cellular manipulation would have to be extremely precise, as well as the fact that ATP levels are inherently self-regulating. Instead, the authors argue in favor of inhibiting P-gp's ability to bind to ATP and/or the therapeutic agent. Denoting the substrate (i.e., the drug) by S , and the intracellular ATP and ADP concentrations by X_{ATP} and X_{ADP} , respectively, the Michelson and Slate model takes the form of a coupled system consisting of one PDE and two ODEs:

$$\frac{\partial S}{\partial t} = \alpha \frac{\partial^2 S}{\partial \ell^2} + F, \quad (2.14)$$

$$\frac{dX_{\text{ATP}}}{dt} = -2F - \frac{V_{\text{ATP}}X_{\text{ATP}}}{K_{\text{ATP}} + X_{\text{ATP}}} + \frac{V_{\text{ADP}}X_{\text{ADP}}}{K_{\text{ADP}} + X_{\text{ATP}}}, \quad (2.15)$$

$$\frac{dX_{\text{ADP}}}{dt} = 2F + \frac{V_{\text{ATP}}X_{\text{ATP}}}{K_{\text{ATP}} + X_{\text{ATP}}} - \frac{V_{\text{ADP}}X_{\text{ADP}}}{K_{\text{ADP}} + X_{\text{ATP}}}, \quad (2.16)$$

$$F = \frac{V_{\text{max}}X_{\text{ATP}}S}{(K_1 + X_{\text{ATP}})(K_2 + S)}. \quad (2.17)$$

Here, ℓ is the distance through the cell membrane, α is a diffusion constant, F is the total flux of S out of the cell, V_{max} is the minimum of the APT and drug binding reaction rates, and $V_{\text{ATP}}, V_{\text{ADP}}, K_{\text{ATP}}, K_{\text{ADP}}$ are Michaelis-Menten type kinetic terms. Note that the minimum is used to account for the slower of the two reactions, and a factor of two appears in the flux terms since ATP concentration is depleted by two molecules per drug molecule.

Through numerical simulation, Michelson and Slate argue that diffusion alone through the cell membrane is incapable of reproducing accurate intracellular drug concentration profiles. However, when including the non-zero flux term F in equa-

tions (2.14)-(2.17), their simulation results agree with biological data. Furthermore, the authors demonstrate that depleting ATP concentrations is not a viable treatment option, as the ATP levels of the cell are sufficiently self-regulated to ensure efflux and thus mediate MDR. An extension of this work by the same authors that includes an inhibitor is given in [92].

For another mathematical model of P-gp mediated drug resistance, see [93], where Panagiotopoulou *et al.* discuss an interesting relationship between Polya's Recurrence Theorem from probability theory and the biophysical properties of MDR. For a model that combines cell-cycle dynamics, metabolism, drug pharmacokinetics, and nutrient and drug diffusion, see [94].

Chapter 3: Cell density, mutations, and drug resistance: a structured population approach

The content of this chapter has been published in [95]. Here we develop a mathematical framework for describing multidrug resistance in cancer. To reflect the complexity of the underlying interplay between cancer cells and the therapeutic agent, we assume that the resistance level is described via a continuous parameter. Our model is written as a system of integro-differential equations that are parameterized by the resistance level, and incorporates cell-density and mutations. Analysis and simulations of the model demonstrate how the dynamics evolve to a selection of one or more traits corresponding to different levels of resistance. The emerging limit distribution with nonzero variance is the desirable modeling outcome as it represents tumor heterogeneity.

3.1 Introduction

Resistance to chemotherapy is a major cause of the failure of cancer treatment. Our current understanding of drug resistance in cancer is that tumor heterogeneity and complex genetic and epigenetic changes contribute to the development of multidrug resistance (MDR). When this occurs, the cell becomes resistant to a variety of

structurally and mechanistically unrelated drugs in addition to the drug initially administered (see [49, 96]).

Mathematical models have been used to study drug resistance in a variety of circumstances. Furthermore, many important questions have been addressed using different computational approaches, such as: How is early detection and therapy connected with the development of drug resistance? What is the probability that at the time of diagnosis, resistant cancer cells are already present? When several drugs are available, how many drugs should be applied? Should these drugs be used in combination, or sequentially? (see the recent review [66]).

Cancer models include features that influence growth, evolution, MDR and/or intra-tumoral heterogeneity, and hence include information about cellular rates. In some cases, these rates can be functions of time, space, density, and/or environmental signals. For instance, consider the natural division rate of cancer cells. The simplest models of cancer growth assume exponential growth (see, e.g., [77, 80]), with the growth rate proportional to the current density of cells. This is useful for describing cells in a certain growth phase, but most likely does not hold over the lifetime of the disease, as it yields unbounded growth. Logistic growth accounts for saturation effects by adding a quadratic competitive interaction term to the growth rate (e.g., see [97–99]). The Gompertz growth function, introduced in 1964 by Laird to account for the sigmoidal growth dynamics often observed in tumors, is another prominent choice (see [100, 101]). And other variations exist (for example, see [102]).

From the above, we observe discrepancies in the description of the division rate. Essentially, these models take different forms for their growth rate as a function

of cellular density. In fact, other rate parameters often also depend on density, and it is an important aspect of modeling drug resistance. Cell density relates to various factors that determine tumor growth, such as cell-cell interactions, cell-matrix interactions, nutrient distribution, survival signals, the penetration of anti-cancer drugs, and the internal tumor pressure. Many experiments using both normal and cancer cells have directly or indirectly measured the relationship between the rates of cell division and death as functions of cell density, both with and without the administration of a chemotherapeutic agent [102–110]). In some studies, the cell division rate was found to be a decreasing function of cell density [102, 105], while in others, the reverse was observed [106, 108, 109]. On the other hand, several independent studies have clearly shown that the exact cell death rate depends on the experimental design [105, 110, 111]. Therefore, a theoretical study should examine the properties of these dependencies.

Furthermore, one must also understand how drug resistance is being characterized so as to choose an appropriate mathematical paradigm to frame the model. Indeed, in the mathematical literature, the problem has been studied for over thirty years. One of the early models of resistance considers the mechanism of point mutations, where cells are in one of two compartments: sensitive or resistant (e.g., see [70, 74, 75, 112–115]). More advanced stochastic models of point mutations are found in the reports of Komarova *et al.* [72, 73], while Kimmel *et al.* [116, 117] study gene amplification through branching processes. Continuum-based models are also widely utilized. For instance, ordinary differential equations (ODEs) are employed by Birkhead *et al.* [118] to study kinetic resistance, and by Tomasetti and Levy [119]

to study point mutations. Michelson and Slate use ODEs to model resistance via increased drug efflux through the ABC transport proteins pathway [91]. Partial differential equation models (PDEs) are also used to describe heterogeneous tumors and vascularization [120], and integro-differential equations (IDEs) are sometimes used to describe mutations between sensitive and resistant cancer cells [121].

Lorz *et al.* [121] proposed a mathematical model that includes the effects of a continuous variable describing the level of resistance a cell has obtained on cell divisions, apoptosis, and mutation rates of both healthy and cancerous populations. They provide a basis for structured population models including a selection process based on an approach previously developed using PDEs [122–130]. Both intratumoral heterogeneity and cellular density were incorporated in the Lorz *et al.* model. Heterogeneity is described via the resistance variable, and density effects were included only on the growth rate of healthy cells, with cancerous cells having no density dependencies.

In this chapter, we aim to develop a mathematical framework for describing MDR by incorporating cell density and mutation dependence. We consider all growth parameters (division, death, and mutation rates, etc.) to be functions of the resistance level (“trait”), which we assume to be a continuous variable. Based on a system of IDEs parameterized by the resistance level, we provide a basis for structured population models designed to estimate the intratumoral heterogeneity over time. Indeed, intratumoral heterogeneity is a widely observed phenomena [33, 131, 132], but many basic models do not account for it (see Section 3.3). Since cellular density has been proven to have a profound impact on cancer cell

dynamics, our study extends the approach of Lorz *et al.* by including these effects on both natural and drug-induced death rates, in addition to the cell division rate of the cancer cells. Moreover, in our model, the spontaneous death rate is considered to vary in different ways than in [121]. This results in significantly different net growth and heterogeneity (Figs. 3.9 and 3.10). Furthermore, in our model, we account for the occurrence of changes that proceed at higher rates than those of genetic mutations, such as epigenetic changes.

The chapter is organized as follows. We first provide an overview of the mathematical model in Section 3.2. In Section 3.3 we explore in detail how the mathematical model corresponds to biological observations. We proceed by exhibiting the different behaviors that can be produced by various formulations of the model, together with biological implications. In particular, we move from a model with no density dependence or genetic variation to one that includes both and demonstrate how these features affect tumor heterogeneity. Concluding remarks are presented in Section 3.4.

3.2 A mathematical model for cancer dynamics

In deriving the mathematical model, we extend the approach of Lorz *et al.* [121]. In [121], both healthy and cancerous cells are modeled using structured population dynamics that incorporate cell divisions with and without mutations. Healthy cells also include a homeostatic density dependence on the cell division rate only, while cancerous cells are assumed to be able to grow without this inhibitive factor. In

Variable	Range	Biological Interpretation
x	$[0, 1]$	Resistance level
μ	\mathbb{R}_+	Time
$n(x, t)$	\mathbb{R}_+	Concentration of cells with trait x at time t
$\rho(t)$	\mathbb{R}_+	Density of cells at time t
ρ_f	\mathbb{R}_+	Density of cells at steady state
$r(x)$	\mathbb{R}_+	Natural division rate of cell with trait x
$d(x)$	\mathbb{R}_+	Natural apoptosis rate of cell with trait x
$c(x)$	\mathbb{R}_+	Drug-induced apoptosis rate of cell with trait x
$D(t)$	\mathbb{R}_+	Concentration of drug applied at time t
$f(\rho)$	\mathbb{R}_+	Density dependence on division rate
$g(\rho)$	\mathbb{R}_+	Density dependence on natural apoptosis rate
$\theta(x)$	$[0,1]$	Proportion of divisions of cells with trait x undergoing mutations
$M(y, x)$	$[0,1]$	Probability a division results in a mutation from state y to state x given a mutation occurs
Δx	$(0, 1)$	Step-size used in numerical simulations

Table 3.1: List of variables

this work, we consider cancer cells only, coupled with a drug, which is assumed to be a chemotherapeutic agent, and assume, as is done in Lorz *et al.* [121], that the resistance level is described by a scalar x , which is normalized such that $x \in [0, 1]$.

Let $n(x, t)$ denote the population density of cancer cells with trait x at time t , and let $\rho(t)$ denote the total number of cells,

$$\rho(t) = \int_0^1 n(x, t) dx. \quad (3.1)$$

The dynamics of the cancer cell population is then described via the selection/mutation

integro-differential equation

$$\begin{aligned} \frac{\partial n(x, t)}{\partial t} = & \left(f(\rho(t))[r(x)(1 - \theta(x)) - h(D(t), x)] - g(\rho(t))d(x) \right) n(x, t) \\ & + f(\rho(t)) \int_0^1 \theta(y)r(y)M(y, x)n(y, t) dy. \end{aligned} \quad (3.2)$$

The rationale for Eq. (3.2) is as follows:

1. $r(x)$ and $d(x)$ denote the natural division rate and death rate of cancer cells with trait x , respectively. We assume that no external forces are influencing the inherent growth and death rates, and hence these rates are constant in time.
2. $D(t)$ represents the dosage of the chemotherapeutic drug at time t , with $h(D(t), x)$ representing the net cytotoxic effect on cancer cells with trait x under dosage $D(t)$. It is common to assume that h increases linearly with drug dosage D , but in certain situations, other forms may be valid. For instance, it may be desirable to include saturation effects in h for large drug concentrations D (for instance, $\lim_{D \rightarrow \infty} h(D, x) = h_1(x) < \infty$, for some function of trait values h_1).
3. The terms $f(\rho(t))$ and $g(\rho(t))$ are included to incorporate density dependence into the division and apoptosis rates. It is well known that both the division and apoptosis rates depend on the cell density (see [66, 104, 105, 133]). The specific functional form of these terms play a key role in the overall dynamics as is demonstrated in Section 3.3.

4. The birth ($r(x)$) and drug-induced apoptosis ($h(D(t), x)$) rates have the same density dependence factor f since we are assuming the drug is cell-cycle specific, and hence affects primarily only cells that are dividing. Hence, the induced death term should be proportional to the division rate. This assumption is widely known as the Norton-Simon Hypothesis [89, 90, 134].
5. All parameters and functions are nonnegative: $r, d, h, g \geq 0$, and $f > 0$ (see Eqs. (3.6) and (3.7) below). We also assume $r, d \in C([0, 1])$. As of now, we make no further assumptions on these terms.
6. We assume that when cells undergo division, they may mutate. That is, if the mother cell has the specific trait x , the daughter cell can have the same trait value x , which we call a faithful division, or it may have a different resistance level y , which we refer to as a mutation. Mutation can either be a permanent genetic change or permanent epigenetic change. The key assumption is that traits are passed down from mother to daughter cell and do not revert back to the original value x after some fixed time.
7. $\theta(x)$ denotes the fraction of cells with trait x undergoing mutation, where $0 \leq \theta(x) \leq 1$, and hence $1 - \theta(x)$ denotes the fraction of cell undergoing faithful division.
8. The last term on the RHS of (3.2) takes into account all mutations during division from different traits y . $M(y, x)$ denotes the probability that given a mutation, a mother cell with trait y will differentiate into a daughter cell with

trait x . As a probability density function, $M(y, x)$ satisfies

$$M(y, x) \geq 0 \quad \forall x, y \in [0, 1], \quad (3.3)$$

$$\int_0^1 M(y, x) dx = 1 \quad \forall y \in [0, 1]. \quad (3.4)$$

Equation (3.4) implies that once a mutation occurs, it must be to some value $x \in [0, 1]$.

The notations used in (3.2) are summarized in Table 3.1. Equation (3.2) is a selection and mutation model, which moves dynamically through the $(x, n(x, t))$ phase space by both Darwinian evolution and mutations. As time progresses, different gene expressions will become advantageous/disadvantageous, and the overall dynamics will be determined both by the rate and mutation parameters, in addition to the initial distribution of cells. It is important to note that Eq. (3.2) can be obtained as an expected value of a stochastic model of cells that obey the division/mutation dynamics described above, where the expected value is taken over the number of cells of type x . For the remainder of the paper, we make the following assumptions:

A1. $D(t) \equiv 1$. This is equivalent to assuming the drug is applied uniformly in time.

A time-dependent $D(t)$ allows one to study treatment protocols, but we do not study this here. Hence, we can write the drug-induced rate term $h(D(t), x)$ as a function of the trait value x only, which we define as $c(x)$ ($c(x) := h(1, x)$). As with the other rate parameters $r(x)$ and $d(x)$, we assume the trait dependence on the cytotoxic rate term is continuous, i.e., $c(x) \in C([0, 1])$.

A2. $\theta(x) \equiv \theta$, a constant. In this work we do not study the effect of the variation of the fraction of mutated divisions on trait values.

Substituting assumptions A1 and A2 into (3.2) yields the simplified equation

$$\begin{aligned} \frac{\partial n(x, t)}{\partial t} = & \left(f(\rho(t))[r(x)(1 - \theta) - c(x)] - g(\rho(t))d(x) \right) n(x, t) \\ & + \theta f(\rho(t)) \int_0^1 r(y)M(y, x)n(y, t) dy. \end{aligned} \quad (3.5)$$

If we define the new timescale $\tau = \int_0^t f(\rho(s)) ds$, then $\frac{\partial n}{\partial t} = \frac{\partial n}{\partial \tau} f(\rho(t))$, and we can rescale (3.5) to become

$$\begin{aligned} \frac{\partial n(x, \tau)}{\partial \tau} = & \left(r(x)(1 - \theta) - c(x) - G(\rho(\tau))d(x) \right) n(x, \tau) \\ & + \theta \int_0^1 r(y)M(y, x)n(y, \tau) dy, \end{aligned} \quad (3.6)$$

where

$$G(\rho) = \frac{g(\rho)}{f(\rho)}. \quad (3.7)$$

Note that $\frac{d\tau}{dt} = f(\rho(t)) > 0$, so that we have not changed the direction of time. Henceforth, we study (3.6), with the notational convention that we replace τ by t , with the change of units understood as above.

Here, we deal with two cases of initial conditions for (3.6):

- (i) Uniform distributions in trait x at time $t = 0$, i.e., $n(x, 0) = n_0$, a constant.
- (ii) An initial distribution that is concentrated around some trait $x = x_*$, i.e., $n(x, 0) = B(x) \exp(-|x - x_*|/\epsilon)$, where $0 < \epsilon$. The smaller the ϵ value, the

more concentrated the initial distribution is about x_* .

Assuming that $n_0, B(x) \geq 0$, the form of (3.6) implies that all solutions satisfy $n(x, t) \geq 0$ for all $t \geq 0$ and $x \in [0, 1]$ (see Appendix A for an elementary proof). Hence, for the remainder of the work, we make these assumptions to ensure positive densities for all time.

3.3 Analysis and simulations

In this section we present analytical and simulation results for Eq. (3.6). In order to study the relative role of different terms in the model, we study different variants of the model. We begin with a trait-based growth model.

3.3.1 Trait-based growth

We first consider the simplified version of (3.6) with no density dependence and no mutations, so that growth depends solely on the cell division and death rates and their response to treatment. Hence, $f \equiv g \equiv 1$, or equivalently, $G(\rho) \equiv 1$, and $\theta = 0$. Thus, Eq. (3.6) can be written as

$$\frac{\partial n(x, t)}{\partial t} = [r(x) - c(x) - d(x)]n(x, t). \quad (3.8)$$

We note that Eq. (3.8) is the model-type used in Lorz *et al.* [121] to describe the dynamics of cancer cells. In writing such a model, it is assumed that the cells do not interact in any way, and growth is dictated solely by the trait x . The solution

to (3.8) is given by

$$n(x, t) = n(x, 0) \exp((r(x) - c(x) - d(x))t), \quad \forall x \in [0, 1], t > 0. \quad (3.9)$$

In this case, it is trivial to see that the dynamics of the cancer cells in (3.9) depend on the relative values of $r(x) - c(x)$ and $d(x)$. Indeed, if $r(x) - c(x) > d(x)$ and $n(x, 0) > 0$, then $n(x, t) \rightarrow \infty$ as $t \rightarrow \infty$. If $r(x) - c(x) < d(x)$, then $n(x, t) \rightarrow 0$ as $t \rightarrow \infty$. If $r(x) - c(x) = d(x)$, then $n(x, t) \equiv n(x, 0)$ for all $t \in \mathbb{R}_+$. Finally, if $n(x, 0) = 0$, then $n(x, t) = 0$ for all $t > 0$. This last case can be ignored by assuming that $n(x, 0) > 0$ for all $x \in [0, 1]$. which is a technical assumption that simplifies the presentation of the following results. Hence, we are left with only two qualitatively distinct scenarios, which are described by the below theorem:

Theorem 1. *Consider the system described by (3.8) with initial condition $n(x, 0)$.*

If $r(x) - c(x) < d(x)$ for all $x \in [0, 1]$, then

$$\rho(t) \xrightarrow[t \rightarrow \infty]{} 0, \quad n(x, t) \xrightarrow[t \rightarrow \infty]{} 0 \quad \forall x \in [0, 1]. \quad (3.10)$$

An analogous statement holds if $r(x) - c(x) \leq d(x)$ with $r(x) - c(x) = d(x)$ at a countable number of points x_ , where we will have $n(x_*, t) \equiv n(x_*, 0)$. On the other hand, if there exists $x \in [0, 1]$ such that $r(x) - c(x) > d(x)$, then the solutions satisfy*

$$\rho(t) \xrightarrow[t \rightarrow \infty]{} \infty, \quad \frac{n(x, t)}{\rho(t)} \xrightarrow[t \rightarrow \infty]{} \sum_{i=1}^m a_i \delta(x - x_i), \quad (3.11)$$

where $\delta(x)$ denotes the Dirac distribution, the $a_i > 0$ are constants such that $\sum_{i=1}^m a_i = 1$, and the x_i satisfy

$$x_i = \arg \max_{0 \leq x \leq 1} (r(x) - c(x) - d(x)). \quad (3.12)$$

For a proof, see Appendix A. The above theorem and proof are analogous to [121, Lemma 2.2]. Note that m denotes the number of points where the maximum is achieved. The constants a_i in the limiting distribution of $n(x, t)/\rho(t)$ must add to one, since this quantity is a probability distribution. Generally, since the initial conditions may be nonhomogeneous, the constants a_i are not identical. In this case, the relative growth in each population will depend on the fraction of initial cells with the corresponding trait. That is, $a_i = n(x_i, 0)/\sum_i^m n(x_i, 0)$.

Theorem 1 implies that (3.8) can only describe two different growth phenomena: extinction or unbounded growth. In modeling the evolution of cells with a resistant gene, both cases have serious limitations. Extinction, while desirable, is unlikely, and unbounded growth is not biologically feasible. It is well known that even cancer cells, which experience less inhibition to growth signals due to cellular pressure and a lack of resources, experience homeostatic forces which slow their division rate (e.g., see [105]). Accordingly, Eq. (3.8) is ill-equipped for studying the growth properties of tumor cells. It is however worth noting that there still is some merit in using such a model when comparing the dynamics of cancerous cells to normal tissue, as cancer cells may exhibit unlimited growth potential in comparison with healthy cells.

Despite with the described limitations, (3.8) can describe a rudimentary model of selection. Theorem 1 implies that such a model selects the traits with the maximum net growth rate. That is, all x that maximize $r(x) - c(x) - d(x)$ are selected. Hence, even though (3.8) provides an overall unrealistic modeling environment, it does capture some important phenomena.

To understand how the dynamics evolve, we provide numerical simulations. For this and all future simulations, we will use a standard collocation method of Sinc basis functions on the interval $[0, 1]$ with an equally spaced partition of $k = 4000$ points (see [135] for a basic introduction to the method). That is, we approximate $n(x, t)$ by $n_k(x, t)$, which are defined by

$$n_k(x, t) := \sum_{i=1}^k S_i(x) n_i(t) \tag{3.13}$$

where $n_i(t) := n(t, x_i = i\Delta x)$, x_i is a fixed point in the partition, and $S_i(x)$ is the function

$$S_i(x) := \begin{cases} \frac{\sin((\pi/\Delta x)(x-i\Delta x))}{(\pi/\Delta x)(x-i\Delta x)} & \text{for } x \neq i\Delta x \\ 1 & \text{for } x = i\Delta x. \end{cases} \tag{3.14}$$

Here Δx is the step-size of the partition ($\Delta x = 1/3999$ in our simulations). Note that $S_i(x_i) = 1$ and $S_i(x_j) = 0$ for $i \neq j$. Substituting the approximation (3.13) into (3.8) (or later on, the more general (3.6)), yields a system of ODEs for the set $\{n_i(t)\}$, which we solve by standard Runge-Kutta methods. For example, in the case of Eq. (3.8), the system of ODEs is given simply by the original equation

Figure	Equation	Parameter Values
3.1	(3.8)	$r(x) = \frac{2}{1.1+2x^5}$, $c(x) = \frac{2}{1+x^2}$, $d(x) \equiv 0.05$
3.2	(3.8)	$r(x) = \frac{2}{1.1+2x^5}$, $c(x) \equiv 0$, $d(x) \equiv 0.05$
3.3	(3.16)	$r(x) = \frac{2}{1.1+2x^5}$, $c(x) = \frac{2}{1+x^2}$, $d(x) \equiv 0.05$, $G(\rho) = \rho(1 - \rho)^2$
3.4	(3.16)	Same as in Fig. 3.3
3.5	N/A	$r(x) - c(x) = -5x^2 + 6x - \frac{6}{5}$, $d(x) \equiv 2$
3.6	(3.16)	$r(x) - c(x) = -5x^2 + 6x - \frac{6}{5}$, $d(x) \equiv 2$, $G(\rho) = \rho(1 - \rho)^2$
3.7	(3.16)	Same as in Fig. 3.6
3.8	N/A	$r(x) - c(x) = -5x^2 + 6x - \frac{6}{5}$, $d(x) = \begin{cases} x & \text{if } x \leq \frac{4}{5} \\ \frac{0.0352}{x^{14}} & \text{if } x > \frac{4}{5} \end{cases}$
3.9	(3.8)	$r(x) - c(x)$ and $d(x)$ as in Fig. 3.8
3.10	(3.16)	$r(x) - c(x)$ and $d(x)$ as previous, $G(\rho) = \rho(1 - \rho)^2$
3.11	(3.16)	Same as in Fig. 3.10
3.12	(3.31)	$r(x) = \frac{2}{1.1+2x^5}$, $c(x) = \frac{2}{1+x^2}$, $d(x) \equiv 0.05$, $G(\rho) = \rho(1 - \rho)^2$, $\epsilon = 0.01$, $\theta = 0.1$, \tilde{M} as in Eq. (3.30)
3.13	(3.31)	Same as in Fig. 3.12
3.14	(3.31)	$\epsilon = 0.01$, $\theta = 0.1, 0.4, 0.7, 1$. All other parameters are identical to Fig. 3.12
3.15	(3.31)	$\epsilon = 0.1$, $\theta = 0.1, 0.4, 0.7, 1$. All other parameters are identical to Fig. 3.12

Table 3.2: Parameter values used in simulations

$$\frac{dn_i(t)}{dt} = [r(x) - c(x) - d(x)]n_i(t). \quad (3.15)$$

In the general IDE (3.6), the analogous expression is more complicated, but remains an ODE with coefficients involving $\int_0^1 S_i(x) dx$, which, once computed initially, are constants.

We would also like to briefly discuss the form of the rate parameters. A common characteristic is for both $r(x)$ and $c(x)$ to decrease as the resistance x increases (see [103, 131]). As x increases, the cell becomes more resistant, so by

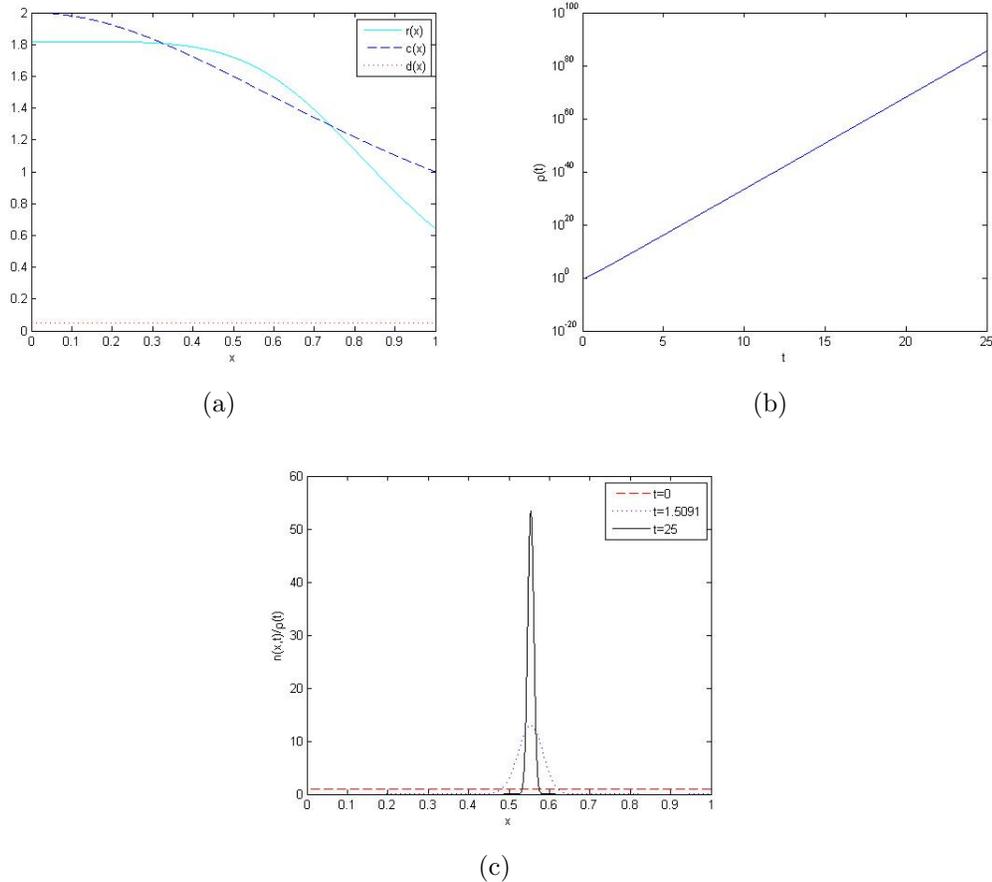


Figure 3.1: Numerical results of (3.8). (a) shows the rate profiles used; (b) a semi-log plot of the evolution of the density ρ ; (c) $n(x, t)/\rho(t)$ as a function of x at three representative times t . Note that in the simulations we have made the time scale change $t \rightarrow t\epsilon^{-1}$, where $\epsilon = 10^{-2}$, since we are primarily interested in studying the long-time behavior of the solution.

definition the drug-induced kill-term $c(x)$ should decrease for resistant cells. It is less obvious how $d(x)$ should vary with x . In some biological situations, such as the accumulation of mutations, $d(x)$ may be increasing. On the other hand, in the case of epigenetic changes that reduce the division rate, the death rate could be reduced. In fact, it may not be a monotone function at all. Hence, when studying such a problem, it is important to identify which mechanism or mechanisms are under investigation, so as to accurately represent the respective rates.

Currently, we assume a simple constant form for $d(x)$, and consider uniform initial conditions, although the results are independent of the latter choice. That is, as long as $n(x, 0) > 0$ for all $x \in [0, 1]$, the same asymptotic behavior is achieved (In this case, since mutations are not allowed, if $n(x, 0) = 0$ for some $x \in [0, 1]$, then growth will never occur at such x , and thus the asymptotic behavior could be different. Such is the case if $n(x_*, 0) = 0$ for a maximizer x_* of $r(x) - c(x) - d(x)$). The explicit functional forms and parameter values used in all simulations are given in Table 3.2. The simulation results are shown in Fig. 3.1. Here, the second part of Theorem 1 applies, where there exists x such that $r(x) - c(x) - d(x) > 0$. Theorem 1 then implies that we should have unbounded growth of the total population size $\rho(t)$, and that the growth should concentrate around the x values that maximize $r(x) - c(x) - d(x)$ (Fig. 3.1(a) shows that this value x_* is in fact unique in this simulation). Figures 3.1(b) and 3.1(c) verify the conclusions of the theorem in this case. That is, $\rho(t)$ in Fig. 3.1(b) demonstrates exponential growth, and $n(x, t)$ in Fig. 3.1(c) appears to approach a Dirac mass about the value x_* as t grows.

For comparison, we also provide simulations for the case of no treatment, i.e., we study (3.8) with identical parameter values, except we set $c(x) \equiv 0$. As in Fig. 3.1 (specifically Fig. 3.1(b)), the density $\rho(t)$ diverges, and in fact diverges more rapidly than in Fig. 3.1. This is due to the net growth rate $(r(x) - d(x))$ being larger than the case where the drug was applied (net growth rate $r(x) - c(x) - d(x)$). That is, the drug inhibits the overall growth of the tumor, as expected. To keep the numerics within MATLAB tolerances, we simulate $n(x, t)/\rho(t)$; the results are shown in Fig. 3.2. We see that the model selects the $x = 0$ trait asymptotically,

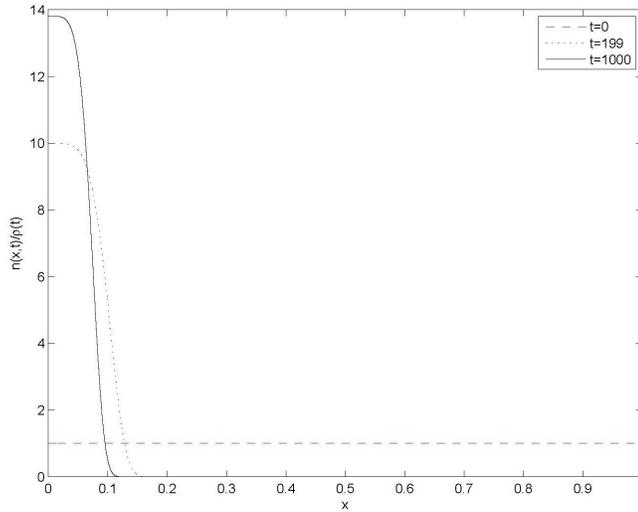


Figure 3.2: Numerical results of (3.8) with $c(x) \equiv 0$, and the same parameters appearing in Fig. 3.1. Shown is $n(x, t)/\rho(t)$ as a function of x at three representative times t .

as predicted from Theorem 1, since the maximum of $r(x) - d(x)$ occurs at $x = 0$. Biologically this is intuitive, since there is no evolutionary advantage to having a higher x value when the drug is applied. We lastly note that the time scale in these simulations is much longer than in Fig. 3.1 because $r(x) - d(x)$ is relatively flat near $x = 0$, and thus it takes longer to visually observe selection.

Clearly these figures do not represent a general case, as multiple maximizers of $r(x) - c(x) - d(x)$ could exist. Even if we assume $r(x)$ and $d(x)$ are both decreasing, there could still be multiple or no points of intersection between them. Given that in general there are no standard assumptions on $d(x)$, the limiting distribution could be any number of Dirac masses, or cell extinction.

3.3.2 Density-dependent model

We next consider a generalization of (3.8) to include competition and/or cooperation among the cells via density effects, while still not allowing for mutations during division. We recall that Lorz *et al.* [121] used a specific form of this model (both with and without mutations) to describe healthy cells, while here we use this more general form to describe cancerous cells. In reference to our original model (3.6), we set $\theta = 0$ and let $G(\rho) \neq 1$ generally. Thus, Eq. (3.6) takes the form

$$\frac{\partial n(x, t)}{\partial t} = [r(x) - c(x) - G(\rho)d(x)]n(x, t). \quad (3.16)$$

Similarly to (3.8), for any x (3.16) is an ODE. However, due to the density dependence via $G(\rho)$, the ODEs in (3.16) are nonlinear and coupled. Due to these two characteristics, Eq. (3.16) has more complex dynamics than the trait-based growth only model (3.8).

We recall that $G(\rho) = g(\rho)/f(\rho)$, where f is the birth-rate density dependence, and g is the natural death-rate density dependence. We assume that in the limit of high density, the growth rate of the cells rapidly decays. More precisely,

$$\lim_{\rho \rightarrow \infty} \frac{f(\rho)}{g(\rho)} = 0^+ \quad \text{or, equivalently,} \quad \lim_{\rho \rightarrow \infty} G(\rho) = \infty. \quad (3.17)$$

The form of Eq. (3.17) guarantees that if the natural death-rate density dependence g decays to zero in the limit of large cell density, the birth-rate f must decay even

more rapidly. This condition hence prohibits unbounded growth as stated in the following Theorem 2, which is proven in Appendix A:

Theorem 2. *Consider the system described by (3.16). If $G(\rho)$ satisfies (3.17), then there exists $\rho_M > 0$ such that*

$$0 \leq \rho(t) \leq \rho_M \quad \forall t \geq 0. \quad (3.18)$$

The conclusion of Theorem 2 is in fact one of the main motivating factors in introducing density effects into our models. It allows for bounded, nonzero populations to exist, which is biologically (and mathematically) desirable when modeling solid tumor growth. This should be compared with Theorem 1, where the only two outcomes were extinction and unbounded growth.

To demonstrate the difference between Eqs. (3.8) and (3.16), we simulate (3.16) with the same trait dependencies as in Fig. 3.1(a). We use the following generic choice for $G(\rho)$:

$$G(\rho) = \rho(\rho - 2)^2. \quad (3.19)$$

Note that (3.19) satisfies (3.17), so that Theorem 2 applies. The simulation results of (3.16) are shown in Figs. 3.3 and 3.4. For a direct comparison with Fig. 3.1, uniform initial conditions are used, although the long-time dynamics would remain unchanged with any initial conditions satisfying $n(x, 0) > 0$ for all $x \in [0, 1]$.

Figures. 3.3(a) and 3.1(c) exhibit the same selection, that is, growth that

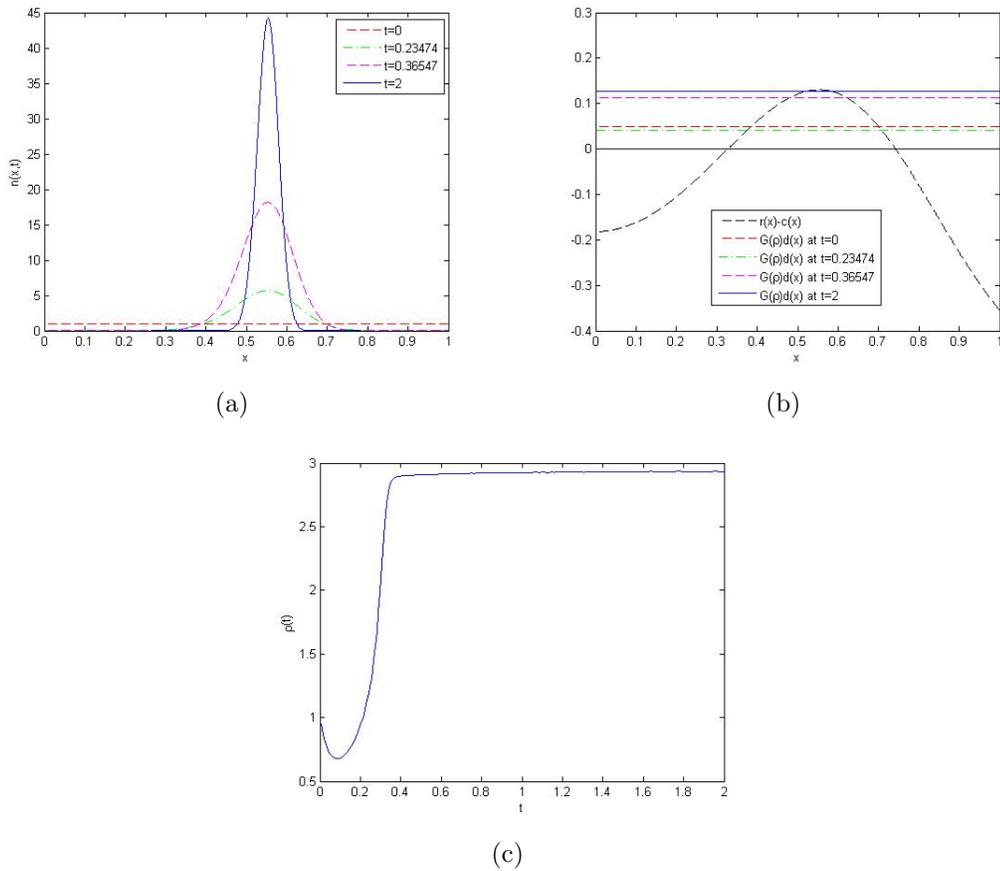


Figure 3.3: Simulations of (3.16) using trait parameters and time scale as in Fig. 3.1. G is given by (3.19). (a) the profiles of $n(x,t)$ at four times t ; (b) the variation of the net-growth rates as the density changes in time, at the same four times used in (a); (c) evolution of the density $\rho(t)$.

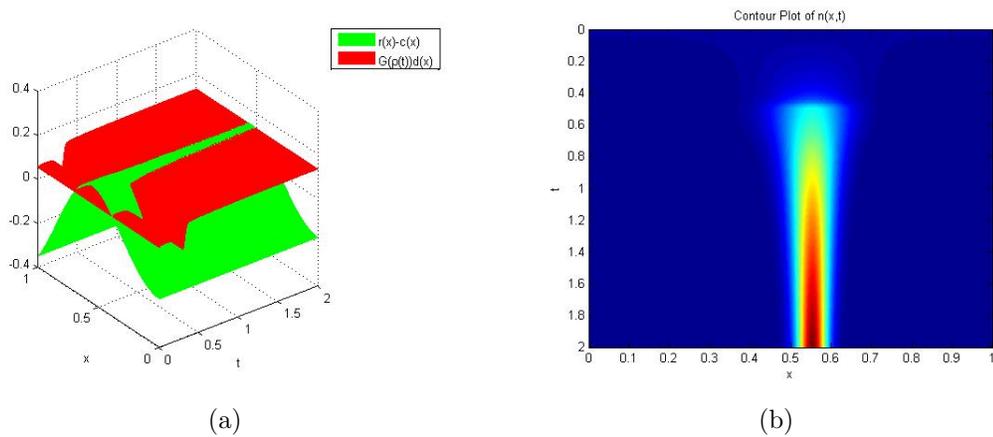


Figure 3.4: Simulations of (3.16) using trait parameters and time scale as in Fig. 3.1 and G is given by (3.19). (a) the net-growth rates as a surface plot over t and x ; (b) is a contour plot of $n(x,t)$.

limits to a monoclonal population with trait x_* . However, comparing Figs. 3.1(b) and 3.3(c), we see an important difference. Namely, instead of unbounded growth in the total population as given by Eq. (3.8), we now have, via Theorem 2, bounded populations. In fact, in Fig. 3.3(c), we see that $\rho(t)$ converges to an asymptotic profile. In contrast to Lorz *et al.* [121], cancer cells exhibit bounded growth, whereas previously, they must either diverge or become extinct. For the brevity and clarity of the presentation, we only show simulations results for $t = 2$. Analogous results are obtained for longer times.

To understand the dynamics portrayed by Fig. 3.4(a) (and Fig. 3.3(b) for snapshots in time), we consider a population with initial density $\rho(0) = 1$. In this case, the growth rate is $r(x) - c(x) - G(1)d(x)$. Due to the initial conditions and the growth parameters $r(x)$, $c(x)$, and $d(x)$, we observe an initial decrease in $\rho(t)$. $G(\rho)$, given by (3.19) increases, pushing the line $G(\rho)d(x)$ upwards. However, the contribution of the populations growing exponentially where $r(x) - c(x) - G(\rho)d(x) > 0$ outweighs the increase of populations with a negative division rate, and we quickly see an increase in $\rho(t)$. The product $G(\rho)d(x)$ decreases until $\rho = 2$, where the growth is most rapid, followed by a slowing down of $\rho(t)$ as $G(\rho)d(x)$ increases as it reaches the peak of $(r - c)(x)$. Note that the limiting behavior adapts to the growth rate, density, and population values satisfying the following relations:

$$\lim_{t \rightarrow \infty} \rho(t) = \rho_*, \quad (3.20)$$

$$G'(\rho_*) > 0, \quad (3.21)$$

$$r(x) - c(x) - G(\rho_*)d(x) < 0 \quad \forall x \neq x_*, \quad (3.22)$$

$$r(x_*) - c(x_*) - G(\rho_*)d(x_*) = 0, \quad (3.23)$$

and

$$n(x, t) \xrightarrow[t \rightarrow \infty]{} \rho_* \delta(x - x_*). \quad (3.24)$$

Properties (3.22) and (3.23) are most readily seen in Fig. 3.3(b) where the line $G(\rho)d(x)$ is essentially tangent to the $(r - c)(x)$ curve at its maximum value x_* when time $t = 2$. Properties (3.20)–(3.24) provide a way to understand the general dynamics of (3.16). An analogous statement holds if there exists multiple x_* satisfying (3.23).

The evolution of the population in this example is then similar to the population described by (3.8), except that the net growth rate adapts to bound the total population, leading to the convergence of $\rho(t)$. Indeed, this is an important biological characteristic that makes (3.16) a more suitable model.

Adding a density dependence can result in even richer dynamics. For example, consider the dynamics of both (3.8) and (3.16) with rate parameters shown in Fig. 3.5, where $G(\rho)$ is given by (3.19). The dynamics of (3.8) are clear: since

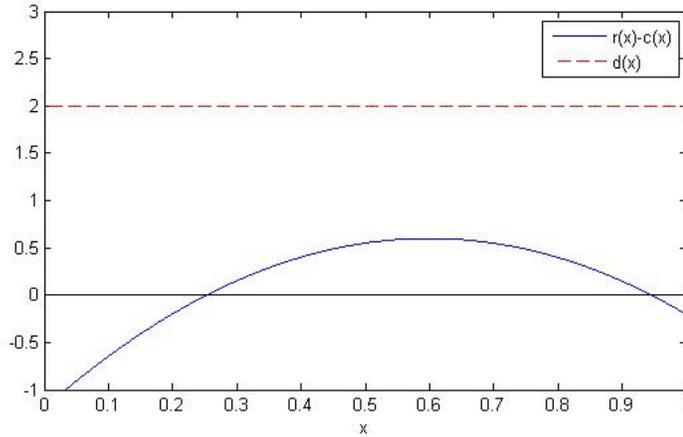


Figure 3.5: $r(x) - c(x)$ and $d(x)$ used for Figs. 3.6 and 3.7. Only the difference between $r(x)$ and $c(x)$ is plotted since that is the term that controls the dynamics in all models. Note that $r(x) - c(x) < d(x) \forall x \in [0, 1]$.

$r(x) - c(x) - d(x) < 0$ for all $x \in [0, 1]$, all the populations go extinct exponentially.

That is: $\rho(t) \rightarrow 0$ and $n(x, t) \rightarrow 0$ for all $x \in [0, 1]$ as $t \rightarrow \infty$.

However, the dynamics of (3.16) are qualitatively different, as shown in Figs. 3.6 and 3.7. Here we observe, as in the case of Eq. (3.8), an initial decay of $n(x, t)$ for all x , and thus a decrease in $\rho(t)$. However, the function $G(\rho)d(x)$ initially increases, causing an even faster decay of $\rho(t)$. As $\rho(t)$ approaches 0, $G(\rho)$ given by (3.19) also decreases. Eventually, by continuity of G and the fact that $G(0) = 0$ and that there exists x such that $r(x) - c(x) > 0$, ρ diminishes sufficiently so that there exists x such that $r(x) - c(x) - G(\rho)d(x) > 0$. Consequently, such cell populations begin to grow, leading to an increase in $\rho(t)$. The entire population is thus prevented from dying out, and the limiting dynamics are again described by (3.20)–(3.24), where the new ρ_* and x_* are determined by the trait and density parameters and by the initial conditions.

It is important to note that the reason the population does not die out in

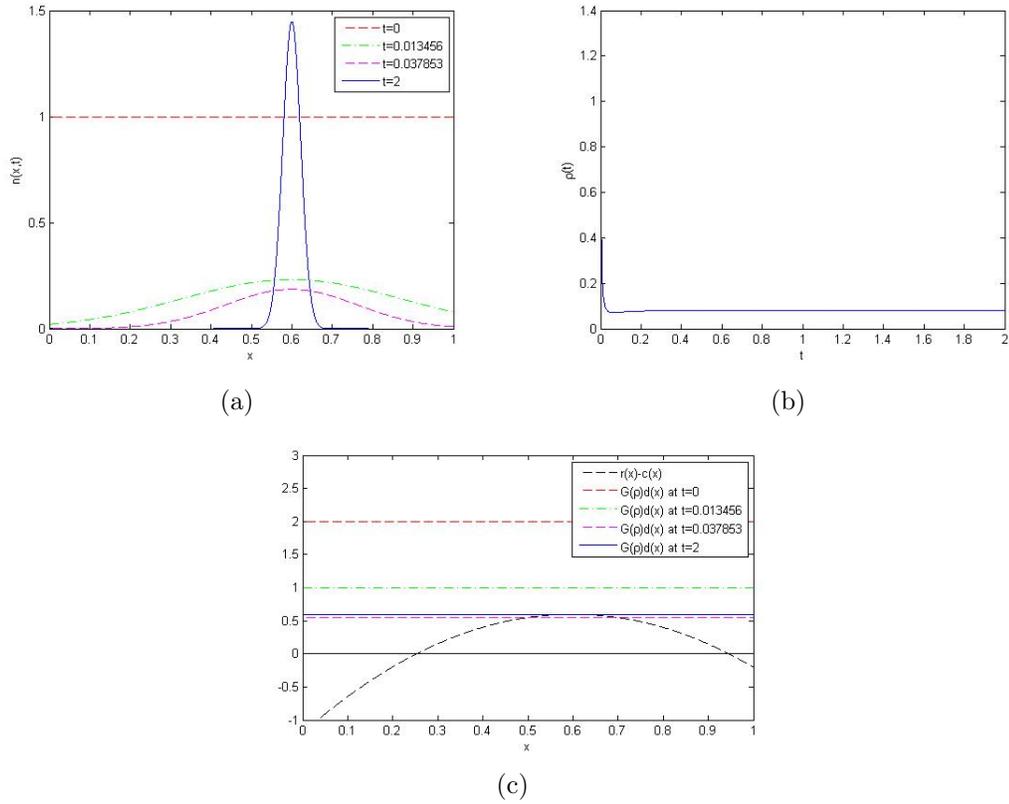


Figure 3.6: Simulations of (3.16) using the trait parameters shown in Fig. 3.5. G is given by (3.19). The time scale as in Fig. 3.1. (a) the profiles of $n(x,t)$ at four representative times; (b) the evolution of the density; (c) the variation of the net-growth rates as the density changes in time, at the same four times used in (a).

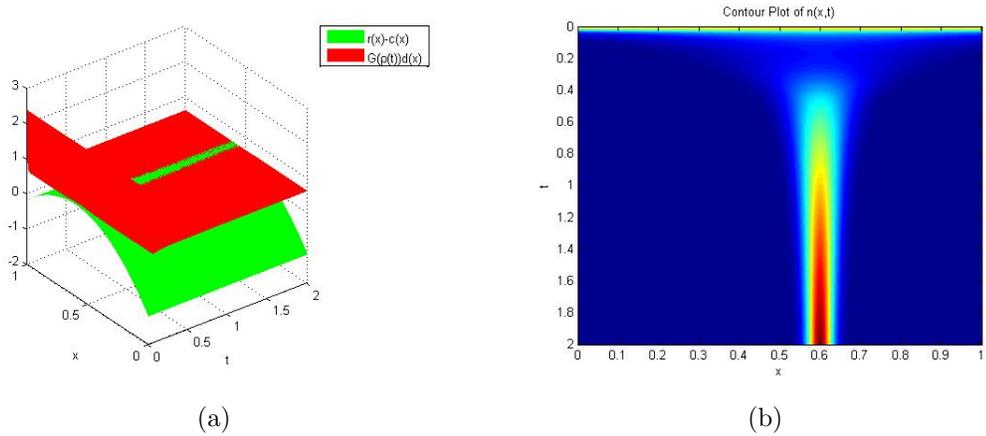


Figure 3.7: Simulations of (3.16) with the trait parameters shown in Fig. 3.5. G is given by (3.19). The time scale are identical to those in Fig. 3.1. (a) a surface plot of the net-growth rates; (b) a contour plot of $n(x,t)$.

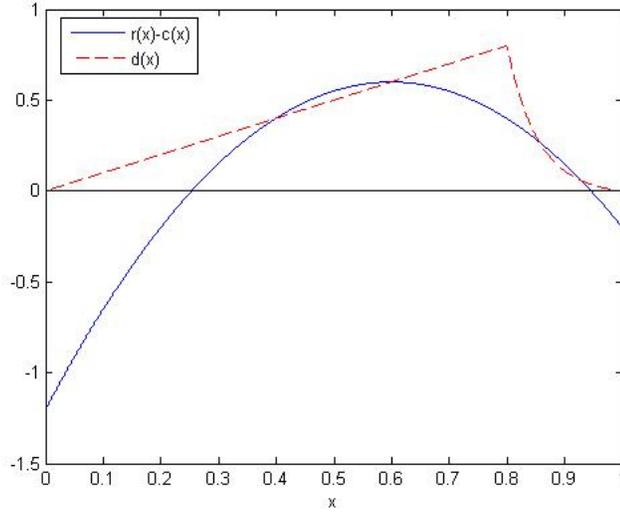


Figure 3.8: $r(x) - c(x)$ and $d(x)$ used for Figs. 3.9–3.11. Note that $\operatorname{argmax}_{0 \leq x \leq 1} (r(x) - c(x) - d(x)) = 0.5$ and that $r(0.5) - c(0.5) - d(0.5) > 0$.

Figs. 3.6 and 3.7 is because $G(0) = 0$. If the population approaches $\rho = 0$, the fact that there exists x where $r(x) - c(x) > 0$ allows us to conclude that there exist cells with a positive growth rate. If, for biological reasons, $G(0) > 0$, then it is entirely possible to have $\rho(t) \rightarrow 0$.

Incorporating a density dependence into the model may also result in a selection of different trait values x . Consider the same density dependence $G(\rho)$ as in (3.19), with $r(x)$, $c(x)$, and $d(x)$ that are shown in Fig. 3.8 (and Table 3.2). For Eq. (3.8), Theorem 1 implies that we have selection toward value(s) x_* that maximize $r(x) - c(x) - d(x)$, which by our choice is unique and given as $x_* = 0.5$, i.e., $\rho(t) \rightarrow \infty$ and

$$\frac{n(x, t)}{\rho(t)} \xrightarrow{t \rightarrow \infty} \delta(x - x_*). \quad (3.25)$$

This is demonstrated in Fig. 3.9. Note that pointwise we still have growth in other

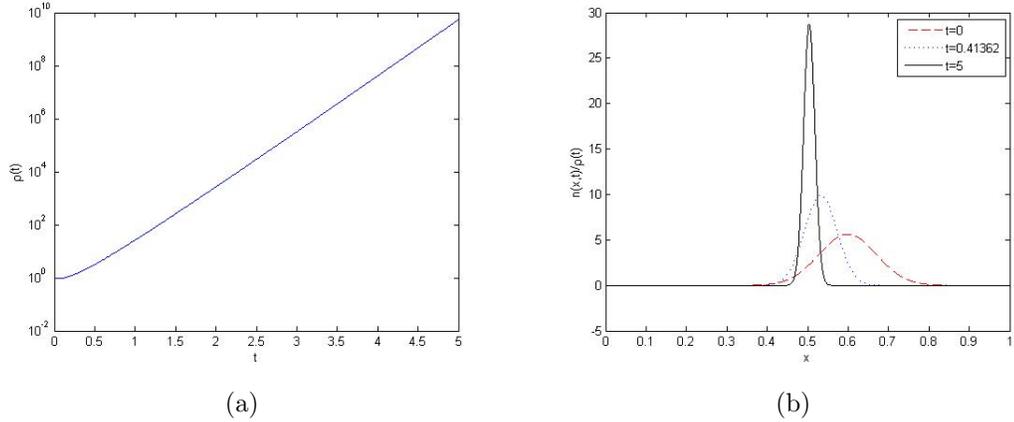


Figure 3.9: Simulations of (3.8) using rates shown in Fig. 3.8. (a) a semi-log plot of the evolution of the density ρ ; (b) the profiles of $n(x, t)$ at three representative times. Note that in the simulations we have made the time scale change $t \rightarrow t\epsilon^{-1}$, where $\epsilon = 10^{-2}$, since we primarily are interested in the long-time behavior.

populations with $x \neq x_*$, but not nearly as fast. For instance, the populations with $x \approx 0.9$ grow exponentially, but not as rapidly as those with $x = x_*$. We also note that uniform initial conditions produce the same qualitative dynamics.

The solution of (3.16), shown in Figs. 3.10 and 3.11, has a different selection strategy. We see that, as in (3.8), initially the trait with the highest growth rate is $x_* = 0.5$, and the population appears to grow monoclally. However, as can be seen in Fig. 3.11(a), there is still a region near $x \approx 0.9$ where $r(x) - c(x) > G(\rho)d(x)$, and thus there, exponential growth is still occurring. Recalling (3.20)–(3.24), the curve $G(\rho)d(x)$ must continue to increase in order for (3.22) to hold for all but a finite number of x . Hence, ρ continues to increase, eventually growing so large that $r(0.5) - c(0.5) < G(\rho)d(0.5)$, causing the death of the $x = 0.5$ cells. In the limit, (3.22) holds for all x except $x_* = 0.8925$, and we have

$$\lim_{t \rightarrow \infty} n(x, t) = \rho_* \delta(x - x_*),$$

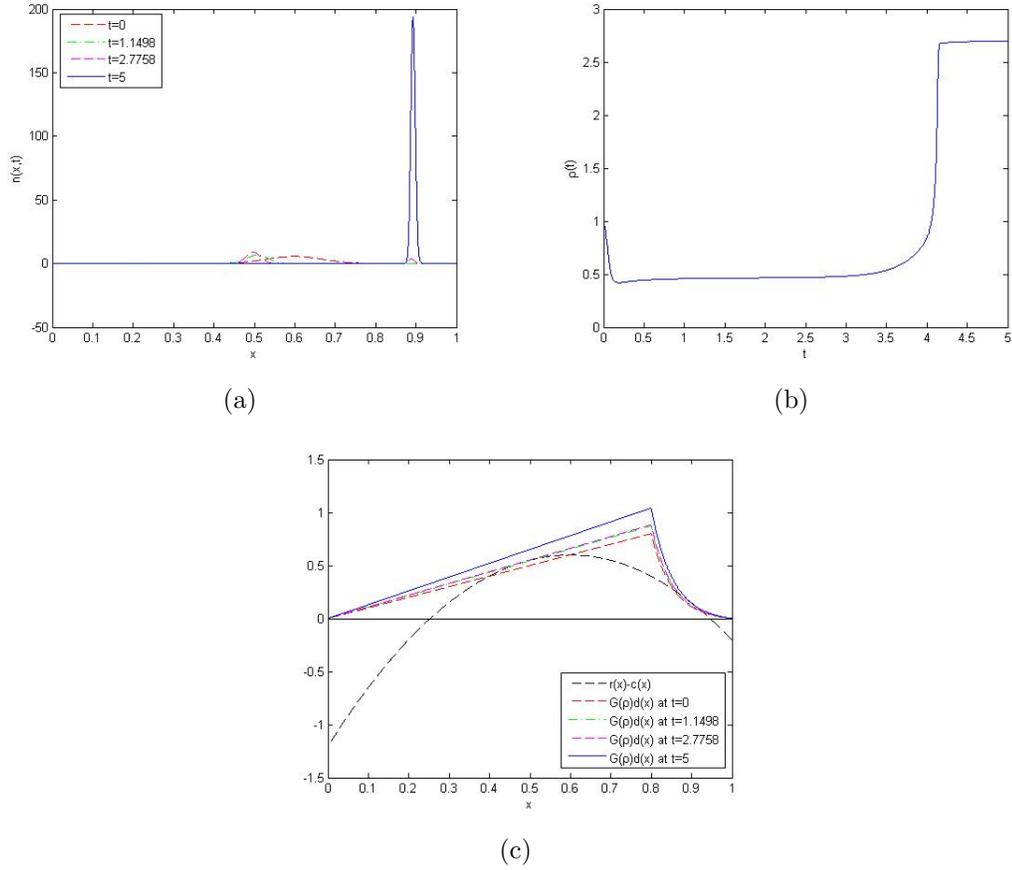


Figure 3.10: Simulations of (3.16) using the trait parameters shown in Fig. 3.8. G is given by (3.19). (a) the profiles of $n(x, t)$ at four representative times; (b) the evolution of the density; (c) the variation of the net-growth rates as the density changes in time, at the same four times used in (a). Note that in the simulations we have made the time scale change $t \rightarrow t\epsilon^{-1}$, where $\epsilon = 10^{-2}$, since we primarily are interested in the long-time behavior.

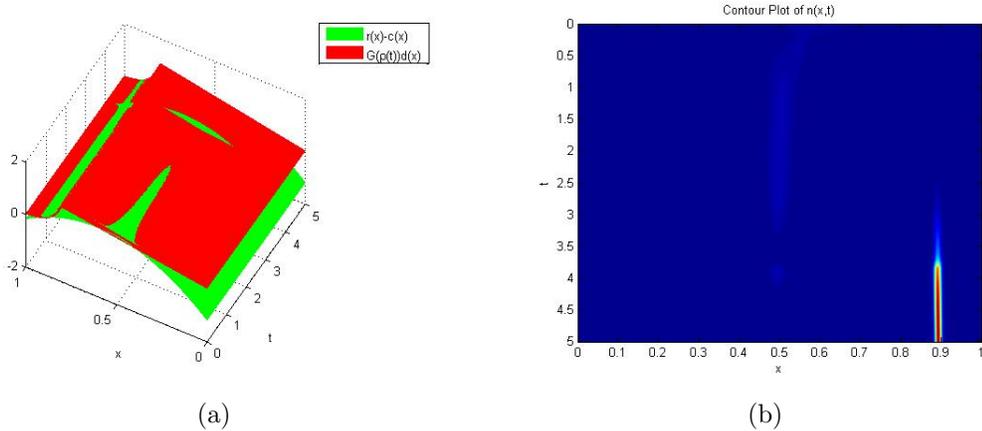


Figure 3.11: Simulations of (3.16) with trait parameters in Fig. 3.8. G is given by (3.19). (a) the net-growth rates as a surface plot over t and x ; (b) a contour plot of $n(x, t)$. Note that in the simulations we have made the time scale change $t \rightarrow t\epsilon^{-1}$, where $\epsilon = 10^{-2}$, since we primarily are interested in the long-time behavior.

as can be seen in Fig. 3.10(a) or Fig. 3.11(b). We see from Fig. 3.10(b) that $\rho_* \approx 2.6$, which means that the inclusion of the density dependence results with the selection of a different trait. Hence, qualitatively different dynamics result in the distinct frameworks.

3.3.3 Selection/mutation model

As discussed in Section 3.3.2, by adding a density dependent cooperation/competition term, we were able to model systems of cells with nonzero yet finite limiting total populations, something that is not possible in the model (3.8). Adding a density dependence was shown to result in richer underlying dynamics. However, the asymptotic behavior of (3.16) always satisfies the following: given that $\exists \rho_* > 0$ such that

$\rho(t) \rightarrow \rho_*$, $\exists m \in \mathbb{N}$, $m < \infty$, $\{x_i\}_{i=1}^m \in [0, 1]$, and $\{a_i\}_{i=1}^m \in \mathbb{R}_+$ such that

$$\lim_{t \rightarrow \infty} n(x, t) = \sum_{i=1}^m a_i \delta(x - x_i),$$

where $\sum_{i=1}^m a_i = \rho_*$. Equivalently,

$$\lim_{t \rightarrow \infty} \frac{n(x, t)}{\rho(t)} = \sum_{i=1}^m a'_i \delta(x - x_i), \quad (3.26)$$

where $0 < a'_i \leq 1$ and $\sum_{i=1}^m a'_i = 1$. In other words, there exists only a finite number of traits that exist asymptotically, with all other populations dying out. Such a result, while an improvement over the dynamics of (3.8), is in contradiction to intratumoral heterogeneity observations. We expect to see a stable distribution centered around a finite number of traits, i.e. distributions with a larger variance. It is worth noting also that the a_i (or a'_i) have a more complex relationship to the initial population values $n(x, 0)$, as opposed to when they were seen in (3.8). This added complexity is due to the density dependence.

With this in mind, we introduce the full model given by Eq. (3.6), which we repeat here for convenience:

$$\begin{aligned} \frac{\partial n(x, t)}{\partial t} &= \left(r(x)(1 - \theta) - c(x) - G(\rho(t))d(x) \right) n(x, t) \\ &+ \theta \int_0^1 r(y)M(y, x)n(y, t) dy. \end{aligned} \quad (3.27)$$

We recall that (3.27) is obtained from Eq. (3.16) by allowing unfaithful divisions, so that a fraction θ of the daughter cells mutate from trait y , to a new trait x . The

rate of transfer is captured in the probability density $M(y, x)$.

We note that if condition (3.17) is satisfied by $G(\rho)$, Theorem 2 remains valid, namely $\rho(t)$ remains bounded for all time t . See Appendix A for a proof.

Typically, when a model includes mutations, a natural assumption is to allow only small mutations to occur during each division. Thus, in this case, the evolution to resistance can be interpreted as the result of a series of many small mutations over a long period of time, which changes the trait space profile of a population of cells. The term small here assesses a close distance in trait space between the original trait x and the new trait y , at least with very high probability. The degree of closeness will be specified by a small parameter ϵ . Accordingly, the mutation kernel can be written in the following form:

$$M(y, x) = \frac{1}{\epsilon} \tilde{M}(y, |y - x|/\epsilon), \quad 0 < \epsilon \ll 1. \quad (3.28)$$

The notation in (3.28) evokes the following: if x and y are within ϵ , then $M(y, x)$ is large, with maximum value $M(y, y)$. Otherwise, $M(y, x) \approx 0$. Hence, mutations are essentially limited to the interval $[y - \epsilon, y + \epsilon]$. In fact, depending on the dependence of \tilde{M} on $|y - x|/\epsilon$, this interval may be smaller (for example, if the dependence were quadratic).

Note that there is still a y dependence outside of the distance term $|y - x|/\epsilon$ in (3.28), due to the fact that (3.4) holds for all $y \in [0, 1]$. In terms of \tilde{M} , this yields

the technical condition that must hold for all y :

$$\int_0^{\frac{1-y}{\epsilon}} \tilde{M}(y, z) dz + \int_0^{\frac{y}{\epsilon}} \tilde{M}(y, z) dz = 1. \quad (3.29)$$

In writing (3.29) we are allowing both forward and backward mutations to occur, as we are considering this model as a basic framework. The above can be easily adapted to consider only forward mutations: simply require $\tilde{M}(y, |y - x|/\epsilon) = 0$ for $y > x$ and ignore the second term on the left-hand side in (3.29).

In our numerical simulations we use a Gaussian mutation kernel of the form

$$\tilde{M}(y, |y - x|/\epsilon) = K(y) \exp(-(|y - x|/\epsilon)^2). \quad (3.30)$$

As always, $K(y)$ is a nonnegative function chosen to ensure that (3.4) holds for all $y \in [0, 1]$.

Since we are interested in studying the asymptotic behavior of (3.27), we can rescale time t by the small parameter ϵ that appears in the mutation kernel. That is, we define the transformation $t \rightarrow t\epsilon^{-1}$ and rewrite Eq. (3.27) in the new time scale together with the form of the mutation kernel given by (3.28) as

$$\begin{aligned} \epsilon \frac{\partial n(x, t)}{\partial t} &= \left(r(x)(1 - \theta) - c(x) - G(\rho(t))d(x) \right) n(x, t) \\ &+ \frac{\theta}{\epsilon} \int_0^1 r(y) \tilde{M}(y, |y - x|/\epsilon) n(y, t) dy. \end{aligned} \quad (3.31)$$

Throughout the remainder of the chapter, by time we will refer to the time that is rescaled by ϵ .

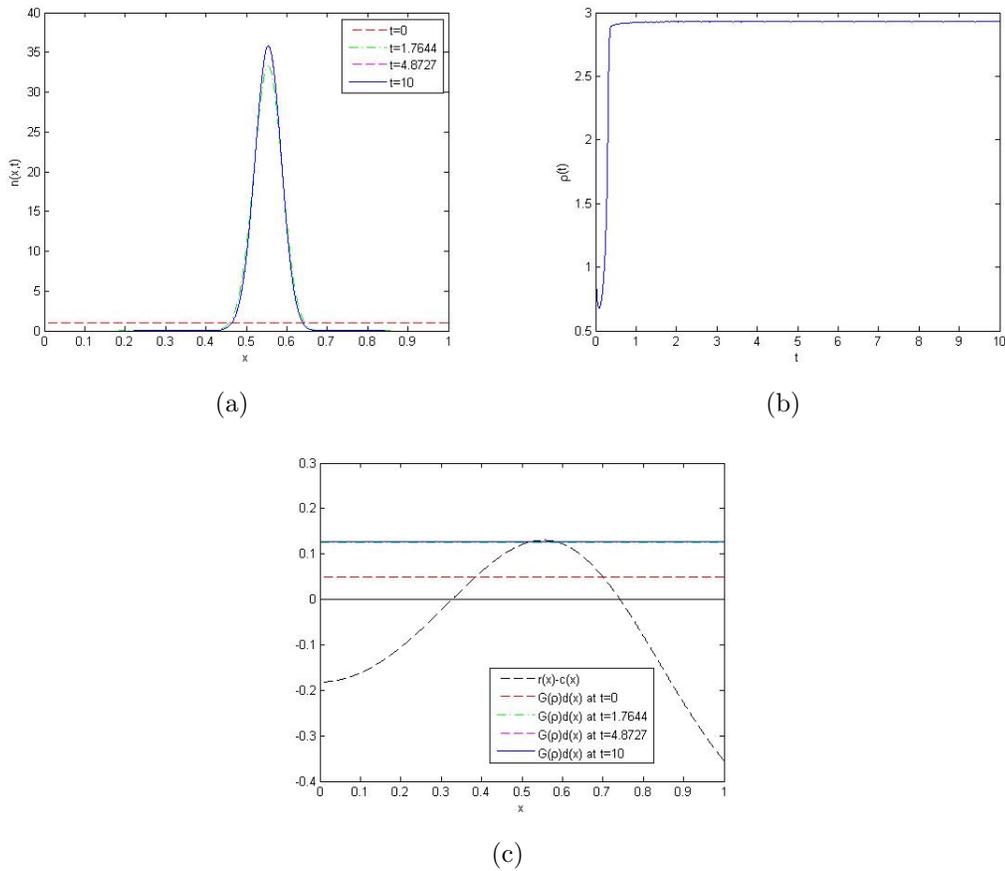


Figure 3.12: Simulations of (3.31), using the rate parameters shown in Fig. 3.1(a). G is given by (3.19). $\epsilon = 0.01$, and $\theta = 0.1$. (a) the profiles of $n(x, t)$ at four representative times; (b) the evolution of the density; (c) the variation of the net-growth rates as the density changes in time, at the same four times used in (a).

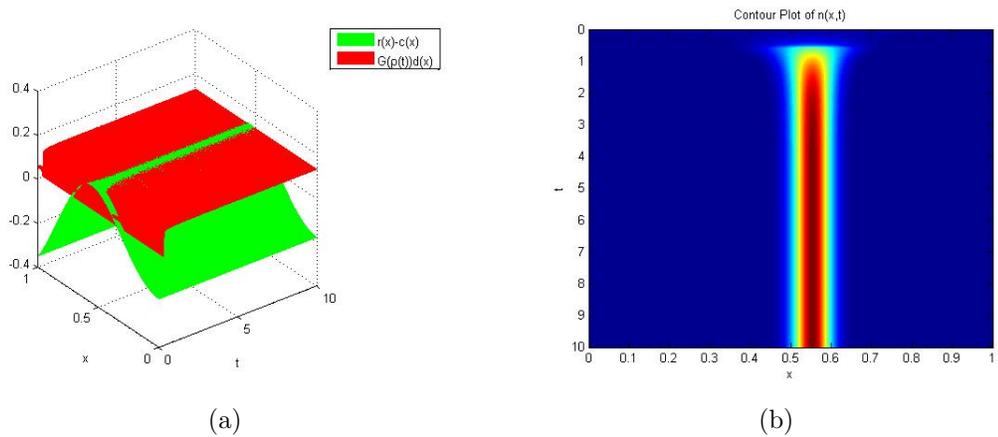


Figure 3.13: Simulations of (3.31), with rate parameters shown in Fig. 3.1(a). G is given by (3.19). $\epsilon = 0.01$, and $\theta = 0.1$. (a) a surface plot of the net-growth rates; (b) a contour plot of $n(x, t)$.

Results of numerical simulations of (3.31) together with (3.30) are shown in Figs. 3.12 and 3.13. In these simulations we use the same rate and density parameters as those that were used in Figs. 3.3 and 3.4. We begin by setting the mutation length (the time-scale parameter) as $\epsilon = 0.01$, and setting the fraction of cells that undergo mutations as $\theta = 0.1$.

The dynamics that we observe from these simulations at first appear similar to what was obtained from Eq. (3.16). From Figs. 3.12(a) and 3.12(b) we see that the greatest density of cells lie with trait $x_* \approx 0.55425$, and the the total population density converges to $\rho_* = 2.929$. However, we see that the cell population densities $n(x, t)$ no longer evolves to a point mass. Instead, they converge to a stable distribution with a finite, nonzero variance, whose mean value will be the same value as given by the locations of the Dirac masses in the corresponding equation (3.16). One can see this either in Fig. 3.12(a), as the curves here corresponding to $t = 4.8727$ and $t = 10$ are vertically identical, or in Fig. 3.13(b), where the population values appear to be stable after $t \approx 2$. Furthermore, the line $G(\rho)d(x)$ does not become tangent to $r(x) - c(x)$ at x_* , so as not to satisfy (3.22) and (3.23). In fact, as seen most readily in Fig. 3.13(a), $G(\rho(t))d(x)$ stays below $r(x) - c(x)$, leaving an x interval such that $r(x) - c(x) > G(\rho(t))d(x)$. We note that this behavior holds for all future time.

Intuitively, it is clear why $n(x, t)$ no longer asymptotically approaches a linear combination of a finite number of Dirac masses, but instead approaches a distribution with nonzero variance. The integral term in (3.31) acts as a diffusion-like operator that balances the growth of the maximal trait with the added death of

mutated daughter cells in regions where $r(x) - c(x) < G(\rho)d(x)$. Despite cells are growing in the regions where $r(x) - c(x) > G(\rho)d(x)$, a fraction θ of the daughter cells are dividing unfaithfully into traits x that satisfy $r(x) - c(x) < G(\rho)d(x)$, and hence are dying. As ρ increases and the region where $r(x) - c(x) > G(\rho)d(x)$ becomes smaller, the gain in faithful divisions is countered by the loss of mutations to regions where cell populations are dying. Our limiting profile in Fig. 3.12(a) is the result of the balance of these two opposed forces. Note that this balance is affected both by the fraction θ of mutating cells and by the mutational distance ϵ .

We would now like to investigate the dependence of the long-time behavior of (3.31) on the parameters θ and ϵ . We perform numerical simulations using the trait parameters as in Fig. 3.1(a) and the density from Eq. (3.19). First, we fix $\epsilon = 0.01$ and vary θ . The results are shown in Fig. 3.14. Note that we only plot n at time $t = 10$, since we have numerically verified that $n(x, t)$ does not significantly change for $t > 10$. We observe two basic phenomena in Fig. 3.14. First, as we allow a higher average fraction of cells to undergo unfaithful division, the distribution of cells about the maximum trait x_* widens. This is intuitive, since increasing the fraction of cells mutating away allows for a greater variability of traits. Secondly, the total mass is nearly constant at $\rho_f := \rho(t = 10) \approx 2.9$, but does slightly decrease monotonically as θ increases. Hence, the main effect in this case seems to be increasing the variance in the trait x of the cell population with mean value x_* . In other words, as θ increases, the heterogeneity in the tumor also increases, with a very slight effect on the overall mass of cells. To understand this, note that ϵ is small in the simulation, so that mutations, when they occur, are localized. Hence, even when a large fraction of

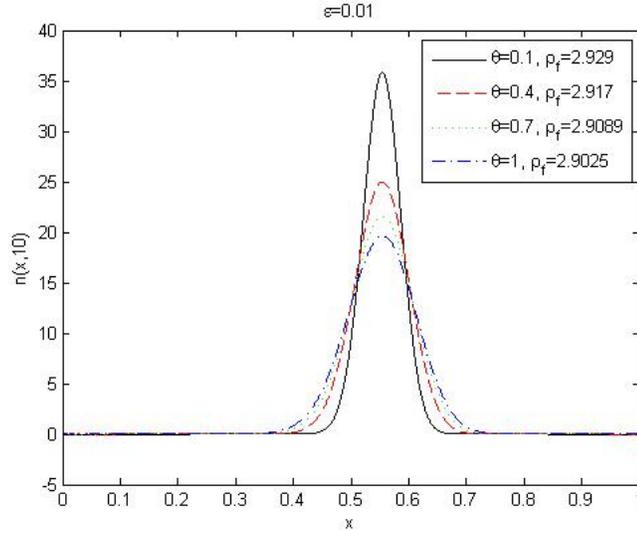


Figure 3.14: Simulations of (3.31), at time $t = 10$. Plotted are $n(x, 10)$ obtained using rate parameters shown in Fig. 3.1(a). G is given by (3.19). $\theta = 0.1, 0.4, 0.7, 1$. For all plots, $\epsilon = 0.01$ is fixed.

cells is mutating (θ larger), they cannot mutate to a distant trait value. By the continuity of the growth/death parameters, the increase in apoptosis to those cells who mutate away from the maximum is small, hence yielding only a small decrease in ρ_f .

We repeat our simulations with a larger value of $\epsilon = 0.1$. We note that in this case, it took a longer amount of time for the simulations to reach the steady state. After $t = 100$, the simulations no longer observably varied, and hence we define $\rho_f := \rho(t = 100)$. The results of these simulations are shown in Fig. 3.15. Once again, the variance of cells about the maximum trait x_* increases as θ increases. However, in this case, the final densities ρ_f decrease much more as θ increases. Since ϵ is larger, daughter cells are allowed to mutate much farther away in an unfaithful division when compared with the smaller ϵ simulations. Hence, these cells jump to regions in trait space x that are more distant from the maximum x_* . Here, the net

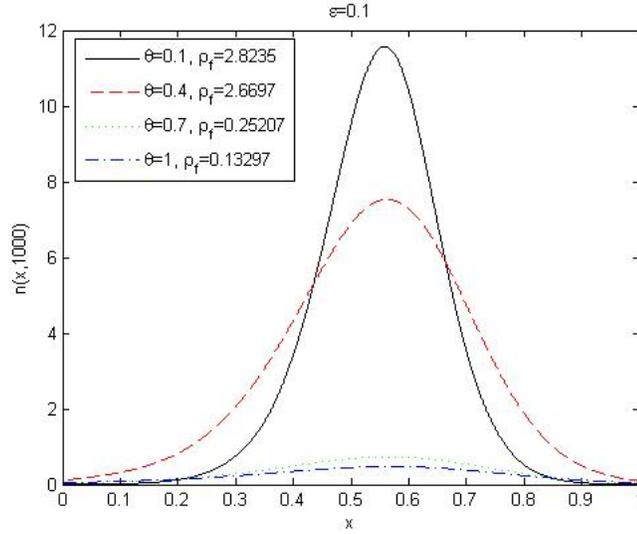


Figure 3.15: Simulations of (3.31) at time $t = 100$. All other parameters are identical to those used in Fig. 3.14.

growth rate is large negative, causing these cells to die quickly and hence have a lower ρ_f when compared to the previous case of $\epsilon = 0.01$. The overall effect leads to a decrease in the overall steady-state populations.

From Figs. 3.14 and 3.15 we conclude that both parameters ϵ and θ affect the degree of heterogeneity observed in the cell population. For a fixed ϵ , we see that as θ increases, the overall population will increase in variance and decrease in steady-state density. The rate of decrease of this density, ρ_f , is strongly affected by the ϵ parameter. Smaller values of ϵ will cause small variation in ρ_f as θ varies, but as ϵ increases, so does the volatility of ρ_f with respect to θ . We note that increasing the fraction of mutating cells can act to decrease the total mass of the tumor. This effect is more pronounced for more widely varying mutations, that is, in models with larger values of ϵ . We finally note that as $\epsilon \rightarrow 0^+$, the dynamics of (3.31) approach the dynamics of (3.16). We refer interested readers to Lorz *et al.* [121], where a

rigorous analysis of such a limit is undertaken. However, here we do not consider such a limit to be appropriate from a modeling point of view, as our goal was to achieve local heterogeneity, which cannot be achieved with Eq. (3.16). Accordingly, our study only considers the case of a finite, nonzero ϵ .

3.4 Discussion

The goal of this chapter was to develop a modeling framework for MDR that can qualitatively capture the complexity of the underlying dynamics. To achieve this goal, we introduce a hierarchy of models, which we have studied analytically and numerically.

Our study starts with a simple uncoupled system of ODEs parameterized by the resistance level x , and given by Eq. (3.8). We demonstrated that this model can capture the selection of a finite number of traits. However, in such models, the total cell mass either diverges or dies off, an outcome that contradicts the biology of cancer cells.

This motivated the introduction of a cooperative/competitive density dependence into the model, i.e. Eq. (3.16). In this case, the emerging dynamics were similar to (3.8), in that the limiting distribution is a finite collection of Dirac masses. Now though, under mild conditions on the density dependence (see (3.17)), the total density remains bounded. Furthermore, we demonstrated that the final selected trait (or traits) can be different between (3.8) and (3.16). While the model (3.16) has clear benefits in comparison with (3.8), the homogeneity of the cell populations

is in contrast to the biological literature, as discussed in Section 3.1.

Finally, we introduced a mutation mechanism into the model via an integral term, given by (3.31). The dynamics of this model were then demonstrated to be similar to those in (3.16), in that the population densities were maximal at the same trait values x , and the population density remains bounded. However, in this mutation/selection model, the Dirac masses are replaced in the limit by distributions with nonzero variance. That is, there exists a continuum of traits that are stable for all time. This is interpreted as tumor heterogeneity, which was the desired outcome.

Mathematical models offer ways to predict the net-growth rate during and after the administration of a treatment. These mathematical models and their outcomes each have their own definitions and assumptions of growth. What function best describes the growth dynamics of cancer, and what are the consequences of using different growth descriptions in MDR? In this work, we have developed an alternative cancer growth model to the exponential, logistic, or Gompertzian growth, which also integrates resistance level. This work has implications in the clinical setting for solid tumor and metastasis studies, and for more details, one should see our recent published works ([136,137]). This issue is a critical step toward the ability to predict the dynamics of drug resistance and intratumoral heterogeneity in cancer during and after therapy. Accordingly, we view the model (3.31) as a basic framework that can be extended to model more realistic MDR mechanisms. In particular, providing experimentally solid estimates for the birth and death rates ($r(x)$, $d(x)$, and $c(x)$) along with the density dependencies ($f(\rho)$ and $g(\rho)$) will be the focus of future work.

Chapter 4: Clinical implications of drug resistance as a continuous variable

In this chapter, we present applications of our adaptive dynamics drug resistance model to clinical scenarios. Specifically, the framework is extended to include dosage strategies as well as epigenetic mutations. Treatment protocols are also examined, and the mathematical models are applied to individual patient data. Both sections in the chapter have been published in biologically oriented journals (*Cancer Research* and *Trends in Molecular Medicine*, respectively).

4.1 The role of cell density and intratumoral heterogeneity in multidrug resistance

The content of this section has been published in [136]. Recent data have demonstrated that cancer drug resistance reflects complex biological factors including tumor heterogeneity, varying growth, differentiation, apoptosis pathways, and cell density. As a result, there is a need to find new ways to incorporate these complexities in the mathematical modeling of multidrug resistance. Here, we derive a novel structured population model that describes the behavior of cancer cells un-

der selection with cytotoxic drugs. Our model is designed to estimate intratumoral heterogeneity as a function of the resistance level and time. This updated model of the multidrug resistance problem integrates both genetic and epigenetic changes, density-dependence, and intratumoral heterogeneity. Our results suggest that treatment acts as a selection process, while genetic/epigenetic alterations rates act as a diffusion process. Application of our model to cancer treatment suggests that reducing alteration rates as a first step in treatment causes a reduction in tumor heterogeneity, and may improve targeted therapy. The new insight provided by this model could help to dramatically change the ability of clinical oncologists to design new treatment protocols and analyze the response of patients to therapy.

4.1.1 Introduction

Resistance to chemotherapy remains a major cause of the failure of cancer treatment. This resistance results from numerous mechanisms. The “traditional” understanding of multidrug resistance (MDR) and its driving mechanisms oversimplifies the complexity of a perturbed cellular cancer network and focuses on several pathways/gene families. From that perspective, drug resistance is associated instead with the induction of drug efflux, activation of DNA repair, variations of target proteins, decreased drug uptake, altered metabolisms, sequestration, and changes in apoptotic pathways [49, 96].

Recently, intratumoral heterogeneity has also been found to be a major facilitator of drug resistance. Intratumoral heterogeneity refers to differences between

cancer cells originating within the same tumor. Many primary human tumors have been found to contain genetically distinct cellular subpopulations reported to be mainly the result of stochastic processes and microenvironment signals (reviewed in [33]). In addition to the genetic differences within a tumor, therapeutic resistance can also be caused by several other non-genetic processes, such as epigenetic changes associated with chromatin modification or DNA methylation [27–30]. One study of these processes was done by Kreso *et al.* [131,132], who found, in a system with a single genetic clone, that there was functional variability among the tumor cells. Clearly, the integration of both genetic and non-genetic assumptions as well as heterogeneity should be included in the design of new experimental and computational models in order to have a better description of and ultimately, a solution to the problem of MDR.

In order to study intratumoral heterogeneity, we consider a mathematical modeling approach. As mentioned above, drug resistance is far from being a black-and-white, resistant or not resistant phenomenon. Accordingly, a continuous variable is a more appropriate way to describe, estimate and measure the resistance level. This estimator is a key variable in any mathematical representation of the MDR system and without it, a comprehensive mathematical model cannot be developed.

Several direct and indirect approaches have been suggested to estimate the drug resistance level, depending on the type of data that is analyzed. For instance, in *in vitro* experiments, the dose-response assay (e.g., the MTT assay) can quantify the number of surviving cells after exposure to different drug concentrations for a certain time period and can be presented by “killing curves”. The 50% inhibitory

concentration (IC_{50}) values can be defined as the drug concentrations required to reduce cell viability to 50% of the untreated control population. Thus, for example, the resistance level can be described here by the IC_{50} value of each clone in the global population. A similar trend in killing curves would be expected, to some extent, for other drugs with similar features (targets, mechanisms, etc.). A linear generalization of such an approach would be the number of different drugs that can be separately applied to those cells and yet the cells still survive, where the level of resistance can be calculated as a score of two variables: the number of drugs and the IC_{50} value of each drug. A non-linear generalization would be the survival percentage of the treated population with drug combinations administered at the same time point. In all of these cases, the higher the score, the higher the resistance level.

Unfortunately, most clinical data do not include the IC_{50} values and the *in vitro* conclusions have not led to success in the clinic [138]. Usually, clinical data include the physiological properties that describe the progress, extent or severity of a tumor (“staging”). All assignments of cancer stage are made at the time of diagnosis, before any treatment is given and thus cannot directly assess the resistance level. Combining clinical data with gene expression and survival data from the same patients can help to categorize them as “good” or “poor” responders, and a score for their resistance level can be calculated. Accordingly, any theoretical model should include subpopulations with resistance levels that can vary within the interval between “good” and “poor” scores.

The number of cellular mutations has also been proposed as a way to charac-

terize resistance level. Due to the stochastic nature of the mutation process, there are mutations that do not necessarily contribute to cancer progression, and are not essential to the resistance level of a tumor. Yet such mutations certainly increase the intratumoral heterogeneity. Of course, once the number of passenger mutations accumulates to a certain level, they can be expected to have a global effect on tumor growth and sensitivity to certain drugs [139]. The number of mutations does not necessarily go hand in hand with the resistance level, but rather the type of mutated pathways affects the evolution of MDR. For instance, mutations in the apoptosis pathway cause a decrease in the death rate [140], mutations in the RAS-RAF pathway cause increased cell proliferation and resistance to apoptosis [141], and mutator genes increase genetic alterations throughout the genome. Moreover, there are certain genes that promote genetic stability, including DNA repair genes, DNA damage sensor genes and cell cycle checkpoint genes. Changes in these stability genes affect the mutation rate and therefore this rate is not necessarily a constant parameter over time and space [142].

Many mathematical models related to MDR have been developed (reviewed in [66]). Most of these models quantify the resistance level as proportional to the number of mutations. Some models quantify the resistance level by considering the expression of a single gene. Our work can be viewed mainly as an extension of the mathematical model of Lorz and colleagues, who developed a physiologically-structured population model that is structured according to a continuous variable that can represent the level of a phenotypic trait that has been selected as relevant to describe population heterogeneity [143]. Lorz and colleagues assumed that the

cell division rate is dependent on the cell density of normal cells, while cancer cells are assumed to grow exponentially with a low mutation rate. Inspired by their model, we propose an extended mathematical model for cancer cells with cell density dependence on cell division and death rates, and account for both genetic and epigenetic changes. Moreover, we describe the drug-induced death rate as a function of the dose and the trait. This work is important for optimizing the efficacy of treatment. It has obvious implications for understanding the complexity of cancer growth dynamics and for the analysis of drug resistance data.

4.1.2 Quick guide to equations and assumptions

In this chapter, we study the same mathematical model introduced in Chapter 3. In particular, the governing equation is (3.2), where we introduce the form

$$h(D(t), x) = c(x, \alpha) \tag{4.1}$$

to account for a constant treatment dosage $D(t) \equiv \alpha$. For more details, we refer the reader to Section 3.2.

We demonstrate in detail how the model can qualitatively predict three common biological observations based on density limitation, but with different assumptions about the cell division, death and alteration rates of cancer cells. In the first case, we study the cell density response as a function of the cell division and death rates, where the drug-induced death rate could be generalized as a function of the drug dose and the resistance level. We show how the relationship between these rates

can create a heterogeneous tumor over time, and illustrate how different treatments could decrease the variation in the heterogeneity by selection, in the effective concentrations range. In the second case, we discuss the effects of genetic/epigenetic permanent alterations that occur in different fixed rates. In the third case, we present the contributions of additional temporal changes that can evolve from specific environmental causes. For full details concerning the parameters and numerical results, see Appendix B.

4.1.3 Results

4.1.3.1 Heterogeneity arising from treatment

The first case we study is a hypothetical scenario in which the system cannot have genetic or epigenetic changes, and its dynamic is solely driven by cell division and death rates (Fig. 4.1A). Since the mass of the tumor is limited by the density constraints (see Theorem 2), the model predicts a stable distribution of cells with different traits after a finite time without any treatment (Fig. 4.1B, C_0 lines for IC_1 and IC_2). The administration of chemotherapy to such a system leads to changes in the original distribution and a selection process that depends on the drug-induced death rate, $c(x, \alpha)$. This rate of death is a function of the drug dose, α , and the resistance level, x .

Drug-induced death rate based on a fixed dose. If the dose is administered at a fixed value α , then $c(x, \alpha)$ may be replaced by $c(x)$. The drug-induced death rates of two qualitative examples of different drugs are plotted in Fig. 4.1A as

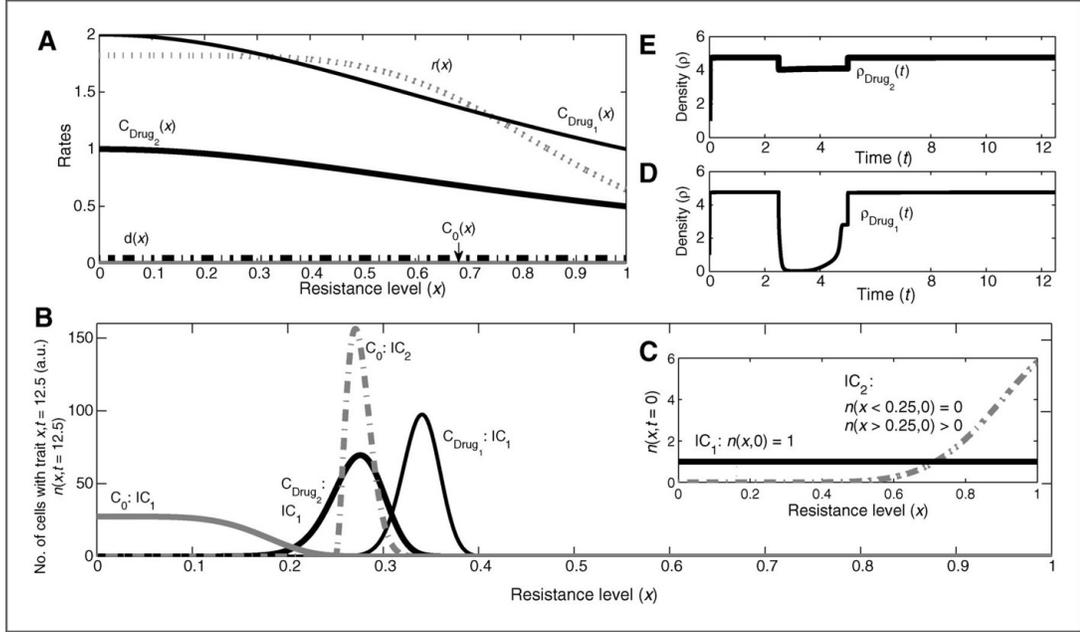


Figure 4.1: Heterogeneity arising from the administration of chemotherapeutic drugs, without genetic/epigenetic changes. (A) rate of cell division $r(x)$; natural death $d(x)$; and drug-induced death for a given dose $c(x)$. All rates are functions of the trait x . The two types of treatments are administered separately, in which drug 1 is a more effective treatment than drug 2. (B) numerical predictions of the net growth ($n(x, t)$) after a certain time ($t = 12.5$), when no alterations (e.g., mutations) occur, with different initial conditions (solid and dashed lines, IC_1 vs. IC_2). In addition to the two treatments, the control (growth without treatment) is also plotted and labeled as C_0 . This plot shows the selection process resulting from treatment with respect to the trait, and demonstrates the time scale it would take to develop a fast-growing tumor. (C) initial conditions, IC_1 : $n(x, t = 0) = 1$ for all x (solid curves); IC_2 : $n(x \leq 0.25, t = 0) = 0$, or $n(x > 0.25, t = 0) > 0$ (dashed curves). (D) and (E), cell density, $\rho(t)$, for the two types of treatments: drugs 1-2, with the initial condition of IC_1 .

C_{Drug_1} and C_{Drug_2} . The spikes in Fig. 4.1B labeled as C_{Drug_1} and C_{Drug_2} show tumor homogeneity, since most cells have similar traits. The C_{Drug_1} curve has a spike at a higher resistance level (i.e., x -value). Fig. 4.1A shows that higher x -values correspond to lower values of $r(x)$, i.e., slower cell division rates, an assumption made when the model was developed. If only a few cells with different resistance levels survive the treatment, that would be sufficient to allow the distribution of intratumoral heterogeneity to return to its original stable profile (Fig. 4.1B, C_0 : IC_1 curve), because the cell division and death rates have not changed. The time needed to reach that control state varies based on the treatment and its drug-induced death rate. Although if after the treatment several subpopulations located in the non-zero range of the original stable profile (Fig. 4.1B, C_0 : IC_1 curve) die out, the system would not restore this distribution (Fig. 4.1B, C_0 : IC_2 curve). Kreso and colleagues provide evidence for a slowly proliferating cell population in primary human colorectal cancer cells that still retains potent tumor propagation potential, thus preferentially driving tumor growth after chemotherapy [131]. A similar mechanism that reduces cell proliferation and induces a dormant non-dividing state was reported by Lewis, in his review on MDR in microbiology [143]. Kreso and colleagues noted that although cancer cells may have a uniform genetic lineage, individual tumor cells are functionally heterogeneous, with a wide variety in terms of growth dynamics and response to therapy. Dividing cells are most likely to be eradicated first, whereas the relatively slower growing/dormant cells are the major population during tumor re-initiation after chemotherapy. Of course, in the study of Kreso and colleagues there could be non-genetic alterations that could change one resistance level to another.

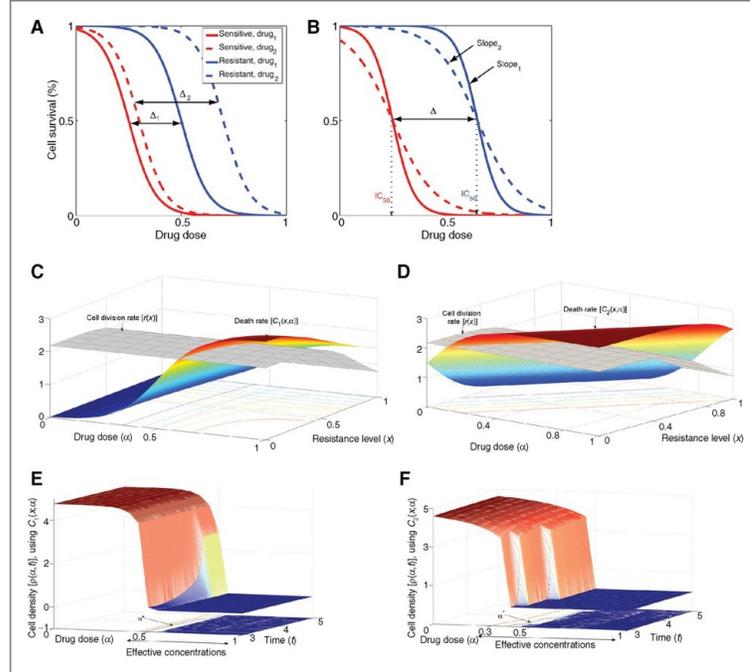


Figure 4.2: Death rate as a function of dose and resistance level. (A) and (B) two common types of treatment (dashed and solid curves) for cells with different resistance levels (red curves for sensitive cells and blue for resistant cells). The cells may react differently to certain drugs. For example, a very toxic drug produces a lower IC_{50} value (solid vs. dashed lines), with a smaller distance between the two curves representing these cells (Δ_1 vs. Δ_2). Also, the slope can demonstrate the response to a treatment, in which the steeper slope is the desired one (slope₁ vs. slope₂). (C) and (D) two examples of death rates ($c_1(x, \alpha)$ vs. $c_2(x, \alpha)$), in which the heights and gradients of the functions vary according to the type of treatment. (E) and (F) the cell density as a function of the dose and duration of a treatment for the cases described in C and D.

In our theoretical case, we show that even if the cells are able to mutate/change to another resistance level, it is possible to have the same dynamic, if only a few slowly proliferating cells that resist the drug remain. Once there is a density limitation, the distribution of cells as a function of the trait would be totally different from the control.

Drug-induced death rate as a function of the dose. Drugs are usually categorized by their biochemical mechanisms of action, targets, molecular structure,

gene expression and in general by the cell response. However, from a mathematical perspective, there is another dimension to drug efficacy, which is the effect of the drug on the shape of the survival curve. Given two sub-populations, resistant and sensitive cells, the efficacy of a treatment could be estimated by the IC_{50} values, the slope of the survival curves and the distance between the two curves of the two cell types (Fig. 4.2A,B). The slopes could be compared at their IC_{50} values (i.e., at their maximum values). The differences in treatment outcomes are not expressed solely by the percentage of cells that survive, but also by the effect on the remaining cells as a function of the dose and the resistance level. A better treatment would result in a lower number of surviving cells and a more homogeneous population. This can be expressed as survival curves with low IC_{50} values (solid vs. dashed lines, Fig. 4.2A), with steeper slopes ($slope_1$ vs. $slope_2$, Fig. 4.2B) and a smaller distance between the two curves (Δ_1 vs. Δ_2 , Fig. 4.2A) with different levels of resistance. These three factors influence the drug efficacy.

Data from survival curves, drug sensitivity assays, and apoptosis assays can be used to estimate the drug-induced death rate in at least two ways when incorporating the features that were mentioned above. Here, we describe two different functions ($c_1(x, \alpha)$ and $c_2(x, \alpha)$) in order to accommodate the two different biological data sets that indirectly measure the drug-induced death rate as a function of the dose (α) or the trait (x). Each data set supplies different pieces of information that are needed to estimate the death rate. Thus, different assumptions should be made when using each type of data set. The first type involves data that measures the percentage of apoptotic cells for a range of resistance levels. Usually, this kind of data is based on

different staining of a tumor from a patient treated with a specific dose. Thus, the data include information on the death rate as a function of the trait, and can be used to infer the dependence of function $c(x, \alpha)$ on the dose. The second type of data set is mainly from *in vitro* experiments. This type of data measure the survival curves for a range of drug concentrations for a few cell populations. The data includes information on the death rate as a function of the dose, and can be used to infer the dependence of function $c(x, \alpha)$ on the trait. In general, both functions $c_1(x, \alpha)$ and $c_2(x, \alpha)$ would be expected to be non-increasing functions of the dose, for a given resistance level. For high doses, the rate is at its maximum value, while for low doses, the rate is at its minimum value. Thus, there is a range of effective doses that this function of the death rate can meaningfully model and a sigmoid function is a very natural function to use in this case. The second variable of the death rate is the trait (i.e., resistance level). For a given dose in the effective range, the rate would be expected to decrease as a function of the trait. Theoretically, with these two types of data and the drug efficacy factors, the drug-induced death rate is illustrated in Fig. 4.2C-D. Fig. 4.2C demonstrates the case where this rate can be written as a product of the death rate for a typical dose and with the dose response as a decreasing sigmoid. The function in Fig. 4.2D, in contrast, demonstrates a case where the survival curves are used to describe the function $c(x, \alpha)$ as a complex, non-separable dependence. All functions are listed in Appendix B.

Using our model, we estimated the cell density as a function of the dose and duration of a treatment (Fig. 4.2E,F). Our results show, in general, the desired sigmoid response to a drug as a function of the dose and describe two factors that

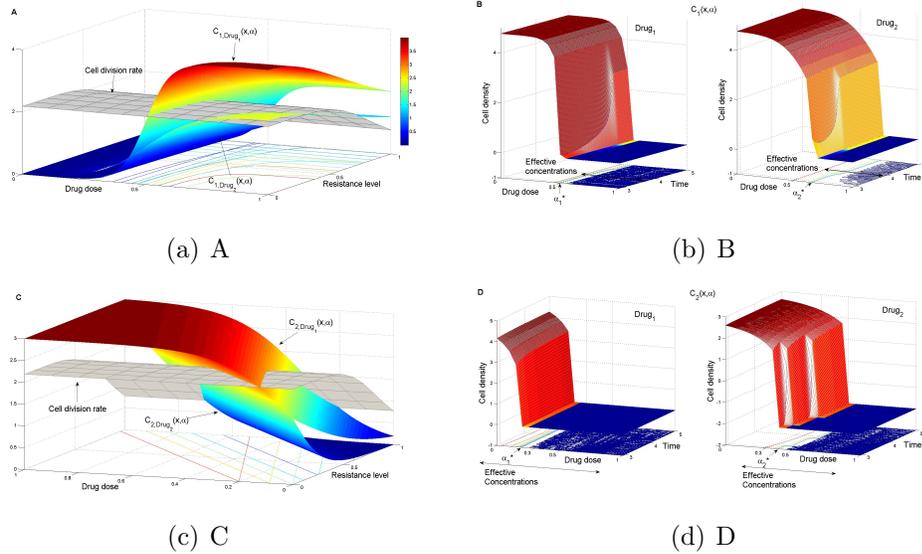


Figure 4.3: Drug efficacy as a function of the dose and resistance level. Panels A and C represent the two common types of drug-induced death rates, $c_1(x, \alpha)$ and $c_2(x, \alpha)$ that are plotted in Fig. 4.2C and D, respectively. Given the exact function of the drug-induced death rate, in each panel (A and C), two different treatments were considered (drug 1 and drug 2), in a range of drug concentrations, for cells with different resistance levels. Panels B and D illustrate the population response during the treatments, and the density is plotted as a function of the drug dose and time. The shapes of the density curves vary according to the type of treatment and type of drug-induced death rate that were assumed. In all cases, a better treatment would give lower cell density at a lower drug dose (α^*), with a wide range of effective concentrations.

are important to the efficacy of a treatment, the effective-concentration range and the critical-drug concentration (α^*). The effective-concentration range is the range of drug concentrations where the cell density is decreased. It is preferable to have as wide a range as possible. The critical drug concentration is the value where the cell density reaches a threshold for the first time (e.g., minimum tumor size that is detectable by the scanning instrument). Clearly, a low threshold is preferred. Furthermore, for each death rate (C_1 or C_2), we compared the cell density for two different drugs (Fig. 4.3A,C). A more efficient treatment has a wider effective-concentration range, a lower critical-drug concentration, and a steeper cell-density slope, as occurs with drug 1 in both cases (Fig. 4.3B,D).

In addition, during treatment the cell density is shown to be a more complex curve than a simple sigmoid curve. The success of a treatment is related to the shape of the corresponding density curve, see for example the density curves based on treatment of drug 1 vs. drug 2 in both Fig. 4.3B,D. The density response is a nonlinear function of the cell division rate, drug-induced death rate and the density limitation in the effective concentrations range. For many drugs, the cell division rate also changes as a function of the dose, with higher doses usually leading to lower division rates. In practice, most drugs that cause a decrease in cell division rate also cause a decrease in the death rate. The strategy of combining multiple drugs, in which the first drug causes an increase in the death rate and the second drug causes a decrease in the cell division rate, may result in a better response, at lower drug concentrations.

4.1.4 Heterogeneity arising from both epigenetic and genetic changes

Spontaneous genetic mutations occur infrequently, but some cancer cells have increased rates of genetic change for various reasons. For example, in breast cancer, mutation of a single gene, *Brca2*, can lead to increased rates of mutations in other genes [144]. The tumor microenvironment can also cause genetic changes. As tumors are exposed to repeated cycles of hypoxia and reoxygenation, DNA repair pathways are downregulated, thus leading to genetic instability [145, 146]. Some of the mechanisms that lead to stable genetic changes are due to epigenetic alterations. Genome-scale genomic and epigenomic analyses have only recently revealed the complexity of the nonlinear relationship between these two types of changes and cancer. While genetic mutations alter the sequence of DNA, epigenetic changes do not (reviewed in [147, 148]). Yet epigenetic changes can be inherited by causing mutations and silencing important genes by methylation (e.g., *MGMT*, *CDKN2B* and *RASSF1A*), with a greater rate of epimutations than the rate of spontaneous genetic mutations [147]. On the other hand, not all epigenetic changes are inherited, so they may lead to relatively rapid phenotype switching. In addition, these two types of changes can evolve from different causes. Genetic mutations can occur as a result of random processes or may be induced by stress (e.g., hypoxia, nutrient starvation, toxic molecules etc.), whereas epigenetic alterations are mainly linked to stress. It has also been shown that a central epigenetic control circuit can be disrupted by a genetic mutation [148]. Clearly, the complexity of genetic and epigenetic networks has not yet been fully characterized, though it is clear that the outcome largely

depends on external stresses. Thus, in the following cases we assume an indirect relationship between both networks that is based on an external stress.

Stable changes. The second case considers the possibility of having epigenetic and/or genetic stable alterations when cells undergo division. Of note, $M(y, x)$ denotes the probability that, given an alteration, a mother cell with trait y will yield a daughter cell with trait x . We considered $M(y, x)$ to be a Gaussian distribution conned to $[0, 1]$ with mean y and variance $\epsilon^2/2$ (see equation (3.30)). The variance is one of the most important factors that must be considered, since it reflects the impact of an alteration on the global system. Hence, the changes in variance can be written as a function of the external stress and time, $\epsilon(\text{stress}, t)$. As time progresses, different traits become advantageous or disadvantageous, and the overall dynamics are determined by various rates, mutation parameters, and the initial distribution of the cells. Fig. 4.4 summarizes several examples of three variances $0 = \epsilon_0 < \epsilon_2 < \epsilon_1$ before (Fig. 4.4A) and after (Fig. 4.4B,C) treatment in a given effective dose. From these examples, it is obvious that different assumptions cause a divergence in predictions related to heterogeneity, even when using a single model. The general trend in these results indicates that mutations/epimutations (even at low rates) can serve as a diffusion process, over x , since the cells with a given trait may spread to an interval of x -values that initially was set to zero. As expected, cases with high variation have high stable heterogeneity, whereas small variation slightly change their dynamic from the non-altered case. All other values and sequences of ϵ are constrained in that range, and affect the time to reach the control state ($\epsilon_0 = 0$).

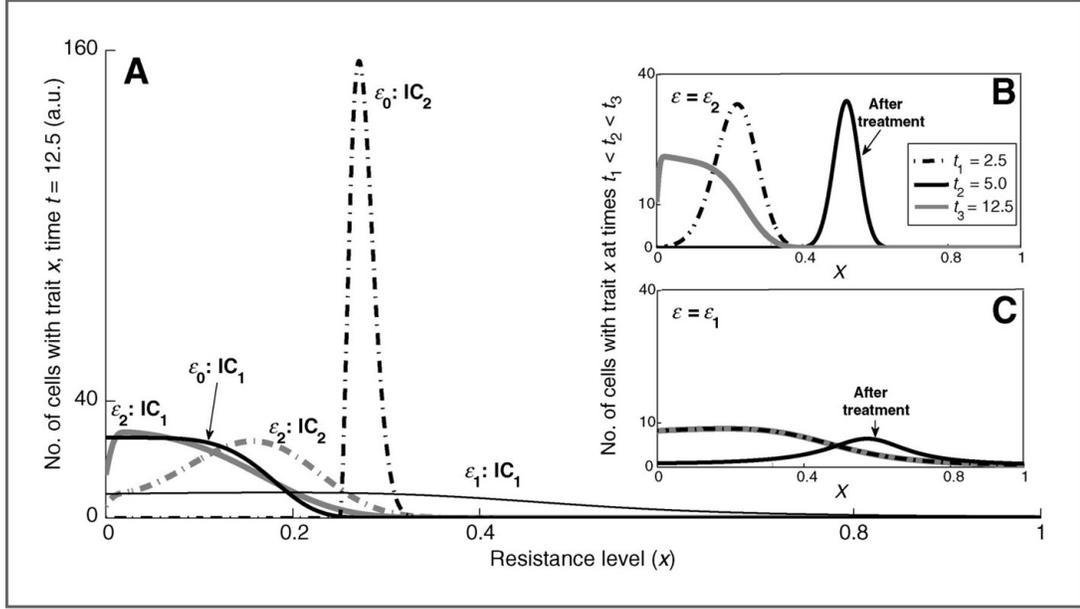


Figure 4.4: Heterogeneity arising from genetic/epigenetic changes. (A) numerical predictions of the net growth ($n(x,t)$) after a certain time ($t = 12.5$), when no chemotherapeutic drugs were administered. The effects of genetic/epigenetic changes act as a diffusion process, where $\epsilon_0 = 0, \epsilon_1 = 0.01, \epsilon_2 = 1$ with two initial conditions (solid and dashed lines, IC_1 vs. IC_2). The initial conditions are shown in Fig. 1C. This graph shows the cases in which the diffusion potentially controls the tumor dynamic and its heterogeneity. (B) and (C) net growth ($n(x,t)$) plotted at three time points: $t_1 = 2.5; t_2 = 5; t_3 = 12.5$, with initial condition IC_1 . The administration of the drug takes place over a certain period of time ($t_1 < t < t_2$), for a given dose. These graphs demonstrate the effect of varying the alteration rates on the dynamic in the presence of anticancer drugs ($\epsilon_1 = 0.01$ in (B), $\epsilon_2 = 1$ in (C)).

Temporal changes. In the third case, the additional contribution of the M function is only valid for a specific period of time, when certain signals stress the tumor. We studied two possible mathematical modifications to system (3.2). The first system assumes an accumulation of changes that occur with different time scales. Thus, we separate the alteration function into two functions. The first type of changes is related to hereditary alterations, and it is therefore applied for all time points, and the second type is applied only to a limited period of time, since it is based on temporal epigenetic variations (marked as M_1 and M_2 , respectively). As mentioned above, there is an indirect assumption of interaction between the M functions, which is qualitatively expressed by the activity timing of M_2 and the ϵ of both functions. We studied the effect of ϵ , and its variation in each M function, on the dynamic with drug for a given dose, $c(x)$. Fig. 4.5 summarizes several examples of different variances for both functions post stress. See Appendix B for more details.

These results highlight the tradeoff between changing the alteration rate and the drug selection process, which in turn affects the heterogeneity of the cancer cell population. If the alteration rate remains low during the entire process, the treatment would be more efficient, and the surviving cells would constitute a more homogenous population (Fig. 4.5B). Once this selection is achieved, it is recommended that the second treatment should be more specifically targeted to the surviving population. However, if the first treatment is known to increase the alteration rate, a combination of drugs should be applied instead, where one of these drugs (such as a drug that affects pattern of gene expression) aims to reduce this rate. On the other

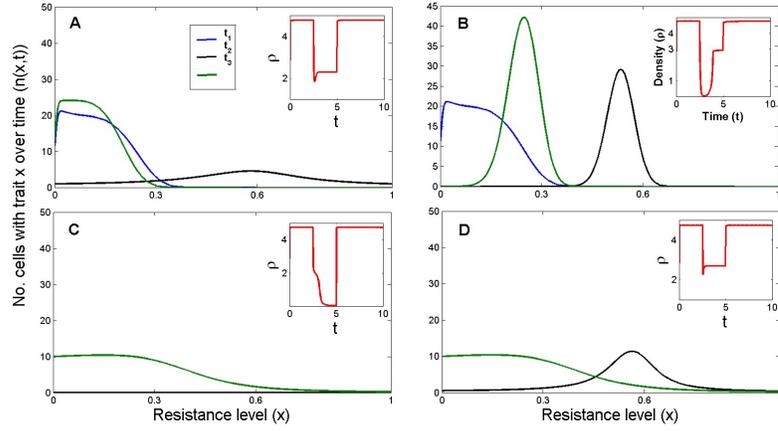


Figure 4.5: Dynamics with variations in the mutation rates over time with/without drugs. In all of these cases, there is an initial alteration (ϵ_1) fixed over time, and an additional alteration rate (ϵ_2) applied only when the drug (C_1) was applied for a certain period of time ($t_1 < t < t_2$, where $t_2 < t_3$). (A) $\epsilon_1 = 0.01, \epsilon_2 = 1$. (B) $\epsilon_1 = 0.01, \epsilon_2 = 0.01$. (C) $\epsilon_1 = 1, \epsilon_2 = 1$. (D) $\epsilon_1 = 1, \epsilon_2 = 0.01$.

hand, if the initial alteration rate is too high and there is no effective treatment to reduce it, the model predicts (see Fig. 4.1A and red curve in Fig. 4.52C) that increasing the alteration rate, at the time of treatment, should also result in a better response to the selective drug. In any event, at the end of the treatment, the rate of alteration must be decreased as much as possible.

The second modification to system includes the temporal epigenetic changes by varying the parameters θ and ϵ based on the external stress ($\theta(\text{stress}, t), \epsilon(\text{stress})$, with a single M function (Appendix B, equation (B.2)). During treatment, the results show that if θ increases, with a constant set of parameters, the density decreases but with a similar Gaussian shape of distribution. The cell density becomes significantly low when θ reaches a certain level (see Appendix B and Figs. 3.14 and 3.15). A possible biological interpretation for the parameter θ is the percentage of proliferating cancer cells based on stress.

4.1.5 Discussion

Enormous progress has been made in understanding the molecular mechanisms leading to cancer, but our perspective on resistance to treatment is still in its infancy. A systems approach combines empirical, mathematical, and computational methods to gain an understanding of this complex phenomenon. Recent data show that the diverse intratumoral heterogeneity is mainly the result of stochastic processes and microenvironment signals. Thus, it is important to quantify the resistance levels in a non-discrete manner. Resistance evolves parallel to the progression of cancer. However, not all abnormal changes contribute to the resistance level. Hence, the next experimental efforts should incorporate a characterization of the intratumoral heterogeneity based on the resistance level, since this classification plays key roles in the development of MDR and in modeling such a system.

Here, we have presented a mathematical analysis of the MDR problem with the assumption that resistance is induced by adaptation to drug and/or environmental signals in a deterministic manner. We have used physiologically structured equations in which a continuous variable represents the resistance level. Such a variable may be generated by integrating a collection of physiological measurements, as discussed in the introduction. Using our model, one can explain how the density plays an important part in tumor dynamics. Adding this density dependence results in a different distribution/heterogeneity.

Many observations of heterogeneous tumors have been reported [33], but not all of those tumors were sampled after drug treatments [149]. Therefore the drug

was not necessarily the only component that affected the heterogeneity and the alteration rate. This rate varies widely across mammalian genomes [150], and this variation has important consequences for our understanding of the global evolutionary process, and in particular concerning cancer. We have sought to investigate such cases by examining heterogeneity as it relates to resistance level and time, where the alteration rate varies over time with or without chemotherapeutic treatment and external stress. We assume that the alteration rate could be affected by either the treatment or the external stress, but the induced-death rate varies only based on the treatment. Certainly, there is a possibility of having a mutation in the apoptosis pathway and therefore affecting the induced-death rate, but there is no clear way of expressing this rate as a function of the trait.

Our results suggest that treatment acts as a selection process in the effective drug concentrations range, while genetic/epigenetic changes act as a diffusion process based on different stress signals. In a case with no changes and no treatment, the dynamic is determined solely by the cell density, cell division and natural death rates as functions of the resistance level. We further examined the effects of two types of mathematical modifications. In this model, the resulting dynamic and response are expressed by only three main elements: density (ρ), heterogeneity ($n(x, t)$), and alteration rate (ϵ).

Thus, the design of a new treatment protocol should also incorporate the individual status regarding these variables that should be estimated at the time of diagnosis. The common approach for selecting a protocol is to find and target genes that are significantly expressed in most cells in that tumor. Successful treatment

is determined by the substantial reduction in the tumor size. The next obvious questions would be: What will happen with the remaining cells? Will these cells evolve into a more aggressive and resistant tumor or not? If these cells are low in number, how can we identify them? The model predicts that it is recommended to develop/bioengineer a treatment that targets the alteration rate of all cells in order to reduce this rate as much as possible before any other treatment, so the heterogeneity will not increase over time. In any event, it is predicted that it is better to first treat the small subset of cells with high alteration rate, than the bulk of the tumor using a specifically targeted therapy but with low alteration rate. In the case where this rate is very high and cannot be affected by external intervention, during treatment it is predicted that the increase in the alteration rate increases the spread to other traits that are not necessarily better in their survival; therefore, the density decreases and the entire system results in a better response. Increasing the alteration rate *per se* is an extreme approach; therefore, there is a need to bioengineer such a modification in a temporary and reversible manner. Another possibility is to increase the percentage of cells that could be exposed to alterations. By the end of the treatment, there is a need to reduce the alteration rate as much as possible.

Our findings imply that at certain levels of resistance, where the division rate is higher than the death rate (including the contribution of the alteration rate), this deterministic approach can be applied. It is possible that when the difference between those rates is low and there is additional stochastic noise, the stochastic results will differ from the deterministic prediction. Also, the study conducted in

this chapter focused on the role of cell density and intratumoral heterogeneity in drug resistance. If several drugs are administered, a full study will be required to consider the multidimensional analog, in which the resistance level is a vector in a space whose dimension is the number of drugs being administered. A complete study of this nature is beyond the scope of this chapter, as it will require detailed information about the cross-reactivity between the various drugs. An alternative approach would be to view the level of drug resistance as the magnitude of such a vector of resistance. In this case, changing the drug can be modeled by shifting the values of the distributions with respect to the trait. A full study addressing these two issues is left for future work.

4.2 Simplifying the complexity of resistance heterogeneity in metastasis

The content of this section has been published in [137]. The main goal of treatment regimens for metastasis is to control growth rates, not eradicate all cancer cells. Mathematical models offer methodologies that incorporate high-throughput data with dynamic effects on net growth. The ideal approach would simplify, but not over-simplify, a complex problem into meaningful and manageable estimators that predict a patient's response to specific treatments. Here, we explore three fundamental approaches with different assumptions concerning resistance mechanisms, in which the cells are categorized into either discrete compartments or described by a continuous range of resistance levels. We argue in favor of modeling

resistance as a continuum and demonstrate how integrating cellular growth rates, density-dependent versus exponential growth, and intratumoral heterogeneity improves predictions concerning the resistance heterogeneity of metastases.

4.2.1 Heterogeneity in primary tumors and in metastasis

Within an individual tumor, and between a primary tumor and its metastases, there are certain genotypic or phenotypic variations. Intratumoral heterogeneity implies that different parts of a tumor may have different properties, including the existence of different degrees of sensitivity to various cancer drugs. This heterogeneity has significant implications when developing clinical protocols (see [33, 34] and the references within). Heterogeneity can lead to variations in different fundamental cellular behaviors as a function of time and stress, regardless of the specific mechanisms that may induce it [131, 132]. Changes in the rates of cell division, death, mutation, migration, etc., have direct effects on the dynamic of growth and initiation of metastatic cells. For example, mutations in the Ras-Raf pathway, a type of Mitogen-Activated Protein Kinases (MAPK) pathway, may cause increased cell proliferation and resistance to apoptosis [141]. In hepatocellular carcinoma LIM and Cysteine-rich Domains-1 (LMCD1) mutations promote cell migration and tumor metastasis [151]. There are also certain genes that promote genetic stability, including DNA repair genes, DNA damage sensor genes, and cell-cycle checkpoint genes. Changes in these stability genes affect the mutation rate [142]. Since these rates are not constant over time, and may vary according to the environment, re-

sistance to chemotherapy becomes a complex dynamic process. Hence, an effective treatment often requires a combination of drugs targeting different resistance mechanisms, based on specific individual genotypic and phenotypic variations.

Along with biological and clinical research, mathematical approaches have been developed to model the development of drug resistance, and have dealt with many of the known aspects of the field [66]. These computational models and their outcomes each have their own definitions of “resistance” for a given treatment, presented in discrete (e.g., two sub-groups of “sensitive” and “resistant”) or continuous (e.g., a range of values between 0 and 1, where 1 represents the maximum resistance level) ways. In addition to the intrinsic-resistance level, there is another factor, the evolution of resistance, which contributes to the dynamics. This factor determines the ability of a cell with a given resistance level to evolve over time and/or space to a different level, based on the processes that were initially assumed. For example, one of these processes is the development of spontaneous or drug-induced mutations.

Here, we discuss the impact of different assumptions concerning resistance level, resistance evolution, and growth limitation on the prediction of the dynamics of cancer growth, heterogeneity, and survival. These concepts are described here through the three mathematical models of Stein *et al.* [1], Lorz *et al.* [121], and Lavi *et al.* [136]. In brief, the Stein model assumes exponential growth for all cells, and includes only the mechanism of intrinsic resistance, with no evolutionary process. In addition, they used a discrete definition of the resistance level that categorized the metastatic cells into two main groups: resistant vs. sensitive. While the Lorz model also assumes exponential growth for all metastatic cells, they in contrast

include resistance evolution with a continuous description of the resistance level. The Lavi model is an extension of the Lorz model, which shares most of the Lorz assumptions, except that we replace the exponential growth with a growth that depends on the cell density. Furthermore, in our model we account for the occurrence of changes (e.g., epi-mutations) that occur at higher rates than those of genetic mutations. Our results show that these realistic assumptions have crucial effects on the predicted level of resistance heterogeneity, and as such integrate differently patient's information into more accurate predicted clinical outcomes.

4.2.2 Resistance level as a discrete variable

Although oncologists are aware of the impact of intratumoral heterogeneity, many clinical trials test only one drug. For a given treatment, the clinical response and the treatment efficacy are determined by basic measurements, such as the quantification of tumor shrinkage and overall survival (OS), that are not designed to reveal the underlying complexity of the response based on varying resistance levels in different cells in a tumor. Interesting metastatic studies with clinical and practical implications have been reported by Stein and colleagues [1, 152, 153, 153, 154]. These studies offer a method of analyzing tumor response and predict survival using surrogate and direct measures of tumor volume while a metastasis patient is receiving therapy in a clinical trial. Their main hypothesis is that for a given treatment, there are two main cell sub-populations, sensitive and resistant. That is, the resistance level can be described by a discrete variable measuring cellular death/survival. Accordingly,

initial tumor shrinkage during treatment is due to a higher portion of sensitive cells in a tumor, and a subsequent increase in the tumor size is a result of the remaining resistant cells that increase in number. The published data of [1, 152, 153, 153, 154] on patients enrolled in clinical trials receiving chemotherapy for mainly prostate cancer and renal cell carcinomas can be divided into four main outcomes: successful treatment, partial response with relapse, no response, or unclear. Stein and colleagues fitted a function given as a sum of two exponentials to estimate tumor size over time ($\rho(t)$). The tumor size was based on the rates of tumor regression (d) and tumor growth (g):

$$\rho(t) = e^{-dt} + e^{gt} + 1. \quad (4.2)$$

ρ was normalized with regards to the initial tumor size on day 0, i.e., $\rho(0) = 1$. They proposed that because the growth rate was found to correlate with OS, it would serve to estimate efficacies of different therapies based on the sampled population or the individual patients. Using these assumptions and method, they concluded that bevacizumab reduces the growth rate constants of renal carcinomas [153], and that sunitinib reduces the tumor growth rate more than interferon type I $_{\alpha}$ (IFN- α) [155]. They argued that during treatment, the rate of tumor re-growth (g) is a constant value, and for certain patients, this indicates that treatment should be continued for a longer period of time since the tumor growth rate is lower than without treatment.

The Stein model describes the net growth of the metastatic cells with no information on cellular rates such as cell division or death rates. Sensitive cells

are assumed to be cells that did not survive the treatment, regardless of the value of their drug-induced death rate. Note that a treatment producing low rates of drug-induced death and cell division could have the same net-growth response as a different treatment with higher rates of drug-induced death and cell division. Therefore, with no further information about cellular growth and death rates, the efficacy of a treatment is only expressed by the net-growth response.

We consider here the approach of Stein *et al.* with a representative example of partial response (published as part of supplementary Figure S1, case 12 in [1]). Their data set serves us throughout this paper as a way to compare three mathematical models, and discuss their assumptions, approaches and the meaning of their results. We estimated the parameters of (4.2) to be $d = 0.0541$ and $g = 0.0005$, using a nonlinear least-squares method (for more details, see Section B.2 in Appendix B). The data points and the curve fitted using (4.2) are plotted in Fig. 4.6A.

4.2.3 Resistance level as a continuous variable

The significant variation that has been observed in intratumoral heterogeneity indicates there is a variation in cell division and death rates of cells within a tumor; thus, drug resistance is far from being a black-and-white, resistant or not resistant phenomenon. In view of that, our hypothesis is that a continuous variable is a more appropriate way to describe, estimate, and measure the resistance level. Several direct and indirect interdisciplinary approaches have been suggested to estimate the drug resistance level, depending on the type of data that is analyzed. For instance,

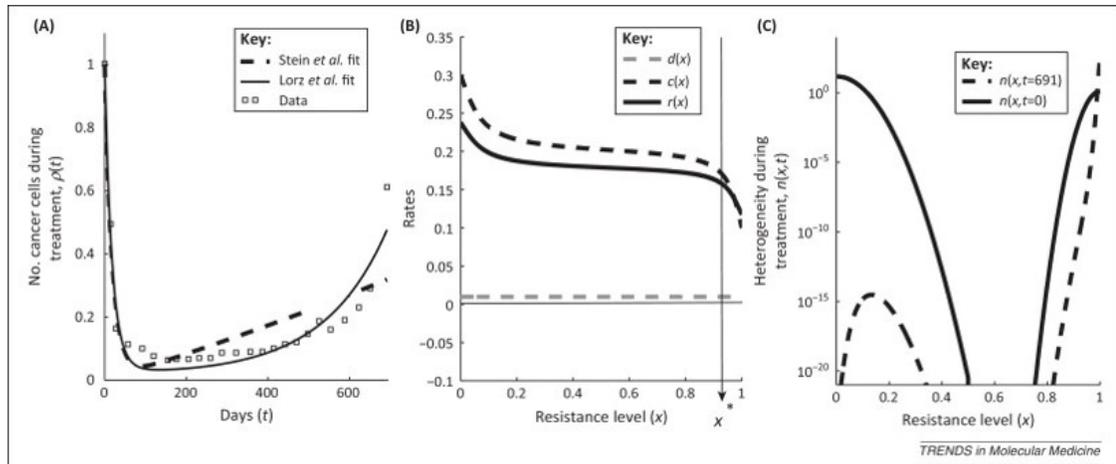


Figure 4.6: Comparison between the Stein and Lorz models, with resistance-level variables that are discrete vs. continuous. (A) Patient response data and two mathematical predictions of the total number of cancer cells during treatment are plotted. The data set (boxes) illustrates a representative relapse response (case 12, Fig. S1 in [1]). The Stein model, given in (4.2), describes the discrete approach of grouping the cells into two categories: resistant and sensitive (dashed line). The Lorz model, given in (4.3), describes the continuous approach to resistance level (solid line). (B) The cellular rates of cell division ($r(x)$, black solid line) and drug-induced death ($c(x)$, black dashed line) are assumed to be sigmoid curves and the natural death rate is assumed to be a low constant value. (C) The trait distributions, $n(x,t)$, at the beginning ($t=0$) and at the end ($t=691$ days) of treatment. This plot is the solution of the exponential growth model described in equation (4.3), where $\theta=0$ and $\epsilon=0$, using initial condition IC_1 (see Figure 4.7(B)) and cellular rates shown in (B).

clinical data that combine clinical staging, gene expression, and survival data from the same patient can help to categorize a patient as a “good” or a “poor” responder in a continuous manner. Such a score can then be used as a more comprehensive method to estimate the different cancer cell phenotypes and their drug sensitivities.

Exponential growth. From a mathematical perspective, Lorz *et al.* [121] proposed a model for the evolution of cancer cells, ρ , over time t

$$\begin{aligned} \frac{\partial n(x, t)}{\partial t} &= \left(r(x)(1 - \theta(x)) - c(x) - d(x) \right) n(x, t) \\ &+ \int_0^1 \theta(y, t) r(y) M(y, x) n(y, t) dy. \end{aligned} \quad (4.3)$$

We note that this is the same equation studied in Chapter 3 with no density dependence, i.e. equation (3.6) with $G \equiv 1$. Their model is based on an approach previously developed in part by several other groups [122–124, 126–128, 130]. It was assumed that the cancer cells had no growth limitations, and grow exponentially. The model was designed to estimate the heterogeneity, n , as a function of the resistance level, x , and time, t . Using this notation, the total number of the cancer cells is described by

$$\rho(t) = \int_0^1 n(x, t) dx. \quad (4.4)$$

Each subpopulation, $n(x, t)$, is based on the rate of cell division, $r(x)$, the spontaneous rate of death, $d(x)$, the rate of drug-induced death, $c(x)$, and the initial conditions, $IC = n(x, 0)$. All rates are functions of the resistance level. In addition, there are contributions from all other subpopulations, depending only on a genetic

mutation function, $M(y, x)$. $\theta(x)$ denotes the fraction of cells with trait x that can carry out new modifications, where $0 \leq \theta(x) \leq 1$. This suggests that the treatment acts as a selection process, while mutations act as a diffusion process.

Once resistance level is included in tumor-growth models, we can study how variation in cell division, death, and alteration rates, as a function of the resistance level, can affect the predicted patient’s response. How is such a model different than having a discrete representation of resistance? How can we quantify the dynamic changes in heterogeneity? To address these questions, we analyzed the data points of Stein *et al.* (squares in Figure 4.6A), followed by a simplified version of the Lorz model where we assume exponential growth with no mutations ($\theta = \epsilon = 0$). We relaxed the assumption that cellular rates are described by step functions, and considered sigmoidal functions instead. The cell division and death rates were estimated using a nonlinear least-squares method (Appendix B.2 and Figure 4.6B), with an initial condition of 96% of the cancer cell population located below the value of $x = 0.5$ (IC₁ in Figure 4.7B). Heuristically, we imagine $x = 0.5$ as the barrier between sensitive and resistant cells. Using the model of Lorz *at el.*, the predicted growth (Figure 4.6A) gives information about changes in heterogeneity over time ($n(x, t)$). Figure 4.6C also gives a more detailed description of the “relapse” case. When starting with approximately 96% “sensitive” cells (described by IC₁), theoretically, after 691 days of treatment, the “sensitive” cells have not totally been eliminated by the drug (dashed curve) and could affect the net-growth over time, after treatment ceases.

We next demonstrate how the net growth (Figure 4.7A) and the heterogeneity

(Figure 4.7C) of four different non-zero initial conditions (IC₁-IC₄, Figure 4.7B), would appear at the end of a treatment (on day 691), using the same rates as in Figure 4.6. It is interesting that the growths of both IC₁ and IC₂ are not significantly different, while their heterogeneous compositions include different levels of resistance, even if the number of cells of certain types are very small. IC₃ and IC₄ resulted in expected curves. IC₄ assumes that initially the majority of cells have higher cell division rates than death rates. Therefore, the tumors grow more than with all the other ICs. IC₃ also resulted in an intuitive curve representing successful treatment. Although the average “resistance level” seems to be high ($x = 0.5$) compared to all possible values of this variable ($x \in [0, 1]$), the important aspect is the relative size of the net death rate (i.e., $c(x) + d(x)$), the cell division rate ($r(x)$) over the x -space, and in particular the trait values where the net-growth rate changes its sign, from negative to positive. A negative net growth would result in decreasing tumor size. In this case, almost all the initial cells lie in this region.

Variation in cellular rates. One should keep in mind that the lower the number of sensitive cells, the smaller their contribution within a short time period. Thus, in the short time period after a treatment is stopped ($t > 691$ days), the results derived from the Lorz model do not necessarily contradict the qualitative conclusions of the discrete model. On the other hand, both models are based on a deterministic description, with fixed parameters, and no stochasticity. A noisy environment could result in a more complex resistance distribution, mainly where the cell division and death rates would be closer in value. In this case, a clear separation between the two sub-populations would no longer be seen and the conclusions of the two models

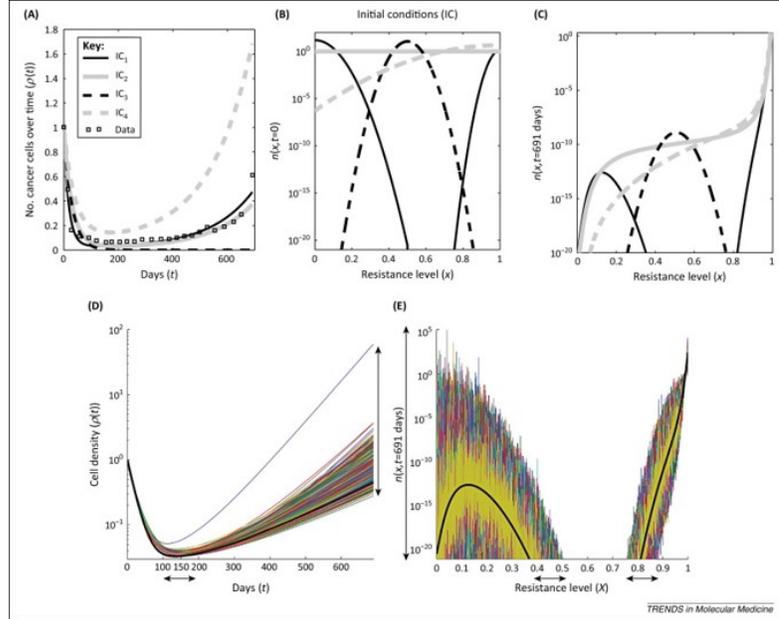


Figure 4.7: The Lorz model of exponential growth with four different initial trait distributions. (A) Four mathematical predictions of the total number of cancer cells over time, $\rho(t)$ where $t \leq 691$ days), based on four different initial conditions (IC₁-IC₄). These curves are the solutions of (4.3), where $\theta = 0$, $\epsilon = 0$ along with the rates in Fig. 4.6B. (B) The four different initial conditions, labeled as IC₁-IC₄. (C) The trait distributions at the end of the treatment, $n(x, t = 691$ days), using the same conditions as in panel A. (D-E) These two panels describe the effect of variations in the net-growth rates on the Lorz model. Using the same conditions as in Figure 4.6 (continuous-resistance model), we illustrate the outcome of the perturbed net-growth rates of equation (4.3) using 1000 simulations. The additional perturbation distribution assumes a normal distribution with mean 0, standard deviation 25% of the net-growth rate in Figure 4.6A applied uniformly over all traits. (D) shows the variation in the growth of cancer cells over time, while (E) shows the variation in the trait profile at the final time (691 days). The different colors are the result of different perturbations corresponding to the aforementioned normal distribution, while the black line corresponds to the simulation in Figure 4.6 (that is, no perturbation). The arrows stress the size of variations.

would be totally different, even for a short period of time.

Variations in cell division and death rates are very common, especially when administering a drug. Using the Lorz model (equation (4.3), Figure 4.6A, solid line), we show in Figures 4.7D-E the outcome in terms of cancer growth over time and in its heterogeneity on the last day of treatment ($t = 691$ days), when fixed perturbations of the net-growth rate as a function of the trait were included as initial conditions. The net-growth rates are sampled from a normal distribution where the mean is the initial “deterministic-growth rate” (which can be estimated from Figure 4.6B), and the standard deviation is 25% of that mean. In that specific example, the maximum growth rate among all traits is 0.0721 (absolute value), and the corresponding maximum standard deviation is 0.018. Given the exponential growth, small perturbations of those rates can result in large differences in cell density (3 – 4 orders of magnitude, see vertical arrow in Figure 4.7D), and in the time to reach the minimum-tumor reduction (a delay of up to 50 days, horizontal arrow in Figure 4.7D). In addition, the heterogeneity at the end of the treatment dramatically varied from the original estimation in the expected traits (x -axis) and in the number of cells with those traits (y -axis) (colored lines versus black line, Figure 4.7E).

Growth with cell-density dependence. Both models, Stein *et al.* [1] and Lorz *et al.* [121] described the total number of metastatic cells, and made an assumption concerning the exponential growth of cancer. However, the total number of cancer cells at any given time can be thought of as an integral of all individual tumors, with different sizes and cell densities: $\rho_{\text{total cancer cells}}(t) = \int_{\text{all tumors}} \rho_{\text{tumor}}(t)$.

These two aforementioned models are thus describing the left-hand side of this equation (i.e., $\rho_{\text{total cancer cells}}(t)$). We can however, consider the more fundamental problem of modeling each tumor individually, and describe the right-hand side of this equation (i.e. $\int_{\text{all tumors}} \rho_{\text{tumor}}(t)$). Each tumor has its own growth rate, which depends on its heterogeneity and environment. Also, at some point in time, cells may migrate and initiate new tumors with a certain probability, depending on the type of tumor, the mutations, and the tumor size. Mathematically this can be modeled, assuming a net-growth rate that depends on the intratumoral cell density in addition to the resistance level. The mathematical formula that best describes the “growth law” of cancer cells has long been debated, though there are three commonly used deterministic functions: exponential, logistic, and Gompertzian [156–161]. In most computational studies, the exponential growth describe only the early stages of tumors or the metastasis. But as the tumor size increases, a slower rate of growth provides a better description. Logistic and Gompertzian laws have been mainly considered for this slower type of growth.

Focusing from a global to a local-tumor level, the rates of cell division and spontaneous death are known to depend on numerous biological parameters, including the local cell density and the resistance level. Cell density relates to various factors that determine tumor growth, such as cell-cell interactions, cell-matrix interactions, the distribution of nutrients, survival signals, the penetration of anticancer drugs, and the pressure within a tumor. In general, many studies have measured the relationship between the cellular rates as functions of cell density in different environmental conditions [33,34,121,131,132,136,141,151,152]. In some studies, the cell

division rate was found to be a decreasing function of cell density [121, 136], while the death rate was described by an increasing function of cell density [33, 34, 131].

We have extended the approach of Lorz *et al.* using the same framework to focus on the single tumor level [95, 136] (see Chapter 3 and Section 4.1 above):

$$\begin{aligned} \frac{\partial n(x, t)}{\partial t} = & \left(f(\rho(t))(r(x)(1 - \theta(x, t)) - c(x)) - g(\rho(t))d(x) \right) n(x, t) \\ & + f(\rho(t)) \int_0^1 \theta(y, t)r(y)M(y, x)n(y, t) dy. \end{aligned} \quad (4.5)$$

The cell division and the spontaneous death rates of the cancer cells we assume to depend on the cell density (functions $f(\rho), g(\rho)$ above). Using the same rates of cell division and death as previously, we examine the heterogeneity and density over time with the administration of a treatment ($t \leq 691$ days), for the four initial conditions (IC₁-IC₄). In all of these cases, the net-growth rates decrease to different degrees and the densities are bounded (Figure 4.8A). The model is flexible enough to include more information on factors that could limit the tumor size/net-growth rate by changing the functions of $f(\rho)$ and $g(\rho)$. Accordingly, the $n(x, t)$ for each IC is different than for exponential growth, even though the same cell division and death rates were assumed (Figure 4.8B versus Figure 4.7C). For instance, after $t = 691$ days with treatment, cells with $x < 0.1$ are still present in this example.

Furthermore, in our model we account for the occurrence of genetic as well as non-genetic changes, which occur at higher rates than the genetic mutations. A small mutation rate was already shown to function as a diffusion process for x , since cells with a given trait may spread to a neighboring interval of x -values via

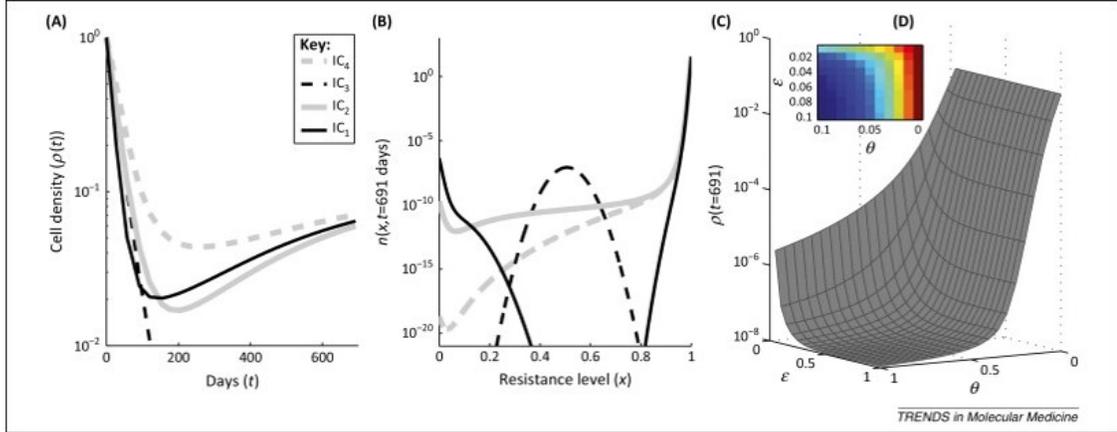


Figure 4.8: Modeling cell growth with density dependence. As in Figure 4.7, the effects of four initial conditions on the density and final trait distribution of the third model are shown. (A) Four mathematical predictions of the total number of cancer cells over time, based on IC_1 - IC_4 . These curves are the solutions of equation (4.5), with $\theta = 0$ and $\epsilon = 0$ and rates as in Figure 4.6B. (B) The trait distributions at the end of the treatment ($t = 691$), using the same conditions as in panel A. (C) Lower plot: numerical predictions of the final cell density, $\rho(t = 691)$, using (4.5), where ϵ and θ vary over a range of parameters $[0, 1]$. All other parameters and functions are the same as in panels A and B. (D) Upper plot: as in panel C, ϵ and θ vary over a limited range of parameters, now in $[0, 0.1]$.

mutations [121]. While the effect of such a low rate has a very limited impact on the total cell density, higher rates are expected to increase the spread of x with different cell division and death rates. Therefore, these cells could die or survive depending on their altered traits. We demonstrate the effect of the alteration rate on the tumor density over different scales (Figures 4.8C-D). There are two parameters that relate to the alteration process, ϵ and θ . The alteration kernel, $M(y, x)$, is considered here to be a Gaussian distribution conned to $[0, 1]$ with mean y and variance $\epsilon^2/2$ (see equation (3.30)). In general, the variance, ϵ , is a function of the external stress and time, $\epsilon(\text{stress}, t)$ (see [136] and Section 4.1). Here, we present a simple case in which these parameters are constant over time. In Figure 4.8C, we plot the cell densities at the end of the treatment ($\rho(t = 691\text{days})$), while varying ϵ and θ in $[0, 1]$. The

upper plot in Figure 4.8D focuses on lower values of ϵ and θ in $[0, 0.1]$. Both plots clearly demonstrate a decrease in cell density as the alteration rate increases.

4.2.4 Concluding remarks and future perspectives

With the advantages of high-throughput data and new automated techniques comes the obvious problem of how to analyze and understand the high-dimensional complexity of a patient's information. Researchers can create complex datasets, but a more challenging task is to develop models that are able to simplify and explain complex data. The ideal mathematical approach would simplify, but not oversimplify, complex datasets into meaningful and manageable estimators that predict a patient's response to specific treatments. Furthermore, there is an urgent need to develop effective methodologies to analyze data from complex diseases such as metastatic cancer, in which tumors are quite heterogeneous, containing cells with different degrees of drug sensitivity, the source of cancer relapses.

In this chapter, we have addressed some of the main aspects relating to the mathematical modeling of drug resistance in metastatic cancer. First, we demonstrated the value of having a model based on a continuous representation of the resistance level. Many primary human tumors have been found to contain genetically distinct cellular sub-populations [33]. Even in a system with a single genetic clone, functional variability has been observed among the tumor cells, with distinct cellular sub-populations and different growth rates [131]. Thus, the definition of "resistance" should incorporate all possible cases in which tumor cells survive. In

addition to the fact that cells could survive the administered treatment, their survival rate should be included in the model. From that perspective, a tumor contains several sub-populations of cancer cells with different net-growth rates. The net-growth (or survival) rate is a combination of cell division, spontaneous death, and drug-induced death rates. In most, if not all, cases there is an additional rate of alteration (e.g., modification/mutation). The better the treatment, the higher the selection, which minimizes heterogeneity due to convergence of the sub-populations with the highest survival rates. However, as was demonstrated here, there may be non-detectable sub-populations that manage to grow at very low survival rates. After a long period of treatment, the global rate of cancer net growth is mainly influenced by the largest sub-populations with the highest survival rates. This is in accord with the discrete-resistance model. However, when a second drug is administered to target the selected “resistant” sub-populations, the undetected cells with different resistance levels appear. Having a method that can predict the underlying dynamic of the heterogeneity might improve the design of treatment protocols for patients that relapse.

The second topic we addressed is the modeling of metastatic cancer cells using a deterministic approach. Metastases are known to spread rapidly, and therefore exponential growth is the commonly used approach in modeling. We demonstrate how small, yet likely, variations in the net-growth rates may result in a large variation in the predicted tumor size and its heterogeneity over time. Thus, it is difficult to target using specific treatments. This prediction is the outcome of assuming exponential growth.

Our view of metastatic cancer is to some extent different. We suggest modeling metastatic cancer cells as a collection of individual tumors with different growth rates, in which a single cell can be considered as a tumor of minimal size. When tumor size increases, a slower growth rate is observed [162]. We show here how cell-density dependence can change the dynamic of the heterogeneity, mainly by scaling the time to reach the heterogeneity of the exponential model. Thus, the heterogeneity of the exponential model is not necessarily identical to the average of all tumors with density dependence, which highlights the advantage of using such a model with density dependence. A full model that includes the probability of initiating new tumors and determines tumor size based on the environment and intratumoral alterations will be the subject of future work.

We would like to emphasize the importance of collecting accurate data for estimating tumor size and its heterogeneity during treatment. This description of tumor heterogeneity should include information on the sub-populations within a tumor, which should give us a way to quantify rates of cell division, death, alteration, migration, etc. Having such data would dramatically accelerate the development of computational models, and increase their ability to assist in the design of optimal treatment protocols.

Chapter 5: Modeling intrinsic heterogeneity and growth of cancer cells

The content of this chapter has been published in [163]. Intratumoral heterogeneity has been found to be a major cause of drug resistance. Cell-to-cell variation increases as a result of cancer-related alterations, which are acquired by stochastic events and further induced by environmental signals. However, most cellular mechanisms include natural fluctuations that are closely regulated, and thus lead to asynchronization of the cells, which causes intrinsic heterogeneity in a given population. Here, we derive two novel mathematical models, a stochastic agent-based model and an integro-differential equation model, each of which describes the growth of cancer cells as a dynamic transition between proliferative and quiescent states. These models are designed to predict variations in growth as a function of the intrinsic heterogeneity emerging from the durations of the cell-cycle and apoptosis, and also include cellular density dependencies. By examining the role all parameters play in the evolution of intrinsic tumor heterogeneity, and the sensitivity of the population growth to parameter values, we show that the cell-cycle length has the most significant effect on the growth dynamics. In addition, we demonstrate that the agent-based model can be approximated well by the more computation-

ally efficient integro-differential equations when the number of cells is large. This essential step in cancer growth modeling will allow us to revisit the mechanisms of multi-drug resistance by examining spatiotemporal differences of cell growth while administering a drug among the different sub-populations in a single tumor, as well as the evolution of those mechanisms as a function of the resistance level.

5.1 Introduction

As discussed in Section 1.2, tumor heterogeneity results both from perturbations in cellular mechanisms (induced heterogeneity), as well as from typical biological processes that are stochastic yet tightly regulated (intrinsic heterogeneity). The former concerns genetic and epigenetic events, while the latter are common to healthy and cancerous cells alike. In this work, we model a specific form of intrinsic heterogeneity, the cell-cycle, and study its effect on tumor growth dynamics.

Two frameworks that are commonly used to design mathematical cancer models which predict cellular behavior are individual-based models and continuous deterministic models. Several different individual-based models of tumor growth have been developed recently (see review [164]). Among them are agent-based models (ABMs). The ABM framework is a powerful simulation method that has seen a variety of applications, including bio-medical research [165–167] and socio-economic modeling [168]. ABMs describe dynamic systems as collections of autonomous decision-making individuals called agents. Each agent assesses its state and makes decisions on the basis of a set of rules. Agents may execute various behaviors appro-

appropriate for the system they represent. ABMs are generally more flexible than deterministic models and may take into account virtually any biological phenomenon. Here, we present two mathematical approaches, the ABM and a corresponding integro-differential (IDE) model, to predict the growth of a single ovarian cell line, OVCAR-8, where the cells can be proliferating, dying, or in quiescence. The novelty of our methods lie in the description of cellular decision-making as a function of the global dynamic cell density, with intrinsic variations of the cell-cycle and death process lengths. Decisions concerning actions are based on how the cell senses its environment, in a probabilistic fashion. We study the robustness of cell growth despite noise in division and natural death rates. The entire system dynamic results from the decisions of individual entities that can cause transient or permanent heterogeneity, generate network effects, and potentially lead to significant deviations from stochastic to deterministic predictions. We demonstrate the existence of fluctuations in cell growth using data of proliferation rates as a function of cellular density. This fundamental framework of cellular growth dynamics is a necessary first step that will allow us to work on more complex co-cultured systems based on geometry, which includes a spatial mechanism of drug resistance that could shed light on the spatiotemporal evolution of intratumoral heterogeneity.

5.2 Agent-based model

The first model we introduce is an ABM, where each cell is distinguished with its own state and behavior. This framework permits a simple way to introduce an age

structure into the model, which is a main focus of this work.

5.2.1 Model construction

The ABM consists of three compartments of cells: proliferative (P), apoptotic (A), and quiescent (Q). See Figure 5.1 for an outline of the transitions between compartments. Q consists of cells that are neither dividing nor dying, and acts mainly as a reservoir for the other two compartments. P consists of cells that are currently in any stage of the cell-cycle. When a cell makes a transition from Q into P, a cell-cycle length, L_P , is chosen. L_P was assumed to be a random variable with normal distribution:

$$L_P \sim \mathcal{N}(\mu, \sigma^2), \quad (5.1)$$

where μ is the mean length of the cell-cycle, and σ is the standard deviation [36]. The value of μ is taken as the doubling time of OVCAR-8 cells, which is estimated to be 24.4416 hours (see Section 5.4.1). Once a cell is in P, it progresses through the cell-cycle for L_P hours, unless a transition occurs to the apoptotic compartment A. Both mother and daughter cells subsequently leave the division stage and become quiescent (Q). The last compartment, A, consists of cells currently in the apoptotic process. Cells in A remain for a random length of time L_A , after which the cell transitions out of the apoptotic compartment to complete cell death. Based on the experimental data of Messam and Pittman [38], we approximate L_A as a gamma-

distributed random variable:

$$L_A \sim \Gamma(\omega, \lambda), \quad (5.2)$$

where ω and λ denote the shape and rate parameters, respectively. See Section 5.4.1 for details.

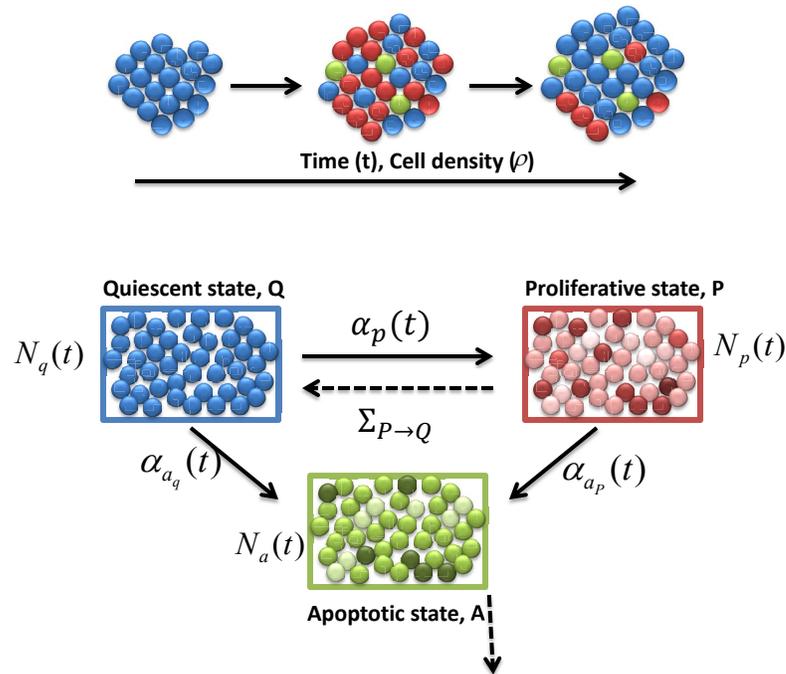


Figure 5.1: Model dynamics. Diagram of transitions between the three cellular compartments in the ABM. Q denotes the quiescent compartment with $N_q(t)$ cells at time t , P denotes the proliferation compartment with $N_p(t)$ cells at time t , and A denotes the apoptosis compartment with $N_a(t)$ cells at time t . Note that $N_q(t)$, $N_p(t)$, and $N_a(t)$ are all stochastic processes. The explicit transition rates between the compartments are shown in solid lines, and are labeled as $\alpha_p(t)$, $\alpha_{a_p}(t)$, and $\alpha_{a_q}(t)$. The implicit transition rates, due to the completion of cellular cycles, are shown in dotted lines, and have no closed-form expression. For example, $\Sigma_{P \rightarrow Q}$ corresponds to the rate of cell-cycle completion. The line originating from compartment A indicates cells that are removed from the simulation.

Lastly, we assumed that transitions between the three compartments are gov-

erved both by the global cellular density, labeled ρ , and the random amount of time spent in P or A (L_P, L_A , respectively). We formulated these transitions in terms of rate parameters. Consider, for example, the transition from Q to P, that is, the probability of a cell entering the cell-cycle from quiescence. We model this using a rate, $\alpha_p(t)$, with the interpretation that for a small amount of time Δt , $\alpha_p(t)\Delta t$ is essentially the probability of one cell making a transition from Q into P at some point in the time interval $[t, t + \Delta t]$. More precisely, we are assuming a first-order dependence in the time step for transition probabilities, and thus the process can be thought of as Poisson, with non-homogeneous intensity $\alpha_p(t)$. Of course, all of this holds only in the limit as $\Delta t \rightarrow 0^+$, as theoretically this is a continuous time Markov chain. In practice however, we simulate using small discrete time steps Δt , and take $\alpha_p(t)\Delta t$ as the exact transition probability per cell. All other explicit transition rates (dark lines in Figure 5.1) have this same interpretation.

The transition rates are functions of β and d (see C.1). One of our fundamental assumptions is that the measurements of β and d did not occur at equilibrium, since the two division fraction data sets do not agree in value (see Figure 5.2(a)). However, the two curves do agree qualitatively in their general trend, as both contain relative maxima $\beta_m \in [0.3, 0.8]$ occurring at some density $\rho_m \in (0, 1)$. Using this observation,

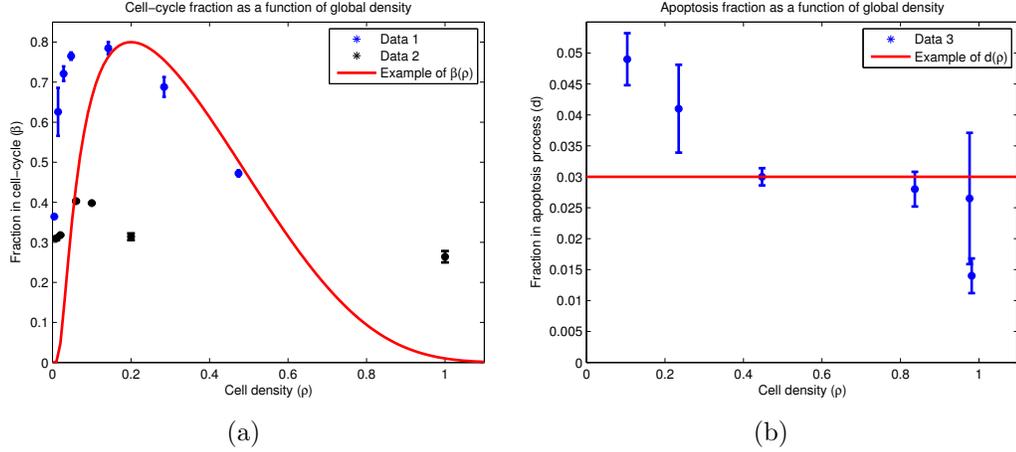


Figure 5.2: Experimental results with functional forms for cell-cycle and apoptosis fractions. Experimentally (blue and black), OVCAR-8 cells are allowed to attach to the plate for twenty-four hours, and then are incubated for a subsequent twenty-four hours. After this incubation, the cells were stained to determine which percentage of cells reside in the proliferation and apoptosis stages. The experiment was performed using two different techniques for cell synchronization: Data 1 and Data 2 (see C.1.1.1 for details). The functional forms (red) correspond to equations (5.4) and (5.3), with free parameters taking values $\beta_m = 0.75$, $\rho_m = 0.15$, $\epsilon = 1$, and $d = 0.03$. (a) Fraction of cells in division stage (P) as a function of the population density on the plate; (b) Fraction of cells in apoptosis stage (A) as a function of the population density on the plate. Note that we allow $\rho > 1$.

we postulated equilibrium distributions $\beta(\rho)$ and $d(\rho)$ with the same basic structure:

$$d(\rho) \equiv d, \quad \text{a constant}, \quad (5.3)$$

$$\beta(\rho) = \begin{cases} \beta_m e^{\frac{-a(\rho-\rho_m)^2}{\rho(1+\epsilon-\rho)^2}} & \text{if } 0 < \rho < 1 + \epsilon, \\ 0 & \text{otherwise,} \end{cases} \quad (5.4)$$

where

$$a = \frac{\epsilon \log(\beta_m/d)}{(1 - \rho_m)^2}. \quad (5.5)$$

There are a total of 4 free parameters in (5.3)-(5.4): $\beta_m, \rho_m, \epsilon$, and d . See Table C.1 in Appendix C for calculated values and explanations, and Figures 5.2(a) and 5.2(b) (red curves) for a sample plot of $\beta(\rho)$ and $d(\rho)$ at an arbitrary parameter set.

$d(\rho)$ is chosen as a constant d since its observed range of values is small ($0.01 \leq d \leq 0.05$), and relative to β , appears essentially constant (see Figure 5.2(b)). However, we do use these values as the lower and upper bound on parameter searches (see Section 5.4.4). One can also check that $\beta(\rho)$ in (5.4) has absolute/relative maximum β_m at $\rho = \rho_m$, which is the aforementioned desired constraint. ϵ is a constant which dictates how rapidly $\beta(\rho)$ decreases for large cellular densities. Condition (5.5) guarantees that $\beta(1) = d$, and furthermore, $\beta(\rho) < d$ for $\rho > 1$. Lastly, $\beta(\rho) = 0$ for $\rho > 1 + \epsilon$. The reason for these choices is as follows: we allow the possibility that $\rho > 1$, since it was observed that OVCAR-8 cells may deform their cell membranes and/or grow upon one another in a two-dimensional culture to complete mitosis. Hence, we allow divisions when $\rho > 1$, but we ensure that death is more likely in this regime. Thus, when $\rho > 1$, a net increase in cells should only occur from cells that previously entered compartment P and successfully completed cell division; no net flow between compartments P and A exists. Furthermore, when the plate becomes dense enough (i.e. $\rho > 1 + \epsilon$), no cells can enter P.

The rates that describe the transitions between the cellular compartments are

given below:

$$\alpha_p(t) = c \frac{(\beta(\rho(t))N(t) - N_p(t))_+}{N_q(t)}, \quad (5.6)$$

$$\alpha_{a_p}(t) = c\gamma \frac{(dN(t) - N_a(t))_+}{N_p(t)}, \quad (5.7)$$

$$\alpha_{a_q}(t) = c(1 - \gamma) \frac{(dN(t) - N_a(t))_+}{N_q(t)}, \quad (5.8)$$

where

$$N(t) = N_q(t) + N_p(t) + N_a(t), \quad \text{and} \quad (5.9)$$

$$\rho(t) = \frac{N(t)}{K}. \quad (5.10)$$

$N_q(t)$, $N_p(t)$, and $N_a(t)$ are the cell counts in compartments Q, P, and A, respectively, at time t , and K represents a constant that defines $\rho = 1$, which should be interpreted as the number of cells which occupy a single layer of the culture. Throughout this work, K was scaled to be 40401, for a 201 cell by 201 cell square environment. $c > 0$ is a per time constant which represents a cellular reaction rate, and $\gamma \in [0, 1]$ is a unitless proportion corresponding to the difference in arrivals to compartment A via compartments P and Q. Note that all quantities are dynamic and stochastic.

5.2.2 Rate derivations

In this section, we provide motivation for the forms employed in equations (5.6)-(5.8). Consider transitions from quiescence to division (Q to P). Our fundamental

assumption is that there exists a theoretical $\beta(\rho)$ (represented by (5.4) with sample visualization appearing as the red curve in Figure 5.2(a)) which yield the fraction of cells that are in compartment P at equilibrium. Hence, all cells on the culture calibrate towards the fraction in this figure. Converting fractions to cell numbers, one can mathematically describe the desired number of proliferative cells as

$$N_{p,\text{desired}}(t) = \beta(\rho(t))N(t), \quad (5.11)$$

where in general, $N_{p,\text{desired}}(t)$ and $N_p(t)$ are not equal, and thus cells transfer compartments. By the interpretation of Poisson rates, the number of cells transferring from Q to P in the time interval $[t, t + \Delta t]$ is approximately $\alpha_p(t)Q(t)\Delta t$, for small Δt . On the other hand, it is natural to assume that the number of transitioning cells should be proportional to the difference of the desired and current states:

$$\alpha_p(t)N_q(t)\Delta t \propto (N_{p,\text{desired}}(t) - N_p(t))_+ = (\beta(\rho(t))N(t) - N_p(t))_+ \quad (5.12)$$

by equation (5.11). The positive part is there to ensure that we do not take a negative number of cells, i.e., that once a cell is dividing, it cannot return to quiescence without successfully completing mitosis. Assuming all cells react to the disparity between $N_{p,\text{desired}}(t)$ and $N_p(t)$ at the same constant rate c , (5.12) can be written as

$$\alpha_p(t)N_q(t)\Delta t = c(\beta(\rho(t))N(t) - N_p(t))_+\Delta t, \quad (5.13)$$

or equivalently, equation (5.6), which is the desired expression for $\alpha_p(t)$, and will be the transfer rate of one cell from compartment Q to compartment P during $[t, t + \Delta t]$. If the population becomes so small that $N_q(t) = 0$ for some time t , then no cells can transfer, and hence we must augment the expression for $\alpha_p(t)$ as

$$\alpha_p(t) = \begin{cases} c \frac{(\beta(\rho(t))N(t) - N_p(t))_+}{N_q(t)} & \text{if } N_q(t) \neq 0, \\ 0 & \text{if } N_q(t) = 0. \end{cases} \quad (5.14)$$

It will be understood that all other rates ($\alpha_{a_p}(t)$ and $\alpha_{a_q}(t)$ in (5.7) and (5.8), respectively) have similar piecewise definitions, which we omit stating explicitly because we believe it will cause no confusion to the reader. We also note that in the simulations for this work, the fraction of dying cells is small enough ($\leq 5\%$) to prevent any cell compartment from becoming extinct. This is described in Section 5.4.

In an analogous way we defined the rates from P to A and from Q to A, which are denoted by $\alpha_{a_p}(t)$ and $\alpha_{a_q}(t)$, respectively. However, Figure 5.2(b) represents the total number of cells entering A, from both compartments P and Q. To separate the transitions between these compartments, we assumed that the relative number of cells transferring between P and Q differs by a constant. Hence, assuming cells respond with the same rate c as above, the death rates can be written as (5.7) and (5.8), where $0 \leq \gamma \leq 1$ represents the arrival difference factor. Furthermore, we assume that dividing cells are more prone to cell death than quiescent cells [169], i.e. $\gamma > \frac{1}{2}$. Note that (5.7) and (5.8) are indeed the correct expressions in our

framework, since in a small time Δt , the total number of cells transitioning into compartment A is given approximately by

$$\begin{aligned}
& \alpha_{a_p}(t)N_p(t)\Delta t + \alpha_{a_q}(t)N_q(t)\Delta t \\
&= c\gamma \frac{(dN(t) - N_a(t))_+}{N_p(t)} N_p(t)\Delta t + c(1 - \gamma) \frac{(dN(t) - N_d(t))_+}{N_q(t)} N_q(t)\Delta t \\
&= c\gamma(dN(t) - N_a(t))_+ + c(1 - \gamma)(dN(t) - N_d(t))_+ \\
&= c\Delta t(dN(t) - N_d(t))_+,
\end{aligned} \tag{5.15}$$

which is the desired total number of cells.

5.2.3 ABM algorithm

We include a brief outline of the steps involved in the algorithm. Suppose at time t we have $N(t)$ cells, each residing in one of the three compartments Q, P, or A, with total cell counts $N_q(t)$, $N_p(t)$, and $N_a(t)$, respectively. We update the system to time $t + \Delta t$ by the following steps:

1. Choose a uniformly random order \mathcal{O} of the cells.
2. Update each cell according to \mathcal{O} :
 - a. If the cell is in compartment Q, do one of the following:
 - i. Enter P with probability $\alpha_p(t)\Delta t$ and choose L_P via equation (5.1).
 - ii. Enter A with probability $\alpha_{a_q}(t)\Delta t$ and choose L_A via equations (5.49) and (5.50).

- iii. Remain in Q with probability $1 - \alpha_p(t)\Delta t - \alpha_{a_q}(t)\Delta t$.
- b. If the cell is in compartment P, do one of the following:
 - i. Enter A with probability $\alpha_{a_p}(t)\Delta t$ and choose L_A via equations (5.49) and (5.50).
 - ii. If the elapsed time of the cell-cycle is L_P , then with probability $1 - \alpha_{a_p}(t)\Delta t$, the cell and its daughter move to compartment Q.
 - iii. If the elapsed time of the cell-cycle is less than L_P , then with probability $1 - \alpha_{a_p}(t)\Delta t$, the cell-cycle age increases by Δt .
 - c. If the cell is in compartment A, do one of the following:
 - i. If the elapsed time of the apoptosis-cycle is L_A , then the cell is removed from the simulation.
 - ii. If the elapsed time of the apoptosis-cycle is less than L_A , then the apoptosis-cycle age increases by Δt .
- 3. Update all cell counts to $N_p(t + \Delta t)$, $N_a(t + \Delta t)$, $N_q(t + \Delta t)$, and $N(t + \Delta t)$.

5.3 Integro-differential equation model

In this work, we study the dynamic mechanisms of cell growth, assuming variations in the cell-cycle and apoptotic periods. The entire system dynamic results from the interactions and decisions of individual cells, which could potentially lead to significant deviations from mean values. However, it is well-known that despite the presence of intrinsic fluctuations, the dynamics of a system with a size that

is sufficiently large may be captured by deterministic equations that reflect mean field limits [170–172]. Thus, we approximate the ABM with a system of integro-differential equations (IDEs) that approximate the expected values of the cellular compartment sizes in time. Details on the derivation and validity of the following system can be found in Section 5.3.2.

5.3.1 Approximation of expected values

We make the notational convention for mean values: $\mathbb{E}[\cdot] := \langle \cdot \rangle$. Following the above discussion, a set of IDEs which approximate the expected value of the ABM random variables is formulated, with corresponding rates evaluated at expected values:

$$\overline{N}_p(t) := \langle N_p \rangle(t), \quad (5.16)$$

$$\overline{\alpha}_p(t) := c \frac{\left(\beta(\langle \rho(t) \rangle) \langle N \rangle(t) - \langle N_p \rangle(t) \right)_+}{\langle N_q \rangle(t)}, \quad (5.17)$$

for example. Note that $\overline{\alpha}_p(t) \neq \langle \alpha_p \rangle(t)$, in general. The IDE system which approximates the mean fields of the three compartment cellular ABM is given below:

$$\begin{aligned} \frac{d}{dt} \overline{N}_p(t) &= \overline{\alpha}_p(t) \overline{N}_q(t) - \overline{\alpha}_{a_p}(t) \overline{N}_p(t) \\ &\quad - \int_0^t f_p(t-t_*; \mu, \sigma) \overline{\alpha}_p(t_*) \overline{N}_q(t_*) \left(1 - \int_{t_*}^t \overline{\alpha}_{a_p}(s) ds\right) dt_*, \end{aligned} \quad (5.18)$$

$$\begin{aligned} \frac{d}{dt} \overline{N}_a(t) &= \overline{\alpha}_{a_q}(t) \overline{N}_q(t) + \overline{\alpha}_{a_p}(t) \overline{N}_p(t) - \int_0^t f_a(t-t_*) \overline{\alpha}_{a_q}(t_*) \overline{N}_q(t_*) dt_* \\ &\quad - \int_0^t f_a(t-t_*) \overline{\alpha}_{a_p}(t_*) \overline{N}_p(t_*) dt_*, \end{aligned} \quad (5.19)$$

$$\begin{aligned} \frac{d}{dt} \overline{N}_q(t) &= -\overline{\alpha}_{a_q}(t) \overline{N}_q(t) - \overline{\alpha}_p(t) \overline{N}_q(t) \\ &\quad + 2 \int_0^t f_p(t-t_*; \mu, \sigma) \overline{\alpha}_p(t_*) \overline{N}_q(t_*) \left(1 - \int_{t_*}^t \overline{\alpha}_{a_p}(s) ds\right) dt_*. \end{aligned} \quad (5.20)$$

Here $f_p(\cdot; \mu, \sigma)$ and $f_a(\cdot)$ denote the probability density functions (PDFs) for the cell-cycle and apoptotic process length random variables, L_P and L_A , respectively.

Recalling equations (5.1) and (5.2), $f_p(\cdot; \mu, \sigma)$ and $f_a(\cdot)$ are

$$f_p(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}, \quad (5.21)$$

$$f_a(x) = \frac{\lambda^\omega}{\Gamma(\omega)} x^{\omega-1} e^{-\lambda x}, \quad x \geq 0, \quad (5.22)$$

where $\Gamma(\cdot)$ is the gamma function.

Equations (5.18)-(5.20) represent a system of delay differential equations (DDEs), where the delay is captured by integrals over time. However, instead of possessing a fixed value, we see that the delays are distributed, with relative contributions to the flow determined by the densities $f_p(\cdot; \mu, \sigma)$ and $f_a(\cdot)$.

5.3.2 The derivation of the IDE

Consider the ABM introduced in Section 5.2. Fix a small time-step $\Delta t > 0$ and a time $t \geq 0$. Consider the system update from t to $t + \Delta t$ in the, say, cell-cycle compartment P. We can decompose the random variable N_p as follows:

$$N_p(t + \Delta t) = N_p(t) + a_p(t, \Delta t) - a_{a_p}(t, \Delta t) - x(t, \Delta t), \quad (5.23)$$

where $a_p(t, \Delta t)$ is the number of cells entering P from Q in the time interval $[t, t + \Delta t]$, $a_{a_p}(t, \Delta t)$ is the number of cells exiting P and entering A in $[t, t + \Delta t]$, and $x(t, \Delta t)$ is the number of cells exiting P and entering Q in $[t, t + \Delta t]$ through successful divisions. Taking the expectation of (5.23) and using linearity we obtain

$$\langle N_p \rangle(t + \Delta t) = \langle N_p \rangle(t) + \langle a_p \rangle(t, \Delta t) - \langle a_{a_p} \rangle(t, \Delta t) - \langle x \rangle(t, \Delta t). \quad (5.24)$$

We continue by finding expressions for the terms on the right-hand side (RHS) of (5.24).

Consider $\langle a_p \rangle(t, \Delta t)$. Using the Poisson interpretation of the ABM, the number of cells transitioning from Q to P in $[t, t + \Delta t]$ is given by the product of the expected probability of transition of one cell and the expected number of cells currently residing in compartment Q:

$$\langle a_p \rangle(t + \Delta t) = \langle \alpha_p(t) \Delta t \rangle \langle N_q \rangle(t) = \left\langle c \frac{(\beta(\rho(t))N(t) - N_p(t))_+}{N_q(t)} \right\rangle \langle N_q \rangle(t) \Delta t \quad (5.25)$$

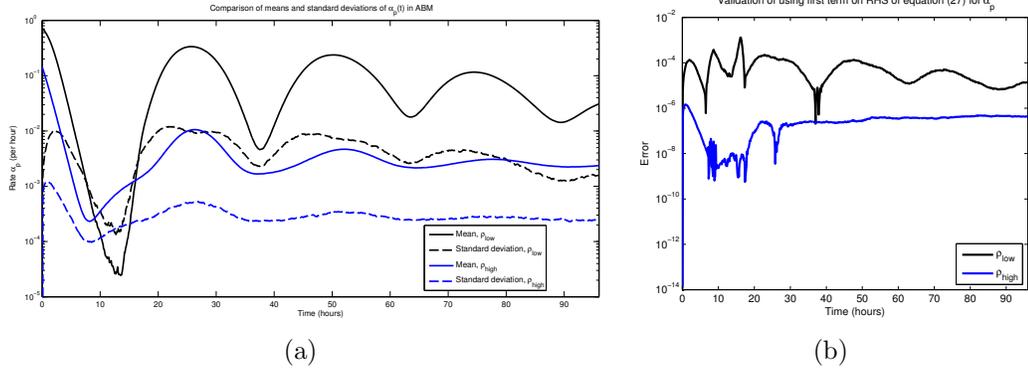


Figure 5.3: (a) The mean and standard deviation of $\alpha_p(t)$ over all times $t \in [0, 96]$ hours, for 1000 Monte Carlo simulations. β and d appear as the red curves in Figures 5.2(a) and 5.2(b) respectively, and $\sigma = 3$ hours, $c = 1$ cell per hour. See equations (5.51) and (5.52) for the two sets of initial conditions. Note that the standard deviations are roughly an order of magnitude smaller than the mean values, for most times. (b) Error obtained in using the first term on the RHS in equation (5.27) as an approximation for $\alpha_p(t)$

Our goal is to simplify the above expression, so as to obtain a closed system of equations; otherwise we must perform Monte-Carlo simulations to compute the above expectations, which would render the deterministic model ineffective. A natural step would then be to reverse the order of the nonlinearity and expectation, while incurring an error, as $\langle f(\mathbf{X}) \rangle \neq f(\langle \mathbf{X} \rangle)$ for a nonlinear f and a random vector \mathbf{X} , in general. However, we recall from Figures 5.6(c) and 5.6(d) that the variances in the ABM are small, and hence we expect the error of approximating \mathbf{X} with its mean to also be small, even through the function f . Indeed, we can expand a general $f(\mathbf{X})$ in a Taylor series about the mean value of \mathbf{X} as

$$f(\mathbf{X}) = f(\langle \mathbf{X} \rangle) + Df(\langle \mathbf{X} \rangle)(\mathbf{X} - \langle \mathbf{X} \rangle) + \frac{1}{2}(\mathbf{X} - \langle \mathbf{X} \rangle)^T D^2 f(\langle \mathbf{X} \rangle)(\mathbf{X} - \langle \mathbf{X} \rangle) + \dots, \quad (5.26)$$

where D^2f denotes the Hessian of f . Upon taking expectations and retaining only these first three terms, we obtain

$$\langle f(\mathbf{X}) \rangle \approx f(\langle \mathbf{X} \rangle) + \frac{1}{2} \text{tr} \left(D^2 f(\langle \mathbf{X} \rangle) \text{Var}[\mathbf{X}] \right). \quad (5.27)$$

Recall that $Df(\langle \mathbf{X} \rangle), D^2f(\langle \mathbf{X} \rangle)$ are constants with respect to the expectation operator. Furthermore, the last term in (5.27) is obtained from the quadratic form appearing in (5.26) (see [173]). Hence, assuming f is sufficiently smooth, our previous variance calculations imply that $\langle f(\mathbf{X}) \rangle \approx f(\langle \mathbf{X} \rangle)$ [174]. To justify the smoothness condition, note that here, $f = \alpha_p$, where α_p is given by equation (5.6). The $Q = 0$ case is not an issue since, as discussed in Section 5.2.2, cell fractions do not become extinct. The only other points where α_p could violate smoothness are times such that $\alpha_p(t) = 0$, since the positive part function is non-differentiable here. Hence, it is sufficient to verify that $\alpha_p(t)$ is bounded away from 0 for all times t . Thus, we ran 1000 Monte Carlo simulations of the ABM, and calculated the mean and standard deviation of $\alpha_p(t)$ at each time point; see Figure 5.3 (a). All parameters values were as in Section 5.4.2 below. Note that for each initial condition ($\rho(0) = 0.1$ and $\rho(0) = 0.8$), the standard deviation is approximately one order of magnitude smaller than that of the mean. As the mean is bounded away from zero, the law of large numbers implies that $\alpha_p(t) = 0$ is statistically improbable, for most times t . The same conclusions can be verified for the other rates $\alpha_{a_p}(t)$ and $\alpha_{a_q}(t)$. Our statement is about “most times” due to the observation that there is a region in Figure 5.3 where the the standard deviation is larger than the mean, and

hence $\alpha_p(t)$ can be 0 with non-zero probability. Indeed, this is observed in realizations. However, this timescale is short and transient, which implies that switching the order of the expectation and the nonlinearity is a valid approximation. We also compare the approximation directly in Figure 5.3(b), and note that for all times, the error is smaller than 10^{-3} . Furthermore, comparing the stochastic ABM to the obtained continuous DEs is another form of validation (see Section 5.4.3).

Using equation (5.25), the previous justification allows us to approximate $\langle a_p \rangle(t, \Delta t)$ as

$$\langle a_p \rangle(t, \Delta t) \approx c \frac{\left(\beta(\langle \rho \rangle(t)) \langle N \rangle(t) - \langle N_p \rangle(t) \right)}{\langle N_q \rangle(t)} \langle N_q \rangle(t) \Delta t = \overline{\alpha_p}(t) \langle N_q \rangle(t) \Delta t, \quad (5.28)$$

where $\overline{\alpha_p}(t)$ is defined as α_p evaluated at the expected values of the compartment populations. In the same way, we obtain an approximation for $\langle a_{a_p} \rangle(t, \Delta t)$:

$$\langle a_{a_p} \rangle(t, \Delta t) \approx \overline{\alpha_{a_p}}(t) \langle N_p \rangle(t) \Delta t. \quad (5.29)$$

Here $\overline{\alpha_{a_p}}(t)$ is defined analogously to $\overline{\alpha_p}(t)$. From this point forward, we drop both the expectation brackets and bars so that, for example, $N_p(t)$ denotes the expected population of the dividing cells at time t , and $\alpha_p(t)$ denotes the transition rate of a cell to compartment P from compartment Q at time t , where the compartment populations are given by their means. This should cause no confusion with realizations of the ABM, and when comparisons are made, it should be understood that the deterministic results represent expectations. Using this notation and the

approximations in (5.28) and (5.29) as equalities, equation (5.24) becomes

$$N_p(t + \Delta t) = N_p(t) + \alpha_p(t)N_q(t)\Delta t - \alpha_{a_p}(t)N_p(t)\Delta t - \langle x \rangle(t, \Delta t). \quad (5.30)$$

To determine $\langle x \rangle(t, \Delta t)$, the expected number of cells that complete cell division in the time interval $[t, t + \Delta t]$, we note the following conditions that must hold for such a cell:

1. It must enter P at some time $t_* \leq t + \Delta t$.
2. Its length of division ℓ must satisfy $t \leq t_* + \ell \leq t + \Delta t$, or equivalently, $t - t_* \leq \ell \leq t - t_* + \Delta t$.
3. It must not enter the apoptotic cycle A at any time in the interval $[t_*, t + \Delta t]$.

In our model formulation, all three of the above events are independent. Thus, to determine $\langle x \rangle(t, \Delta t)$, we must add up (i.e. integrate over) all times t_* , where $0 \leq t_* \leq t + \Delta t$, the product of the expected number of cells that enter P at t_* and the probability of a cell choosing the correct division length ℓ while not entering A during $[t_*, t + \Delta t]$. We recall that the division length is distributed according to (5.1), with corresponding Gaussian probability density function (PDF) $f_p(\cdot; \mu, \sigma)$. The expected number of cells that enter P from Q between times t_* and $t_* + \Delta t_*$ for small Δt_* is given approximately by $\alpha_p(t_*)N_q(t_*)\Delta t_*$, and hence the expected number yielding the correct division length is approximately

$$\alpha_p(t_*)N_q(t_*)\Delta t_* \int_{t-t_*}^{t-t_*+\Delta t} f_p(x; \mu, \sigma) dx. \quad (5.31)$$

However, we must take only the fraction of these cells that remain in P during $[t_*, t + \Delta t]$, which we obtain by multiplying (5.31) by the probability \mathcal{P} of a cell not entering A throughout $[t_*, t + \Delta t]$. Again, an independence assumption between cell-cycle length and the event of entering the apoptosis cycle is used. To calculate \mathcal{P} , we note that

$$\mathcal{P} = 1 - \mathcal{P}', \quad (5.32)$$

where \mathcal{P}' is the probability of entering A at some point during $[t_*, t + \Delta t]$. Dividing $[t_*, t + \Delta t]$ into non-overlapping subintervals $[t_* + (i-1)\Delta t'_*, t_* + i\Delta t'_*]$ via small $\Delta t'_*$, the event corresponding to \mathcal{P}' can be realized as the disjoint union of sub-events, where each sub-event corresponds to entering A in a subinterval $[t_* + (i-1)\Delta t'_*, t_* + i\Delta t'_*]$ with corresponding probability $\alpha_{a_1}(t_* + (i-1)\Delta t'_*)\Delta t'_*$. Mutual exclusivity and the fact that these sub-events cover the event corresponding to \mathcal{P}' implies that

$$\mathcal{P}' \approx \sum_{i=1}^n \alpha_{a_p}(t_* + (i-1)\Delta t'_*)\Delta t'_* \xrightarrow{\Delta t'_* \rightarrow 0^+} \int_{t_*}^{t+\Delta t} \alpha_{a_p}(s) ds, \quad (5.33)$$

where in the limit $\Delta t'_* \rightarrow 0^+$, the expression becomes exact. Here $n \in \mathbb{Z}$ is such that $t_*(n-1)\Delta t_* = t + \Delta t$. Combining (5.31), (5.32), and (5.33) yields an approximation for the number of cells that enter at a fixed t_* and carry out a division in $[t, t + \Delta t]$:

$$\alpha_p(t_*)N_q(t_*)\Delta t_* \int_{t-t_*}^{t-t_*+\Delta t} f_p(x; \mu, \sigma) dx \left(1 - \int_{t_*}^{t+\Delta t} \alpha_{a_p}(s) ds \right). \quad (5.34)$$

Adding contributions from every $0 \leq t_* \leq t + \Delta t$ produces $\langle x \rangle(t, \Delta t)$. In the limit

as $\Delta t_* \rightarrow 0^+$, the summation becomes an integral, and we obtain

$$\langle x \rangle(t, \Delta t) = \int_0^{t+\Delta t} \alpha_p(t_*) N_q(t_*) \int_{t-t_*}^{t-t_*+\Delta t} f_p(x; \mu, \sigma) dx \left(1 - \int_{t_*}^{t+\Delta t} \alpha_{a_p}(s) ds \right) dt_*. \quad (5.35)$$

Note that in this limit, (5.35) is exact. Substituting this expression into (5.30) yields, upon rearrangement and division by Δt ,

$$\begin{aligned} & \frac{N_p(t + \Delta t) - N_p(t)}{\Delta t} \\ &= \alpha_p(t) N_q(t) - \alpha_{a_p}(t) N_p(t) \\ & - \int_0^{t+\Delta t} \alpha_p(t_*) N_q(t_*) \frac{1}{\Delta t} \int_{t-t_*}^{t-t_*+\Delta t} f_p(x; \mu, \sigma) dx \left(1 - \int_{t_*}^{t+\Delta t} \alpha_{a_p}(s) ds \right) dt_*. \end{aligned} \quad (5.36)$$

Taking the limit of (5.36) as $\Delta t \rightarrow 0^+$, we have, via the Fundamental Theorem of Calculus,

$$\begin{aligned} \frac{d}{dt} N_p(t) &= \alpha_p(t) N_q(t) - \alpha_{a_p}(t) N_p(t) \\ & - \int_0^t f_p(t - t_*; \mu, \sigma) \alpha_p(t_*) N_q(t_*) \left(1 - \int_{t_*}^t \alpha_{a_p}(s) ds \right) dt_*. \end{aligned} \quad (5.37)$$

Equation (5.37) is an integro-differential equation (IDE), where a delay is represented by the integral terms. Indeed, if cells divided after a fixed length of time ℓ , instead of randomly, then f_p takes the form of a δ -function, and (5.37) reduces to a

delay-differential equation (DDE):

$$\begin{aligned} \frac{d}{dt} \langle N_p \rangle(t) &= \alpha_p(t)N_q(t) - \alpha_{a_p}(t)N_p(t) \\ &\quad - \alpha_p(t-\ell)N_q(t-\ell) \left(1 - \int_{t-\ell}^t \alpha_{a_p}(s) ds \right). \end{aligned} \quad (5.38)$$

Hence, (5.37) represents the expected value of system of DDEs with distributed time delays.

We can similarly derive equations for the expected value of $N_a(t)$ and $N_q(t)$:

$$\begin{aligned} \frac{d}{dt} N_a(t) &= \alpha_{a_q}(t)N_q(t) + \alpha_{a_p}(t)N_p(t) - \int_0^t f_a(t-t_*)\alpha_{a_q}(t_*)N_q(t_*) dt_* \\ &\quad - \int_0^t f_a(t-t_*)\alpha_{a_p}(t_*)N_p(t_*) dt_*, \end{aligned} \quad (5.39)$$

and

$$\begin{aligned} \frac{d}{dt} N_q(t) &= -\alpha_{a_q}(t)N_q(t) - \alpha_p(t)N_q(t) \\ &\quad + 2 \int_0^t f_p(t-t_*; \mu, \sigma)\alpha_p(t_*)N_q(t_*) \left(1 - \int_{t_*}^t \alpha_{a_p}(s) ds \right) dt_*. \end{aligned} \quad (5.40)$$

where $f_a(\cdot)$ is a gamma PDF for the apoptotic process length (see Section 5.4.1 and Figure 5.5(d)). Note that there is no inner integral in (5.39) since we assumed that once cells enter apoptosis, they cannot escape eventual cell death. Also, a factor of two appears in (5.40), since the original cell and its daughter both re-enter state Q from P. We can also write an equation for the expected number of the total number

of cells, $N(t)$, by summing all of (5.37), (5.39), and (5.40):

$$\begin{aligned} \frac{d}{dt}N(t) = & \int_0^t f_p(t-t_*; \mu, \sigma)\alpha_p(t_*)N_q(t_*) \left(1 - \int_{t_*}^t \alpha_{a_p}(s) ds\right) dt_* \\ & - \int_0^t f_a(t-t_*)\alpha_{a_q}(t_*)N_q(t_*) dt_* - \int_0^t f_a(t-t_*)\alpha_{a_p}(t_*)N_p(t_*) dt_*. \end{aligned} \quad (5.41)$$

Equations (5.37), (5.39), and (5.40) correspond to the system of equations presented in Section 5.3.1. Standard theory implies existence and uniqueness of these equations [175].

5.3.3 Numerical methods

We provide a brief outline of the numerical algorithm used in solving (5.18)-(5.20). The scheme used was the four-step explicit Adams-Bashforth method [176], which for a general system

$$\frac{dy}{dt} = f(t, y(t)), \quad (5.42)$$

discretizes via interpolation of f through four previous points, and hence generates a fourth order accurate method. For an explicit, uniform time-step Δt , the method takes the form

$$y_{n+1} = y_n + \frac{\Delta t}{24}(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3}), \quad (5.43)$$

where $t_n := n\Delta t$, $y_n := y(t_n)$, and $f_n := f(t_n, y_n)$. For the first 3 steps, (5.43) is augmented with standard initializations of lower degree Adams-Bashforth methods.

Regarding the calculation of the integrals appearing in (5.18)-(5.20), we consider integrals of the form

$$\int_0^t h_1(t, t_*) \left(1 - \int_{t_*}^t h_2(s) ds \right) dt_*, \quad (5.44)$$

for arbitrary functions h_1 and h_2 and time t . These integrals are discretized with the same step-size Δt used in the numerical algorithm described previously. Denoting the inner integral in (5.44) by $I(t, t_*)$, $I(t + \Delta t, t_*)$ is related to $I(t, t_*)$ via the approximation

$$\begin{aligned} I(t + \Delta t, t_*) &= \int_{t_*}^{t+\Delta t} h_2(s) ds = \int_{t_*}^t h_2(s) ds + \int_t^{t+\Delta t} h_2(s) ds \\ &\approx \int_{t_*}^t h_2(s) ds + h_2(t)\Delta t \\ &= I(t, t_*) + h_2(t)\Delta t. \end{aligned} \quad (5.45)$$

Using (5.45) to calculate the inner integral requires only one function evaluation and no quadrature, and hence is more numerically efficient than re-integration. Note that we must store values of $I(t, \cdot)$ at all time points in the discretization. To calculate the complete integral (5.44), the composite trapezoidal rule with uniform grid spacing Δt was used:

$$\begin{aligned} &\int_0^t h_1(t, t_*) \left(1 - \int_{t_*}^t h_2(s) ds \right) dt_* \\ &\approx \frac{\Delta t}{2} \sum_{k=1}^N \left(h_1(t, t_k)(1 - I(t, t_k)) + h_1(t, t_{k+1})(1 - I(t, t_{k+1})) \right) \end{aligned} \quad (5.46)$$

where t_k is the k th point, and N is the size of, the discretization. Updating the

expression from $t \rightarrow t + \Delta t$ requires this calculation at each step, because here, unlike h_2 , h_1 depends explicitly on t , and thus prevents an approximation analogous to (5.45). Equations (5.45) and (5.46) give a complete characterization of the calculation of all integrals appearing in this work.

5.4 Results

Cell survival is determined by many factors, including the ability of the cell to transition between the cell-cycle and quiescence, and vice versa. These dynamics can be described mathematically, as both an ABM and a system of IDEs, as discussed previously. Using the OVCAR-8 cell line as a model, we performed three fundamental biological experiments, and estimated the following: 1) cell growth and density over 4 days (i.e. $\rho(t)$), 2) cell proliferation fraction as a function of cell density after 24 hours (i.e. $\beta(\rho)$), and 3) cell death fraction as a function of cell density after 24 hours (i.e. $d(\rho)$). See Appendix C.1 for details and the experimental results.

To describe the biological results, we introduce two mathematical models: an ABM and a set of IDEs. Detailed descriptions of our methods can be found in Sections 5.2 and 5.3. The ABM is specified on the level of the cell, which can reside in one of three compartments: proliferative (P), apoptotic (A), and quiescent (Q), with corresponding populations N_p , N_a , and N_q (see Figure 5.1). Transitions between the compartments are dictated by both the rates α_p , α_{a_p} , and α_{a_q} , and the amount of time spent in the the cell-cycle and the apoptotic process. The rate functions depend on the total cellular density ρ and the equilibrium distributions

$\beta(\rho)$ and $d(\rho)$, while the time spent in P is assumed to be a normal random variable with mean length μ hours, and standard deviation σ hours. Similarly, the amount of time spent in A is modeled with a gamma distribution, with shape parameter ω and rate parameter λ per hour. The IDE system is a set of three equations governing the dynamics of approximations to the expected values of the compartment populations.

5.4.1 Distribution of cell-cycle and apoptotic lengths

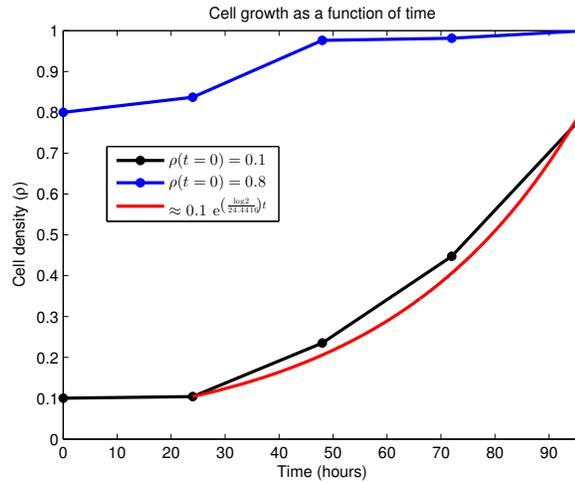


Figure 5.4: Growth dynamics of OVCAR-8 human ovarian carcinoma cell line. Global density is measured on the vertical axis as a function of time. Both initial densities ($\rho(0) = 0.1$ and $\rho(0) = 0.8$) were generated by random seeding on a 24-well plate and cultured 24 hours prior to the first image. Data was collected every 24 hours, as indicated by stars. No data exists between the points, and the connecting line segments are added merely for viewing elucidation. The estimated density percentage is based on the average of two complete wells. The red curve represents a least-squares fit to the region of exponential growth, with doubling time $\mu = 24.4416$ hours.

To calculate the mean length of the cell-cycle μ , we used the experimental data appearing in Figure 5.4. See C.1.1.2 for details. We observed essentially exponential growth from day 2 to day 4 in the low density seeding (denoted ρ_{low}). Standard

least-squares techniques fit

$$\rho_{\text{low}}(t = 24)e^{kt} \tag{5.47}$$

to the data in this region, and k was related to the mean length of the cell-cycle via the doubling time:

$$\mu = \frac{\log 2}{k}. \tag{5.48}$$

See the red curve in Figure 5.4 for the results of the calculation, which conclude $\mu = 24.4416$ hours.

The calculation of the distribution of L_A was based on the results of [38]. In their work, Messam and Pittman observed that apoptosis could be characterized by three morphological phases. In particular, once apoptosis is triggered, cells begin by exhibiting normal morphology (phase 1), then plasma membrane bubbling (phase 2), and finally whole cell body blebbing and apoptosis execution (phase 3). Furthermore, the authors obtained experimental results, in the form of frequency distributions, for how long cells remain in each of these three phases. Using their frequency distributions, we fit gamma distributions to each phase, and then independently sum the distributions to obtain the overall apoptosis process length L_A as

a random variable. Denoting the length of phase i ($i = 1, 2, 3$) by L_{A_i} , this becomes

$$L_{A_i} \sim \Gamma(\omega_i, \lambda_i), \quad (5.49)$$

$$L_A = L_{A_1} + L_{A_2} + L_{A_3}. \quad (5.50)$$

Here ω_i and λ_i are the shape and rate parameters of phase i , respectively. To find the distributions of each phase, we used the MATLAB function “fitdist.m,” which uses standard maximum likelihood estimation, to find the parameters characterizing the distribution. Moreover, we restricted the search domain to gamma distributions. The results of this, along with the experimental results appearing in [38], can be found in Figure 5.5. With these computations, we find a probability density function (PDF) for each L_{A_i} , and via equations (5.49) and (5.50), are able to realize a value for L_A for each cell entering A. We are further able to fit a gamma distribution to 10^5 realizations of (5.50), which yields (5.2) with $\omega = 4.9436$ and $\lambda = 0.19117$ (Figure 5.5(d)).

5.4.2 ABM simulation

We provide a sample experiment of the ABM, which demonstrates the behavior of the model. We used $\beta(\rho)$ and d appearing as the red curves in Figures 5.2(a) and 5.2(b) respectively, $\sigma = 3$ hours, $c = 1$ cell per hour, and $\gamma = 0.70$. We emphasize that these parameters are not a fit to the experimental data appearing in Figure 5.4. Before computing, we specify the initial conditions, which are used in the models throughout this work, unless explicitly stated otherwise. Experimentally, all cells

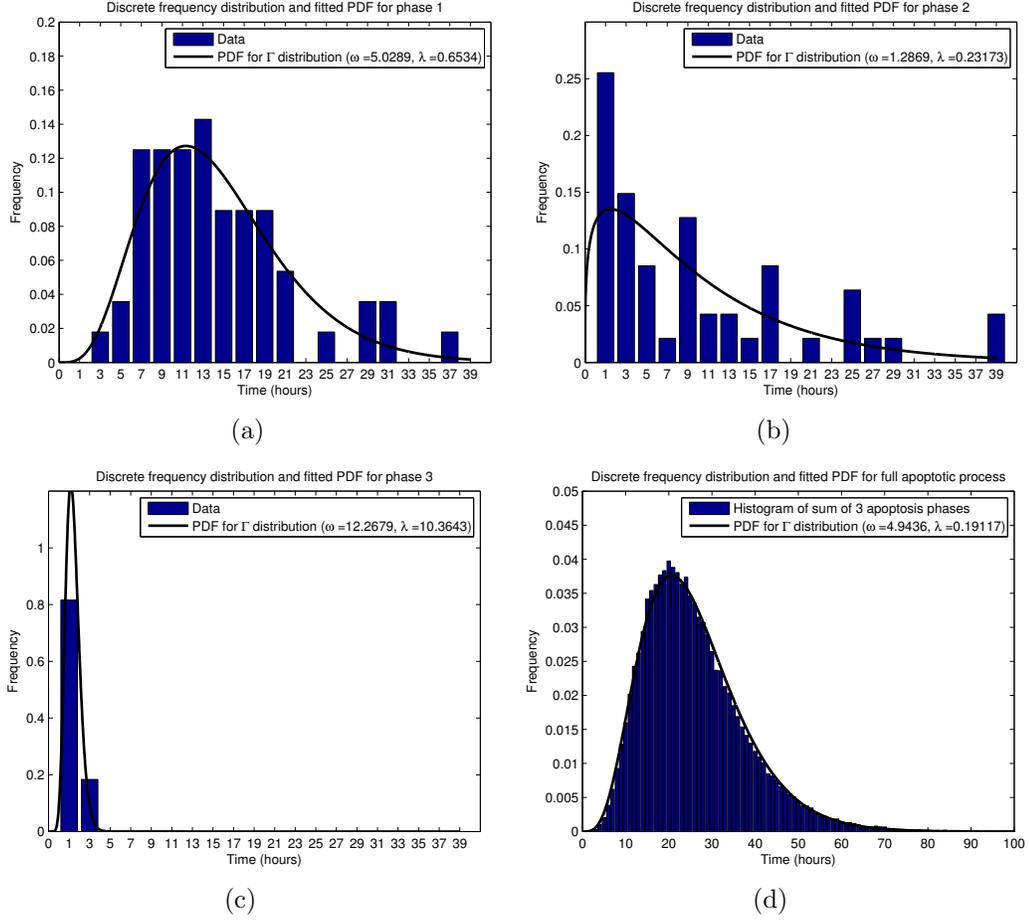


Figure 5.5: Frequency distributions for the apoptosis process. The x -axis represents the time in the respective phase, while the y -axis labels a fraction of cells. For the experimental results (histogram in (a)-(c)), the x value indicates a range of times, so that for $x = i$ hours, the corresponding y -value (height of the bar) counts the relative frequency of cells with phase length between $i - 1$ and $i + 1$ hours. For the theoretical distribution (black curve), we allowed the phase length to be a continuous random variable, and fitted a gamma distribution with probability density function (PDF) using MATLAB’s “fitdist.m” function. In (d), we used the PDFs obtained in (a)-(c) to obtain a single gamma distribution for the entire apoptotic process. (a) Phase 1, with $n = 56$ total cells measured; (b) Phase 2, with $n = 47$ total cells measured; (c) Phase 3, with $n = 49$ total cells measured; (d) Total length of time spent in apoptosis, using $n = 10^5$ simulations of (5.50).

are initially synchronized as quiescent, so we define

$$N_p(0) = 0, \tag{5.51}$$

$$N_a(0) = 0.$$

Using the data from Figure 5.4 as motivation, we simulate two sets of initial conditions, $\rho_{\text{low}}(0) = 0.1$ and $\rho_{\text{high}}(0) = 0.8$. Equations (5.10) and (5.51) imply that

$$\begin{aligned} N_{q_{\text{low}}}(0) &= \text{round}(0.1K), \\ N_{q_{\text{high}}}(0) &= \text{round}(0.8K). \end{aligned} \tag{5.52}$$

One thousand Monte Carlo simulations were then performed, the results of whose mean and variance calculations appear in Figure 5.6. The main aspect to note is the expected density dynamics in Figure 5.6(a), with faster growth for smaller population values coupled with a flattening as ρ approaches 1. In the same figure, we also observe periodic prominences of density increase located approximately 24 hours apart, which coincide with the mean division time μ . In Figure 5.6(b) we see periodic population fluctuations between the compartments Q and P, as cells complete division and reenter quiescence. A larger fraction of cells migrate to compartment P initially in the simulation satisfying $\rho_{\text{low}}(0) = 0.1$, as β is larger here at time $t = 0$. We also do not observe much fluctuation in compartment A, which is due to the fact that $d(\rho)$ is taken as a constant. Figures 5.6(c) and 5.6(d) are included mainly to quantify the stochasticity in the model (on the order of 10^{-6}). We note that the small variation between realizations is due to the large number of cells in the computation ($\mathcal{O}(10^5)$). Note also that at time $t = 0$ the variances evaluate to machine 0, as the initial conditions are identical in every realization.

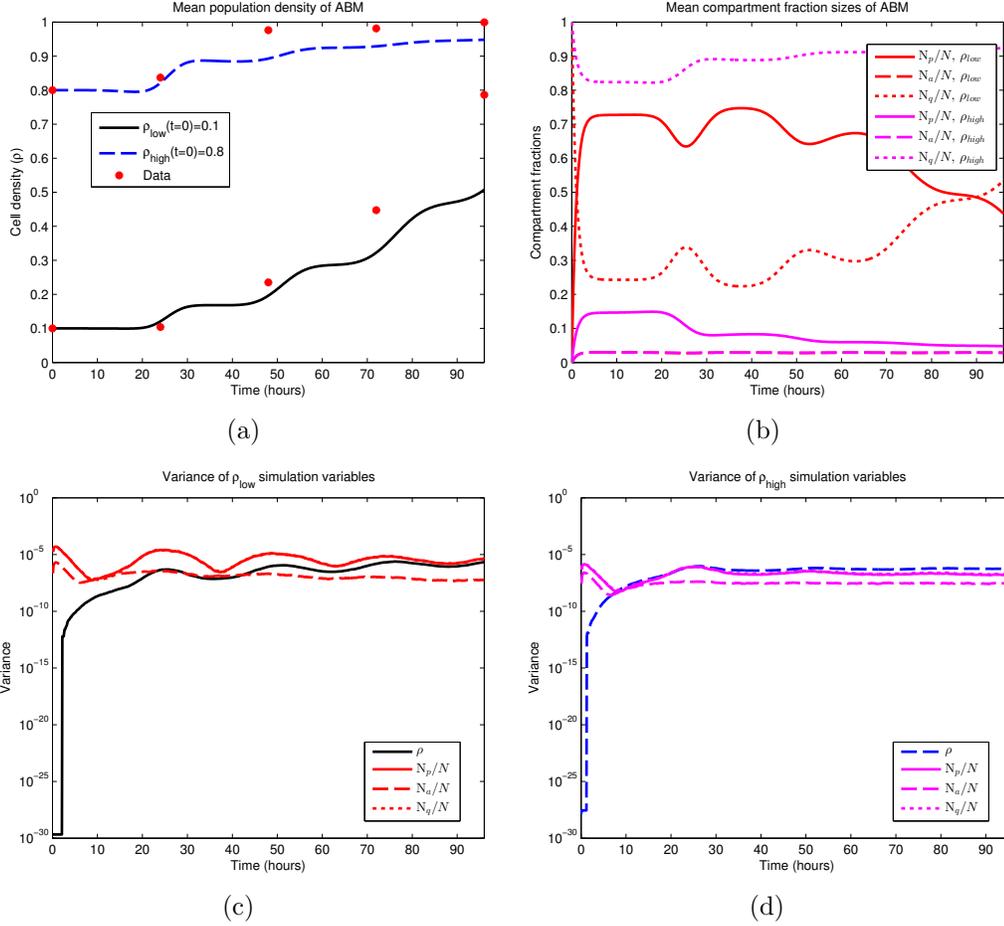


Figure 5.6: First and second moments of the ABM. Shown are the temporal dynamics of the mean and variance of 1000 Monte Carlo simulations. β and d appear as the red curves in Figures 5.2(a) and 5.2(b) respectively, and $\sigma = 3$ hours, $c = 1$ cell per hour, $\gamma = 0.70$. Note that these parameters do not yield an optimal fit. See equations (5.51) and (5.52) for the two sets of initial conditions. In (b), the values of N_a for both sets of initial conditions essentially overlap, making them difficult to distinguish. The same is true for the variances of N_p and N_q in (c) and (d). (a) Sample means for the population density; (b) Sample means for the three cellular compartments; (c) Variances for simulations that satisfy $\rho(0) = 0.1$; (d) Variances for simulations that satisfy $\rho(0) = 0.8$.

5.4.3 Stochastic and deterministic comparison

We investigate the relationship between realizations of the ABM and the IDE. Due to the large number of cells in the simulation, the small variances observed in Fig-

ures 5.6(c) and 5.6(d) imply that individual ABM simulations can be approximated well by their mean value, and the derivation in Section 5.3.2 (together with Figure 5.3) further implies that the mean value can be approximated by the solution of the system of IDEs appearing in Section 5.3. Here, we quantify these approximations. Data from the ABM simulations appearing in Section 5.4.2 is compared to the numerical solution of the IDE system (5.18)-(5.20) with corresponding parameter values and initial conditions. The results appear in Figure 5.7. The system of IDEs is solved using a four-step explicit Adams-Bashforth method. For more details, see Section 5.3.3.

Figure 5.7(a) demonstrates the growth of the global cell density ρ for both the low and high density seedings. All 1000 ABM realizations are plotted (blue), together with the IDE solution (black). Here, we observe both the small variance in the ABM simulations and the accuracy of the IDE approximation. Measuring the error $\mathcal{E}_{\text{IDE,mean}}$ of the IDE to the ABM mean via the maximum of the distances induced by the supremum norm $\|\cdot\|_\infty$ for each initial condition, we obtain

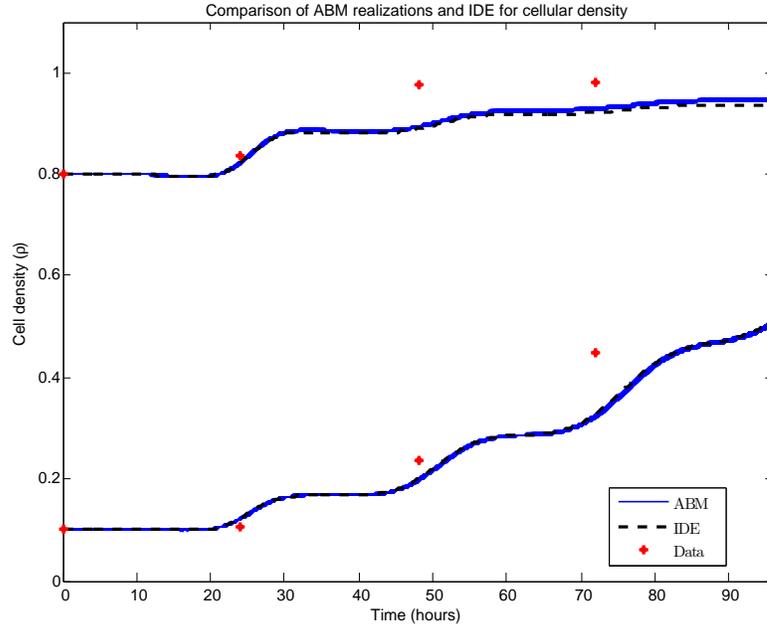
$$\begin{aligned} \mathcal{E}_{\text{IDE,mean}} &:= \max \left\{ \|\rho_{\text{IDE}}^{0.1} - \rho_{\text{ABM,mean}}^{0.1}\|_\infty, \|\rho_{\text{IDE}}^{0.8} - \rho_{\text{ABM,mean}}^{0.8}\|_\infty \right\} \\ &\approx \max \{0.0051, 0.0108\} = 0.0108. \end{aligned} \tag{5.53}$$

As observed in Figures 5.6(c) and 5.6(d), ρ has a variance on the order of 10^{-6} , and hence a standard deviation on the order of 10^{-3} . By Chebyshev's inequality [177], 93.75% of the realizations lie within four standard deviations of the mean; thus the

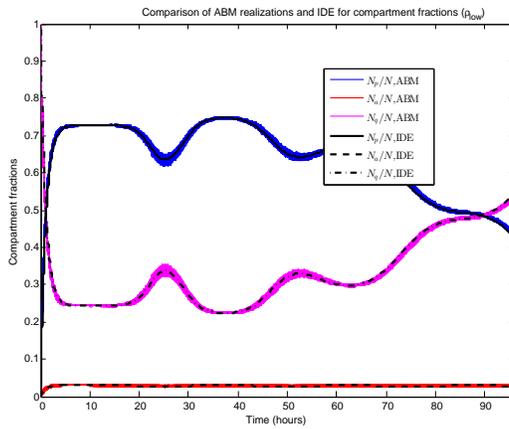
incurred error between the IDE and such a realization can be bounded by

$$\begin{aligned}
|\rho_{\text{ABM}}(t) - \rho_{\text{IDE}}(t)| &\leq \|\rho_{\text{ABM}} - \rho_{\text{IDE}}\|_{\infty} \\
&\leq \|\rho_{\text{ABM}} - \rho_{\text{ABM,mean}}\|_{\infty} + \|\rho_{\text{ABM,mean}} - \rho_{\text{IDE}}\|_{\infty} \quad (5.54) \\
&\leq 4\sigma + \mathcal{E}_{\text{IDE,mean}} \leq 4 \cdot 10^{-3} + 0.0108 = 0.0148.
\end{aligned}$$

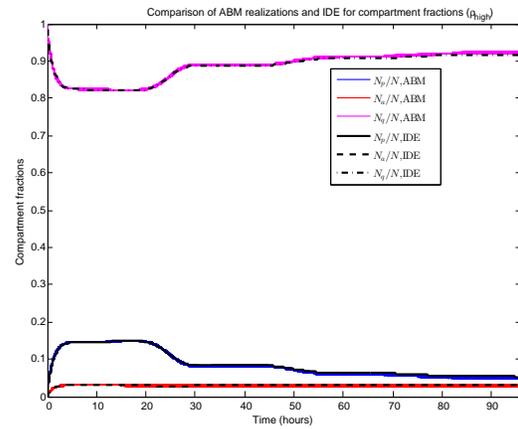
Furthermore, this is a conservative estimate, as the sample expectation of the LHS of (5.54) at any time t is computed to be 0.003. Similar conclusions are drawn in regards to the compartment populations appearing in Figures 5.7(b)-5.7(c). These calculations justify using the deterministic IDE system in the remainder of the work when performing numerical experiments and determining parameter values. We lastly note that for simulations involving small populations, a much higher variance is observed, and the IDE is no longer a valid approximation to ABM realizations. This was investigated by varying the plate size K in both the ABM and IDE simulations. We found that a small number of cells (e.g. $K \approx 100$) ensures a small error between the IDE and the expected value of the ABM. More precisely, we observe the L^{∞} norm (with respect to time) of the error to be bounded above by 0.0216. However, to guarantee that the expected value accurately approximates individual realizations of a stochastic experiment, a larger number of cells is required. For simulations with $K = 100$, we calculated a relative standard deviation of the ABM simulations to be on the order of 0.14, while for $K = 1000$, this value decreases to 0.035. Thus we conclude that this system requires K to be on the order of 1000 cells for the IDE to accurately approximate individual ABM realizations.



(a)



(b)



(c)

Figure 5.7: Comparison of ABM and IDE simulations. 1000 Monte Carlo realizations are plotted together with the corresponding IDE. β and d appear as the red curves in Figures 5.2(a) and 5.2(b) respectively, and $\sigma = 3$ hours, $c = 1$ cell per hour, $\gamma = 0.70$. (a) Cellular density for both sets of initial conditions ($\rho(0) = 0.1$ and $\rho(0) = 0.8$); (b) Compartment fractions for $\rho(0) = 0.1$; (c) Compartment fractions for $\rho(0) = 0.8$.

5.4.4 Parameter estimation and variation

To calculate parameter values that model the measured cellular growth dynamics (Figure 5.4), we minimized the distance between the experimental data and the

deterministic IDE. Denoting the vector of parameters $p := (c, \gamma, d, \beta_m, \rho_m, \epsilon, \sigma)$ and defining

$$E_\rho(t; p) := \rho_{\text{IDE}}(t; p) - \rho_{\text{data}}(t), \quad (5.55)$$

we used MATLAB's nonlinear least-squares solver "lsqnonlin.m" to minimize the ℓ^2 norm of (5.55) at all measured times:

$$\left(E_\rho(0; p), E_\rho(24; p), \dots, E_\rho(96; p) \right) \quad (5.56)$$

Here t is measured in hours. The minimization of (5.56) is performed over the parameter set p , using a constrained trust region algorithm. The constraints were generated both from biological interpretations and also experimental data; they appear below in Table 5.1, together with their justification. As the routine locates only local minima, 100 random initial parameter sets were used to seed the algorithm, yielding 100 (non-unique) local minimizing parameter sets. The results of this computation appear partially in Figure 5.8(a)-5.8(b).

Table with parameters for optimization (and ranges)

Figure 5.8(a) displays the low and high density cultures, together with the IDE simulations corresponding to four local minimizing parameter sets calculated in the numerical experiment; Figure 5.8(b) plots the corresponding equilibrium division fraction distributions $\beta(\rho)$. The four sets are selected as minima of the local minima which display qualitatively different behavior while producing an ℓ^2 norm of (5.56)

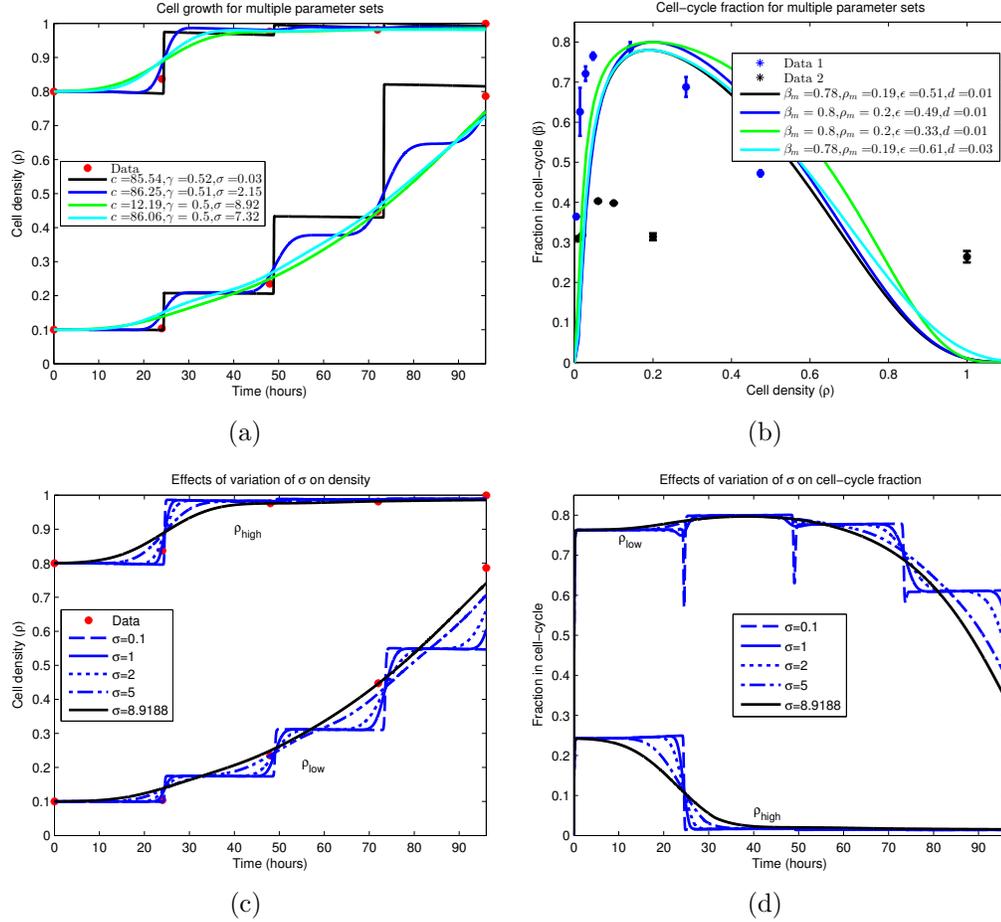


Figure 5.8: Cellular growth dynamics for varying parameter sets. (a) A nonlinear least-squares algorithm was used to locate parameter values which minimize the difference between the experimental data and the IDE (equations (5.18)-(5.20), together with (5.9)-(5.10)). Due to the sparsity of the data (red), multiple local minimums were found. We note that all curves plotted have residuals with ℓ^2 norm $\in [0.0583, 0.0980]$, and are displayed in descending order by this measure. Similar plot styles in (a) and (b) correspond to the same parameter set. (a) Growth of cellular density in time; (b) Division (β) and death fractions ($d = \beta(1)$) for the fitted parameter sets. (c)-(d) Dependence of IDE solutions on cell-cycle standard deviation σ . σ is varied, while all other parameters are fixed from a local minimizing set located in (a) and (b). We plot the global density growth and cell-cycle fraction for each case. (c) Cellular density ρ for σ variation; (d) Proliferative compartment for σ variation.

in $[0.0583, 0.0980]$. Hence we conclude that no global minimum can be determined, as disparate parameter sets produce equally valid solutions. Indeed, as there are only 8 true data points (initial conditions are constrained) and 7 parameters, uniqueness

Parameter	Min. Value	Max. Value	Comments
c	0	∞	Rate of transition is always non-negative
γ	0.5	1	See Section 5.2.2
d	0.01	0.05	See Section 5.2.1
β_m	0.3	0.8	See Figure 5.2(a)
ρ_m	0.05	0.2	See Figure 5.2(a)
ϵ	0	∞	Ensures $\beta(\rho)$ has a relative maximum
σ	0	10	Lower bound: standard deviation is non-negative Upper bound: essentially excludes possibility of negative cell-cycle length

Table 5.1: List of minimization parameters and ranges

should not be expected.

Nevertheless, we can still draw biological and experimental conclusions from the results. We observe that the most important parameter which dictates the qualitative structure of the population growth dynamics is the standard deviation of the cell-cycle length σ . For $\sigma = 0.03$ hours, the dynamics are essentially piecewise-constant, with steps of growth occurring at multiples of the mean cycle length $\mu = 24.4416$ hours. As σ increases, these steps become smoother, with $\sigma = 8.92$ hours showing no visible step structure. See Figures 5.8(c)-5.8(d) for further elucidation of the σ dependence. Biologically this dependence on σ is intuitive, since as the population becomes more homogeneous (σ decreases), the overall structure should increase, resulting in more regular growth patterns. Lastly, we note that $\sigma = 0.03$ hours seems biologically improbable, as the density curve exhibits a large sensitivity to the measurement time. Indeed, for cells governed by such a σ , differing the measurement time by less than half an hour could yield an approximate 100% change in the measurement value, which is an unrealistic observation.

Other parameter values were determined with more certainty. Examining Figure 5.8(b), we observe that the optimal parameter sets yield a maximal dividing fraction of $\beta_m \in [0.78, 0.80]$ at a density of $\rho_m \in [0.19, 0.20]$, both lying in small upper regions of their constraints (see Table 5.1). $\gamma \in [0.50, 0.52]$ and $d \in [0.01, 0.03]$ consistently, with $d \approx 0.01$ in most cases. The rate of cellular transfer c and the division fraction parameter ϵ are more varied, and we investigated the dependence of the model on these parameters; the results appear partially in Figures 5.9 and 5.10. Here we vary only the parameter of interest, while leaving the others fixed. In Figure 5.9(a), we see minimal variation in the cellular growth as c varies over a large range. In particular, we observe that as long as c is not sufficiently small ($c > 1$ cell per hour), the overall growth dynamics do not vary substantially, and that c 's principal effect is on the rate which cells migrate into the division compartment P (Figure 5.9(c)). Note that for larger c , the earlier transitions into P occur; however after this initial variation, all parameterizations are relatively synchronized. Contrast this with Figure 5.10, where the ϵ dependence is examined. We observe a large variation in the dynamics for small variations in ϵ , especially in regards to high density growth.

5.5 Conclusion

Understanding the mechanisms of cellular growth is an essential step in studying cancer progression and the evolution of drug resistance; thus, it is the subject of the current work. Our main hypothesis is that variations in both the cell-cycle

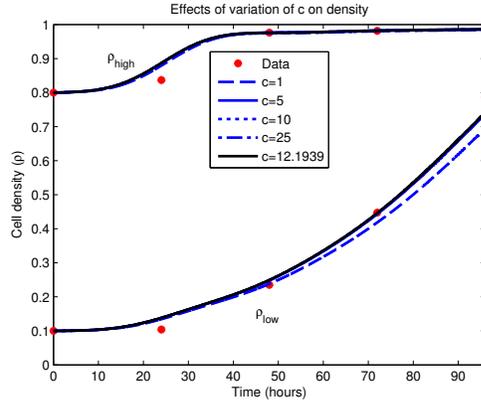
and the apoptotic lengths are central contributing factors to the overall dynamics. We postulate that intrinsic heterogeneity, in the form of distributed cell-cycle and apoptotic process lengths, are fundamental aspects of cell growth, and cannot be ignored. Here, we introduced an ABM, and demonstrated the best approximation of the model to the data. An ABM permits a straightforward procedure for introducing an age-structure on a population, while simultaneously avoiding the technical details of macroscopic limits and population sizes, and hence continuum, age-structured models. Furthermore, our understanding of the biology is cell-driven, so it is natural to specify behavior on the cellular level, as opposed to the population as a whole. Such models are also quite flexible and can take into account complex cell-cell and cell-environment interactions in a straightforward manner, which will be the subject of future work.

Formulating an ABM requires information of a cell behavior as a function of the environment. In our model it requires specific transition rates based on a given environmental condition. Our experimental results yield information on the distribution of cells in the cell-cycle and apoptosis process at various global densities, but do not include explicit information on the desired rates. A novel feature of this work is the manner in which the experimental results were incorporated into the ABM, translating the experimental distributions into probabilistic transitions on the cellular level.

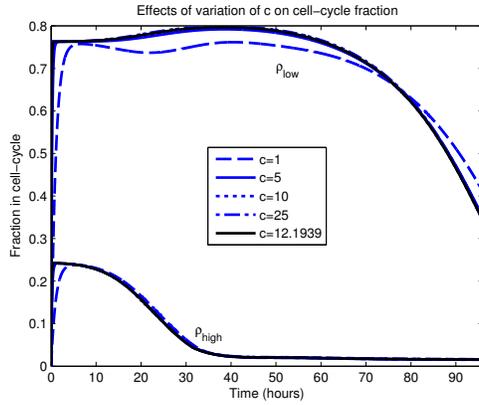
It is a difficult task to fit a stochastic model, such as an ABM, to biological data, as different realizations produce different results. However, since the population size is relatively high, we observed that the overall variance of all random

variables is small. Thus, the model can be viewed as essentially deterministic, with realized values that are given by their corresponding means. Hence, a system of IDEs which describe the expected values of the dynamic compartment populations was derived. Using this IDE system, we performed standard optimization on these deterministic variables to calculate the model parameters in the least-squared sense. Our results suggest that the IDE accurately describes the dynamics of the ABM, and both models emulate the experimental data for a large range of parameters.

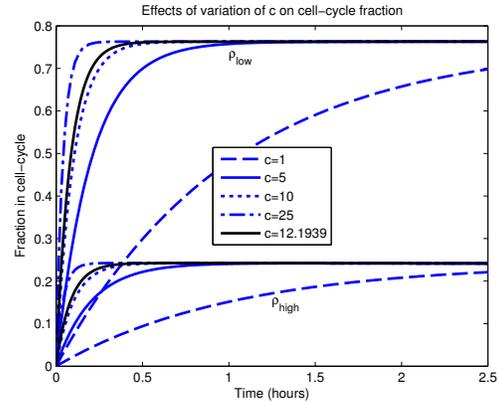
The application of our framework to cancer research is expected to lead to the development of systematic methods for determining treatment strategies, where cellular behavior is governed by random, as opposed to deterministic, events. Furthermore, an expanded model that includes drug effects will allow us to revisit the mechanism of multi-drug resistance by studying the spatiotemporal mechanisms of intratumoral heterogeneity. Intratumoral heterogeneity includes many tumor sub-populations with different cellular dynamic characteristics, such as proliferation rates. Each sub-population may include intrinsic and/or induced heterogeneity mechanisms. Our work may be expanded to incorporate and study the differences between induced vs. intrinsic heterogeneity, and their effects on the resistance level, so that a global understanding and estimation of the total tumor evolution may be achieved.



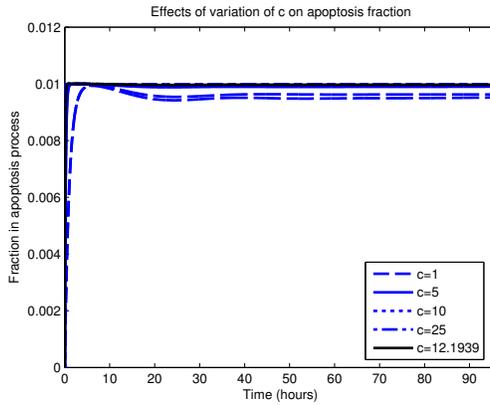
(a)



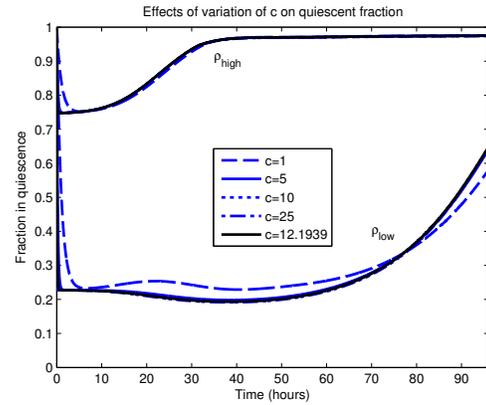
(b)



(c)



(d)



(e)

Figure 5.9: Dependence of IDE solution on c . All parameters, excluding c , are fixed and obtained from an optimization appearing in Figure 5.8 (green curves). The parameter values are $\beta_m = 0.8$, $\rho_m = 0.2$, $\epsilon = 0.33$, $d = 0.01$, $\gamma = 0.5$ and $\sigma = 8.92$ hours. The black curve represents the local minimizer. (a) Cellular density ρ ; (b) Proliferative compartment; (c) Magnification of $t = 0$ dynamics of proliferative compartment; (d) Apoptotic compartment; (e) Quiescent compartment.

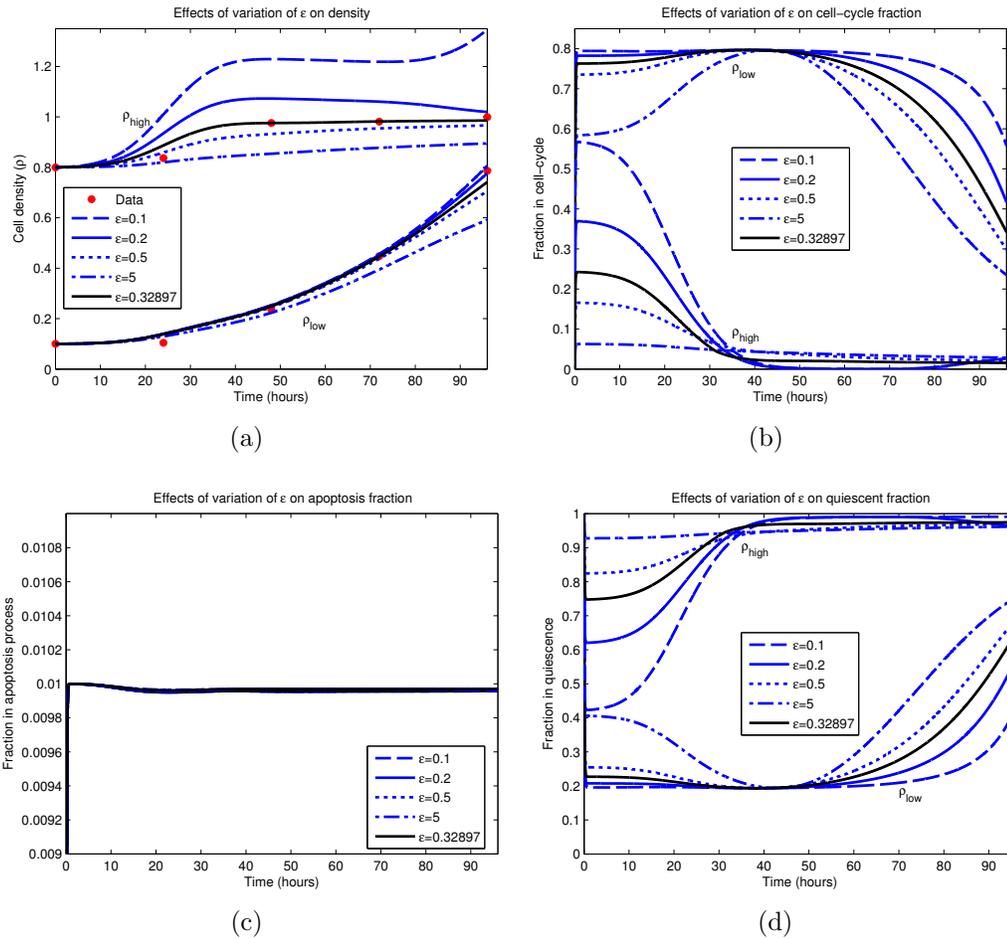


Figure 5.10: Dependence of IDE solution on ϵ . All parameters, excluding ϵ , are fixed and obtained from an optimization appearing in Figure 5.8 (green curves). The parameter values are $c = 12.19$ cells per hour, $\beta_m = 0.8$, $\rho_m = 0.2$, $d = 0.01$, $\gamma = 0.5$ and $\sigma = 8.92$ hours. The black curve represents the local minimizer. (a) Cellular density ρ ; (b) Proliferative compartment; (c) Apoptotic compartment; (d) Quiescent compartment.

Chapter 6: The cell-cycle and drug resistance - a spatial mechanism

In this chapter, we present preliminary work on an extension of the model introduced in Chapter 5. Namely, we propose a spatial mechanism to describe the variation in the cell-cycle length which was taken as an intrinsic distribution in the previous ABM. We show that local density considerations can describe the observed variation in cell-cycle length, and also introduce a biologically realistic mechanism to describe the effect of treatment by a specific chemotherapeutic agent. Lastly, the model is used to qualitatively explain experimental data collected by Dr. Lavi at the NIH on the connection between cell density and drug resistance.

6.1 Introduction

In Chapter 5, we introduced a model of cancer cell growth which incorporated an intrinsic form of heterogeneity via the length of the cell-cycle and the apoptosis process. Specifically, the amount of time a cell spent in the proliferative compartment P was assumed to follow an *a priori* normal distribution with mean 24.4416 hours and variance σ^2 , *independent of any local measurements or other properties of the cell*. That is, all cells experienced the same “force” of heterogeneity, with explicit variation modeled independently of a cell’s environment. Using such a dis-

tribution, we were able to determine the effect on the growth dynamics of the model parameters, as well as emulate experimental data. We also note that the model was non-spatial (i.e. assumed a homogenous population), and thus all density dependencies in the transition rates were based on global quantities. We primarily focused on the cell-cycle, but similar methodology was applied to the apoptosis process, with [38] providing the distributional data.

In this chapter, we propose a spatial mechanism to address the observed variance in the cell-cycle. We assume that a cell has each local environment, in which its local density can be evaluated. Furthermore, we no longer assume a randomly generated cell-cycle length upon entering the proliferating compartment P, but rather a “cell-cycle age,” which varies its rate of progression according to the cell’s local density. Since one of our main objectives is to study the effect of chemotherapy, we also incorporate a mechanistic effect of a specific chemotherapy into the modeling framework. This effect models the action of the chemotherapeutic agent Paclitaxel, the agent used in the experiments of Dr. Lavi at the NIH.

As discussed in Chapter 1, the cell-cycle is a fundamental mechanism of intrinsic drug resistance. It is observed to be stochastic, yet tightly regulated in nature, and is present in both healthy and cancerous cells. Furthermore, as many chemotherapeutic drugs target the cell-cycle, understanding the cell-cycle dynamics can provide invaluable insight on this form of intrinsic drug resistance, as well as guidance for designing optimal treatment strategies. We hypothesize that this stochastic nature as well as the emerging drug resistance can be explained by the presence of local variations in cell density.

This chapter is organized as follows. In Section 6.2, we present an extension of the ABM from Section 5.2, which incorporates spatial variables, local density, and the effect of chemotherapy. In Section 6.3, we study the prediction of the model and apply it to experimental data. In particular, we determine a suitable range for some of the parameters, study the distribution of the cell-cycle, and show a qualitative agreement between experimental and theoretical survival curves as a function of dosage. Concluding remarks are presented in Section 6.4.

6.2 A mathematical model of cancer cell dynamics

In this section, we describe the mathematical/computational model of cancer cell growth in a two-dimensional spatial environment. The model is formulated as an agent-based model (ABM), which incorporates spatial aggregation/repulsion forces, as well as transitions through three cellular states: proliferation (P), apoptosis (A), and quiescence (Q). Also incorporated are the effects of a chemotherapeutic agent, which acts only on cells in the proliferative compartment P. Below we describe each component of the model in detail.

6.2.1 Motility

We assume that cell movement is governed by a stochastic differential equation (SDE), with deterministic components modeling aggregation and repulsion, and a random term modeled by a standard Wiener process. We note that the random effects are included to account for the experimentally observed random walk-like

motion exhibited by cells in low density regimes, and that the movement is not constrained to a lattice. This framework is chosen to account for the observed phenomena of cellular aggregation, and is adapted from the works of Morale *et al.* [178–180]. Letting $X_k(t) \in \mathbb{R}^2$ denote the position of the k th particle ($k \in \mathbb{N}$), the motion is governed by

$$dX_k(t) = F_a[\tilde{X}_k(t)](X_k(t))dt + F_r[\tilde{X}_k(t)](X_k(t))dt + \sigma dW_k(t), \quad (6.1)$$

where $\sigma \geq 0$ is the mean free path, and $\tilde{X}_k(t)$ is the local empirical measure (i.e. the local spatial distribution of particles at time t):

$$\tilde{X}_k(t) = \frac{1}{N_k(t)} \sum_{k=1}^{N_k(t)} \epsilon_{X_k(t)}. \quad (6.2)$$

Here $\epsilon_{X_k(t)}$ is the indicator measure for the point mass $X_k(t)$, defined by

$$\epsilon_{X_k(t)}(B) = \begin{cases} 1 & \text{if } X_k(t) \in B, \\ 0 & \text{if } X_k(t) \notin B, \end{cases} \quad (6.3)$$

$\forall B \in \mathcal{B}(\mathbb{R}^2)$, and $N_k(t)$ is the number of cells which the k th particle can detect at time t . Thus the total force is obtained through superposition of pairwise interactions. Note that this differs from most other works, in that we use a *local*, as opposed to *global*, probability measure for the attraction and repulsion forces. That is, all quantities are based on local measurements, so that cells have no global information on the total number of cells $N(t)$; rather, only cells in a local neighborhood, with

radius ϵ_a , affect the cellular dynamics.

The attraction and repulsion forces (F_a and F_r , respectively), are defined as gradient operators with opposite signs:

$$F_a[\tilde{X}_k(t)](X_k(t)) = [\nabla G_a * \tilde{X}_k(t)](X_k(t)), \quad (6.4)$$

$$F_r[\tilde{X}_k(t)](X_k(t)) = -[\nabla G_r * \tilde{X}_k(t)](X_k(t)), \quad (6.5)$$

where G_a and G_r are even real-valued functions on \mathbb{R}^2 that exhibit a global maximum at 0, with different compact supports. Throughout this work, G_a and G_r will be referred to as *potentials*. As discussed above, the support of G_a is the radius ϵ_a , the local neighborhood in which a cell can detect its neighbors. G_r represents the net repulsive interaction between two cells, which acts to prevent the cellular mass centers from coalescing. Thus, this force models volume exclusion, and its support corresponds to an approximate cellular diameter. This value was measured experimentally, and was chosen throughout to coincide with the average value of the measurements, relative to the plate size. We non-dimensionalized the ABM spatial units so that $\epsilon_r = 1$. See Figure 6.1 for a plot of G_a (with $\epsilon_a = 5$) and G_r . Finally, we note that forces are not uniform throughout the attraction and repulsion detection radii, but rather depend on the distance between any two particles. This is readily observed via the gradients of the potentials shown in Figure 6.1.

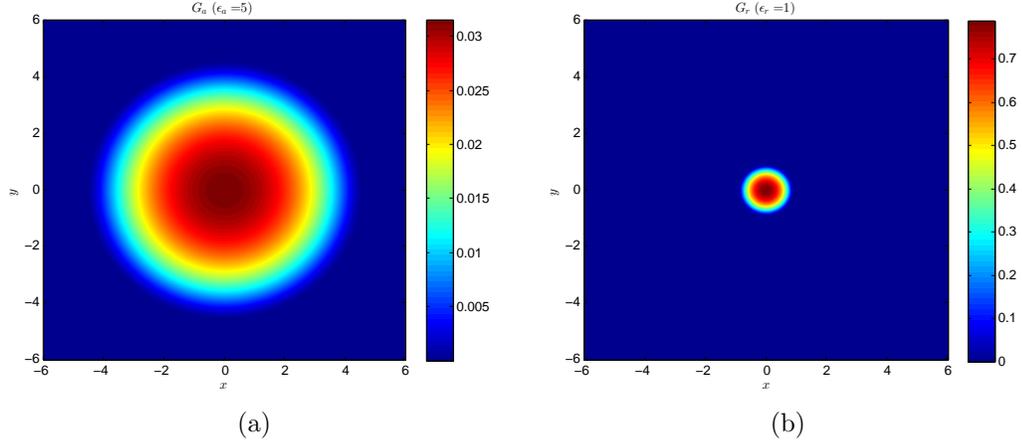


Figure 6.1: Attraction and repulsion potentials used in the SDE model (6.1), via equations (6.4) and (6.5). (a) Attraction force potential, with local detection radius $\epsilon_a = 5$. (b) The repulsion potential used throughout this chapter.

6.2.2 Chemotherapy

In this chapter, we consider a drug that is uniformly distributed throughout the two-dimensional environment, and is in equilibrium. Thus, the drug concentration is assumed constant, $c(t) \equiv c$. Furthermore, the treatment used in the experimental setup is Paclitaxel (Taxol, Abraxane), which is a mitotic inhibitor. More precisely, its primary chemotherapeutic mechanism acts as a cytoskeletal drug which targets tubulin [181]. During cell division, chromosomal microtubules are necessarily destabilized and broken down as the mother cell divides. Paclitaxel acts as a stabilizer for these microtubules, thus preventing metaphase spindle configuration. The cell, unable to complete mitosis, remains at the mitotic checkpoint of the cell-cycle for a prolonged period of time. Eventually, due to the normal signalling pathways of the cell-cycle, the prolonged activation of the mitotic checkpoint triggers apoptosis, with no successful cell division. Thus, for our model formulation, we assume that

the drug does not result with a direct cell death, but rather acts to increase the cell-cycle length (or equivalently, to slow the progression throughout the cell-cycle). Furthermore, we model the activity of the mitotic checkpoint by assuming an upper bound on the length of time that can be spent in P, after which the cell commits apoptosis by transferring to the A compartment. Note that this implies that the drug only effects dividing cells. See Section 6.2.3.2 below for details.

6.2.3 Compartments in the ABM model

We next describe the compartments (or states) in which a cell may reside. We also discuss the transitions occurring between these states, which, in general, depend on local measurements and specific cellular variables. In general, the transition rates are either *explicit* and *implicit*. The explicit rate transitions are interpreted as probabilities per unit time, i.e., continuous time Markov chain transition rates, while the implicit rates depend on state variables specific to individual cells. The states diagram for the model is shown in Figure 6.2.

6.2.3.1 Quiescence (Q)

Cells in quiescence are neither dividing nor dieing. Instead they are in a resting phase. Often this state is referred to as G_0 , and is the compartment where all dividing cells originate. Cells enter the proliferating compartment P from Q based

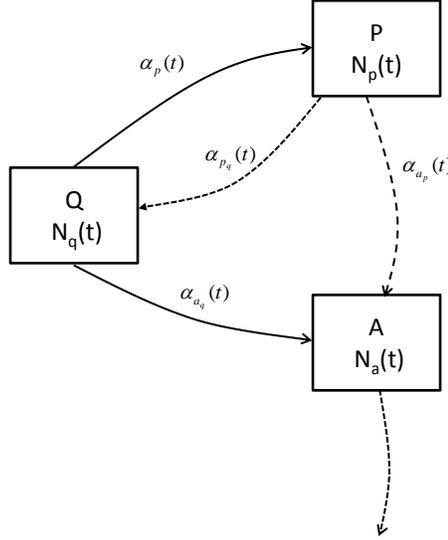


Figure 6.2: Schematic representation of division/apoptotic/quiescent process. The model includes three compartments: proliferating cells (P), quiescent cells (Q), and apoptotic cells (A). Cells move stochastically through the three phases, with explicit rates represented by solid lines, and implicit rates given by dashed lines.

on local density measurements, ρ_{loc} , which is defined for each cell as

$$\begin{aligned} \rho_{\text{loc}} &= \frac{\text{total area of all cells in ball of radius } \epsilon_a}{\text{Area}(B_{\epsilon_a})} = \frac{\sum_{i=1}^{N_{B_{\epsilon_a}}} (\pi \epsilon_r^2)}{\pi \epsilon_a^2} \\ &= \frac{N_{B_{\epsilon_a}} \cdot \epsilon_r^2}{\epsilon_a^2}. \end{aligned} \quad (6.6)$$

Here ϵ_a is as defined in Section 6.2.1, B_{ϵ_a} is the ball of radius ϵ_a centered at the cell's location, and $N_{B_{\epsilon_a}}$ is the number of cells in B_{ϵ_a} . Cells are considered hard-spheres, so that each has area $\pi \epsilon_r^2$. We normalized the simulation space so that $\epsilon_r = 1$, and the count in the numerator of (6.6) only considers center-of-mass detection, so that no partial areas are possible.

Transitions into the cell-cycle are explicit, and thus are determined by prob-

abilistic rates. We model this rate α_p as a quadratic function in ρ_{loc} with finite support, to qualitatively agree with the global density in Chapter 5. Specifically,

$$\alpha_p(\rho_{\text{loc}}) = \begin{cases} -\frac{\alpha_{\text{max}}}{(\rho_m - \rho_r)^2}(\rho_{\text{loc}} + \rho_r - 2\rho_m)(\rho_{\text{loc}} - \rho_r), & \text{if } 0 \leq \rho_{\text{loc}} \leq \rho_r, \\ 0, & \text{otherwise.} \end{cases} \quad (6.7)$$

Note that since $\rho_{\text{loc}} = \rho_{\text{loc}}(t)$, α_p is a function of time. ρ_r determines the value of ρ_{loc} above which no cell will enter division, and ρ_m is the local density value at which α_p obtains its global maximum α_{max} .

Analogously, we define the transition rate of natural death, α_{a_q} . This is an explicit probabilistic rate which, in principle, should be a function of a cell's local density ρ_{loc} . However, as discussed in Chapter 5, the experimental data does not seem to support a direct dependence on density. In fact, since this rate is assumed small (between one and five percent of the population at any time), any functional form for α_{a_q} would not alter the dynamics significantly. Thus, we assume

$$\alpha_{a_q} \equiv d, \quad (6.8)$$

where $d \in [0.01, 0.05]$. See Figure 6.3 for a plot of the Q transition rates for sample parameter values.

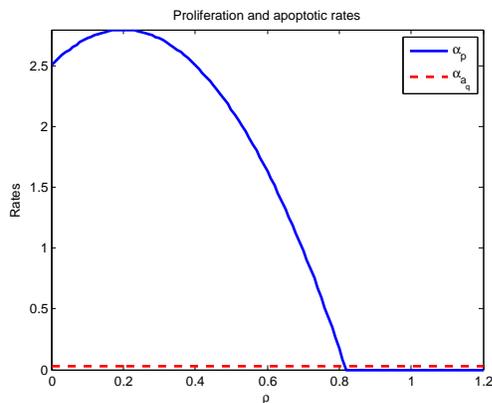


Figure 6.3: Transition rates α_p and α_{aq} , with parameters values $\alpha_{\max} = 2.8$, $\rho_m = 0.2$, $\rho_r = 0.85$, and $d = 0.03$.

6.2.3.2 Proliferation (P)

Cells that are in the proliferation state P are progressing through the cell-cycle. Once this cycle is complete, both mother and daughter cells transition to the quiescent compartment Q. This $P \rightarrow Q$ transition is represented by the transition function α_{pq} (see Figure 6.2), which is implicit in the sense that it depends on the time the compartment P was entered, and on the rate in which the cell moves through the cell-cycle. That is, it is non-Markovian. In this work, we assume that all cells have a base cell-cycle length μ , after which division occurs. However, due to local density constraints, proliferating cells do not move through the cell-cycle at a constant speed, but instead decelerate as their local density (ρ_{loc}) increases. Specifically, we assume a base length of $\mu = 15$ hours and a state variable age $a(t)$, which measures the cell's division progress in both absolute and relative terms. Here t is the amount of absolute time in which a cell has spent in P, which upon entering is normalized to 0 hours. $a(t)$, on the other hand, is the cell's relative age in the

cell-cycle, which increases via a progression rate $\omega(\rho_{\text{loc}}, c)$. Note that ω depends on the drug concentration c , which we discuss in Section 6.2.2 below. In the absence of any treatment (i.e. the control case), ω takes the form

$$\omega(\rho_{\text{loc}}, 0) = e^{\log(p)\rho_{\text{loc}}}. \quad (6.9)$$

Here, $p \in [0, 1]$ is a parameter which dictates the relative speed (compared to a maximum value of 1 when $\rho_{\text{loc}} = 0$) through the cell-cycle in the dense microenvironment $\rho_{\text{loc}} = 1$. The cell-cycle age $a(t)$ updates from absolute time t to $t + \Delta t$, via

$$a(t + \Delta t) = a(t) + \omega(\rho_{\text{loc}}(t + \Delta t), 0)\Delta t, \quad (6.10)$$

where we assume the local density measurements have previously been updated. The cell divides when t_* is such that $a(t_*) = \mu$, with the elapsed t_* measuring the amount of “real” time spent in the cell-cycle. Thus, the population of proliferating cells produces a distribution of measurable times spent in the cell-cycle. Furthermore, this distribution is determined by spatial configurations, and is a dynamic quantity. We note that the minimum possible observed cell-cycle length is μ hours, and we think of the maximum cell-cycle to be approximately μ/p , since a proliferating cell in a dense environment ($\rho_{\text{loc}} = 1$) will take μ/p hours to satisfy $a = \mu$.

To account for the drug, we follow Section 6.2.2 and assume that the drug acts to increase the absolute time spent in the cell-cycle. Using our age-based formu-

lation, this corresponds to decreasing ω as the drug concentration c increases. We assume that the local density, ρ_{loc} , and drug concentration, c , influence ω independently, and thus with (6.9) in mind, we define $\omega(\rho_{\text{loc}}, c)$ as

$$\omega(\rho_{\text{loc}}, c) = \exp\left(\log(p)\rho_{\text{loc}} - \omega_c \frac{c^2}{c_0^2 + c^2}\right). \quad (6.11)$$

Note that the dependence on c in (6.11) is a second-order Michaelis-Menten term, since the drug is experimentally observed to be inefficient at low dosages (see Figure 6.12 below). Here c_0 is the drug concentration at which its effect is half-saturated of the asymptotic value ω_c . To update the age $a(t)$ in the presence of an amount of drug c , $\omega(\rho_{\text{loc}}, c)$ replaces $\omega(\rho_{\text{loc}}, 0)$ in equation (6.10). A typical surface plot of $\omega(\rho_{\text{loc}}, c)$ is shown in Figure 6.4.

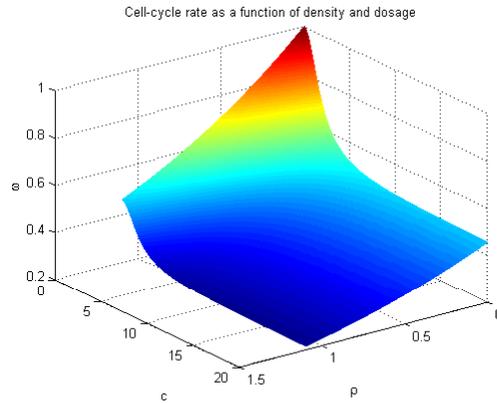


Figure 6.4: The cell-cycle rate (equation (6.11)) for parameter values $p = 0.5$, $c_0 = 2.29$, and $\omega_c = 0.7931$.

To completely specify the dynamics of the proliferation compartment, we must lastly implement the mitotic checkpoint discussed in Section 6.2.2, i.e., to determine the implicit rate α_{a_p} . Since the longest cell-cycle takes approximately μ/p hours to

complete ($\rho_{\text{loc}} = 1$ all throughout the process), we assume that once a cell has spent 95% of this time in P, it transitions to the apoptotic compartment A. That is, if t is such that $t \geq 0.95\mu/p$ and $a(t) < \mu$, the cell enters apoptosis. This is a reasonable assumption, since we expect increased cell death in very dense environments. The value of 0.95 is chosen to be slightly less than 1, since it is very unlikely that cells will enter division if $\rho_{\text{loc}} \geq 1$ (see Section 6.2.3.1). This transition is the dotted arrow originating from the compartment P in Figure 6.2. We also note that since death is observed at high values of treatment, independent of the density, a bound on ω_c can be obtained. In this c limit, the fastest speed in which a cell can move through the cell-cycle is given by $\omega(0, c \rightarrow \infty) = e^{-\omega_c}$. The time in which such a cell resides in P must still be greater than the apoptotic checkpoint $0.95\mu/p$, so that the apoptotic checkpoint is triggered. Thus, we must have

$$\frac{\mu}{\omega(0, c \rightarrow \infty)} \geq 0.95\frac{\mu}{p}, \quad (6.12)$$

$$\omega_c \geq \log\left(\frac{0.95}{p}\right). \quad (6.13)$$

6.2.3.3 Apoptosis (A)

Cells that are undergoing apoptosis are destined to complete cell death; that is, once a cell enters compartment A there are no transitions back to P or Q. Once completed, the cell is removed from the simulation (dotted arrow in Figure 6.2 originating from A). The amount of time spent in A is random, and is determined via the gamma distribution utilized in Chapter 5. That is, we do not study mechanisms which affect

the apoptotic process length, but instead assume that each cell selects a random length of time to spend in A, L_A , where

$$L_A \sim \Gamma(\omega, \lambda). \quad (6.14)$$

Following Chapter 5, we set $\omega = 4.9436$ and $\lambda = 0.19117$.

6.2.4 The ABM algorithm

We include a brief outline of the steps involved in the ABM. Suppose that at time t we have $N(t)$ cells, each residing in one of the three compartments Q, P, or A. Each cell in P has two unique state variables, \tilde{t} (cell-cycle absolute time) and $a(\tilde{t})$ (cell-cycle age). We update the system to time $t + \Delta t$ as follows:

1. Move all cells simultaneously via equation (6.2).
2. Update local densities to $\rho_{\text{loc}}(t + \Delta t)$ using (6.6).
3. Choose a uniformly random order \mathcal{O} of the cells.
4. Update the state of each cell according to \mathcal{O} :
 - a. If the cell is in compartment Q, do one of the following:
 - i. Enter P with probability $\alpha_p(t)\Delta t$.
 - ii. Enter A with probability $\alpha_{a_q}(t)\Delta t$ and choose L_A via equation (5.2).
 - iii. Remain in Q with probability $1 - \alpha_p(t)\Delta t - \alpha_{a_q}(t)\Delta t$.

- b. If the cell is in compartment P, update the cell-cycle time from \tilde{t} to $\tilde{t} + \Delta t$, and the cell-cycle age according to (6.10).
 - i. If $\tilde{t} \geq 0.95\mu/p$, enter A and choose L_A via equation (5.2).
 - ii. If $a(\tilde{t}) \geq \mu$, cell division successfully occurs, and both mother and daughter cell are transferred to compartment Q. Their spatial locations are chosen from a uniform distribution on the unit ball centered at the mother cell's previous position.
 - iii. If $a(\tilde{t}) < \mu$, the mother cell remains in P, and continues the cell-cycle.
- c. If the cell is in compartment A, do one of the following:
 - i. If the elapsed time of the apoptosis-cycle is L_A , then the cell is removed from the simulation.
 - ii. If the elapsed time of the apoptosis-cycle is less than L_A , then the apoptosis-cycle age increases by Δt , and the cell remains in A.

6.3 Results

6.3.1 Control parameter estimation

Our model has ten parameters (see Table 6.1 for reference). In this section, we derive bounds for the model parameters in the case of the control experiment (i.e. no drug). As discussed in Section 6.2.3.1, the d parameter has a narrow range, and thus we fix $d = 0.03$.

We first use experimental data to obtain bounds on the parameters determin-

Variable	Biological Interpretation
ϵ_r	Normalized cell radius
ϵ_a	Radius of cell microenvironment
p	Relative cell-cycle speed when $\rho_{\text{loc}} = 1$
α_{max}	Global maximum of α_p
ρ_m	Density at which $\alpha_p = \alpha_{\text{max}}$
ρ_r	Density above which $\alpha_p = 0$
d	Constant value of α_{a_q}
μ	Minimum cell-cycle length
c_0	Drug concentration at which cell-cycle rate ω is half saturated
ω_c	Asymptotic value of ω

Table 6.1: List of variables for spatial ABM algorithm

ing cellular movement, ϵ_a and θ . The experimental data contains two sets of initial and final data cellular configurations over a period of 24 hours. Using the initial configurations as input to the ABM, we simulate the model for a range of ϵ_a and θ parameters. Final simulation configurations were compared with those of the experiment, and for each value of θ , a minimizing ϵ_a was calculated. Configurations were compared using the ℓ^1 error, after the experimental image data was both adjusted in size and mapped to a binary image. A period of one day was used to reduce the number of completed cell divisions, so as to initially analyze the motion dynamics only. See Figure 6.5 for sample initial and final configurations of the model, and Figure 6.6 for the corresponding experimental data images.

Results of this parameter search are shown in Figure 6.7. We observe from both plots that the error is small only when $\sigma \leq 0.25$, so we use this upper bound on the random diffusion coefficient in (6.1). Furthermore, it is also clear that the absolute minimums do not occur for the same values of ϵ_a and θ . Given the stochastic nature of a single experiment, as well as the uncertainty in the initial conditions

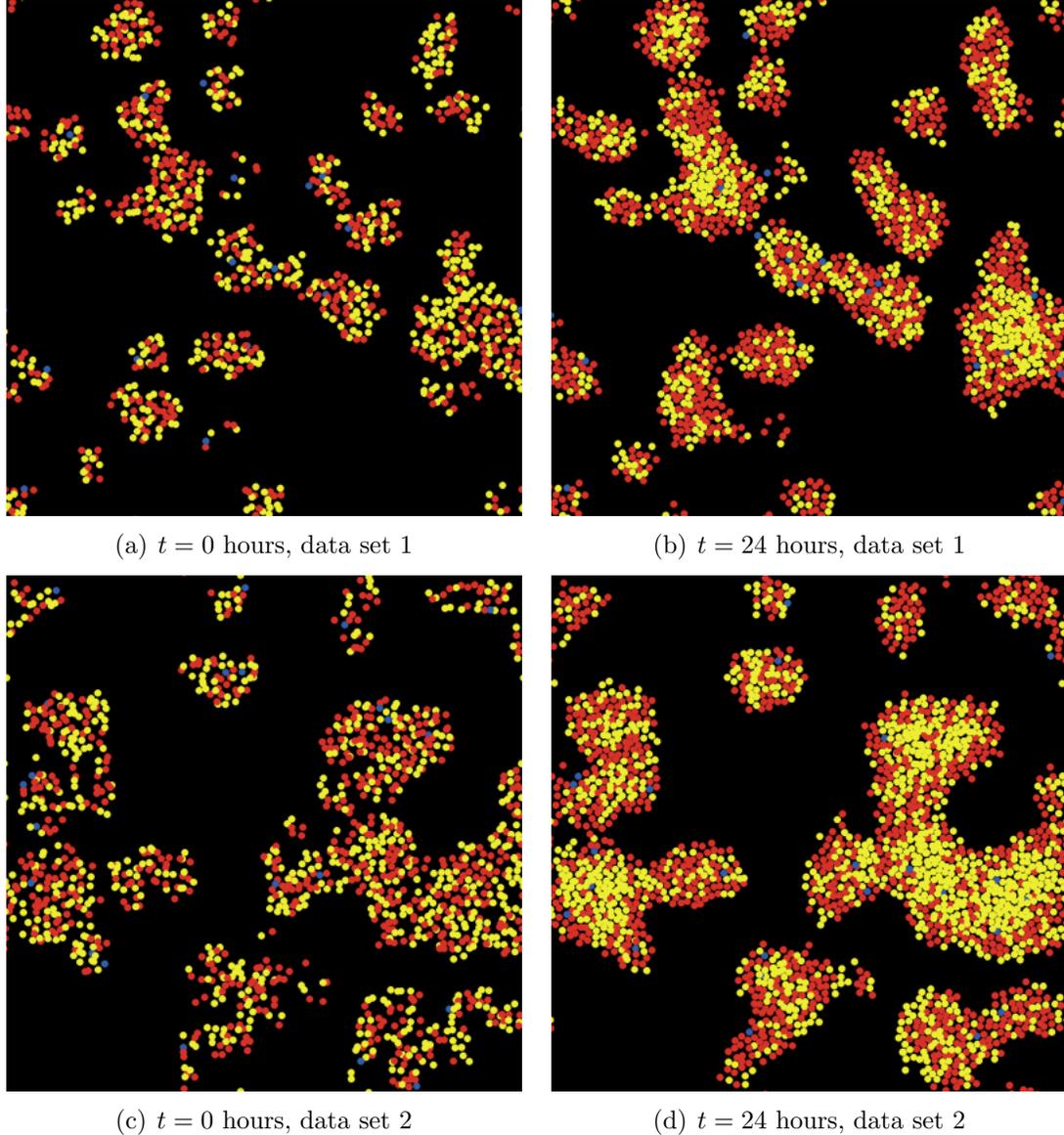


Figure 6.5: Sample initial and final configurations of the ABM used for comparison with 24 hour image control data. (a) and (c) are derived from the experimental data at $t = 0$ hours (Figures 6.6(a) and 6.6(c)), while (b) and (d) are calculated from the ABM after a period of 24 hours, and are compared to corresponding experimental data (Figures 6.6(b) and 6.6(d)). (b) corresponds to the initial conditions (a), while (d) corresponds to the initial conditions (c). The color scheme represents the three cellular compartments: yellow denotes quiescence, red denotes proliferation, and blue denotes apoptosis. $\epsilon_a = 5, \theta = 0.2, p = 0.5$, and all other parameters appear as in Figure 6.3.

(for example, in the relative fraction of cells in the different compartments, or in the cell-cycle synchronization), this inconsistency is expected. Furthermore, as our

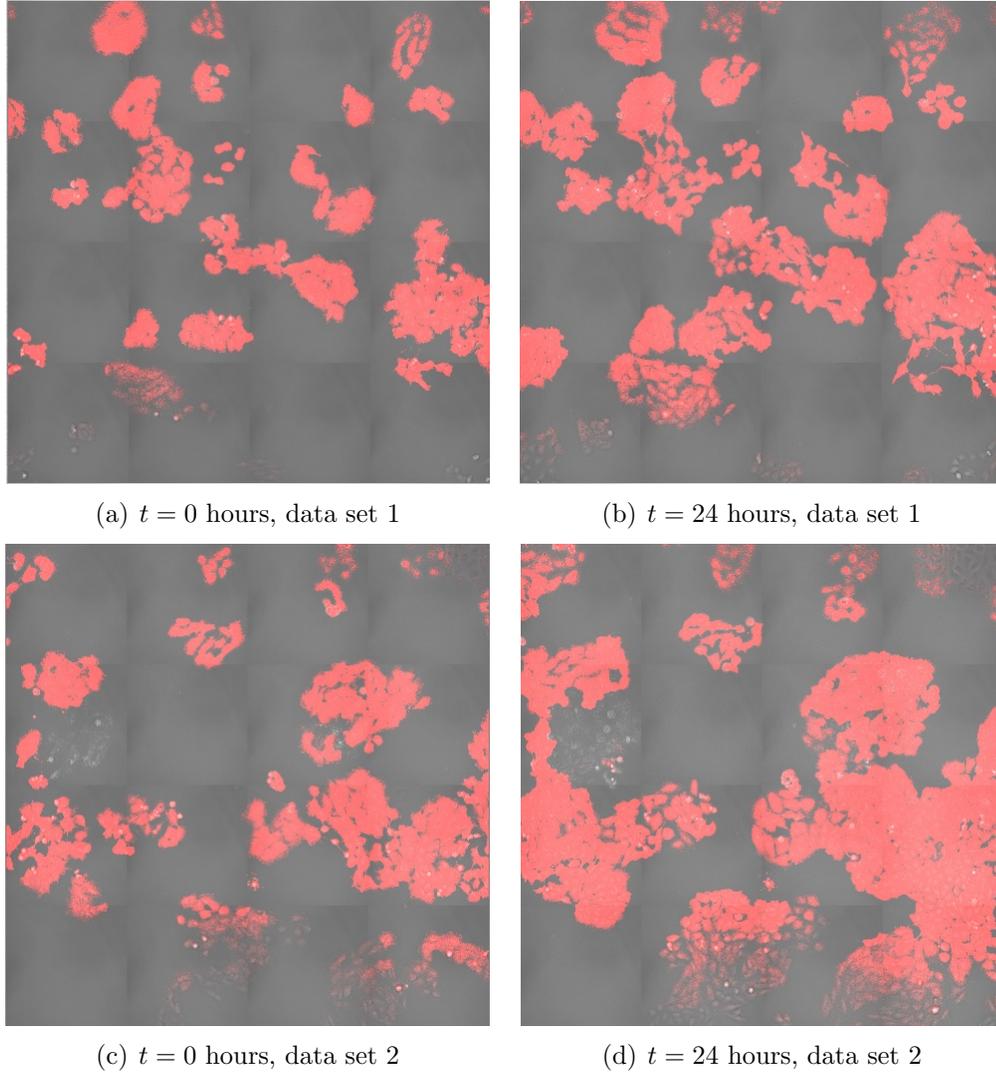


Figure 6.6: Initial and final configurations of the experimental data used for model comparison. Note that all cells are marked with red fluorescence, independently of their position in division or apoptosis. (a) Data set 1 initial conditions. (b) Data set 1 after 24 hours of cell growth. (c) Data set 2 initial conditions. (d) Data set 2 after 24 hours of cell growth.

investigation is qualitative, approximate parameter values will suffice. Calculating the least-squares fit for each data set (solid lines in Figure 6.7), and averaging their values at $\sigma = 0.2$, we obtain $\epsilon_a \approx 5$, which we take as an equality for the remainder of this chapter.

We also realize parameter values that produce the cell growth data presented

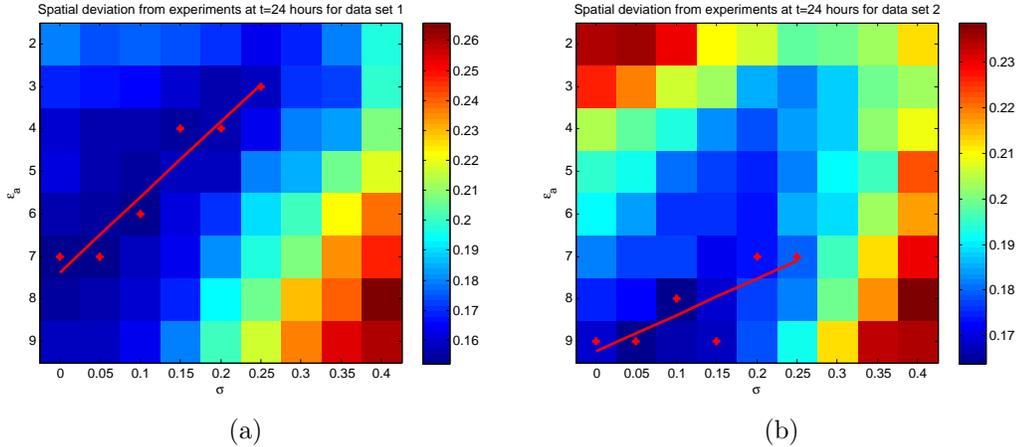


Figure 6.7: Spatial deviation parameter search. We plot the ℓ^1 spatial pixel distance between the experimental data and the simulations for a range of parameters ($\epsilon_a \in [2, 9], \theta \in [0, 0.4]$) for each data set. Also plotted are the minimums for each value of θ producing a small error (dots) and the least-squares line $\tilde{\epsilon}_a$ connecting these minimums (line). (a) Data set 1: $\tilde{\epsilon}_a = -17.714\sigma + 7.381$; (b) Data set 2: $\tilde{\epsilon}_a = -8.571\sigma + 9.24$.

in Figure 5.4. We recall that this data records the net cellular density dynamics over a four day period, for two sets of initial conditions, $\rho(0) = 0.1$ and $\rho(0) = 0.8$. Both cases consider a homogeneous configuration of cells in space, with an initial population that is entirely quiescent. The estimated parameter values are: $\beta_m = 2.8, \rho_m = 0.2, \rho_r = 0.85$, and $p = 0.5$. These values may not be unique or optimal, nevertheless, they do provide a good agreement with the experimental data, and hence we fix their values as such. Figure 6.8 compares the model output to the data for 10 Monte Carlo simulations of the ABM. Note that at least during the time simulated, the density curves for the different simulations remain relatively close to each other. The model closely fits the data in both cases.

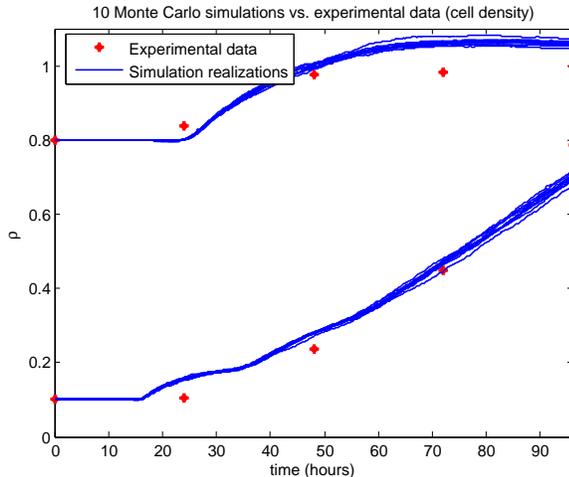


Figure 6.8: Comparison of growth data of OVCAR-8 human ovarian carcinoma cell line to model output. Experimental data (red) is compared to 10 Monte Carlo simulations with growth parameter values $\beta_m = 2.8$, $\rho_m = 0.2$, $\rho_r = 0.85$, and $p = 0.5$.

6.3.2 Cell-cycle distribution

We now investigate the effect of imposing a deterministic cell-cycle structure based on the spatial location of the cell. In particular, we are interested in the distribution of the cell-cycle as a function of time. Recall also that we are still analyzing the control growth dynamics in the absence of the drug, so we assume $c = 0$.

Instead of simulating both data sets (as in Figure 6.8), and analyzing two cell-cycle distributions, we consider a simulation starting from the low density value $\rho(0) = 0.1$, and allow the population of cells to grow until a time t such that a fully dense plate is reached, i.e., $\rho(t) = 1$. As in Section 6.3.1, the entire population of cells is initially assumed to be quiescent, and is uniformly distributed in space.

The dynamics of the total number of cells in the three compartments that were obtained from 10 Monte Carlo simulations are presented in Figure 6.9. In

Figure 6.9(a), we observe a population of cells growing logically, as well as a close correspondence between the model and the experimental data. In Figure 6.9(b), we plot the fractions of each cellular compartment (P, Q, and A) as a function of time. Note the aperiodic nature of the proliferative and apoptotic compartments, as the average local density increases and the cells transition in and out of the cell-cycle. We also observe that N_a/N appears to be relatively stable, except for large time simulation when its value increases. This is explained by Figure 6.9(c), where we plot the proportion of apoptotic transitions from both compartments Q and P. Initially, the overall density is low, so that the mitotic transition mechanism (see Section 6.2.3.2) does not take effect, and cells transfer from Q to P uniformly via (6.8). However, as the population grows, local density effects imply increased cell-cycle lengths to values near the transition length ($0.95\mu/p$), thus increasing the overall transition rate into apoptosis.

Figure 6.10 illustrates the effect of space on the cell-cycle distribution (equations (6.9)-(6.10)). In Figure 6.10(a), we plot a histogram for the lengths of successful divisions for 10 hour windows in a single simulation. For example, the first histogram labeled 15 on the “occurrence interval (hours)” axis denotes the distribution for all divisions occurring in simulation time t such that $t \in [15, 25]$. We observe a traveling wave of normal-like distributions, starting with small mean at low densities which gradually increases as the population grows (Figure 6.10(b)). Averaging over all time intervals yields a distribution with mean ≈ 23 , similar to the value used in Chapter 5 (24.4116). In Figure 6.10(c), we see that the variance of the distribution has interesting temporal dynamics, with a relative maximum oc-

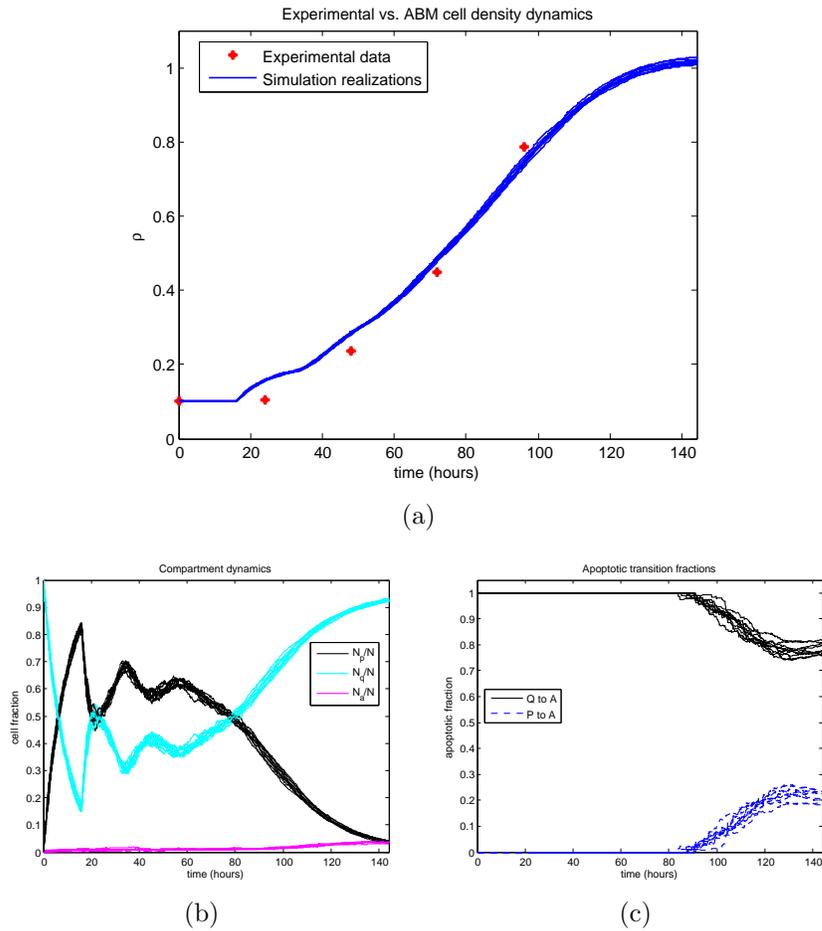
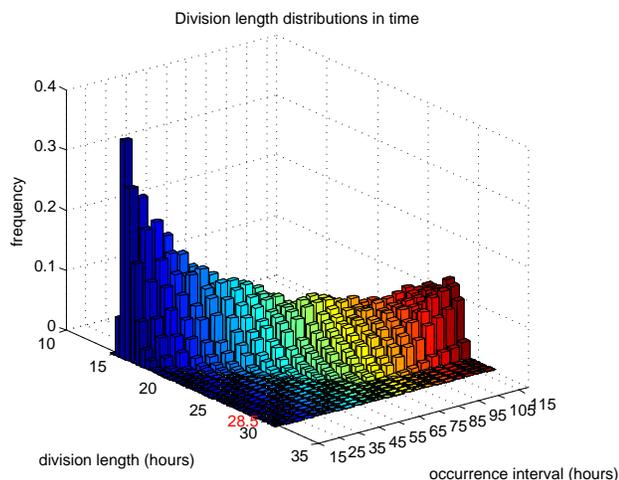
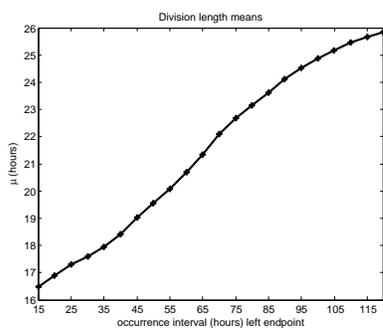


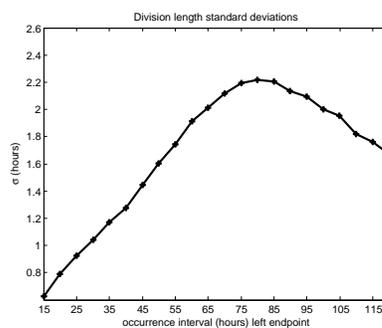
Figure 6.9: 10 Monte Carlo long-time simulations of the ABM. (a) Cellular density growth analogous to Figure 6.8, where the mathematical model satisfies $\rho(0) = 0.1$ and $\rho(t = 145 \text{ hours}) = 1$. (b) The dynamics of the different compartments. Note that the apoptotic fraction (N_a/N) remains relatively constant, except for large time, when the mitotic checkpoint begins to take effect. (c) Fraction of apoptotic cells coming from Q (rate α_{a_q}) and P (rate α_{a_p}). Initially all apoptotic transitions originate from compartment Q, while as time and density increase, the percentage of transitions from P increases as well.



(a)



(b)



(c)

Figure 6.10: Cell-cycle distributions generated by local density. (a) Histogram for cell-cycle lengths based on the time interval of division. Division intervals are 10 hours in length. The red 28.5 denotes mitotic threshold length ($0.95\mu/p$) after which cells must transfer from P to A. (b) Mean of division lengths as a function of the occurrence interval. (c) Standard deviation of division lengths as a function of the occurrence interval.

curing at approximately $t = 80$ hours. This can be explained by observing that initially, all cells experience a low density environment, so that we expect a relatively homogenous local cell density. Similarly, when the cell population is large (i.e., t is large), we expect a high average local cell density, which is also homogenous. Hence, the variance is expected to have a relative maximum at some intermediate time, which is precisely what we observe.

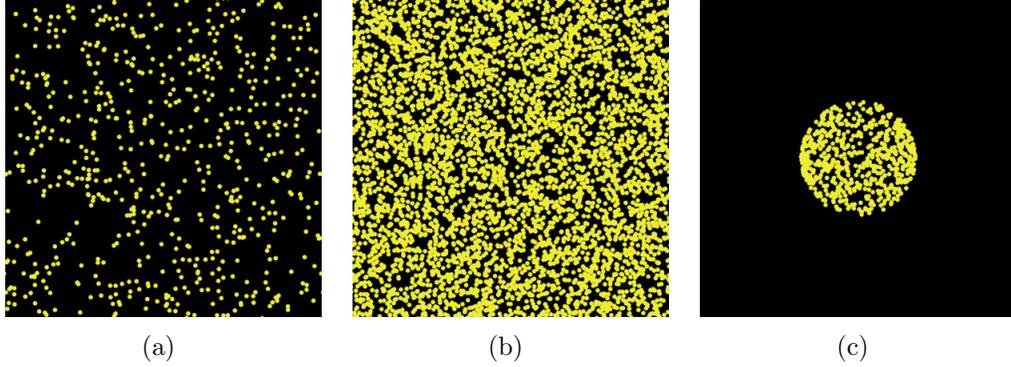


Figure 6.11: Initial *in vitro* spatial configuration of cells for the killing curve experiment. The figures show the distribution of cells in a two-dimensional environment initialized in the ABM, which qualitatively agrees with the experimental seeding. All cells initially reside in compartment Q (yellow). (a) Configuration 1: $\rho(0) = 0.1$, uniform random on entire plate. (b) Configuration 2: $\rho(0) = 0.8$, uniform random on entire plate. (c) $\rho(0) = 0.1$, uniform random on compact ball.

6.3.3 Density-dependent drug resistance

We now utilize the ABM to provide preliminary results with regards to the implementation of the ABM on experimental data describing the dynamics of *in vitro* OVCAR-8 cell lines, subject to the drug Paclitaxel in different geometrical configurations. This experimental data is given as killing curves for three different spatial configurations of cells:

1. $\rho(0) = 0.1$, spatially distributed uniformly on the plate
2. $\rho(0) = 0.8$, spatially distributed uniformly on the plate
3. $\rho(0) = 0.1$, spatially distributed uniformly in a compact ball

See Figure 6.11 for the ABM implementations of the above initial conditions. Killing curves provide the percentage of cell that survive 72 hours of treatment as a function of the dosage of the drug. In the experiment, the cells are seeded (according to the

geometrical configuration 1, 2, or 3) as entirely quiescent, and are left to grow without the drug for one day. The drug Paclitaxel (at a given dosage) is then administered to the plate, and after a period of three days, the number of remaining cells are counted. This result is then normalized with respect to four days of control cell growth to obtain a survival fraction as a function of dosage, $S_i(c)$, $i = 1, 2, 3$.

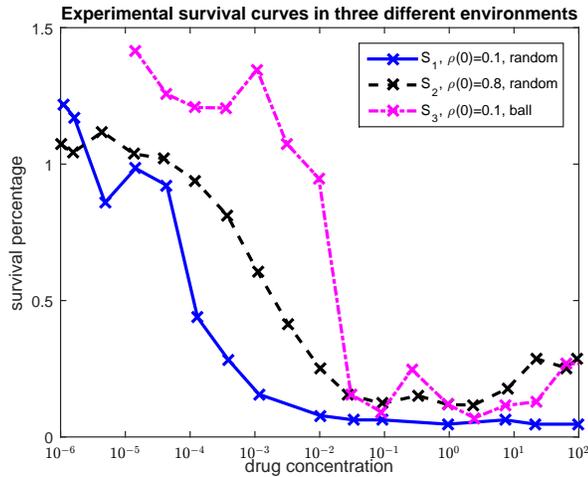


Figure 6.12: Experimental killing curves for three geometrical configurations 1, 2, and 3. Relative cell survival (compared to the control case) is measured after three days of treatment as a function of dosage (x -axis). Note that the dense geometries (both on the entire plate 2 and locally in 3, see Figure 6.11) experience on overall greater survival rate, and thus are drug resistant.

The experimental killing curves are shown in Figure 6.12. We recall that we are only interested in qualitative aspects of the experimental data, as the model relies on cellular parameters whose values are not available. The salient features of the three curves are listed below:

- (a) $S_2(c)$ and $S_3(c)$ exhibit a higher survival percentages (i.e. resistant to treatment) than $S_1(c)$ for large c .
- (b) $S_1(c)$ decays with a steeper slope than $S_2(c)$.

- (c) $S_3(c)$ is greater than S_1 for small c , with $S_3(c)$ approaching S_1 for large c .
- (d) $S_2(c)$ and $S_3(c)$ should begin decreasing for a larger c value than $S_1(c)$.

Biologically, these results imply that dense microenvironments exhibit a greater cell survival than their less dense counterparts. Thus, the geometrical configuration of cells impacts the overall efficiency of treatment, which is an intrinsic form of drug resistance. Our hypothesis is that this resistance can be traced to the cell-cycle dynamics.

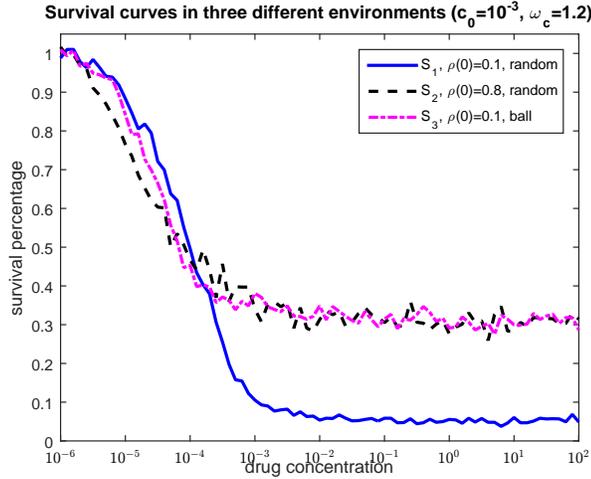


Figure 6.13: Killing curves for three geometrical configurations 1, 2, and 3 computed from the ABM. Relative cell survival (compared to the control case) is computed after three days of treatment as a function of dosage (x -axis). Note that the dense geometries (2 and 3) experience on overall greater survival rate, and thus are drug resistant. The qualitative features of the model output presented here should be compared to the corresponding curves in Figure 6.12.

Using the ABM model, we simulate the geometrical configurations 1,2, and 3. We estimate the c_0 and ω_c parameters in equation (6.11) to approximately match the experimental dynamics of case 1, and thus fix their values at $c_0 = 10^{-3}$ and $\omega_c = 1.2$. The results are shown in Figure 6.13. We observe that, qualitatively, the

simulation results agree with properties (a), (b), and (c) of the experimental data. Namely, the dense configurations experience a higher relative survival percentage for most drug concentrations, and experience a more gradual decrease for increasing drug concentrations. The simulations also reproduce the *local* (3) vs. *global* high density (2) results, with the compact configuration exhibiting more drug resistance for low drug concentrations (c), and approximately equal resistance as c increases.

Although the model successfully accounts for this intrinsic form of drug resistance, it does not completely match the experimental data. For instance, in Figure 6.13, we observe that all three killing curves begin decreasing at the same approximate drug concentration ($c \approx 10^{-5}$), which is in contrast to property (d). This is most likely due to the independence relation assumed between ρ_{loc} and c in equation (6.11). Indeed, this was a simplifying approximation, as the current understanding suggests that local density is more important for early stages of the cell-cycle ($G1 \rightarrow S \rightarrow G2$), while treatment is more important in the M phase. We are currently investigating other relations that consider these biological observations more accurately. Furthermore, it also appears that the less dense environment is most resistant for small values of c in Figure 6.13, which we do not observe experimentally. One explanation could be that for high local densities, the spatial distribution of the drug is not uniform, so that the crowding in the high density environments prevent some cells from being exposed to the drug. Future work will model the spatial dynamics of the treatment.

6.4 Discussion

In this chapter, we introduced a spatial mechanism for studying the distribution in the cell-cycle lengths. This distribution arises via the local interaction of cells, namely to decrease the rate of progression through the cell-cycle as the number of neighbors increases. We showed that this mechanism can explain experimental cell growth data. We also used it to generate a traveling wave of normal distributions describing the cell-cycle. Motivated by the method of action of Paclitaxel, we incorporated in the model an explicit mechanism for transitioning from division to apoptosis by prolonging a mitotic checkpoint. Taken together, this framework is shown to successfully predict an intrinsic form of drug resistance based on the cellular microenvironment, independent of any specific mutation. Thus, this model represents an important first step for deriving mathematical models that include specific mutational models of drug resistance.

Chapter 7: Conclusion

In this dissertation, we have mathematically studied drug resistance arising in cancer chemotherapy, together with the related notion of tumor heterogeneity. Drug resistance is a complex phenomenon, with a multitude of confounding factors. For example, it is well known that tumor cells exhibit a distribution of cellular growth, differentiation, and apoptosis rates due to genetic variation. The spatial arrangement of cells also determines local density measurements, which affects basic cellular processes, such as cell-cycle mechanics and nutrient uptake/distribution. Such influences imply that a tumor constitutes an extremely heterogeneous environment, and that the effect of treatment is non-uniform on such a population. In the first two chapters, we reviewed tumor heterogeneity and drug resistance from a biological perspective (Chapter 1), and reviewed part of the literature on mathematical models for drug resistance (Chapter 2).

In Chapter 3, we introduced a mathematical framework for describing multi-drug resistance to cancer. Heterogeneity was modeled using a continuous parameter, i.e. a “trait,” upon which cell division, natural apoptosis, and drug-induced apoptosis were all dependent. Specifically, we derived a system of integro-differential equations (IDEs), structured by the trait, which also included density effects as well

as mutations. We studied the limiting behavior of this model, both analytically and numerically, and discovered how one or more traits are selected corresponding to different levels of resistance. Furthermore, we discussed the validity of the work in comparison with experimental data, and discovered that the emerging limiting distribution with nonzero variance is the desirable modeling outcome, as it represents tumor heterogeneity. In Chapter 4, we studied extensions of the IDE model, and applied it to treatment protocols. For example, in the case of a high mutation rate, our results suggest that if this rate cannot be effectively reduced, increasing its value is expected to result in a better treatment outcome, a non-intuitive result that can lead to novel treatment strategies.

We next studied the cell-cycle and its effect on cell growth in Chapter 5. We formulated a stochastic agent-based model (ABM) that describes the growth of cancer cells as a dynamic transition between proliferative, quiescent, and apoptotic states. In addition, we demonstrated that the agent-based model can be well approximated by the more computationally efficient integro-differential equations when the number of cells is large. These models predict variations in growth as a function of the intrinsic heterogeneity emerging from the durations of the cell-cycle and apoptosis, and also include cellular density dependencies. By examining the role all parameters play in the evolution of intrinsic tumor heterogeneity, and the sensitivity of the population growth to parameter values, we showed that the cell-cycle length has the most significant effect on the growth dynamics. This work also included novel implementations of experimental data into the transition rates.

Finally, in Chapter 6, we considered the impact of space as a mechanism to ac-

count for the intrinsic distribution assumed in Chapter 5. We assumed the transition rates of each cell no longer depended on the global density, but instead are governed by a local density that is measured in a cellular microenvironment. For cells in the proliferative state, the cell-cycle length is assumed to increase with local density, and a mitotic checkpoint mechanism is utilized to govern transitions from proliferation to apoptosis. We exhibited close time-averaged agreement with the distribution used in Chapter 5, and showed that full distribution of cell-cycle lengths resembles a traveling wave of normal distributions with time-dependent mean and variance. We also incorporated a specific cytotoxic mechanism of action for a specific drug (Paclitaxel), and survival curves were calculated and shown to qualitatively agree with experimental measurements in different densities and geometries. Ongoing and future work will consider treatment strategies based on tumor configurations, as well as explicit spatial modeling of the drug and nutrients.

Appendix A: Chapter 3

Proposition 1. *Consider the integro-differential equation (3.6). If $n(x, 0) \geq 0$ for all $x \in [0, 1]$, then*

$$n(x, t) \geq 0 \quad \forall t \in \mathbb{R}_+.$$

Proof. The global existence of continuously differentiable solutions of (1) can be obtained following standard arguments (see [129]). For the positivity of n , since $n \geq 0$ at $t = 0$, due to its continuity, if there exists a t_* for which $n(x, t_*) = 0$ for some $x \in [0, 1]$ (and $n(x, t) > 0$ for $t < t_*$) then $n(y, t_*) \geq 0$ for all y . Hence by (3.6),

$$\frac{\partial n(x, t_*)}{\partial t} = \theta \int_0^1 r(y)M(y, x)n(y, t_*) dy \geq 0, \quad (\text{A.1})$$

which implies that $n(x, t_*)$ is nondecreasing at t_* . Consequently, $n(x, t)$ cannot pass through 0, as stated. \square

Proof of Theorem 1. We follow the proof of [121, Lemma 2.2]. Consider system (3.8). If $r(x) - c(x) - d(x) < 0$ for all $x \in [0, 1]$, then due to the positivity of $n(x, t)$, $\frac{\partial n(x, t)}{\partial t} < 0$ for all $x \in [0, 1]$. Hence, $n(x, t) \xrightarrow[t \rightarrow \infty]{} 0$ in $[0, 1]$. By Lebesgue's

Dominated Convergence Theorem, this implies that $\rho(t) \rightarrow 0$ as $t \rightarrow \infty$.

If, on the other hand, there exists x_* such that $r(x_*) - c(x_*) - d(x_*) = 0$, then these are fixed points of (3.8), and hence,

$$n(x_*, t) = n(x_*, 0), \quad \forall t \in \mathbb{R}_+.$$

Now suppose that there exists $x \in [0, 1]$ such that $r(x) - c(x) - d(x) > 0$. By continuity of the growth parameters and the compactness of $[0, 1]$, $r(x) - c(x) - d(x)$ achieves its maximum, say at $\{x_i\}_{i=1}^m$. We note that it is possible that $m = \infty$, or even that the set $\{x_i\}_{i=1}^m$ is uncountable (in which case our notation should be altered).

To see that $\rho(t) \xrightarrow[t \rightarrow \infty]{} \infty$, fix $x_j \in \{x_i\}_{i=1}^m$ such that $n(x_j, 0) > 0$ (as the points where $n = 0$ do not contribute to the growth). Then for all $0 < \lambda < r(x_j) - c(x_j) - d(x_j)$, exists γ_λ such that

$$r(x) - c(x) - d(x) \geq \lambda > 0 \quad \forall x \in [x_j - \gamma_\lambda, x_j + \gamma_\lambda].$$

Let $h_\lambda(t) := \int_{x_j - \gamma_\lambda}^{x_j + \gamma_\lambda} n(x, t) dx$. Then

$$\frac{d}{dt} h_\lambda(t) = \int_{x_j - \gamma_\lambda}^{x_j + \gamma_\lambda} [r(x) - c(x) - d(x)] n(x, t) dx \geq \lambda \int_{x_j - \gamma_\lambda}^{x_j + \gamma_\lambda} n(x, t) dx = \lambda h_\lambda(t),$$

so that, for a positive constant $h(0)$,

$$h_\lambda(t) \geq h(0)e^{\lambda t}. \quad (\text{A.2})$$

As $\rho(t) = \int_0^1 n(x, t) dx \geq h_\lambda(t)$, (A.2) implies that $\rho(t) \xrightarrow[t \rightarrow \infty]{} \infty$, as desired. To find the limiting distribution, note that for $x \notin \{x_i\}_{i=1}^m$, choose λ such that

$$r(x) - c(x) - d(x) < \lambda < r(x_j) - c(x_j) - d(x_j).$$

Then,

$$\frac{n(x, t)}{\rho(t)} = \frac{n(x, 0)e^{[r(x)-c(x)-d(x)]t}}{\rho(t)} \leq \frac{n(x, 0)}{h(0)} e^{[r(x)-c(x)-d(x)-\lambda]t} \xrightarrow[t \rightarrow \infty]{} 0. \quad (\text{A.3})$$

As $\int_0^1 \frac{n(x, t)}{\rho(t)} dx = 1$, for all time t , we have the desired result, namely

$$\lim_{t \rightarrow \infty} \frac{n(x, t)}{\rho(t)} = \sum_{i=1}^m a_i \delta(x - x_i),$$

with $\sum_{i=1}^m a_i = 1$. If the number of maximizers is uncountable, a similar result will hold for a continuous measure. \square

Proof of Theorem 2. Let $n(x, t)$ satisfy (3.16). By Proposition 1, since $n(x, t) \geq 0$, $\rho(t) \geq 0$ as well. Let r_M, c_m , and d_m be constants such that $r(x) \leq r_M$, $c(x) \geq c_m >$

0, and $d(x) \geq d_m > 0$, and recall that $G \geq 0$. Hence, $\rho'(t)$ can be bounded above by

$$\frac{d\rho(t)}{dt} = \int_0^1 [r(x) - c(x) - G(\rho(t))d(x)]n(x, t) dx \leq [r_M - c_m - G(\rho(t))d_m]\rho(t).$$

Since $G(\rho) \xrightarrow{\rho \rightarrow \infty} \infty$, $\exists \rho_M$ such that $\rho(0) \leq \rho_M$ and $r_M - c_m - G(\rho_M)d_m < 0$. This implies that at $\rho = \rho_M$, $\rho'(t) < 0$, and hence $\rho(t) \leq \rho_M$. □

Appendix B: Chapter 4

B.1 Supplemental material for Section 4.1

As mentioned in Section 4.1.4, there are two types of modifications that could occur, heredity or temporal changes. $\theta(x)$ should be thought of as the summation of two separate parameters: $\theta_1(x)$ and $\theta_2(x)$. When a similar effect is applied to the modification (M) function, the system (3.6) can be rewritten as:

$$\begin{aligned} \frac{\partial n(x, t)}{\partial t} = & \left(r(x) \left(1 - \sum_{i=1,2} \theta_i(x) \right) - c(x, \alpha) - G(\rho(t))d(x) \right) n(x, t) \\ & + \sum_{i=1,2} \int_0^1 \theta_i(y) r(y) M_i(y, x) n(y, t) dy. \end{aligned} \tag{B.1}$$

Specific phenomena studied include modeling M functions with different timescales (see Fig. 4.5), as well as utilizing a single M function which varies its distribution parameters as a function of treatment. In the latter, treatment acts as a “stress,” modifying the number of cells mutated (θ) and the range of mutation (ϵ). In this case, the governing equation takes the form

$$\begin{aligned} \frac{\partial n(x, t)}{\partial t} = & \left(r(x) \left(1 - \theta(\text{stress}, x, t) \right) - c(x, \alpha) - G(\rho(t))d(x) \right) n(x, t) \\ & + \int_0^1 \theta(\text{stress}, y, t) r(y) M(y, x) n(y, t) dy. \end{aligned} \tag{B.2}$$

For results using this equation, see the last paragraph in Section 4.1.4.

For exact numerical parameters implemented, we refer the reader to the Supplementary Data section of our Cancer Research article [136].

B.2 Supplemental material for Section 4.2

We briefly review the optimization procedure used to obtain the cellular growth/death rates appearing in Figure 4.6B. We chose to represent $r(x) - c(x)$ as a sigmoid curve with four parameters, and the natural death rate $d(x)$ as a constant:

$$r(x) - c(x) = \frac{p_1 - p_2}{\pi} \left(\frac{\pi}{2} - \tan^{-1} \frac{p_3(x - p_4)}{x(1 - x)} \right) + p_2, \quad (\text{B.3})$$

$$d(x) \equiv p_5. \quad (\text{B.4})$$

The total density $\rho(t)$ is then obtained via integration over the trait space as in equation (4.4). We solved the linear system described in (4.3), with no mutations ($\theta = \epsilon = 0$) subject to IC_1 . To integrate over x , we discretized $[0, 1]$ uniformly, with step size $h = 1/3999$, and used a standard trapezoidal rule on $n(x, t)$. To find the parameters that best fit the 25 patient data points of the form $(t, \rho(t))$ in Stein *et al.* [1], we first performed a random parameter search over the five free parameters p_1, p_2, p_3, p_4, p_5 , constraining them to lie within physically realistic values, and solved for $\rho(t)$. This resulted with parameter values that were then used as an initial estimate for a nonlinear least-squares algorithm. The MATLAB built-in function “lsqcurvefit”

was used to obtain a local minimizer for the norm squared error:

$$LSQ(p_1, \dots, p_5) = \sum_{i=1}^{25} (\rho_{\text{theory}}(t; p_1, \dots, p_5) - \rho_{\text{data}}(t_i))^2. \quad (\text{B.5})$$

The parameter values obtained are approximately $p_1 = -0.0624$, $p_2 = 0.0173$, $p_3 = 0.0846$, $p_4 = 0.6144$, and $p_5 = 0.0097$. We then made a qualitative choice, that both $r(x)$ and $c(x)$ should appear sigmoid as well. We chose the same type of functional form for $c(x)$, with a corresponding set of parameters $\{a_i\}_{i=1}^4$ (there is no fifth parameter accounting for $d(x)$), with $a_1 = 0.3$, $a_2 = 0.1$, $a_3 = 0.1$, and $a_4 = 0.6$. Then, we can simply find $r(x)$ using the above two formulas and the fact that $r(x) = [r(x) - c(x)] + [c(x)]$. Plotting this with the initial condition $n(x, 0)$ above yields Figure 4.6B.

The same method was used to estimate the parameters of the growth (g) and the regression (d) rates in the original Stein model, with an initial estimate given by the rates found in [1]. The values we obtained for their parameters are $d = 0.0541$ and $g = 0.0005$ using equation (4.2).

Appendix C: Chapter 5

C.1 Experiments

C.1.1 Experimental Design

This work is aimed at studying the heterogeneity arising from the variation in the cell-cycle and apoptosis length. With this in mind, three sets of experiments were performed: a study of cell growth and density, an analysis of cell proliferative percentage, and an analysis of cell apoptotic percentage. The results of these experiments appear in Figure 5.2.

C.1.1.1 Analysis of Proliferation Percentage

One goal is to understand the role of global density on cellular division. The results appear as means and standard deviations in Figure 5.2(a). We assume that, for a given cell density, there exists an equilibrium distribution of cells in the cell-cycle, which we define as $\beta(\rho)$. To determine $\beta(\rho)$, two series of cell-cycle arrest experiments were performed, each followed by KI-67 measurements.

Cell-Cycle Arrest. OVCAR-8 cells were seeded in different cell densities (see Figure 5.2(a)) on a 6-well plate with culture medium (RPMI-1640 medium +

10%FBS + 100U/ml Penicillin-Streptomycin + 2 mM Glutamine) for 24 hours prior to cell cycle arrest. To arrest the cells, 2 experiments were performed by changing the starvation medium and the starvation duration. In both experiments, cells were washed 3 times with 1x PBS and replenished with starvation medium. Experiment 1) The medium of RPMI-1640 without Penicillin-Streptomycin, and Glutamine was used for cell cultured in 37°C + 5% CO₂ for 24 hours. The data set from this experiment was referred in the text and figures as “Data 1”. Experiment 2) The medium of RPMI-1640 without Glucose, Penicillin-Streptomycin, and Glutamine was used for cell cultured in 37°C + 5% CO₂ for 48 hours. The data set from this experiment was referred in the text and figures as “Data 2”. Arrested cells from both experiments were allowed to re-enter cell cycle by washing the cells with PBS 3 times and then replenished with culture medium for 16 hours. The materials: Roswell Park Memorial Institute (RPMI)-1640 medium, Fetal Bovine Serum, Penicillin-Streptomycin, and Glutamine, were purchased from Life Technologies (Grand Island, NY).

Cell-Cycle Analysis by Ki-67 Immunolabeling. OVCAR-8 cells were seeded on 6-well plate with culture medium. Cells were harvested by trypsinization and re-suspended with PBS. Diameter of individual cells was measured by Nexcelom Cellometer (Lawrence, MA). To permeabilize cells, 10⁶ cells were washed with 1x PBS and incubated with -20°C ethanol overnight. For Ki-67 labeling, cells were centrifuged and re-suspended with warm 1x PBS for 15 minutes. Then, cells were incubated with IMDM medium supplemented with 5% FBS and IgG2a-FITC (0.06 μg) or Ki-67-FITC (0.06 μg) at 37°C for 30 minutes. The fluorescence intensity of labeled cells was analyzed by Flow cytometry (BD Biosciences, San Jose, CA).

Results were analyzed by FlowJo version 7.6.4 (Ashland OR). IgG2a-FITC and Ki67-FITC were purchased from eBioscience (San Diego, CA), and Iscove's Modified Dulbecco's Medium (IMDM) was purchased from Life Technologies (Grand Island, NY).

C.1.1.2 Evaluation of Cell Growth and Density

A second goal is to understand the individual cellular mechanics that govern the macroscopic dynamics in cancer cells. Thus, time series data is needed. Experiments were performed which measure the global cellular density over a period of 96 hours in two different seedings: low and high density. See Figure 5.4 for the results. All experiments were performed using the parental OVCAR-8 and the OVCAR-8-DsRed2 human ovarian carcinoma cell lines [103]. In this study, all of our applicable cell-based assays utilize timelines of 96 hours or less. The cell density curves were generated by randomly seeding 10^4 and 10^5 cells per well on 24-well plates and culturing 24 hours prior to the first imaging. The plates were imaged on a Zeiss 710 Confocal microscope on both green and red fluorescent channels, and images were captured every 24 hours for 96 hours. The estimated density percentage is based on the average of two complete wells.

C.1.1.3 Analysis of Apoptosis Percentage

As in C.1.1.1, we hypothesize that an equilibrium distribution $d(\rho)$ exists for the fraction of cells in the apoptosis process. An analogous experiment was conducted

in the aid of determining $d(\rho)$. Percentage of apoptosis-mediated cell death as a function of cell density was measured by a double staining method using the Dead Cell Apoptosis Kit with Annexin V Alexa Fluor[®] 488 & PI (Invitrogen) according to the manufacturer's instructions. 10^4 and 10^5 parental OVCAR-8 cells were randomly seeded on 24-well plates in drug-free medium and incubated at 37°C for 24 hours to allow cells to attach. The percentages of apoptosis and necrosis were then evaluated by FACS (LSR II). Data from 20,000 gated events per sample were collected. Cells in early stages of apoptosis were positively stained with Annexin V, whereas cells in late apoptosis and necrosis were positively stained with both Annexin V and PI. The number of cells positively stained with Annexin V was used to estimate the death rate. Experiments were repeated three times; see Figure 5.2(b) for the results.

C.2 Variable table

Variable/ Parameter	Range (units)	Interpretation
t	[0, 96] (hours)	Time
Δt	0.1 (hours)	Time step used in ABM simulations
$N_q(t)$	[0, ∞)	Number of cells in quiescence at time t
$N_p(t)$	[0, ∞)	Number of cells in cell-cycle at time t
$N_a(t)$	[0, ∞)	Number of cells in apoptosis at time t
$N(t)$	[0, ∞)	Total number of cells at time t
μ	24.4416 (hours)	Mean length of cell-cycle
σ	[0, 10] (hours)	Standard deviation of cell-cycle
ω_1	5.0289	Shape parameter of apoptosis phase 1
λ_1	0.6534 (per hour)	Rate parameter of apoptosis phase 1
ω_2	1.2869	Shape parameter of apoptosis phase 2
λ_2	0.23173 (per hour)	Rate parameter of apoptosis phase 2
ω_3	12.2679	Shape parameter of apoptosis phase 3
λ_3	10.3643 (per hour)	Rate parameter of apoptosis phase 3
ω	4.9436	Shape parameter of entire apoptosis process
λ	0.19117 (per hour)	Rate parameter of entire apoptosis process
$\rho(t)$	[0, ∞)	Density of cells at time t
K	40401	Number of cells defining full plate ($\rho = 1$)
$d(\rho), d$	[0.01, 0.05]	Fraction of cells in apoptosis as a function of plate density
$\beta(\rho)$	[0, 1]	Fraction of cells in cell-cycle as a function of plate density
β_m	[0, 1]	Relative/absolute maximum of $\beta(\rho)$
ρ_m	[0, 1]	Maximizing density of $\beta(\rho)$
ϵ	(0, ∞)	Parameter governing shape of $\beta(\rho)$
$\alpha_p(t)$	[0, ∞)	Rate of transition from Q to P
$\alpha_{a_p}(t)$	[0, ∞)	Rate of transition from P to A
$\alpha_{a_q}(t)$	[0, ∞)	Rate of transition from Q to A
c	[0, ∞) (per hour)	Cellular reaction rate
γ	[0, 1]	Rate difference between α_{a_p} and α_{a_q}
$f_p(\cdot; \mu, \sigma)$	[0, ∞)	PDF of $\mathcal{N}(\mu, \sigma^2)$ cell-cycle length
$f_a(\cdot)$	[0, ∞)	PDF of Γ distributed apoptosis process length
E_ρ	$(-\infty, \infty)$	Error vector between IDE and experimental data

Table C.1: List of parameters and variables

Bibliography

- [1] Wilfred D Stein, William Doug Figg, William Dahut, Aryeh D Stein, Moshe B Hoshen, Doug Price, Susan E Bates, and Tito Fojo. Tumor growth rates derived from data for patients in a clinical trial correlate strongly with patient survival: a novel strategy for evaluation of clinical trial data. *The Oncologist*, 13(10):1046–1054, 2008.
- [2] Rebecca Siegel, Deepa Naishadham, and Ahmedin Jemal. Cancer statistics, 2013. *CA: A Cancer Journal for Clinicians*, 63(1):11–30, 2013.
- [3] Hassan Jumaa, Rudolf W. Hendriks, and Michael Reth. B cell signaling and tumorigenesis. *Annual Review of Immunology*, 23(1):415–445, 2005. PMID: 15771577.
- [4] Ziqiang Zhang, Zhihong Chen, Yuanlin Song, Pinghai Zhang, Jie Hu, and Chunxue Bai. Expression of aquaporin 5 increases proliferation and metastasis potential of lung cancer. *The Journal of Pathology*, 221(2):210–220, 2010.
- [5] Scott W. Lowe and Athena W. Lin. Apoptosis in cancer. *Carcinogenesis*, 21(3):485–495, 2000.
- [6] Eric R. Fearon and Bert Vogelstein. A genetic model for colorectal tumorigenesis. *Cell*, 61(5):759 – 767, 1990.
- [7] Masakazu Yashiro, John M. Carethers, Luigi Laghi, Koji Saito, Premysl Slezak, Edgar Jaramillo, Carlos Rubio, Koichi Koizumi, Kosei Hirakawa, and C. Richard Boland. Genetic pathways in the evolution of morphologically distinct colorectal neoplasms. *Cancer Research*, 61(6):2676–2683, 2001.
- [8] M. Ilyas, J. Straub, I.P.M. Tomlinson, and W.F. Bodmer. Genetic pathways in colorectal and other cancers. *European Journal of Cancer*, 35(3):335 – 351, 1999.
- [9] J.M. Bishop, R.A. Weinberg, and inc Scientific American. *Scientific American Molecular Oncology*. Scientific American introduction to molecular medicine. Scientific American, 1996.
- [10] Jiro Inagaki, Victorio Rodriguez, and Gerald P. Bodey. Causes of death in cancer patients. *Cancer*, 33(2):568–573, 1974.

- [11] M.B. Sporn. The war on cancer. *Lancet*, 347(9012):1377–1381, 1996.
- [12] Michael J Tisdale. Cachexia in cancer patients. *Nature Reviews Cancer*, 2(11):862–871, 2002.
- [13] NCI. What is cancer? <http://www.cancer.gov/cancertopics/cancerlibrary/what-is-cancer>.
- [14] Douglas Hanahan and Robert A Weinberg. The hallmarks of cancer. *Cell*, 100(1):57 – 70, 2000.
- [15] Douglas Hanahan and Robert A Weinberg. Hallmarks of cancer: The next generation. *Cell*, 144(5):646–674, 2011.
- [16] Robert S Kerbel. A cancer therapy resistant to resistance. *Nature*, 390(6658):335–336, 1997.
- [17] Vincent T. DeVita and Edward Chu. A history of cancer chemotherapy. *Cancer Research*, 68(21):8643–8653, 2008.
- [18] Vikas Malhotra and Michael C Perry. Classical chemotherapy: mechanisms, toxicities and the therapeutic window. *Cancer biology and therapy*, 2:1–3, 2003.
- [19] Christopher P Leamon and Joseph A Reddy. Folate-targeted chemotherapy. *Advanced drug delivery reviews*, 56(8):1127–1141, 2004.
- [20] Renaud Capdeville, Elisabeth Buchdunger, Juerg Zimmermann, and Alex Matter. Glivec (sti571, imatinib), a rationally developed, targeted anticancer drug. *Nature Reviews Drug Discovery*, 1(7):493–502, 2002.
- [21] Robert J Motzer, Thomas E Hutson, Piotr Tomczak, M Dror Michaelson, Ronald M Bukowski, Olivier Rixe, Stéphane Oudard, Sylvie Negrier, Cezary Szczylik, Sindy T Kim, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *New England Journal of Medicine*, 356(2):115–124, 2007.
- [22] Radcliffe Infirmary CTSU. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet*, 365:1687–1717, 2005.
- [23] H. A. Azim, E. de Azambuja, M. Colozza, J. Bines, and M. J. Piccart. Long-term toxic effects of adjuvant chemotherapy in breast cancer. *Annals of Oncology*, 2011.
- [24] Andriy Marusyk and Kornelia Polyak. Tumor heterogeneity: Causes and consequences. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1805(1):105 – 117, 2010.

- [25] So Yeon Park, Mithat Gönen, Hee Jung Kim, Franziska Michor, Kornelia Polyak, et al. Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. *The Journal of clinical investigation*, 120(2):636–644, 2010.
- [26] Fabienne Kerangueven, Tetsuro Noguchi, François Coulier, Florence Allione, Véronique Wargniez, Joelle Simony-Lafontaine, Michel Longy, Jocelyne Jacquemier, Hagay Sobol, François Eisinger, et al. Genome-wide search for loss of heterozygosity shows extensive genetic diversity of human breast carcinomas. *Cancer research*, 57(24):5469–5474, 1997.
- [27] Victoria Sanz-Moreno, Gilles Gadea, Jessica Ahn, Hugh Paterson, Pierfrancesco Marra, Sophie Pinner, Erik Sahai, and Christopher J Marshall. Rac activation and inactivation control plasticity of tumor cell movement. *Cell*, 135(3):510–523, 2008.
- [28] Sabrina L Spencer, Suzanne Gaudet, John G Albeck, John M Burke, and Peter K Sorger. Non-genetic origins of cell-to-cell variability in trail-induced apoptosis. *Nature*, 459(7245):428–432, 2009.
- [29] Piyush B Gupta, Christine M Fillmore, Guozhi Jiang, Sagi D Shapira, Kai Tao, Charlotte Kuperwasser, and Eric S Lander. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell*, 146(4):633–644, 2011.
- [30] Sreenath V Sharma, Diana Y Lee, Bihua Li, Margaret P Quinlan, Fumiyuki Takahashi, Shyamala Maheswaran, Ultan McDermott, Nancy Azizian, Lee Zou, Michael A Fischbach, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*, 141(1):69–80, 2010.
- [31] Gloria H Heppner. Tumor heterogeneity. *Cancer research*, 44(6):2259–2265, 1984.
- [32] Isaiah J Fidler. Tumor heterogeneity and the biology of cancer invasion and metastasis. *Cancer research*, 38(9):2651–2660, 1978.
- [33] Nicholas A Saunders, Fiona Simpson, Erik W Thompson, Michelle M Hill, Liliana Endo-Munoz, Graham Leggatt, Rodney F Minchin, and Alexander Guminski. Role of intratumoural heterogeneity in cancer drug resistance: molecular and clinical perspectives. *EMBO molecular medicine*, 4(8):675–684, 2012.
- [34] Vanessa Almendro, Andriy Marusyk, and Kornelia Polyak. Cellular heterogeneity and molecular evolution in cancer. *Annual Review of Pathology: Mechanisms of Disease*, 8:277–302, 2013.
- [35] Angela Oliveira Pisco, Amy Brock, Joseph Zhou, Andreas Moor, Mitra Mojtabehi, Dean Jackson, and Sui Huang. Non-darwinian dynamics in therapy-induced cancer drug resistance. *Nature communications*, 4, 2013.

- [36] Ping Wang, Lydia Robert, James Pelletier, Wei Lien Dang, Francois Taddei, Andrew Wright, and Suckjoon Jun. Robust growth of escherichia coli. *Current biology*, 20(12):1099–1103, 2010.
- [37] Amit Tzur, Ran Kafri, Valerie S LeBleu, Galit Lahav, and Marc W Kirschner. Cell growth and size homeostasis in proliferating animal cells. *Science*, 325(5937):167–171, 2009.
- [38] Conrad A Messam and Randall N Pittman. Asynchrony and commitment to die during apoptosis. *Experimental cell research*, 238(2):389–398, 1998.
- [39] Paraskevi Giannakakou, Rick Gussio, Eva Nogales, Kenneth H Downing, Daniel Zaharevitz, Birgit Bollbuck, George Poy, Dan Sackett, KC Nicolaou, and Tito Fojo. A common pharmacophore for epothilone and taxanes: molecular basis for drug resistance conferred by tubulin mutations in human cancer cells. *Proceedings of the National Academy of Sciences*, 97(6):2904–2909, 2000.
- [40] Thomas O’Hare, Christopher A Eide, and Michael WN Deininger. Bcr-abl kinase domain mutations, drug resistance, and the road to a cure for chronic myeloid leukemia. *Blood*, 110(7):2242–2249, 2007.
- [41] William Pao, Theresa Y Wang, Gregory J Riely, Vincent A Miller, Qiulu Pan, Marc Ladanyi, Maureen F Zakowski, Robert T Heelan, Mark G Kris, and Harold E Varmus. Kras mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS medicine*, 2(1):e17, 2005.
- [42] Sandra Misale, Rona Yaeger, Sebastijan Hobor, Elisa Scala, Manickam Janakiraman, David Liska, Emanuele Valtorta, Roberta Schiavo, Michela Buscarino, Giulia Siravegna, et al. Emergence of kras mutations and acquired resistance to anti-egfr therapy in colorectal cancer. *Nature*, 2012.
- [43] Robert T Schimke. Gene amplification, drug resistance, and cancer. *Cancer research*, 1984.
- [44] Mercedes E Gorre, Mansoor Mohammed, Katharine Ellwood, Nicholas Hsu, Ron Paquette, P Nagesh Rao, and Charles L Sawyers. Clinical resistance to sti-571 cancer therapy caused by bcr-abl gene mutation or amplification. *Science*, 293(5531):876–880, 2001.
- [45] Jeffrey A Engelman, Kreshnik Zejnullahu, Tetsuya Mitsudomi, Youngchul Song, Courtney Hyland, Joon Oh Park, Neal Lindeman, Christopher-Michael Gale, Xiaojun Zhao, James Christensen, et al. Met amplification leads to gefitinib resistance in lung cancer by activating erbb3 signaling. *Science*, 316(5827):1039–1043, 2007.
- [46] SPC Cole, G Bhardwaj, JH Gerlach, JE Mackie, CE Grant, KC Almquist, AJ Stewart, EU Kurz, AM Duncan, and RG Deeley. Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science*, 258(5088):1650–1654, 1992.

- [47] Mohammad Azam, Robert R Latek, and George Q Daley. Mechanisms of autoinhibition and sti-571/imatinib resistance revealed by mutagenesis of *bcr-abl*. *Cell*, 112(6):831–843, 2003.
- [48] June H Carver, Gerald M Adair, and Daniel L Wandres. Mutagenicity testing in mammalian cells: Ii. validation of multiple drug-resistance markers having practical application for screening potential mutagens. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 72(2):207–230, 1980.
- [49] JP Gillet and MM Gottesman. Mechanisms of multidrug resistance in cancer. *Methods in molecular biology (Clifton, NJ)*, 596:47, 2010.
- [50] Yehuda G Assaraf. Molecular basis of antifolate resistance. *Cancer and Metastasis Reviews*, 26(1):153–181, 2007.
- [51] Jing Zhang, Frank Visser, Karen M King, Stephen A Baldwin, James D Young, and Carol E Cass. The role of nucleoside transporters in cancer chemotherapy with nucleoside drugs. *Cancer and Metastasis Reviews*, 26(1):85–110, 2007.
- [52] James J Farrell, Hany Elsaleh, Miguel Garcia, Raymond Lai, Ali Ammar, William F Regine, Ross Abrams, A Bowen Benson, John Macdonald, Carol E Cass, et al. Human equilibrative nucleoside transporter 1 levels predict response to gemcitabine in patients with pancreatic cancer. *Gastroenterology*, 136(1):187–195, 2009.
- [53] Hristos Glavinas, Peter Krajcsi, Judit Cserepes, and Balázs Sarkadi. The role of abc transporters in drug resistance, metabolism and toxicity. *Current drug delivery*, 1(1):27–42, 2004.
- [54] R. L. Juliano and V Ling. A surface glycoprotein modulating drug permeability in chinese hamster ovary cell mutants. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 455(1):152–162, 1976.
- [55] Kazumitsu Ueda, Carol Cardarelli, Michael M Gottesman, and Ira Pastan. Expression of a full-length cDNA for the human "mdr1" gene confers resistance to colchicine, doxorubicin, and vinblastine. *Proceedings of the National Academy of Sciences*, 84(9):3004–3008, 1987.
- [56] Matthew D Hall, Mitsunori Okabe, Ding-Wu Shen, Xing-Jie Liang, and Michael M Gottesman. The role of cellular accumulation in determining sensitivity to platinum-based chemotherapy*. *Annu. Rev. Pharmacol. Toxicol.*, 48:495–535, 2008.
- [57] Macus Tien Kuo, Helen HW Chen, Im-Sook Song, Niramol Savaraj, and Toshihisa Ishikawa. The roles of copper transporters in cisplatin resistance. *Cancer and Metastasis Reviews*, 26(1):71–83, 2007.

- [58] Muralikrishna Duvvuri and Jeffrey P Krise. Intracellular drug sequestration events associated with the emergence of multidrug resistance: a mechanistic review. *Front Biosci*, 10(2):1499–1509, 2005.
- [59] Roohangiz Safaei, Barrett J Larson, Timothy C Cheng, Michael A Gibson, Shinji Otani, Wiltrud Naerdemann, and Stephen B Howell. Abnormal lysosomal trafficking and enhanced exosomal export of cisplatin in drug-resistant human ovarian carcinoma cells. *Molecular cancer therapeutics*, 4(10):1595–1604, 2005.
- [60] Kevin G Chen, Julio C Valencia, Barry Lai, Guofeng Zhang, Jill K Paterson, Francois Rouzaud, Werner Berens, Stephen M Wincovitch, Susan H Garfield, Richard D Leapman, et al. Melanosomal sequestration of cytotoxic drugs contributes to the intractability of malignant melanomas. *Proceedings of the National Academy of Sciences*, 103(26):9903–9907, 2006.
- [61] Kristy J Gotink, Henk J Broxterman, Mariette Labots, Richard R de Haas, Henk Dekker, Richard J Honeywell, Michelle A Rudek, Laurens V Beerepoot, René J Musters, Gerrit Jansen, et al. Lysosomal sequestration of sunitinib: a novel mechanism of drug resistance. *Clinical Cancer Research*, 17(23):7337–7346, 2011.
- [62] Danyelle M Townsend and Kenneth D Tew. The role of glutathione-S-transferase in anti-cancer drug resistance. *Oncogene*, 22(47):7369–7375, 2003.
- [63] Lainie P Martin, Thomas C Hamilton, and Russell J Schilder. Platinum resistance: the role of dna repair pathways. *Clinical Cancer Research*, 14(5):1291–1295, 2008.
- [64] Jann N Sarkaria, Gaspar J Kitange, C David James, Ruth Plummer, Hilary Calvert, Michael Weller, and Wolfgang Wick. Mechanisms of chemoresistance to alkylating agents in malignant glioma. *Clinical Cancer Research*, 14(10):2900–2908, 2008.
- [65] Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1-3):389–422, 2002.
- [66] Orit Lavi, Michael M Gottesman, and Doron Levy. The dynamics of drug resistance: a mathematical perspective. *Drug Resistance Updates*, 15(1):90–97, 2012.
- [67] Salvador E Luria and Max Delbrück. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, 28(6):491, 1943.
- [68] Yoh Iwasa, Martin A Nowak, and Franziska Michor. Evolution of resistance during clonal expansion. *Genetics*, 172(4):2557–2566, 2006.

- [69] Franziska Michor, Martin A Nowak, and Yoh Iwasa. Evolution of resistance to cancer therapy. *Current pharmaceutical design*, 12(3):261–271, 2006.
- [70] JH Goldie and AJ Coldman. A mathematic model for relating the drug sensitivity of tumors to their spontaneous mutation rate. *Cancer treatment reports*, 63(11-12):1727–1733, 1978.
- [71] Cristian Tomasetti and Doron Levy. Role of symmetric and asymmetric division of stem cells in developing drug resistance. *Proceedings of the National Academy of Sciences*, 107(39):16766–16771, 2010.
- [72] Natalia L Komarova and Dominik Wodarz. Drug resistance in cancer: principles of emergence and prevention. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9714–9719, 2005.
- [73] Natalia Komarova. Stochastic modeling of drug resistance in cancer. *Journal of theoretical biology*, 239(3):351–366, 2006.
- [74] JH Goldie. Rationale for the use of alternating non-cross-resistant chemotherapy. *Cancer Treat Rep*, 66:439–449, 1982.
- [75] AJ Coldman and JH Goldie. A model for the resistance of tumor cells to cancer chemotherapeutic agents. *Mathematical Biosciences*, 65(2):291–307, 1983.
- [76] Jeng-Huei Chen, Ya-Hui Kuo, and Hsing Paul Luh. Optimal policies of non-cross-resistant chemotherapy on goldie and coldmans cancer model. *Mathematical biosciences*, 245(2):282–298, 2013.
- [77] John Carl Panetta. A mathematical model of drug resistance: heterogeneous tumors. *Mathematical biosciences*, 147(1):41–61, 1998.
- [78] Roger S Day. Treatment sequencing, asymmetry, and uncertainty: protocol strategies for combination chemotherapy. *Cancer Research*, 46(8):3876–3885, 1986.
- [79] Rachel Roe-Dale, David Isaacson, and Michael Kupferschmid. A mathematical model of breast cancer treatment with cmf and doxorubicin. *Bulletin of mathematical biology*, 73(3):585–608, 2011.
- [80] Urszula Ledzewicz and Heinz Schattler. Drug resistance in cancer chemotherapy as an optimal control problem. *Discrete and Continuous Dynamical Systems Series B*, 6(1):129, 2006.
- [81] Lev Semenovich Pontryagin. *Mathematical theory of optimal processes*. CRC Press, 1987.

- [82] Marek Kimmel and Andrzej Swierniak. Control theory approach to cancer chemotherapy: Benefiting from phase dependence and overcoming drug resistance. In *Tutorials in mathematical biosciences III*, pages 185–221. Springer, 2006.
- [83] A Polanski, M Kimmel, and A Swierniak. Qualitative analysis of the infinite-dimensional model of evolution of drug resistance. *Advances in Mathematical Population Dynamics-Molecules, Cells and Man (eds. Arino, Axelrod, Kimmel) World Scientific, Singapore*, pages 595–612, 1997.
- [84] A Świerniak, A Polański, Marek Kimmel, Adam Bobrowski, and J Śmieja. Qualitative analysis of controlled drug resistance model-inverse laplace and and semigroup approach. *Control and Cybernetics*, 28:61–73, 1999.
- [85] Jarosla Smieja, Andrzej Swierniak, and Zdzislaw Duda. Gradient method for finding opitimal scheduling in infinite dimensional models. *Computational and Mathematical Methods in Medicine*, 3(1):25–36, 2000.
- [86] Andrzej Swierniak, Marek Kimmel, and Jaroslaw Smieja. Mathematical modeling as a tool for planning anticancer therapy. *European journal of pharmacology*, 625(1):108–121, 2009.
- [87] Andrzej Swierniak and Jaroslaw Smieja. Analysis and optimization of drug resistant and phase-specific cancer chemotherapy models. *Mathematical biosciences and engineering: MBE*, 2(3):657–670, 2005.
- [88] Helen C Monro and Eamonn A Gaffney. Modelling chemotherapy resistance in palliation and failed cure. *Journal of theoretical biology*, 257(2):292–302, 2009.
- [89] L Norton and R Simon. Tumor size, sensitivity to therapy, and design of treatment schedules. *Cancer treatment reports*, 61(7):1307, 1977.
- [90] Larry Norton and Richard Simon. The norton-simon hypothesis revisited. *Cancer treatment reports*, 70(1):163, 1986.
- [91] Seth Michelson and Doris Slate. A mathematical model of the p-glycoprotein pump as a mediator of multidrug resistance. *Bulletin of mathematical biology*, 54(6):1023–1038, 1992.
- [92] Seth Michelson and Doris Slate. A mathematical model for the inhibition of the multidrug resistance-associated p-glycoprotein pump. *Bulletin of mathematical biology*, 56(2):207–223, 1994.
- [93] Vasiliki Panagiotopoulou, Giles Richardson, Oliver E Jensen, and Cyril Rauch. On a biophysical and mathematical model of pgp-mediated multidrug resistance: understanding the space–time dimension of mdr. *European Biophysics Journal*, 39(2):201–211, 2010.

- [94] Raja Venkatasubramanian, Michael A Henson, and Neil S Forbes. Integrating cell-cycle progression, drug penetration and energy metabolism to identify improved cancer therapeutic strategies. *Journal of theoretical biology*, 253(1):98–117, 2008.
- [95] James Greene, Orit Lavi, Michael M Gottesman, and Doron Levy. The impact of cell density and mutations in a model of multidrug resistance in solid tumors. *Bulletin of mathematical biology*, 76(3):627–653, 2014.
- [96] Valentina Fodale, Mariaelena Pierobon, Lance Liotta, and Emanuel Petricoin. Mechanism of cell adaptation: when and how do cancer cells develop chemoresistance? *The Cancer Journal*, 17(2):89–95, 2011.
- [97] URSZULA Forys and Anna Marciniak-Czochra. Logistic equations in tumour growth modelling. *International Journal of Applied Mathematics and Computer Science*, 13(3):317–326, 2003.
- [98] R Schuster and H Schuster. Reconstruction models for the ehrlich ascites tumor of the mouse. *Mathematical population dynamics*, 2:335–348, 1995.
- [99] Andrew M Stein, Tim Demuth, David Mobley, Michael Berens, and Leonard M Sander. A mathematical model of glioblastoma tumor spheroid invasion in a three-dimensional in vitro experiment. *Biophysical journal*, 92(1):356–365, 2007.
- [100] John Carl Panetta and K Renee Fister. Optimal control applied to competing chemotherapeutic cell-kill strategies. *SIAM Journal on Applied Mathematics*, 63(6):1954–1971, 2003.
- [101] Anna Kane Laird. Dynamics of tumour growth. *British journal of cancer*, 18(3):490, 1964.
- [102] Evgeniy Khain and Leonard M Sander. Dynamics and pattern formation in invasive tumor growth. *Physical review letters*, 96(18):188103, 2006.
- [103] Kyle R Brimacombe, Matthew D Hall, Douglas S Auld, James Inglese, Christopher P Austin, Michael M Gottesman, and King-Leung Fung. A dual-fluorescence high-throughput cell line system for probing multidrug resistance. *Assay and drug development technologies*, 7(3):233–249, 2009.
- [104] Rama Grantab, Shankar Sivananthan, and Ian F Tannock. The penetration of anticancer drugs through tumor tissue as a function of cellular adhesion and packing density of tumor cells. *Cancer research*, 66(2):1033–1039, 2006.
- [105] Maria Hakanson, Stefan Kobel, Matthias P Lutolf, Marcus Textor, Edna Cukierman, and Mirren Charnley. Controlled breast cancer microarrays for the deconvolution of cellular multilayering and density effects upon drug responses. *PloS one*, 7(6):e40141, 2012.

- [106] Hong Long, Hong Han, Baofeng Yang, and Zhiguo Wang. Opposite cell density-dependence between spontaneous and oxidative stress-induced apoptosis in mouse fibroblast l-cells. *Cellular Physiology and Biochemistry*, 13(6):401–414, 2003.
- [107] Motoo Nagane, Frank Coufal, Hong Lin, Oliver Bögler, Webster K Cavenee, and HJ Su Huang. A common mutant epidermal growth factor receptor confers enhanced tumorigenicity on human glioblastoma cells by increasing proliferation and reducing apoptosis. *Cancer research*, 56(21):5079–5086, 1996.
- [108] Liang Qiao and Geoffrey C Farrell. The effects of cell density, attachment substratum and dexamethasone on spontaneous apoptosis of rat hepatocytes in primary culture. *In Vitro Cellular & Developmental Biology-Animal*, 35(7):417–424, 1999.
- [109] Kumiko Saeki, Akira Yuo, Mitsuyasu Kato, Kohei Miyazono, Yoshio Yazaki, and Fumimaro Takaku. Cell density-dependent apoptosis in hl-60 cells, which is mediated by an unknown soluble factor, is inhibited by transforming growth factor β 1 and overexpression of bcl-2. *Journal of Biological Chemistry*, 272(32):20003–20010, 1997.
- [110] Valerie M Weaver, Sophie Lelievre, Johnathon N Lakins, Micah A Chrenek, Jonathan CR Jones, Filippo Giancotti, Zena Werb, and Mina J Bissell. β 4 integrin-dependent formation of polarized three-dimensional architecture confers resistance to apoptosis in normal and malignant mammary epithelium. *Cancer cell*, 2(3):205–216, 2002.
- [111] Nastaran Zahir and Valerie M Weaver. Death in the third dimension: apoptosis regulation and tissue architecture. *Current opinion in genetics & development*, 14(1):71–80, 2004.
- [112] AJ Coldman and JH Goldie. Role of mathematical modeling in protocol formulation in cancer chemotherapy. *Cancer treatment reports*, 69(10):1041–1048, 1985.
- [113] AJ Coldman and JH Goldie. A stochastic model for the origin and treatment of tumors containing drug-resistant cells. *Bulletin of mathematical biology*, 48(3-4):279–292, 1986.
- [114] JH Goldie and AJ Coldman. Quantitative model for multiple levels of drug resistance in clinical tumors. *Cancer treatment reports*, 67(10):923–931, 1983.
- [115] James H Goldie and Andrew J Coldman. *Drug resistance in cancer: mechanisms and models*. Cambridge University Press, 2009.
- [116] LE Harnevo and Z Agur. The dynamics of gene amplification described as a multitype compartmental model and as a branching process. *Mathematical biosciences*, 103(1):115–138, 1991.

- [117] Marek Kimmel and David E Axelrod. Mathematical models of gene amplification with applications to cellular drug resistance and tumorigenicity. *Genetics*, 125(3):633–644, 1990.
- [118] Brian G Birkhead, Elaine M Rankin, Stephen Gallivan, Leanne Dones, and Robert D Rubens. A mathematical model of the development of drug resistant to cancer chemotherapy. *European Journal of Cancer and Clinical Oncology*, 23(9):1421–1427, 1987.
- [119] Cristian Tomasetti and Doron Levy. An elementary approach to modeling drug resistance in cancer. *Mathematical biosciences and engineering: MBE*, 7(4):905, 2010.
- [120] Trachette L Jackson and Helen M Byrne. A mathematical model to study the effects of drug resistance and vasculature on the response of solid tumors to chemotherapy. *Mathematical biosciences*, 164(1):17–38, 2000.
- [121] Alexander Lorz, Tommaso Lorenzi, Michael E Hochberg, Jean Clairambault, and Benoit Perthame. Populational adaptive evolution, chemotherapeutic resistance and multiple anti-cancer therapies. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47(02):377–399, 2013.
- [122] Àngel Calsina and Silvia Cuadrado. A model for the adaptive dynamics of the maturation age. *Ecological Modelling*, 133(1):33–43, 2000.
- [123] Àngel Calsina and Sílvia Cuadrado. Small mutation rate and evolutionarily stable strategies in infinite dimensional adaptive dynamics. *Journal of mathematical biology*, 48(2):135–159, 2004.
- [124] Nicolas Champagnat, Régis Ferrière, and Sylvie Méléard. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. *Theoretical population biology*, 69(3):297–321, 2006.
- [125] Laurent Desvillettes, Pierre Emmanuel Jabin, Stéphane Mischler, Gaël Raoul, et al. On selection dynamics for continuous structured populations. *Communications in Mathematical Sciences*, 6(3):729–747, 2008.
- [126] Odo Diekmann, Pierre-Emanuel Jabin, Stéphane Mischler, and Benoit Perthame. The dynamics of adaptation: an illuminating example and a hamilton–jacobi approach. *Theoretical population biology*, 67(4):257–271, 2005.
- [127] Alexander Lorz, Sepideh Mirrahimi, and Benoit Perthame. Dirac mass dynamics in multidimensional nonlocal parabolic equations. *Communications in Partial Differential Equations*, 36(6):1071–1098, 2011.
- [128] P Magal and G.F. Webb. Mutation, selection and recombination in a model of phenotype evolution. *Discrete Contin. Dynam. Systems*, 6(1):221–236, 2000.

- [129] Benoît Perthame. *Transport equations in biology*. Springer Science & Business Media, 2006.
- [130] Benoit Perthame and Guy Barles. Dirac concentrations in lotka-volterra parabolic pdes. *Indiana Univ. Math. J.*, 57(7):3275–3301, 2008.
- [131] Antonija Kreso, Catherine A O’Brien, Peter van Galen, Olga I Gan, Faiyaz Notta, Andrew MK Brown, Karen Ng, Jing Ma, Erno Wienholds, Cyrille Dunant, et al. Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science*, 339(6119):543–548, 2013.
- [132] A Marusyk and K Polyak. Cancer.(2013) cancer cell phenotypes, in fifty shades of grey. *Science*, 339:528–529.
- [133] Ivana Bozic, Benjamin Allen, and Martin A Nowak. Dynamics of targeted cancer therapy. *Trends in molecular medicine*, 18(6):311–316, 2012.
- [134] Larry Norton and Richard Simon. Growth curve of an experimental solid tumor following radiotherapy. *Journal of the National Cancer Institute*, 58(6):1735–1741, 1977.
- [135] N Bellomo. Nonlinear models and problems in applied sciences from differential quadrature to generalized collocation methods. *Mathematical and Computer Modelling*, 26(4):13–34, 1997.
- [136] Orit Lavi, James M Greene, Doron Levy, and Michael M Gottesman. The role of cell density and intratumoral heterogeneity in multidrug resistance. *Cancer research*, 73(24):7168–7175, 2013.
- [137] Orit Lavi, James M Greene, Doron Levy, and Michael M Gottesman. Simplifying the complexity of resistance heterogeneity in metastasis. *Trends in molecular medicine*, 20(3):129–136, 2014.
- [138] Robert W Robey, Paul R Massey, Laleh Amiri-Kordestani, and Susan E Bates. Abc transporters: unvalidated therapeutic targets in cancer and the cns. *Anti-cancer agents in medicinal chemistry*, 10(8):625, 2010.
- [139] Christopher D McFarland, Kirill S Korolev, Gregory V Kryukov, Shamil R Sunyaev, and Leonid A Mirny. Impact of deleterious passenger mutations on cancer progression. *Proceedings of the National Academy of Sciences*, 110(8):2910–2915, 2013.
- [140] Thomas A Buchholz, Darren W Davis, David J McConkey, W Fraser Symmans, Vicente Valero, Anuja Jhingran, Susan L Tucker, Lajos Pusztai, Massimo Cristofanilli, Francisco J Esteva, et al. Chemotherapy-induced apoptosis and bcl-2 levels correlate with breast cancer response to chemotherapy. *The Cancer Journal*, 9(1):33–41, 2003.

- [141] Kwong-Kwok Wong. Recent developments in anti-cancer agents targeting the ras/raf/mek/erk pathway. *Recent patents on anti-cancer drug discovery*, 4(1):28–35, 2009.
- [142] Clement R Boland et al. Tumor hypoxia and genetic alterations in sporadic cancers. *Journal of Obstetrics and Gynaecology Research*, 37(2):85–98, 2011.
- [143] Kim Lewis. Persister cells, dormancy and infectious disease. *Nature Reviews Microbiology*, 5(1):48–56, 2007.
- [144] Andrew NJ Tutt, Conny Th M van Oostrom, Gillian M Ross, Harry van Steeg, and Alan Ashworth. Disruption of brca2 increases the spontaneous mutation rate in vivo: synergism with ionizing radiation. *EMBO reports*, 3(3):255–260, 2002.
- [145] Thomas J Klein and Peter M Glazer. The tumor microenvironment and dna repair. In *Seminars in radiation oncology*, volume 20, pages 282–287. Elsevier, 2010.
- [146] Toni Y Reynolds, Sara Rockwell, and Peter M Glazer. Genetic instability induced by the tumor microenvironment. *Cancer Research*, 56(24):5754–5757, 1996.
- [147] Oliver J Rando and Kevin J Verstrepen. Timescales of genetic and epigenetic inheritance. *Cell*, 128(4):655–668, 2007.
- [148] Hui Shen and Peter W Laird. Interplay between the cancer genome and epigenome. *Cell*, 153(1):38–55, 2013.
- [149] Marco Gerlinger, Andrew J Rowan, Stuart Horswell, James Larkin, David Endesfelder, Eva Gronroos, Pierre Martinez, Nicholas Matthews, Aengus Stewart, Patrick Tarpey, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *New England Journal of Medicine*, 366(10):883–892, 2012.
- [150] Alan Hodgkinson and Adam Eyre-Walker. Variation in the mutation rate across mammalian genomes. *Nature Reviews Genetics*, 12(11):756–766, 2011.
- [151] CY Chang, SC Lin, WH Su, CM Ho, and YS Jou. Somatic lmc1d1 mutations promoted cell migration and tumor metastasis in hepatocellular carcinoma. *Oncogene*, 31(21):2640–2652, 2012.
- [152] Wilfred D Stein, James L Gulley, Jeff Schlom, Ravi A Madan, William Dahut, William D Figg, Yang-min Ning, Phil M Arlen, Doug Price, Susan E Bates, et al. Tumor regression and growth rates determined in five intramural nci prostate cancer trials: the growth rate constant as an indicator of therapeutic efficacy. *Clinical Cancer Research*, 17(4):907–917, 2011.

- [153] Wilfred D Stein, James Yang, Susan E Bates, and Tito Fojo. Bevacizumab reduces the growth rate constants of renal carcinomas: a novel algorithm suggests early discontinuation of bevacizumab resulted in a lack of survival advantage. *The oncologist*, 13(10):1055–1062, 2008.
- [154] Krastan B Blagoev, Julia Wilkerson, Wilfred D Stein, Robert J Motzer, Susan E Bates, and A Tito Fojo. Sunitinib does not accelerate tumor growth in patients with metastatic renal cell carcinoma. *Cell reports*, 3(2):277–281, 2013.
- [155] Wilfred D Stein, Julia Wilkerson, Sindy T Kim, Xin Huang, Robert J Motzer, Antonio Tito Fojo, and Susan E Bates. Analyzing the pivotal trial that compared sunitinib and ifn- α in renal cell carcinoma, using a method that assesses tumor regression and growth. *Clinical Cancer Research*, 18(8):2374–2381, 2012.
- [156] Richard Simon and Larry Norton. The norton-simon hypothesis: designing more effective and less toxic chemotherapeutic regimens. *Nature Clinical Practice Oncology*, 3(8):406–407, 2006.
- [157] MW Retsky, DE Swartzendruber, RH Wardwell, and PD Bame. Is gompertzian or exponential kinetics a valid description of individual human cancer growth? *Medical hypotheses*, 33(2):95–106, 1990.
- [158] Caterina Guiot, Piero Giorgio Degiorgis, Pier Paolo Delsanto, Pietro Gabriele, and Thomas S Deisboeck. Does tumor growth follow a universal law? *Journal of theoretical biology*, 225(2):147–151, 2003.
- [159] D Hart, E Shochat, and Z Agur. The growth law of primary breast cancer as inferred from mammography screening trials data. *British journal of cancer*, 78(3):382, 1998.
- [160] John A Spratt, D Von Fournier, John S Spratt, and Ernst E Weber. Decelerating growth and human breast cancer. *Cancer*, 71(6):2013–2019, 1993.
- [161] Susan E Clare, Faina Nakhlis, and John Carl Panetta. Molecular biology of breast metastasis: The use of mathematical models to determine relapse and to predict response to chemotherapy in breast cancer. *Breast Cancer Research*, 2(6):430, 2000.
- [162] Paul L Crispen, Rosalia Viterbo, Stephen A Boorjian, Richard E Greenberg, David YT Chen, and Robert G Uzzo. Natural history, growth kinetics, and outcomes of untreated clinically localized renal tumors under active surveillance. *Cancer*, 115(13):2844–2852, 2009.
- [163] James M Greene, Doron Levy, King Leung Fung, Paloma S Souza, Michael M Gottesman, and Orit Lavi. Modeling intrinsic heterogeneity and growth of cancer cells. *Journal of theoretical biology*, 367:262–277, 2015.

- [164] Alexander RA Anderson and Vito Quaranta. Integrative mathematical oncology. *Nature Reviews Cancer*, 8(3):227–234, 2008.
- [165] Monika Joanna Piotrowska and Simon D Angus. A quantitative cellular automaton model of in vitro multicellular spheroid tumour growth. *Journal of theoretical biology*, 258(2):165–178, 2009.
- [166] Bryan C Thorne, Alexander M Bailey, and Shayn M Peirce. Combining experiments with multi-cell agent-based modeling to study biological tissue patterning. *Briefings in bioinformatics*, 8(4):245–257, 2007.
- [167] Le Zhang, Zihui Wang, Jonathan A Sagotsky, and Thomas S Deisboeck. Multiscale agent-based cancer modeling. *Journal of mathematical biology*, 58(4-5):545–559, 2009.
- [168] Eric Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 3):7280–7287, 2002.
- [169] JC Panetta, WE Evans, and MH Cheok. Mechanistic mathematical modelling of mercaptopurine effects on cell cycle of human acute lymphoblastic leukaemia cells. *British journal of cancer*, 94(1):93–100, 2005.
- [170] Crispin Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences (Springer Series in Synergetics)*. Springer, 4th edition, 2009.
- [171] L. P. Pitaevskii and E.M. Lifshitz. *Physical Kinetics: Volume 10 (Course of Theoretical Physics S)*. Butterworth-Heinemann, 1st edition, 1981.
- [172] Adam Shwartz and Alan Weiss. *Large Deviations for Performance Analysis: QUEUES, Communication and Computing (Stochastic Modeling Series)*. Chapman and Hall/CRC, 1st edition, 1995.
- [173] A.M. Mathai and Serge B. Provost. *Quadratic Forms in Random Variables (Statistics: A Series of Textbooks and Monographs)*. CRC Press, 1st edition, 1992.
- [174] Alfredo H-S. Ang and Wilson H. Tang. *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering (v. 1)*. Wiley, 2nd edition, 2006.
- [175] V. Lakshmikantham and M.R.M. Rao. *Theory of Integro-Differential Equations (Stability and Control)*. CRC Press, 1st edition, 1995.
- [176] Uri M. Ascher. *Numerical Methods for Evolutionary Differential Equations (Computational Science and Engineering)*. Society for Industrial and Applied Mathematics, 1st edition, 2008.

- [177] Geoffrey R. Grimmett and David R. Stirzaker. *Probability and Random Processes*. Oxford University Press, 3rd edition, 2001.
- [178] Daniela Morale. Modeling and simulating animal grouping: individual-based models. *Future Generation Computer Systems*, 17(7):883–891, 2001.
- [179] Martin Burger, Vincenzo Capasso, and Daniela Morale. On an aggregation model with long and short range interactions. *Nonlinear Analysis: Real World Applications*, 8(3):939–958, 2007.
- [180] Daniela Morale, Vincenzo Capasso, and Karl Oelschläger. An interacting particle system modelling aggregation behavior: from individuals to populations. *Journal of mathematical biology*, 50(1):49–66, 2005.
- [181] Alastair JJ Wood, Eric K Rowinsky, and Ross C Donehower. Paclitaxel (taxol). *New England Journal of Medicine*, 332(15):1004–1014, 1995.