UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Dipartimento di Matematica "Tullio Levi-Civita"

Corso di Laurea Triennale in Fisica

Tesi di Laurea

A driver-passenger stochastic model for cancer progression

Relatore

Dott. Marco Formentin

Correlatore

Dott.ssa Anna Tovo

Laureando

Andrea Perin

Anno Accademico 2017/2018

Contents

Introduction

It should come as no surprise that the concept of mutation is of primary importance when discussing cancer. To anyone not knowing the rudiments of cancer phenomenology, it should suffice to know that cancer is originated when one or more healthy cells undergo some key somatic mutations to their chromosomes, and consequently become significantly different from the surrounding, healthy cells.

The insurgence of cancer disrupts the cellular environment: healthy tissues form a network of interactions that are finely tuned in order to avoid any possible deleterious effect, such as the extinction of one particular type of cell. Cancer, however, introduces many uncontrolled interactions with the healthy tissues. As a result, its impact on the network is almost always dramatic.

It is possible to think of the different types of cells as species in an ecologic environment: different types of healthy cells behave like different species interacting with each other. Cancer can be thought of as an additional species, which interacts with the network by favoring some species, damaging others, or both at the same time. However such interactions may be shaped, they are generally disruptive, in that they tend to prevent the ecologic system from attaining the "healthy" equilibrium [2].

By taking into account these effects on the cellular environment, one could summarize cancer's hallmarks as follows [4]:

- a self-sustained growth signalling caused by the so called *oncogenes*. Tumors can proliferate without needing any external stimuli;
- insensibility to growth-inhibition signals caused by *oncosuppressors*. In healthy cells, excessive proliferation is countered by specific molecules, which inhibit growth. Tumors, however, may not respond to these molecules;
- evasion of apoptosis, that is, the cell's "programmed death" mechanism;
- an unlimited replicative potential;
- alteration of metabolism and inducted *angiogenesis*, the physiological process by which new blood vessels are formed from preexisting ones;

- the ability of evading the immunitary system and invading healthy tissues, causing *metastasis*, the spread of cancer from the primary site to a secondary one; for example, a colorectal cancer might spread to the lungs through blood vessels.

All of these common traits of cancer are originated from chromosomal mutations, the effect of which can significantly vary from one another. All in all, there is a wide spectrum of possible mutations; some are more impactful than others and, most importantly, not all of them have a strictly positive effect on the cancer cell.
It is worth noting from now that, since the evolutionary process has been a long one, somatic alterations are far more likely to cause harm to the cell rather than bringing beneficial traits: a mutation, being a largely random event, will most likely disrupt the existent and well-oiled mechanisms rather than introduce a novel, advantageous feature.
The most interesting distinction one can make when discussing these alterations, is to classify them strictly on behalf of their effect on the cell's fitness, that is, the cell's reproductive advantage (or disadvantage). The easiest way to carry out this distinction is to coarsely divide between deleterious alterations and advantageous ones. Little further distinction is then given to alterations of analogous effect but with different magnitude.
With these concepts in mind, one can then define two different types o mutations: drivers and passengers [1]. The names are rather self-explaining: a driver mutation is a rare, advantageous mutation that confers its bearer an improved phenotype (that is, the set of an organism's observable characteristics); the aforementioned increased birth rate, for example, could be the result of a driver mutation having taken place. Passengers mutations, on the other hand, are alterations which have deleterious effects on the cell's fitness: they could provoke an immunitary response (effectively neutralising one of cancer's hallmarks, as discussed above), or make it more weak to pharmacological treatment.
The discovery and cataloguing of driver mutations has been one of the most sought after topics in genome-wide cancer sequencing. Drivers compose only a minimal fraction of all the possible (and observed) mutations. Moreover, they are somewhat recurring: different patients show many similar mutations, involving the same genes, loci (fixed positions on a chromosome) or pathways (an ensemble of molecular regulators that interact with each other and with other elements of a cell in order to regulate gene expression, that is, how much of the information carried by a gene is used by the cell itself).
Passengers, on the other hand, have not shared this same spot in the limelight. As they are not as easily trackable, are much more likely to happen and are far less incisive on cancer progress, they have been generally neglected.

Figure 1: A schematic overview of the steps leading to *CRC*. These key mutations can be thought of as drivers. To be precise, they are actually the first to arise, being the responsibles for the initial outbreak. Image taken from Kumar, Abbas, Aster, "Robbins and Cotran Pathologic Basis of Disease", Elsevier-Saunders, 2014, 9^{th} ed. [4]

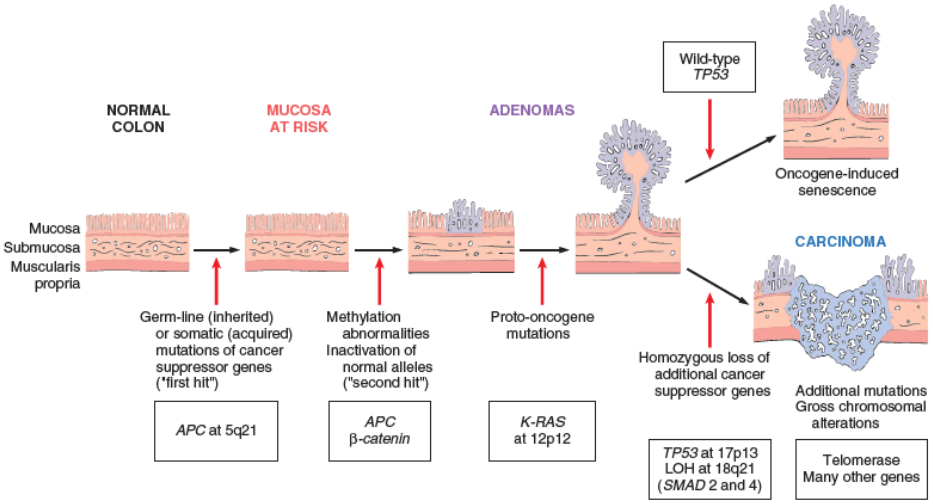


Figure 7-42 Molecular model for the evolution of colorectal cancers through the adenoma-carcinoma sequence. Although *APC* mutation is an early event and loss of *TP53* occurs late in the process of tumorigenesis, the timing for the other changes may be variable. Note also that individual tumors may not have all of the changes listed. *Top right*, cells that gain oncogene signaling without loss of *TP53* eventually enter oncogene-induced senescence. LOH, loss-of-heterozygosity.

Take colorectal cancer (CRC), for instance. The mutations responsible for its arisal have been described rather precisely; they follow a somewhat schematic order.

1. The first mutation involves the *APC* gene, disrupting its preventive function and resulting in the accumulation of the protein β-catenin;

2. The accumulation results in allele deactivation. It is then the gene *K-RAS* that gets mutated, leading to increased proliferation;

3. It is then the gene *TP-53* to be mutated; as a result, defective cells are no longer killed.

What about passengers? While their single effect is small, they greatly outnumber drivers; and yet, they are often left out of the picture. The question, then, is the following: *how do passengers influence cancer development?*

The situation we intend to study is that of a generic, although inspired by CRC, established cancer population of small size. We do not intend to study the initial setting of the neoplastic growth, but rather its evolution, under the assumption that two main types of mutation are possible, namely drivers and passengers [1].

The study will be carried out through stochastic simulations that will take into account the features described until now. Cancer cells will be subject to the following events:

- *reproduction*: the produced offspring will inherit *all and only* the features exhibited by the parent cell. This means that a cancer cell will simply double itself;
- *mutation*: a mutation event can occur, simultaneous to birth. Both advantageous and deleterious mutations will be possible, but of the two newborn cell, only one will be mutated, and with only one mutation;
- *death*: a cell might simply die, leaving its space to a luckier fellow.

In this thesis, two different scenarios will be studied using this general idea: in the first place, cancer will be studied as a standalone population; then, such a population will be studied as part of a network.

In the single population cancer model, the general biological model will be more precisely discussed. Stochastic simulations will be then described, and some preliminary outcomes will be shown. An estimation of a key factor in the evolution of cancer, the *critical population size* $N_{crit.}$, will follow. Lastly, possible cures based on the critical size will be tested.

In the second part, cancer will be part of a biological network where interactions of cancer with healthy cells will be present. This is loosely based on bone tissue, where this kind of dynamics has been previously observed and studied [2]. A brief discussion of the new dynamics will follow.

The parameters of the model are inferred from actual data.

Chapter 1

The single population cancer model

1.1 The outlines of the biological model

The stochastic model that is going to be studied from now on is the one briefly introduced above [1]. The neoplastic population is composed of a certain number of individuals. Each of them is characterized by the number of driver and passenger mutations. These mutations can be acquired while reproducing, and are inherited by the cell's offspring. There are then 4 key events we are going to consider; each of them will happen with a certain rate. Once again, to recap, a single cell can:

- give birth to two identical copies of itself; the rate for this happening depends on the cell's fitness, as will be shown later;
- give birth to two cells, one of which will be an identical copy, while the other one will present an additional passenger mutation. This event's rate is not dependent on the cell's parameters;
- give birth to two cells, one of which will be an identical copy, while the other one will present an additional driver mutation. This event's rate is not dependent on the cell's parameters;
- die, with a rate linked to the population's size.

The likelihood of both birth and death events depends on the number of driver and passenger mutations in each cell; it is also directly linked to the effect of the environment (for example, a better access to blood vessels should grant a better proliferation; on the other hand, less oxygenation should hamper the cell's progress).

In order to capture these first observations, the first and simplest hypothesis is to assume that all drivers (conversely, all passengers) confer the same advantage (conversely, disadvantage). If that is the case, then it becomes possible to characterize both birth and death rates for each cell via the number of drivers, d, and passengers, p, as well as the population size N:

$$B(d, p, N) \qquad\qquad D(d, p, N)$$

Let us name s_p the disadvantage conferred by each passenger mutation, and s_d the advantage of a driver. Then, by posing yet another assumption, which is aggregating the genetic component into the birth function, and only leaving the environmental effects to factor into the death function, which is an admittedly gross oversimplification of many size-related variables that influence and determine cell death, the following expressions can be used:

$$B(d, p) = \frac{(1 + s_d)^d}{(1 + s_p)^p} \qquad\qquad D(N) = \frac{N}{K} \tag{1.1}$$

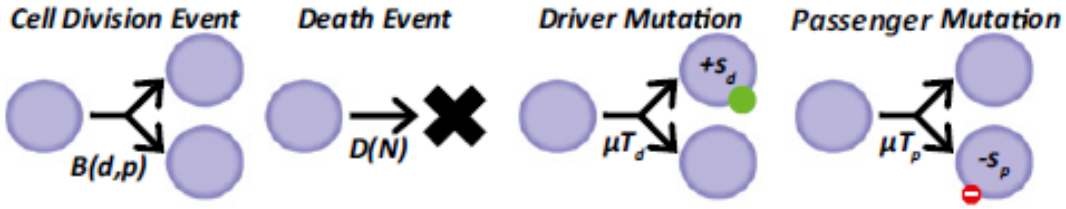
where K is the initial equilibrium size (reflecting the cancer microenvironment at the moment of the onset). Such a choice of functions may seem arbitrary, and this suspicion is partly true. However, the specific analytic formulas used have little consequence on the resulting dynamics. If the rates are chosen so that an additional driver (conversely, passenger) mutation increases (decreases) birth rate, then the qualitative behaviour of the resulting model will be similar to the one employed here.

Some other important parameters that will be used are:

- the *mutation rate* μ, which measures how likely to occur a mutation of any type is; measured in mutations per nucleotide (nt) per division, it is assumed constant, with a value $\mu \approx 10^{-8} nt^{-1} division^{-1}$;

- the *target sizes*, one for drivers, T_d, and one for passengers, T_p. These two parameters take into account the number of loci that, if altered, give rise to either a driver or a passenger mutation. As only a "handful" of mutations can result in actual benefits to the cell, whereas deleterious mutations have many more ways to arise, we can expect T_d to be significantly lower than T_p. A rough estimate gives $T_d \approx 700\, nt$, while $T_p \approx 5 \times 10^6\, nt$.

Figure 1.1: A visual representation of the four events that each cell can undergo. Juxtaposed to each event is the rate at which it happpens.
Birth rate $B(d,p)$ (eq. 1.1) depends on both the number of mutations and their strength, so that more drivers equate to an increased rate, while more passengers decrease it.
Death rate $D(N)$ (eq. 1.1) is only dependent on the population size N and on the equilibrium size, K; as population increases, so does the death rate, as to simulate environmental effects.
The rates for the two types of mutations, μT_d and μT_p, are only dependent on the mutation rate μ and the target sizes T_d, T_p. The former measures how likely a mutation is to occur, while the latter measure the number of nucleotides which, if mutated, lead to either a driver or passenger mutation.



It's worth noting that, if we consider a population at the equilibrium size, composed of inividuals with no mutations, we get a result that, hopefully, is convincent enough, as both death and birth functions are equal to 1: the equilibrium is, at least intially, preserved.

1.2 Description of the stochastic model

The stochastic model for cancer evolution which was described above has been given in terms of *probability rates*. Thus, it is worth to spend a few lines in order to explain how rates relate to the dynamics of the model [6].

One may think of the occurrence of cell division, death and mutation events as points on the real half line $\mathbb{R}^+$, representing time. Let us denote with $X_i, i = 1,2,3,4$ the random position of the time points where a transition of type $i = 1,2,3,4$ takes place. Here, types 1,2,3,4 correspond respectively to cell division, death, driver and passenger accumulation.

In order to shorten notations, let us set

$$\lambda_1 = B(d,p), \tag{1.2}$$
$$\lambda_2 = D(N), \tag{1.3}$$
$$\lambda_3 = \mu T_d, \tag{1.4}$$
$$\lambda_4 = \mu T_p. \tag{1.5}$$

Stating that a generic event Z has rate λ_z means that

$$\mathcal{P}(Z \in dt) = \lambda_z dt + \mathcal{O}(dt^2), \tag{1.6}$$

with an infinitesimal dt.

In order to achieve such a process, a possible approach (based on the Colouring Theorem, [5]) is to simulate a single process on $\mathbb{R}^+$, which we shall denote as X, with rate

$$\Lambda = \sum_{k=1}^{4} \lambda_k.$$

Such a process will only determine where one generic event occurs, but not the type of transition involved (Figure 1.2).

Figure 1.2: A representation of how the points of a generic process Z may be distributed on the real half line $\mathbb{R}^+$. There are no distinction between events.

To differentiate between the four possible types of event, we compute the following probability:

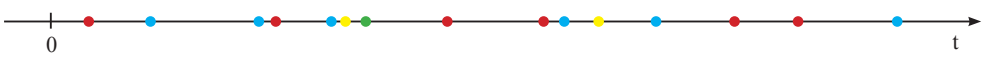
$$p_i = \frac{\lambda_i}{\sum_{k=1}^{4} \lambda_k}.$$

By doing so, each event will be of type i with probability p_i. As a consequence, the following holds true:

$$\begin{aligned} \mathcal{P}(X_i \in dt) &= \mathcal{P}(X \in dt) p_i && (1.7) \\ &= \left(\Lambda dt + \mathcal{O}(dt^2)\right) \frac{\lambda_i}{\Lambda} && (1.8) \\ &= \lambda_i dt + \mathcal{O}(dt^2) && (1.9) \end{aligned}$$

The result can be visualized as in Figure 1.3.

Figure 1.3: A representation of how the points of four generic processes $Z_i, i = 1, 2, 3, 4$ may be distributed on the real half line $\mathbb{R}^+$. The different colors represent the different types of events. Their overall distibution is the one determined by the process Z in Figure 1.2.



To recap, the model is as follows:

- choose a time interval dt on $\mathbb{R}^+$;
- evaluate how many events occurred in the time interval;
- choose a type for each event, based on the probabilities p_i.

Simulations

The most crucial aspect for simulations lies in calculating the rates. Birth and death functions have already been introduced; birth depends on the genetic outfit of the cell, while death follows population size linearly. The mutation events, however, are somewhat different: they are not dependent on the cell or on the population at large. As one could expect, the probability of developing a mutation is largely independent from either population size or previous mutations . As a consequence, by using the parameters introduced before, these fixed rates are:

- $\mu T_p \approx 5 \times 10^{-2} division^{-1}$ for passengers;
- $\mu T_d \approx 7 \times 10^{-6} division^{-1}$ for drivers;

As a result, mutations are more frequent in large populations. If by evolution we mean the process which leads cancer to change its overall genetic composition, accumulating driver and/or passenger mutations, then evolution becomes faster and faster as population size grows.

In the end, the four rates are $B(d, p)$ for birth, $D(N)$ for death, μT_p for passenger mutations and μT_d for driver mutations.

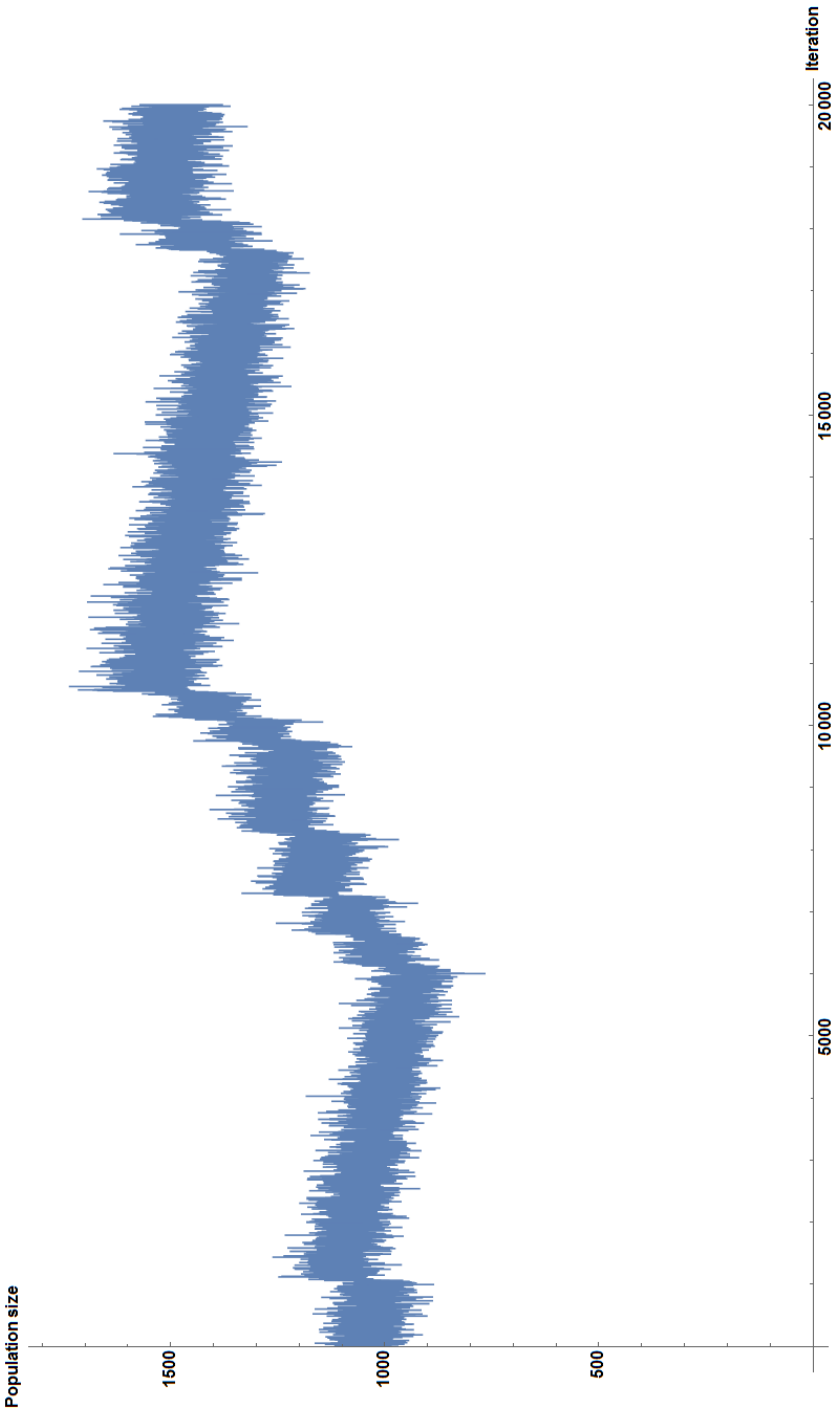
The first test has been carried out with the following parameters:

Table 1.1: The parameters used.

Parameter	Value
N	1000
K	1000
s_p	0.001
s_d	0.1
μT_p	$5 \times 10^{-2} division^{-1}$
μT_d	$7 \times 10^{-6} division^{-1}$

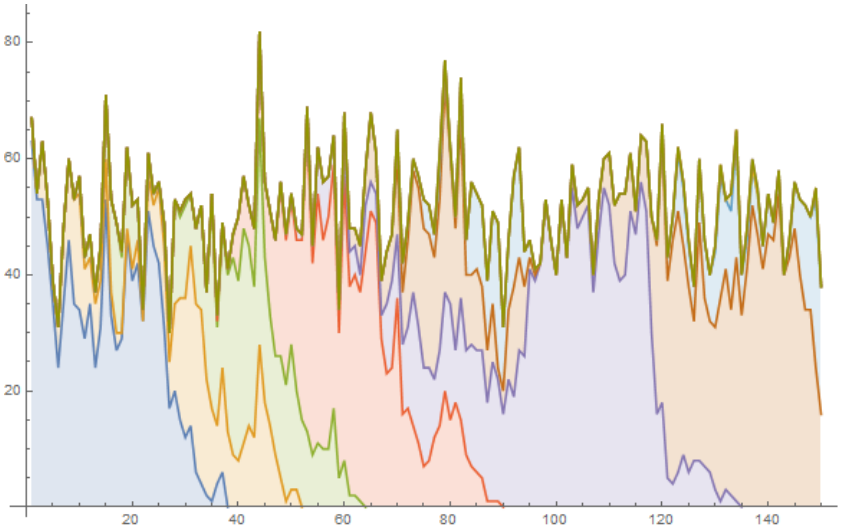
Results are shown in Figure 1.4. The dynamics of the evolution is interesting. There are some steep increases in population size, intertwined with longer, although less steep, decreases. An increase in population size means that the birth function, for a brief transient, is "overpowering" the death function. We can then infer that these increases are due to the appearance of a driver mutation in one cell.

Figure 1.4: An example of the behaviour of the population. The sawtooth shape is due to driver and passenger mutations: the appearance of a driver causes a steep increase in population size, while passenger accumulation causes a less steep decrease. These two processes are interwined.
Parameters in Table 1.1.



This particular individual will then be able to reproduce more than its peers; the resulting *clonal expansion*, which is the progressive take over of the whole population by one particular mutated group of cells, will increase the population size N. As a consequence, cells that have not developed a similar increase in fitness will find themselves struggling to stay alive, as the death function, being linear in N and equal for all cells, takes its greatest toll on these, now comparatively weaker, cells. Passengers, on the other hand, tend to accumulate in all cells. Figure 1.5 describes this accumulation process. The depicted section takes place between two driver-caused expansions. As iterations progress, the population becomes more and more filled with passengers, as the varying colour proves.

Figure 1.5: In x axis are iterations, while in y axis is population size. At each iteration, the overall population (represented by the green , bold line) is composed of cells with varying numbers of passengers. Each different colour represents the number of individuals with a certain number of passengers. Each of these "classes" of cells expands and then disappears, as passenger mutations are accumulated at every iteration.
Population size and number of iterations are deliberately small, as not to cramp the figure.



A result of the clonal expansion is the "hitchhiking" of passengers; the ones in the cell which developed the driver will then be present in all the cells of the population. Clonal expansion, and as a consequence, hitchhiking, can be visualized by analyzing the population in terms of number of cells per "passenger class", that is, the number of cells that have a precise number of passenger mutations. To do so, population has been divided into these classes before, during and after the fixation of a driver mutation. The results show that, before and after the driver-caused expansion, population is rather dishomogeneous: there are many passenger classes,

most of them are composed of a sizeable number of individuals. However, during the expansion, the number of classes is strongly reduced: the population becomes almost homogeneous, as the cell which has developed the driver mutation takes over the whole population, replacing the individuals which have not developed such a mutation. This situation is in good agreement with well studied cancer populations [7]: amongst different classes, the one with the higher fitness prevails.

Figure 1.6: The simulation from which the analysis of passenger classes was carried out. The driver-caused expansion which was analyzed is the one at (approximately) iteration 5000.

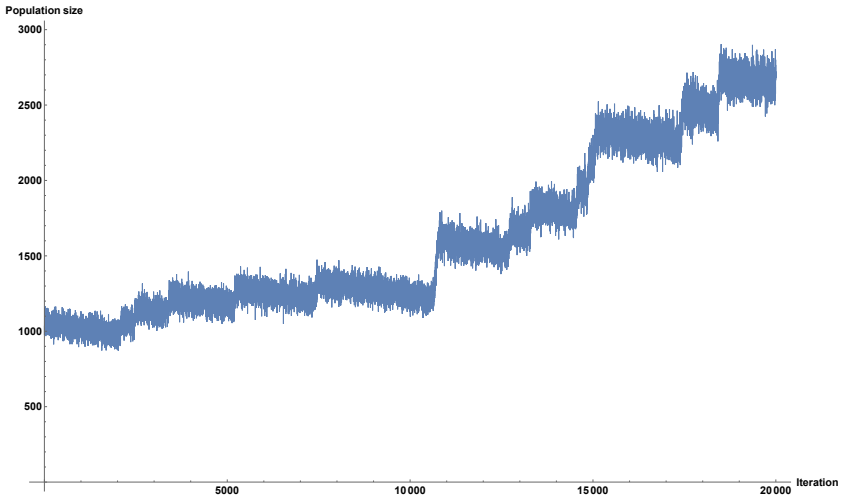


Figure 1.7: The composition of the population, in terms of number of passenger mutations, at iteration 5180, right before the expansion. Note how many different peaks are present; the population is composed of many subtypes of cancer cells.

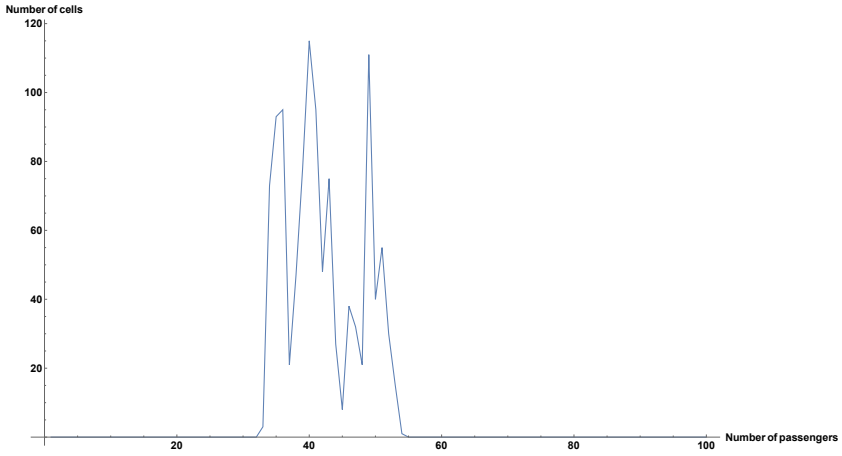


Figure 1.8: The composition of the population, in terms of number of passenger mutations, at iteration 5280, during the clonal expansion. Note how there is only one peak: population is now homogeneous, as there are few, very similar subtypes.

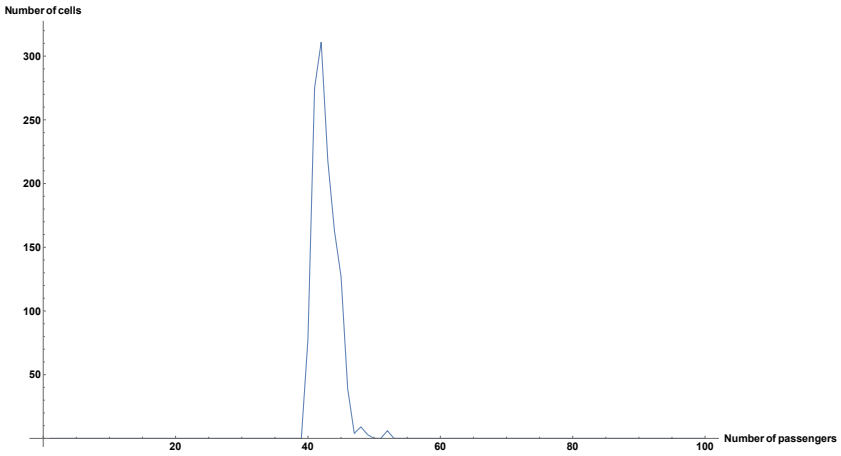
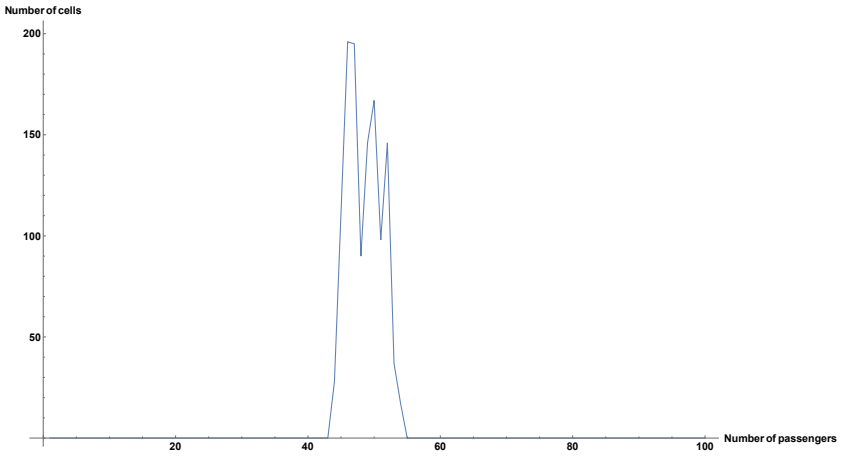


Figure 1.9: The composition of the population, in terms of number of passenger mutations, at iteration 5500, long after the expansion has ceased. Note how the population is once again dishomogeneous, as many peaks are once again present.



Reiterating simulations without changing any parameter leads to different kinds of outcome. Growth can be very fast, as seen in Figure 1.10; a large number of driver mutations arise quickly; the population grows accordingly. Growth can also be much slower, (Figure 1.11); if no driver mutation appears, population will decrease because of diminished fitness, a result of passenger accumulation. A burst of driver mutations makes so that, in the end, cancer grows back to its original size.

Figure 1.10: Parameters for this simulation in Table 1.1. In this simulation, the population develops many driver mutations in a short period of time. As a result, the increased population size makes it more likely that further mutations appear. As the simulation ended, cancer had grown to approximately four times its original size.

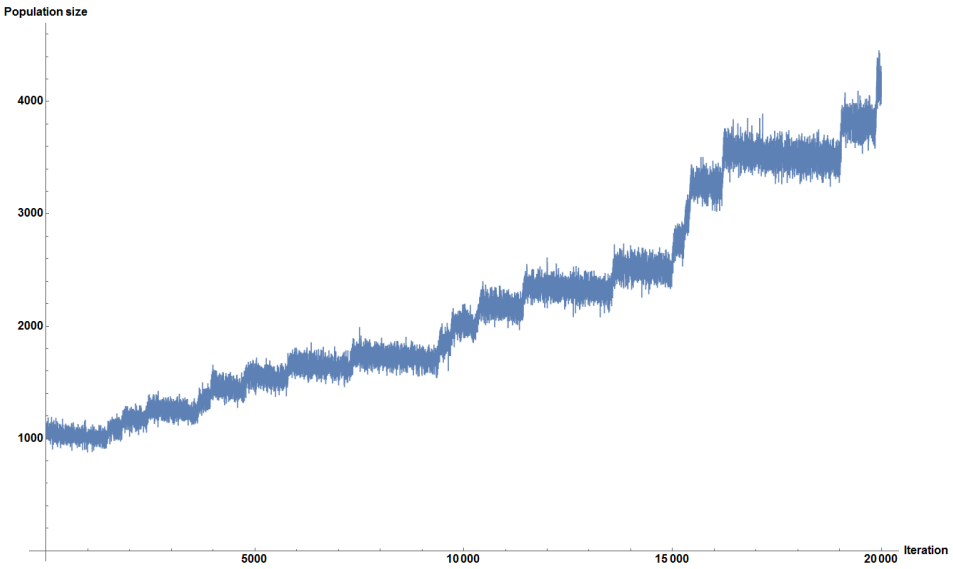
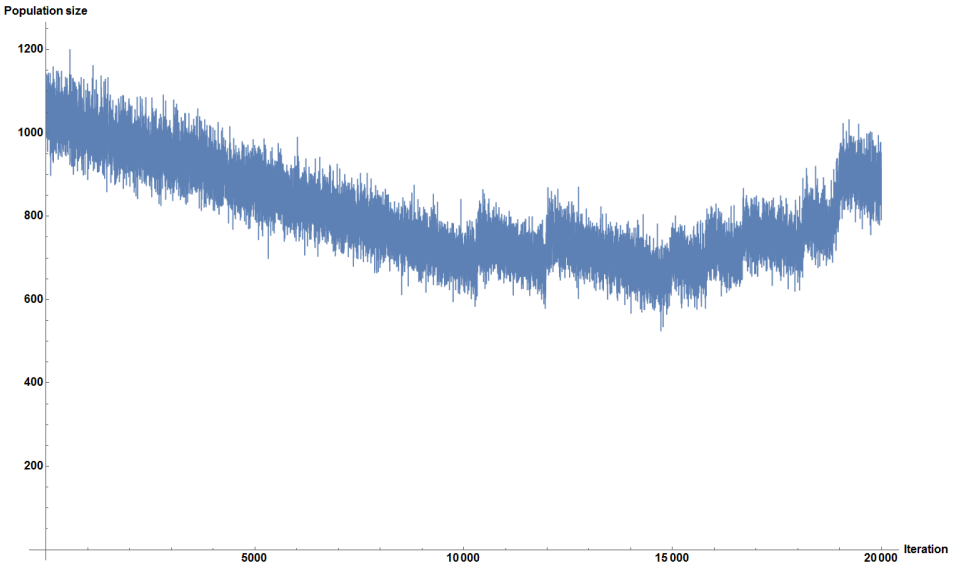


Figure 1.11: Parameters for this simulation in Table 1.1. For the first half of this simulation, population does not develop driver mutations. As the size decreases, so does the likelihood of further mutations appearing. Only in the end does cancer seem to avoid extinction, thanks to many driver mutations appearing in a short lapse, increasing birth rate and consequently population size.



Comparation of different models

Having established the first rough properties of this model, the next step is to variate some more intrinsic features.

The first modification was carried out by relaxing the hypothesis on s_p and s_d. By not assuming them to be the same for each mutation, they were instead drawn each time from a probability distribution. The ones that were tested are the *Lognormal* (Figure 1.12), the *Exponential* (Figure 1.13) and the *Gamma* (Figure 1.14) distributions. These distributions were chosen because a value drawn from them will always be positive; as such, they constitute good examples of the effect of randomly chosen s_p and s_d. Analytical forms and the chosen parameters can be found in Table 1.2.

Table 1.2: The distributions used and their parameters.

Distribution	Formula	Parameters
Lognormal	$\mathcal{LN}(x\|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma x}\exp\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right)$	$\mu_{pass.} = 0.001$ $\sigma = 1$
Exponential	$\exp(x\|\mu) = e^{-\mu x}\mu$	$\mu_{driver} = 0.1$ $\mu_{pass.} = 0.001$
Gamma	$\Gamma(x\|\theta,k) = \frac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^k\Gamma(k)}$	$\theta_{driver} = 0.1$ $\theta_{pass.} = 0.001$ $k = 2$

Qualitatively, the results are not significantly different from the initial model. The most evident difference lies, of course, in the magnitude of the "jumps" that the population is subject to after the arisal of each driver. Since the effect of these mutations is not fixed, such aleatory behaviour is not surprising. When it comes to the decline caused by the accumulation of passengers, however, the fixed effect model and the variable one are pretty much identical. This is not unexpected, though. When a particularly deleterious passenger arises, its effect on the cell's fitness is so dramatic that such a mutation is quickly weeded out by selection. On the other hand, if a passenger's effect is small enough, it is effectively neutral. Only slightly deleterious mutations cause a visible decrease in the population's size.

Figure 1.12: Except for s_p, parameters can be found in Table 1.1.
Simulation with s_p drawn from a Lognormal distribution, $\mathcal{LN}_p(x|\mu = s_p, \sigma = 1)$; formula and parameters in Table 1.2. The simulation is qualitatively similar to the one with fixed s_p.

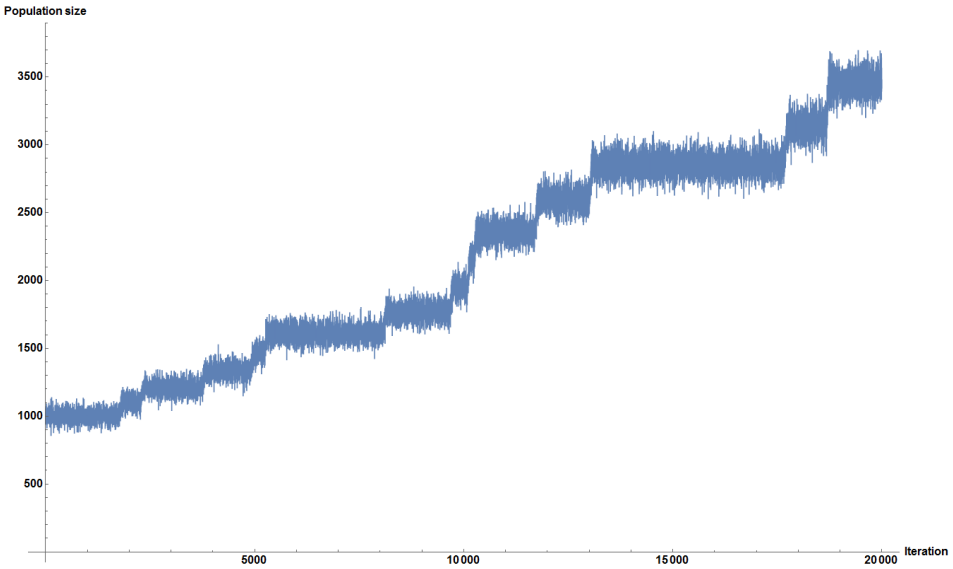


Figure 1.13: Except for s_p and s_d, parameters can be found in Table 1.1.
Simulation with s_p and s_d drawn from an exponential distribution, $\exp_{d,p}(x|\mu = s_{p,d})$; formula and parameters in Table 1.2. The simulation is qualitatively similar to the one with fixed s_p and s_d; however, since each driver mutation gives a random advantage, the increases in population size caused by driver mutations have random heights.

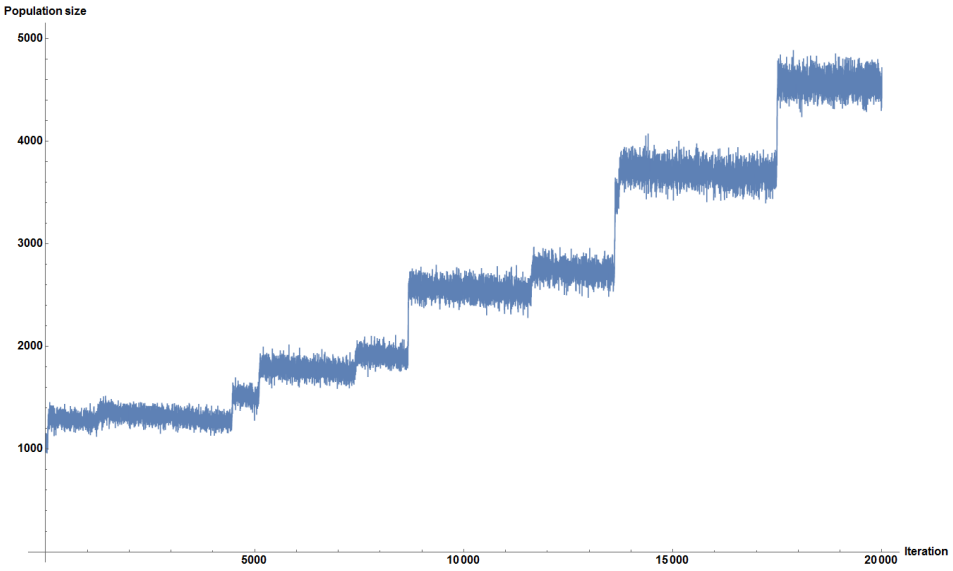
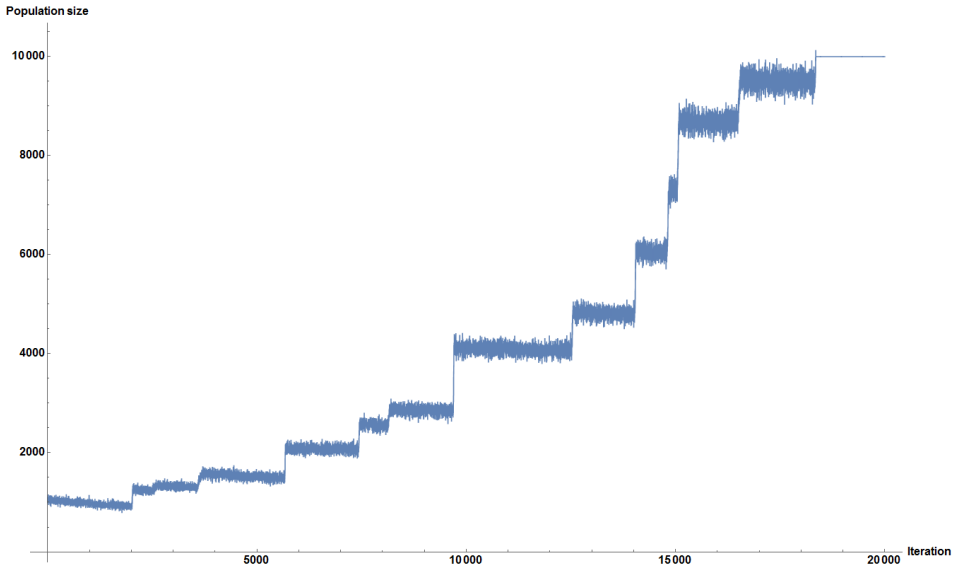


Figure 1.14: Except for s_p and s_d, parameters can be found in Table 1.1. Simulation with s_p and s_d drawn from a Gamma distribution, $\Gamma_{p,d}(x|k = 2, \theta = s_{p,d})$; formula and parameters in Table 1.2. The simulation is qualitatively similar to the one with fixed s_d and s_p; differences can be found in the magnitude of the increases in population size due to driver mutations. In the last portion of this simulation, cancer population exceeded the limit size of 10000 individuals; as a consequence, the run was stopped, thus producing the straight segment.



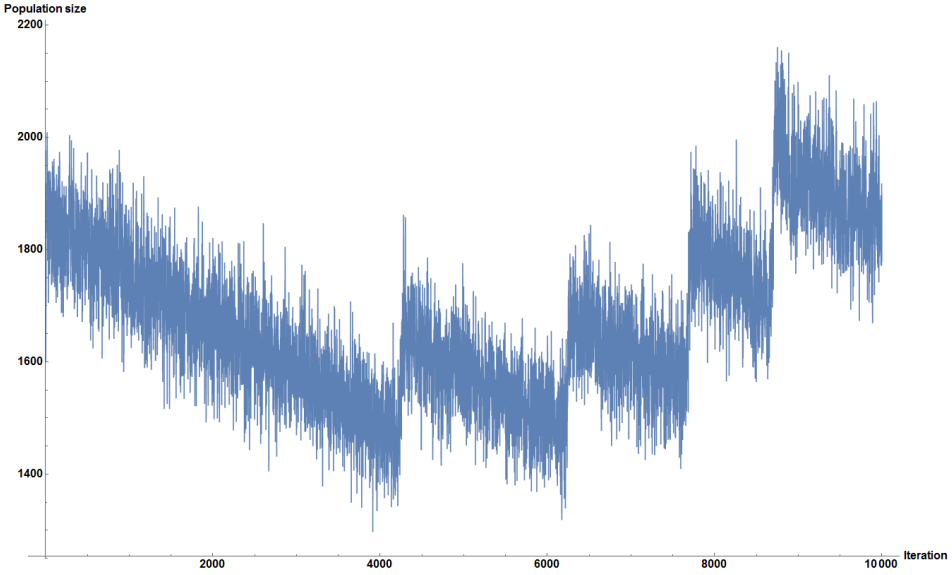
Another modification was then to change the death function. A *Gompertz*-like function was tested, with the following analytical form:

$$D(N) = \log\left(1 + \frac{N}{K}\right)$$

This function takes into account finite resources for the population. For small values of $\frac{N}{K}$, it will behave almost linearly. Greater differences are observed for higher values of the population size N. However, these differences are not qualitatively interesting: the dynamics stays the same (Figure 1.15).

The starting model, then, seems to capture the essential features of a driver-passenger cancer model: population evolves through quick expansions, caused by driver mutations, and through slow decreases, caused by passenger accumulation. The main advantage of the starting model lies in its relative simplicity: by assuming s_p and s_d constant, an approximated analytical tractation can be carried out.

Figure 1.15: Parameters can be found in Table 1.1. Simulation with Gompertzian death rate, $D(N) = \log\left(1 + \frac{N}{K}\right)$. The results are qualitatively similar to the ones where $D(N)$ is linear.



Exploring the parameters' ranges

Having established the stochastic model, some parameters were varied and their effect on the evolution was recorded. Amongst all possible variations, the most interesting ones involve the following:

- the equilibrium size K: this value is extremely important in determining the equilibrium levels between driver-caused expansions;
- the mutation strength s_p, more evidently linked to the birth function.

The parameter s_d was initially deemed not as interesting; its effects, being directly observable, are also more predictable (it will be shown in (1.55), however, that this is not exactly the case). The initial tests were centered around testing the model for different values of s_p. The effect of this change is visible in the steepness of the decrease between the driver-caused expansions.
For very low values of s_p, the decrease is almost unnoticeable (Figure 1.16); in this situation, passengers are effectively neutral. Their accumulation will not have a dramatic impact.
As the values of s_p increase, so does the steepness. A kind of sweetspot is found for s_p in the ranges from 0.003-0.01: the decrease is stronger in this spectrum of values (Figure 1.18). Interestingly, in most iterations, the population appears to be headed towards extinction (Figure 1.17).

Figure 1.16: Except for s_p, all parameters can be found in Table 1.1. In this simulation, $s_p = 0.0001$; notice how there is no downward slope between driver-caused expansions, as passenger mutations have such a small disadvantage that they can be considered effectively neutral mutations.

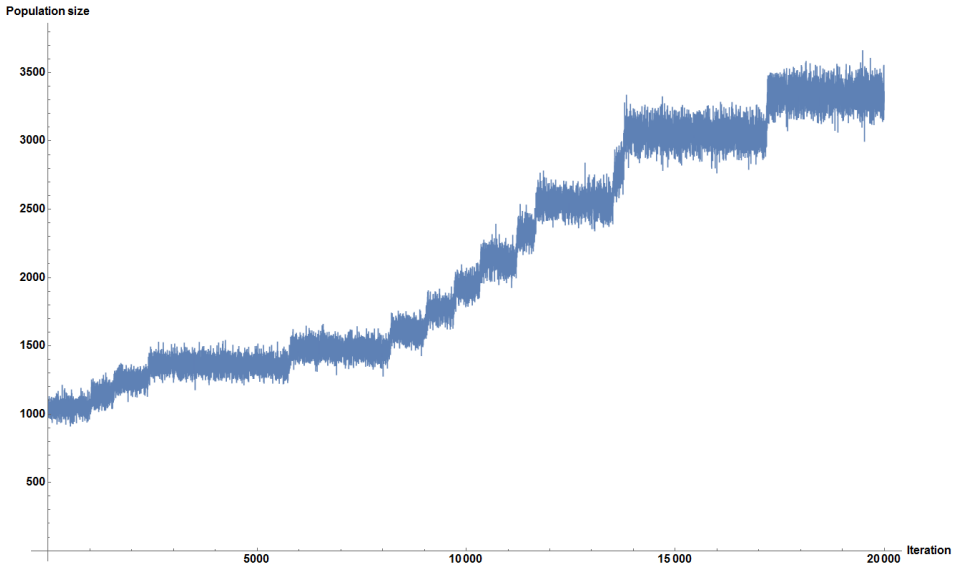


Figure 1.17: Except for s_p, all parameters can be found in Table 1.1. In this simulation, $s_p = 0.003$; there is a fast decline, and the population settles on a very low level. Cancer has spontaneously disappeared: as such, spontaneous remission is achieved.

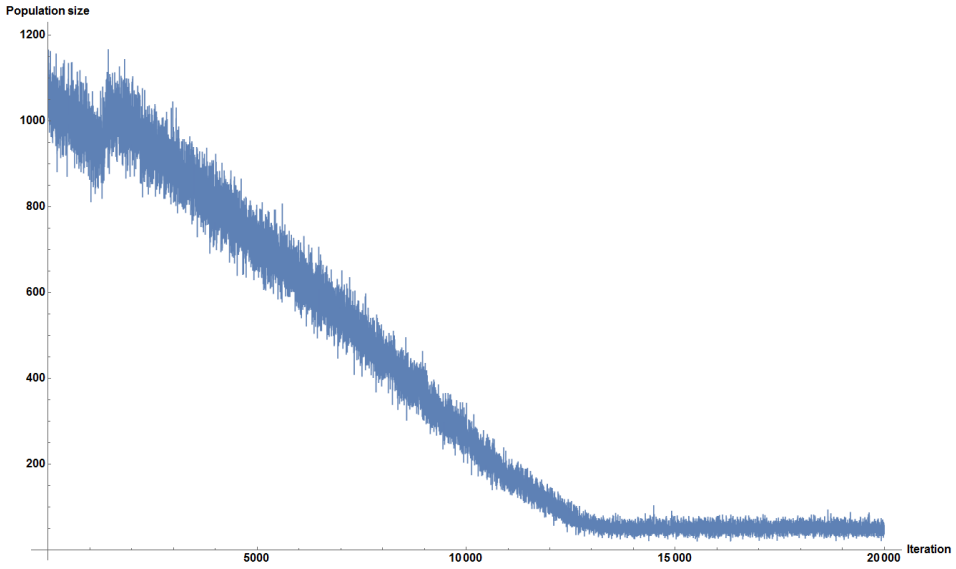


Figure 1.18: Except for s_p, all parameters can be found in Table 1.1. In this simulation, $s_p = 0.007$. Decrease is not as sharp as in the previous simulation; however, there is no explosive growth either. The steepness of the decrease caused by passenger accumulation is very high.

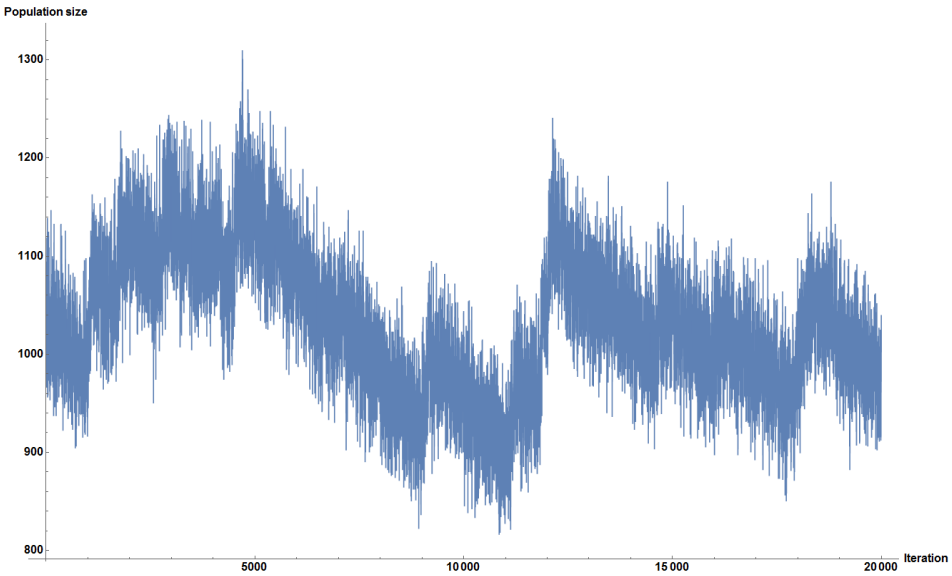
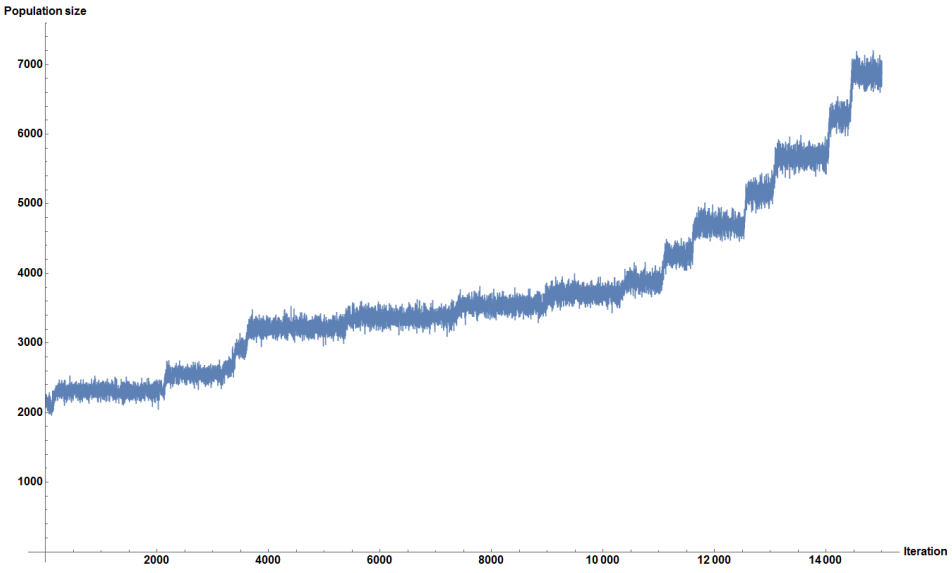


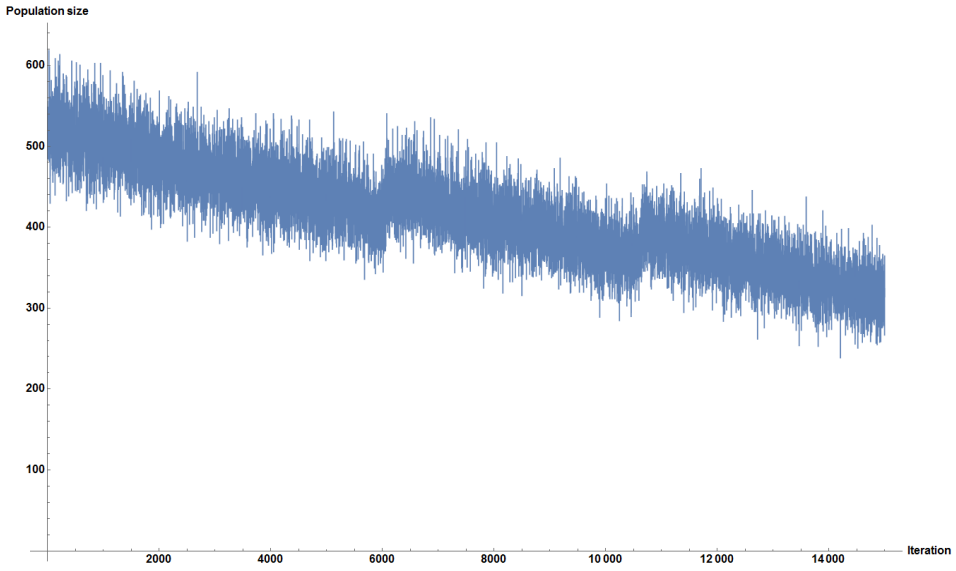
Figure 1.19: Except for s_p, all parameters can be found in Table 1.1. In this simulation, $s_p = 0.05$. Despite the great disadvantage conferred by passengers, in this simulation cancer thrived as if the disadvantage were close to none. This can be explained in terms of excessive disadvantage: as each passenger is so damaging, cells with a passenger mutation are eliminated too quickly to expand.



However, the trend that seemed to directly link higher values of s_p to faster decreases does not hold forever. As a certain threshold is surpassed, the steepness will once again become less and less noticeable (Figure 1.19).

By modifying K, while keeping all other parameters as they appear in Table 1.1, similar changes reappear (Figure 1.20). Population slowly progresses towards extinction, and the few driver mutations that arise through the simulation are not sufficient to avoid this fate.

Figure 1.20: Except for s_p, all parameters can be found in Table 1.1. Simulation with $K = 500$; just by changing the equilibrium size, population seems to be headed towards extinction, as driver mutations are too few to counter passenger accumultion.



Such a behaviour suggests that some kind of critical size exists, which roughly determines the evolution at large.

1.3 Estimation of the critical population size $N_{crit.}$

As it was previously shown, some combinations of the parameters s_p and K seem to be linked to spontaneous remission. It can be shown, using basic population dynamics and many simplifying assumptions, that under a *critical population size* the evolution will mostly progress towards spontaneous extinction [1].

The first step in order to formulate an analytical form for the critical population is invoking a rather harsh assumption:

$$B(d, p) = D(N) \tag{1.10}$$

This condition means that, at a given time, birth and death events are even. In this case, the process takes the shape of the so called "Moran process".

The Moran process

Consider a population of finite size N, composed of individuals (potentially of different types) which are subject to random and independent birth and death events. In each time step, an individual is randomly chosen for reproduction and one for death. In doing so, the offspring of the former replaces the latter, leading to a stochastic model with constant size. This model makes analytical treating feasible, at the cost of a general simplification.

In this model, the accumulation of driver and passenger mutations are considered as independent processes. By doing so, we can write the variation in population size as:

$$\frac{dN}{dt} = v_d + v_p \tag{1.11}$$

where v_d and v_p represent the change due to driver and passenger fixation [1] respectively. The main goal is now to find meaningful expressions for these velocities, reducing them to functions of the model's parameters.

Rates

Drivers and passengers arise respectively with rates $\mu T_d N$ and $\mu T_p N$; this means that drivers' and passengers' rates differ only because of the target size.

Fixation probabilities

Fixation probability is, intuitively, the probability of reaching a state in which the population is entirely composed of a single type of cells; in our model, it is the probability of a particular mutation taking over the whole population.

For a more formal definition, and the derivation of an expression for the fixation probability, let us consider a birth-death process of fixed size N, with two different types of cells, named A and B, which can die and reproduce *without* mutation: A produces A, while B produces B [3].

The only stochastic variable is the number of A-type cells, i. The number of B-type is therefore $N - i$. The resulting Moran process is defined on the state space $i = 0, ..., N$. In each time step, the state variable i can change to either $i-1$ or $i+1$, or remain unchanged. Let us denote α_i the probability of a transition from i to $i+1$, and β_i the probability of a transition from i to $i-1$. We have

[1] This term will be properly defined later.

$\alpha_i + \beta_i \leq 1$, so that the probability of remaining in state i is given by $1 - \alpha_i - \beta_i$. The states $i = 0$ and $i = N$ are called "absorbing states", meaning that once the process reaches either of them, no further change can be achieved (unless, of course, a new mutation arises). Therefore, we have $\alpha_0 = 0$ and $\beta_N = 0$. Denoting by x_i the probablity of reaching state N when starting from i, we have

$$x_i = \begin{cases} 0, & \text{if } i = 0 \\ \alpha_i x_{i+1} + \beta_i x_{i-1} + (1 - \alpha_i - \beta_i) x_i, & \text{if } 1 \leq i \leq N-1 \\ 1, & \text{if } i = N \end{cases} \tag{1.12}$$

If we now introduce the variables

$$y_i = x_i - x_{i-1} \tag{1.13}$$

$$\gamma_i = \frac{\beta_i}{\alpha_i} \tag{1.14}$$

for $i = 1, ..., N$, and substitute them in 1.12, we find

$$y_{i+1} = \frac{\beta_i}{\alpha_i} y_i = \gamma_i \ y_i \tag{1.15}$$

By using (1.15) recursively, we find $y_1 = x_1 - x_0 = x_1$, $y_2 = \gamma_1 x_1$, $y_3 = \gamma_1 \gamma_2 x_1$ and so forth; as a consequence,

$$y_j = x_1 \prod_{k=1}^{j-1} \gamma_k \tag{1.16}$$

Since

$$\sum_{j=1}^{N} y_j = x_1 - x_0 + x_2 - x_1 + ... + x_N - x_{N-1} = x_N - x_0 = 1, \tag{1.17}$$

it holds that

$$x_1 \left(1 + \gamma_1 + \gamma_1 \gamma_2 + \gamma_1 \gamma_2 \gamma_3 + ...\right) = 1, \tag{1.18}$$

from which

$$x_1 = \frac{1}{1 + \sum_{i=1}^{N-1} \prod_{j=1}^{i} \gamma_j}. \tag{1.19}$$

Note that this expression is already sufficient to the problem at hand, as it gives the probability that a mutation fixates starting from 1 individual. A further generalization, however, is possible. Indeed, let us consider the following equation:

$$x_i = \sum_{j=1}^{i} y_j = y_1 + \sum_{j=2}^{i} y_j = x_1 + \sum_{j=2}^{i} y_j \tag{1.20}$$

Inserting (1.16) into this last sum, we find

$$\sum_{j=2}^{i} y_j = \sum_{j=2}^{i} x_1 \prod_{k=1}^{j-1} \gamma_k = x_1 \left(\sum_{j=1}^{i-1} \prod_{k=1}^{j} \gamma_k \right) \tag{1.21}$$

which, substituted into (1.20) leads us to the following expression for x_i:

$$x_i = x_1 \left(1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} \gamma_k \right). \tag{1.22}$$

Using (1.19), we have

$$x_i = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} \gamma_k}{1 + \sum_{i=1}^{N-1} \prod_{j=1}^{i} \gamma_i} \tag{1.23}$$

thus giving the more general fixation probability starting from i individuals. As we have already pointed out, we are interested in the fixation probability ρ_A of one A-type cell, taking over $N-1$ B-type cells. Thus, the formula we are going to use is

$$\rho_A = x_1 = \frac{1}{1 + \sum_{j=1}^{N-1} \prod_{j=1}^{i} \gamma_j}. \tag{1.24}$$

For a simple Moran process, assuming both types of cells have equal fitness, the probabilities α_i and β_i can be easily derived. In the state i (meaning that there are i A-type cells), the probability of choosing an A-type cell (for either birth or death) is given by $\frac{i}{N}$, whereas for a B-type it is given by $\frac{N-i}{N}$. As a result, we have four possible outcomes at any given time step:

- two A-type cells are chosen for death and birth; this happens with probability $\left(\frac{i}{N}\right)^2$. The resulting transition is from state i to state i;
- two B-type cells are chosen for death and birth; this happens with probability $\left(\frac{N-i}{N}\right)^2$. Once again, the resulting transition is from state i to state i;
- an A-type cell is chosen for death, whereas a B-type cell is chosen for birth; this happens with probability $\frac{i(N-i)}{N^2}$. The resulting transition is from state i to state $i-1$;
- a B-type cell is chosen for death, whereas an A-type cell is chosen for birth; this happens with probability $\frac{i(N-i)}{N^2}$. The resulting transition is from state i to state $i+1$.

In this case, $\alpha_i = \beta_i = \frac{i(N-i)}{N^2}$, with the obvious consequence of having $\gamma_i = 1$ for every i. When substituted in 1.19, we find

$$x_1 = \frac{1}{1 + \sum_{j=1}^{N-1} \prod_{j=1}^{i} \gamma_j} = \frac{1}{1 + \sum_{j=1}^{N-1} 1} = \frac{1}{N} \tag{1.25}$$

This result is not surprising; since all individuals reproduce and die in the same rates, the probability of one particular individual taking over the whole population is precisely $\frac{1}{N}$. However, we are interested in a more realistic situation: what if the A-type, as a consequence of its mutation, obtained a different fitness value? To be more precise, let us suppose that B-type cells have fitness 1, while A-type cells have fitness r. Depending on the value of r, we can distinguish three possible scenarios:

- $r > 1 \rightarrow$ selection favors A;
- $r = 1 \rightarrow$ neutral drift;
- $r < 1 \rightarrow$ selection favors B.

In order to reflect this fitness change, we have to modify the probabilities of choosing A or B for reproduction:

- an A-type is chosen for reproduction with probability $\frac{ri}{(ri)+N-i}$;
- a B-type is chosen for reproduction with probability $\frac{N-i}{(ri)+N-i}$;
- an A-type is chosen for death with probability $\frac{i}{N}$;
- a B-type is chosen for death with probability $\frac{N-i}{N}$.

We can now calculate α_i and β_i for this process. A simple calculation, analogous to the one performed above, gives the following expressions:

$$\alpha_i = \frac{ri}{(ri) + N - i} \frac{N-i}{N} \tag{1.26}$$

$$\beta_i = \frac{N-i}{(ri) + N - i} \frac{i}{N}, \tag{1.27}$$

leading to

$$\gamma_i = \frac{1}{r}. \tag{1.28}$$

Therefore, by using (1.23), the probability of being absorbed in state N when starting in state i is given by

$$x_i = \frac{1 - \frac{1}{r^i}}{1 - \frac{1}{r^N}}, \tag{1.29}$$

and the fixation probability of a single A-type cell in a population of $N-1$ B-type individuals is

$$\rho_A = x_1 = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^N}}. \tag{1.30}$$

For an advantageous mutation, in the limit of $N \gg 1$, the following approximation will prove useful:

$$\rho_A \xrightarrow{N\to\infty} 1 - \frac{1}{r}. \tag{1.31}$$

In our model, the increased fitness that comes with a driver mutation (conversely, the decrease produced by a passenger mutation) is represented by the parameters s_d (respectively, s_p). By definition, we have

$$r = 1 + s_d \tag{1.32}$$
$$r = 1 - s_p \tag{1.33}$$

The latter, in particular, can be further approximated: since in the model at hand passengers are only slightly deleterious, $r \approx 1$. As a consequence, we can finally compute the fixation probabilities of driver (π_d) and passenger (π_p) mutations:

$$\pi_d = \frac{s_d}{1 + s_d} \approx s_d, \tag{1.34}$$
$$\pi_p \approx \frac{1}{N}. \tag{1.35}$$

Variation in population size

Once fixated, a mutation leads to a variation in the population size, which can be quantified by invoking the homeostasis condition already introduced in (1.10). The alterations in the population size are:

$$\Delta N_d = N_{d+1} - N_d, \tag{1.36}$$
$$\Delta N_p = N_{p+1} - N_p, \tag{1.37}$$

where the pedices identify the population sizes at equilibrium when $d+1$ or d drivers are present, or, similarly, when $p+1$ or p passengers are fixated.
To find ΔN_d, let us invoke condition (1.10) in states d and $d+1$:

$$B(d, p; N_d) = D(N_d) \tag{1.38}$$
$$B(d+1, p; N_{d+1}) = D(N_{d+1}). \tag{1.39}$$

Recalling the analytical forms of $B(d, p; \cdot)$ and $D(N)$, it follows that

$$B(d+1, p; N_{d+1}) = \frac{(1+s_d)^{d+1}}{(1+s_p)^p} = B(d, p; N_d) \cdot (1+s_d) \tag{1.40}$$
$$B(d+1, p; N_{d+1}) = D(N_d) \cdot (1+s_d) \tag{1.41}$$

Since

$$B(d+1, p; N_{d+1}) = D(N_{d+1}) \tag{1.42}$$

it holds true that

$$D(N_{d+1}) = D(N_d) \cdot (1+s_d) \tag{1.43}$$

and thus, expanding the right hand member with respect to N to the first order, we have

$$D(N_{d+1}) \approx D(N_d) + \frac{\partial D(N_d)}{\partial N} \Delta N_d. \tag{1.44}$$

Since $D(N) = \frac{N}{K}$, this expansion can be written as

$$D(N_{d+1}) \approx D(N_d) + \frac{\Delta N_d}{K} \tag{1.45}$$

By equating this last expression to the one obtained in (1.43), we have

$$D(N_d) + \frac{\Delta N_d}{K} \approx D(N_d) + s_d D(N_d) \tag{1.46}$$
$$\frac{\Delta N_d}{K} \approx s_d \frac{N}{K} \tag{1.47}$$
$$\Delta N_d \approx N s_d \tag{1.48}$$

By applying the same reasoning to ΔN_p, but this time using the following relation:

$$D(N_{p+1}) = B(d, p+1; N_{p+1}) = \frac{(1+s_d)^d}{(1+s_p)^{p+1}} = \frac{B(d, p; N_p)}{(1+s_p)} = \frac{D(N_p)}{(1+s_p)} \tag{1.49}$$

and once again expanding the leftmost member with respect to N to the first order, we find

$$\Delta N_p \approx -N s_p. \tag{1.50}$$

Estimation of the critical population size

We are now ready to find the analyitical expression for the critical population size. We have:

- the fixation probabilities:

$$\pi_d \approx s_d \qquad\qquad \pi_s \approx \frac{1}{N}$$

- the rates at which mutations arise:

$$\mu T_d N \qquad\qquad \mu T_p N$$

- the variation in the population size caused by the fixation of a mutation:

$$\Delta N_d \approx N s_d \qquad\qquad \Delta N_p \approx N s_p$$

The variation in the population size in a time unit is then the product of these factors; as a consequence, we can write

$$v_d = \mu T_d N \cdot s_d \cdot N s_d \tag{1.51}$$

$$v_p = \mu T_p N \cdot \frac{1}{N} \cdot (-N s_p)\,. \tag{1.52}$$

By substituting these expressions in equation (1.11), we find

$$\frac{dN}{dt} = \mu T_d s_d^2 N^2 - \mu T_p s_p N \tag{1.53}$$

$$= \mu T_p s_p N \left(\frac{N}{N_{crit.}} - 1 \right) \tag{1.54}$$

where

$$N_{crit.} = \frac{T_p s_p}{T_d s_d^2}. \tag{1.55}$$

$N_{crit.}$ represents an unstable fixed point: the population will decline for $N < N_{crit.}$, and will, on the other hand, increase if $N > N_{crit.}$.
Of course, one should not expect this formula to be precise; the number of approximations used makes so that this is just a rough indication of where the threshold actually is.

Testing the formula

Formula (1.55) can be tested. The interesting parameters (s_p and K) were once again varied, in order to collect many samples around the predicted critical size.

As the algorithm is time-expensive, simulations were not carried out until the eventual remission, nor were they unrestricted: a limit of 15000 iterations and a maximum population size of 10000 individuals were established, in order to get a rough estimate of the behaviour. The results seem to prove that approximation (1.55) for $N_{crit.}$ is rough but meaningful. Populations well under that size manifested a tendency of either stationary behaviour or, more rarely, explosive growth. Populations above, instead, are almost always headed towards unrestricted growth.

1.4 Further testing and possible cures

As demonstrated, the appearance of a critical size is roughly described by equation 1.55:

$$N_{crit.} = \frac{T_p s_p}{T_d s_d^2}.$$

This suggests some possible testing: by modifying s_p and μ, the critical size would change accordingly, hopefully leading to spontaneous remission even in situations where cancer size would leave few hopes.

In order to prove this, the model was modified: the evolution followed the same parameters as before, but, halfway, μ was changed first to 5 times its original value, and then to 50. In the first case, the effect is the opposite of the intended one, as population quickly expands. In the second, the outcome is completely different: size quickly increases, only to then become stable.

A similar procedure was then followed, but this time the modified parameter was s_p, which, halfway the simulation, was set to 5 times its starting value (Figure 1.23).

Such a great success was not always repeated when a surgery-like event was simulated. To do so, the population was halved at once; however, the success of this operation strictly depends on the size that was reached by cancer at the time of the surgery. Roughly twice the critical size is required for the survival of the cancer. As a consequence, remission is not always granted.

Figure 1.21: All parameters can be found in Table 1.1. In this simulation, μ was changed to five times its original value. However, this had an unwanted effect on the evolution of cancer, as it was made faster by the change.

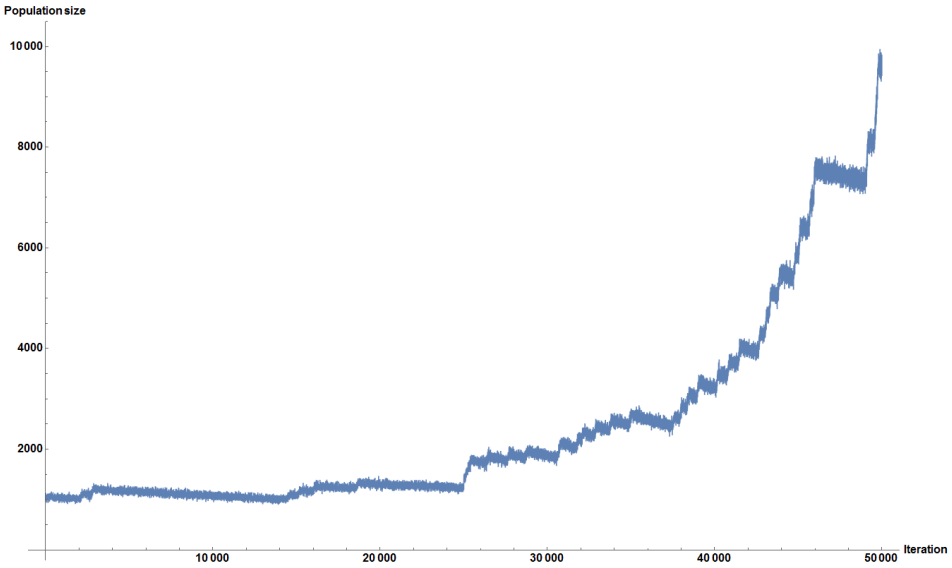


Figure 1.22: All parameters can be found in Table 1.1. In this simulation, μ was changed to fifty times its original value. However, this had an unexpected effect: the population became stationary, despite reaching a higher value. Despite the increased likelihood of mutations happening, the population does not increase nor does it decrease; cancer seems to have reached a chronic but stable condition.

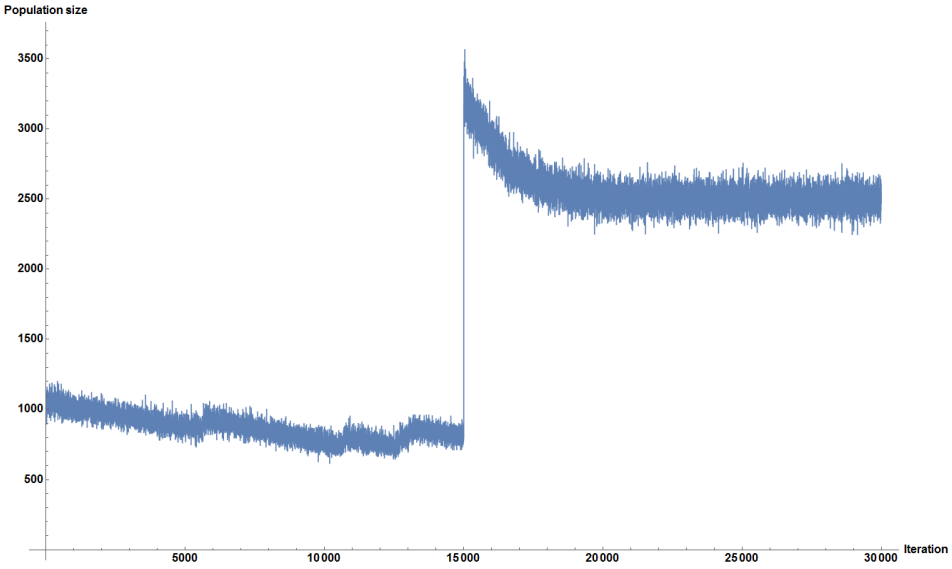


Figure 1.23: All parameters can be found in Table 1.1. In this simulation, s_p was changed to five times its original value. This is clearly more effective than the aforementioned change in μ (Figures 1.21 and 1.22).

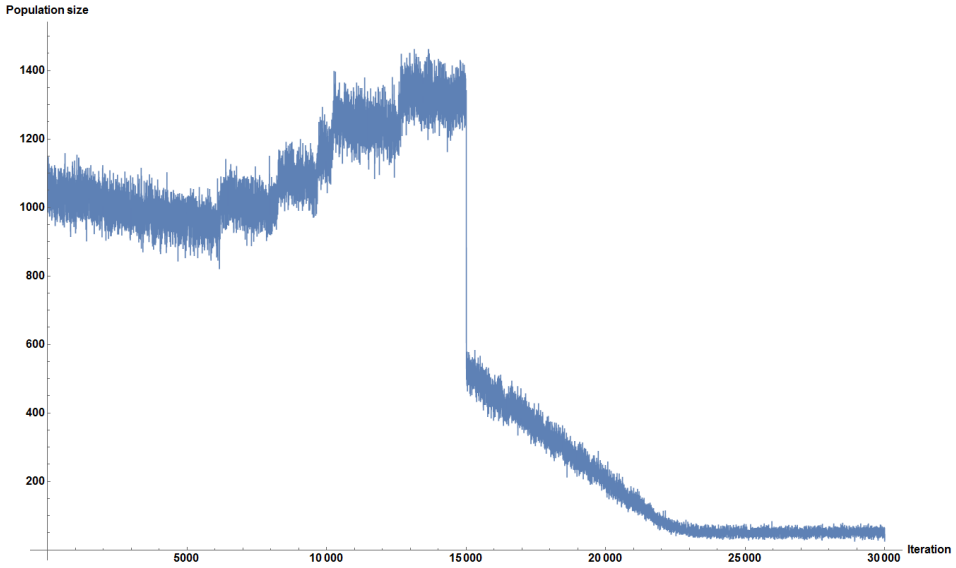


Figure 1.24: In this simulation, population was halved halfway through the evolution. As the cancer had not grown more than twice the critical size, this led to a hinted remission.

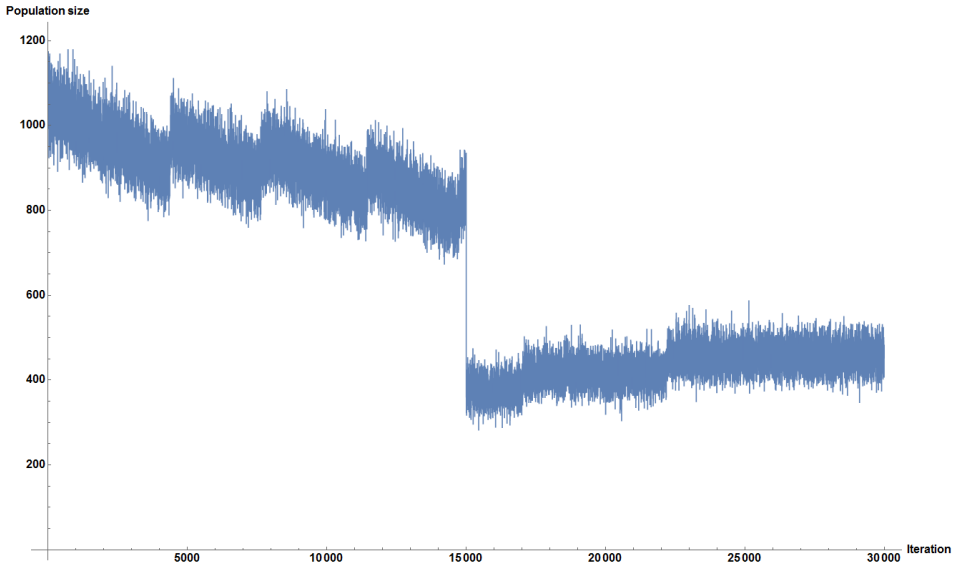
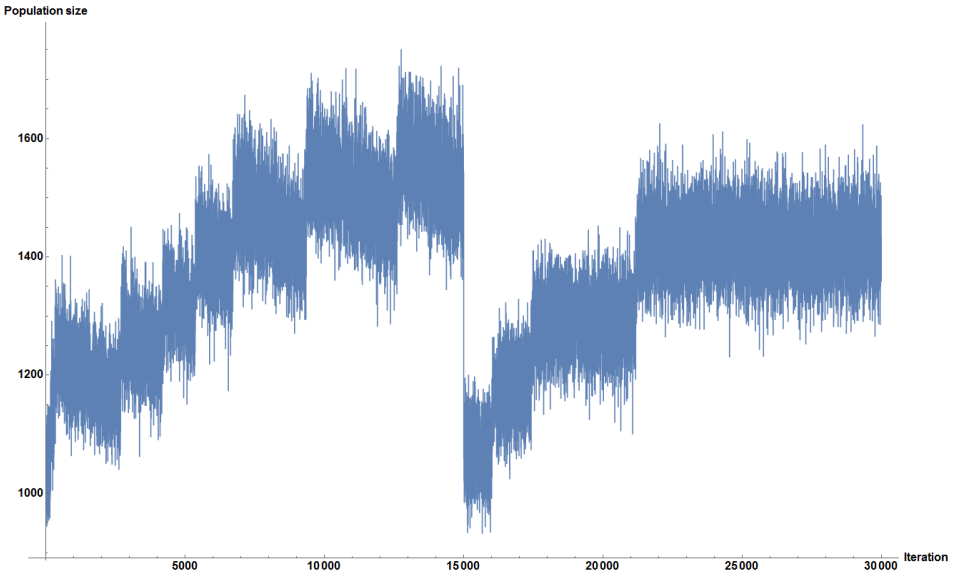


Figure 1.25: In this simulation, cancer had already outgrown the limit of $2N_{crit.}$. The operation was quickly forgotten, as cancer continued its growth.



By looking at these results, a hypothetical cure where passengers are targeted seems the most promising one.

Chapter 2

Cancer in a network with healthy cells

2.1 The biological framework

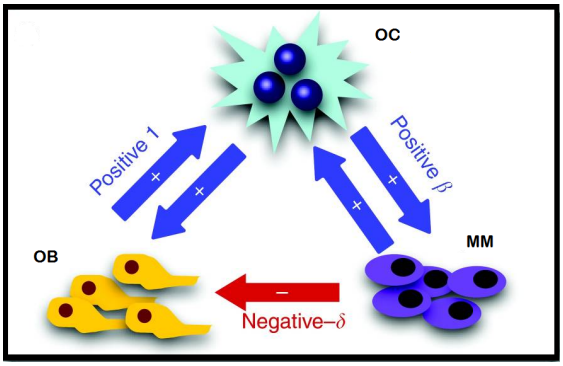
In the following section, a neoplastic population the likes of which have been extensively described above, is placed into an interacting network. The biggest inspiration for such a model comes from *multiple myeloma bone disease (MM)* [2]. As the name suggests, this particular kind of cancer arises in the bone tissue. The main feature of this microenvironment is the presence of two main types of interactive cells, namely *osteoblasts (OB)* and *osteoclasts (OC)*[1]. Simply put, osteoblasts are responsible for the building of the bone, while osteoclasts are the ones that destroy it. This kind of relationship between these two components is responsible for the bone's exceptional properties of self reparation, renewal and adaptability, as old tissue is continuously replaced by newer one. Furthermore, the relation between the two populations is symbiotic: each of them promotes the development of the other.

Multiple myeloma bone disease disrupts this refined mechanism: the cancer cells, named *MM*, act as a third (and rather unwated) actor, and interact with OC and OB cells in a somewhat counterintuitive way. They maintain a symbiotic relatioship with OC cells, promoting their growth and being favored in turn. On the other hand, they damage OB cells, receiving no harm from them in turn.

As strange as it may seem, however, game theory studies on this model suggest that the symptoms of multiple myeloma are caused by the overtaking of MM cells on the general population, rather than by the deleterious effect on OB cells. An

[1] *Osteocytes*, the cells which compose tha large majority of the bone's living matter, are not as interesting to consider, as their role is rather "static": they are produced and destroyed by OCs and OBs, but they do not directly take part in the network.

Figure 2.1: A visual representation of the biological model at hand. Image adapted from [2].



attempt will be made to capture the features of the deterministic model, and at the same time maintaining the key driver-passenger duality.

2.2 Description of the new stochastic model

The new stochastic model is an extension of the former. Identifying the population sizes by their respective names, and putting $Tot = OB + OC + MM$, birth and death rates for each group of cells will be modelled as in Table 2.1. Each death rate follows the linear formula already used for cancer cells; $K_{<pop>}$ is each group's equilibrium size. The interaction parameter α describes the normal dynamics between healthy cells, whereas β describes the interaction between MMs and OCs. Finally, δ describes the damage caused by MMs to OBs.

Table 2.1: The new rates.

Cell pop.	Birth rate	Death rate
OC	$1 + \frac{\alpha \cdot OB + \beta \cdot MM}{Tot}$	$\frac{OC}{K_{OC}}$
OB	$1 + \frac{\alpha \cdot OC - \delta \cdot MM}{Tot}$	$\frac{OB}{K_{OB}}$
MM	$B(d,p) + \frac{\beta \cdot OC}{Tot}$	$\frac{MM}{K_{MM}}$

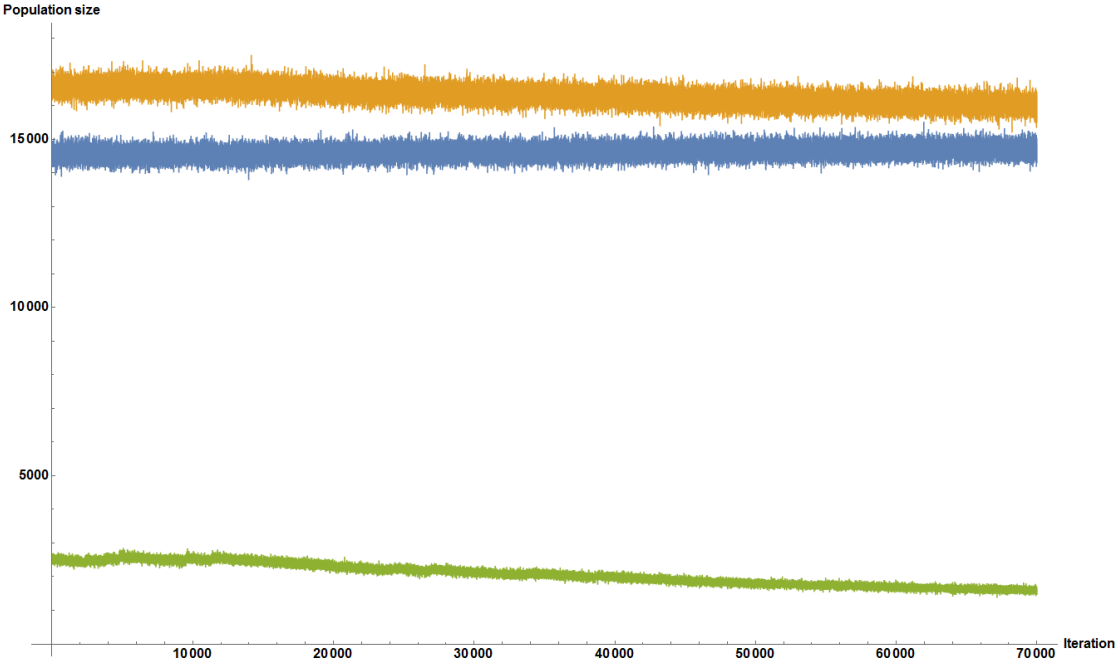
Table 2.2: The parameters used.

Parameter	Value
s_d	0.1
s_p	0.003
K	1000
α	1
β	2
δ	0.5
$N_{crit.}$	≈ 2100

As birth rate is now increased because of the interaction term $\frac{\beta \cdot OC}{Tot}$, the previous critical population size estimate $N_{crit.}$ is not expected to hold true any longer. A

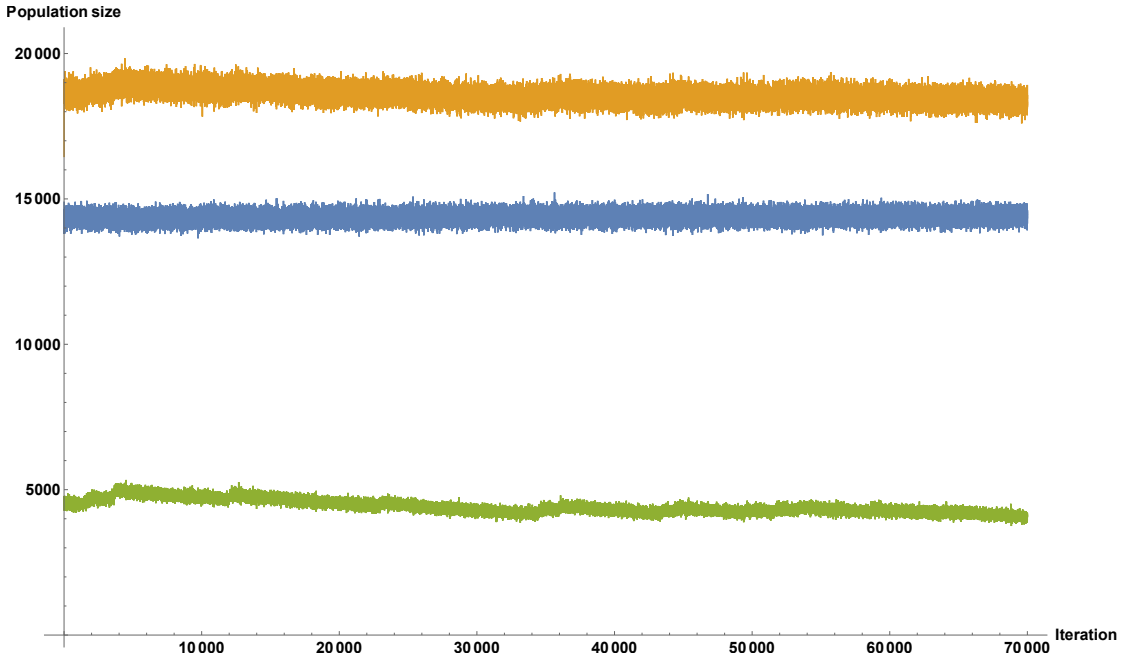
test is carried out (Figure 2.2), with a combination of parameters s_p and K that, in the single cancer population, would correspond to $N < N_{crit.}$; interaction terms have weights α, β and δ that correspond to an enhanced interaction between MMs and OCs when compared to the one between OCs and OBs, and to a weak deleterious effect of MMs on OBs. Specific values are collected in Table 2.2.

Figure 2.2: The blue line represents OB cells, the orange one represents OC cells and the green one represents MM cells. In this simulation, cancer progress is slowly but surely heading towards remission. The small "bumps" that can be seen along the trajectory are driver mutations arising; however, their effect is trumped by the overall reduction caused by passenger accumulation. The birth rate of MM cells is increased by the interactions with the healthy cells; however, these interactions are not sufficient to avoid the reduction that cancer is undergoing.



Further testing with increasingly higher interaction parameters led to no progress in terms of observing an eventual cancer growth. For instance, in Figure 2.3, $\beta = 4$, while all other parameters were left unchanged. The resulting evolution bears a strong resemblance to the one in Figure 2.2.

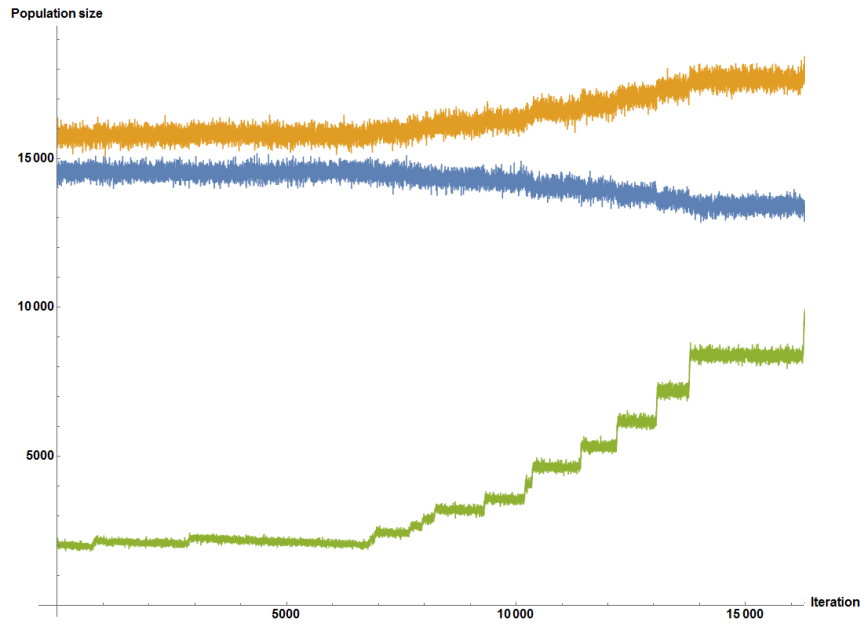
Figure 2.3: The blue line represents OB cells, the orange one represents OC cells and the green one represents MM cells. In this simulation, all parameters can be found in Table 2.2, except for β, which was set to 4. Behaviour is qualitatively similar to the one depicted in Figure 2.2, as cancer size is relatively constant throughout the simulation. Drivers and passengers alike do not influence evolution as they previously did in the single population model.



By trying to decrease the interaction parameters and by simultaneously doubling the mutation strength of both drivers and passengers, however, cancer behaved similarly to what has been observed previously in the single population model. In Figure 2.4, the parameters used are those that can be found in Table 2.2. However, now $s_p = 0.006$ and $s_d = 0.2$. The resulting dynamics allows cancer to reach the threshold size of 10000 individuals in a relatively brief interval. This time, unlike in Figure 2.3, driver mutations are significant enough to cause clonal expansions. The reason for doubling s_p and s_d is to be found in the relative advantage that is conferred by drivers and passengers. In the single population model, both $B(d, p)$ and $N(K)$ were initially close to 1. The appearance of either a driver or a passenger had a relative impact on the fitness of the order of, respectively, 10% and 0.1%. In this model, however, the interaction terms are added to what we may refer to as the "intrinsic" birth rate, that is, the function $B(d, p)$. As can be seen in these simulations, the interaction terms make so that the intial equilibrium between birth and death events for cancer is set at approximately 5000 individuals

(Figure 2.3). As a result, we can infer that the rates of both birth and death are approximately equal to 5. In this scenario, an increase of approximately 0.1 in the intrinsic birth rate will lead to no appreciable increase. Similarly, passengers' impact will be less noticeable in this situation, and they will be effectively neutral. Thus, both the absence of sharp increases in cancer population size and the avoidance of an eventual remission are explained in terms of how s_d and s_p relate to birth and death rates as a whole. By resetting the interaction terms to the ones in Table 2.2, and simultaneously doubling s_p and s_d, a situation similar to the initial single population model was achieved. The effect of mutations is now much clearer.

Figure 2.4: The blue line represents OB cells, the orange one represents OC cells and the green one represents MM cells. In this simulation, all parameters can be found in Table 2.2, except for s_p and s_d, which were set to 0.006 and 0.2 respectively. Behaviour of cancer is qualitatively similar to the ones depicted in the single population model. Note how OCs and OBs are influenced by the increasingly higher cancer population. The last, flat portion of this graph is due to the reaching of the threshold size of 10000 individuals by cancer; as a result, cancer was deemed to be growing to the eventual demise of the hyopthetical patient, and the simulation was stopped.



The results of simulations the likes of which have been depicted above (Figures 2.2 and 2.3), however, suggest a hypothetical "chronic" scenario [7]. Cancer, in these simulations, is not able to grow over the size that interactions dictated. As

a consequence, in this conditions, chronicity is attained.

It seems then that a necessary condition for a driver-passenger influenced cancer to grow in a network is for mutation strength s_d and interaction terms to be in a ratio such that the fitness increase granted by drivers is approximately 10% of the total, as was determined in Figure 2.4. This is not, however, a lower bound; it is merely a sufficient ratio.

Conclusions

The model discussed in this thesis resulted to adequately describe the progress of cancer. Its key features, namely the occurrence of spontaneous remission, dormancy and growth are observed in cancer; also, the fact that, at least initially, cancer has a somewhat homogeneous composition, can be found in this model, as after a driver-caused expansion, the population is primarily composed of cells with that particular driver mutation. Moreover, this model provides an explanation for the "hitchhiking" phenomenon. Passenger mutations which rise in a cell where a driver has just appeared will spread to the whole population because of the effect of such driver, despite not giving their host any advantage, and rather being deleterious.
All of these features are consequence solely of the driver-passenger duality: no other mechanism was necessary to have these peculiarities arise. Indeed, one may investigate whether all of the features of the present model are necessary and, if that were not the case, which features could be removed without disrupting the present qualitative behaviour. For instance, since cells with a different number of drivers are regarded as different species, the model studied in this thesis has potentially infinitely many species. If such complication could be avoided, an analytical approach could be more feasible.

The study of a cancer population in a network has shown significant difference when compared to a single cancer population model. The most interesting one is that the formula for the critical size that was found for a single cancer population does not hold true any longer.

One possible research direction could consist in a coarse graining of this model, by approaching it via a branching-process with only three species. By dividing the cancer population in three different classes on behalf of their birth rate, one could schematically represent the distinction between cells where passengers' effects are predominant, cells which have an approximate balance between the two types of mutation's effects, and lastly cells for which drivers' effects are dominant. By doing so, a computationally tractable model might be attained; however, the actual agreement between such a model and the one studied in this thesis should be tested beforehand.

Bibliography

[1] C. D. McFarland, K. S. Korolev, G. V. Kryukov, S. R. Sunyaev, L. A. Mirny, *Impact of deleterious passenger mutations on cancer progression*, www.pnas.org/cgi/doi/10.1073/pnas.1213968110, 2013

[2] J. M. Pacheco, F. C. Santos, D. Dingli, *The ecology of cancer from an evolutionary game theory perspective*, Interface Focus 4:20140019, http://dx.doi.org/10.1098/rsfs.2014.0019, 2014

[3] M. A. Nowak, *Evolutionary dynamics*, Harvard University Press, 2006

[4] V. Kumar, A. K. Abbas, J. C. Aster, *Robbins & Cotran Pathologic Basis of Disease*, Elsevier-Saunders, 9^{th} ed., 2014

[5] J. F. C. Kingman, *Poisson Processes*, Oxford University Press, 1993

[6] P. Brémaud, *Markov Chains, Gibbs fields, Monte Carlo simulations and queues*, Springer, 1998

[7] A. Brioli, L. Melchor, M. Cavo, G. J. Morgan, *The impact of intra-clonal heterogeneity on the treatment of multiple myeloma*, British Journal of Haematology, 2014, 165, 441-454, doi:10.1111/bjh.12805