

Politecnico di Torino

Master's Degree in Physics of
Complex Systems



Master's Degree Thesis

A Minimal Model of
Inoculum-dependent Growth in Cancer
Cell Cultures

Supervisors:
Carla Bosia, Andrea De Martino

Candidate:
Mattia Tarchini Bojczuk

December 2022

Abstract

Through fission cells are known to increase their numbers following a well-known cycle of growth. When this process is recreated in the laboratory several factors and initial conditions are in control of the culturer, among which the initial density of the population: the inoculum size. Inoculum dependent traits have been observed in bacterial [6, 9], insect [10] and plant cells [12, 14] and in many cases such dependencies should be expected: when population size is limited by the environment, its growth potential can be drastically reduced if the starting density is close to the maximum one. Several studies have also shown that growth rates of mammalian cell populations can exhibit positive correlations with the inoculum size while in some cases growth is completely absent for initial densities under a certain threshold [6, 19]. In [2] a comprehensive assay of how inoculum size affects growth profiles of cancer cells is provided: by following the growth of hundreds of populations of cancer cells, the authors found that the time they need to adapt to the environment decreases as the initial cell density increases. Moreover, the population growth rate showed a maximum at intermediate initial densities. Traditional models of growth are not able to explain the observed non-monotonic dependence of maximum growth rate with starting cell count.

The aim of this thesis is to develop a minimal model that contains known features of mammalian cell population growth and is able to provide testable predictions of inoculum dependent growth-rates. Since growth hormones stimulation is essential to mammalian cell proliferation, and its presence has been involved in observations of cell density dependent growth rates [35], one of the main processes we attempt to include in our new description is a form of cooperation through chemical signaling. With stochastic simulations, growth trajectories are obtained for different initial densities and qualitatively similar growth rate modulations are measured.

Contents

1	Introduction	9
1.1	Physics and the Cell	9
1.2	Cell Cultures	11
1.2.1	Fitting the Growth Curve	11
1.2.2	Models of growth	13
1.2.3	Inoculum density effects	15
1.3	Initial Cell Density Encodes Proliferative Potential in Cancer Cell Populations	18
1.4	Aim of the Thesis	21
2	Methods	22
2.1	Stochastic Chemical Reactions Formalism	22
2.2	Master Equation for the process	23
2.3	The Gillespie algorithm	24
2.4	Linear Birth Death process	25
2.4.1	Testing the SSA	27
3	Model Formulation and Analysis	30
3.1	Limited Environment and Multiple Populations	30
3.1.1	Results	37
3.2	Chemical Signaling	44
3.2.1	Analysis	46
3.2.2	Results	48
4	Conclusions	54
4.1	Comparison with Experimental Data	55
4.2	Discussion	58

5	Appendix	59
5.0.1	A	59

List of Figures

1.1	A growth curve	12
1.2	(a) Comparison between the behavior of λ_{max} (maximum value of the growth rate) vs N_0 (the size of the inoculum) for the Logistic and the weakly cooperative Allee models. b is the exponent that defines how strong the Allee effect is (a) N_0 axis is logarithmic (b) Same comparison but the N_0 axis is in linear scale	15
1.3	(a) Representative growth curve. The fitting parameters are t_{lag} , A and λ_{max} . The maximum growth rate λ_{max} corresponds to the slope of the tangent to the inflection point (orange line) of the fitting curve (gray curve). The intersection of such a tangent with the time-axis gives the lag time t_{lag} , while its intersection with the line $\ln(N/N_0) = A$ yields the time of exit from the exponential phase, t_{log} . (b) Empirical distribution of the maximum growth rate. (c) Maximum growth rate as a function of the inoculum density N_0 . Orange dots represent parameter estimates from individual experiments with their standard errors, while red dots represent the mean values of λ_{max} . The dashed vertical line marks the value of the mean carrying capacity. The blue line denotes the behavior of λ_{max} vs N_0 expected on the basis of a purely logistic model with $k = 8.6 \times 10^6$ and intrinsic maximal growth rate $r = 0.029$ defined in section 1.2.2	19
2.1	Average population number	28
2.2	Variance	29

3.1	Minimum Carrying Capacities We show as a function of the ratio r_1/r_2 the quantity k_{mi} as defined in equation (3.26) , for different values of N_0 ($N_{01} = N_{02} = 1/2N_0$)	36
3.2	Comparison Due to execution times that scale like $t_{exec} \sim \exp(N_{tot})$ the carrying capacity is set to $k = 10^4$. Growth trajectories are obtained for two different values of initial conditions $N_{01} = N_{02} = 10$ and $N_{01} = N_{02} = 500$. a) $\langle N(t) \rangle$ is an average over 1 sample trajectory b) average is over 10 samples trajectories generated by the algorithm, it is possible to see a closer similarity between the results of the two methods for larger N_0 c) 100 sample trajectories	38
3.3	Exponential Phase Comparison: numerical solutions for the system (3.12) for $N_{01}, N_{02} = 10, 100, 500$ with parameters $k = 10^5$ $r_2 = 0.1, r_1 = 1.5r_2$ in (a-b) ,and $k = 10^5$ $r_2 = 0.1, r_1 = 5r_2$, in (d-e). In (c) is shown difference between the two measurements for $N_{01} = 10$ and $N_{02} = 10$, $r_2 = 0.1, r_1 = 1.5r_2$ (f) $k = 10^5$, $r_2 = 0.1, r_1 = 5r_2$	40
3.4	Initial Conditions: Colored curves represent the set of points for which $N_{01} + N_{02} = N_0$. The markers are the actual samplings performed on these curves to obtain the initial conditions of our simulations.	42
3.5	Λ_{max} vs N_0 :results of simulations obtained for different values of $r_1 - r_2$, the intrinsic maximal growth rates of the two subpopulations with. Error bars represent standard deviations of the average values	43
3.6	Varying k: we show results for same parameters values $r_1 = r_2 = 0.1$ but different carrying capacity k . a) $k = 10^6$. b) $k = 10^5$	44
3.7	Growth Rates Numerical solutions,growth curves,and comparison between Λ and Λ_t for $N_{01}, N_{02} = 10, 100, 500$ with parameters $k = 10^6$ $r_1 = 0.3, r_2 = 0.1, r_3 = 0.7$ in (a-b-c) ,and $k = 10^6$ $r_1 = 0.5, r_2 = 0.1, r_3 = 0.7$ in (d-e-f). In (c-f) is shown difference between the two measurements for $N_{01} = 10$ and $N_{02} = 10$	48

3.8	Selected Results we plot for the fixed combination of parameters values $r_1 = 0.6, r_2 = 0.1, r_3 = 0.9, k = 10^6$ measurements of $\Lambda_{max}(N_0)$ for a closely spaced range of values of the "contagion" rate, c_7 . (a) $c_7 = 1 \times 10^{-4}$ (b) $c_7 = 7 \times 10^{-5}$ (c) $c_7 = 3 \times 10^{-5}$ (d) $c_7 = 1 \times 10^{-5}$	50
3.9	Effects of varying carrying capacity and contagion rate on $\Lambda_{max}(N_0)$ profile In each figure we plot three different $\Lambda_{max}(N_0)$ curves that correspond to a different value of k , the carrying capacity.(a) $c_7 = 1 \times 10^{-4}$ (b) $c_7 = 7 \times 10^{-5}$ (c) $c_7 = 3 \times 10^{-5}$ (d) $c_7 = 1 \times 10^{-5}$	51
3.10	Varying r_1 In each sub-figure for a fixed values of contagion rate, carrying capacity, and r_3 we vary the values of the "phenotype" of sub-population N_1 : its intrinsic maximal growth rates r_1 . Each different curve correspond to a different value of $r_1 \in \{0.2, 0.3, 0.4, 0.5, 0.6\}$.(a) $c_7 = 1 \times 10^{-4}$ (b) $c_7 = 7 \times 10^{-5}$ (c) $c_7 = 3 \times 10^{-5}$ (d) $c_7 = 1 \times 10^{-5}$	52
3.11	Varying r_3 In each sub-figure for a fixed values of contagion rate, carrying capacity, and r_1 we vary the values of the "phenotype" of sub-population N_3 : its intrinsic maximal growth rates r_3 . Each different curve correspond to a different value of $r_3 \in \{0.2, 0.3, 0.4, 0.5, 0.6\}$.(a) $c_7 = 1 \times 10^{-4}$ (b) $c_7 = 7 \times 10^{-5}$ (c) $c_7 = 3 \times 10^{-5}$ (d) $c_7 = 1 \times 10^{-5}$	53
4.1	a) Λ_{max} is the average maximal growth rate measured and averaged over Gillespie generated samples. Bars represent the standard deviation of the measurement. The dotted line marks the average maximum recorded in vitro. b) Experimental data: orange dots represent parameter estimates from individual experiments with their standard errors, while red dots represent the mean values of λ_{max} . In both (a-b) the blue line denotes the behavior of λ_{max} vs N_0 expected on the basis of a purely logistic model with $k = 8.6 \times 10^6$ and intrinsic maximal growth rate $r = 0.029$	56

Chapter 1

Introduction

1.1 Physics and the Cell

Physics and biology have for a long time entertained a fruitful relationship culminating in the synthesis of prolific new fields of research such as biophysics and bioengineering. In addition to the fundamental technologies that advanced the study of life in the laboratory, physics, with its approach based on fundamental quantitative principles, has in the last half-century significantly contributed to the understanding of living matter.

Cells are the fundamental units of life and their basic functions are to grow, replicate, and multiply. On one side these biological processes are of physical interest because some of their parts include inherently physical problems. Molecular motors and mechanical sensing are some examples of areas in the study of a cell where energy conversion into work, diffusion of chemical species, and action of physical forces are the characterizing phenomena.

On the other hand, even in its constitutive elements, biological complexity is often unsormountable. With new high-throughput methods, we are now able to measure protein levels and concentrations of DNA transcripts, correlating their possible values to different physiological states of the cell (its phenotypes). The metabolic activity can thus be in principle measured and converted in massive amounts of data, but if we wanted to understand a single cellular function, (for instance, the ability of a eukaryotic cell to move along a substrate) we would have to deal with a very intricate and often incomprehensible biochemical reaction network. In this regard, the demand to reduce complexity is high, and a coarse-grained modeling approach that ab-

stracts from the underlying biological details is poised to provide the needed simplified functional picture of biological pathways.

In addition to the deterministic laws that might govern the inner workings of a cell, one has to take into account the fact that virtually all cellular processes are affected by noise. In this regard, cellular growth involves stochastic phenomena on scales ranging from the molecular to the macroscopic. The growth dynamics of a population can be naturally described by a master equation and understood in a framework akin to the statistical physics one. If indeed there is a possibility to derive complex macroscopic behaviors from a simple set of microscopic rules, one would spontaneously be oriented to use the tools and techniques developed in statistical mechanics. In this thesis project, we set out to explore such possibility in the context of cancer cell growth, trying to understand in which way the long-term evolution of a system comprised of self-replicating agents might be affected by the interaction among its constitutive elements and their initial densities.

1.2 Cell Cultures

In biology, cell culture is one of the fundamental techniques employed in the laboratory. It is generally intended as the practice of fostering the proliferation of biological material in an artificial environment. Typically, the successful culturing of cells is subject to conducive environmental conditions such as a suitable temperature, the presence of a substrate (a surface cells need to be anchored to for survival), and an appropriate growth medium. The growth medium is a liquid or a gel generally composed of nutrients and other chemical compounds (e.g. amino acids, vitamins, glucose) that support and regulate the growth cycle, and can maintain an optimal level of pH. Specifically for the *in vitro* cultivation of mammalian cells, there is a serum requirement: the medium must contain a source of growth factors (specialized and diffusable signaling proteins that stimulate growth). Artificial media, the chemical composition of which is precisely known, are preferred in place of natural media: naturally occurring biological fluids that do not allow for reproducibility. For example, among serum-containing artificial media, one the most commonly used is medium with the addition of Fetal Bovine Serum. Essentially, experimental work aimed at growing cells in culture is punctuated by a few main steps. An initial amount of cells, the *inoculum*, is sampled from a culture in a known pre-growth phase, then seeded on new culture plates where it begins to grow after some time. The total population number is then assayed and monitored, with the technique of choice, at regular intervals. By plotting the logarithm of this number as a function of time one obtains a typical sigmoid curve, known as a growth curve. It is possible, by looking at a growth curve, to subdivide the growth cycle into four different phases.

1.2.1 Fitting the Growth Curve

Given growth data one can in principle measure the *specific growth rate* of the population:

$$\lambda = \frac{BiomassProduced}{Time} = \frac{\log(N/N_0)}{t} \quad (1.1)$$

The value of λ (measured in h^{-1}) is expected to grow from zero to a maximal value λ_{max} in a given period of time, commonly referred to as *Lag Time*. Eventually the specific growth rate will reach again a value of zero, so

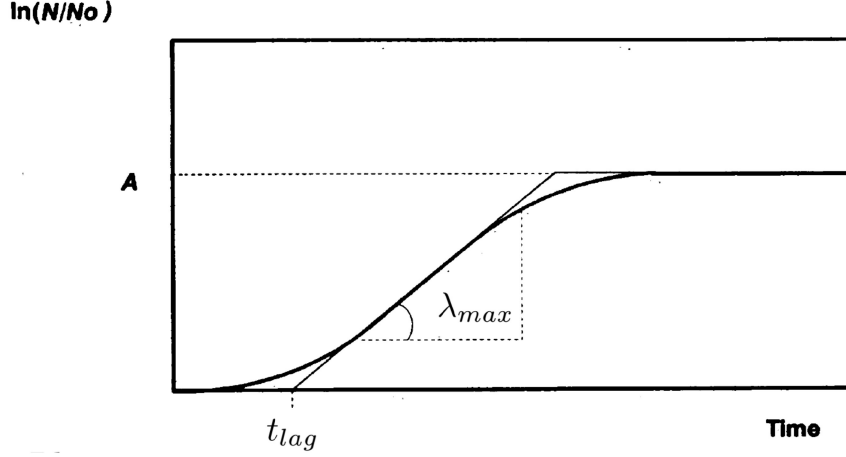


Figure 1.1: A growth curve

that the logarithm of the population will reach the asymptotic value:

$$A = \log\left(\frac{N(\infty)}{N_0}\right) \quad (1.2)$$

Under the assumption that the population is growing exponentially at the time when the growth rate is maximum, λ_{max} is usually defined as the value of the slope of the tangent line at the inflection point of the growth curve. The intersect of the tangent with the time axis is commonly identified as the value of the *Lag Time* t_{lag} . The maximum growth rate is also used often to obtain the cell doubling time (or generation time, GT): the time required for a cell population to double its size $GT = \frac{\log(2)}{\text{growthrate}}$.

To fit properly the data to a model, sigmoid curves are reparametrized in order to include variables that represent meaningful biological quantities as they are defined in the chosen model. An example is the sigmoid:

$$\ln \frac{N}{N_0} = \frac{A}{1 + \exp \left[\frac{4\lambda_{max}}{A} (t_{lag} - t) + 2 \right]} \quad (1.3)$$

Equation 1.3 is a sigmoid modified for a logistic model, with fitting parameters t_{lag} , A and λ_{max} . $A = \log(k/N_0)$, where k is the carrying capacity

in the logistic model and λ_{max} would corresponds to formula 1.7 for $b = 0$ defined in the next section.

Given the equation 1.3 one can for example apply a non linear least-squares method to infer growth parameters that encapsulate the relevant information describing the different phases of growth.

Some of the models focus exclusively on the log phase or log and stationary phase while others take also take into account the lag phase. In fact it is possible to include in a model variables that encode the physiological state of the cell, and express the ratio of the lag time to the generation time as a function of these variables [1]. To describe the influence of the environment on growth, (through temperature, nutrient concentration, pH) secondary models are developed. Again, some of these models treat the lag time and the maximum growth rate as independent quantities while others are based on the general assumption that the lag time is proportional to the generation time [3].

In the following we review the most commonly adopted models of growth, detailing , as it will be relevant in the next sections, which of these includes a prediction of the influence of the initial size of the population on the maximum specific growth rate.

1.2.2 Models of growth

The log (or exponential) phase of cellular growth has been extensively and successfully modeled under different conditions, including cases in which the population is not only multiplying but also possibly cooperating or competing for resources.

Some of these models are able to predict how the growth rate of a population growing exponentially will change while varying the size of the initial population N_0 .

Exponential - Malthus Simple exponential growth is governed by the differential equation

$$\frac{dN(t)}{dt} = rN(t) \quad (1.4)$$

with solution

$$\log\left(\frac{N(t)}{N_0}\right) = rt \quad (1.5)$$

In this case the growth rate remains constant and the population grows exponentially indefinitely, unaffected by a change in N_0 .

Logistic and Allee Models The logistic model adds a carrying capacity k , effectively slowing growth and reproducing the effects of limited space and resources. The Allee model introduces a "cooperative" effect by inserting a positive dependence on population size. Both models can be described by:

$$\frac{dN(t)}{dt} = Nr \left(1 - \frac{N(t)}{k} \right) \left(\frac{N(t)}{k} \right)^b \quad (1.6)$$

where r is the intrinsic maximal population growth rate (the same of Malthusian growth) and $b \geq 0$ is a numerical exponent (returning the logistic model for $b = 0$ and the weakly cooperative Allee model for $b > 0$). The right hand side of equation (1.6) is a function of N , which we will define as $\lambda(N(t))$. Thus for a value N^* $\lambda(N)$ reaches a maximum $\lambda_{max}(N^*)$. This value can be found by setting $\frac{d}{dN} \lambda(N)$ evaluated at N^* equal to 0, and corresponds to the population density at the inflection point of the growth curve, namely $N^* = \frac{kc}{1+b}$. For initial conditions $N_0 < N^*$, the inflection point occurs at some $t > 0$ and λ_{max} is simply obtained by evaluating $\lambda(N)$ at $N = N^*$. If N_0 is larger than N^* , however, the inflection point occurs at a negative time, implying that the growth rate is maximum at $t = 0$. In summary,

$$\lambda_{max} = \begin{cases} r \frac{b^b}{(1+b)^{1+b}} & \text{for } N_0 \leq \frac{b}{1+b}k \\ r \left(1 - \frac{N_0}{k} \right) \left(\frac{N_0}{k} \right)^b & \text{otherwise} \end{cases} \quad (1.7)$$

for general $b \geq 0$, while $\lambda_{max} = r(1 - N_0/k)$ for the purely logistic model ($b = 0$). From a qualitative viewpoint the two models are however very similar since, in both cases, λ_{max} is largest at small N_0 and decreases (monotonically for $b = 0$, following a plateau for $b > 0$) as N_0 increases.

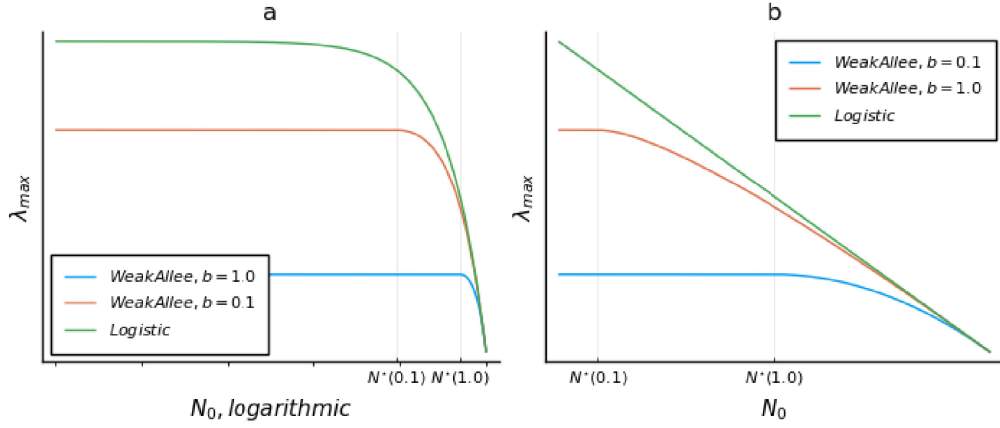


Figure 1.2: (a) Comparison between the behavior of λ_{max} (maximum value of the growth rate) vs N_0 (the size of the inoculum) for the Logistic and the weakly cooperative Allee models. b is the exponent that defines how strong the Allee effect is (a) N_0 axis is logarithmic (b) Same comparison but the N_0 axis is in linear scale

1.2.3 Inoculum density effects

In addition to the expected dependencies described in the previous section, several studies have shown that growth rates of mammalian populations exhibit also a positive correlation with the size of their initial density. In some cases it is found that below a critical threshold of initial seeding growth is not observed ([5],[17]). Rein and Rubin pioneered the field when they grew chick embryo fibroblasts (a type of cell that contributes to the formation of connective tissue) for varying initial densities in different media.

In a serum-treated medium, the growth rate of the population decreased as the inoculum density was reduced while in a conditioned medium (containing secretions of cells that were previously grown on it) there was no observed effect dependent on the inoculum size.

Tumor growth curves usually assumed to be the result of exponential growth, exhibited inoculum-dependent patterns in studies where low cell densities resulted in increased cancer cell generation times and decreased specific growth rates [15]. In the same study, breast cancer primary and metastatic tumor take rates (the percentage of tumor samples that survives) were found to be higher when the number of implanted cells was smaller. Likewise, tumor onset times and doubling times turned out to be negatively correlated with N_0 ,

so smaller inocula implied longer latency periods and slower growth rates. Both cancer and non-cancer cells are known to self-produce and share growth factors that, as we will discuss further in the following, are believed to be the biological basis for the aforementioned effects.

Individual Cells and Noise Because of the technical limitations in observing the growth of individual cells, conventional cell growth studies rely on large population numbers for their measurements and fail to consider the possible effects of single-cell heterogeneities on population growth. With modern experimental setups developed and perfected in the last twenty years, it is now possible to measure these behavioral differences and inspect whether they have any influence on the dynamics of the total population.

The processes that underlie the reproduction of the cell are fundamentally regulated by cell metabolism, gene expression, and regulation: a complex system of chemical reactions that are stochastic in nature. How this intrinsic source of noise and variability propagates through the microorganism and consequently in a population has been an intense subject of research in recent years. For example, it has been observed that for small numbers of seeding phenotypic heterogeneity inside a genetically uniform population can lead to significant differences in population growth behaviors [9]. It has also been shown that fluctuations in the expression of enzymes central to the cell metabolism can cause growth fluctuations that propagate back into the single cells, generating additional stochasticity in gene expression [37]. Phenotypic variability can be also directly incorporated in expressions for population growth parameters [24] and identified as a fundamental tool of fitness maximization [25].

The origins and evolutionary advantages of these population distributions have recently been elucidated by applying statistical principles such as entropy maximization to very general models of cellular growth [32]. Because large numbers are expected to average out single-cell heterogeneities, it is not surprising that in the context of stochastic variability, significant fluctuations around these averages and a considerable change of their value are observed when dealing with population sizes that fall below a particular threshold. Stochastic models trained on empirical distributions of single cell growth parameters are able to circumvent experimental limits and can predict how phenotypic variability can significantly increase for small inocula

[18], [8] and completely disappear for larger ones [9].

Theoretical frameworks of stochastic character have been developed in order to understand whether such inoculum-dependent behaviors are due to fluctuations in the population or to deterministic cooperative mechanisms [21].

1.3 Initial Cell Density Encodes Proliferative Potential in Cancer Cell Populations

In [2] the authors provide a study that characterizes the growth of a cell population with quantitative accuracy across a broad range of initial densities. They performed a set of batch culture experiments, recording the growth dynamics of 217 populations (Jurkat and K562, two commonly used cancer cell lines) starting from initial densities ranging over 5 orders of magnitude (from $N_0 \approx 102 \text{ cells/ml}$ to $N_0 \approx 7106 \text{ cells/ml}$). Growth curve parameters, representing respectively the lag time t_{lag} , the maximum growth rate λ_{max} and the carrying capacity k through the quantity $A = \log(k/N_0)$, are then inferred from the growth curves using the same fitting function described in section 1.2.1.

We present in the following the results obtained for the lag time and the maximum growth rate.

lag time

- **mean behavior:** the mean $\langle t_{lag} \rangle$ (averaged over experiments) was found to have a marked decreasing linear dependence on $\log(N_0)$.
- **fluctuations:** the fluctuations of the parameter were then quantified by the empirical coefficient of variation (CV):

$$CV = \frac{Std(t_{lag})}{\langle t_{lag} \rangle} \quad (1.8)$$

that remained approximately constant across 4 orders of magnitude in N_0 .

maximum growth rate

- **mean behavior:** In a regime of low initial densities (i.e. $N_0 \ll k$) the mean growth rate, measured as the maximum slope of the growth curve, was found to be roughly constant, in agreement with the predictions of the Logistic and Allee models. What was surprisingly observed instead, is the presence of a maximum of λ_{max} at intermediate initial population size. Finally, as N_0 approaches the carrying capacity, λ_{max} decreases approximately linearly with N_0 , in agreement with the logistic and Allee models.

- **fluctuations:** the distribution of values of lambda max presented significant variability (up to about 50%). For small enough inocula sample-to-sample is generically small, i.e. different populations grow at similar rates. As N_0 increased, it was observed a significant enhancement of fluctuations, which noticeably only occur above the reference level given by the asymptotic growth rate.

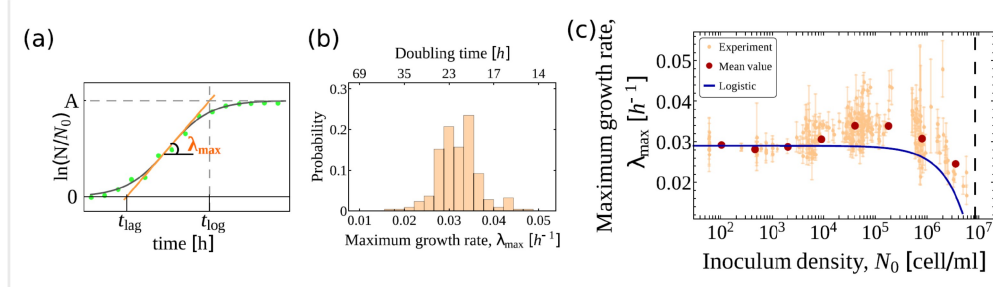


Figure 1.3: (a) Representative growth curve. The fitting parameters are t_{lag} , A and λ_{max} . The maximum growth rate λ_{max} corresponds to the slope of the tangent to the inflection point (orange line) of the fitting curve (gray curve). The intersection of such a tangent with the time-axis gives the lag time t_{lag} , while its intersection with the line $\ln(N/N_0) = A$ yields the time of exit from the exponential phase, t_{log} . (b) Empirical distribution of the maximum growth rate. (c) Maximum growth rate as a function of the inoculum density N_0 . Orange dots represent parameter estimates from individual experiments with their standard errors, while red dots represent the mean values of λ_{max} . The dashed vertical line marks the value of the mean carrying capacity. The blue line denotes the behavior of λ_{max} vs N_0 expected on the basis of a purely logistic model with $k = 8.6 \times 10^6$ and intrinsic maximal growth rate $= 0.029$ defined in section 1.2.2

λ_{max} profile The peculiar non-monotonic modulation of λ_{max} cannot be captured by conventional deterministic models like the logistic or Allee model. The positive feedback between N_0 and λ_{max} suggests that cooperative effects are involved. After excluding mechanical interaction as the mechanism at the heart of cooperation the authors point to the fact that intracellular chemical signaling seems to be the best next candidate, specifically via growth factors. This hypothesis is reinforced by the presence of a significant lag time (that

suggests some degree of conditioning in the medium of the mother population).

It is then argued by the authors that two mechanisms are at play in shaping the profile of λ_{max} : weakly cooperative interaction mediated by growth factors and the limitation on growth imposed by the carrying capacity on higher densities. It is speculated that if the carrying capacity is augmented the maximum growth rate versus N_0 will achieve a plateau before decreasing.

Relationship between parameters Through a minimal model of delayed exponential growth for a bipartite population and cues from the fitting function, the authors obtain a relation linking the two main growth parameters.

$$\lambda_{max} t_{lag} \simeq \ln \left[\left(\frac{k}{N_0} \right)^\alpha - 1 + p \right] - \ln p \quad (1.9)$$

where $\alpha \equiv (1 + e^2)^{-1} \simeq 0.12$ and p is the value of the fraction of “fast adapters”, the portion of individuals that are assumed to begin growth right after inoculation. Empirical results are captured remarkably well for $p \approx 0.4$ but not their variability. Nonetheless, it is remarkable how differences in single-cell adaptation times are crucial in explaining the interrelations between the two parameters describing the population.

Population Ensemble Trade-Off An additional result that was also reproduced by the model described above is a peculiar relationship between the mean lag time and the standard deviation of the maximal growth rate estimated across multiple populations.

It suggests the presence of a trade-off between the relaxation times and the magnitude of fluctuations of λ_{max} . The former generally encodes the ability of a cell to adapt to a new environment while the latter is believed to be a useful tool in the use of its resources.

1.4 Aim of the Thesis

As it has been explained so far, traditional deterministic models of exponential growth are unable to predict the inoculum dependence observed by Bena and colleagues. While parts of the trend can be thought to be the result of more thoroughly understood causes (limitation of resources for N_0 close to the carrying capacity), the increase of lambda max with initial cell density Is more difficult to explain. The biological basis of known positive feedbacks between N_0 and λ_{max} has been in several studies ([5],[17],[35])identified in the production of autocrine (i.e. secreted by the cells themselves) growth factors.

Because both Jurkat and K562 are known to produce such growth factors, a positive correlation between λ_{max} and N_0 can in this case be expected.

In this work we are thus interested in developing a minimal stochastic model, based on a birth-death process, that takes into account the finiteness of the environment, through a carrying capacity, and can give quantifiable fluctuations.

More importantly: we also aim at including the microscopic signaling mechanism we mentioned in section 1.3, by linking the amount of growth factor present in the environment to the size of the population, we formulate a model that incorporates the essential biological aspects of the problem.

We run simulations of the model using the Gillespie algorithm, exploring the possible relationships among model parameters.

We then compare qualitatively the results with the behaviors observed in [2], in an attempt to gain useful insights into the dynamics of a population of cancer cells.

In the next section we discuss the tools and the methods employed in this project.

Chapter 2

Methods

We'll develop the model starting from the most basic BD process. These processes involve the generation (birth) and destruction (death) of a certain quantity (the number of cells in our case) in a totally analogous way to chemical species of reactants. We will then first develop a chemical reaction formalism in order to describe these processes.

2.1 Stochastic Chemical Reactions Formalism

We can assume a generic chemical reaction to be characterized by a reaction probability per unit time. *Example* : Given a reaction



we say it exists a constant c_1 , and $c_1 \times dt$ is the average probability that any particular pair of species S_1 and S_2 will react like R_1 in the infinitesimal interval $(t, t+dt)$.

Then, the probability that the reaction R_1 happens in the volume V containing an amount of species N_1 and N_2 for species S_1 and S_2 respectively, in the infinitesimal interval $(t, t+dt)$ will be

$$N_1 N_2 c_1 dt \quad (2.2)$$

In general, in a volume V (assumed to be spatially homogeneous), there are N_i (numbers of molecules of species S_i where $i=1,..L$) that can interact

according to M specific reactions $R_\mu (\mu = 1, 2, \dots, M)$.
For each of these reactions, we have a constant

$$c_\mu dt \quad (2.3)$$

And in the volume

$$n_\mu c_\mu dt = a_\mu dt \quad (2.4)$$

where n_μ is the product of all the numbers of molecules N_i involved in the reaction R_μ , and a_μ is the full volumic reaction rate.

2.2 Master Equation for the process

If we were to write a Master Equation for the process it would be:

$$\frac{d}{dt} P(\vec{N}, t) = \sum_{\mu=1}^M P(\vec{N}', t) c_\mu n'_\mu - \sum_{\mu=1}^M c_\mu n_\mu P(\vec{N}, t) \quad (2.5)$$

Where $\vec{N}' = (N'_1, \dots, N'_L)$ is the configuration that can reach $\vec{N} = (N_1, \dots, N_L)$ through a reaction R_μ and M is the number of reactions allowed in the system.

n'_μ, n_μ are the products of all the amounts N'_i and N_i involved in a reaction R_μ .

Solving the master equation would amount to obtain the full probability distribution of the process. With it, it's possible to compute the dynamics of any moment of the distribution and of functions depending on them, thus obtaining a full description of the quantities involved in the processes. It is not however always the case that the master equation can be solved analytically, and several approximation methods are often employed.

Another equally valid alternative is using stochastic simulation methods based on the extraction of the random sequence of events that make up the process.

In this way, the algorithm can give us distinct realizations of the process, effectively sampling the distribution. We can use these samples to obtain the empirical moments and their dynamics. The method we will employ in our discussion is the Gillespie algorithm.

2.3 The Gillespie algorithm

We define a *reaction probability density function* $P(\tau, \mu)d\tau$ as the probability that given the state $\vec{N} = (N_1, \dots, N_L)$ the next reaction will occur in the infinitesimal time interval $(t + \tau, t + \tau + d\tau)$ **and** the reaction will be an R_μ reaction. An analytic expression for $P(\tau, \mu)d\tau$ is the following:

$$P(\tau, \mu)d\tau = P_0(\tau)a_\mu d\tau \quad (2.6)$$

where $P_0(\tau)$ is the probability that given the state $\vec{N} = (N_1, \dots, N_L)$ at time t no reaction will occur in the time interval $(t, t + \tau)$ and $a_\mu d\tau = n_{mu}c_{mu}d\tau$ is the probability that an R_μ reaction will occur in the volume V in the time interval $(t + \tau, t + \tau + d\tau)$. It's possible to show that $P_0(\tau)$ takes the form

$$P_0(\tau) = \exp \left[- \sum_{\nu=1}^M a_\nu \tau \right] \quad (2.7)$$

With the aid of a random number generator it is now possible to reproduce the subsequent extraction of the two pieces of information we need to simulate the system:

- τ : the time of the next reaction
- μ : the specific reaction that will occur in the volume.

If r_1 and r_2 are the random numbers we extracted from a uniform distribution U on the interval $[0, 1)$ then for $a_0 \equiv \sum_{\nu=1}^M a_\nu \equiv \sum_{\nu=1}^M h_\nu c_\nu$ we have :

$$\tau = (1/a_0) \ln (1/r_1) \quad (2.8)$$

and μ is the integer for which

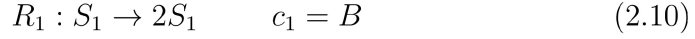
$$\sum_{\nu=1}^{\mu-1} a_\nu < r_2 a_0 < \sum_{\nu=1}^{\mu} a_\nu \quad (2.9)$$

Subsequently, we update the values of the configuration $\vec{N} = (N_1, \dots, N_L)$ according to reaction R_μ and update the current value of t as $t = t + \tau$.

We repeat the process until the moment t_{end} we want to simulate the system.

2.4 Linear Birth Death process

We begin by describing a linear Birth-Death stochastic process, for the growth of a population N . When formulated as chemical reactions the processes of Birth (creation) and Death (destruction) read as follows:



We write the Master equation for the process

$$\frac{d}{dt}P(N, t) = \sum_{\mu=1}^M P(N', t) n'_\mu c_\mu - \sum_{\mu=1}^M n_\mu c_\mu P(N, t) \quad (2.12)$$

In this case $M = 2$ and we have

$$\frac{d}{dt}P(N, t) = B(N-1)P(N-1, t) + D(N+1)P(N+1, t) - NP(N, t)(B+D) \quad (2.13)$$

Over time the mean of the population will evolve like

$$\frac{d\langle N(t) \rangle}{dt} = \frac{d}{dt} \sum_{N=0}^{\infty} NP(N, t) = \frac{d}{dt} \sum_{N=1}^{\infty} NP(N, t) \quad (2.14)$$

and by using the master equation we would obtain

$$\begin{aligned} \frac{d\langle N \rangle}{dt} &= \sum_{N=1}^{\infty} (BN(N-1)P(N-1, t)) + \sum_{N=1}^{\infty} (DN(N+1)P(N+1, t)) - \\ &\quad - \sum_{N=1}^{\infty} ((B+D)N^2P(N, t)) \end{aligned} \quad (2.15)$$

using

$$\sum_{N=1}^{\infty} BN(N-1)P(N-1) = \sum_{N=0}^{\infty} B(N+1)NP(N) \quad (2.16)$$

and

$$\sum_{N=1}^{\infty} DN(N+1)P(N+1) = \sum_{N=0}^{\infty} D(N-1)NP(N) \quad (2.17)$$

we get

$$\frac{d\langle N \rangle}{dt} = \sum_{N=0}^{\infty} P(N) (BN(N+1) + DN(N-1) - N^2(B+D)) \quad (2.18)$$

And finally

$$\frac{d\langle N(t) \rangle}{dt} = (B - D)\langle N(t) \rangle \quad (2.19)$$

that for $(B - D) = r$ corresponds to the equation describing Malthusian growth. Given an initial condition $P(N = N_0, 0) = 1$ we can solve the differential equation for the average and obtain

$$\langle N \rangle = n_0 \exp(B - D)t \quad (2.20)$$

Using the same approach we find the evolution in time of the average of N^2

$$\frac{d\langle N^2 \rangle}{dt} = 2(B - D)\langle N^2(t) \rangle + (B + D)\langle N(t) \rangle \quad (2.21)$$

This is an ODE that for $\langle N^2 \rangle = n_0^2$ at $t = 0$ has solution

$$\langle N^2 \rangle = n_0 \frac{B + D}{B - D} \exp[(B - D)t] (\exp((B - D)t) - 1) + n_0^2 \exp[2(B - D)t] \quad (2.22)$$

We can now compute the variance:

$$Var(N) = \langle N^2 \rangle - \langle N \rangle^2 = n_0 \frac{B + D}{B - D} \exp[(B - D)t] (\exp[(B - D)t] - 1). \quad (2.23)$$

For $B > D$ and sufficiently long times the variance increases as $Var(N) \sim \exp[2(B - D)t]$

so that the Standard Deviation

$$Std(N) = \sqrt{Var(N)} \sim \exp[(B - D)t] \quad (2.24)$$

increases in time in the same order of $\langle N \rangle$.

2.4.1 Testing the SSA

A simple pseudocode that implements the Gillespie algorithm simulating a linear Birth-Death process would read:

- $t = t_0, N(t) = N_0, n = N_0$
- while $t < t_{end}$
- extract time of the next event $t + \delta t$ according to:

$$\delta t = (1/(BN + DN)) \ln(1/r_1) \quad (2.25)$$

- update growth curve: $N(t, t + \delta t) = n$
- extract random number r_2 from $U[0, 1)$
- if $r_2 > Dn/(Bn + Dn) : n = n + 1$
 else: $n = n - 1$
- update time: $t = t + \delta t$
- end while

In figures 2.1 and 2.2 we show results matching the dynamical evolution of the average number of cells predicted by the master equation: the average in the following refers to averages over the samples generated by the algorithm.

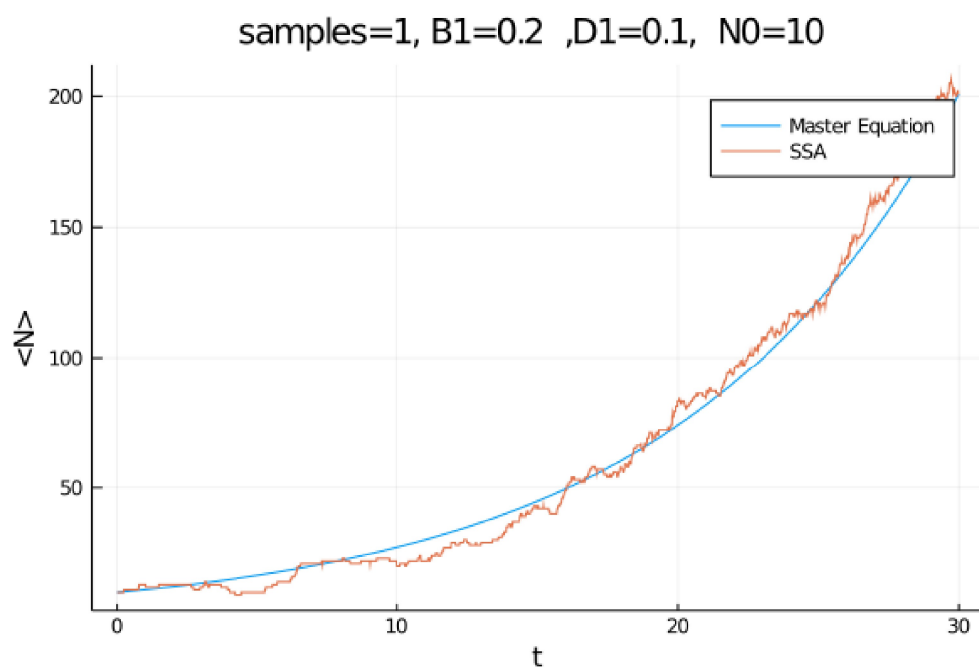


Figure 2.1: Average population number

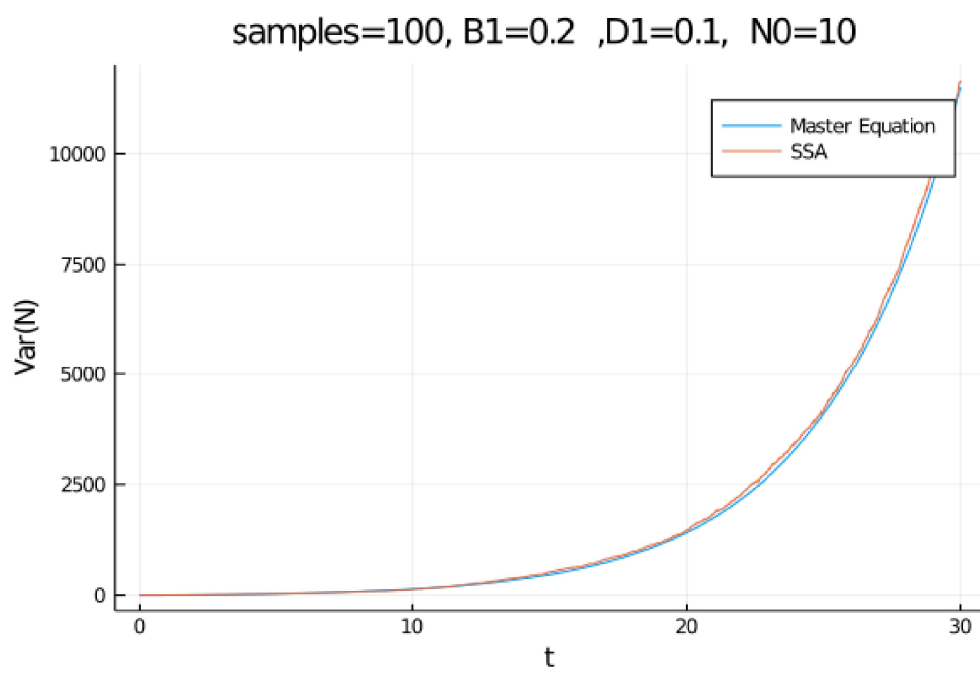


Figure 2.2: Variance

Chapter 3

Model Formulation and Analysis

3.1 Limited Environment and Multiple Populations

Limited Environment The starting point of our model is a linear birth-death process that we modify with the intent to incorporate the effects of a limited environment.

In an way analogous to how a Malthusian growth model is modified to contain a carrying capacity, we suggest the following new birth rates and death rates for the single cell.

$$R_1 : S_1 \rightarrow 2S_1 \quad c_1 = B_1(1 - \frac{N_1}{k_{eff}}) \quad (3.1)$$

$$R_2 : S_1 \rightarrow 0 \quad c_2 = D_1(1 + \frac{N_1}{k_{eff}}) \quad (3.2)$$

We cannot solve the master equation for this new process like we did in section 2.4, because the carrying capacity term introduces a non-linearity. Nonetheless, since, as detailed in section 2.4, the evolution in time of the average of the total population for a linear BD process reads:

$$\frac{d\langle N(t) \rangle}{dt} = (B - D)\langle N(t) \rangle \quad (3.3)$$

we expect that by modifying birth and death rates as in equations (3.1) and 3.2 we would observe a dynamical evolution for $\langle N(t) \rangle$ that reproduces approximately that one entailed by the following deterministic differential equation:

$$\frac{dN_1}{dt} = (B_1(1 - \frac{N_1}{k_{eff}}) - D_1(1 + \frac{N_1}{k_{eff}}))N_1 \quad (3.4)$$

By setting to 0 equation (3.4), one finds that the effective carrying capacity k_{eff} is

$$k_{eff} = (\frac{R+1}{R-1})k \quad (3.5)$$

Where $k = N_{tot,max}$ is the maximum value the total population can reach, and $R = \frac{B_1}{D_1}$.

Equation (3.4) can be conveniently reduced to a logistic equation:

$$\frac{dN_1}{dt} = r_1(1 - \frac{N_1}{k})N_1 \quad (3.6)$$

where $r_1 = B_1 - D_1$ and

$$k = k_{eff}(\frac{R-1}{R+1}) \quad (3.7)$$

We then know the system has a well understood steady state at $N_1 = k$. We also know the maximum growth rate can be defined as in section 1.2.2 , and assume (we will verify it in the following) the same dependence Λ_{max} on N_0 . Because we are mainly interested in the behavior of $\langle N_1 \rangle$ and not in the full probability distribution of the process we will limit ourselves to comparing results between stochastic simulations generated by the Gillespie algorithm and numerical solutions of the corresponding differential equations. The intrinsic differences between the two approaches will be addressed in the next section.

Multiple Populations The natural extension of our model to a situation where two sub-populations are competing for resources reads:

$$R_1 : S_1 \rightarrow 2S_1 \quad c_1 = B_1(1 - \frac{(N_1 + N_2)}{k_{eff}}) \quad (3.8)$$

$$R_2 : S_2 \rightarrow 2S_2 \quad c_2 = B_2(1 - \frac{(N_1 + N_2)}{k_{eff}}) \quad (3.9)$$

$$R_3 : S_3 \rightarrow 2S_3 \quad c_3 = B_3(1 - \frac{(N_1 + N_2)}{k_{eff}}) \quad (3.10)$$

$$R_4 : S_1 \rightarrow 0 \quad c_4 = D_1(1 + \frac{(N_1 + N_2)}{k_{eff}}) \quad (3.11)$$

It can be described by the system of ODE:

$$\begin{aligned} \frac{dN_1}{dt} &= (B_1(1 - \frac{N_1 + N_2}{k_{eff}}) - D_1(1 + \frac{N_1 + N_2}{k_{eff}}))N_1 \\ \frac{dN_2}{dt} &= (B_2(1 - \frac{N_1 + N_2}{k_{eff}}) - D_2(1 + \frac{N_1 + N_2}{k_{eff}}))N_2 \end{aligned} \quad (3.12)$$

Again, k_{eff} can be found by setting to zero the time evolution of the total population. In this case we have:

$$k_{eff} = (\frac{R_N + 1}{R_N - 1})k \quad (3.13)$$

where now $R(N)$ is equal to:

$$R_N = \frac{B_1N_1 + B_2N_2}{D_1N_1 + D_2N_2} \quad (3.14)$$

k and 0 are the values of the total population at the two steady states of the system. The system has no analytic solution.

For later use we define:

$$\lambda_1 = B_1(1 - \frac{(N_{tot})}{k_{eff}}) - D_1(1 + \frac{(N_{tot})}{k_{eff}}) \quad (3.15)$$

$$\lambda_2 = B_2(1 - \frac{(N_{tot})}{k_{eff}}) - D_2(1 + \frac{(N_{tot})}{k_{eff}}) \quad (3.16)$$

Extinction When a second population is introduced in a model with a limited environment, population and sub-population dynamics are greatly affected by the presence or absence of death events. It is possible to show that regardless of the initial conditions the sub-population with larger fitness (r) will eventually dominate growth, while the other will go extinct. A much different situation is described by the system:

$$\begin{aligned}\frac{dN_1}{dt} &= r_1\left(1 - \frac{N_1 + N_2}{k}\right)N_1 \\ \frac{dN_2}{dt} &= r_2\left(1 - \frac{N_1 + N_2}{k}\right)N_2\end{aligned}\tag{3.17}$$

Since for initial conditions $N_{01} < k$ and $N_{02} < k$ both rates of change cannot be negative it is easy to see how both populations will at most, stop growing once $N_1 + N_2 = k$, but neither of them will reduce its size or go extinct.

Analysis

Measuring the Growth Rate Which type of total population dynamics can we expect to be described by the model so far?

We know that for $N_1 + N_2 = k$ the system will stop growing, and we wonder until then what will be the observed growth rate. The sub-population with higher fitness (r_i), even far from the carrying capacity will grow faster than the other two. We can expect then that Λ will increase in time from a starting initial value of $\Lambda_p(t_0) = \frac{(\lambda_1 N_{01} + \lambda_2 N_{02})}{N_0}$, to possibly a maximal value of $\Lambda_{p,max} = \lambda_1$. We want to know if this maximal value is reached before the total population starts feeling the effects of the carrying capacity (i.e. $N_{tot} \ll k$), and how this might depend on the initial conditions, the value of k and the product $tr_1 r_2$.

In the limit $k \rightarrow \infty$ the evolution in time of the total population described by the system of ode 3.12 is

$$N(t) = N_1(t) + N_2(t) = N_{01} \exp(r_1 t) + N_{02} \exp(r_2 t) \quad (3.18)$$

where N_{01}, N_{02} are the initial values of the two sub-populations.

If we were to measure the growth rate Λ of the total population, as we have seen in section 1.2.1, we would have

$$\Lambda = \frac{\log(N/N_0)}{t} = \frac{1}{t} [\log(\frac{N_{01}}{N_0}) + r_1 t + \log(1 + \frac{N_{02}}{N_{01}} \exp((r_2 - r_1)t))] \quad (3.19)$$

For $r_1 > r_2$ we can predict that the sub-population N_1 will grow faster than N_2 , and equation (3.19) tells us that for a given combination of values of $N_{01}, N_{02}, r_1, r_2, t$ we would observe a growth rate equal to r_1 .

We can find approximately these values by observing that for

$$\frac{N_{02}}{N_{01}} \exp[(r_2 - r_1)t] = \frac{N_2(t)}{N_1(t)} = \frac{1 - f}{f} \approx 0 \quad (3.20)$$

($f(t) = N_1(t)/N_{tot}$) then equation (3.27) reduces to

$$\Lambda = \frac{\log(N/N_0)}{t} \approx \frac{1}{t} [\log(\frac{N_{01}}{N_0}) + r_1 t] \approx r_1 \quad (3.21)$$

By using lower bounds for the values of f , N_{tot} , we can estimate the minimum value of the carrying capacity that allows us to observe during growth a value

$\Lambda = r_1$, for different values of initial conditions N_{01} and N_{02} .
Thus, imposing the lower bound for $f(t) = N_2(t)/N_{tot}$:

$$\frac{1-f}{f} = 10^{-3} \quad (3.22)$$

we find that the time t^* at which such condition is satisfied is equal to

$$t^* = \frac{1}{r_1 - r_2} \log\left(\frac{N_{02}(f)}{N_{01}(1-f)}\right) \quad (3.23)$$

We also impose that our "pure exponential growth" approximation is still valid for

$$\frac{N_{tot}}{k_{eff}} = 10^{-3} \quad (3.24)$$

Altogether we have that:

$$N_{tot}(t^*)10^3 = k_{eff,min} \quad (3.25)$$

If $k_{eff} > k_{eff,min}$ we can expect to observe a growth rate approximately equal to r_1 before the effect of the carrying capacity kicks in. Substituting for $k_{eff,min}$:

$$k_{min} = N_{tot}(t^*)10^3 \left(\frac{R_N - 1}{R_N + 1}\right) \quad (3.26)$$

In figure we 3.1 we show k as a function of r_1 and r_2 , for different values of N_0 . As we can see even for relatively low values of N_0 , the minimum carrying capacity required is quite large.

We conclude we will not always be able to explore such conditions, which forces us to reconsider how to measure the growth rate. Another way to approach the problem of measuring the growth rate would be to measure the instantaneous growth rate

$$\Lambda_t = \frac{d \log(N/N_0)}{dt} = \frac{d}{dt} \left[\log\left(\frac{N_{01}}{N_0}\right) + r_1 t + \log\left(1 + \frac{N_{02}}{N_{01}} \exp((r_2 - r_1)t)\right) \right] \quad (3.27)$$

which corresponds to the *population average* growth rate Λ_p . Of course the two measurements in the limit described above correspond to Λ . In the case instead in which the carrying capacity is not infinite and the combination of parameters r_1, r_2, t and initial conditions N_{01}, N_{02} do not lead to an observable constant growth for values of $N_{tot} < k$ it is still possible to observe

a "pure" exponential phase of the population (expressed as an observed approximately constant value of Λ). Growth in fact will accelerate as the faster growing population conquers a larger fraction of the total population, until the "deceleration" imposed by the carrying capacity counterbalances its effect, yielding for a short period of time an approximately constant growth rate. In this regime Λ_t and Λ differ greatly. In the following section we show qualitatively the difference between the two measurements.

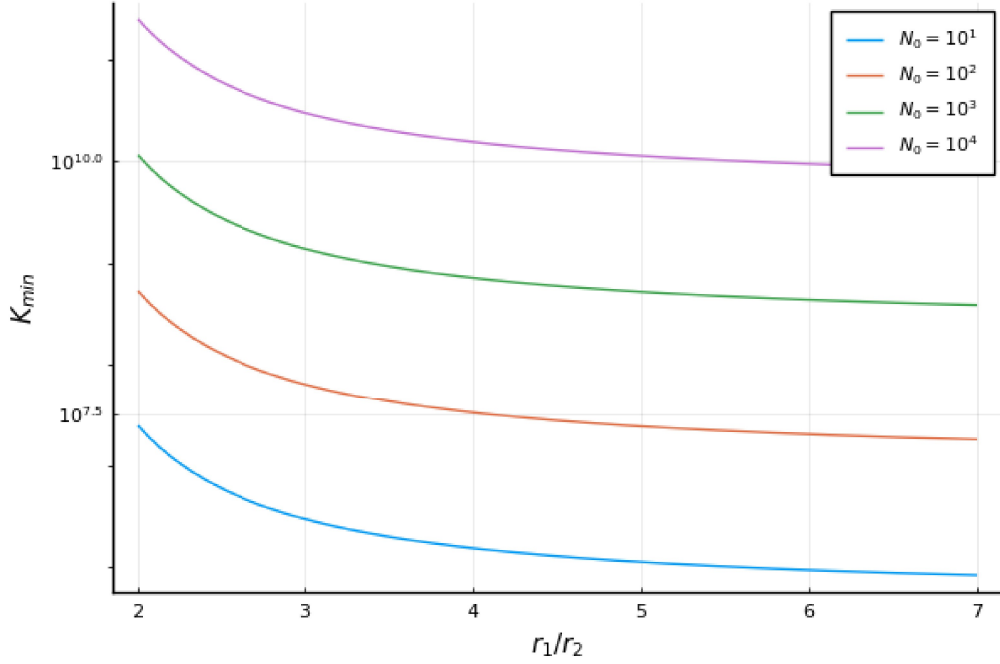


Figure 3.1: **Minimum Carrying Capacities** We show as a function of the ratio r_1/r_2 the quantity k_{mi} as defined in equation (3.26) , for different values of N_0 ($N_{01} = N_{02} = 1/2N_0$)

3.1.1 Results

Numerical solutions were obtained with the ODE solver package DifferentialEquations, in Julia v 1.8.2.

The stochastic process trajectories were generated using an implemented direct Gillespie method. Because the number of operations grows exponentially with the size of the population, we kept for every simulation a carrying capacity $k \leq 10^6$.

In order to speed up the end of the process still, we modified our code so that the computation would stop once the value $N_{tot} = k$ was reached. Another necessary modification was to automatically set a lower bound of zero for the rate of birth of each species

$$c_i = B_i(1 - \frac{(N_1 + N_2)}{k_{eff}}) \quad (3.28)$$

Indeed for initial conditions $N_0 \gg k$, such birth rate turns negative, and a negative reaction probability rate has no meaning. Furthermore and when it dominates the total sum of the volumic rates, the δt of the next event becomes negative (see pseudocode in section 2.4).

An important difference between the two approaches is the existence of *absorbing states* in the stochastic process. While the logistic equation for $B_1 > D_1$ will deterministically lead to a steady state $N_1(\infty) = k$, there is always a non-zero probability that the stochastic "logistic" B-D process will lead to the extinction of the population. The smaller N_0 is the higher the chance that a fluctuation (an improbable sequence of death events) crystallizes the total population (or the sub-population) to an extinct state. In figure 3.2 we can see minor differences in growth trajectories obtained with the two different methods : these tend to disappear as we increase the value of N_0 .

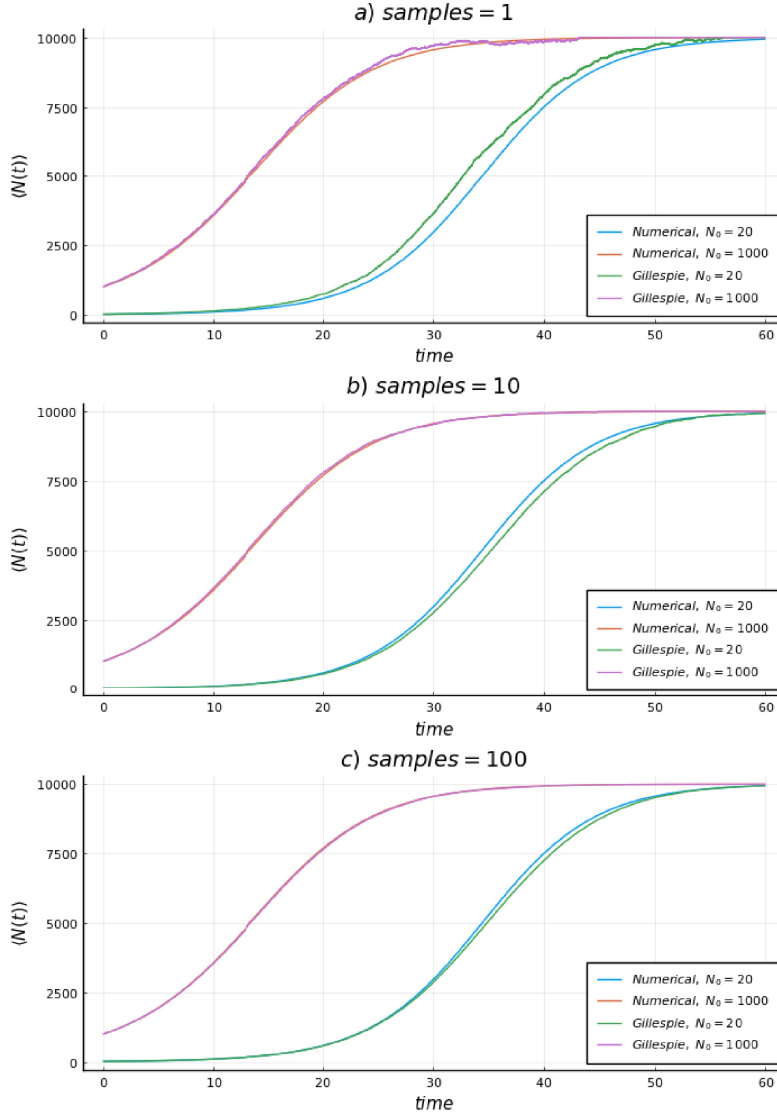


Figure 3.2: **Comparison** Due to execution times that scale like $t_{exec} \sim \exp(N_{tot})$ the carrying capacity is set to $k = 10^4$. Growth trajectories are obtained for two different values of initial conditions $N_{01} = N_{02} = 10$ and $N_{01} = N_{02} = 500$. a) $\langle N(t) \rangle$ is an average over 1 sample trajectory b) average is over 10 samples trajectories generated by the algorithm, it is possible to see a closer similarity between the results of the two methods for larger N_0 c) 100 sample trajectories

Growth Rates

Performing a stochastic simulation of birth death processes for multiple populations is possible to record at all times the values of the distinct sub-population sizes. We can therefore define and measure the quantity:

$$\Lambda_p = \frac{\lambda_1 N_1(t) + \lambda_2 N_2(t)}{N_{tot}(t)} \quad (3.29)$$

Where each λ_i is the growth rate expressed by each sub-population. Using instead data from the trajectory in time of N_{tot} we measure

$$\Lambda = \frac{\log(N_{tot}/N_0)}{t} \quad (3.30)$$

and

$$\Lambda_t = \frac{d\log(N_{tot}/N_0)}{dt} \quad (3.31)$$

To compute Λ_t we first fit $\log(N_{tot}/N_0)$ with a spline and then compute its derivative numerically. For all these measurements, the maximum value of the growth rate Λ_{max} is obtained through the standard function *findmax()* in Julia v1.8.2.

In figure 3.3 one can see how the maximum occurs in a phase of growth that appears exponential, and how different the values given by Λ and Λ_t are.

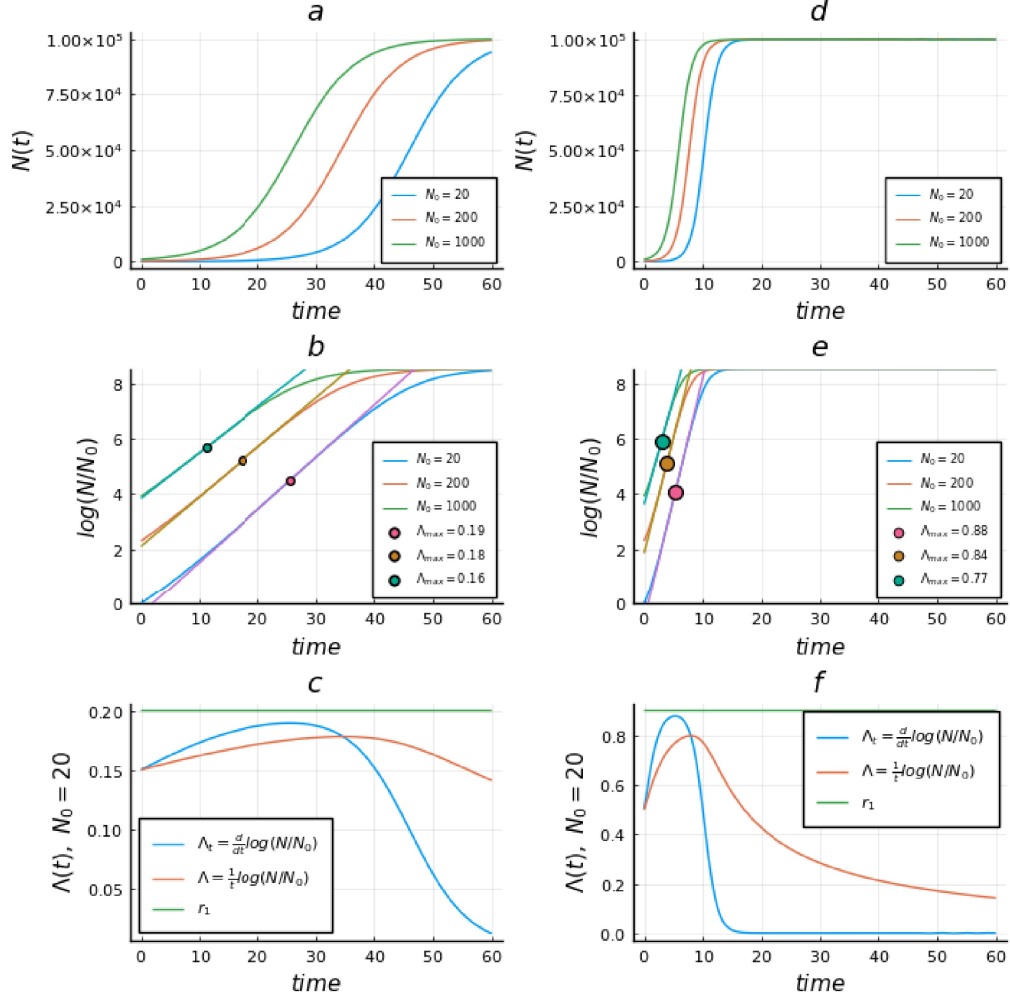


Figure 3.3: **Exponential Phase Comparison:** numerical solutions for the system (3.12) for $N_{01}, N_{02} = 10, 100, 500$ with parameters $k = 10^5$, $r_2 = 0.1$, $r_1 = 1.5r_2$ in (a-b), and $k = 10^5$, $r_2 = 0.1$, $r_1 = 5r_2$, in (d-e). In (c) is shown difference between the two measurements for $N_{01} = 10$ and $N_{02} = 10$, $r_2 = 0.1$, $r_1 = 1.5r_2$ (f) $k = 10^5$, $r_2 = 0.1$, $r_1 = 5r_2$

Λ_{max} vs N_0

We attempt to recreate batch culture experiments similar to the ones presented in section 1.3. Given a set of values of model parameters (r_1, r_2, k) we obtain growth trajectories *in silico* changing initial conditions in the following manner: 5 sample growth curves are obtained for each of 10 selected values of N_0 distributed logarithmically in the interval $[10^{1.5}, 10^6]$. Each of the 5 samples per N_0 value, has a different value of $f_0 = N_{01}/N_0$. f_0 is drawn each time from a uniform distribution on the interval $[0, 1)$, and N_{02} is set according to $N_{02} = N_0 - N_{01}$, as the two initial conditions are constrained by the sum $N_0 = N_{01} + N_{02}$ (examples in figure 3.4). For each growth curve so generated, its instantaneous growth rate Λ_t is computed, and its maximum value Λ_{max} is obtained through a standard search-of-max function in Julia. One finally obtains 5 different $\Lambda_{max}(N_0)$ curves, that are averaged together resulting in the trend we plot in the following figures. The errorbars are the standard deviations of Λ_{max} for each N_0 value.

In general we can observe that for $r_1 > r_2$ Λ_{max} decreases as N_0 increases, understandably because for the same value of f_0 it takes longer for the faster growing sub-population to dominate growth. Importantly for $r_1 = r_2$ our system behaves like one with a single population, and in fact we can see in fig 3.5 (a) the typical logistic dependence of Λ_{max} with N_0

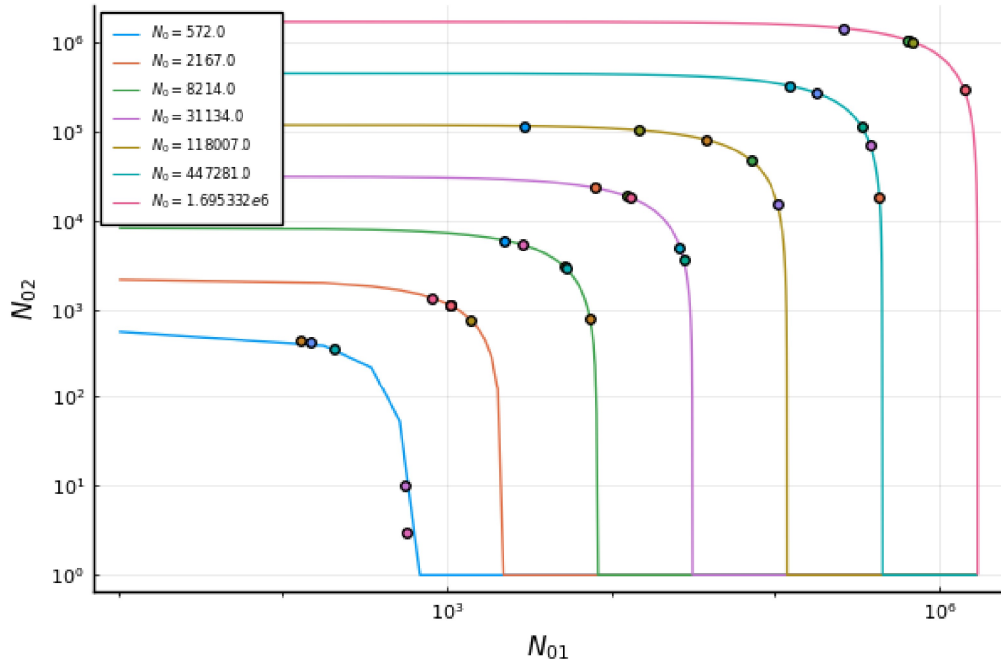


Figure 3.4: **Initial Conditions:** Colored curves represent the set of points for which $N_{01} + N_{02} = N_0$. The markers are the actual samplings performed on these curves to obtain the initial conditions of our simulations.

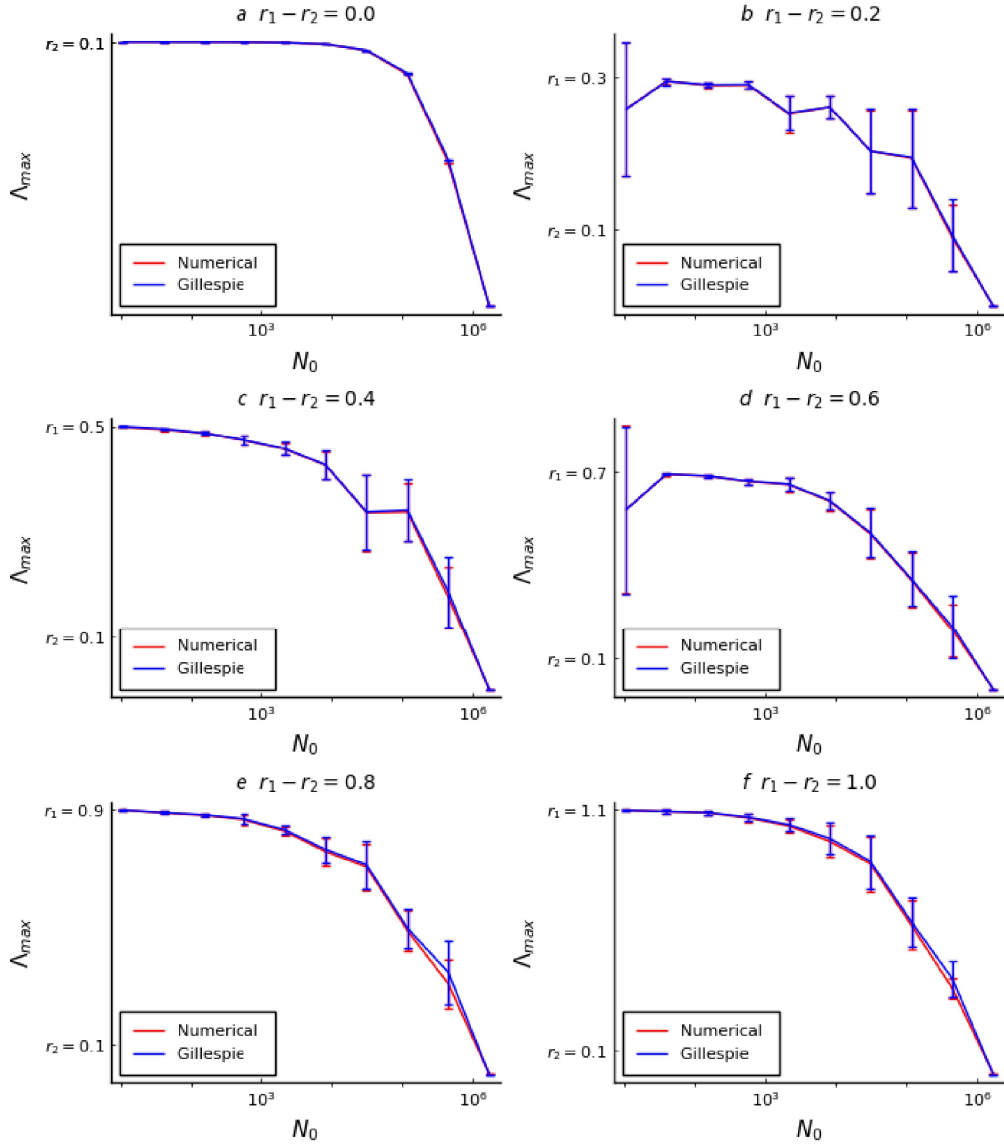


Figure 3.5: Λ_{max} vs N_0 : results of simulations obtained for different values of $r_1 - r_2$, the intrinsic maximal growth rates of the two sub-populations with. Error bars represent standard deviations of the average values

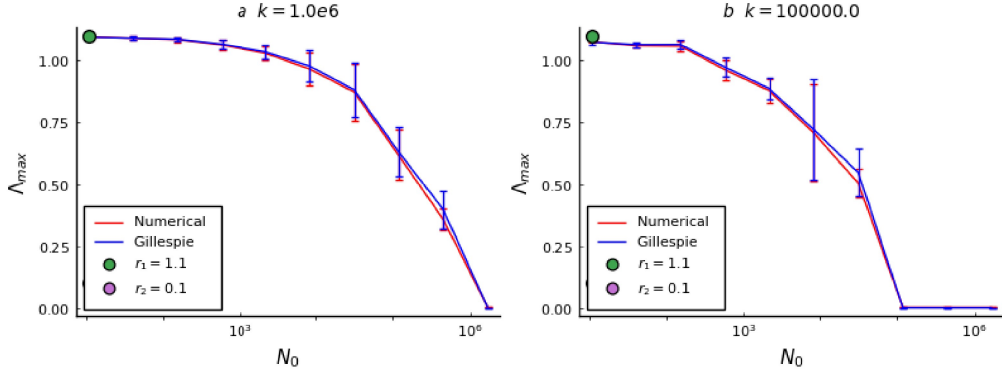


Figure 3.6: **Varying k** : we show results for same parameters values $r_1 = r_2 = 0.1$ but different carrying capacity k . a) $k = 10^6$. b) $k = 10^5$

3.2 Chemical Signaling

We now want to introduce in our model a growth-signaling mechanism.

In its definition we will make use of the clear picture outlined in [35], where growth factor producers are not *consumers* of the substance. Our interest is to find out whether such extreme limit case of a more general approach to single-cell cooperation via growth factors, is able to reproduce the results we discussed in section 1.3. The authors of the study present first a common feature of cancer cell cultures: cooperation in a heterogeneous population through the sharing of a resource.

They record the growth of two genotypically different β -tumor cell lines derived from insulinomas of (Rip1Tag2) mice. A non -mutated strain ("producer" cells, (+/+)), and a mutated one carrying a homozygous deletion of the IGF-II gene ["non-producer" cells (-/-)]. The non-mutated gene entails the production of the growth hormone IGF-II (commonly up-regulated in many tumors) that sustains growth.

Individually the two populations grow at different rates with the IGF-II producing population exhibiting the highest.

But when growing the non-productive genotype on a medium previously conditioned by the presence of the +/+ genotype, the former showed an increased growth rate, higher than the latter.

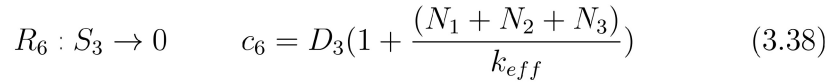
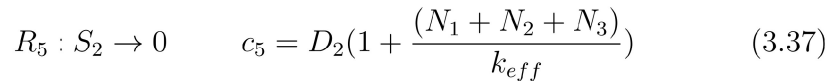
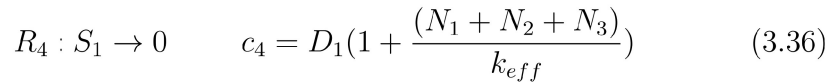
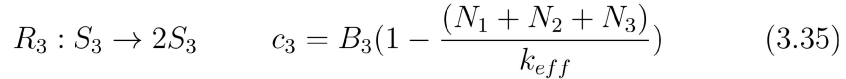
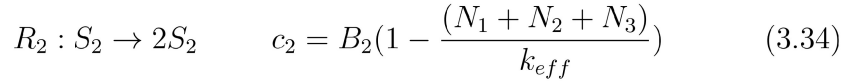
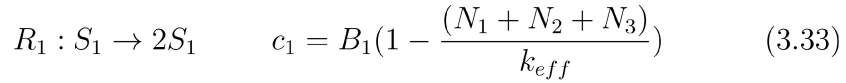
When the two genotypes were grown together, the growth rate of the total population was found to be a Hill function of the concentration of the endogenously produced growth rate.

Assuming that the concentration of growth rate available is proportional to the size of the population that produces it (species: S_1), and that the population "consuming" it (species S_2) can boost its fitness (encoded in the value of the intrinsic Malthusian growth rate) without a production cost, we can postulate the following reaction to occur in the volume:



The reaction entails the creation of a new species S_3 with intrinsic maximal growth rate $r_3 > r_2 > r_1$. An intuition of the mechanism behind this "contagion" process is the following: when the growth factor binds to the free receptor of a cell of the non-producer species S_2 , that same cell is now "signaled" to reproduce faster, implying $r_3 > r_2$. A classic contagion model would see the infected species gain the ability to infect other individuals. In our model, this possibility is excluded because we constraint the signaling to stimulate growth and not turn the cell into a new producer: in other words, the "infected" cell is exclusively consuming a resource that is produced by another species. The reaction rate c for the reaction (3.32) is a free parameter of the model.

Summarizing, the complete set of reactions happening in the volume is:



The deterministic version of the model reads:

$$\begin{cases} \frac{dN_1}{dt} = \lambda_1 N_1 \\ \frac{dN_2}{dt} = \lambda_2 N_2 - c_7 N_1 N_2 \\ \frac{dN_3}{dt} = \lambda_3 N_3 + c_7 N_1 N_2 \end{cases} \quad (3.40)$$

where

$$\lambda_1 = B_1 \left(1 - \frac{(N_{tot})}{k_{eff}}\right) - D_1 \left(1 + \frac{(N_{tot})}{k_{eff}}\right) \quad (3.41)$$

$$\lambda_2 = B_2 \left(1 - \frac{(N_{tot})}{k_{eff}}\right) - D_2 \left(1 + \frac{(N_{tot})}{k_{eff}}\right) \quad (3.42)$$

$$\lambda_3 = B_3 \left(1 - \frac{(N_{tot})}{k_{eff}}\right) - D_3 \left(1 + \frac{(N_{tot})}{k_{eff}}\right) \quad (3.43)$$

$N_{tot} = N_1 + N_2 + N_3$ and the effective carrying capacity k_{eff} is

$$k_{eff} = \left(\frac{R+1}{R-1}\right)k \quad (3.44)$$

where $k = N_{tot,max}$ is the maximum value the total population can reach, and $R = \frac{B_1 N_1 + B_2 N_2 + B_3 N_3}{D_1 N_1 + D_2 N_2 + D_3 N_3}$.

3.2.1 Analysis

The time evolution of the total population is:

$$\frac{dN_{tot}}{dt} = \lambda_1 N_1 + \lambda_2 N_2 + \lambda_3 N_3 \quad (3.45)$$

and it reaches a steady state for $N_{tot} = k$. As we have seen in the analysis of the model without contagion, measurements of the growth rate Λ_t coincide with measurements of Λ_p that now reads:

$$\Lambda_p = \frac{\lambda_1 N_1 + \lambda_2 N_2 + \lambda_3 N_3}{N_{tot}} \quad (3.46)$$

We can thus expect that the presence, during exponential growth, of a new population with a new phenotype λ_3 , would affect the aforementioned measurements. How much of an increase in the maximum growth rate is observed

depends on **if** and **when** the "contagion" process begins during growth. If the birth of the first S_3 individual happens outside of the growth time window, we would not see its effect on the growth rate. We will explore this idea further by looking at the deterministic version of the model.

N_0 dependency of Contagion Rate

Assuming that at the beginning of growth there are no "infected" individuals ($N_{03} = 0$), for each growth trajectory we can then expect that for a certain period of time starting from $t = t_0$, (the beginning of our simulations) $N_3 \approx 0$.

In this time interval case the time evolution of our system reads:

$$\begin{cases} \frac{dN_1}{dt} = \lambda_1 N_1 \\ \frac{dN_2}{dt} = \lambda_2 N_2 - c_7 N_1 N_2 \\ \frac{dN_3}{dt} = c_7 N_1 N_2 = c_7 (1 - f(t))(f(t)) N_{tot}^2(t) \end{cases} \quad (3.47)$$

Where $f(t) = \frac{N_1(t)}{N_{tot}(t)}$.

If we assume $f(t) = f$, constant in time, we can see how for low enough values of c_7 , we will observe that $N_3(t) \approx 1$ the sooner N_{tot} reaches a certain threshold. If that threshold value of N_{tot} is close enough to the carrying capacity of the model we would not measure a significant increase in growth rate, since by that time the value of $\lambda_3(N_{tot}(t))$ expressed by N_3 would be lower than the maximum growth rate observed by then. Additionally, $f(t)$ does not remain constant in time, but increases until species S_3 begins its growth, so that two different inocula, characterized by the same initial value of $f_0 = f(t = 0)$ but a different value of N_0 , would reach the same value of N_{tot} with very different values of $f(t)$: the inoculum with lower N_0 will grow up to $N_{tot}(t)$ exhibiting at that time a much higher value of $f(t)$.

In some cases $f(t)$ reaches its maximum value $f_{max} = 1$ (thus reducing the last equation of system (3.47) to zero), before N_3 is significantly present in the population. The rate at which f increases depends in turn, on the difference between $r_2 = B_2 - D_2$ and $r_1 = B_1 - D_1$. The larger the difference the faster $f(t)$ will grow up to 1, extinguishing the sub-population N_2 .

In the following subsections we test these predictions.

3.2.2 Results

As we did before we first look at some growth curves obtained as explained in section 3.1.1, making sure we are measuring a maximum growth rate Λ_{max} during an exponential phase.

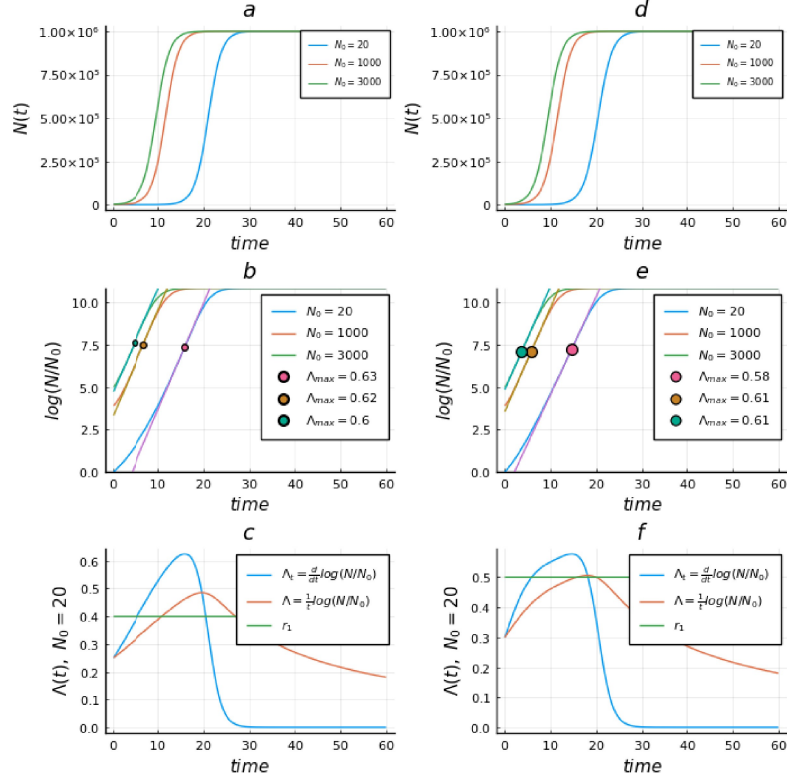


Figure 3.7: **Growth Rates** Numerical solutions, growth curves, and comparison between Λ and Λ_t for $N_{01}, N_{02} = 10, 100, 500$ with parameters $k = 10^6$ $r_1 = 0.3, r_2 = 0.1, r_3 = 0.7$ in (a-b-c), and $k = 10^6$ $r_1 = 0.5, r_2 = 0.1, r_3 = 0.7$ in (d-e-f). In (c-f) is shown difference between the two measurements for $N_{01} = 10$ and $N_{02} = 10$

Results: $\Lambda_{max}(N_0)$

We attempt to recreate the experiments presented in section 1.3. Given a set of values of model parameters (r_1, r_2, r_3, k, c_7) we obtain growth trajectories *in silico* changing initial conditions in the following manner: 10 sample growth curves are obtained for each of the 15 selected values of N_0 distributed logarithmically in the interval $(10^{1.5}, 10^6)$. Each of the 10 samples per N_0 value, has a different value of $f_0 = N_{01}/N_0$. N_{03} is always set to zero. f_0 is drawn each time from a uniform distribution on the interval $[0, 1)$, and N_{02} is set according to $N_{02} = N_0 - N_{01}$, as the two initial conditions are constrained by $N_0 = N_{01} + N_{02}$.

For each growth curve so generated, its instantaneous growth rate Λ_t is computed, and its maximum value Λ_{max} is obtained through a standard search-of-max function in Julia. One finally obtains 10 different $\Lambda_{max}(N_0)$ curves, that are averaged together resulting in the trend we plot in the figures. The errorbars are the standard deviations of Λ_{max} for each N_0 value.

Parameters Sweep We obtain results for a wide range of key parameters values. We do it solving numerically the system of ODE describing our model, because a rough estimation of the time required to run as many Gillespie simulations returns a value $t_{simulation} \approx 600Hrs$. Specifically, r_1 ranges in the set $\{0.3, 0.4, 0.5, 0.6, 0.7\}$, r_3 in the set $0.3, 0.4, 0.5, 0.7, 0.8, 0.9, 1.0$, $c_7 \in \{0.001, 0.0001, 7e-5, 3e-5, 1e-5, 5e-6, 1e-6, 6e-7, 1e-7, 0.0\}$ and $k \in \{10^5, 10^{5.5}, 10^6\}$. Parameter r_2 was kept constant at value 0.1. In figure 3.8 we show selected results for parameters values that seemed most to capture qualitatively the experimental behavior.

varying c_7 and K As we speculated in the previous section increasing the value of the contagion rate c_7 appears to lower the threshold for the onset of observable "contagion". Eventually for large enough contagion rate we can see its effect for the smallest values of N_0 . Increasing the carrying capacity brings about instead, an overall shift of the "cutoff" value of N_0 , towards, in similar fashion to what would be observed in a logistic model. In our case this permits large inocula to manifest their higher contagion potential with respect to smaller ones for the same value of c_7 . In figure 3.9 one can observe these effects.

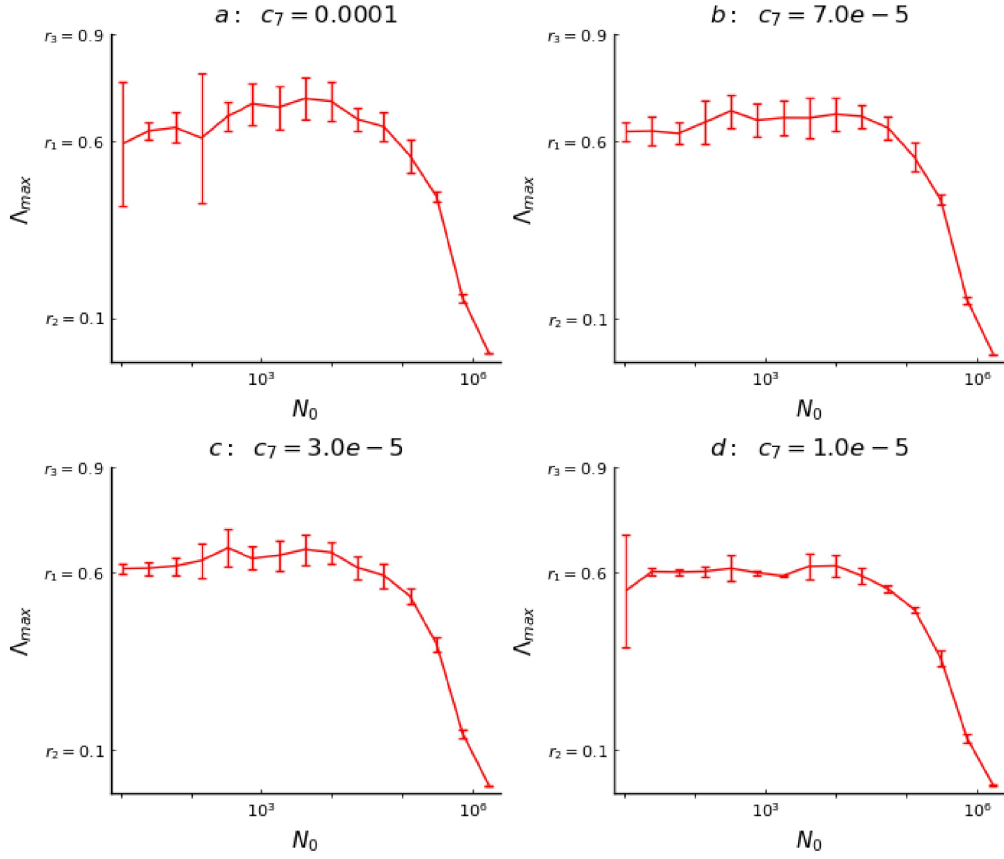


Figure 3.8: **Selected Results**we plot for the fixed combination of parameters values $r_1 = 0.6, r_2 = 0.1, r_3 = 0.9, k = 10^6$ measurements of $\Lambda_{max}(N_0)$ for a closely spaced range of values of the "contagion" rate, c_7 . (a) $c_7 = 1 \times 10^{-4}$ (b) $c_7 = 7 \times 10^{-5}$ (c) $c_7 = 3 \times 10^{-5}$ (d) $c_7 = 1 \times 10^{-5}$

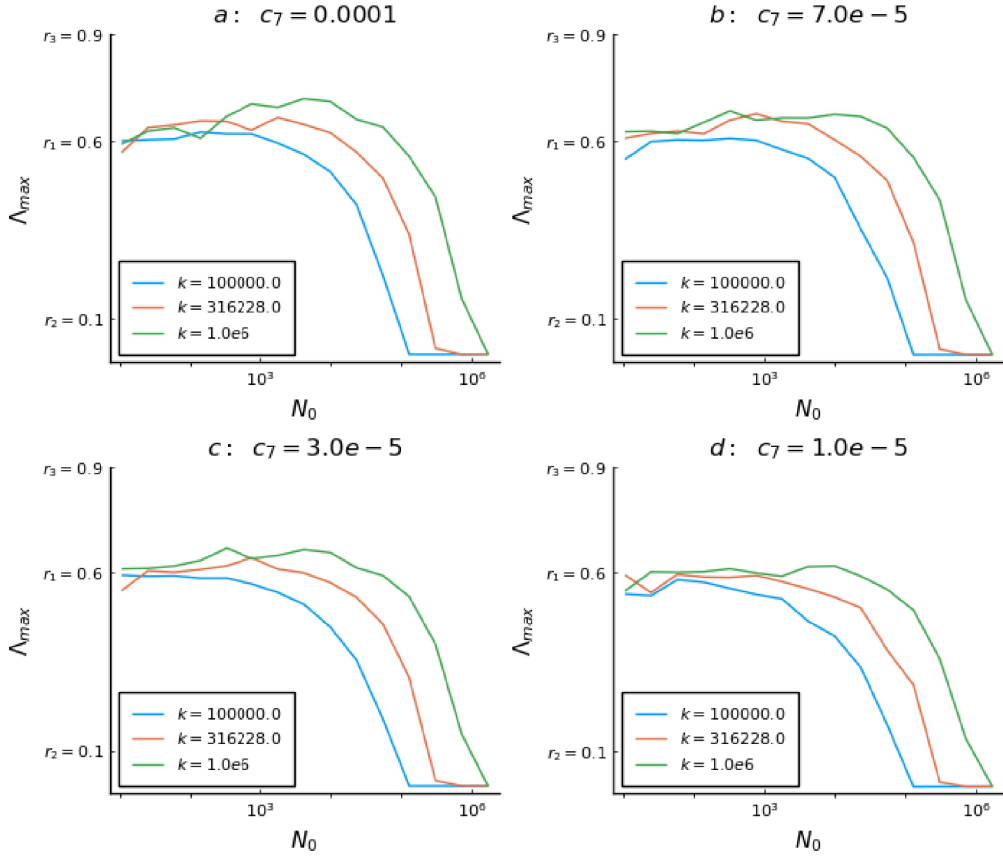


Figure 3.9: **Effects of varying carrying capacity and contagion rate on $\Lambda_{max}(N_0)$ profile** In each figure we plot three different $\Lambda_{max}(N_0)$ curves that correspond to a different value of k , the carrying capacity. (a) $c_7 = 1 \times 10^{-4}$ (b) $c_7 = 7 \times 10^{-5}$ (c) $c_7 = 3 \times 10^{-5}$ (d) $c_7 = 1 \times 10^{-5}$

Varying intrinsic growth rates (phenotypes) Increasing r_1 , the intrinsic maximal growth rate of the "producer" sub-population N_1 has a double effect: the larger it is than r_2 , "consumer" species phenotype, the sooner sub-population N_2 will go extinct, preventing any observable form of contagion. On the other hand if r_1 is too close to r_3 we won't observe any actual increase in the maximum growth rate. In figures 3.10 and 3.11 we explore these effects.

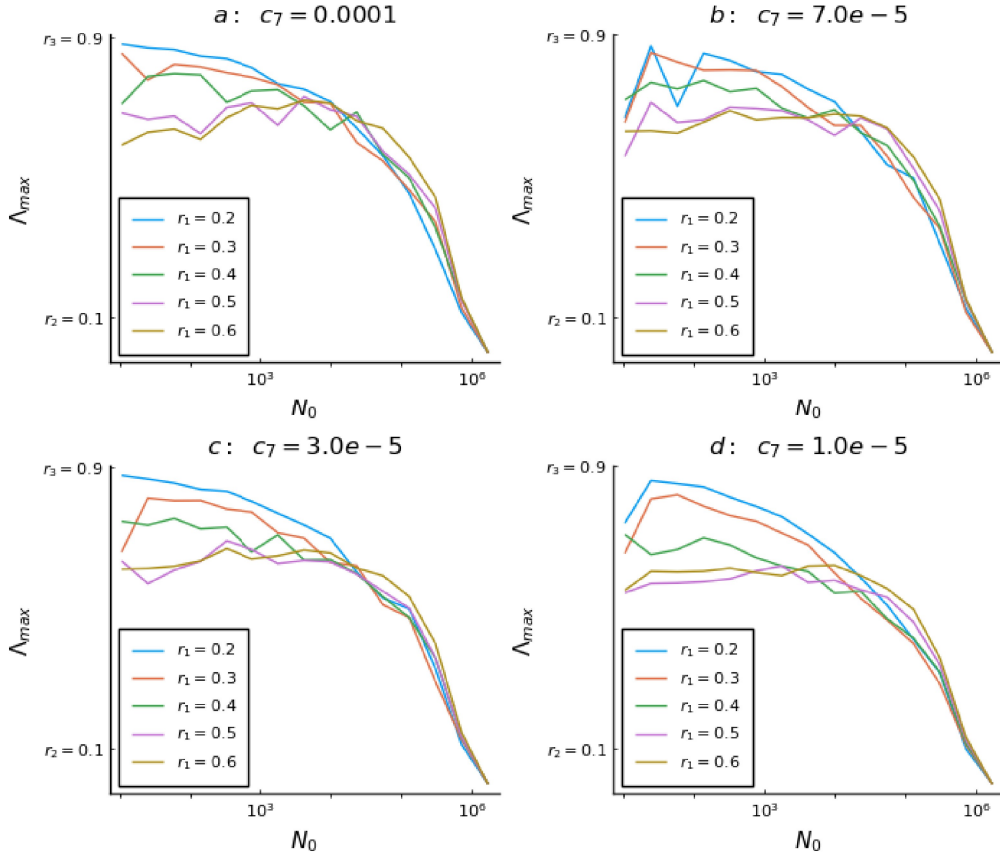


Figure 3.10: **Varying r_1** In each sub-figure for a fixed values of contagion rate, carrying capacity, and r_3 we vary the values of the "phenotype" of sub-population N_1 : its intrinsic maximal growth rates r_1 . Each different curve correspond to a different value of $r_1 \in \{0.2, 0.3, 0.4, 0.5, 0.6\}$. (a) $c_7 = 1 \times 10^{-4}$ (b) $c_7 = 7 \times 10^{-5}$ (c) $c_7 = 3 \times 10^{-5}$ (d) $c_7 = 1 \times 10^{-5}$

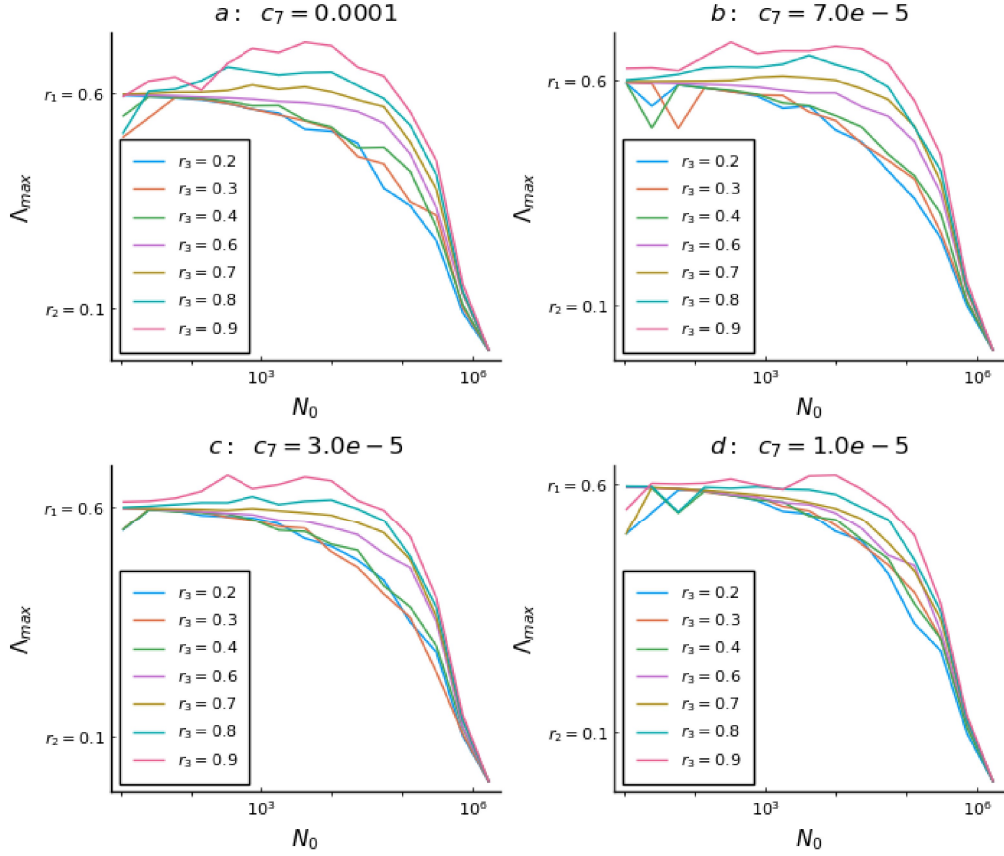


Figure 3.11: **Varying r_3** In each sub-figure for a fixed values of contagion rate, carrying capacity, and r_1 we vary the values of the "phenotype" of sub-population N_3 : its intrinsic maximal growth rates r_3 . Each different curve correspond to a different value of $r_3 \in \{0.2, 0.3, 0.4, 0.5, 0.6\}$. (a) $c_7 = 1 \times 10^{-4}$ (b) $c_7 = 7 \times 10^{-5}$ (c) $c_7 = 3 \times 10^{-5}$ (d) $c_7 = 1 \times 10^{-5}$

Chapter 4

Conclusions

The goal of this thesis project was to develop a minimal toy-model of cancer cell chemical cooperation in exponential growth, with the aim of comparing qualitatively its predictions with observed experimental behaviors.

In the first chapter a brief description of the mechanisms we wanted to embed in our model was given. We first introduce the reader to the basics of cell culturing, then overview traditional approaches to capture their growth quantitatively. Subsequently a review of the relevant literature is , to better understand the experimental work that is at the basis of this project. Summarizing: inoculum-dependent maximum growth rate is observed , and its dependence on N_0 is not explained by classical models. In the second chapter the tools and the formalism adopted in the development of the model are introduced, together with the techniques employed in its analysis. Finally, in the third chapter we begin the modeling work. First we show how the limitations imposed by a finite environment can be included in a classic birth-death process, and we discuss how the possibility of multiple sub-populations might be approached. Secondly we try to encode a complex biochemical interaction in a fairly simple set of microscopic events. After combining all the elements together in a single growth model for multiple and interacting cell populations, we set out to explore qualitatively its predictions. A short analysis of the “contagion” term is conducted, outlining the expected effects of model parameters on the dynamics. Using stochastic simulations and numerical solvers, we explore the model predictions, mimicking the experimental procedures that were employed in 1.3.

4.1 Comparison with Experimental Data

For an accurate qualitative comparison with the experiments we fix to constant values the following parameters of the model: k is set equal to 10^7 , which is the average value of carrying capacity measured over the populations in the experimental work. The general picture outlined in chapter 3 tells us that for low N_0 we can observe an approximately constant value of Λ_{max} that corresponds to the intrinsic maximal growth rate of the producer population: for this reason we set r_1 to a value of 0.03, as it was approximately the constant one recorded by Bena and colleagues for low inocula. The parameters that remain are the two intrinsic maximal growth rates that characterize the consumer population before and after the signaling (or the contagion) and the contagion rate. Initial conditions again are picked following the procedures outlined in the previous chapter: N_0 spans a range of 15 values distributed logarithmically in $(10^1, 10^7)$, and for each value 10 growth samples are generated, each with a different value of $f_0 = \frac{N_1}{N_{tot}}$, for a total of 150 simulated growth curves.

Contagion rate By now we understood that a large value of contagion rate fosters the birth of the fitness-boosted population at lower inoculum densities. Since k is now an order of magnitude larger than the values we tested before we can expect its counterbalancing effect to manifest only at larger inocula (the smaller k , the lower is the maximum value the producer population reaches before declining and end up going extinct, preventing any contagion from happening). We then look at results for c_7 , the contagion rate, ranging between 10^{-7} to 10^{-8} .

Intrinsic maximal growth rates r_2, r_3 As r_1 is now kept fixed we have less freedom to vary the difference $\delta_{1,2} = r_1 - r_2$. We know that for $\delta_{1,2} \approx 0$ there will be no suppressing of contagion for low inocula, while the greater the difference between the two rates, the larger N_0 needs to be to prevent rapid extinction of sub-population N_2 . The chosen value of r_2 is 0.008 so that $\delta_{1,2} = 0.022$.

For what concerns r_3 , if $r_3 \gg r_1$ then we expect Λ_{max} to jump to the r_3 value as soon as N_0 is large enough. For $r_3 \approx r_1$ instead, we would not observe any significant increase: we explore a range $r_3 \in [0.04, 0.08]$. Best results are obtained for $c_7 = 5 \times 10^{-8}$ $r_3 = 0.075$, we show them in comparison with laboratory results in figure 4.1.

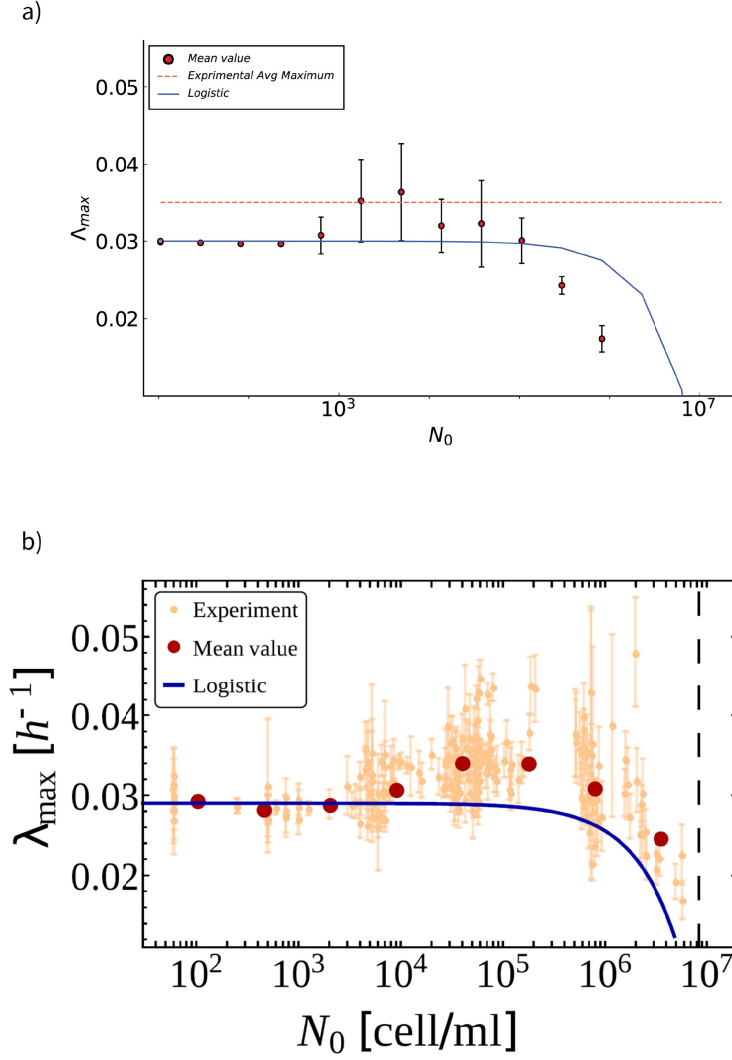


Figure 4.1: a) λ_{max} is the average maximal growth rate measured and averaged over Gillespie generated samples. Bars represent the standard deviation of the measurement. The dotted line marks the average maximum recorded in vitro. b) Experimental data: orange dots represent parameter estimates from individual experiments with their standard errors, while red dots represent the mean values of λ_{max} . In both (a-b) the blue line denotes the behavior of λ_{max} vs N_0 expected on the basis of a purely logistic model with $k = 8.6 \times 10^6$ and intrinsic maximal growth rate $r = 0.029$

Comparison

Mean Behavior In general our model predicts the average behavior of Λ_{max} fairly well. For approximately the same range of small inocula, we measure the constant value $\Lambda_{max} \approx 0.03$. At around $N_0 \approx 10^3$, the maximum growth rate of the population increases until it reaches a maximum of ≈ 0.035 , very close to the maximum average recorded during the in vitro experiments. The gradual decrease towards zero due to $N_0 \rightarrow k$ appears to be smoother in the trend obtained in [2] while our curve is characterized by a steeper negative slope.

Fluctuations Qualitatively fluctuations follow the same trend: their size positively correlates with N_0 , until the correlation becomes negative for inocula sizes that approach the carrying capacity. Quantitatively, the absence of significant variations in the constant section of the trend predicted by our model, might suggest that the source of these variations is ultimately linked with the appearance of contagion, implying that as long as N_0 is too small to prevent the consumer population to go extinct before being “infected”, f_0 , has no relevant effect on the observed growth rate. In fact we always measure approximately the maximal intrinsic growth rate of the producer population. When instead the new growth-boosted population N_3 finds suitable conditions to grow, is able to amplify the fluctuations in the partition of the initial condition (f_0) and translate them in growth rate fluctuations: a smaller f_0 would see an onset of contagion for smaller values of the total population, so that $\lambda_3 = B_3(1 - \frac{N_{tot}}{k_{eff}}) - D_3(1 + \frac{N_{tot}}{k_{eff}})$ is less affected by the finite environment.

4.2 Discussion

In summary our analysis tells us that the simple model we developed seems to capture qualitatively quite well some of the inoculum-dependent behaviors described in 1.3. Fierce competition for resources and autocrine growth factors production are known to be key elements in tumor population growth [35], and in this case they play a fundamental role in shaping the dynamics of population growth in relation to initial density size. Similar results might be achieved reducing the number of free parameters of the model, for example deriving the value of the contagion rate from the known and quantified physical and biochemical processes it encapsulates. Another step in the direction of biological verisimilitude would be to limit the effect of growth signaling to a period of time, when instead was assumed to be everlasting in our extremely simplified modeling: it is reasonable to expect the increase in fitness to fade over time. Further improvement could also be achieved exploring the possibility of fixing the value of f_0 to an approximately constant value. In fact, in [35] the authors found that the two competing populations are able to coexist in a well established equilibrium, suggesting there might be a value of f_0 that (assuming a well mixed population) would reduce part of the uncertainty defining the initial conditions. This is especially intriguing considering that the delayed growth bipartite model tested in [2] successfully predicts relations between the lag time and Λ_{max} assuming the total population is a bipartite one, characterized by two different adaptation times. Since death events play a central role in the model one could also explore the its dependence on growth factor. In [40] it was found that cells growing in high concentrations of growth factor had an increased susceptibility to cell death upon growth factor withdrawal.

A simple experimental test of some of the ideas considered in this thesis would consist of performing the same batch cultures in preconditioned media, to understand whether chemical signaling in combination with inoculum size is actually responsible for the modulation of growth parameters. Finally, considering the more general aim of developing reliable and predictive tumor growth models, essential features of population growth should be considered: including a description of the lag phase and phenotypic variability are the first two steps in this direction.

Chapter 5

Appendix

5.0.1 A

Calculations for k_{eff}

For $N_{tot} = K$ the total population must stop growing, there fore we set

$$\frac{dN_{tot}}{dt} = \lambda_1(K)N_1 + \lambda_2(K)N_2 + \lambda_3(K)N_3 = 0 \quad (5.1)$$

where

$$\lambda_1(K) = B_1(1 - \frac{K}{k_{eff}}) - D_1(1 + \frac{K}{k_{eff}}) \quad (5.2)$$

$$\lambda_2(K) = B_2(1 - \frac{K}{k_{eff}}) - D_2(1 + \frac{K}{k_{eff}}) \quad (5.3)$$

$$\lambda_3(K) = B_3(1 - \frac{K}{k_{eff}}) - D_3(1 + \frac{K}{k_{eff}}) \quad (5.4)$$

Explicitly we have

$$(B_1N_1 + B_2N_2 + B_3N_3)(1 - \frac{K}{k_{eff}}) - (D_1N_1 + D_2N_2 + D_3N_3)(1 + \frac{K}{k_{eff}}) = 0 \quad (5.5)$$

Yielding for $R = \frac{(B_1N_1 + B_2N_2 + B_3N_3)}{(D_1N_1 + D_2N_2 + D_3N_3)}$ and $A = \frac{K}{k_{eff}}$

$$R - RA = 1 + A \quad (5.6)$$

and eventually

$$\frac{R - 1}{R + 1} = \frac{K}{k_{eff}} \quad (5.7)$$

Bibliography

- [1] Baranyi, J. Roberts, T. A. A dynamic approach to predicting bacterial growth in food. *Int. J. Food Microbiol.* 23, 277–294 (1994).
- [2] Enrico Bena, C., Del Giudice, M., Grob, A. et al. Initial cell density encodes proliferative potential in cancer cell populations. *Sci Rep* 11, 6101 (2021).
- [3] Delignette-Muller, M. L. Relation between the generation time and the lag time of bacterial growth kinetics. *Int. J. Food Microbiol.* 43, 97–104 (1998).
- [4] Dufrenne, J., Delfgou, E., Ritmeester, W. Notermans, S. The effect of previous growth conditions on the lag phase time of some foodborne pathogenic micro-organisms. *Int. J. Food Microbiol.* 34, 89–94 (1997).
- [5] Rein, A. Rubin, H. Effects of local cell concentrations upon the growth of chick embryo cells in tissue culture. *Exp. Cell Res.* 49, 666–678 (1968).
- [6] Postma, J., Hok-A-Hin, C. Oude Voshaar, J. Influence of the inoculum density on the growth and survival of rhizobium leguminosarum biovar trifolii introduced into sterile and non-sterile loamy sand and silt loam. *FEMS Microbiol. Lett.* 73, 49–57 (1990).
- [7] Coleman, M., Tamplin, M., Phillips, J. Marmer, B. Influence of agitation, inoculum density, pH, and strain on the growth parameters of *Escherichia coli* o157:h7-relevance to risk assessment. *Int. J. Food Microbiol.* 83, 147–160. [https://doi.org/10.1016/S0168-1605\(02\)00367-7](https://doi.org/10.1016/S0168-1605(02)00367-7) (2003).
- [8] Irwin, P. L., Nguyen, L.-H.T., Paoli, G. C. Chen, C.-Y. Evidence for a bimodal distribution of *Escherichia coli* doubling times below a threshold initial cell concentration. *BMC Microbiol.* 10, 207 (2010).

- [9] Koutsoumanis, K. P. Lianou, A. Stochasticity in colonial growth dynamics of individual bacterial cells. *Appl. Environ. Microbiol.* 79, 2294–2301. <https://doi.org/10.1128/AEM.03629-12> (2013).
- [10] Marteiijn, R. C. L., Oude-Elferink, M. M. A., Martens, D. E., de Gooijer, C. D. Tramper, J. Effect of low inoculation density in the scale-up of insect cell cultures. *Biotechnol. Prog.* 16, 795–799. <https://doi.org/10.1021/bp000104d> (2008).
- [11] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, vol. 81, no. 25, pp. 2340–2361, (1977).
- [12] Gulik, W., Nuutila, A., Vinke, K., ten Hoppen, H. Heijnen, S. Effect of co2 airflow rate and inoculation density on the batch growth of catharanthus roseus cell suspensions in stirred fermentors. *Biotechnol. Prog.* <https://doi.org/10.1021/bp00027a015> (1994).
- [13] Kanokwaree, K. Doran, P. M. Effect of inoculum size on growth of atropa belladonna hairy roots in shake flasks. *J. Ferment. Bioeng.* 84, 378–381 (1997).
- [14] Carvalho, E. B. Curtis, W. R. The effect of inoculum size on the growth of cell and root cultures of *Hyoscyamus muticus*: implications for reactor inoculation. *Biotechnol. Bioprocess Eng.* 4, 287–293 (1999).
- [15] Gregório, A. C. et al. Inoculated cell density as a determinant factor of the growth dynamics and metastatic efficiency of a breast cancer murine model. *PLoS ONE* 11(11), 1–19. <https://doi.org/10.1371/journal.pone.0165817> (2016).
- [16] Rodriguez, N. E., Perez, M., Casanova, P. Martinez, L. Effect of Seed Cell Density on Specific Growth Rate Using CHO Cells as Model (2001).
- [17] Ozturk, S. S. Palsson, B. O. Effect of initial cell density on hybridoma growth, metabolism, and monoclonal antibody production. *J. Biotechnol.* 16, 259–278 (1990).
- [18] Pin, C. Baranyi, J. Kinetics of single cells: observation and modeling of a stochastic process. *Appl. Environ. Microbiol.* 72, 2163–2169. <https://doi.org/10.1128/AEM.72.3.2163-2169.2006> (2006).

- [19] Augustin, J. C., Brouillaud-Delattre, A., Rosso, L. Carlier, V. Significance of inoculum size in the lag time of *listeria monocytogenes*. *Appl. Environ. Microbiol.* 66, 1706–1710 (2000).
- [20] Robinson, T. P. et al. The effect of inoculum size on the lag phase of *listeria monocytogenes*. *Int. J. Food Microbiol.* 70, 163–173. [https://doi.org/10.1016/S0168-1605\(01\)00541-4](https://doi.org/10.1016/S0168-1605(01)00541-4) (2001).
- [21] Johnson, K. E. et al. Cancer cell population growth kinetics at low densities deviate from the exponential growth model and suggest an allee effect. *PLoS Biol.* 17, e3000399 (2019).
- [22] De Martino, D., Capuani, F. De Martino, A. Quantifying the entropic cost of cellular growth control. *Phys. Rev. E* 96, 010401 (2017).
- [23] Taheri-Araghi, S., Brown, S. D., Sauls, J. T., McIntosh, D. B. Jun, S. Single-cell physiology. *Annu. Rev. Biophys.* 44, 123–142. <https://doi.org/10.1146/annurev-biophys-060414-034236> (2015).
- [24] Jafarpour, F. et al. Bridging the timescales of single-cell and population dynamics. *Phys. Rev. X* 8, 021007. <https://doi.org/10.1103/PhysRevX.8.021007> (2018).
- [25] Hashimoto, M. et al. Noise-driven growth rate gain in clonal cellular populations. *Proc. Natl. Acad. Sci.* 113, 3251–3256. <https://doi.org/10.1073/pnas.1519412113> (2016).
- [26] Stephens, P. A., Sutherland, W. J. Freckleton, R. P. What is the Allee effect?. *Oikos* 87, 185–190 (1999).
- [27] MacIver, N. et al. Glucose metabolism in lymphocytes is a regulated process with significant effects on immune cell function and survival. *J. Leukoc. Biol.* 84(4), 949–957. <https://doi.org/10.1189/jlb.0108024> (2008).
- [28] Kotte, O., Volkmer, B., Radzikowski, J. L. Heinemann, M. Phenotypic bistability in *escherichia coli*’s central carbon metabolism. *Mol. Syst. Biol.* 10, 736 (2014).
- [29] Chu, D. Barnes, D. J. The lag-phase during diauxic growth is a trade-off between fast adaptation and high growth rate. *Sci. Rep.* 6, 1–15 (2016).

- [30] Basan, M. et al. A universal trade-off between growth and lag in fluctuating environments. *Nature* 584, 470–474 (2020).
- [31] Heinemann, M., Basan, M. Sauer, U. Implications of initial physiological conditions for bacterial adaptation to changing environments. *Mol. Syst. Biol.* 16, e9965 (2020).
- [32] De Martino, D., Capuani, F. De Martino, A. Growth against entropy in bacterial metabolism: the phenotypic trade-off behind empirical growth rate distributions in *E. coli*. *Phys. Biol.* 13, 036005 (2016).
- [33] De Martino, D. Masoero, D. Asymptotic analysis of noisy fitness maximization, applied to metabolism growth. *J. Stat. Mech: Theory Exp.* 2016, 123502 (2016).
- [34] De Martino, A., Gueudré, T. Miotto, M. Exploration-exploitation tradeoffs dictate the optimal distributions of phenotypes for populations subject to fitness fluctuations. *Phys. Rev. E* 99, 012417 (2019).
- [35] Archetti, M. C. G. Ferraro, D. A. Heterogeneity for igf-ii production maintained by public goods dynamics in neuroendocrine pancreatic cancer. *PNAS* 112(6), 1833–1838. <https://doi.org/10.1073/pnas.1414653112> (2015).
- [36] Wang, P. et al. Robust growth of *Escherichia coli*. *Curr. Biol.* 20, 1099–1103. <https://doi.org/10.1016/j.cub.2010.04.045> (2010).
- [37] Kiviet, D. J. et al. Stochasticity of metabolism and growth at the single-cell level. *Nature* 514, 376 (2014).
- [38] Matera, L. et al. Prolactin is an autocrine growth factor for the jurkat human t-leukemic cell line. *J. Neuroimmunol.* 79, 12–21. [https://doi.org/10.1016/S0165-5728\(97\)00096-9](https://doi.org/10.1016/S0165-5728(97)00096-9) (1997).
- [39] Avalos, B. et al. K562 cells produce and respond to human erythroid-potentiating activity. *Blood* 71, 1720–1725 (1988).
- [40] Vander Heiden, Matthew G. et al. Growth Factors Can Influence Cell Growth and Survival through Effects on Glucose Metabolism. *Molecular and Cellular Biology*, Sept. 2001, p. 5899–5912

- [41] Zwietering, M. H., Jongenburger, I., Rombouts, F. M. van't Riet, K. Modeling of the bacterial growth curve. *Appl. Environ. Microbiol.* 56, 1875–1881 (1990).