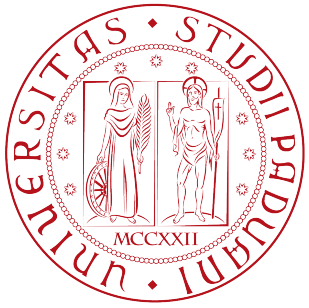

UNIVERSITY OF PADUA
DEPARTMENT OF MATHEMATICS "TULLIO LEVI-CIVITA"
MASTER DEGREE IN MATHEMATICS

MASTER THESIS IN MATHEMATICS

A STOCHASTIC MODEL FOR IMMUNOTHERAPY OF CANCER

Supervisor
Prof. Marco Formentin

Author
Marina Gionfriddo
Student Number
1108499

2020-2021
2021 April 23

Contents

Introduction

In recent years, the use of mathematics in cancer research has caught on, with the rapid accumulation of data and applications of mathematical methodologies [56]. The application of mathematics in cancer research is known as *mathematical oncology*. Mathematical oncology, starting from theoretical studies, tries to design clinical experiments with mathematical models. Mathematical models represent an useful tool for an interdisciplinary approach to cancer research. Indeed, various points of view, coming from several scientific areas, are fundamental to face the complexity of cancer evolution. In this setting, cancer ecology arises as promising quantitative approach. Cancer ecology looks at different groups of cells in an organism as interacting species of an ecosystem. From this perspective, cancer cells are a new species appearing in a stable ecosystem. Cancer cells represent a harmful and invasive species, abling to influence and change the interaction among the different types of healthy cells, that represent the pre-existing species. To promote their growth, cancer cells trigger a struggle for survival, which can lead to the extinction of certain types of cells and, in the worst cases, to the collapse of the entire ecosystem. The stochastic models (specially interacting particle systems) fit the noisy dynamics of cancer. One of the main problematic features of cancer is therapy resistance. Among therapies, immunotherapy has some notable characteristics [68, 67]:

- it enables the immune system to recognize and target cancer cells;

- it can train the immune system to remember cancer cells, leading to an eventually immune response in longer-lasting remissions. That is, cancer may be less likely to return;

- compared to other treatments, it causes fewer side effects which vary according to each therapy and how it interacts with the body. This is because it only targets the immune system and not all cells;

- It can help other cancer treatments work better;

- Immunotherapy may work when other treatments don't; some cancers (like

skin cancer) don't respond well to radiation or chemotherapy but begin to disappear after immunotherapy.

Thus, immunotherapy represents a potential and effective treatment for patients with certain types of resistant cancer. Moreover a wide range of cancers are currently treated using immunotherapy; in particular melanoma has a good response to this kind of treatment. Therefore, in this thesis, we have chosen a stochastic model for immunotherapy of cancer. In particular the proposed model allows to simulate different treatment protocols, highlighting some counter-intuitive results: under some particular conditions therapy could work in favour of cancer resistance. Accordingly, the same type of cancer acts in very different ways with respect to the affected person and used therapy. Therefore, mathematical oncology could play a decisive role in the future of personalized medicine. In fact, patient-specific mathematical modelling, analysis and collection of clinical data could represent effective tools to develop patient-specific adaptive therapies and to face therapy resistance.

Chapter 1 introduce the biological setting necessary to have a better understanding of the evolutive processes and the interactions among the primary types of molecule and cells involved in cancer genesis and evolution. Indeed, understanding the mechanism of cancer evolution is essential to breed ground for an efficient mathematical model; even if any mathematical model is self-consistent and independent from the biological features of the process which describes. The first part of the chapter gives an idea of DNA structure, replication and mutations. This because errors in DNA coding represent the starting point of onset of cancer. In fact, mutations at DNA-level start in one cell and affect the whole offspring. Furthermore, special attention is placed on the hallmarkers of cancer, a small set of characterizing traits that are in common to almost all types of cancer. Finally, an intuitive motivation about the necessity to introduce a stochastic framework is given.

Chapter 2 begins by briefly presenting the main results of the theory of Markov chains, providing a rigorous theoretical foundations to the chosen model. Starting from definitions of discrete-time and continuous time Markov chains and Poissonian representation of the infinitesimal generator, we studied the mean-field limit of some interacting particle systems (e.g. Voter and Curie-Weiss models). Interacting Particle Systems (IPS) are Markov chains valued in a space S^Λ , where S is a finite set and Λ is countable set. Thus, assuming $N = |\Lambda| < \infty$, an arbitrary state, of an IPS $\big(\eta(t)\big)_{t\geq 0}$, has the following configuration

$$(\eta_1, \eta_2, \dots, \eta_N) = \big(\eta(i)\big)_{i\in\Lambda} \in S^\Lambda \qquad (\eta_i := \eta(i) \in S)$$

However, under the mean-field hypothesis, as $N \to \infty$, the dynamics of the empirical means $\left(\text{i.e. } \bar{\eta}_i^N := \frac{1}{N}\sum_{j=1}^N \delta_i(\eta_j)\right)$ can be approximated by a system of ODE (i.e. the mean-field limit). A modification of Voter model is used to study and simulate the process defined in Chapter 3, which describes the growth of melanoma treated with immunotherapy.

Firstly, Chapter 3 marks the fundamental role of mathematics in cancer research, giving examples of models with different mathematical backgrounds (e.g. population dynamics, game theory and stochastic models). The kernel of this chapter is the analysis of the stochastic model for immunotherapy of cancer proposed in the article [7]. This model describes melanoma evolution under treatment with respect to an ecological perspective. Namely, an organism affected by melanoma can be considered as a stable ecosystem, in which a new species springs up. The new interactions between pre-existing and rising species unbalance the equilibrium state of the ecosystem, leading to a dramatic change in the behaviour of all species. The immunotherapy, which consist of injection of T-cells, can be represent the appearing of a further new species, which competes for survival against the species represented by cancer. Thus, interactions among various types of cells are fundamental to provide a model for cancer growth under therapy which is close enough to reality. In fact, the presented stochastic model for immunotherapy of melanoma is an extension of Voter Model (VM), which takes into account interactions among species. In the context of melanoma under T-cell therapy, the main types of interacting particles are skin cancer cells, T-cells and cytokine (in the model others interactions are neglected). Hence, the local state space of the process describing the dynamics of these interacting particles is the following discrete and finite set

$$\mathcal{X} := \mathcal{G} \times \mathcal{P} \,\dot\cup\, \mathcal{Z} \,\dot\cup\, \mathcal{W}$$

where

- $(g, p) \in \mathcal{G} \times \mathcal{P}$ represents a cancer cell with genotype g and phenotype p;
- $z \in \mathcal{Z}$ denotes a T-cell of type z;
- $w \in \mathcal{W}$ expresses a cytokine of type w.

An arbitrary state of the population can be described by the random measure

$$\nu^K(t) := \frac{1}{K}\sum_{x\in\mathcal{X}} \nu_x(t)\delta_x,$$

where

- $K \in \mathbb{N}$ is a fixed parameter that allows to scale the population size and is usually called *carrying-capacity* of the environment;

- $\nu_x(t)$ is the number of individuals of type x at time t;

- δ_x denotes the Dirac measure at x.

Additionally, in order to analyse the model (qualitatively), we have showed two examples, based on the evidence of the experiments of Landsberg et al. [41], about melanoma in mice treated with T-cells. The first example the stochastic process depicting the evolution of a population has trait space

$$\mathcal{X} = \{x = (g, p); y = (g, p')\} \dot{\cup} \{z_x\} \dot{\cup} \{w\}.$$

The dynamics of the population is described by the transitions in Figure 1

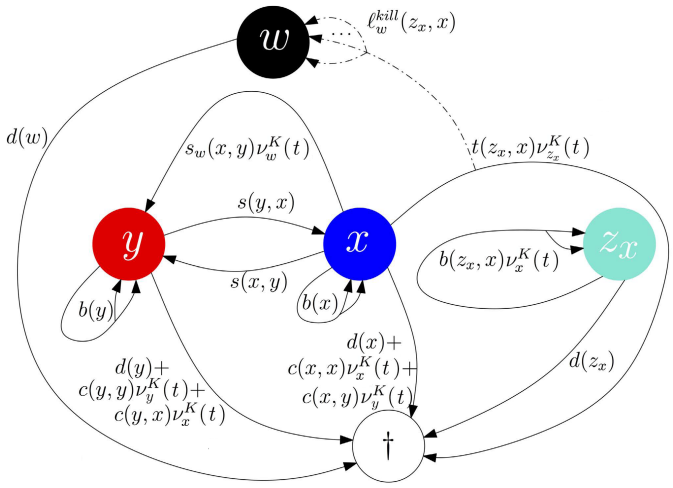

Figure 1: Representation of the transitions for a monomorphic population with two possible types of cancer cell $x = (g, p)$ and $y = (g, p')$, one type of lymphocyte z_x and one type of cytokine w [7].

The process described in this first example can be approximated by the follow-

ing ODE system

$$\begin{aligned}
\dot{\mathfrak{n}}_x &= \mathfrak{n}_x\big(b(x) - d(x) - s(x,y) - c(x,x)\mathfrak{n}_x - c(x,y)\mathfrak{n}_y - t(z_x,x)\mathfrak{n}_{z_x} - s_w(x,y)\mathfrak{n}_w\big) \\
&+ \mathfrak{n}_y s(y,x) \\
\dot{\mathfrak{n}}_y &= \mathfrak{n}_x\big(s(x,y) + s_w(x,y)\mathfrak{n}_w\big) + \mathfrak{n}_y\big(b(y) - d(y) - s(y,x) - c(y,x)\mathfrak{n}_x - c(y,y)\mathfrak{n}_y\big) \\
\dot{\mathfrak{n}}_{z_x} &= \mathfrak{n}_{z_x}(-d(z_x) + b(z_x,x)\mathfrak{n}_x) \\
\dot{\mathfrak{n}}_w &= \mathfrak{n}_x\big(l_w^{kill}(z_x,x)t(z_x,x) + l_w^{prod}(z_x,x)b(z_x,x)\big)\mathfrak{n}_{z_x} - \mathfrak{n}_w d(w)
\end{aligned}$$

where for each trait $\chi \in \mathcal{X}$, $\mathfrak{n}_\chi = \mathfrak{n}_\chi(t)$ represents the fraction of individuals of trait χ at time t. In particular the switching parameters (e.g. $s(x,y), s_w(y,x)$) points out that the melanoma can switch its phenotype in order to evade an immune response. This phenotypic plasticity of melanoma cells in an inflammatory microenvironment contributes to tumour relapse after initially successful T-cell immunotherapy. The second example describe a new protocol proposed by Landsberg et al. [41] in order to avoid tumour relapse. Finally, the analysis of rare mutations in large population underline that purely stochastic events may help to understand the resistance of tumours to therapeutic approaches [7]. Thus, this kind of stochastic models may underline some results that are not immediately evident by laboratory experiments. They may guide the creation of new tumour treatment protocols, supporting the design of experimental treatments and, thus, reducing the time of experimentation.

Chapter 1

The Biology of Cancer

To start this thesis, it is basic introducing the biological framework where a mathematical model for immunotherapy of cancer (see chapter 3) is inserted. Indeed, the following chapter argues about DNA structure, replication and mutations, that represent the starting point to understand onset of cancer (tumorigenesis). DNA molecule is the store of all genetic information for most cells and living organism. Mutations at DNA-level can arise either due to errors during DNA replication or from exposure to genotoxic agents and they could have various and serious effects on health of the involved organism, leading to genetic diseases (like cancer) in the worst cases [25, 50]. Moreover, cancer genetics and the main biological processes involved in cancer evolution are briefly discussed. Cancer (or malignant tumour) is a disease characterizing multicellular organisms. It results from somatic alterations, that disturb the normal cooperative behaviour of cells. Somatic alterations (i.e mutations at DNA-level) start in one cell and affect all the eventually future descendant cells. The latter become invasive cells through very quick replications and mutations, thereby producing abnormal cellular growth, which lead to an increasing of the size and invasiveness of the tumour. Furthermore, cancer cells induce many uncontrolled interactions within healthy tissues, therewith destroying the surrounding cellular environment. Summing up, cancer is a complex tissue which evolves in mutual influence with its environment. The evolution of cancer brings mostly irreparable damage to the cellular network, almost always carrying dramatic consequences for the affected organism [10, 26, 41, 57].

1.1 DNA Structure and Mutations

The ability to store and transmit genetic information from one generation to the next is a fundamental condition for life [50]. Deoxyribonucleic acid (DNA) is the repository for the hereditary information of most living organisms and cells.

Namely, DNA contains the master blueprint for the production of proteins and for its own replication. Proteins are the fundamental bricks of life; they are building materials for living cells, build tissue, appear in the structures inside cell and within cell membrane, carry oxygen, copy DNA for the next generation, etc. The DNA molecule is constituted by a double helix formed by two complementary chains twisted around each other. Each chain is a linear oriented sequence of four-type nucleotides. Generally, for the construction of nucleotides a set of five nitrogenous bases is used (see Figure 1.1): Adenine (A), Guanine (G), Cytosine (C), Thymine (T), and Uracil (U). Moreover, to make nucleotides, these bases attach to a pentose sugar (desoxyribose for DNA), along with a phosphate group. DNA contains no uracil because deoxyribose does not couple with it. This last nitrogen basis is present in the single-stranded structure of RNA (ribonucleic acid) instead of thymine (ribose, the pentose sugar of RNA, does not couple with thymine). There are different forms of RNA; specifically, messenger RNA (mRNA) carries the DNA information to the entire cell trough a process called transcription, responsible for the formation of proteins, which are essential for most cellular processes.

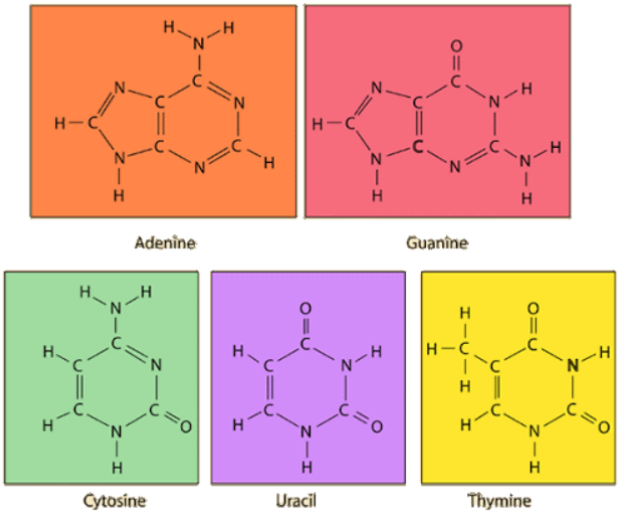

Figure 1.1: The set of five nitrogenous bases used to construct nucleotides, which in turn build up the nucleic acids, like DNA and RNA [49].

The nitrogen bases are crucially important because their sequencing, in nucleic acids, is the way information is stored. The bases can be attached in any order, giving an huge number arrangement possibilities which are achievable in the genetic code. However, DNA sequences can not be considered as mere combinations of four letters, randomly repeated multiple times; chemistry plays an remarkable role in determining the structure and function of nucleic acids. Indeed, in both DNA and

RNA, cytosine pairs with guanine by means of three hydrogen bonds ($C \equiv G$), while adenine pairs with thymine (in a DNA sequence) or with uracil (in RNA sequence), by means of two hydrogen bonds ($A = T$ and $A = U$ respectively). The double bond is weaker than the triple one; namely, it breaks more easily. The only existence of hydrogen bonds between complementary bases makes possible the separation of the two strands of DNA (see Figure 1.2), along from each of which it is possible to synthesize a complementary strand, giving origin to two copies of the original DNA.

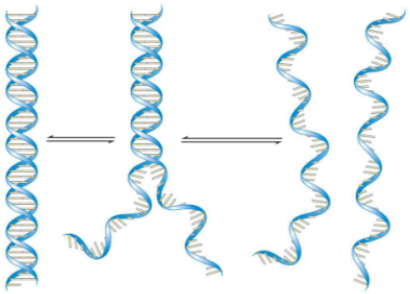

Figure 1.2: The hydrogen bonds between nitrogen bases, since they are not strong chemical bonds, let the separation of the two strands of DNA be possible [50].

The functional segments of DNA that code for the transfer of genetic information are called genes. The genes are transcribed into mRNA, which subsequently is translated into proteins [23, 49, 50].

However, the DNA in any cell can be altered by many different factors. Somatic mutations may occur during the process of DNA replication or reparation, or through exposure to exogenous mutagens (some chemical substances, radiation emanating, for example, from X, Gamma and UV rays) or endogenous ones (DNA oxidation, which generally is an intrinsic process in living molecules life), etc. Thus, some genes could be affected by random or driven alterations, that could lead to a change in DNA bases sequence and bases mismatching. These alterations obviously affect the genotype (the set of all genes in DNA) of the involved individual, but may not produce discernible changes in the set of functional and observable characteristics (phenotype). Two main classes of mutation can be distinguished:

- *Germ-line cell mutations*, which involve gametes (reproductive cells). These mutations can be passed to the organism's offspring.

- *Somatic mutations*, that can occur in any organism's cell different from germinal cells. Such mutations are passed to daughter cells during the cell replication, but they are not passed to offspring conceived via sexual reproduction.

One can note that it is sufficient only a single base change to create a devastating genetic disease or beneficial adaptation, or it may seemingly have no apparent effect [16, 26]. Also, at nucleotide level, it is possible to characterize various typologies of mutations, that are listed according to the nature of changes occurred in the nucleic sequence ([69], see figure 1.3):

- *point mutation* is a change in one base in the DNA sequence;
- *substitution* happens when one or more bases in the sequence is replaced by the same number of different bases;
- *inversion* realizes an end-to-end reversion of a DNA segment;
- *insertion* is the addition of a base in the sequence;
- *deletion* is when a base is deleted from the sequence.

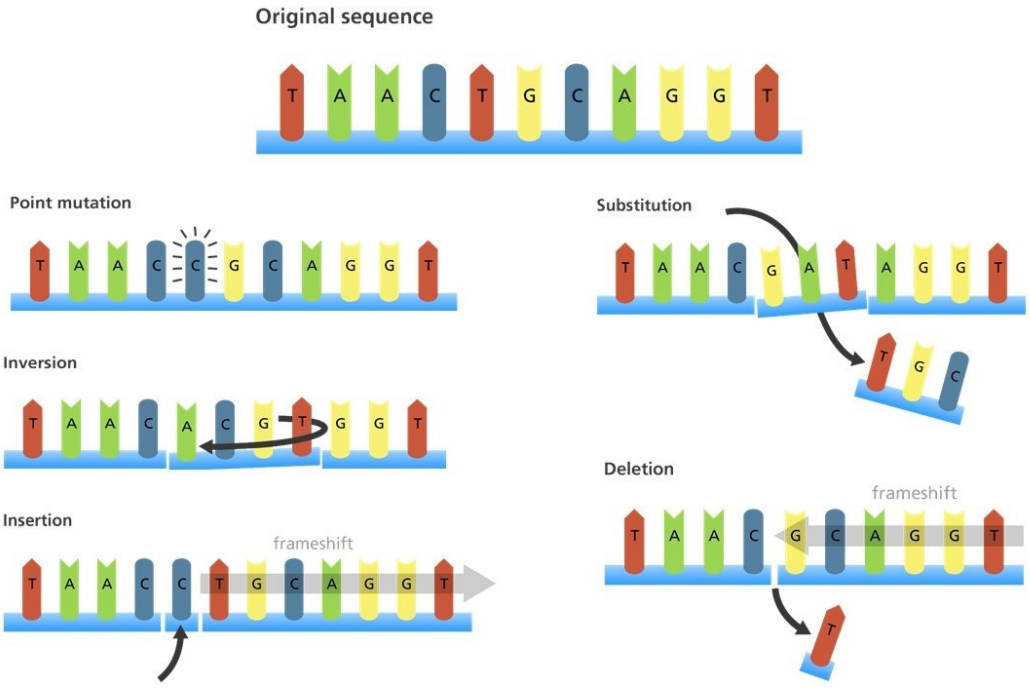

Figure 1.3: Representation of different types of Nucleic sequences alterations [69].

DNA is known to be subject to continuous damage, but the living organism has an automatic mechanism for dealing with such injuries, primarily by recognizing and repairing mutations. Most alterations undergo such a reparation mechanism, but others do not. Summing up, mutations or deficiencies in repair can lead to

catastrophic consequences, causing a wide range of human diseases (like cancer). In spite of this, mutations represents fundamental processes to life and evolution [25].

1.2 The Genetics of Cancer

One of the most common genetic disease affecting multicellular organisms is cancer [57]. It arises from the acquisition of somatic mutations by single cell. However, not all the somatic mutations, that can occur, work in the cancer development. Some of them are deleterious and lead the clone straight to extinction. Others, called passenger mutations, are neutral, without functional consequences; i.e. they are somatic mutations which do not confer any advantages to the mutated cell. Contrariwise, acquisition of driver mutations provides a fitness advantage which lead to improved replication or resistance to apoptosis (programmed cell death in cells). Nevertheless, both drivers and passengers mutations mark the history of the cancer cell. Then, each driver mutation allows the mutant cell to ride a wave of clonal expansion since, being somatic mutations, they are passed to the offspring of the mutated cell. Accordingly, the normal cooperative behaviour of cells is disturbed [10]. Finally, cancer main feature is the lack of growth control by the cell cycle regulatory mechanism, due to gene alterations. The consequent abnormal cell expansion can invade beyond normal tissue boundaries and metastasize to distant organs [39, 57]. In particular, the mutations feeding tumour primarily affect two types of genes, *proto-oncogenes* and *oncosuppressors*;

- proto-oncogenes normally help cell growth, but when they are affected by a mutation (e.g. point mutation, translocation, insertion, deletion) or an alteration, they became over-present or can overturn their role thus promoting an abnormal cell proliferation. Mutated proto-oncogenes are called *oncogenes*;

- oncosuppressors (or tumour suppressor genes) are genes that typically code for proteins that promote apoptosis, repress cell cycle regulation or are involved in DNA damage repair; i.e. they protect the cells. When mutations occur, there is a loss or a reduction in their functions in favour of cancer progression [39].

The Hallmarks of Cancer

One of the main challenge in recent cancer research is to simplifies the complexity of cancer study. In order to reach a logic, rigorous and common framework for all cancer types is essential making the tumour analysis understandable in terms of a small number of underlying principles. The latter aim still is an open problem in

cancer research field. Anyhow, research over the past decades has revealed (with support of biological evidences) a small number of molecular, biochemical, and cellular traits, which are the linkers for most, perhaps all, typologies of human cancer [28]. These traits are functional capabilities that healthy cells acquired during multistep process, called tumorigenesis, that leads them to become malignant. In the article [28], it is suggested that these alterations in cell physiology are essentially six (see Figure 1.4):

- *Self-sufficiency in growth signals*: in order to proliferate, healthy cells require signals that stimulate replication. Oncogenes simulate the effect of growth signals, inducing a self-sustained mechanisms, which lead to a chronic and unregulated cancer cell proliferation.

- *Insensitivity to anti-growth signals*: fundamental to cellular quiescence is the anti-growth signal, which block proliferation guaranteeing a correct tissue homoeostasis. Cancer cells can deactivate growth-inhibitory signal; many oncosuppressors (deputed to limit cancer cells reproduction), become deactivated.

- *Evading apoptosis*: every cell in living organisms are subjected to apoptosis (a self-programmed process that drives cells to death). Cancer cells get more resistant to apoptosis, slowing down this natural death process.

- *Limitless replicative potential*: for all mammalian cell types there is a self-autonomous program that induces senescence, limiting the number of successive cycles of cell growth and division, until cell division ceases. Oncogenes, deactiving the above surveillance neoplasms mechanism, let cancer cells do not have a limited number of successive cell growth-and-division cycles. This trait is also termed immortalization.

- *Sustained angiogenesis*: oxygen and nutrients are essential for the survival of cells. They are carried throughout the circulatory system of organisms, which has a dense network of capillary, reaching all cells. Normally, the process of growth of a blood vessel, termed angiogenesis, is transitory and carefully regulated. By contrast, when cancer occurs, there is an alterated signal, which causes new vessels to continually sprout, thus providing nutrients and oxygen that give cancer cells advantage in growth power.

- *Tissue invasion and metastasis*: Sustained angiogenesis has another determinant role in cancer evolution. Cancer cells use the rich network of blood vessels to flow towards new tissues, both adjacent and distant. In these way, the expansion of cancer is favoured through tissue invasion and birth of secondary neoplasms, responsible for 90% of cancer deaths.

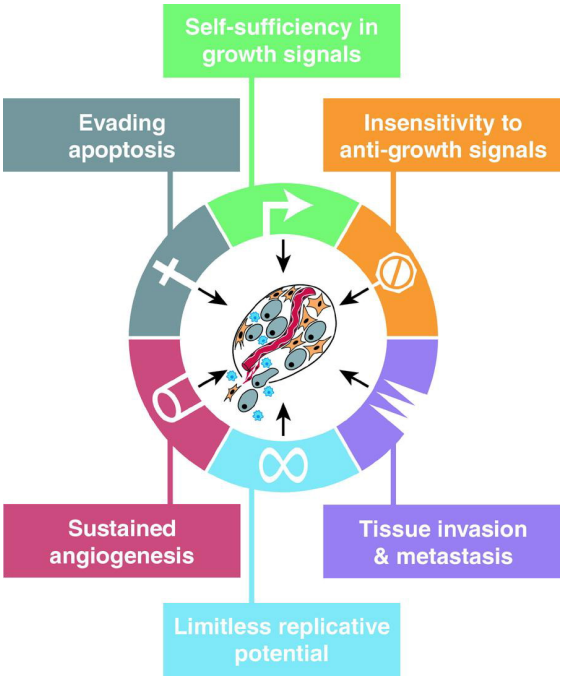

Figure 1.4: Set of novel capabilities acquired by cells during tumour development [28].

Two other traits were added to the list of main hallmarks afterwards, because to their functional importance in cancer development [29]:

- *Avoiding immune destruction*: in an organisms affected by cancer the immune system try to face and overcome formation and progression of incipient neoplasia, already formed tumour masses and metastasis. Unfortunately, cancer cells are able to hide from immune system, limiting the effectiveness of immunological killing.

- *Deregulating of cellular energy*: healthy cells need oxygen to make their metabolism work. A well functioning metabolism produces the energy necessary for the survival, growth and replication of cells. In tumour masses, cells can modify their metabolism, acquiring the ability to produce energy in an anaerobic context (without oxygen). This fact give an advantage to cancer domination because, in general, the absence of oxygen prevents healthy cells from producing energy.

Furthermore, multistep cancer progression can be viewed as a sequence of clonal expansions. Each clonal expansion is triggered by the random acquisition of an enabling mutant genotype. Two main enabling characteristics have been highlighted in [29], because they are essential conditions for the acquisition of the listed hallmarks:

- *Genome instability and mutation*: tumour cells express a mutator phenotype which gives cancer cells better efficiency in acquiring mutations; i.e. the mutation rate in the cancer cells is much greater than that in normal cells. [43]. This feature deceive oncosuppressores, increasing the rates of mutation. The insufficiency of the surveillance systems lead to an acceleration of the accumulation of mutations and force genetically damaged cells into either senescence or apoptosis, bringing serious tissue injuries and advantaging of cancer expansion.

- *Tumour-promoting inflammation*: inflammation associated to cancer is the response of immune system, that try to face cancer. Paradoxically, inflammation promote tumour progression, indeed contribute to multiple hallmark capabilities by supplying bioactive molecules to the tumour microenvironment.

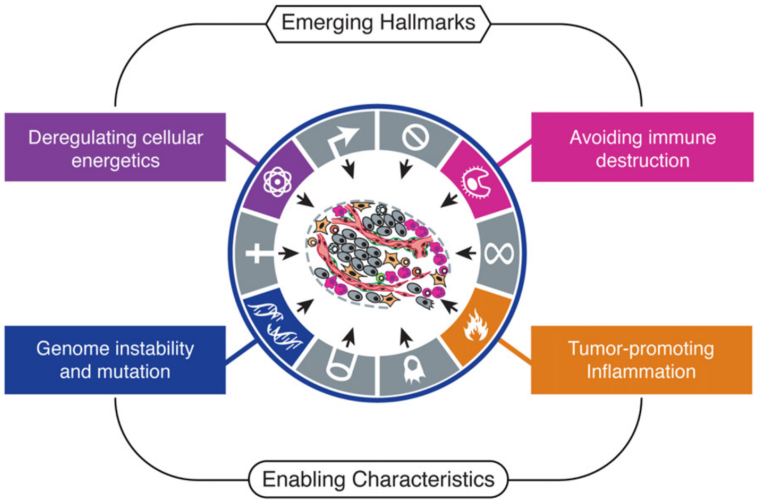

Figure 1.5: Emerging hallmarks and enabling characteristics [29].

One important turning point in cancer research is that the dialogue between tumour cells and their environment is critical to understanding the complexity of cancer evolution. Accordingly, it is crucial to recognize cancer as complex tissue composed of multiple distinct cell types that interact with one another [9, 29]. More precisely, "the tumour mass consists not only of a heterogeneous population of cancer cells but also a variety of resident and infiltrating host cells, secreted factors and extracellular matrix proteins, collectively known as the tumour microenvironment" [48]. That is, cancer cannot be considered as an mutated isolated mass in a organism. Consequently, cancer progression is strictly correlated

to interaction with its microenvironment, that ultimately determine its evolution [29, 48].

1.3 The Role of Randomness

The role of randomness in tumorogenesis (i.e. cancer initialization) is still a reason for disagreement in literatures [39]. Recently, a paper [62], supporting the existence of a strong correlation between tissue-specific cancer risk and the lifetime number of tissue-specific stem-cell divisions, is published;

> Some tissue types give rise to human cancers millions of times more often than other tissue types. Although this has been recognized for more than a century, it has never been explained. Here, we show that the lifetime risk of cancers of many different types is strongly correlated (0.81) with the total number of divisions of the normal self-renewing cells maintaining that tissue's homeostasis. These results suggest that only a third of the variation in cancer risk among tissues is attributable to environmental factors or inherited predispositions. The majority is due to "bad luck", that is, random mutations arising during DNA replication in normal, noncancerous stem cells. This is important not only for understanding the disease but also for designing strategies to limit the mortality it causes. [62]

This paper shook a lot since it contradict the well-established role of environmental factors and lifestyle for cancer incidence [39]. As a response to article [62] a subsequent paper [66] was published in the same year. Its authors support that the correlation exists, it cannot properly distinguish between intrinsic and extrinsic factors [39];

> Recent research has highlighted a strong correlation between tissue-specific cancer risk and the lifetime number of tissue-specific stem-cell divisions. Whether such correlation implies a high unavoidable intrinsic cancer risk has become a key public health debate with the dissemination of the 'bad luck' hypothesis. Here we provide evidence that intrinsic risk factors contribute only modestly (less than about $10-30\%$ of lifetime risk) to cancer development. First, we demonstrate that the correlation between stem-cell division and cancer risk does not distinguish between the effects of intrinsic and extrinsic factors. We then show that intrinsic risk is better estimated by the lower bound risk controlling for total stem-cell divisions. Finally, we show that the rates of endogenous mutation accumulation by intrinsic processes are

> not sufficient to account for the observed cancer risks. Collectively, we conclude that cancer risk is heavily influenced by extrinsic factors. These results are important for strategizing cancer prevention, research and public health. [66]

Nevertheless, in the construction of mathematical cancer-evolution models, tuomorigenesis problem can be overcame, treating factors triggering the cancer initialization as random influences external to the model. This argument can be justified by following topic. Randomness play a fundamental role in modelling biological dynamical system. Indeed, there may be parts of dynamic, which are not explicitly included in modelling for their arduous management (e.g. sub-dynamics that either are not predictable or understandable). However, to be realistic, biological dynamical system models should include the influences of these excluded parts of the dynamics; a way of doing it is to consider the effects due to influences external to the model as random results. Thus, biological dynamical systems evolve under stochastic lows, where stochasticity suggests the presence of random variables and behaviours in the model. Generally, stochastic effects may slightly or even dramatically change the dynamic behaviour of the system [20]. Consequently, omitting stochasticity makes the proposed model inefficient. Therefore, basics of stochastic processes are introduced, in Chapter 2. Indeed, they represent the mathematical background needed to rigorously construct a model for cancer progression (e.g. model presented in Chapter 3).

Chapter 2

Basics of Markov Chains

Aim of the present chapter is to review few fundamental results of Markov chain theory. Discrete-time and continuous time Markov chains are defined and their Possonian construction is reviewed. The latter will be useful later on to study and simulate the process defined in Chapter 3. This model describes a melanoma under T-cell treatment, it comes from the class of models which depict the immunotherapy of malignant tumours and it represents the core of this thesis. Such a process belong to a class of processes called Interacting Particle System (IPS). Loosely speaking IPS are system comprised of a large number of units (particles) whose dynamics are Markovian and shaped by the interaction among the particles of the system. In this chapter we consider mean-field IPS. In mean-field IPS each particle interacts with the others trough the empirical mean. Moreover, the empirical mean is a sufficient statistic for the system. It is showed (see Section 2.4, Theorem 2.18) that in the thermodynamic limit (i.e. when the number of particles grows to infinity) the empirical mean behaves deterministically and its time evolution can be described by an Ordinary Differential Equation (ODE). The chapter ends with two standard example: the Curie-Weiss and the Voter models.

2.1 Discrete-Time Markov Chains

Let $S \subseteq \mathbb{R}^\nu$, with $\nu \in \mathbb{N} := \{1, 2, 3, \dots\}$, be a finite set ($|S| < \infty$) and let $(\Omega, \mathcal{F}, P)$ be a probability space. A random variable X from $(\Omega, \mathcal{F}, P)$ to S is called *S-valued* random variable and S is called *state space*.

Definition 2.1. *A stochastic process is a sequence of S-valued random variables* $(X_n)_{n\in\mathbb{N}}$.

In particular, S is called state space of the process.

Definition 2.2 (Markov Chain)**.** *A stochastic process $(X_n)_{n\in\mathbb{N}}$ is a Markov Chain (MC) if*

$$P\big(X_{n+1}=j|X_n=i,\dots,X_1=i_1,X_0=i_0\big)=P\big(X_{n+1}=j|X_n=i\big),$$

$\forall n\geq 1$ *and* $\forall j,i,i_0,i_1,\dots,i_{n-1}\in S$.

If $\forall i,j\in S,\quad P\big(X_{n+1}=j|X_n=i\big)$ does not depend on n, we say the chain is *time-homogeneous*. In this case the matrix

$$\mathcal{P}:=(p_{ij})_{i,j\in S}\quad with\quad p_{ij}:=P\big(X_{n+1}=j|X_n=i\big)$$

is called the *transition matrix* (p_{ij} is called *transition probability*) and it has the fallowing properties:

- $p_{ij}\geq 0,\quad \forall i,j\in S$;
- $\sum_{j\in S}p_{ij}=1$.

An important fact is that the law of the process is completely determined by the initial distribution $\pi_0(i):=P\big(X_0=i\big),\ i\in S$, and the transition matrix $\mathcal{P}$:

$$P\big(X_n=i_n,X_{n-1}=i_{n-1},\dots,X_1=i_1,X_0=i_0\big)=\pi_0(i_0)p_{i_0i_1}\dots p_{x_{n-1}i_n}$$

2.2 Continuous-Time Markov Chains

Let S be a finite metrizable space, let $\mathcal{S}$ be the Borel$-\sigma$ field on S and $t\in[0,+\infty)$.

Definition 2.3. *An S-valued stochastic process is a collection $(X(t))_{t\geq 0}$ of measurable map from a common probabilistic space $(\Omega,\mathcal{F},P)$ to $(S,\mathcal{S})$.*

Thus, the stochastic process $(X(t))_{t\geq 0}$ gives an S-valued random element $X(t)$ on $(\Omega,\mathcal{F},P),\ \forall t\geq 0$, and fixing $\omega\in\Omega$ the maps $t\to X(t)(\omega)$ on $\mathbb{R}$ are called the *trajectories*, or *sample paths*. The sample paths are functions from $\mathbb{R}$ to S, i.e. elements of $S^{\mathbb{R}}$. Hence, we can view $(X(t))_{t\geq 0}$ as a random element of the function space $S^{\mathbb{R}}$.

Definition 2.4. *A collection of $(\mathcal{F}_t)_{t\geq 0}$ of sub-σ-algebras of $\mathbb{F}$ is called a filtration if $\mathcal{F}_s\subseteq\mathcal{F}_t,\quad\forall\ s\leq t$.*

Taking a stochastic process $(X(t))_{t\geq 0}$, it is always possible define a filtration $(\mathcal{F}_t^X)_{t\geq 0}$ defined by $\mathcal{F}_s^X = \sigma(X_s : s \leq t)$. This filtration is called the *filtration generated* by $(X(t))_{t\geq 0}$ (or the *natural filtration*). Intuitively, the natural filtration of a stochastic process keeps track of the "history" of the process. Moreover a stochastic process $(X(t))_{t\geq 0}$ is always adapted to its natural filtration, i.e. $\forall t \geq 0$, the random variable $X(t)$ is $\mathcal{F}_t$-measurable.

Definition 2.5 (Markov Chain)**.** *A stochastic process $(X(t))_{t\geq 0}$ is a Markov chain if $\forall j \in S$ and $0 \leq s \leq t$*

$$P\big(X(t) = j|\mathcal{F}_s^X\big) = P\big(X(t) = j|X(s)\big).$$

Let $(X_t)_{t\geq 0}$ be a *time-homogeneous* Markov chain (HMC),
i.e.

$$P\big(X(t+s) = j|X(s) = i\big) = P\big(X(t) = j|X(0) = i\big), \quad \forall s, t \geq 0,$$

and let $(\mathcal{P}_t)_{t\geq 0}$ be a family of operators acting on the functions $f : S \to \mathbb{R}$ as follows

$$\mathcal{P}_t f(i) := E\big(f(X(t))|X(0) = i\big) = \sum_{j\in S} f(j)P\big(X(t) = j|X(0) = i)\big).$$

The following standard results are valid (e.g. a proof of them are in [11, 42]):

(i) $\mathcal{P}_t : \mathbb{R}^{|S|} \to \mathbb{R}^{|S|}$ is a linear operator represented by the matrix with entries

$$p_{ij}(t) := P\big(X(t) = j|X(0) = i\big) \qquad i, j \in S \tag{2.1}$$

which is a stochastic matrix, indeed (by definition (2.1)) it results that

$$p_{ij}(t) \geq 0 \quad \text{and} \quad \sum_{j\in S} p_{ij}(t) = 1 \quad (t \geq 0);$$

(ii) $(\mathcal{P}_t)_{t\geq 0}$ is a semigroup (by Markov's property), that is

$$\mathcal{P}_0 = \mathbb{I} \tag{2.2}$$

$$\mathcal{P}_{t+h} = \mathcal{P}_t\mathcal{P}_h, \tag{2.3}$$

where $\mathbb{I}$ is the identity matrix;
i.e.

$$p_{ij}(0) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

and

$$p_{ij}(t+h) = \sum_{k\in S} p_{ik}(t)p_{kj}(h)$$

for $i, j \in S$;

(iii) $(\mathcal{P}_t)_{t\geq 0}$ is a *contraction* semigroup,
i.e.

$$||f||_\infty := \sup_{i\in S}|f(i)| \leq 1 \quad \Longrightarrow \quad ||\mathcal{P}_t f||_\infty := \sup_{i\in S}|\mathcal{P}_t f(i)| \leq 1;$$

(iv) $(\mathcal{P}_t)_{t\geq 0}$ is a *strongly continuous* semigroup,
i.e.

$$\lim_{t\to 0}\mathcal{P}_t f = f \overset{\text{DEF}}{\Longleftrightarrow} ||\mathcal{P}_t f - f||_\infty \xrightarrow{t\to 0} 0;$$

Note that the last point implies that $\mathcal{P}_t f$ converges point-wise to f:

$$\lim_{t\to 0}\mathcal{P}_t f(i) = f(i) \quad \forall i \in S$$

that is

$$\lim_{t\to 0}\mathcal{P}_t = \mathbb{I}.$$

$(\mathcal{P}_t)_{t\geq 0}$ is called *continuous transition semigroup* of the HMC. One can note that, since the states space S is finite, $\mathcal{P}_t = (p_{ij}(t))_{i,j\in S}$ and $f = (f(i))_{i\in S}$ can be considered as a matrix and a vector respectively;

$$\mathcal{P}_t f(i) = \sum_{j\in S} f(j)p_{ij}(t) \quad \forall i \in S,\ t \geq 0.$$

Besides, for $t \geq 0$, let

$$\mu_i(t) := P\big(X(t) = i\big), \quad \forall i \in S \qquad \textit{and} \qquad \mu(t) = (\mu_1(t), \mu_2(t), \ldots, \mu_{|S|}(t))$$

then

$$\begin{aligned}\mu_j(t) = P\big(X(t) = j\big) &= \sum_{i\in S} P\big(X(0) = i\big)p_{ij}(t) \\ &= \sum_{i\in S}\mu_i(0)p_{ij}(t) \\ &= (\mu(0)P_t)_j\end{aligned}$$

that is,

$$\mu(t) = \mu(0)\mathcal{P}_t. \tag{2.4}$$

This means that the process is completely determined by the initial distribution μ and the transition semigroup matrix $\mathcal{P}_t$. Furthermore, the following important property holds for $0 =: t_0 < t_1 < t_2 < \cdots < t_n$

$$\begin{aligned}&P\big(X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i_n\big) \\ &= \mu_{i_0}(0)(i_0)p_{i_0 i_1}(t_1 - t_0)\ldots p_{i_{n-1}i_n}(t_n - t_{n-1}),\end{aligned} \tag{2.5}$$

where $\forall\ k \in \{0, 1, 2, \dots n\},\ i_k \in S$ and $\mu_{i_0}(0) = P\big(X(0) = i_0\big)$. For any HMC it is possible to define another operator: $\mathcal{L} : \mathbb{R}^{|S|} \to \mathbb{R}^{|S|}$;

$$\mathcal{L}f := \lim_{t\to 0} \frac{\mathcal{P}_t f - f}{t}\ , \quad \forall (f : S \to \mathbb{R}) \in \mathcal{D}(\mathcal{L}) \tag{2.6}$$

$\mathcal{L}$ is called the *infinitesimal generator* of the HMC and $\mathcal{D}(\mathcal{L})$ is its domain. More precisely

$$\mathcal{D}(\mathcal{L}) := \left\{ f : S \to \mathbb{R} \;\middle|\; \exists \lim_{t\to 0} \frac{\mathcal{P}_t f - f}{t} \right\},$$

therefore

$$f \in \mathcal{D}(\mathcal{L}) \iff \left|\left| \frac{\mathcal{P}_t f - f}{t} - \mathcal{L}f \right|\right|_\infty \xrightarrow{t\to 0} 0.$$

This means that if $f \in \mathcal{D}(\mathcal{L})$ than one has also the pointwise convergence

$$\mathcal{L}f(i) = \lim_{t\to 0} \frac{\mathcal{P}_t f(i) - f(i)}{t} \qquad \forall i \in S. \tag{2.7}$$

The infinitesimal generator $\mathcal{L}$ is informally the derivative at time $t = 0$ of the continuous semigroup $\mathcal{P}_t$, this may be heuristically understood from the following computation;

$$\mathcal{L}f = \lim_{t\to 0} \frac{\mathcal{P}_t f - f}{t} = \lim_{t\to 0} \frac{\mathcal{P}_t - \mathbb{I}}{t} f$$

since, it is valid $\forall (f : S \to \mathbb{R}) \in \mathcal{D}(\mathcal{L})$, from (2.2) it follows that

$$\mathcal{L} = \lim_{t\to 0} \frac{\mathcal{P}_t - \mathbb{I}}{t} = \dot{\mathcal{P}}_0, \quad \text{in} \quad \mathcal{D}(\mathcal{L}).$$

Let $(q_{ij})_{i,j\in S}$ be the matrix associated to the linear operator $\mathcal{L}$, thus

$$q_{ij} = \lim_{t\to 0} \frac{p_{ij}(t) - \delta_{ij}}{t}, \qquad \forall i, j \in S \tag{2.8}$$

by definition. The entries q_{ij} are named *local characteristics* of the continuous HMC and the following properties hold (see Theorem 2.1 in [11])

$$\exists\ q_i := -q_{ii} = \lim_{t\to 0} \frac{p_{ij}(t) - \delta_{ij}}{t} \in [0, \infty] \qquad \forall i \in S \tag{2.9}$$

and

$$\exists\ q_{ij} = \lim_{t\to 0} \frac{p_{ij}(t)}{t} \in [0, \infty), \qquad \forall i, j \in S : i \neq j \tag{2.10}$$

Moreover, since $|S| \leq \infty$,

$$\sum_{j\in S} q_{ij} = \sum_{j\in S} \lim_{t\to 0} \frac{p_{ij}(t) - \delta_{ij}}{t} = \lim_{t\to 0} \sum_{j\in S} \frac{p_{ij}(t) - \delta_{ij}}{t} = 0$$

this and property (2.10) implies that

$$q_i = \sum_{j \neq i} q_{ij} < \infty. \tag{2.11}$$

This means, the HMC related to the local characteristics q_{ij} is *stable* and *conservative*. Thus, the local characteristics determine some particular features of the related semigroup $(\mathcal{P}_t)_{t\geq 0}$; more precisely:

Definition 2.6 (Stability and Conservation)**.** *If* $\forall\ i \in S$,

- $q_i < \infty, \quad (\mathcal{P}_t)_{t\geq 0}$ *is called* ***stable****;*
- $q_i = \sum_{i\in S} q_{ij}, \quad (\mathcal{P}_t)_{t\geq 0}$ *is called* ***conservative****.*

Equivalently, a HMC is said stable and/or conservative, if its transition group $(\mathcal{P}_t)_{t\geq 0}$ is stable and/or conservative. When a HMC is stable and conservative (i.e. condition (2.11) holds), it is possible to infinitesimally describe the transition probabilities (in accordance with definition (2.8)):

$$\begin{cases} p_{ii}(t) = 1 - q_i t + 1 + o(t) \\ p_{ij}(t) = q_{ij} t + o(t) \end{cases}$$

Furthermore, from the semigroup properties (2.2) and (2.3), it follows that

$$\mathcal{P}_{t+h} - \mathcal{P}_t \ = \ \mathcal{P}_t \mathcal{P}_h - \mathcal{P}_t \ = \ \mathcal{P}_t(\mathcal{P}_h - \mathbb{I}),$$

consequently

$$\frac{\mathcal{P}_{t+h} - \mathcal{P}_t}{h} = \mathcal{P}_t \ \frac{\mathcal{P}_h - \mathbb{I}}{t}$$
$$\frac{\mathcal{P}_{t+h} - \mathcal{P}_t}{h} = \frac{\mathcal{P}_h - \mathbb{I}}{h} \ \mathcal{P}_t.$$

Now, let $\mathcal{P}(t) := \mathcal{P}_t$ and, since the states space S is finite ($|S| < +\infty$), the passage to the limit, for $h \to 0$, is allowed; the *Kolmogorov's equations* emerge:

$$\dot{\mathcal{P}}(t) = \mathcal{P}(t)\ \mathcal{L} \qquad (forward)$$
$$\dot{\mathcal{P}}(t) = \mathcal{L}\ \mathcal{P}(t) \qquad (backward);$$

or equivalently, $\forall\ i, j \in S$,

$$\frac{d}{dt} p_{ij}(t) = -p_{ij}(t) q_j + \sum_{\substack{k\in S \\ k\neq j}} p_{ik}(t) q_{kj} \qquad (forward)$$

$$\frac{d}{dt} p_{ij}(t) = -q_i p_{ij}(t) + \sum_{\substack{k\in S \\ k\neq i}} q_{ik} p_{kj}(t) \qquad (backward),$$

writing them explicitly. Again, since $|S| < +\infty$, considering the initial condition $\mathcal{P}_0 = \mathbb{I}$,

$$\mathcal{P}_t = e^{t\mathcal{L}} = \sum_{n=0}^{+\infty} \frac{(t\mathcal{L})^n}{n!}, \qquad t \geq 0 \tag{2.12}$$

results the unique solution for both the forward and backward equations (see Theorem 2.1 of Appendix in [11]). In this setting, Kolmogorov's equations highlight how the infinitesimal generator is univocally linked to semigroup and initial distribution.

Regular Jump Homogeneous Markov Chain

Among HMC there is a very general class of stable and conservative process (generally defined on a countable state space S, which is not necessary finite), the regular jump HMC (see Theorem 3.4. in [11]).

Definition 2.7 (Regular Jump Process)**.** *An S-valued stochastic process $(X(t))_{t\geq 0}$ is called jump process if for almost all $\omega \in \Omega$ and all $t \geq 0$, $\exists\ \epsilon(t,\omega) > 0$ such that*

$$X(t+h)(\omega) = X(t)(\omega) \quad \forall h \in \big[t, t+\epsilon(t,\omega)\big). \tag{2.13}$$

It is called Regular Jump Process (RJP) if in addition, for almost all $\omega \in \Omega$, the set $A(\omega)$ of discontinuities of the function $t \to X(\omega)(t)$ is σ-discrete, that is,

$$\big|A(\omega) \cup [0,c]\big| < \infty \tag{2.14}$$

It is interesting point out that for a general jump process $(X(t))_{t\geq 0}$, there exist a sequence of times $(\tau_k)_{k\geq 0}$ (*transition times sequence*) and a sequence $(X_k)_{k\geq 0}$ (*embedded process*) such that

$$0 = \tau_0 < \tau_1 < \cdots < \tau_k < \dots$$

and

$$X(t) = X_k \qquad \forall\ t \in [\tau_k, \tau_{k+1}), \quad k \geq 0.$$

Indeed, by definition the process is simple; namely, there are no two infinitesimally close jumps (condition 2.13). Then

$$m, n \in S : m \neq n \Rightarrow \tau_m \neq \tau_n.$$

Moreover, if the jump process $(X(t))_{t>0}$ is also regular, then the process is not explosive, that means there is not an infinite number of jumps in any finite time interval (i.e. condition (2.14)). Hence,

$$\tau_\infty := \lim_{k\to\infty} \tau_k = \infty \quad a.s.$$

where τ_∞ is named *explosion time*.

Definition 2.8. *A **regular jump homogeneous Markov chain** is by definition a continuous-time HMC that is also a RJP.*

Consequently, the study of RJP can be focused on its local description trough the infinitesimal generator (i.e. transitions are analysed in infinitesimal times), since the statistical properties of any continuous-time MC are determined by its semigroup after fixing an initial distribution. In simple words, two regular jumps HMCs with the same infinitesimal generator have the same transition group [11]. Often, it is useful representing the infinitesimal generator as a linear operator from $\mathbb{R}^{|S|}$ to $\mathbb{R}$:

$$\mathcal{L}f(i) = \sum_{j \in S} q_{ij} f(j) = \sum_{j \neq i} q_{ij}[f(j) - f(i)], \tag{2.15}$$

where the last equality follows from (2.11). It is important to mark that "there exist bona fide continuous-time HMCs that are not regular jump HMCs, however, for which all the states are unstable !" [11].

2.3 Poisson Construction of Markov Chains

To lay the ground for the simulations of the model analysed in this thesis, it is advantageous to introduce the Poisson construction of Markov process. Let S be a metrizable space which is also $\sigma-$compact (i.e. exists a countable collection of compact sets $S_i \in S$ s.t. $\cup_i S_i = S$), $\mathcal{S}$ the Borel-σ-field on S and $(\Omega, \mathcal{F}, P)$ an underlying probability space.

Definition 2.9. *A random measure on S is a function $\xi : \Omega \times \mathcal{S} \to [0, \infty]$ s.t.*

- *$\forall$ fixed $\omega \in \Omega$, the function $\xi(\omega, \cdot)$ is a locally finite function on $(S, \mathcal{S})$, i.e. $\xi(\omega, K) < \infty$, $\forall$ compact $K \in S$;*
- *$\forall$ fixed $A \in \mathcal{S}$, the function $\xi(\cdot, A)$ is measurable.*

As pointed out in Chapter 2 of [59], it is helpful to think of ξ as a random variable evaluated in the space of locally finite measures on $(S; \mathcal{S})$, endowed with the σ-field generated by the maps $\mu \mapsto \mu(A)$, with $A \in \mathcal{S}$ and μ locally finite measure. Consequently, for all measurable $f : S \to [0, \infty]$

- the integral $\int f d\xi$ defines a $[0, 1]$-valued random variable,
- there exists a unique measure,

$$\begin{aligned} E\xi : \mathcal{S} &\to [0, +\infty] \\ A &\mapsto E[\xi(A)] \end{aligned}$$

called the *intensity* of ξ, such that

$$\int f \, dE\xi \;=\; E\Big[\int f d\xi\Big].$$

Where

$$E[X] := \int_\Omega X dP$$

is the *mean* (or *expectation* or *expected value*) of a random variable X.

Proposition 2.10. *Let μ be a locally finite measure on $(S, \mathcal{S})$. Then there exists an unique in distribution random measure ξ s.t. for any family of disjoint set $\{A_1, \dots, A_n\} \subset \hat{\mathcal{S}} := \{A \in \mathcal{S} : \bar{A}$ is compact$\}$, the random variables $\xi(A_1), \dots, \xi(A_n)$ are independent and $\xi(A_i)$ is Poisson distributed with mean $\mu(A_i), \quad \forall\, i = 1, \dots, n$.*

Such random measure ξ is called *Poisson point measure* with *intensity* μ. Indeed, it results $E\xi = \mu$. Furthermore,

$$\xi(A) \in \mathbb{N} \quad \forall\, A \in \hat{\mathcal{S}},$$

such measures are called **counting measures**. In general, each local finite counting measure ν on S has the following form [59]:

$$\nu = \sum_{x \in Z} n_x \delta_x,$$

where

- $Z := \text{supp}(\nu) = \{x \in S : \nu(x) \neq 0\} \subset S$ is a locally finite set (i.e. $\forall\, x \in Z,\ \exists\, U \subset Z$ open with $x \in U : |U \cap S| < \infty$);
- δ_x is the delta-measure at $x \in Z$;
- $n_x \in \mathbb{N},\ \forall x \in Z$; moreover, if $\forall x \in Z,\ n_x = 1,\ \nu$ is said *simple*.

Definition 2.11. *An integer-valued random measure ξ on S is called* ***point process*** *on S.*

Consequently, a Poisson point measure ξ is a point process.

Lemma 2.12. *Let ξ be a Poisson point measure with locally finite intensity μ. Then*

$$\xi \text{ is a.s. simple} \iff \mu \text{ is atomless} \Big(\overset{\text{DEF}}{\iff} \mu(\{x\}) = 0,\ \forall x \in S \Big).$$

Therefore, Poisson point measure ξ with atomless intensity μ is characterized by its support; *supp*(ξ) is called *Poisson point set* with intensity μ. Another important feature of Poisson point measures is that *the sum* $\xi_1+\xi_2$ *of two independent Poisson point measures* ξ_1 *and* ξ_2 *with intensities* μ_1 *and* μ_2 *respectively, is a Poisson point measure with intensity* $\mu_1 + \mu_2$ (see Lemma 2.3 in [59]). Finally, the key result to Poisson construction of Markov processes is the following lemma;

Lemma 2.13 (Poisson points on the halfline)**.** *Let* $(\tau_k)_{k>0}$ *be real random variables such that* $\tau_0 = 0$ *and* $\sigma_k := \tau_k - \tau_{k+1} > 0, \quad k \geq 1$ *. Then* $\tau := \{\tau_k : k \geq 1\}$ *is a Poisson point set on* $[0; \infty)$ *with intensity* $c\ell$ *(where* ℓ *is the Lebesgue measure) iff the random variables* $(\sigma_k)_{k\geq 1}$ *are i.i.d. exponentially distributed with mean* c^{-1}*.*

A simple Poisson point measure ξ on $[0; \infty)$ with intensity $c\ell$, where $c > 0$ and ℓ denote the Lebesgue measure, is called *Homogeneous Poisson measure* with rate c. Then, Lemma 2.13 is equivalent to the following Proposition (see Chapter 10 in [31])

Proposition 2.14. *Let* ξ *be a simple point process on* $\mathbb{R}_0^+ := [0, +\infty)$ *with atoms at* $\tau_1 < \tau_2 < \cdots < \tau_k < \ldots$ *and put* $\tau_0 = 0$*. Then* ξ *is homogeneous Poisson with rate* $c > 0$ *iff* $\tau_k - \tau_{k-1}$ *are i.i.d exponentially distributed with mean* c^{-1}*.*

Poisson Construction of Markov Processes

The first step for the Poisson construction of MC is defining a stochastic flow in terms of a Poisson point set. Then, let the following setting be considered:

- let S be a finite set;
- let $\mathcal{G} := \{m : S \to S\}$ be the set of self maps of S;
- let $(r_m)_{m\in\mathcal{G}} \in [0, +\infty)$ be a sequence of non negative constant;
- considering the Borel-σ-field $\mathcal{B}(\mathbb{R})$ and the Lebesgue measure ℓ, the following function
$$\rho(\{m\} \times A) := r_m \ell(A), \quad \forall m \in \mathcal{G},\ A \in \mathcal{B}(\mathbb{R}) \tag{2.16}$$
provides a measure to the space $\mathcal{G} \times \mathbb{R}$.

In particular ρ is locally finite, by construction (see definition (2.16)); indeed the Lebesgue measure ℓ is locally finite on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Therefore, by Proposition 2.10, it is possible to define, a Poisson point set σ with intensity ρ. Then

$$\nu := \sum_{(m,t)\in\sigma} \delta_t$$

is a Poisson point process on $\mathbb{R}$ with intensity $r\ell$, where

$$r := \sum_{m \in \mathcal{G}} < \infty \tag{2.17}$$

(r is a finite sum since $\{S \text{ finite} \Rightarrow \ \mathcal{G} \text{ finite}\}$). Condition 2.17 implies that $r\ell$ is an atomless local finite measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (again, ℓ is atomless local finite measure), hence ν is simple and local finite. Therefore,

- $\forall\ t \in \mathbb{R}$, exists at most one $m \in \mathcal{G} : (m, t) \in \sigma$, by simplicity of ν;
- $\sigma = \text{supp}(\nu)$ is a local finite set by definition.

This implies that

$$\sigma_{s,u} = \{(m, t) \in \sigma : t \in (s, u]\} \qquad (s \leq u)$$

is a finite set; then it is possible to give an order to its element:

$$\sigma_{s,u} = \{(m_1, t_1), (m_2, t_2), \dots, (m_n, t_n)\} \quad \text{for} \ \ t_1 < t_2 < \cdots < t_n, \ \ n \in \mathbb{N}. \tag{2.18}$$

Starting from (2.18), it is possible to define a stochastic flow $(\mathbf{X}_{s,t})_{s \leq t}$ in terms of the poisson point set ω: let $(\mathbf{X}_{s,u})_{s \leq u}$ be a family of random maps such that

$$\mathbf{X}_{s,u} := m_n \circ m_{n-1} \circ \cdots \circ m_1,$$

where $m_k \in \mathcal{G} : (m_k, t_k) \in \omega_{s,u} \quad \forall k = 1, \dots, n$. Then, by definition

$$id_S = \mathbf{X}_{s,u} \text{ if } \sigma_{s,u} = \emptyset,$$

i.e. the identity map id_S corresponds to no maps composition and

$$\mathbf{X}_{s,t} \to \mathbf{X}_{s,s} = 1 \ \ \text{as} \ \ t \to s; \tag{2.19}$$

$$\mathbf{X}_{t,u} \circ \mathbf{X}_{s,t} = \mathbf{X}_{s,u}, \qquad s \leq t \leq u. \tag{2.20}$$

Moreover, $\mathbf{X}_{s,t}$ is right-continuous in both s and t and

$$\mathbf{X}_{t_0,t_1}, \ \mathbf{X}_{t_1,t_2}, \ \dots, \ \mathbf{X}_{t_{n-1},t_n} \text{ are independent } \forall\ t_0 < t_1 < \cdots < t_n,$$

i.e. the flow $(\mathbf{X}_{s,t})_{s \leq t}$ has *independent increments*.

Proposition 2.15. *Let $(\mathbf{X}_{s,t})_{s \leq t}$ be a stochastic flow defined in terms of a Poisson point set σ. Let $X_0 := X(0)$ be an S-valued random variable, independent of σ. Then*

$$X(t) := \mathbf{X}_{0,t}(X_0), \qquad t > 0 \tag{2.21}$$

defines a HMC $(X(t))_{t \geq 0}$ characterized by the generator

$$\mathcal{L}f(i) := \sum_{m \in \mathcal{G}} r_m \Big(f\big(m(i)\big) - f(i) \Big), \qquad i \in S \tag{2.22}$$

Proof. The process $(X(t))_{t\geq 0}$, defined in (2.21), has piecewise constant right-continuous sample paths; indeed for any fixed $\omega \in \Omega$

$$X(t)(\omega) = \mathbf{X}_{0,t}\big(X_0(\omega)\big), \qquad t \geq 0$$

represents a right continuous step function $t \to X(t)(\omega)$, by construction. Let

$$\mathcal{P}_{ij}(t) := P\big(\mathbf{X}_{s,s+t}(i) = j\big) = P\big(\mathbf{X}_{0,t}(i) = j\big), \qquad t \geq 0. \tag{2.23}$$

One can note that it is possible to choose $s = 0$ in definition (2.23) because the point process associated to σ is invariant under time translations. Moreover, the distributions of the process $(X(t))_{t\geq 0}$ satisfy property (2.5), since X_0 is independent from σ and the flow $(\mathbf{X}_{s,t})_{s\leq t}$ has independent increments. Furthermore flow's properties (2.19) and (2.20) imply that $(\mathcal{P}_t)_{t\geq 0}$ has semigroup properties; thus the process $(X(t))_{t\geq 0}$ is a HMC, defined by $(\mathcal{P}_t)_{t\geq 0}$. It remains to show that the infinitesimal generator of $(X(t))_{t\geq 0}$ is given by (2.22); by properties of Poisson processes, it follows that

$$P\big(|\sigma_{0,t}| \geq 2\big) = o(t) \quad \text{as} \quad t \to 0$$

and

$$P\big(|\sigma_{0,t}| = 1\big) = P\big(\sigma_{0,t} = \{(m,s)\};\ \exists\ s \in (0,t]\big) = r_m t + o(t) \quad \text{as} \quad t \to 0.$$

This implies that $\forall\ f : S \to \mathbb{R}$ and $i \in S$

$$\begin{aligned}
\mathcal{P}_t f(i) &= E\Big(f\big(\mathbf{X}_{0,t}(i)\big)\Big) \\
&= \sum_{m\in\mathcal{G}} f(m(i)) P\big(\mathbf{X}_{0,t}(i) = m(i)\big) \\
&= f(i) + t\sum_{m\in\mathcal{G}} r_m\Big(f\big(m(i)\big) - f(i)\Big) + o(t) \quad \text{as} \quad t \to 0
\end{aligned}$$

That is

$$\mathcal{P}_t f = f + t\mathcal{L}f + o(t) \quad \text{as} \quad t \to 0 \iff \mathcal{L}f = \lim_{t\to 0} \frac{\mathcal{P}_t f - f}{t}$$

□

The representation (2.22) of an infinitesimal generator $\mathcal{L}$ is called *random mapping representation*. Furthermore in a random mapping representation the collection of non negative constants $r_m,\ m \in \mathcal{G}$, are called rates. They point out the probability that map m is applied. In fact, by its regular jumps structure, the

process $(X(t))_{t\geq 0}$ can be approximated by a discrete-time MC, where time is increased in steps of size dt, then $r_m\,dt$ is the probability that the map m is applied during the time interval $(t, t+dt]$. In particular, rates say with which Poisson intensity the local map m should be applied to the configuration $X(t)$ [59].

In this section, the results, that are not proven, fallow from Chapter 10 of [31] (as briefly indicated in [59]).

2.3.1 Example of Poisson Representations

It is interesting to point out that the random mapping representation of an infinitesimal generator is not unique. This fact is showed by the following example:

let $S := \{0, 1\}$ be the considered state space and let $\mathcal{L}$ be an infinitesimal generator of an S-valued HMC, such that $\mathcal{L}$ is defined by the following local characteristics:

$$q_{01} = 2 \text{ and } q_{10} = 1. \tag{2.24}$$

Hence the jumps of the considered process are

$0 \mapsto 1$ with rate 2;

$1 \mapsto 0$ with rate 1.

Moreover, one can note that local characteristics (2.24) are sufficient to define the generator $\mathcal{L}$, since

$$q_{00} = -q_{01} \text{ and } q_{10} = -q_{10}$$

by property (2.11) of HMC, defined in a finite space. Considering (2.15), the generator $\mathcal{L}$ can be write as a linear operator from $\mathbb{R}^2$ to $\mathbb{R}$:

$$\mathcal{L}f(i) = \sum_{j\neq i} q_{ij}\big(f(j) - f(i)\big) \qquad \forall i \in S. \tag{2.25}$$

In detail,

$$\mathcal{L}f(0) = 2\big(f(1) - f(0)\big) \tag{2.26}$$
$$\mathcal{L}f(1) = f(0) - f(1). \tag{2.27}$$

Let $\mathcal{G} := \{\mathsf{down},\ \mathsf{up},\}$ be a set of self maps of the state space S, defined as follow:

$$\left.\begin{array}{rcl} \mathsf{down}(i) & = & 0 \\ \mathsf{up}(i) & = & 1 \end{array}\right\} \qquad \forall i \in S.$$

Thus, the generator $\mathcal{L}$ has the following random mapping representation that is equivalent to (2.25):

$$\mathcal{L}f(i) := r_{\mathsf{down}}\big(f(\mathsf{down}(i)) - f(i)\big) + r_{\mathsf{up}}\big(f(\mathsf{up}(i)) - f(i)\big), \qquad \forall i \in S, \tag{2.28}$$

where $r_{\mathsf{down}} = 1$ and $r_{\mathsf{up}} = 2$. Generator's representation (2.28) is equivalent to (2.25); indeed:

$$\begin{aligned} i = 0 \;&\Rightarrow\; f(\mathsf{down}(i)) - f(i) = 0 \\ i = 1 \;&\Rightarrow\; f(\mathsf{up}(i)) - f(i) = 0. \end{aligned}$$

That is,

$$\mathcal{L}f(i)_{|_{i=0}} = r_{\mathsf{up}}\big(f(\mathsf{up}(i)) - f(i)\big)_{|_{i=0}} = 2\big(f(1) - f(0)\big)$$

and

$$\mathcal{L}f(i)_{|_{i=1}} = r_{\mathsf{down}}\big(f(\mathsf{down}(i)) - f(i)\big)_{|_{i=1}} = f(0) - f(1).$$

Furthermore, let $\mathsf{swap} : S \to S$ such that $\mathsf{swap}(i) := 1 - i, \quad \forall i \in S$. It is possible give another equivalent random mapping representation of $\mathcal{L}$ considering the maps set $\mathcal{G}' := \{\mathsf{swap}, \mathsf{up}\}$;

$$\mathcal{L}f(i) := r'_{\mathsf{swap}}\big(f(\mathsf{swap}(i)) - f(i)\big) + r'_{\mathsf{up}}\big(f(\mathsf{up}(i)) - f(i)\big), \qquad \forall i \in S, \tag{2.29}$$

where $r'_{\mathsf{swap}} = 1 = r'_{\mathsf{up}}$. Indeed, from representation (2.29), one has

$$\mathcal{L}f(i)_{|_{i=0}} = (r'_{\mathsf{swap}} + r'_{\mathsf{up}})\big(f(\mathsf{swap}(i)) - f(i)\big)_{|_{i=0}} = 2\big(f(1) - f(0)\big).$$

Hence, the total rate of the jump $0 \to 1$ is $r'_{\mathsf{swap}} + r'_{\mathsf{up}} = 2 = q_{01}$. Again,

$$i = 1 \;\Rightarrow\; f(\mathsf{up}(i)) - f(i) = 0 \;\Rightarrow\; \mathcal{L}f(1) = f(0) - f(1).$$

The two different random mapping representations, (2.28) and (2.29), highlight that starting from two different Poisson point sets, $\omega \in \mathcal{G} \times \mathbb{R}$ and $\omega' \in \mathcal{G}' \times \mathbb{R}$, and defining the respective stochastic flows, it is possible to give two different constructions of the same HMC (see Figure 2.1).

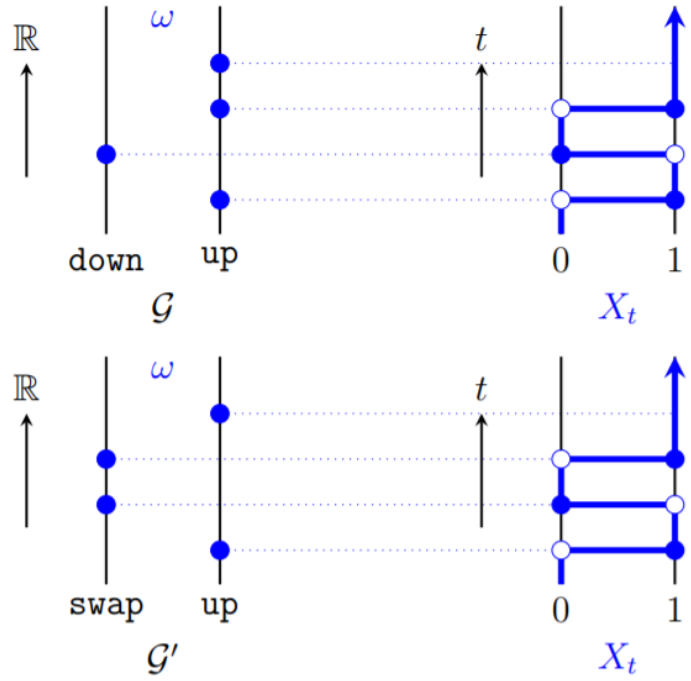

Figure 2.1: Two stochastic flows representing the same homogeneous Markov process [59].

2.4 Interacting Particle Systems

The model of cancer dynamics studied in this thesis belong to the class of *mean-field Interacting Particle System*. Mean-field assumption is a simplification of the dynamics. However the mean-field hypothesis allows for some interesting analytical results.

Definition 2.16. *Let S be a finite sent and Λ a countable set. An Interacting Particle System (IPS) is a continuous-time Markov processes $(X(t))_{t\geq 0}$ with a state space of the form S^Λ, where S is called local state space and Λ lattice.*

From now on, let $|\Lambda| < +\infty$. Thus, an interacting particle system is a MC $(X(t))_{t\geq 0}$ such that the state of the process, for every fixed time $t \geq 0$, is of the form

$$X(t) = (X_i(t))_{i\in\Lambda}\ , \qquad \text{where } X_i(t) \in S,\ \forall i \in \Lambda.$$

Positions $i \in \Lambda$ are also often named sites and $X_t(i)$ is called *local state* of $(X(t))_{t\geq 0}$ at time t and at the site i. Furthermore, interacting particle systems are defined in terms of *local maps* $(m : S^\Lambda \to S^\Lambda)$; as usual, the dynamic of this kind of processes

is described by the infinitesimal generator G, for which it is possible to choose a random mapping representation (2.22) of the form

$$Gf(x) = \sum_{m \in \mathcal{G}} r_m \{f(m(x)) - f(x)\}, \qquad x \in S^\Lambda$$

where $\mathcal{G} := \{local\ maps,\ m : S^\Lambda \to S^\Lambda\}$ and $(r_m)_{m \in \mathcal{G}} \in [0, +\infty)$ are the rates. Often, the lattice Λ has the structure of an (undirected) graph (V, E). In this case, each node in V represent a particle of the considered system and the edge $ij := (i, j) \in E$ exists if and only if between i and j there is some kind of interaction (relevant to the study of the system). The set $\mathcal{N}_i := \{j \in \Lambda : ij \in E\}$ is the neighbourhood of the site i. Two paradigmatic examples of Interacting Particle System are the "Ising model" and "the Voter model" (the latter is useful with the prospect of application to the model analysed in this thesis). Furthermore, to set out a rigorous description of the models is necessary to explain the following two theorem about functions of Markov chain [59].

Proposition 2.17. *Let $(X(t))_{t \geq 0}$ be an S-valued MC with generator G and semigroup $(\mathcal{P}(t))_{t \geq 0}$. Let T be a finite set and $f : S \to T$ be a function. For each present state $X(t) = x \in S$ and $\tilde{y} \in T$ such that $f(x) \neq \tilde{y}$, the total rate of the jump $f(x) \mapsto \tilde{y}$ defined as*

$$\mathcal{H}(x, \tilde{y}) := \sum_{\tilde{x} : f(\tilde{x}) = \tilde{y}} G(x, \tilde{x})$$

is, by assumption, of the form $\mathcal{H}(x, \tilde{y}) = H(f(x), \tilde{y})$, with H infinitesimal generator of some T-valued MC. Let $Y(t) := f(X(t))\ \forall t \geq 0$, then $(Y(t))_{t \geq 0}$ is a MC with generator H.
Conversely, if $(Y(t))_{t \geq 0}$ is a T-value MC with generator H, for all possible initial low of the process $(X(t))_{t \geq 0}$, then $\mathcal{H}(x, \tilde{y}) = H(f(x), \tilde{y}), \quad x \in S, \tilde{y} \in T$.

Proof. Let $\mathcal{H}(x, \tilde{y}) = H(f(x), \tilde{y})$ be true $\forall x \in S$ and $\tilde{y} \in T$ s.t. $f(x) \neq \tilde{y}$, than it also holds for $f(x) = \tilde{y}$

$$\begin{aligned} H(f(x), f(x)) &= -\sum_{\tilde{y} : \tilde{y} \neq f(x)} H(f(x, \tilde{y}) = -\sum_{\tilde{y} : \tilde{y} \neq f(x)} \sum_{\tilde{x} : f(\tilde{x}) = \tilde{y}} G(x, \tilde{x}) \\ &= -\sum_{\tilde{x} : f(\tilde{x}) \neq f(x)} G(x, \tilde{x}) = \sum_{\tilde{x} : f(\tilde{x}) = f(x)} G(x, \tilde{x}) \end{aligned}$$

where it is used property (2.11) applied to the infinitesimal generators H and G. Thus the initial hypothesis is equivalent to

$$\mathcal{H}(x, \tilde{y}) = H(f(x), \tilde{y}), \qquad \forall\ x \in S \text{ and } \tilde{y} \in T. \tag{2.30}$$

Let $(\mathcal{Q}(t))_{t\geq 0}$ be the semigroup generated by H, the claim is to show that equation (2.30) is equivalent to

$$\mathcal{Q}_t(f(x),\tilde{y}) = \sum_{\tilde{x}:f(\tilde{x})=\tilde{y}} \mathcal{P}_t(x,\tilde{x}) \ , \qquad t\geq 0,\ x\in S,\ \tilde{y}\in T. \tag{2.31}$$

where $\mathcal{Q}_t(i,j) := (\mathcal{Q}(t))_{i,j}$ and $\mathcal{P}_t(h,k) := (\mathcal{P}(t))_{h,k}$, for $t\geq 0,\ i,j\in S,\ h,k\in T$. Indeed, let $g: T\to\mathbb{R}$ be an arbitrary function,

$$\begin{aligned} G(f\circ g)(x) &= \sum_{\tilde{x}\in S} G(x,\tilde{x})g(f(\tilde{x})) = \sum_{\tilde{y}\in T}\sum_{\tilde{x}:f(\tilde{x})=\tilde{y}} G(x,\tilde{x})g(\tilde{y}), \\ (Hg)\circ f(x) &= \sum_{\tilde{y}\in T} H(f(x),\tilde{y})g(\tilde{y}). \end{aligned}$$

Thus, for $t\geq 0,\ x\in S,\ g: T\to\mathbb{R}$,

$$\begin{gathered} G(f\circ g)(x) = (Hg)\circ f(x) \\ \Updownarrow \\ \sum_{\tilde{y}\in T}\sum_{\tilde{x}:f(\tilde{x})=\tilde{y}} G(x,\tilde{x})g(\tilde{y}) = \sum_{\tilde{y}\in T} H(f(x),\tilde{y})g(\tilde{y}) \\ \Updownarrow \\ \mathcal{H}(x,\tilde{y}) = H(f(x),\tilde{y}) \qquad (\tilde{y}\in T). \end{gathered}$$

Namely,

$$G(f\circ g) = (Hg)\circ f \ , \qquad g: T\to\mathbb{R}, \tag{2.32}$$

is equivalent to hypothesis (2.30). Moreover,

$$\mathcal{P}_t(f\circ g) = (\mathcal{Q}_t g)\circ f \ , \qquad g: T\to\mathbb{R},\ t\geq 0, \tag{2.33}$$

is equivalent to assumption (2.31), fallowing the same computation, but using $\mathcal{P}_t$ and Q_t instead of G and H, respectively. The last step to reach the wanted claim is to show that equations (2.32) and (2.33) are equivalent:

- by (2.32) fallows

  $$G^2(f\circ g) = G(Hg)\circ f) = (H^2 g)\circ f$$

  taking it by inductive hypothesis, it springs up that

  $$G^n(f\circ g) = (H^n g)\circ f \qquad \forall n\geq 0$$

  by induction on n. The latter equation implies (2.33), by using representation (2.12) for the semigroups $\mathcal{P}_t$ and $\mathcal{Q}_t$, i.e.

  $$\mathcal{P}_t = \sum_{n=0}^{\infty}\frac{1}{n!}t^n G^n \quad \text{and} \quad \mathcal{Q}_t = \sum_{n=0}^{\infty}\frac{1}{n!}t^n H^n \tag{2.34}$$

- conversely, if equation (2.33) holds $\forall t \geq 0$, then by (2.34)

$$\sum_{n=0}^{\infty} \frac{1}{n!} t^n G^n (f \circ g) = \left(\sum_{n=0}^{\infty} \frac{1}{n!} t^n H^n g \right) \circ f. \tag{2.35}$$

  When it is stopped to the firs order as $t \to 0$, equivalence (2.35) must continue to be true, thus it fallows (2.32).

Ultimately, assumption (2.30) is equivalent to (2.31). Thus assuming by hypothesis that (2.31) holds, then, by (2.4), the finite dimensional distributions of Y are given by

$$\begin{aligned} P\big(Y(t) = \tilde{y}\big) &= \sum_{\tilde{x}: f(\tilde{x}) = \tilde{y}} P\big(X(t) = \tilde{x}\big) = \sum_{\tilde{x}: f(\tilde{x}) = \tilde{y}} \left(\sum_{x \in S} P\big(X(0) = x\big) \mathcal{P}_t(x, \tilde{x}) \right) \\ &= \sum_{x \in S} P\big(X(0) = x\big) \sum_{\tilde{x}: f(\tilde{x}) = \tilde{y}} \mathcal{P}_t(x, \tilde{x}) \\ &= \sum_{x \in S} P\big(X(0) = x\big) Q_t(f(x), \tilde{y}) \\ &= \sum_{y_0 \in f(S)} \sum_{x_0: f(x_0) = y_0} P\big(X(0) = x_0\big) \mathcal{Q}_t(y_0, \tilde{y}) \\ &= \sum_{y_0 \in f(S)} P\big(Y(0) = y_0\big) \mathcal{Q}_t(y_0, \tilde{y}) \end{aligned}$$

Again, by (2.4), since the semigroup $\mathcal{Q}_t$ is generated by H.
Conversely, if Y is a MC with generator H for all possible initial low of the process $(X(t))_{t \geq 0}$, then

$$\begin{aligned} P\big(Y(t) = \tilde{y}\big) &= \sum_{y_0 \in f(S)} P\big(Y(0) = y_0\big) \mathcal{Q}_t(y_0, \tilde{y}) \\ &= \sum_{y_0 \in f(S)} \sum_{x_0: f(x_0) = y_0} P\big(X(0) = x_0\big) \mathcal{Q}_t(y_0, \tilde{y}) \\ &= \sum_{x \in S} P\big(X(0) = x\big) \mathcal{Q}_t(f(x), \tilde{y}) \end{aligned}$$

and

$$\begin{aligned} P\big(Y(t) = \tilde{y}\big) &= \sum_{\tilde{x}: f(\tilde{x}) = \tilde{y}} P\big(X(t) = \tilde{x}\big) \\ &= \sum_{\tilde{x}: f(\tilde{x}) = \tilde{y}} \left(\sum_{x \in S} P\big(X(0) = x\big) \mathcal{P}_t(x, \tilde{x}) \right) \\ &= \sum_{x \in S} P\big(X(0) = x\big) \sum_{\tilde{x}: f(\tilde{x}) = \tilde{y}} \mathcal{P}_t(x, \tilde{x}) \end{aligned}$$

In particular,

$$\sum_{x\in S} P\big(X(0)=x\big) \sum_{\tilde{x}:f(\tilde{x})=\tilde{y}} \mathcal{P}_t(x,\tilde{x}) = \sum_{x\in S} P\big(X(0)=x\big)\, \mathcal{Q}_t(f(x),\tilde{y})$$

thus, it follows equation (2.31), that is equivalent to

$$\mathcal{H}(x,\tilde{y}) = H(f(x),\tilde{y}), \quad x\in S, \tilde{y}\in T.$$

□

Put simply, given a function of a S-valued Markov Process, $(f(X(t)))_{t\geq 0}$, if its jump rates are function of the present state only (i.e. for $t \geq 0$ fixed time and $x = X(t)$), $f\big(X(t)\big)$ jumps with rates that depend exclusively on x, and not on the other states in S, then $\big(f\big(X(t)\big)\big)_{t\geq 0}$ is itself a Markov process. The other important result explains under which hypotheses it is feasible to approximate a IPS by solutions of a differential equation. Following the way showed in Paragraph 3.4 of [59], the chosen setting looks like as follows, $\forall N > 0$:

- let $(X^N(t))_{t\geq 0}$ be a MC with finite state space S_N, generator G_N and semigroup $(P_N(t))_{t\geq 0}$;
- let $f_N : S_N \to \mathbb{R}$ be a function;
- let the function α_N be the *quadratic variation*

$$\alpha_N(x) := \sum_{\tilde{x}\in S_N} G_N(x,\tilde{x})(f_N(\tilde{x}) - f_N(x))^2 \; ; \tag{2.36}$$

- let the function β_N be the *drift*

$$\beta_N(x) := \sum_{\tilde{x}\in S_N} G_N(x,\tilde{x})(f_N(\tilde{x}) - f_N(x)) \; . \tag{2.37}$$

Following [59], if $\alpha_N \xrightarrow{N\to\infty} 0$ and β_N approximates an appropriate, Lipschitz continuous function of $Y^N(t) := f_N(X^N(t))$, then, using [18, Thm 4.1], it is possible to show that the process $(Y^N(t))_{t\geq 0}$, should in the limit be describe by the solution of a differential equation. To rigorously formalize this idea, it is necessary to introduce some other assumption (as it is was done in [59]):

- $\forall N > 0, \quad f_N : S_N \to I \subseteq \mathbb{R}$,
  where I is a closed interval s.t. $I_- := \inf I \; , \; I_+ := \sup I \in [-\infty, +\infty]$

- $\exists b : I \to \mathbb{R}$ globally Lipschitz continuous function s.t. β is uniformly approximated by $b \circ f_N$, i.e.

$$\sup_{x \in S_N} |\beta_N(x) - b(f_N(x))| \xrightarrow{N\to\infty} 0 \tag{2.38}$$

and

$$I_- > -\infty \;\Rightarrow\; b(I_-) \geq 0 \quad \text{and} \quad I_+ < +\infty \;\Rightarrow\; b(I_+) \leq 0 \tag{2.39}$$

Consequently, for each initial state $y_0 \in I$ the differential equation

$$\dot{y}(t) = b(y(t)) \qquad (t \geq 0)$$

has a unique I-valued solution $(y(t))_{t\geq 0}$.

Theorem 2.18. *Let $f_N(X^N(0))$ converge in probability to $y(0)$ and, apart from (2.38), suppose that*

$$\sup_{x \in S_N} \alpha_N(x) \xrightarrow{N\to\infty} 0. \tag{2.40}$$

Then

$$P\big(\big|f_N(X^N(t)) - y(t)\big| \leq \epsilon \;\forall t \in [0,T]\big) \xrightarrow{N\to\infty} 1, \quad \forall\, T < \infty \text{ and } \epsilon > 0\ .$$

Proof. Let $t < \infty$, $\epsilon > 0$ and $y_0 \in I$ be fixed. Let L denote the Lipschitz constant of the function b. Since f_N takes value in a closed interval $I \in \mathcal{R}$ then the more general condition

$$\forall x \in S_N \text{ and } t \leq t_0, \quad |f_N(x) - y(t)| < \epsilon \;\Rightarrow\; f_N(x) \in I,$$

which is necessary to use Theorem 4.1 of [18], is directly satisfied. Set $\delta := \frac{\epsilon}{3} e^{-LT}$ and let the following events be considered

$$\begin{aligned}
\Omega_0 &:= \big\{\big|f_N(X(0)) - y(0)\big| \leq \delta\big\},\\
\Omega_1 &:= \left\{\int_0^T \big|\beta_N(X(t)) - b(f_N(X(t)))\big|\, dt \leq \delta\right\},\\
\Omega_{K,2} &:= \left\{\int_0^T \alpha_N(X(t))\, dt \leq KT\right\}, \quad K > 0.
\end{aligned}$$

Then Theorem 4.1 of [18] implies that

$$P\left(\sup_{t\in[0,T]} \big|f_N(X(t)) - y(t)\big| > \epsilon\right) \leq 4KT\delta^{-2} + P\big({}^c\Omega_0 \cup^c \Omega_1 \cup^c \Omega_{K,2}\big) \tag{2.41}$$

and assumption $f_N(X^N(0)) \xrightarrow{N\to\infty} y(0)$ implies $P\big({}^c\Omega_0\big) \xrightarrow{N\to\infty} 0$. Moreover,

$$\int_0^T \big|\beta_N(X(t)) - b(f_N(X(t)))\big|\,dt \le \sup_{x\in S_N} \big|\beta_N(x) - b(f_N(x))\big| - T \le \delta,$$

from hypothesis (2.38) fallows that $P\big({}^c\Omega_1\big) = 0$, for N sufficiently large. Again,

$$\int_0^T \alpha_N(X(t))\,dt \le \sup_{x\in S_N} \alpha_N(x) \quad \Rightarrow \quad P\big({}^c\Omega_{K,2}\big) = 0\ \forall N > 0,$$

by assumption (2.40). Finally, choosing $K = \sup_{x\in S_N} \alpha_N(x)$ in the inequality (2.41), then

$$P\bigg(\sup_{t\in[0,T]} \big|f_N(X^N(t)) - y(t)\big| > \epsilon\bigg) \xrightarrow{N\to\infty} 0.$$

□

Hence, the reached goal is that theorem (2.18) provides sufficient conditions for the I-valued processes $\big(f_N\big(X^N(t)\big)\big)_{t\ge 0}$ to approximate a solution of the differential equation.

2.4.1 Curie-Weiss Model

The Curie-Weiss model is the mean-field approximation of the Ising model. Sites are interpreted as atoms in a crystal structure, which is depicted by a graph, and the local state space is $S = \{-1, +1\}$. Traditionally, the local state $x_i \in \{-1, +1\}$ of a site i is usually called the spin at i, because it is interpreted as the direction of magnetic field of the atoms [59]. According to Glauber dynamics, sites update their spin values with rate one, loosing "memory" of previous state. For each site $i \in \Lambda$, the choice of new spin value depends on the spin values of neighbors, trough a real parameter β. In order to study the mean-field limit, it is necessary to adapt the model on the complete graph Λ_N (that is the lattice) with $|\Lambda_N| = N$ vertices (that correspond to the sites). Hence, each site is a neighbour of each other site and, for simplicity, a site is considered as a neighbour of itself. Consequently, the resulting process $(X(t))_{t\ge 0}$ has the following structure, for an arbitrarily fixed time $t \ge 0$:

- an arbitrary state of the system $x = X(t)$ is configured as

$$x = (x_1, x_2, \ldots, x_i, \ldots, x_N), \qquad x_k \in \{-1, +1\} \qquad (k \in \{1, \ldots, N\});$$

- the dynamic of the process has the following infinitesimal description:

$$P(X(t+dt) = x^i \mid X(t) = x) \;=\; \frac{e^{\frac{\beta}{N} N_x(-x_i)}}{\sum_{s\in S} e^{\frac{\beta}{N} N_x(s)}} dt + o(dt) \tag{2.42}$$

where the transition from x to $x^i := (x_1, x_2, \ldots, -x_i, \ldots, x_N)$ means that the i-th spin is flipped and

$$N_x(s) := \sum_{j=1}^{N} \mathbf{1}_{\{x_j = s\}}, \quad s \in S.$$

Besides, if $\beta > 0$ the model is called *ferromagnetic* and sites prefer to have spin values that agree with as many neighbours as possible, otherwise, when $\beta < 0$, sites are inclined to have the opposite spin values with respect to the main part of their neighbours and the model is recognized as antiferromagnetic [59]. One can note that in a ferromagnetic setting ($\beta > 0$),

- if x_i has different sign with respect to the major part of spin values, then $N_x(-x_i)$ high and consequently so is the transition rate;

- if x_i has sign equal to the ones of most sites, then the probability to change value is low.

Now, the key point is to consider the *average magnetization*

$$m_N(t) := \frac{1}{N} \sum_{j=1}^{N} X_j(t) \qquad t \geq 0,$$

which is a function of the process $(X(t))_{t\geq 0}$, taking values in the space

$$S_N := \left\{ -1, -1 + \frac{2}{N}, \ldots, 1 - \frac{2}{N}, 1 \right\}.$$

Let $t \geq 0$ be fixed and $\bar{x} := m_N(t)$ a chosen state at time t. The process $(m_N(t))_{t\geq 0}$ jumps, with probability (2.42), from $\bar{x}$ to $\bar{x} + \frac{2}{N}$, every time that a spin is flipped to 1, and to $\bar{x} - \frac{2}{N}$, when a spin with positive values flips its sign. That is

$$\bar{x} \quad \text{switch to} \quad \bar{x} + \tfrac{2}{N} \qquad \text{with rate} \qquad N_x(-1) \frac{e^{\frac{\beta}{N} N_x(1)}}{e^{\frac{\beta}{N} N_x(-1)} + e^{\frac{\beta}{N} N_x(1)}}$$

$$\bar{x} \quad \text{switch to} \quad \bar{x} - \tfrac{2}{N} \qquad \text{with rate} \qquad N_x(1) \frac{e^{\frac{\beta}{N} N_x(-1)}}{e^{\frac{\beta}{N} N_x(-1)} + e^{\frac{\beta}{N} N_x(1)}}.$$

Furthermore, one can observe that, for a considered state x, the number of sites with spin value 1 is

$$\begin{aligned} N_x(1) &= \#\{i \in \Lambda_N : x_i = 1\} \\ &= \sum_{i=1}^{N} \mathbf{1}_{x_i=1} = \sum_{i=1}^{N} \frac{1+x_i}{2} = \frac{N}{2}\sum_{i=1}^{N} \frac{1+x_i}{N} \\ &= \frac{N}{2}(1+m_N), \end{aligned}$$

similarly, the number of sites with spin value -1 is

$$\begin{aligned} N_x(-1) &= \#\{i \in \Lambda_N : x_i = -1\} \\ &= \frac{N}{2}\sum_{i=1}^{N} \frac{1-x_i}{N} \\ &= \frac{N}{2}(1-m_N) \end{aligned}$$

As consequence, it results that the possible jumps for average magnetization process are:

$\bar{x} \mapsto \bar{x} + \frac{2}{N}$ with rate $r_+(\bar{x}) := \frac{N}{2}(1-m_N)\frac{e^{\beta(1+m_N)/2}}{e^{\beta(1-m_N)/2}+e^{\beta(1+m_N)/2}}$

$\bar{x} \mapsto \bar{x} - \frac{2}{N}$ with rate $r_-(\bar{x}) := \frac{N}{2}(1+m_N)\frac{e^{\beta(1-m_N)/2}}{e^{\beta(1-m_N)/2}+e^{\beta(1+m_N)/2}}$.

These rates depend on the present state $\bar{x}$ only, therefore applying theorem (2.17) it results that $(m_N(t))_{t\geq 0}$ is a Markov process with infinitesimal generator

$$\begin{aligned} \mathcal{G}_N f(m_N) &= r_-(\bar{x})\left(f\left(m_N - \frac{2}{N}\right) - f(m_N)\right) + r_+(\bar{x})\left(f\left(m_N + \frac{2}{N}\right) - f(m_N)\right) \\ &= \frac{N}{2}(1+m_N)\frac{e^{-\beta m_N/2}}{e^{-\beta m_N/2}+e^{\beta m_N/2}}\left(f\left(m_N - \frac{2}{N}\right) - f(m_N)\right) \\ &+ \frac{N}{2}(1-m_N)\frac{e^{\beta m_N/2}}{e^{-\beta m_N/2}+e^{\beta m_N/2}}\left(f\left(m_N + \frac{2}{N}\right) - f(m_N)\right). \end{aligned}$$

In order to approximate $(m_N(t))_{t\geq 0}$ by a differential equation, choosing $f_N = id$ (i.e. $f(\bar{x}) = \bar{x}$), hence, the quadratic variation and the (local) drift rate of the process (according to definitions (2.36) and (2.37)) are respectively

$$\begin{aligned} \alpha_N(m_N) &= r_-(m_N)\left(-\frac{2}{N}\right)^2 + r_+(m_N)\left(\frac{2}{N}\right)^2 \\ &= \frac{2}{N}\frac{(1+m_n)e^{-\beta m_N/2} + (1-m_N)e^{\beta m_N/2}}{e^{-\beta m_N/2}+e^{\beta m_N/2}} \\ &= \frac{2}{N}\left(1 + m_N\frac{e^{-\beta m_N/2} - e^{\beta m_N/2}}{e^{-\beta m_N/2}+e^{\beta m_N/2}}\right) \end{aligned}$$

$$\begin{aligned} g_\beta(m_N) &= \mathcal{G}_N f_N(m_N) = \mathcal{G}_N(m_N) \\ &= r_-(m_N)\left(-\frac{2}{N}\right) + r_+(m_N)\frac{2}{N} \\ &= -(1+m_N)\frac{e^{-\beta m_N/2}}{e^{-\beta m_N/2}+e^{\beta m_N/2}} + (1-m_N)\frac{e^{\beta m_N/2}}{e^{-\beta m_N/2}+e^{\beta m_N/2}} \\ &= \frac{(1-m_n)e^{\beta m_N/2} - (1+m_N)e^{-\beta m_N/2}}{e^{-\beta m_N/2}+e^{\beta m_N/2}} \\ &= \frac{e^{\beta m_N/2} - e^{-\beta m_N/2}}{e^{\beta m_N/2}+e^{-\beta m_N/2}} - m_N = tanh\left(\frac{\beta}{2}m_N\right) - m_N \end{aligned}$$

Thus, it tour out that $\alpha_N(x) \xrightarrow{N\to\infty} 0$, namely assumptions (2.40) is verified. Moreover, g_β does not explicitly depend on N, thus selecting $b := g_\beta$, hypothesis (2.38) is verified too; in particular, tanh is a Lipschitz continuous function. Lastly, applying theorem (2.18) the process $(m_N(t))_{t\geq 0}$ can be approximate by a solution of the differential equation

$$\dot{m}(t) = g_\beta(m(t)) \qquad (t \geq 0) \tag{2.43}$$

and the validity of (2.39) guarantees the unicity of solution for each initial state $m(0) = m_0$.

Analysis of the model

Studying the behaviour of a model, it is significant highlight its *phase transitions*, which are an abrupt changes in behaviour of the system, due to particular values, called critical points, assumed by a parameter governing the dynamics. Paying attention to Curie Weiss model, the fixed points of the differential equation (2.43) are the solution of the equation

$$\tanh\left(\frac{\beta}{2}m\right) = m \qquad \big(\iff \dot{m} = 0 = g_\beta(m)\big).$$

Taking advantage of the following facts

$$|\tanh(x)| < |x|, \qquad \lim_{x\to 0}\frac{\tanh(x)}{x} = 1, \qquad \lim_{x\to\pm\infty}\tanh(x) = \pm 1, \quad x \in \mathbb{R},$$

one has that,

- if $\beta = 2$, then the differential equation (2.43) has an unique fixed point in $m_0 = 0$ (as in Figure 2.2(a));

- for $\beta < 2$ and $x \in \mathbb{R}$ it holds that $\left|\tanh\left(\frac{\beta}{2}x\right)\right| < \left|\frac{\beta}{2}x\right| < |x|$. Thus, the differential equation (2.43) still has a single fixed point at $m_0 = 0$ (see Figure 2.2(a));

- otherwise, $\beta > 2$, the following laying arises

  $\left|\tanh\left(\frac{\beta}{2}x\right)\right| \geq |x|$ near the zero;

  $\left|\tanh\left(\frac{\beta}{2}x\right)\right| < |x|$ for some $x > \frac{1}{2}$.

  Therefore, there are three fixed points of (2.43), $m_- < 0 < m_+$ (see Figure 2.2(b)).

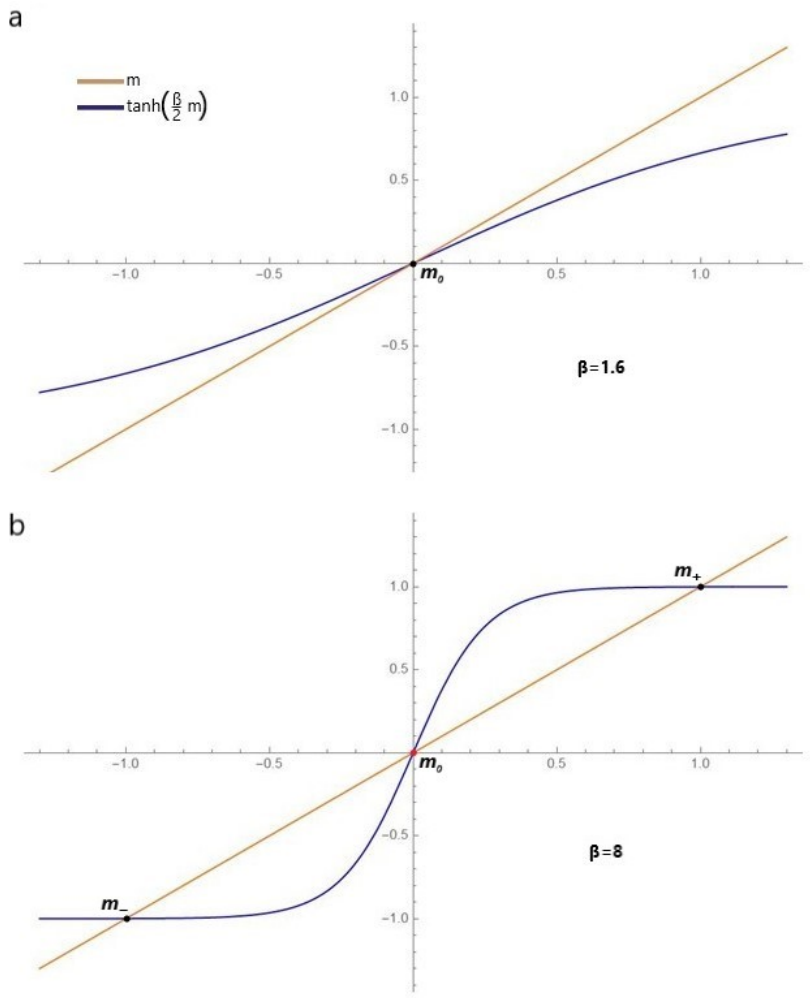

Figure 2.2: The graphic solutions of equation $\tanh\left(\frac{\beta}{2}m\right) = m$, for $\beta = 1.6$ (a) and $\beta = 8$ (b).

Consequently, it emerges that

- for $\beta \leq 2$ and for any choice of the initial state, $m(0) = m_0 \in \mathbb{R}$, the differential equation (2.43) has an unique solution converges to the stable equilibrium $m(t) = 0$, as $t \to \infty$;

- for $\beta > 2$, if $m(0) < 0$, then the solutions of (2.43) converge to the stable fixed point m_-, as time t increases, else if solutions start in $m(0) > 0$, then they converge to the stable fixed point m_+. Finally, the fixed point $m_0 = 0$ results to be unstable.

2.4.2 Voter Model

Initially, Voter model is used to describe the evolution of opinions in a population [59]:

- sites in a lattice Λ represent people;

- states in a finite local state space S represent political opinions and are called *types*;

- each voter is influenced only by a fixed set of neighbours;

- the rules of the dynamics is the following: with rate one, an individual becomes unsure what political party to vote for and copies the opinion of a randomly chosen neighbour;

- this update is repeated until the finite population of N individuals reaches consensus.

The Voter model is often used to model biological populations, where organisms with different genetic types occupy sites in space, with rate 1, the organism living at a given site dies and is replaced by a descendant, chosen with equal probability its neighbours [59]. This last application is very workable to the model studied in this thesis, where

- sites represent cells with the same genotype,

- states identify the different phenotypes which come from the same genotype,

- with probability 1, a cell is randomly picked, forgets its phenotype and changes it choosing, with equal probability, among the phenotypes of its neighbours (i.e. the transition rates correspond to the mean values of cells for each phenotype).

To rigorously describe the model, let the local state set be $S := \{0, 1\}$ and let η_i be the type at the site $i \in \Lambda_N$. The lattice Λ_N has the structure of the complete graph with N vertices (as in Curie-Weiss model), hence the process $(\eta(t))_{t\geq 0}$ has following structure:

- fixed a time t, an arbitrary state of the system is configured as

$$\eta = (\eta_1, \eta_2, \ldots, \eta_i, \ldots, \eta_N), \qquad \eta_i \in \{0, 1\};$$

- the dynamic of the process is infinitesimally described by

$$P(\eta(t + dt) = \eta_{s,i} \mid \eta(t) = \eta) = \overline{\eta}_s \, dt + o(dt), \qquad s \in \{0, 1\}$$

where

$$\eta_{s,i} := (\eta_1, \ldots, \eta_{i-1}, s, \eta_{i+1} \ldots, \eta_N) \quad \text{and} \quad \bar{\eta}_s := \bar{\eta}_{s,i} = \frac{1}{N} \sum_{i=1}^{N} \delta_s(\eta_j), \quad \forall i \in \Lambda_N.$$

An example is depicted by Figure 2.3, where downward arrow indicates the 0-type and upward arrow individuates the 1-type.

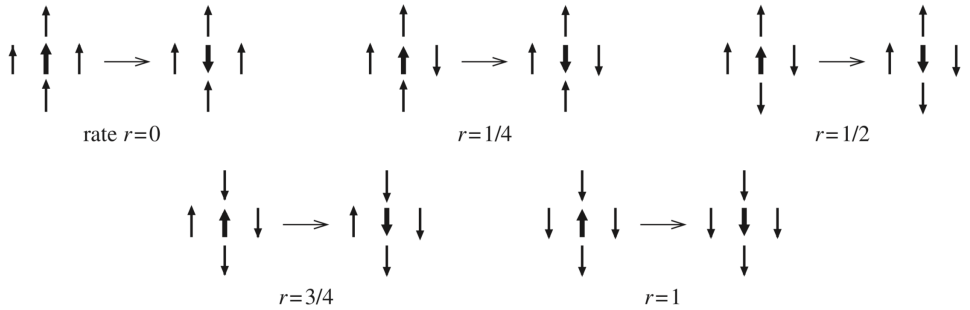

Figure 2.3: The switch rates of the central site on the square lattice Λ_2 [59].

It is helpful to notice that, taking in account the *empirical mean*

$$\overline{\eta} = \bar{\eta}^N := \frac{1}{N} \sum_{j=1}^{N} \delta_1(\eta_j) \tag{2.44}$$

then, the infinitesimal generator of $(\eta(t))_{t\in\mathbb{R}^+}$ can be write as

$$\mathcal{L}_N f(\eta) = \sum_{j=1}^{N} (1 - \overline{\eta}) \big(f(\eta_{0,j}) - f(\eta)\big) + \overline{\eta} \big(f(\eta_{1,j}) - f(\eta)\big)$$

where $1-\overline{\eta}=\overline{\eta}_0$ and $\overline{\eta}=\overline{\eta}_1$. Thus, it is possible to define the infinitesimal generator for the Markovian dynamic induced by the empirical mean $\bar{\eta}$:

$$\mathcal{G}_N f(\overline{\eta}) = N\bar{\eta}(1-\overline{\eta})\left(f\left(\overline{\eta}-\frac{1}{N}\right)-f(\overline{\eta})\right)+N(1-\overline{\eta})\bar{\eta}\left(f\left(\overline{\eta}+\frac{1}{N}\right)-f(\overline{\eta})\right),$$

Indeed, the process $(\bar{\eta}(t))_{t\geq 0}$ is a function (see (2.44)) of the MC $(\eta(t))_{t\geq 0}$, with jumps

$\bar{\eta}\to\bar{\eta}-\frac{1}{N}$ every time that type 0 switches to 1

$\bar{\eta}\to\bar{\eta}+\frac{1}{N}$ for each type 1 that switches to 0,

additionally, from (2.44) it results

$$\begin{aligned}\#\{i\in\Lambda_N:\eta_i=1\} &= \sum_{i=1}^{N}\delta_1(\eta_i)=N\bar{\eta}\\ \#\{i\in\Lambda_N:\eta_i=0\} &= \sum_{i=1}^{N}\delta_0(\eta_i)=N(1-\bar{\eta}).\end{aligned}$$

This implies that, the rates

$$r_-:=N\bar{\eta}(1-\overline{\eta}) \quad \text{and} \quad r_+:=N(1-\overline{\eta})\bar{\eta}$$

are dependent only on the present state $\bar{\eta}$, hence the process $(\bar{\eta}(t))_{t\geq 0}$ has the Markov property (by theorem 2.17). Now, considering $f_N\equiv id$ it is a very simple calculus to determine the quadratic variation α and the drift β:

$$\begin{aligned}\alpha_N(\bar{\eta}) &= r_-(\bar{\eta})\frac{1}{N^2}+r_+(\bar{\eta})\frac{1}{N^2} = \frac{2}{N}\bar{\eta}(1-\bar{\eta})\\ \beta_N(\bar{\eta}) &= r_-(\bar{\eta})\frac{1}{N}+r_+(\bar{\eta})\frac{1}{N} = 0.\end{aligned}$$

Applying Theorem 2.18, as $N\to\infty$ the process $(\bar{\eta}^N(t))_{t\geq 0}$ is approximately constant as a function of t, whereas

$$\dot{\bar{\eta}}(t)=0 \qquad t\geq 0. \tag{2.45}$$

Moreover, it is interesting to highlight an immediate consequence of ODE (2.45):

$$\lim_{t\to\infty}\lim_{N\to\infty}\bar{\eta}^N(t)=\bar{\eta}_0\ , \qquad \forall \text{ initial state } \bar{\eta}_0. \tag{2.46}$$

Contrariwise, for finite fixed $N>0$, the process $\bar{\eta}^N(t)$, which takes values in the state space

$$\left\{0,\frac{1}{N},\ldots,\left(1-\frac{1}{N}\right),1\right\},$$

has two adsorbing states 0 and 1. Thus, starting from an arbitrary initial state $\bar{\eta}_0$, $\bar{\eta}^N(t)$ fluctuates and falls in one of the two absorbing states, as the time t increases, i.e., for any initial state $\bar{\eta}_0$,

$$\lim_{N\to\infty}\lim_{t\to\infty}\bar{\eta}^N(t) \in \{0,1\}. \tag{2.47}$$

Therefore, the two limit in the equation (2.46), can not be interchanged. The reason why the double limits (2.46) and (2.47) can coexist is that the process $\bar{\eta}^N(t)$ has a substantial slowdown as $N \to \infty$. Namely, the process reach one of the absorbing state (i.e. it goes far from an initial mean $\bar{\eta}_0 \notin \{0,1\}$) in a very long time.
Voter model represents the skeleton of the stochastic model for immunotherapy of cancer discussed in Chapter 3.

Chapter 3

A Stochastic Model for Immunotherapy of Cancer

Firstly, the present chapter highlights the fundamental role of mathematics to understand the complexity of cancer phenomena. In the last years, within the framework of cancer ecology, many arising research projects focus on quantitative stochastic models of cancer evolution and statistics of cancer. An important aim of these researches is to give useful tools to take part to design effective therapy. According to an ecological point of view, an organism affected by cancer can be considered as a stable ecosystem, in which a new species springs up. The new interactions between pre-existing and rising species (which represent healthy and cancer cells respectively) unbalance the equilibrium state of the ecosystem, triggering a struggle for survival, that can lead to ecosystem collapse. The kernel of this chapter is the analysis of the stochastic model for immunotherapy of cancer proposed in the article [7]. This model is an extension of Interacting Particle System (IPS), in particular it is a modification of Voter Model (VM), which takes into account interactions among species. Interactions among various types of cells are fundamental to set up a lifelike model describing a cancer behaviour under therapy. In fact, due to the accumulation of driver and passenger mutations, cancer cells form an heterogeneous population [39]. Moreover, because of therapy and immune response, behaves and interactions of lymphocytes and cytokines must be consider, apart from that of cancer cells. Following [7], the considered model is applied to the example of melanoma under T-cell therapy. Understanding purely stochastic events, which cannot be obtained with deterministic models, may help to understand the resistance of tumours to therapeutic approaches and may have non-trivial consequences on tumour treatment protocols [7]. Additionally, numerical simulations supports the obtained results.

3.1 Mathematics of Cancer

> One of the primary aims of mathematical modelling is to make the system being studied more understandable. This often means defining the system as simply as possible, and not making it more complex than reality. [9]

In last decades, as cancer therapy has moved towards personalized treatment, mathematical modelling approaches have acquired significant attention in cancer research, representing an important tool for quantitative description of physiopathological phenomena (e.g. cancer evolution) [1, 2, 4, 8]. Above all, quantitative models represent practical tools to clarify cancer mechanisms and to provide quantitative predictions, that can be validated by experimental data. Furthermore, quantitative models can complement experimental and clinical studies, because they allow for a better understanding of cancer biological processes like cancer initiation, progression and metastases as well as intra-tumour heterogeneity, treatment responses and resistance. Thus, quantitative models generate useful individual clinical predictions, for instance for a cancer personalized therapeutic management [4, 63]. In particular, cancer ecology comes up in the setting of mathematical cancer research, as a promising quantitative approach [4, 9, 38, 46, 52]. To understand how an ecological view of cancer may be beneficial, it is crucial to consider the interactions among mutated cells and their microenviroment. Taking in account that an ecosystem consists of individuals and the physical environment they inhabit, it is possible to think of the different types of cell as species in an ecological environment. The different types of healthy cell in an organism behave like different interacting species in an ecosystem and cancer cells constitute an additional species, which spreads in the environment and favours some species and/or damages others. Namely, cancer cells evolve, adapt to and change the environment in which they live [9]. Thus, from an ecological perspective, cancer is a dynamical and heterogeneous disease, that continuously evolves and diversifies as an adaptive Darwinian system. Darwinian point of view unites genetics with population biology and biodiversity [30]. As a consequence, cancer modelling should include approaches from ecology, population dynamics and evolutionary game theory.

3.1.1 Population Dynamics Models

First approaches for tumour growth modelling came from population dynamics field. In this framework, cancer dynamics is approximated by a deterministic system, namely cancer growth curves is described by ordinary differential equations (ODEs). The aim of ODE models is to predict the rate of change in the cancer volume with respect to the changes in time [60, 63]. The *exponential model* is a

well-known ODE model used to approximate tumours in the earlier part of their observed growth period [55]. It is assumed that cancer growth rate (i.e. rate of volume change) is proportional to the volume of cancer mass. That is, the model is described by the following ODE

$$\dot{V} = aV \ ,$$

where $V = V(t)$ is the volume of cancer cells and a is the intrinsic rate of natural increase [65] of tumour ($a = b - d$ where b and d are the birth and death rates respectively). As time increases, exponential growth is possible only when infinite natural resources are available, thus it does not depict a real biological system definitively. Indeed, the exponential model fails at the last stages, when tumour colonial expansion starts to play a role and resources, provided by the host organism, become depleted. To model the reality of limited resources, the *logistic model* was developed by ecologist. Therefore, using the logistic model to describe cancer growth represents a tentative to face the limitations of exponential model to predict the long-term growth rate of cancer cell proliferation [51, 60]. The logistic differential equation is

$$\dot{V} = aV\left(1 - \frac{V}{b}\right)$$

where a is the intrinsic growth rate and b is the carrying capacity (that is the maximum volume size that can be reached by the population in its environment). One can note that if V is very small (i.e. $\frac{V}{b} \to 0$), than population is not influenced by carrying capacity and it approximately has an exponential growth. Conversely, when V is large (i.e. it is close to b) the growth rate is close to zero, which means that population growth is slowed greatly or even stopped. Thus, population growth is dramatically slowed in large populations by the carrying capacity b [51]. Summing up, populations with unlimited resources exhibit exponential growth, resulting in a J-shaped curve. When resources are limited, population expansion decreases as resources become scarce and the growing stops when the carrying capacity of the environment is reached. Thus populations realize logistic growth, resulting in an S-shaped curve ([51], see figure 3.1) .

However, exponential and logistic models failed to fit the experimental data, whereas the *Gompertz model* shows excellent descriptive power [63]. Gompertz model is a generalization of logistic model and it is widely used to describe cancer growth. Indeed, Gompertz model generates very good fits with data [60, 63]. The mathematical equation for the model is the following ODE

$$\dot{V} = aV \ln\left(\frac{b}{V}\right)$$

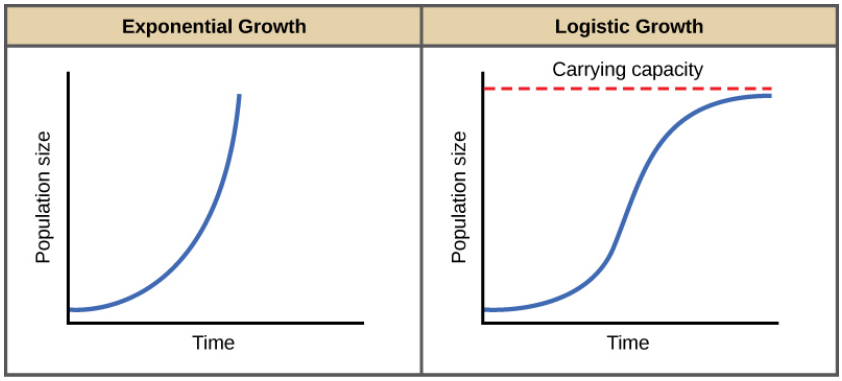

Figure 3.1: The logistic model depicts an S-shaped growth curve (on right side), and it is a more realistic model of population growth than exponential growth, that is described by a J-shaped curve (on left side) [51].

where a is the intrinsic growth rate and b is the carrying capacity. One can note that both logistic and Gompertz models approximate an exponential growth in the early state of the system (i.e. for small value of V). Lastly, in paper [47] are listed, analysed and compared several commonly used ODE models, apart from exponential, logistic and Gompertz models. The authors highlight that does not exist a model better than another, but the crucial fact is the choices of growth model, parameters and initial conditions with respect to the problem under consideration. Furthermore, they studies lead to the following conclusion:"the model that best fits experimental data might not be the model that best predicts future growth." [47]

3.1.2 Game Theory Models

Game theory is among useful mathematical tools to study evolution in ecosystems. This ecological subfield in game theory is known as *Evolutionary Game Theory* (EGT). A key aspect, of games studied in game theory, is that an isolated strategy does not produce good or bed results. The outcome affecting a player, can be defined good or bed if the strategy used is compared with the strategies employed by the other players [9]. This game set fits the evolution of tumour, where cancer cells "play" with the surrounding healthy cells. Cells of each type, sharing their phenotype, compete for available resources. The cell phenotype corresponds to its growth strategy in the game model and, as time increases, it changes with respect to microenviroment alterations. Typically, such a game is formulated as a table that ascribes fitness values (pay-offs) to every pair of interactions between cell phenotypes (strategies); individuals with best phenotype will spread in the population (i.e. it wins the game). Essentially, with the support of EGT, the

long-term proliferation of different-type cells, in an organism, can be analysed in a game context [4, 7]. In the paper [9] it is analysed the following evolutionary game with two strategy: let hawk and dove be considered as two different types of individual of the same species, which have aggressive and meek strategies to resolve disputes over food, respectively. The following pay-off table shows the game behaviour

	hawk	**dove**
hawk	$\frac{V-C}{2}$	V
dove	0	$\frac{V}{2}$

Accordingly, three scenarios arise:

i. When two hawks dispute over food they fight and the victor takes the spoils (V), whereas the loser is assumed to be severely harmed (-C).

ii. When two doves have to share food (which we will refer as V) they just divide it into two halves (each getting $\frac{V}{2}$).

iii. In the third scenario, when a hawk and a dove meet, the dove balks away from the fight leaving all the food to the hawk.

EGT provides a few things about this population:

- a population made of dove-like individuals (meek phenotype) is susceptible to be invaded by a few hawks (aggressive phenotype);

- in many cases a population made of hawks is unlikely to be immune to invasion by a handful of doves (the wildness of the hawk phenotype is counterproductive for the hawks themselves, in fact they fight each other for the spoil, keeping their growth rate low and consequently favouring the growth of doves).

Furthermore with some additional informations (e.g. the wight of average injury and the influence of a given resource on growth rate), it is possible to establish what the proportion of aggressive versus meek individuals would be in the long term. Thus, hawk-dove game can be used to study, for example, a cancer cell populations with cells that move away when confronted with scarce resources (motile) and cells that stay to use them (proliferative) [9]. In Section 1.2 it is highlight that tumour cells acquire a number of new phenotypical capabilities on the path towards

malignancy. Additionally, article [7] points out that cancer cells can switch their phenotypes (e.g. from differentiated to dedifferentiate) as strategy to avoid the immune response. The study of different tumour phenotypes with EGT leads to investigate the possible progressive steps that characterize cancer progression [9]. For a mathematical analysis of the evolutionary dynamics of an hawk-dove-like game (i.e. a game with two strategies) applied to cancer, the following setting (proposed in [61]) is consider:

- let A and B be the two different phenotypes shearing in the cell population; for example A and B could be two different phenotypes related to the same genotype in cancerous cell (as is in the simplification of model analysed in Section 3.2 of this thesis) or A could represent the phenotypes of cancerous cells and B could be the phenotypes of healthy cells;

- it is assumed that the population size is constant in time;

- let x_k denote the frequency of individuals shearing phenotype (i.e. adopting strategy) $k \in \{A, B\}$. Moreover,

$$x_A + x_B = 1 \tag{3.1}$$

  since only two strategy are allowed (by assumption);

- the payoff matrix for the game is

	A	B
A	a	b
B	c	d

  hence, strategy A player receives payoff a when playing against another strategy A player, and payoff c when playing against a strategy B player. A strategy B player would receive payoffs b and d when playing against A and B players, respectively;

- the fitness of A and B are defined as

$$f_A := ax_A + bx_B$$
$$f_B := cx_A + dx_B$$

The game dynamics is described by the following non-linear differential equations, that are called the *replicator equations* [3]:

$$\dot{x}_A = x_A(f_A - \phi) \tag{3.2}$$
$$\dot{x}_B = x_B(f_B - \phi) \tag{3.3}$$

where $\phi := f_A x_A + f_B x_B$ is the *average fitness* of the population. The replicator equations (3.2) and (3.3) describe a deterministic selection process, where $f_k - \phi$ is the per capita rate of growth for strategy $k \in \{A, B\}$. Moreover, assumption (3.1) implies that

$$\dot{x}_A = x_A(1 - x_A)(f_A - f_B)$$

and

$$f_A - f_B = (a - c)x_A + (b - d)(1 - x_A).$$

That is, the game behaviour is described by the following differential equation

$$\dot{x} = x(x - 1)(\alpha x - \beta) \tag{3.4}$$

where $x = x_A$ and $\begin{cases} \alpha = b - d + c - a \\ \beta = b - d \end{cases}$.

Therefore, the equilibrium solutions of (3.4) are $x = 0$, $x = \frac{\beta}{\alpha}$ (with $0 \leq \frac{\beta}{\alpha} \leq 1$) and $x = 1$ and the following generic outcome arise:

i. *A dominates B.* If $a > c$ and $b > d$, then the entire population will eventually consist of A phenotypes. The only stable equilibrium is $x = 1$. A is a strict Nash equilibrium, i.e. an *Evolutionarily Stable Strategy* (ESS), while B is not.

ii. *B dominates A.* If $a < c$ and $b < d$, then the situation is similar to case *i.* with A and B exchanged.

iii. *A and B coexist in stable equilibrium.* If $a < c$ and $b > d$, then the interior equilibrium $x = \frac{\beta}{\alpha}$ is stable. Neither A nor B is a Nash equilibrium. This is the strategy behaviour of the aforementioned *Hawk-Dove game.*

iv. *A and B are bi-stable.* If $a > c$ and $b < d$, the equilibrium point in the interior where $x = \frac{\beta}{\alpha}$ is unstable, and the two boundary points where $x = 0$ or $x = 1$ are attracting.

Moreover, if $a = c$ and $b = d$, then $f_A = f_B$ for all frequencies. In this singular case, the two strategies are equivalent. The frequency distribution remains the same for all generations.

3.1.3 Stochastic Models

ODEs models assume that the observed dynamics are driven exclusively by internal, deterministic mechanisms. However, real biological

> systems will always be exposed to influences that are not completely understood or not feasible to model explicitly. Ignoring these phenomena in the modelling may affect the analysis of the studied biological systems. Therefore there is an increasing need to extend the deterministic models to models that embrace more complex variations in the dynamics. A way of modelling these elements is by including stochastic influences or noise. [20]

In a complex ecosystem the interplay among different species (i.e. the interaction network) is the key feature of its dynamics [58]. To model such large families of interactive unites (cells) evolving trough noisy dynamics the well-suited choice for the theoretical framework is the IPS setting (see Section 2.4). IPS simulate populations, or systems of populations, as being composed of discrete individual organisms. Each individual has a set of state variables and behaviours. State variables can include spatial location, physiological and behavioural traits. These attributes vary among the individuals and can change through time. Each individual is assumed to undergo behaviours with rates changing in time. Behaviours can include growth, reproduction, habitat selection, foraging, and dispersal[19]. Some example are branching and Moran processes.

Branching and Moran Processes

The branching processes are largely used to model the stochastic growth of cell populations [5, 22, 27, 35]. In particular, they are powerful tools to realistically describe cancer growth and mutation accumulation, taking into account the demographic and environmental stochasticities, typical of the evolutionary process of cancer. Indeed, for example, the effect of environmental factors cannot be underestimate (it is stressed in Section 1.2 that cancer evolves by continuous interactions with its microenvironment); thus it should be included in the dynamical description [27]. Therefore, within this stochastic context, the branching processes clarify interaction mechanisms among cancer cells and healthy cells. In a mathematical framework, branching processes are Markov chains based on the hypothesis that cellular events (replication, mutation and death) do not influence each other [53]. Namely, each cell events occur at given rates, without depending by population size or composition, or by point in time [4]. Accordingly, at any time, each cell is fully described by cell-intrinsic proliferation, mutation and death rates. Let $n_k(t)$ the number of cells harbouring k mutation, whose birth and death rates per cell are λ_k and μ_k, for $k \in \{1, 2\}$ and $\forall\ t > 0$. The mutation rate from the first type (one mutation) to the second type (two mutations) is u. By assumption, any individual cell of the population at any time t produces a random number of offspring at a later time $t+h$ (where $h = dt$). To simplify the notation, let $n(t) := \big(n_1(t), n_2(t)\big)$,

then the transition probabilities of those two cell types after an infinitesimal time interval h are:

$$\begin{aligned}
P\big(n(t+h)=(i-1,j)\big|n(t)=(i,j)\big) &= \mu_1 ih + o(h)\\
P\big(n(t+h)=(i+1,j)\big|n(t)=(i,j)\big) &= \lambda_1(1-u)ih + o(h)\\
P\big(n(t+h)=(i,j-1)\big|n(t)=(i,j)\big) &= \mu_2 jh + o(h)\\
P\big(n(t+h)=(i,j+1)\big|n(t)=(i,j)\big) &= (\lambda_2+\lambda_1 u)jh + o(h)
\end{aligned}$$

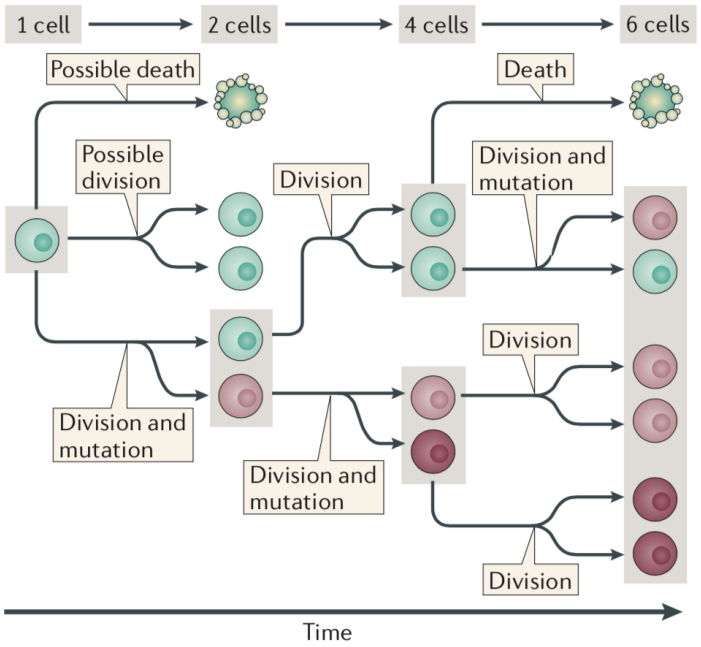

Figure 3.2: Branching process: realization of three time steps [4].

The Moran process has a similar dynamics, but the average population size is constant in time; let it be $N \in \mathbb{N}$;

- let $n \in \mathbb{N}$ be the number of different types of individual,
- let N_k represent the number of individuals of type $k \in \{1, 2, \ldots, n\}$;
- any type k has fitness value f_k, (may be true that $f_k = f_h$ with $k \neq h$).

During each time step an individual of type k is chosen to reproduce with a probability proportional to f_k, and subsequently, a random individual is chosen to die. As a consequence, the number of k-type individuals increases and the amount of l-type individuals decreases according to the following probability

$$P\big(N_k(t+dt) = i+1, N_l(t+dt) = j-1 \big| N_k(t) = i, N_l(t) = j\big) \;=\; \frac{if_k}{if_k + jf_l}\frac{j}{N} + o(dt)$$

This process can be theoretically enlarged to consider random mutations, nonrandom death proportional to weakness (inverse fitness) or time-dependent fitness as well [4].

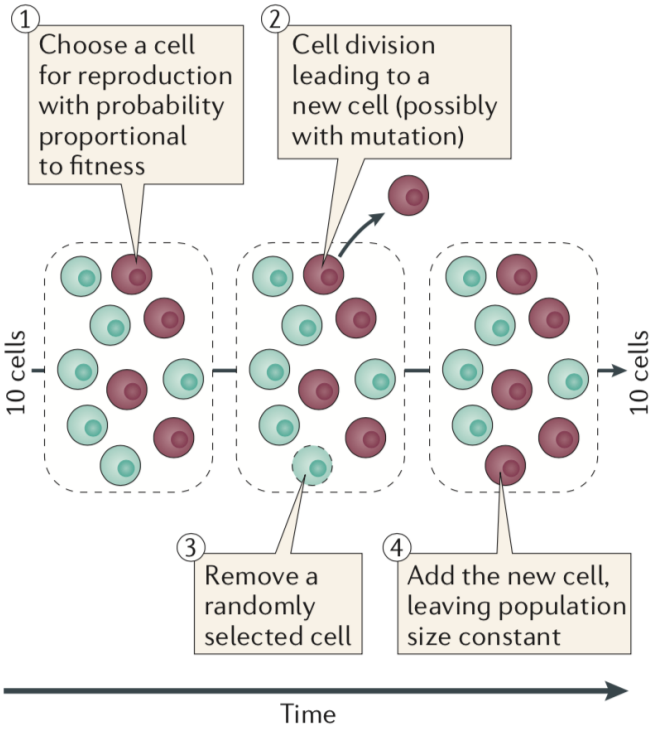

Figure 3.3: Moran process: realization of one time step [4].

Branching and Moran processes have been used to analyse the accumulation of passenger mutations and driver mutations during tumour growth. The key point for a possible quantitative interpretation is that selectively neutral passenger mutations may also arise in healthy tissues. Thus, half or more of the somatic mutations found in a tumour may arise before cancer initiation. Consequently, this approach leads to the "bad luck" tumorigenesis (see Section 1.3), i.e. it suggests that cancer initialization is predominantly the result of error accumulation during stochastic stem cell divisions [4].

3.2 Description of the Model

In this Section, the stochastic model for immunotherapy of cancer, proposed in [7], is investigate. In last decades, immunotherapy emerged among therapy approaches against tumour in both medical and mathematical field. The main issue that can arise during any kind of therapy (e.g. chemo-, radio-, immunotherapy) is resistance; despite an initial efficacy of the therapy, very often a relapse occurs. Heterogeneity, both in genotype and phenotype, which can increase during therapy (as defensive response first and counter-attack later), is the primary feature of tumours that enhances resistance (see [7] and references therein). The chosen quantitative mathematical model described is inspired by experiments of Landsberg *et al.* [41], which test Adoptive Cell Transfer (ACT) therapy on genetically engineered mice melanoma;

- melanoma is a tumour involves epidermal cells,
- ACT therapy is a type of immunotherapy that uses specially injections of cytotoxic T-cells [41].

Cancer under therapy essentially can be viewed as a large family of different type of interactive cells undergoing noisy dynamics, thus choosing an IPS setting gives advantage in the construction of a model which fits to the considered phenomena. Indeed the proposed stochastic model comes from IPS family. In particular, it is an extension of VM (see Section 2.4), which takes into account interactions among species; the main limit of VM (in its standard version) is that it does not allow for interaction between different types of particles. In the context of melanoma under T-cell treatment, the main types of interacting particles are:

- skin cancer cells,
- T-cells,
- TNF-α (*Tumour Necrosis Factor*), a particular type of cytokine.

Due to immune response, other lymphocytes and cytokines are also present. However, their influence can be neglected in the context of the considered phenomena; this result is confirmed by control experiments [7]. Before starting an rigorous mathematical description of the model under consideration, it is necessary clarify the principal steps of biological mechanism of melanoma evolution under T-cell therapy [41]:

- the injected T-cells, recognizing the melanocyte-specific antigens, are able to destroy melanoma cells with differentiated phenotype;

- due to T-cell attack against cancer cells, the involved microenvironment undergoes inflammation (in response to T-cell-driven inflammatory stimuli).
- microenvironment's inflammation favours cancer cells phenotypic switch; their phenotypes passing from a differentiated types to dedifferentiated;
- T-cells are poorly capable to recognize dedifferentiated cancer cells, consequently, melanocytic antigens are down-regulated in the dedifferentiated type;
- cancer develop therapy resistance.

That is, T-cell therapy achieves of remissions in patients with metastatic melanomas, but tumours frequently relapse. Relapses occurs because melanomas acquire ACT resistance. Furthermore, a generic state of the tumour has to be considered as a mixture of differentiated and dedifferentiated cells [7]. Moreover, it is important to mark that

- T-cell dysfunction is improved by the presence of pro-inflammatory cytokine (TNF)-α, that intensifies the phenotype switching;
- phenotype switching is reversible (namely cells can pass from differentiated to dedifferentiated phenotype and vice versa), it does not require cell divisions, and is not induced by a mutation (i.e. it is a purely phenotypic change and does not involve genotype).

Finally, the advice that the authors of the article [41] give, to improve the ATC protocol, is to inject two types of T-cells in order to simultaneously target melanocytic and non-melanocytic antigens to ensure broad recognition of both differentiated and dedifferentiated melanoma cells.

3.2.1 Mathematical Description of the Model

Let the finite space $\mathcal{X} := \mathcal{G} \times \mathcal{P} \dot{\cup} \mathcal{Z} \dot{\cup} \mathcal{W}$ be the local state space. Any local state $x \in \mathcal{X}$ is called trait, hence $\mathcal{X}$ is named trait space. In detail:

- the pair

  $$(g, p) \in \mathcal{G} \times \mathcal{P} := \{g_1, \dots, g_{|\mathcal{G}|}\} \times \{p_1, \dots, p_{|\mathcal{P}|}\}, \qquad |\mathcal{G}|, |\mathcal{P}| < \infty,$$

  represents a cancer cell with genotype g and phenotype p. Every cancer cell can reproduce by cell division, with or without mutation (namely, the two or more cells resulted by division can have or not the same genotype of the parent cell) and it can switch its phenotype; by assumption, the switched phenotype is inherited from any daughter cell. Eventually, a cancer cell may die due to multiple factors: ageing, competition, therapy;

- the trait
$$z \in \mathcal{Z} = \{z_1, \dots, z_{|\mathcal{Z}|}\}, \qquad |\mathcal{Z}| < \infty,$$
denotes a T-cell of type z. T-cells can reproduce (by division), die and produce cytokines;

- the trait
$$w \in \mathcal{W} = \{w_1, \dots, w_{|\mathcal{W}|}\}, \qquad |\mathcal{W}| < \infty,$$
expresses a cytokine of type w. Cytokines play the role of influencing phenotype switching of cancer cells and they can melt.

Figure 3.4 clarifies how the three types of presented particles (cancer cells, T-cell, cytokine) behave and interact. Particularly, it depicts the simplified dynamics of the IPS (without mutation) modelling the experiment described in [41]: x denotes differentiated melanoma cells, y dedifferentiated melanoma cells, z_x T-cells and w TNF$-\alpha$. At each arrow the rate for occurrence of the corresponding event is indicated (e.g. birth is illustrated with two arrowheads and death with an arrow directed to †).

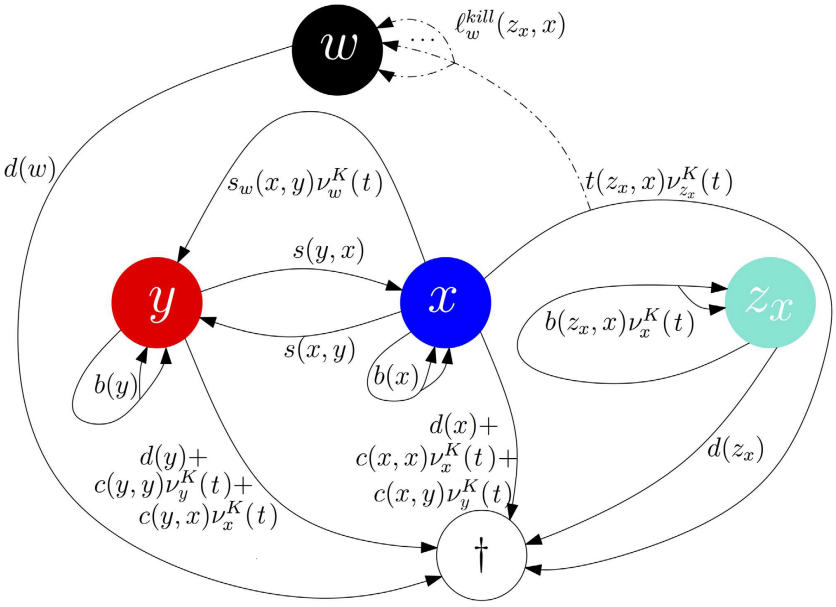

Figure 3.4: Representation of the transitions for a population with trait space $\mathcal{X} = \{x = (g,p); y = (g,p')\} \dot\cup \{z_x\} \dot\cup \{w\}$ [7].

To go forward in developing the model, let $t > 0$ be a fixed time, a population at time t is represented by the following measure acting on the trait space $\mathcal{X}$:

$$\nu^K(t) := \frac{1}{K} \sum_{x \in \mathcal{X}} \nu_x(t) \delta_x,$$

where

- $K \in \mathbb{N}$ is a fixed parameter that allows to scale the population size and is usually called *carrying-capacity* of the environment;

- $\nu_x(t)$ is the number of individuals of type x at time t;

- δ_x denotes the Dirac measure at x.

Hence, for each trait $\tilde{x} \in \mathcal{X}$,

$$\nu^K_{\tilde{x}}(t) := \frac{1}{K} \sum_{x \in \mathcal{X}} \nu_x(t) \delta_x(\tilde{x}) = \frac{1}{K} \nu_{\tilde{x}}(t)$$

indicates the numberer of individuals of type $\tilde{x}$ rescaled by K. One can note that for small trait space the entire population involved in the process at time t can be also viewed as a vector; for example, considering the trait space $\mathcal{X}$ described in figure (3.4), the process at time $t > 0$ can be rewrite as fallows

$$(\nu^K(t)) = (\nu^K_x(t), \nu^K_y(t), \nu^K_{z_x}(t), \nu^K_w(t)).$$

For any arbitrarily fixed $K \in \mathbb{N}$, the dynamics of the $\mathcal{M}^K(\mathcal{X})-$valued continuous-time MC $(\nu^K(t))_{t \geq 0}$ depicts the time evolution of the population under consideration, where

$$\mathcal{M}^K(\mathcal{X}) := \left\{ \frac{1}{K} \sum_{i=1}^{n} \delta_{x_i} \ : \ n \in \mathbb{N}, \ x_1, x_2, \ldots, x_n \in \mathcal{X} \right\},$$

is the set of finite counting measures on $\mathcal{X}$ rescaled by K. Let $\nu^K(0) \in \mathcal{M}(\mathcal{X})$ be a fixed initial state of the process, which represents the beginning population, then low of $(\nu^K(t))_{t \geq 0}$ is described by the following infinitesimal generator [7]:

$$
\begin{aligned}
\mathcal{L}^K\phi(\nu^K) \;=\; & \sum_{(g,p)\in\mathcal{G}\times\mathcal{P}} \left(\phi\left(\nu^K + \frac{\delta_{(g,p)}}{K}\right) - \phi(\nu^K)\right) \\
& \times\, (1-\mu_g)\left\lfloor b(p) - \sum_{\tilde{p}\in\mathcal{P}} c_b(p,\tilde{p})\nu^K_{\tilde{p}}\right\rfloor_+ K\nu^K_{(g,p)} \\
+ & \sum_{(g,p)\in\mathcal{G}\times\mathcal{P}} \left(\phi\left(\nu^K - \frac{\delta_{(g,p)}}{K}\right) - \phi(\nu^K)\right) \\
& \times \left(d(p) + \sum_{\tilde{p}\in\mathcal{P}} c(p,\tilde{p})\nu^K_{\tilde{p}} + \left\lfloor b(p) - \sum_{\tilde{p}\in\mathcal{P}} c_b(p,\tilde{p})\nu^K_{\tilde{p}}\right\rfloor_-\right) K\nu^K_{(g,p)} \\
+ & \sum_{(g,p)\in\mathcal{G}\times\mathcal{P}}\sum_{z\in\mathcal{Z}} \left(\phi\left(\nu^K - \frac{\delta_{(g,p)}}{K} + \sum_{w\in\mathcal{W}} l^{kill}_w(z,p)\frac{\delta_w}{K}\right) - \phi(\nu^K)\right) \\
& \times\, t(z,p)\eta(z)K\nu^K_{(g,p)} \\
+ & \sum_{(g,p)\in\mathcal{G}\times\mathcal{P}}\sum_{\tilde{p}\in\mathcal{P}} \left(\phi\left(\nu^K + \frac{\delta_{(g,\tilde{p})}}{K} - \frac{\delta_{(g,p)}}{K}\right) - \phi(\nu^K)\right) \\
& \times \left(s^g(p,\tilde{p}) + \sum_{w\in\mathcal{W}} s^g_w(p,\tilde{p})\nu^K_w\right) K\nu^K_{(g,p)} \\
+ & \sum_{z\in\mathcal{Z}}\sum_{p\in\mathcal{P}} \left(\phi\left(\nu^K + \frac{\delta_z}{K} + \sum_{w\in\mathcal{W}} l^{prod}_w(z,p)\frac{\delta_w}{K} - \phi(\nu^K)\right) b(z,p)\nu^K_p K\nu^K_z \right. \\
+ & \sum_{z\in\mathcal{Z}} \left(\phi\left(\nu^K + \frac{\delta_z}{K}\right) - \phi(\nu^K)\right) b(z)K\nu^K_z \\
+ & \sum_{z\in\mathcal{Z}} \left(\phi\left(\nu^K - \frac{\delta_z}{K}\right) - \phi(\nu^K)\right) d(z)K\nu^K_z \\
+ & \sum_{w\in\mathcal{W}} \left(\phi\left(\nu^K - \frac{\delta_w}{K}\right) - \phi(\nu^K)\right) d(w)K\nu^K_w \\
+ & \sum_{(\tilde{g},\tilde{p})\in\mathcal{G}\times\mathcal{P}}\sum_{(g,p)\in\mathcal{G}\times\mathcal{P}} \left(\phi\left(\nu^K + \frac{\delta_{(\tilde{g},\tilde{p})}}{K}\right) - \phi(\nu^K)\right) \\
& \times\, \mu_g m((g,p)(\tilde{g},\tilde{p}))\left\lfloor b(p) - \sum_{p'\in\mathcal{P}} c_b(p,p')\nu^K_{p'}\right\rfloor_+ K\nu^K_{(g,p)}
\end{aligned}
\tag{3.5}
$$

where $\lfloor\star\rfloor_-$ and $\lfloor\star\rfloor_+$ are the negative and the positive parts and, by assumption, the generator $\mathcal{L}^K$ acts on real valued bounded measurable functions ϕ, for any

arbitrary state $\nu^K = \nu^K(t)$ of the process at time t. Moreover in (3.5) it is used the following short hand notation:

$$\nu_p^K \equiv \sum_{g\in\mathcal{G}} \nu_{(g,p)}^K.$$

The infinitesimal generator contains all the informations about events that cancer cells, T-cells and cytokines undergo. Accordingly, each present individual at any time t has several exponential clocks with intensities depending on its trait $x \in \mathcal{X}$ and on the current state of the system $\nu^K(t)$ [7]:

(i) Each present cancer cell of trait $(g,p) \in \mathcal{G}\times\mathcal{P}$ has

- a *clonal reproduction* clock with rate

$$(1-\mu_g)\left\lfloor b(p) - \sum_{\tilde{p}\in\mathcal{P}} c_b(p,\tilde{p})\nu_{\tilde{p}}^K(t)\right\rfloor_+$$

and a *mutant reproduction* clock with rate

$$\mu_g\left\lfloor b(p) - \sum_{p'\in\mathcal{P}} c_b(p,p')\nu_{p'}^K(t)\right\rfloor_+,$$

where $\mu_g \in [0,1]$ denotes the probability that a mutation occurs in a birth event of a cancer cell with genotype g (thus $1-\mu_g$ is the probability that a new cancer cell born without genotypic mutation), $b(p) \in \mathbb{R}_{\geq 0}$ is the *natural birth rate* and $K^{-1}c_b(p,\tilde{p}) \in \mathbb{R}_{\geq 0}$ is the *birth-reducing competition kernel* which lowers the birth rate of a cancer cell of phenotype p in presence of a $\tilde{p}$-type cancer cell. Furthermore, if the total birth rate $\sum_{p\in\mathcal{P}} b(p) = 0$ the birth-reducing competition $K^{-1}c_b(p,\tilde{p})$ as an additional death rate of the existing p-type cancer cells;

- a *natural mortality* clock with rate

$$d(p) + \sum_{\tilde{p}\in\mathcal{P}} c(p,\tilde{p})\nu_{\tilde{p}}^K(t) + \left\lfloor b(p) - \sum_{\tilde{p}\in\mathcal{P}} c_b(p,\tilde{p})\nu_{\tilde{p}}^K(t)\right\rfloor_-,$$

where $d(p) \in \mathbb{R}_{\geq 0}$ is the *natural death rate*, $K^{-1}c(p,\tilde{p}) \in \mathbb{R}_{\geq 0}$ is another *competition kernel* which increases the death rate, $b(p)$ and $K^{-1}c_b(p,\tilde{p})$ as above;

- a *therapy mortality* clock with rate

$$\sum_{z\in\mathcal{Z}} t(z,p)\nu_z^K(t),$$

where $K^{-1}t(z,p) \in \mathbb{R}_{\geq 0}$ is a therapy kernel that represent an additional cancer death rate for p-type cell due to the presence of z-type T-cell;

- a *natural and cytokine-induced switch* clock with rate

$$\sum_{\tilde{p}\in\mathcal{P}}\left(s^g(p,\tilde{p})+\sum_{w\in\mathcal{W}}s^g_w(p,\tilde{p})\nu^K_w(t)\right),$$

where $s^g(p,\tilde{p})\in\mathbb{R}_{\geq 0}$ and $s^g_w(p,\tilde{p})\nu^K_w(t)\in\mathbb{R}_{\geq 0}$ are the *natural and cytokine-induced switch kernels* respectively; they both switch a cancer cell from (g,p)-type to $(g,\tilde{p})$-type.

Any times that one of these clocks rings, any cancer cell of trait (g,p) undergoes corresponding event; whenever

- clonal reproduction clock rings, an additional cancer cell of the same trait (g,p) appears;
- mutant reproduction clock rings, a cancer cell of trait $(\tilde{g},\tilde{p})$ appears according to the kernel $m\big((g,p),(\tilde{g},\tilde{p})\big)\in[0,1]$, which represents the probability that, whenever a mutation event occurs in cancer cell, the (g,p)-type cell generates a $(\tilde{g},\tilde{p})$-type cell. Moreover, $m\big((g,p),(\tilde{g},\tilde{p})\big)$ acts as a probability matrix, indeed by definition it is assumed that $\sum_{(\tilde{g},\tilde{p})}m\big((g,p),(\tilde{g},\tilde{p})\big)=1$ and $m\big((g,p),(g,p)\big)=0$
- natural mortality clock rings, the cancer cell disappears;
- therapy mortality clock rings, the cancer cell disappears and an amount of $l^{kill}_w(z,p)\in\mathbb{N}_0$ cytokines of type w are deterministically produced (according to weights $t(z,p)\nu^K_z(t)$) at each cancer killing event;
- natural and cytokine-induced switch clock rings, this cancer cell disappears and a new cancer cell of trait $(g,\tilde{p})$ appears according to the weights $s^g(p,\tilde{p})+\sum_{w\in\mathcal{W}}s^g_w(p,\tilde{p})\nu^K_w(t)(w)$.

(ii) Each present T-cell of trait $z\in\mathcal{Z}$ has

- a *natural birth* clock with rate $b(z)\in\mathbb{R}_{\geq 0}$
- a *natural mortality* clock with rate $d(z)\in\mathbb{R}_{\geq 0}$
- a *reproduction* clock with rate

$$\sum_{p\in\mathcal{P}}b(z,p)\nu^K_p(t)$$

Thus, any time that natural birth and mortality clocks ring, one T-cell of trait z sprigs up and disappears, respectively. Whenever the reproduction clock of a T-cell rings, due the presence of a cancer phenotype p, an additional T-cell appears with the same trait z with rate $K^{-1}b(z,p)\in\mathbb{R}_{\geq 0}$, and $l^{prod}_w(z,p)\in\mathbb{N}$ cytokines of type w arise according to the weights $b(z,p)\nu^K_p(t)$.

(iii) Each present cytokine $w \in \mathcal{W}$ has

- a *mortality* clock with rate $d(w)$.

Moreover, as highlighted in point (i), an amount of $l_w^{kill}(z,p) \in \mathbb{N}_0$ of trait w are produced every time a T-cell of trait z kills a cancer cell of phenotype p, and a number of $l_w^{prod}(z,p)$ of cytokines of the same trait are generated every time a T-cell of trait z is produced in the presence of a cancer cell of phenotype p.

Thus, taking in account these jump rates, the evolution of population can be approximate by the solution of the following quadratic dynamics system

$$\begin{aligned}
\dot{\mathfrak{n}}_{(g,p)} &= \mathfrak{n}_{(g,p)}\Bigg((1-\mu_g)\Big\lfloor b(p) - \sum_{(\tilde{g},\tilde{p})\in\mathcal{G}\times\mathcal{P}} c_b(p,\tilde{p})\mathfrak{n}_{(\tilde{g},\tilde{p})}\Big\rfloor_+ \\
&\quad - \Big\lfloor b(p) - \sum_{(\tilde{g},\tilde{p})\in\mathcal{G}\times\mathcal{P}} c_b(p,\tilde{p})\mathfrak{n}_(\tilde{g},\tilde{p})\Big\rfloor_- - d(p) - \sum_{(\tilde{g},\tilde{p})\in\mathcal{G}\times\mathcal{P}} c(p,\tilde{p})\mathfrak{n}_{(\tilde{g},\tilde{p})} \\
&\quad - \sum_{z\in\mathcal{Z}} t(z,p)\mathfrak{n}_z - \sum_{\substack{\tilde{p}\in\mathcal{P}\\ \tilde{p}\neq p}}\Big(s^g(p,\tilde{p}) + \sum_{w\in\mathcal{W}} s_w^g(p,\tilde{p})\mathfrak{n}_w\Big)\Bigg) \\
&\quad + \sum_{\tilde{p}\in\mathcal{P}} \mathfrak{n}_{(\tilde{g},\tilde{p})}\Big(s^g(p,\tilde{p}) + \sum_{w\in\mathcal{W}} s_w^g(p,\tilde{p})\mathfrak{n}_w\Big) \\
&\quad + \sum_{(\tilde{g},\tilde{p})\in\mathcal{G}\times\mathcal{P}} \mathfrak{n}_{(\tilde{g},\tilde{p})}\Big(\mu_{\tilde{g}} m\big((g,p),(\tilde{g},\tilde{p})\big)\Big\lfloor b(p) - \sum_{p'\in\mathcal{P}} c_b(p,p')\mathfrak{n}_{p'}\Big\rfloor_+\Big) \\
\dot{\mathfrak{n}}_z &= \mathfrak{n}_z\Big(b(z) - d(z) + \sum_{(g,p)\in\mathcal{G}\times\mathcal{P}} b(z,p)\mathfrak{n}_{(g,p)}\Big) \\
\dot{\mathfrak{n}}_w &= -\mathfrak{n}_w d(w) + \sum_{(g,p)\in\mathcal{G}\times\mathcal{P}} \mathfrak{n}_{(g,p)} \sum_{z\in\mathcal{Z}} \big(l_w^{kill}(z,p)t(z,p) + l_w^{prod}(z,p)b(z,p)\big)\mathfrak{n}_z
\end{aligned}$$

$$\forall\ (g,p)\in\mathcal{G}\times\mathcal{P},\ z\in\mathcal{Z},\ w\in\mathcal{W}. \tag{3.6}$$

The idea is the same of Theorem 2.18, used to study the mean field limit of the Voter model, but in the case a more general result is used; that is, let $d = |\mathcal{G}||\mathcal{P}| + |\mathcal{Z}| + |\mathcal{W}|$ and $\nu^K(0) \xrightarrow{K\to\infty} \mathfrak{n}(0)$ $(a.s.)$, i.e. it is supposed that the initial conditions converge almost surely to a deterministic limit, then the sequence of rescaled processes $(\nu^K(t))_{t\geq 0}$ converges almost surely as $K \to \infty$ to the d-dimensional deterministic process which is the unique solution $\mathfrak{n}(t)$ to the quadratic

ODEs system (3.6). This result comes from Theorem 2.1 in Chapter 11 of [24]. One of the most important features of these model (like VM) is its double sided nature of stochastic and deterministic model [7]:

- The stochastic side allows the possibility of extinctions (which deterministic side can not predict); e.g. the extinction either of a set of χ-type cells, for a specific trait $\chi \in \mathcal{X}$, or of the whole population (i.e. irreversible organism collapse). Indeed, let K be a fixed carring capacity, due to fluctuations essentially two facts lead to extinction; first, population sizes are typically transient states and when the process jumps in a low minimum size, random fluctuations can lead to extinction (that corresponds to an absorbing state) before the population reaches its equilibrium. Second, when the process is in a finite equilibrium, after a long enough time, the fluctuations around it drive the process to extinction of population (see simulations in Figure 3.6). Moreover, one can note that the extinction of a specific type of cells can strongly influences the long-term behaviour of the entire population; e.g the equilibria points could change, with respect to the remaining cell types. In both cases, the value of K, determining the amplitude of fluctuations, characterizes the probability of extinction. Furthermore, the relevant mutations in the setup of cancer therapy are driver mutations, but they are invisible in the deterministic limit, indeed the mutation probabilities $\mu_g \equiv \mu_g^K \xrightarrow{K\to\infty} 0$.
- The deterministic side describes the logistic part of the process, concerning the necessary and available resources during evolution of the population. Moreover, it highlight prey-predator dynamics, that occur between T-cells and cancer cells, but also hawk-dove dynamics that take place among cancer cells of different phenotype (mutation and switch parts).

To improve the analysis of the model, presented in this section, and to clarify cancer evolution mechanisms, some simplifications are necessary; presence of a large number of switches and mutations makes difficult the analysis of ODE system. Thus, considering few types of cell (i.e. a small trait space $\mathcal{X}$) is a good starting point. Consequently, fitting to this simplification, the modelling of experiment described in the article [41] represents a good example.

3.2.2 Therapy with T-cells of One Specificity

Modelling experiment explained in paper [41] the following setting arise; let

- $G = \{g\}$: all cancer cells have the same genotype g; mutations are not considered in the model, since the experiment does not investigate them. Thus, the probability that during a birth event a mutation occurs is zero, i.e. $\mu_g = 0$;

- $P = \{p_x, p_y\}$: each cancer cell phenotype can be either p_x or p_y; where p_x and p_y are the differentiated and dedifferentiated phenotypes respectively. Any cancer cell can switch its phenotype from p_x to p_y and also in the reverse order (namely $p_x \to p_y$ and $p_y \to p_x$). Additionally, birth-reduction competition are not considered, that is $c_b(\rho, \tilde{\rho}) = 0, \ \forall \rho, \tilde{\rho} \in \mathcal{P}$;

- $Z := \{z_x\}$: there is only one type of T-cell z_x attacking exclusively cells of phenotype p_x; therapy with T-cells of one specificity. Therefore, there is not a type of T-cell attacking phenotype p_y, i.e. $t(z_x, p_y) = 0$, and the presence of cancer cells of phenotype p_y does not stimulate the production of T-cell, $b(z_x, p_y) = 0$ (due to the fact that dedifferentiated phenotype is not recognized by T-cells of type z_x). Moreover, the natural birth rate of the considered T-cells is $b(z_x) = 0$, because in the model are included only injected T-cells, whose reproduction is stimulated only in presence of cancer cells of x-type.

- $W = \{w\}$: the type w of cytokine represents the TNF$-\alpha$ proteins.

Therefore, the considered trait space is $X = \{x := (g, p_x), y := (g, p_y), z_x, w\}$. Since the dimension of this trait space is small, it is useful to think about the process $(\nu(t))_{t\geq 0}$ as a vector as well as the solution of associated ODE system, that is an arbitrary configuration of the process is

$$\eta = (\eta_x, \eta_y, \eta_{z_x}, \eta_w),$$

where, for fixed time $t > 0$ and carrying capacity $K \in \mathbb{N}$,

$$\eta := \nu^K(t) \ \text{ and } \ \eta_\chi = \nu^K_\chi(t), \ \forall \ \chi \in \mathcal{X}$$

and the solution of the ODE system results

$$\mathfrak{n}(t) = \big(\mathfrak{n}_x(t), \mathfrak{n}_y(t), \mathfrak{n}_{z_x}(t), \mathfrak{n}_w(t)\big),$$

Additionally, since the cancer cells are characterized only by phenotype, it is possible to do a little abuse of notation substituting in equation (3.6) p_x and p_y with x and y (respectively). Furthermore, let $s := s^g$ and $s_w := s^g_w$, and consider that $\lfloor f \rfloor_+ - \lfloor f \rfloor_- = f, \quad \forall$ $\mathbb{R}$-valued functions f; the resulting ODE system is

$$
\begin{aligned}
\dot{\mathfrak{n}}_x &= \mathfrak{n}_x\big(b(x) - d(x) - s(x,y) - c(x,x)\mathfrak{n}_x - c(x,y)\mathfrak{n}_y - t(z_x,x)\mathfrak{n}_{z_x} - s_w(x,y)\mathfrak{n}_w\big) \\
&+ \mathfrak{n}_y s(y,x) \\
\dot{\mathfrak{n}}_y &= \mathfrak{n}_x\big(s(x,y) + s_w(x,y)\mathfrak{n}_w\big) + \mathfrak{n}_y\big(b(y) - d(y) - s(y,x) - c(y,x)\mathfrak{n}_x - c(y,y)\mathfrak{n}_y\big) \\
\dot{\mathfrak{n}}_{z_x} &= \mathfrak{n}_{z_x}(-d(z_x) + b(z_x,x)\mathfrak{n}_x) \\
\dot{\mathfrak{n}}_w &= \mathfrak{n}_x\big(l_w^{kill}(z_x,x)t(z_x,x) + l_w^{prod}(z_x,x)b(z_x,x)\big)\mathfrak{n}_{z_x} - \mathfrak{n}_w d(w)
\end{aligned}
\tag{3.7}
$$

3.2.3 Therapy with T-cells of Two Specificities

The phenotypic plasticity of melanoma cells in an inflammatory microenvironment contributes to tumour relapse after initially successful T-cell immunotherapy [41]. On the basis of their work, authors of the article [41] have proposed to improve future ACT protocols with

- T-cell abling to target melanocytic and non-melanocytic antigens to ensure broad recognition of both differentiated and dedifferentiated melanoma cells,
- strategies to sustain T-cell effector functions by blocking immune-inhibitory mechanisms in the tumour microenvironment.

Thus the following example, presented in [7], represents an useful tool to analyse the kind of new protocol proposed in [41]. Let z_y represent that type of T-cell recognizing and attacking dedifferentiated (y-type) cancer cells. The setting for a new model including a therapy with z_y T-cells is the same of model described by ODE system (3.7), if not specified: $\mathcal{X} := \{x, y, z_x, z_y, w\}$ is trait space, z_x and z_y are the two traits characterizing a therapy with T-cells of two specificities. As in therapy with T-cells of one specificity, the natural birth rates of the considered types of T-cell are $b(z_x) = 0$ and $b(z_y) = 0$, because in the model are included only injected T-cell, which production are stimulated only in presence of cancer cells. In addition, a T-cell of specific type interacts only with cancer cells of the corresponding type, i.e.

$$
\begin{aligned}
b(z_x, y) = 0 = b(z_y, x) \\
t(z_x, y) = 0 = t(z_y, x).
\end{aligned}
$$

The ODE system describing the considered model is

$$\begin{aligned}
\dot{\mathfrak{n}}_x &= \mathfrak{n}_x\big(b(x) - d(x) - s(x,y) - c(x,x)\mathfrak{n}_x - c(x,y)\mathfrak{n}_y - t(z_x,x)\mathfrak{n}_{z_x} - s_w(x,y)\mathfrak{n}_w\big) \\
&+ \mathfrak{n}_y s(y,x) \\
\dot{\mathfrak{n}}_y &= \mathfrak{n}_x\big(s(x,y) + s_w(x,y)\mathfrak{n}_w\big) \\
&+ \mathfrak{n}_y\big(b(y) - d(y) - s(y,x) - c(y,x)\mathfrak{n}_x - c(y,y)\mathfrak{n}_y - t(z_y,y)\mathfrak{n}_{z_y}\big) \\
\dot{\mathfrak{n}}_{z_x} &= \mathfrak{n}_{z_x}(-d(z_x) + b(z_x,x)\mathfrak{n}_x) \\
\dot{\mathfrak{n}}_{z_y} &= \mathfrak{n}_{z_y}(-d(z_y) + b(z_y,y)\mathfrak{n}_y) \\
\dot{\mathfrak{n}}_w &= \mathfrak{n}_x\big(l_w^{kill}(z,x)t(z_x,x) + l_w^{prod}(z_x,x)b(z_x,x)\big)\mathfrak{n}_{z_x} \\
&+ \mathfrak{n}_y\big(l_w^{kill}(z_y,x)t(z_y,y) + l_w^{prod}(z_y,y)b(z_y,y)\big)\mathfrak{n}_{z_y} - \mathfrak{n}_w d(w)
\end{aligned} \tag{3.8}$$

3.2.4 Qualitative Analysis and Simulations

To underline the qualitative features of systems (3.7) and (3.8), the choice of parameters, carrying capacity and initial condition is crucial:

- for a qualitative study of the experiment in [41], thus for the simulation of the ODE system (3.7), in [7] were proposed the following parameters:

$$\begin{array}{llll}
b(x) = 3 & b(y) = 3 & & \\
d(x) = 1 & d(y) = 1 & d(z_x) = 3 & d(w) = 15 \\
s(x,y) = 0.1 & s(y,x) = 1 & & s_w(x,y) = 4 \\
c(x,x) = 0.3 & c(y,x) = 0 & b(z_x,x) = 8 & l_w^{prod}(z_x,x) = 0 \\
c(x,y) = 0 & c(y,y) = 0.3 & t(z_x,x) = 28 & l_w^{kill}(z_x,x) = 1
\end{array} \tag{3.9}$$

and

$$\begin{aligned}
K &= 2 \cdot 10^2 \\
\mathfrak{n}(0) &= (2, 0, 5 \cdot 10^{-2}, 0)
\end{aligned} \tag{3.10}$$

are the fixed carrying capacity and initial conditions, respectively;

- for the qualitative analysis of the model for therapy with T-cells of two specificities, described by ODE system (3.8), in addition to parameters (3.9), the following parameters were chosen:

$$b(z_y, y) = 14 \quad l_w^{prod}(z_y, y) = 0$$
$$t(z_y, y) = 28 \quad l_w^{kill}(z_y, y) = 1 \quad d(z_y) = 3 \tag{3.11}$$

  with carrying capacity $K = 2 \cdot 10^2$ and initial conditions

$$\mathfrak{n}(0) = (2, 0, 5 \cdot 10^{-2}, 2 \cdot 10^{-1}, 0) \tag{3.12}$$

Thus, parameters, initial conditions and carrying capacity were chosen in order to highlight the influence of the randomness and the possible behaviours of the systems [7].

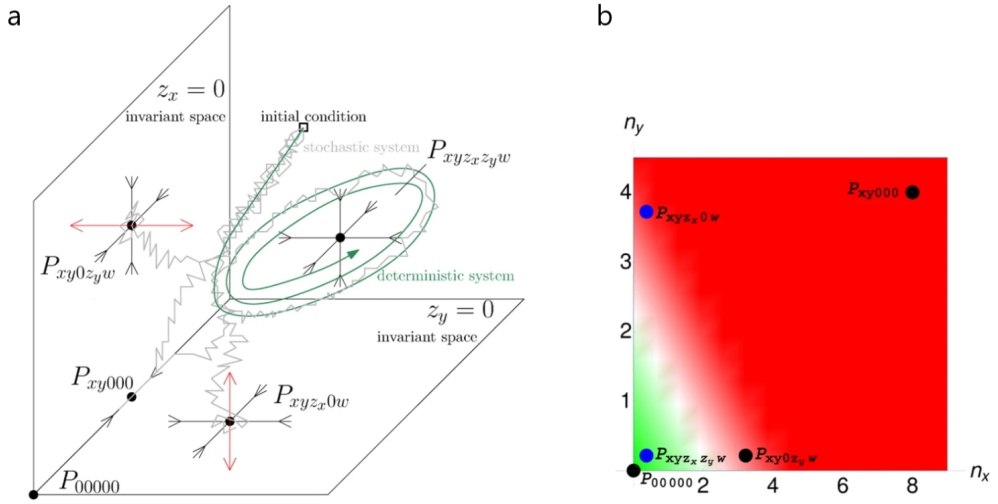

Figure 3.5: (a) Sketch of the invariant subspaces, stability of the fixed points. (b) Projection of the fixed points onto $\mathfrak{n}_x$ and $\mathfrak{n}_y$. The black dots show the fixed points of system (3.7) and the blue points represent the two additional fixed points of system (3.8). The green area contains the fixed points which correspond to cure or remission and the red area contains those describing relapse [7].

An essential step in the analysis of the deterministic systems is the determination of equilibrium points and the study of their stability (an equilibrium point is stable if the eigenvalues of the Jacobian matrix of the system at that point have all strictly negative real parts). The simulations of the stochastic systems depict

several types of behaviour, that can occur with certain probabilities, as Figure 3.6 shows.

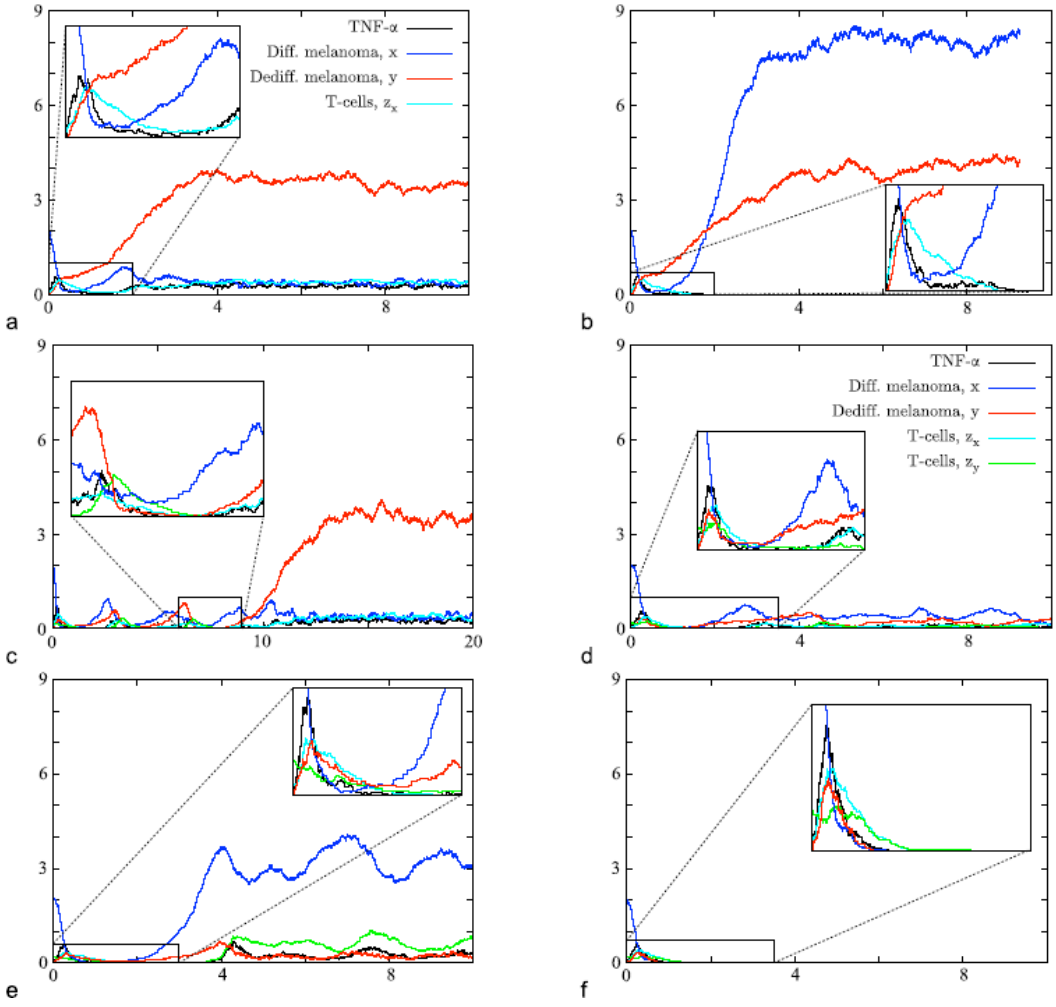

Figure 3.6: The graphs show the number of individuals divided by $K = 200$ (carrying capacity) versus time. Two scenarios are possible for therapy with T-cells of one specificity: (a) T-cells z_x survive and the system is attracted to P_{xyz_xw}, or (b) T-cells z_x die out and the system is attracted to P_{xy00}. Three additional scenarios are possible for therapy with T-cells of two specificities: (d) T-cells z_x and z_y survive and the system stays close to $P_{xyz_xz_yw}$, (e) T-cells z_x die out and the system is attracted to P_{xy0z_yw}, (f) the tumour is eradicated (corresponding to P_{00000}). (c) Transition between cases (a,d) [7].

Therapy with T-cells of One Specificity

The equilibrium points of ODE system (3.7) are the solutions of the following system

$$\begin{cases} 0 &= \mathfrak{n}_x\big(b(x)-d(x)-s(x,y)-c(x,x)\mathfrak{n}_x-c(x,y)\mathfrak{n}_y-t(z_x,x)\mathfrak{n}_{z_x} \\ & \quad -s_w(x,y)\mathfrak{n}_w\big)+\mathfrak{n}_y s(y,x) \\ 0 &= \mathfrak{n}_x\big(s(x,y)+s_w(x,y)\mathfrak{n}_w\big)+\mathfrak{n}_y\big(b(y)-d(y)-s(y,x)-c(y,x)\mathfrak{n}_x \\ & \quad -c(y,y)\mathfrak{n}_y\big) \\ 0 &= \mathfrak{n}_{z_x}(-d(z_x)+b(z_x,x)\mathfrak{n}_x) \\ 0 &= \mathfrak{n}_x\big(l_w^{kill}(z_x,x)t(z_x,x)+l_w^{prod}(z_x,x)b(z_x,x)\big)\mathfrak{n}_{z_x}-\mathfrak{n}_w d(w). \end{cases} \tag{3.13}$$

There are three acceptable solutions for the system (3.13), with parameters (3.9) (see Figure 3.5(b)):

- $P_{0,0,0,0}$: there is not type of cell;
- $P_{x,y,0,0}$: T-cells and TNF-α are absent and both differentiated and dedifferentiate melanoma cells are present;
- $P_{x,y,z_x,w}$: all types of cell are present.

Among them, the only stable point is $P_{x,y,z_x,w}$, whereas $P_{x,y,0,0}$ result stable in absence of T-cell (i.e. in the invariant space $\{\mathfrak{n}_{z_x}=0\}$). Therefore, for initial condition (3.10) the deterministic system (3.13) is attracted to $P_{x,y,z_x,w}$; that is, the four types of cell coexist in a equilibrium state. In general, the deterministic system is attracted to the stable point $P_{x,y,z_x,w}$ for initial conditions such that:

- the number of differentiated melanoma cells is large, i.e. $\mathfrak{n}_x(0)$ is far from zero;
- the number of injected T-cells is small, i.e. $\mathfrak{n}_{z_x}(0)$ is close to zero;
- the numbers of dedifferentiated melanoma cells and TNF-α molecules are small or equal to zero, i.e. $\mathfrak{n}_y(0)$ and $\mathfrak{n}_w(0)$ are close or equal to zero;

From the biological point of view, the large initial amount of x-type cancer cells induces the T-cells of z_x-type to increase, TNF-α is secreted, and the number of differentiated melanoma cells (consequently, $\mathfrak{n}_x$) shrinks due to killing and TNF-α induced switching, whereas the population of dedifferentiated melanoma cells,

$\mathfrak{n}_y$, grows. Furthermore, one can note that the parameters were chosen in other to respects the biological features of the system: the minimum amount of T-cells during remission is low, and such that the equilibrium value of melanoma of type x (differentiated cancer cell) in presence of T-cells is low, whereas equilibrium values of both melanoma types in absence of T-cells are high [7].

For the stochastic system, two types of behaviour can occur with certain probabilities:

- the trajectory stays close to that of the deterministic system and, finally, the system reaches a neighbourhood of the fixed point P_{xyz_xw} (see Figure 3.6(a));
- the system reaches a neighbourhood of P_{xy00}, due to the death of T-cell population, $\nu_{z_x}^K$ (see Figure 3.6(b)).

Biologically, after the extinction of the T-cells ($\nu_{z_x}^K = 0$), the number of cytokines, ν_w^K, gradually decreases until the extinction of the entire TNF-α population. Thus, the population of differentiated melanoma cells, ν_x^K, favoured by the lack of lymphocytic attack, increases. Moreover, without TNF-α stimulation, the switch from x to y stops and a relapse, which mainly consists of differentiated cells, occurs. Again, it is fundamental to remark the importance of parameters choice (in particular switching, therapy or cross-competition), which causes a variety of different possible behaviour [7].

Therapy with T-cells of Two Specificities

By construction, the model for therapy with T-cells of two specificities, defined by the ODE system (3.8), contains one more predator-prey term $t(z_y, y)$ between dedifferentiated melanoma cells (of type y) and the specific T-cell z_y, witch recognize and attacking them. This adds two new equilibrium points (see the blue dots on Figure 3.5(b)) beside those of ODE system (3.7):

- $P_{xyz_xz_yw}$: coexistence of all types population,
- P_{xy0z_yw}: absence of the z_x-type of T-cell population.

The new stable point is $P_{xyz_xz_yw}$ and, in this setting, the invariant subspaces are $\{\mathfrak{n}_{z_x} = 0\}$, $\{\mathfrak{n}_{z_y} = 0\}$ and $\{\mathfrak{n}_{z_x} = 0\} \cap \{\mathfrak{n}_{z_y} = 0\}$, in which P_{xy0z_yw}, P_{xyz_x0w} and P_{xy000} are stable, respectively. One can note that P_{xyz_x0w}, P_{xy000} and P_{00000} correspond to P_{xyz_x0w}, P_{xy00} and P_{0000} (fixed points of ODE system (3.7)) respectively, but P_{xyz_x0w} loses stability in the enlarged space and also P_{xy000} does not preserve stability in the hyperplane $\{\mathfrak{n}_{z_x} = 0\}$; indeed they result stable only considering the intersection with the hyperplane $\{\mathfrak{n}_{z_y} = 0\}$, that is, during therapy there is

not injection of T-cell attacking dedifferentiated (y-type) melanoma cells. Indeed, therapy with T-cells of one specificity model results a particular case of the model for therapy with T-cells of two specificities such that the initial amount of the specific T-cells for y-type cancer cells is zero, $\mathfrak{n}_{z_y}(0) = 0$, and null birth and death rates, i.e. $b(z_y, y) = 0$ and $d(z_y) = 0$. Whereas, considering as initial conditions a small amount of T-cell of types z_y and z_x ($\mathfrak{n}_{z_y}(0) > 0$ and $\mathfrak{n}_{z_x}(0) > 0$), a large quantity of melanoma cells $\mathfrak{n}_x(0)$ and small or null amount of dedifferentiated melanoma cells ($\mathfrak{n}_y(0)$), the deterministic system is attracted to the stable equilibrium point $P_{xyz_xz_yw}$ (see simulations in Figure 3.7(B)). Biologically, the presence of target x induces the growth of T-cell population $\mathfrak{n}_{z_x}$, TNF-α is secreted, and the differentiated melanoma population $\mathfrak{n}_x$ shrinks due to killing and switching, the population of dedifferentiated melanoma $\mathfrak{n}_y$ grows. The increasing both differentiated and dedifferentiated melanoma populations is regulated and kept at a low level by the presence of the T-cells of the two specificities z_x and z_y [7]. In particular, the system (3.8), with parameters (3.9) and (3.11), has different solutions with to respect the particular choice of the initial condition; for example:

a) $\mathfrak{n}(0) = (2, 0, 5 \cdot 10^{-2}, 0, 0)$, the system is attracted to the equilibrium point P_{xyz_x0w} (see Figure 3.7(A));

b) $\mathfrak{n}(0) = (2, 0, 5 \cdot 10^{-2}, 2 \cdot 10^{-1}, 0)$, the system is attracted to the fixed point P_{xyz_x0w} (see Figure 3.7(B)). This show that P_{xyz_x0w} is not stable (i.e. with initial condition $\mathfrak{n}_{z_y} > 0$, chosen in a neighbourhood of the zero, the whole system go far from P_{xyz_x0w}, as the time t increases);

c) $\mathfrak{n}(0) = (2, 0, 0, 2 \cdot 10^{-1}, 0)$, the system is attracted to the fixed point P_{xy0z_yw} (see Figure 3.7(C)), which is not stable by b);

d) $\mathfrak{n}(0) = (2, 0, 0, 0, 0)$, the system is attracted to the fixed point P_{xy000} (see Figure 3.7(D)), that is not stable too.

Moreover, graphs (A) and (B) in Figure 3.7 depict also the solution of ODE system (3.7) with parameters (3.9) and initial conditions $\mathfrak{n}(0) = (2, 0, 5 \cdot 10^{-2}, 0)$ and $\mathfrak{n}(0) = (2, 0, 0, 0)$, respectively.

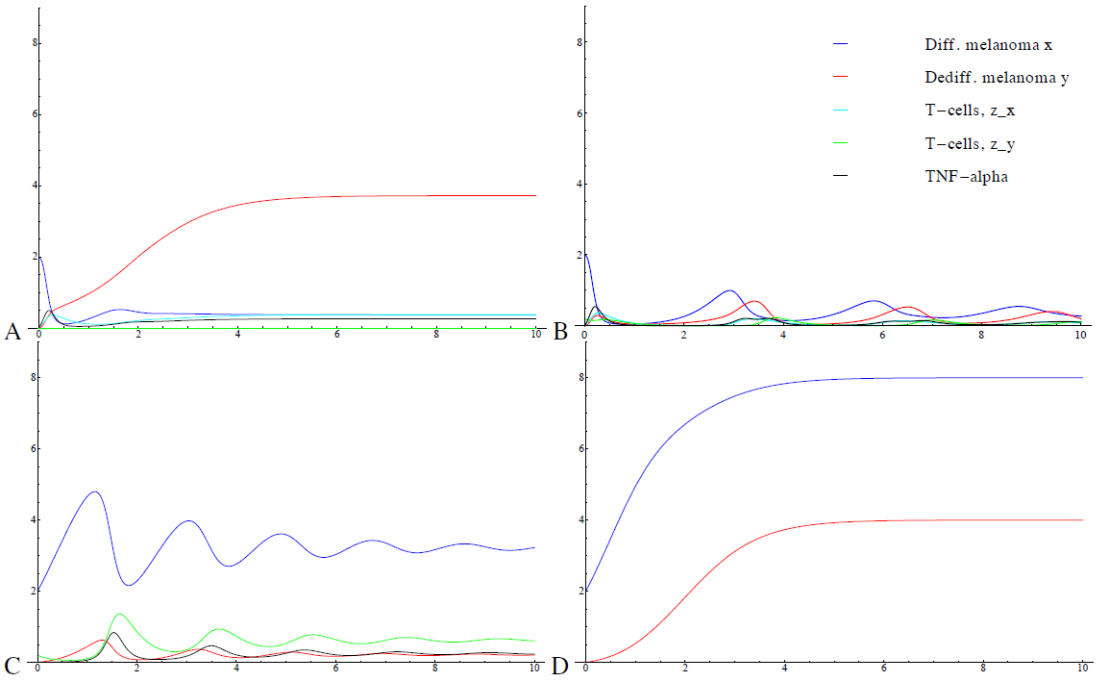

Figure 3.7: Solutions of the deterministic system (3.8) with numerical parameters (3.9) and (3.11) and different choices of initial conditions [7].

Analysing the simulation of the stochastic model for protocol therapy with two types of T-cell (see Figure 3.6), five main different scenarios arise:

- T-cells of type z_y die out (a);
- the extinction of the two types (z_x and z_y) of T-cell population; i.e. the melanoma can grow until exhaustion of the host body's resources (that is until the collapse of the organism) (b);
- all populations can survive for some time fluctuating around their joint equilibrium (d);
- T-cells of type z_x vanish (e);
- the extinction of both differentiated and dedifferentiated (types x and y) melanoma cells, due to the simultaneous attack of the two different types of T-cell, that corresponds to a cure, i.e. the tumour is eradicated. Moreover, without the melanoma targets (x and y) the stimuli for reproduction of T-cells and TNF-α stop, therefore, also the lymphocyte and the cytokine populations vanish (f).

Furthermore, in figure 3.6(c), a transition from case (d) to case (a) is highlight, but others transitions between the different scenarios can be observed; for example, system could pass from case (d) to (e) and then to (b).

In this setting (therapy protocol with two types of T-cell), authors of paper [7] have chosen parameters such that the minima of the two types of T-cells during remission is low, so that they have a large enough probability to die out in the stochastic process, in line with the real biological behaviour of the system. Additionally, some parameters were estimated from experimental data (provided by article [41]). Generally, at the beginning of therapy there is a null or small quantity of dedifferentiated melanoma cells, therefore the population of T-cells of type z_y starts growing only later. This is an other biological feature that could drive to the choice of injecting a higher initial amount of these T-cells, to avoid their early extinction. It is important to stress that, the deterministic system converges to the stable equilibrium $P_{xyz_xz_y}$, when a particular set of initial conditions is fixed, e.g. initial conditions (3.12), but the stochastic process, going under fluctuation, can reach (before the equilibrium $P_{xyz_xz_y}$) an invariant hyperplane, thus it can be absorbed to an equilibrium point different from $P_{xyz_xz_y}$, as Figure 3.5 shows. Thus, stochastic analysis of the model is relevant for therapy success, indeed it shows that a cure scenario is possible (see simulation in Figure 3.6(f)).

3.2.5 Mutations

The examples of model for immunotherapy with T-cell (see Subsection 3.2.2 and Subsection 3.2.3) highlight the essential role of phenotypic switches, which don't involve mutations, in cancer defence mechanisms. Nevertheless, mutations are fundamental to really understand the behaviour of cancer evolution, especially the driver mutations that promote cancer development. In particular, it is interesting to discuss briefly the appearance of rare mutations in large cell populations. In this setting it is necessary to redefine the concept of *fitness value*, which is basic in the study of the stochastic population models. With this purpose authors of [7] have showed some examples (see simulations in Figure 3.8), placed in the context of interaction of rare mutations and fast switches in the case of pure tumour evolutions (that is in a setup without therapy, ignoring the T-cells and the TNF-α proteins):

- the aim is to study the invasion of a mutant that has just appeared in a population close to a stable fixed point;

- let $\mathcal{X}$ be a set of traits of a system, whose evolution can be described by infinitesimal generator (3.5);

- let $M \subset \mathcal{X}$ represent a population in a stable equilibrium such that, $\forall y \in M$, the trait y characterizes one among its subpopulations;

- in the classical case the invasion fitness it is defined by

$$f(x, M) := b(x) - d(x) - \sum_{y \in M} c(x, y)\bar{\mathfrak{n}}_y \tag{3.14}$$

  where biological parameters $b(x), d(x)$ and $c(x, y)$ have the usual meaning, natural birth and death rates and competition kernel, respectively

- $f(x, M)$ represents the growth rate of a population consisting of a single individual with trait $x \notin M$ in the presence of the equilibrium population $\bar{\mathfrak{n}}$ on M. $f(x, M) > 0$ implies that a mutant appearing with phenotype x from the equilibrium population on M has a positive probability (uniformly in K) to grow to a population of size of order K, conversely $f(x, M) < 0$ implies that such a mutant population will die out with probability tending to one, as $K \to \infty$, before its size reaches K.

- in [7] it is proposed to use multi-type branching processes to describe a mutant population of genotype g including all its associated phenotypes (i.e all the traits are of the form $(g, p_i),\ p_i \in \mathcal{P},\ \forall i \in \{1, \ldots, |\mathcal{P}|\}$), when the mutant arrives in a resident population at equilibrium and to redefine the concept of fitness value (3.14). Let an initial population of genotype g be considered, such that it is able to mutate at rate μ_g^K to another genotype g', associated with different phenotypes $p'_j \in \mathcal{P}'$, for $i = 1, \ldots, |\mathcal{P}'|$. The new definition of fitness value of mutant g' is strictly linked to the switching parameters of its phenotypes thought the following relative fitness values: $\forall j \in \{1, ...|\mathcal{P}'|\}$,

$$f_j := b(g', p'_j) - d(g', p'_j) - \sum_{h=1}^{|\mathcal{P}|} c(p'_j, p_h) - \sum_{k=1}^{|\mathcal{P}'|} s^{g'}(p'_j, p'_k) \tag{3.15}$$

  where the biological parameters are the same of the classical definition (3.14) with the adjoint of switching parameters $s'_{jk} := s^{g'}(p'_j, p'_k)$, related to cancer cell switch from (g', p'_j)-type to (g', p'_k)-type;

- the simulations showed in Figure 3.8 depict a population holding just a trait (g, p), which is able to mutate at rate μ_g^K to another genotype g', associated with two different phenotypes:

  (a) The parameters were chosen in order to have $f_2 < 0$ and $f_1 > 0$ (according to definition (3.15)), moreover, in this case, p'_2 can switch to

p_1' ($s_{21}' = 2$), but the back-switch is very weak ($s_{12}' = 10^{-1}$). The global fitness of the genotype g' (as defined in [7]) results positive and close to f_1;

(b) Both phenotypes, p_1' and p_2', have a negative initial growth rate, for the chosen parameters, i.e. $f_1, f_2 < 0$. Instead, the global fitness of the genotype g' is positive, due to the cooperation (i.e. due the switch parameters $s_{12} = 2 = s_{21}$) of the two phenotypes. This is possible because an outgoing switch is a loss of a particle for a phenotype, but not for the whole genotype.

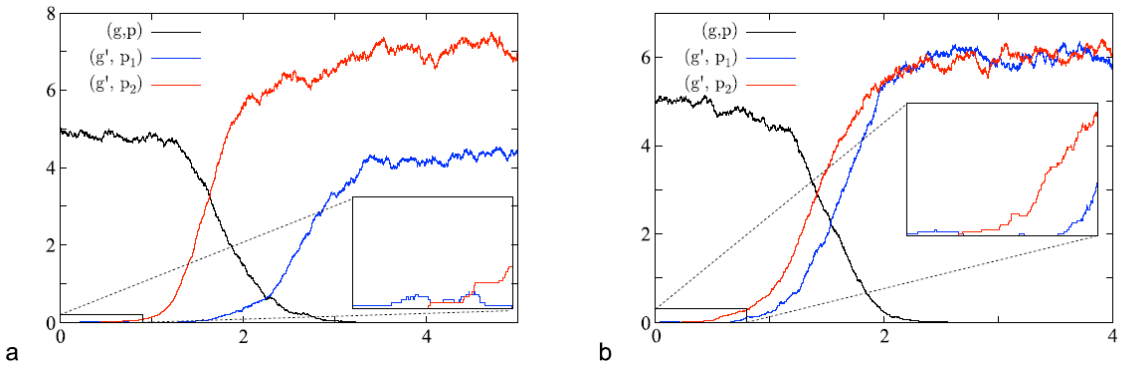

Figure 3.8: Simulations for rare mutations in combination with fast switching, where the number of individuals divided by 200 is plotted versus time [7].

To have a detailed and rigorous explication of this topics one can see for example [13, 14] and, especially for multi-type branching processes [6, 32, 33, 34].

Interplay of Mutation and Therapy

In the context of cancer therapy, drug resistance is a complex phenomenon and appears as a serious problem; indeed failures in the therapy occur during cancer invasion and metastasis related to drug resistance [45]. There are several mechanisms that cause the resistance to tumour therapy, the mutations are among them. The common cancer treatments are surgery, radiation therapy, chemotherapy, combination therapy and laser therapy and almost all of them consist of cycles of treatment. Thus, during a therapy there are phases when populations shrink and regrow due to treatment and relapse phenomena. This may lead to a fixation of a "super-resistant mutant" [7, 45]. The simplest setup to study this effect in the model for immunotherapy of cancer (3.5) has the following stucture:

- melanoma population is monomorphic, with trait (g, p) which can mutate to one fitter trait (g', p');

- switching is excluded, therefore, all parameters concern with it and cytokines are set to zero (the presence of TNF-α only stimulates phenotype switches)

- only birth-reducing competition (BRC) is present (competition kernel increasing the death rate is set to zero, by assumption);

- to include the effect of therapy in the simplest way possible, only one type of T-cells, z, targeting cancer cells of type (g, p) but not of type (g', p'), is considered.

It is crucial to stress that the case under analysis concern with the case of rare mutations in large populations (namely, mutation probability $\mu_g^K \to 0$ as $K \to \infty$) on a timescale such that a population reaches equilibrium before a new mutant appears. Consequently, mutation term does not appear in the limiting deterministic system

$$\begin{aligned}
\dot{\mathfrak{n}}_{(g,p)} &= \mathfrak{n}_{(g,p)}\big(b(p) - d(p) - c_b(p,p)\mathfrak{n}_{(g,p)} - c_b(p,p')\mathfrak{n}_{(g',p')} - t(z,p)\mathfrak{n}_z\big) \\
\dot{\mathfrak{n}}_{(g',p')} &= \mathfrak{n}_{(g',p')}\big(b(p') - d(p') - c_b(p',p)\mathfrak{n}_{(g,p)} - c_b(p',p')\mathfrak{n}_{(g',p')}\big) \\
\dot{\mathfrak{n}}_z &= \mathfrak{n}_z(-d(z) + b(z,p)\mathfrak{n}_{(g,p)})
\end{aligned} \tag{3.16}$$

which describes the interactions of the populations holding traits $(g, p), (g', p')$ and z. Furthermore, the effects of BRC on mutation birth are intrinsically stochastic and happen on time-scales diverging with K [7]. To start a qualitative analysis it is basic to find the equilibrium points of system (3.16) without the mutant population ((g', p')-type); namely, it is necessaty to investigate the solution of the following system

$$\begin{cases}
0 &= \mathfrak{n}_{(g,p)}\big(b(p) - d(p) - c_b(p,p)\mathfrak{n}_{(g,p)} - t(z,p)\mathfrak{n}_z\big) \\
0 &= \mathfrak{n}_z(-d(z) + b(z,p)\mathfrak{n}_{(g,p)})
\end{cases}.$$

There are at least three equilibrium points:

- null population size (both cancer cells and T-cells are absent)

$$(\mathfrak{n}_{(g,p)}, \mathfrak{n}_z) = (0, 0) \tag{3.17}$$

- T-cell population is absent and cancer cells reaches an equilibrium dependent on BRC parameter

$$(\mathfrak{n}_{(g,p)}, \mathfrak{n}_z) = \left(\frac{b(p) - d(p)}{c_b(p,p)}, 0 \right) \tag{3.18}$$

  where it is supposed $b(p) - d(p) \neq 0$

- both cancer cell and T-cell populations are present

$$(\mathfrak{n}_{(g,p)}, \mathfrak{n}_z) = \left(\frac{d(z)}{b(z,p)} \ , \ \frac{b(p) - d(p) - c_b(p,p)\frac{d(z)}{b(z,p)}}{t(z,p)} \right) \tag{3.19}$$

  where $\frac{d(z)}{b(z,p)} \neq \frac{b(p)-d(p)}{c_b(p,p)}$.

To study the stability of these fixed point, the eigenvalues of the following Jacobian matrix (associated to the system (3.16) without the equation describing the mutant population) are investigated:

$$J = \begin{pmatrix} b(p) - d(p) - 2c_b(p,p)\mathfrak{n}_{(g,p)} - t(z,p)\mathfrak{n}_z & -t(z,p)\mathfrak{n}_{(g,p)} \\ b(z,p)\mathfrak{n}_z & b(z,p)\mathfrak{n}_{(g,p)} - d(z) \end{pmatrix}$$

Evaluating the matrix at any equilibrium point, it results:

- $(0,0)$ is unstable, when $b(p) - d(p) > 0$; because its related eigenvalues are the real values

$$\lambda_1^0 = b(p) - d(p) \in \mathbb{R} \text{ and } \lambda_2^0 = -d(z);$$

- the equilibrium point (3.18) results stable under conditions

$$b(p) - d(p) > 0 \text{ and } \frac{d(z)}{b(z,p)} > \frac{b(p) - d(p)}{c_b(p,p)},$$

  since its corresponding eigenvalues are

$$\lambda_1 = d(p) - b(p) \text{ and } \lambda_2 = b(z,p)\frac{b(p) - d(p)}{c_b(p,p)} - d(z) \ (\in \mathbb{R})$$

- the equilibrium point (3.19) is stable if

$$\frac{b(p) - d(p)}{c_b(p,p)} > \frac{d(z)}{b(z,p)} \quad \left(\Leftrightarrow \frac{d(z)}{b(z,p)} \Big(b(z,p)\big(b(p) - d(p)\big) - c_b(p,p)d(z) \Big) > 0 \right) \tag{3.20}$$

because the corresponding eigenvalues have negative real part:

$$\begin{aligned}\lambda_{\pm} &= -\frac{c_b(p,p)d(z)}{2b(z,p)} \\ &\pm \sqrt{\frac{c_b(p,p)^2 d(z)^2}{4b(z,p)^2} - \frac{d(z)}{b(z,p)}\Big(b(z,p)\big(b(p)-d(p)\big) - c_b(p,p)d(z)\Big)}\end{aligned}$$

Now, to study the influence of BRC on mutation events, it is necessary to consider the behaviour of the stochastic process $(\nu^K_{(g,p)}(t))_{t\geq 0}$ (which is described by the infinitesimal generator (3.5), adapted to the case under analysis). The total mutation rate of the population of type (g,p) at time $t>0$ is given by the following function of the Markov process $(\nu^K_{(g,p)}(t))_{t\geq 0}$:

$$\mathfrak{m}(\nu^K_{(g,p)}(t)) := \mu_g^K \big\lfloor b(p) - c_b(p,p)\nu^K_{(g,p)}(t)\big\rfloor_+ \nu^K_{(g,p)}(t)K, \tag{3.21}$$

since there is only a BRC (no others competition types are involved). The function $\mathfrak{m}(\nu^K_{(g,p)}(t))$ is strictly positive $\forall\ \nu^K_{(g,p)}(t) \in \left[0, \frac{b(p)}{c_b(p,p)}\right]$; it represents a parabola opened downwards and attaining its maximum at $\frac{b(p)}{2c_b(p,p)}$ (see Figure 3.9).

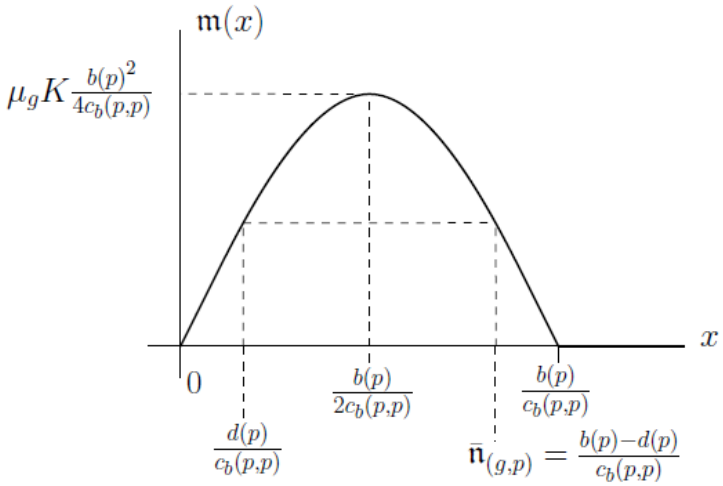

Figure 3.9: Shape of the initial total mutation rate of the population (g,p) [7].

Without or before therapy and before the first mutant appears, the melanoma population $\nu^K_{(g,p)}(t)$ can be approximated by the solution of the deterministic system

$$\dot{\mathfrak{n}}_{(g,p)} = \mathfrak{n}_{(g,p)}\big(b(p) - d(p) - c_b(p,p)\mathfrak{n}_{(g,p)}\big) \tag{3.22}$$

Thus, let the melanoma population be close to its equilibrium (see Figure 3.10(b)), i.e.

$$\nu^K_{(g,p)}(t) = \bar{\mathfrak{n}}_{(g,p)} = \frac{b(p) - d(p)}{c_b(p,p)}$$

(which corresponds to (3.18)). Then the time until a mutation occurs is exponentially distributed with approximate parameter equal to

$$\mathfrak{m}(\bar{\mathfrak{n}}_{(g,p)}) = K\mu_g^K \big(b(p) - c_b(p,p)\bar{\mathfrak{n}}_{(g,p)}\big)\bar{\mathfrak{n}}_{(g,p)} = K\mu_g^K d(p)\bar{\mathfrak{n}}_{(g,p)}.$$

Moreover, looking at Figure 3.9, it is immediate establish that

- the total mutation rate at the equilibrium point $\bar{\mathfrak{n}}_{(g,p)}$ is not maximal, with the following parameters choice

  $$d(p) < \frac{b(p)}{2} \;\Rightarrow\; \bar{\mathfrak{n}}_{(g,p)} > \frac{b(p)}{2c_b(p,p)};$$

  indeed

  $$\mathfrak{m}(\bar{\mathfrak{n}}_{(g,p)}) < \mu_g K \frac{b(p)^2}{4c_b(p,p)};$$

- population with a selected small size have an higher total mutation rate (see Figure 3.10(a));
  i.e.

  $$\nu^K_{(g,p)}(t) \in \left[\frac{d(p)}{c_b(p,p)}\ ,\ \bar{\mathfrak{n}}_{(g,p)}\right] \;\Rightarrow\; \mathfrak{m}(\nu^K_{(g,p)}(t)) > \mathfrak{m}(\bar{\mathfrak{n}}_{(g,p)}) \tag{3.23}$$

Finally, to grasp the interaction between mutation and therapy, let $(\tilde{\mathfrak{n}}_{(g,p)}, \tilde{\mathfrak{n}}_z)$ be the stable equilibrium reached from population under therapy before a mutant birth occurs (i.e. the equilibrium point (3.19) of system (3.16) without the mutant population under stability condition (3.20)). The waiting time for a mutation, starting from equilibrium $(\tilde{\mathfrak{n}}_{(g,p)}, \tilde{\mathfrak{n}}_z)$, is exponentially distributed with approximate parameter

$$\mathfrak{m}(\tilde{\mathfrak{n}}_{(g,p)}) = K\mu_g^K \big(b(p) - c_b(p,p)\tilde{\mathfrak{n}}_{(g,p)}\big)\tilde{\mathfrak{n}}_{(g,p)} = K\mu_g^K \left(b(p) - c_b(p,p)\frac{d(z)}{b(z,p)}\right)\tilde{\mathfrak{n}}_{(g,p)}.$$

Stability condition (3.20) points out that melanoma population at equilibrium has smaller size during therapy, indeed

$$\tilde{\mathfrak{n}}_{(g,p)} = \frac{d(z)}{b(z,p)} < \frac{b(p) - d(p)}{c_b(p,p)} = \bar{\mathfrak{n}}_{(g,p)}.$$

Additionally, if it is assumed that

$$\frac{d(p)}{c_b(p,p)} < \frac{d(z)}{b(z,p)}$$

then

$$\tilde{\mathfrak{n}}_{(g,p)} = \frac{d(z)}{b(z,p)} \in \left[\frac{d(p)}{c_b(p,p)}\ ,\ \bar{\mathfrak{n}}_{(g,p)}\right] \ \Rightarrow\ \mathfrak{m}(\tilde{\mathfrak{n}}_{(g,p)}) > \mathfrak{m}(\bar{\mathfrak{n}}_{(g,p)}) \ \ \text{(by (3.23))}.$$

That is, for a specific choice of parameters, the mutation rate of genotype (g, p) is larger during the treatment with T-cells (see Figure 3.10(c)).

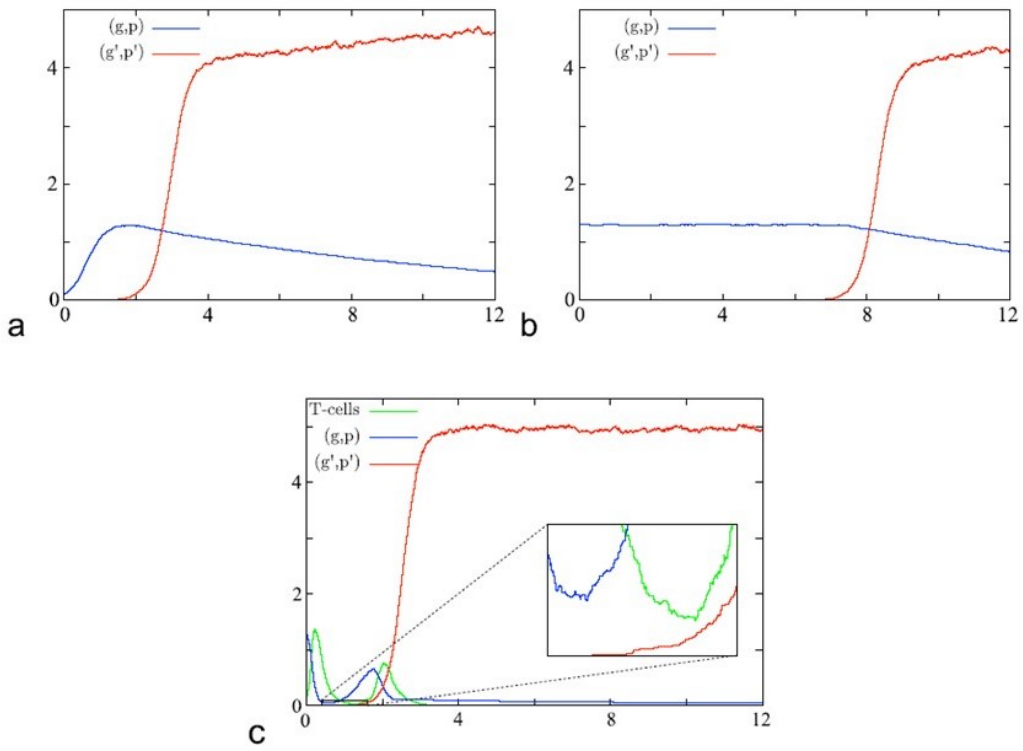

Figure 3.10: The number of individuals divided by 1000 is plotted versus time: Effect for an initial population which is small (a), or at equilibrium (b) or under therapy (c) [7].

Summing up, during immunotherapy, firstly the injection of T-cell works by lowering the melanoma population, but instead to increase the remission probability, it paradoxically increases the probability for melanoma to mutate to a potentially fitter and pathogenic genotype, which is not affected by the T-cells [7].

Conclusions

Cancer is a term which comprises an huge number of different diseases, with various intrinsic dynamics and microenviroments. In this thesis, we have focused on the properties of melanoma, a specific type of skin cancer, to provide an effective example of mathematical model describing cancer growth dynamics. Furthermore, immunotherapy was selected to study the behaviour of melanoma under treatment, instead of the common cancer therapies (such as surgery, radiation therapy, chemotherapy, combination therapy and laser therapy). This choice was made on the basis that immunotherapy works effectively on those tumours that are difficult to deal with traditional therapies [67]. The model proposed in this thesis is a stochastic model for immunotherapy of cancer, that allows to simulate different treatment scenarios. The examples and simulations presented are related to melanoma, and they are based on the experiments presented by Landsberg J., et al. [41] and on a model proposed by Baar M., et al. [7]. Additionally, examples are constructed to point out the intrinsic abilities of melanoma to dedifferentiate its phenotypes, in order to evade the immune system and to mutate, developing resistance to therapy during treatment. In particular, the proposed simulations about interplay of fast phenotypic switches and rare driver mutations, highlight how treatment itself can lead to a smart resistance to immunotherapy. Indeed, the choice of a particular dose of injected T-cell (besides others initial conditions and parameters choices) can induce an earlier mutation in a small population at equilibrium after a cycle of treatment. Consequently, for any cycle of treatment, the probability that a mutation arises before a new equilibrium can be high. This implies that a multiple-cycle treatment protocol could be inefficient due to an higher probability of the cancer to acquire therapy resistance. Moreover, the comparison to experimental data so far looks very promising [7]. The investigation of the interplay between mutation and therapy points out the basilar role of stochastic nature of the chosen model. Indeed, in the considered setting of rare mutation in large population, influence of the mutation rate disappears in the deterministic limit ($\mu_g \to 0$ as $K \to \infty$). Furthermore, the stochastic analysis of the model allows a total remission of cancer as possible scenario. Due to fluctuations, the system can catch the point P_{00000} (the absence of all populations), before reach-

ing a different equilibrium. On the contrary the success of therapy cannot be expected from a pure deterministic analysis of the model. Starting from non-zero initial conditions, the deterministic system goes far from the equlibrium P_{00000} as time increases, since it is not stable. However, the model cannot fully describe the complexity of the interactions among melanoma cells, immune system, microenvironment and therapy; numerical computations could be impractical and theoretical aspects could become too complex to be well understood. Therefore, the model presents some intrinsic limitations by construction; in particular it does not take into account the geometry of the considered biological system (i.e. it does not describe the three-dimensional structure of the tumour and its microenvironment). In addition, examples and simulations were made on the basis of further simplifications, e.g. just one cancer genotype with only two possible phenotypes was considered to study melanoma behaviour without mutation, interactions between cytokines and cells were considered at an individual level, and interplay of mutation and therapy was analysed starting from a monomorphic population with only one possibility of trait choice for the eventually mutation. Another aspect that cannot be overlooked is the choice of parameters; actually, the model parameters are not known well enough and are adjusted to reproduce the experimental findings [7]. Recapping:

- Drug resistance represents the main issue in cancer research. Although chemotherapy has remained the backbone of cancer treatment for many tumour types, almost all failures in the chemotherapy are during the invasion and metastasis of cancers related to drug resistance [21, 45]. A multidisciplinary research team, providing various points of view, is fundamental to deal with this very complex problem.

- The mathematical approach proposed in this thesis represents a useful tool to support the development of new treatment protocols, especially counter-intuitive results, hardly to be evident by laboratory experimentation. In particular, the chosen stochastic model for immunotherapy of cancer applied to melanoma, leads to the conclusion that therapy itself can favour cancer developing. In fact, the examples presented in Chapter 3 show that:

    - phenotypic plasticity of melanoma cells in an inflammatory microenvironment contributes to tumour relapse, after initially successful T-cell immunotherapy;
    - the choice of a particular dose of injected T-cell, under some particular initial conditions, can induce an earlier mutation after any cycle of treatment; the probability of the appearance of a new mutation increases after each cycle.

Consequently, treatment could enhance the cancer acquisition of therapy resistance.

- The limitations of the model presented in this thesis highlight what may be some new challenges for mathematical oncology:
    - the construction of new models, enabling to manage more information without loosing efficiency;
    - the inclusion in the model of the geometry of analysed biological systems;
    - the increase of the number of parameters that are possible to determine experimentally (which may help to calibrate model).

Finally, in the last twenty years, oncology was revolutionized by examples of successful personalized cancer treatments [21]. In this context, the simulations may help in the selection of laboratory experiments [7] necessary during the phase of creation of a personalized therapy. Hence, the time used for laboratory experimentation may be reduced, leading to faster achievement of an effective personalized treatment. Thus, mathematical oncology may play a fundamental role in designing personalized treatment protocols. The hope is that mathematical models, such as the one presented in this thesis, can be improved to become a powerful tool to support the feasibility analysis of a treatment, which is minimally invasive and as effective as possible.

Glossary

The intent of this glossary is to clarifies the meaning of some terms, used in this thesis, that are specific to the field of study (different from Mathematics) to which they belong (Genetics, Biology, Medicine). The following definitions are not meant to be comprehensive, but they are useful tools to understand terms that are not so intuitive or well-known. The latter are essential to understand the genetic and biological framework of the model discussed in this thesis.

- **Apoptosis**
  An energy-requiring physiological process that leads to cell death without exciting an inflammatory response, unlike necrosis. Apoptosis is distinct from programmed cell death although the terms are often treated as interchangeable.

- **Biodiversity**
  The range of genetic, taxonomic, and ecosystem differences that exist in a given area or environment; this can, of course, extend to the whole planet.

- **Bioactivation**
  The conversion of a *xenobiotic* substance to a more toxic or active derivative within the body. The term *xenobiotic* defines any substance found in an organism but that is not produced by that organism and is not a normal constituent of its diet. More often it is used to describe substances foreign to an entire biological system, artificial substances that did not exist in nature before being synthesized by humans.

- **Clone**
  A population of cells or organisms derived from a single progenitor and therefore genetically identical.

- **Cytokine**
  A rather loose category of small proteins that are released by cells and that affect the behaviour of other cells.

- **Endogenous-Exogenous**
  A product or an activity that arises in the body or cell, in contrast to exogenous agents or stimuli that come from outside.

- **Gamete**
  A haploid cell produced by meiosis and involved in sexual reproduction. Male gametes (spermatozoa) are small, motile, and produced in large numbers, whereas female gametes (oocytes) are larger and nonmotile.

- **Genetic code**
  The relationship between the sequence of bases in nucleic acid and the sequence of amino acids in the polypeptide encoded by that DNA. Each amino acid is specified by at least one triplet of bases (a codon), although there is degeneracy in the code and some amino acids are specified by more than one codon.

- **Genetic information**
  the information contained in a sequence of nucleotide bases in a nucleic acid molecule [36]. Heritable characteristics.

- **Genetics**
  The scientific discipline dealing with

  - the study of inheritance and variation of biological traits, and
  - the study of genes, including their structure, function, variation, and transmission [36].

- **Genotoxic**
  A substance, setting, or process that is toxic or harmful to the genetic material. An agent or process that interacts with cellular DNA, either directly or after metabolic biotransformation, resulting in alteration of DNA structure. DNA-adduct formation is one type of genotoxicity [54].

- **Genotype**
  The genetic constitution of an organism, as opposed to the expressed features, the phenotype.

- **Germinal cell**
  cells that produce gametes by meiosis: e.g., oocytes in females and spermatocytes in males [36].

- **Growth Factor**
  A diverse group of proteins that are important in the regulation of cell proliferation (growth) and differentiation. The distinction between growth factors

and cytokines is blurred since some cytokines act as growth factors and some cytokines, originally described as important in the haematopoietic lineages, act on a broader range of cell types. Autonomous growth factor production or altered responsiveness to growth factors is a common characteristic of many neoplastic cells which thereby lose growth control. In particular, the *Trasforming Growth Factor* is family of growth factors secreted by transformed cells that induce the phenotypic characteristics of cell transformation, but do not cause hereditable changes.

- **Immune system**
  the organs (e.g., thymus, lymph nodes, spleen), tissues (e.g., hematopoietic tissue of bone marrow, mucosal and cutaneous lymphoid tissues), cells (e.g., thymocytes, blood and tissue lymphocytes, macrophages), and molecules (e.g., complement, immunoglobulins, lymphokines) responsible for immunity (protection against foreign substances) [36].

- **Homeostasis**
  The maintenance of constancy. Homeostatic mechanisms keep the properties of the internal environment of organisms within fairly well-defined limits and generally require a sensor, a control centre, and positive or negative feedback regulation.

- **Hydrogen bond**
  an association between an electronegative atom, e.g. fluorine, oxygen, nitrogen, or sulfur, and a hydrogen atom attached to another such electronegative atom. Although hydrogen bonding is due to interaction between dipoles, the force of attraction is large enough to permit formation of aggregates of small molecules or to stabilize the conformation of many macromolecules. The spatial relation of the donor and acceptor atoms is such that the hydrogen atom lies very close to the straight line between them [12].

- **Metabolism**
  The complete set of chemical changes that maintain life.

- **Mitosis**
  The process of nuclear division in the somatic cells of eukaryotes in which the genomic information, is distributed equally between two daughter cells so that each contains a *diploid* set of chromosomes identical to that of the parent cell. The adjective *diploid* describes a cell that has two copies of genomic information. Diploid cells have pairs of homologous chromosomes and are usually described as being $2n$ where n is the haploid chromosome number. In mammals, only gametes are haploid.

- **Monomorphic population**
  A population showing only one trait (of potentially variable expression) due to fixation of one allelic form of the gene responsible for that trait [36]. Thus, population holding the same genotype and phenotype.

- **Neoplasia**
  Literally, a term meaning on "new growth" but referring to abnormal new growth that persists in the absence of the original stimulus. The term covers both tumours, where there is an actual swelling, and other proliferative disorders, such as leukaemias, all colloquially referred to as "cancer", although this term strictly refers to carcinoma.

- **Neoplasm**
  any new and morbid formation of tissue; a tumour [12].

- **Nucleoside**
  A purine base (adenine, guanine) or pyrimidine base (cytosine, thymine, uridine) linked glycosidically to ribose or deoxyribose, but lacking the phosphate residues that would make it a nucleotide. The major ribonucleosides are adenosine, guanosine, cytidine, and uridine.

- **Nucleotide**
  A phosphate ester of a nucleoside.

- **Oncogene**
  A normal cellular gene (a proto-oncogene) that is mutated or is overexpressed so that normal restraints on proliferation and sometimes of positional control are lost. The mutation may (e.g.) make the gene product constitutively active, or insensitive to normal regulation.

- **Phenotype**
  The observable characters, including morphology and behaviour of an organism, regardless of the actual genotype of the organism. Identical genotypes do not necessarily produce identical phenotypes.

- **Phenotypic plasticity**
  a phenomenon in which a given genotype may develop different states for a character or group of characters in different environments [36].

- **Physiological**
  normal; not pathological or pharmacological [12].

- **Physiology**
  the study of the dynamic processes of living organisms [36].

- **Population Biology**
  the study of the patterns in which organisms are related in space and time. Such disciplines as ecology, taxonomy, ethology, population genetics, and others that deal primarily with the interactions of organisms or groups of organisms (demes, species, etc.) are included under this term [36].

- **Tumour Necrosis Factor (TNF)**
  A pro-inflammatory cytokine (TNF-α, cachectin, 157 aa). Although it kills tumour cells it also has a wide range of pro-inflammatory actions. Soluble TNF-α is released from the cell surface by the action of TACE (TNF-α converting enzyme). TNF-α, it is secreted in a conventional manner from activated T and B cells.

- **Tumor progression**
  The process that is thought to occur in the course of development of a tumour. Implicitly, the idea that more than one change must occur to cause full malignancy and that initiation must be followed by other changes. In many tumours heterogeneity develops as a result of further mutational events.

- **Tumor suppressor (or Oncosuppressor)**
  Generally a gene (antioncogene, cancer susceptibility gene) encoding a negative regulator of the cell cycle, e.g. an inhibitor of a growth factor signalling system, that must be mutated or otherwise inactivated for unregulated proliferation (neoplasia). There are negative regulators of tumour suppressors that, when overexpressed, increase susceptibility to tumours.

Definition are taken from dictionary [40], where is not specified.

Bibliography

[1] Agur Z., S. Vuk-Pavlovic, Mathematical modeling in immunotherapy of cancer: Personalizing clinical trials, *Mol Ther*, 20(1):1-2, 2012, DOI:10.1038/mt.2011.272.

[2] Agur Z., M. Elishmereni, Y. Kheifetz, Personalizing oncology treatments by predicting drug efficacy, side-effects, and improved therapy: mathematics, statistics, and their integration, Wiley Interdiscip, *Rev Syst Biol Med*, 6(3):239-53, 2014, DOI:10.1002/wsbm.1263.

[3] Altrock P.M. and A. Traulsen, Deterministic evolutionary game dynamics in finite populations, *Phys. Rev. E Stat. Nonlin*, Soft Matter Phys, 80:011909, 2009.

[4] Altrock P.M., L. Liu, F. Michor, The mathematics of cancer: integrating quantitative models. *Nat. Rev. Cancer*, 15:730-745, 2015, DOI:10.1038/nrc4029

[5] Antal T., P.L. Krapivsky, Exact solution of a two-type branching process: models of tumor progression, *J. Statist. Mechan.-Theory Exper*, 2011, DOI:10.1088/1742-5468/2011/08/P08018.

[6] Athreya K.B. and P.E. Ney, Branching Processes (Springer, 1972).

[7] Baar M., et al. A stochastic model for immunotherapy of cancer, *Sci. Rep.* 6:24169, 2016, DOI:10.1038/srep24169.

[8] Barbolosi D., J. Ciccolini, B. Lacarelle, F. Barlesi, N. AndrÃ©, Computational oncology-mathematicalmodelling of drug regimens for precision medicine, *Nat Rev Clin Oncol* 13(4):242-254, 2016, DOI:10.1038/nrclinonc.2015.204.

[9] Basanta D., A.R.A. Anderson, Exploiting ecological principles to better understand cancer progression and treatment, *Interface Focus*, 3:20130020, 2013, DOI:10.1098/rsfs.2013.0020.

[10] Beerenwinkel N., T. Antal, D. Dingli , A. Traulsen, K.W. Kinzler, V.E. Velculescu , et al. Genetic Progression and the Waiting Time to Cancer, *PLoS Comput Biol*, 3(11):e225, 2007, DOI:10.1371/journal.pcbi.0030225.

[11] Bremaud P., Markov chains: Gibbs fields, Monte Carlo simulation, and queues, *Springer Science & Business Media*, 2013.

[12] Cammack R., T. Atwood, P. Campbell, H. Parish, A. Smith, F. Vella, and J. Stirling, Oxford Dictionary of Biochemistry and Molecular Biology, *Oxford University Press*, 2006 (published on line 2008), DOI:10.1093/acref/9780198529170.001.0001.

[13] Champagnat N., A microscopic interpretation for adaptive dynamics trait substitution sequence models, *Stochastic Processes and their Applications*, 116:1127-1160, 2006, DOI:10.1016/j.spa.2006.01.004.

[14] Champagnat N. and S. MÃ©lÃ©ard, Polymorphic evolution sequence and evolutionary branching, *Probab. Theory Relat. Fields* 151:45-94, 2011, DOI: 10.1007/s00440-010-0292-9.

[15] Chang H., N. Pannunzio, N. Adachi, et al., Non-homologous DNA end joining and alternative pathways to double-strand break repair, *Nat Rev Mol Cell Biol*, 18:495-506, 2017, DOI:10.1038/nrm.2017.48

[16] Clancy, S., Genetic mutation, *Nature Education*, 1(1):187, 2008.

[17] D'Auria B., Notes, lecture notes, unpublished, March 20, 2012.

[18] Darling R.W.R., J.R. Norris, Differential equation approximations for Markov chains, *Probab Surveys*, 5:37-79, 2008.

[19] DeAngelis D.L., W.M. Mooij, Individual-Based Modeling of Ecological and Evolutionary Process, *Annu. Rev. Ecol. Evol. Syst.*, 36:147-168, 2005, DOI:10.1146/annurev.ecolsys.36.102003.152644.

[20] Ditlevsen S. and A. Samson, Introduction to Stochastic Models in Biology, lecture note in Mathematics, *Springer Verlag*, 2058, DOI:10.1007/978-3-642-32157-3-1.

[21] Dumbrava E.I., F. Meric-Bernstam, Personalized cancer therapy-leveraging a knowledge base for clinical decision-making, *Cold Spring Harb Mol Case Stud.*, 4(2):a001578, 2018 April 2, DOI: 10.1101/mcs.a001578.

[22] Durrett R., Branching Process Models of Cancer, *Springer*, 2015, DOI:10.1007/978-3-319-16065-8.

[23] Durrett R., Probability models for DNA sequence evolution, *Springer Science and Business Media*, 2008.

[24] Ethier S.N. and T.G. Kurtz, Markov processes, *John Wiley and Sons, Inc.*, 1986.

[25] Friedberg E., DNA damage and repair, *Nature*, 421:436-440, 2003, DOI:10.1038/nature01408

[26] Greenman C., P. Stephens, R. Smith, et al., Patterns of somatic mutation in human cancer genomes, *Nature*, 446:153-158, 2007, DOI:10.1038/nature05610.

[27] Haccou, P. Branching Processes: Variation, Growth, and Extinction of Populations, *Cambridge Univ. Press*, 2005, DOI:10.1017/CBO9780511629136.

[28] Hanahan D. and R.A. Weinberg, The hallmarks of cancer, *Cell*, 100:2000, DOI: 10.1016/S0092-8674(00)81683-9.

[29] Hanahan D. and R.A. Weinberg, Hallmarks of cancer: the next generation, *Cell*, 5:646-674, 2011, DOI:10.1016/j.cell.2011.02.013.

[30] Hua X. and L. Bromham, Darwinism for the Genomic Age: Connecting Mutation to Diversification, *Front. Genet.*, 8:12, 2017, DOI: 10.3389/fgene.2017.00012

[31] Kallenberg O., Foundations of modern probability, *Springer*, New York, 1997.

[32] Kesten H. and B.P. Stigum, A limit theorem for multidimensional Galton-Watson processes, *Ann. Math. Stat.* 37:1211-1223, 1966.

[33] Kesten H. and B.P. Stigum, Additional limit theorems for indecomposable multidimensional Galton-Watson processes, *Ann. Math. Stat.* 37:1463-1481, 1966.

[34] Kesten H. and B.P. Stigum, Limit theorems for decomposable multidimensional Galton-Watson processes, *J. Math. Anal. Appl.*, 17:309-338, 1967.

[35] Kimmel, M. and D.Axelrod, Branching Processes in Biology, *Interdisciplinary Applied Mathematics*, Springer, 2015.

[36] King R.C., W.D. Stansfield, and P.K. Mulligan, A Dictionary of Genetics, *Oxford University Press*, 2007, DOI:10.1093/acref/9780195307610.001.0001.

[37] Klenke A., Probability theory: a comprehensive course, *Springer Science & Business Media*, 2013, DOI:10.1007%2F978-1-84800-048-3.

[38] Korolev K.S., J.B. Xavier, J. Gore, Turning ecology and evolution against cancer, *Nat. Rev. Cancer*, 14.5:371-380, 2014.

[39] La Porta C.A.M., S. Zapperi. Physics of Cancer, *Cambridge University Press*, Cambridge, 2017.

[40] Lackie J., A dictionary of biomedicine, *Oxford University Press*, 2010.

[41] Landsberg J., et al., Melanomas resist t-cell therapy through inflammation-induced reversible dedifferentiation, *Nature*, 490:412-416, 2012, DOI:10.1038/nature11538.

[42] Liggett T.M., Continuous Time Markov Processes, *American Mathematical Society*, 2010.

[43] Loeb L. A., J. H. Bielas, R. A. Beckman, Cancers exhibit a mutator phenotype: Clinical implications, *Cancer Research*, 68:3551-3557, 2008.

[44] Longley D.B., P.G. Johnston, Molecular mechanisms of drug resistance, *J Pathol*, 205(2):275-92, 2005,. DOI: 10.1002/path.1706.

[45] Mansoori B. et al., The Different Mechanisms of Cancer Drug Resistance: A Brief Review, *Advanced pharmaceutical bulletin*, 7,3: 339-348, 2017, DOI:10.15171/apb.2017.041.

[46] McFarland C.D. et al., Impact of deleterious passenger mutations on cancer progression, *Proceedings of the National Academy of Sciences*, 2013, DOI:10.1073/pnas.1213968110.

[47] Murphy H., H. Jaafari, H.M. Dobrovolny, Differences in predictions of ODE models of tumor growth: a cautionary example, *BMC Cancer*, 16:163, published online 2016 Feb 26, DOI:10.1186/s12885-016-2164-x.

[48] The tumour microenvironment, Nature, *Nature Reviews Cancer*, Collection, 2018 September 28, https://www.nature.com/collections/khylqkxqbr.

[49] Nave C.R., HyperPhysics, Georgia State University (Department of Physics and Astronomy), 2016, http://hyperphysics.phy-astr.gsu.edu/hbase/Organic/base.html.

[50] Nelson D.L,. M.M. Cox, Introduzione alla biochimica di Lehninger, *Zanichelli*, Bologna, 2003.

[51] Environmental limits to population growth, OpenStax College, Biology, CC BY 4.0., http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@10.12.

[52] Pacheco J.M., F.C. Santos, D. Dingli, The ecology of cancer from an evolutionary game theory perspective, *Interface Focus*, 4.4:20140019, 2014, DOI:10.1098/rsfs.2014.0019.

[53] Parzen E., Stochastic processes, *Society for Industrial and Applied Mathematics*, 1999.

[54] Porta M., A Dictionary of Epidemiology, *Oxford University Press*, 2014 (published online 2016).

[55] Steel G.G., L.F. Lamerton, The growth rate of human tumours, *Br J Cancer.*, 20(1):74-86, 1966, DOI:10.1038/bjc.

[56] Rockne R.C., A. Hawkins-Daarud , K.R. Swanson, et al., The 2019 mathematical oncology roadmap, *Phys Biol*, 16(4):041005, 2019 Jun 19, DOI:10.1088/1478-3975/ab1a09.

[57] Stratton M., P. Campbell, P. Futreal, The cancer genome, *Nature*, 458:719-724, 2009, DOI:10.1038/nature07943.

[58] Suweis S., F. Simini ,J. Banavar, et al., Emergence of structural and dynamical properties of ecological mutualistic networks, *Nature*, 500:449-452, DOI:10.1038/nature12438.

[59] Swart J.M., A Course in Interacting Particle Systems, lecture notes, unpublished, December 16, 2019.

[60] Tabassum S., et al, Mathematical Modeling of Cancer Growth Process: A Review, *J. Phys.: Conf. Ser.*, 1366:012018, 2019, DOI:10.1088/1742-6596/1366/1/012018

[61] Taylor C., et al., Evolutionary Game Dynamics in Finite Populations, *Bulletin of Mathematical Biology*, 66:1621-1644, 2004, DOI:10.1016/j.bulm.2004.03.004.

[62] Tomasetti C. and B. Vogelstein, Variation in cancer risk among tissues can be explained by the number of stem cell divisions, *Science*, 347(6217):78-81, 2015 January 2, DOI:10.1126/science.1260825

[63] Vaghi C., A. Rodallec, R. Fanciullino, et al., Population modeling of tumor growth curves and the reduced Gompertz model improve prediction of the age of experimental tumors, *PLoS Comput Biol*, 16(2):e1007178, Published online, 2020 Feb 25, DOI:10.1371/journal.pcbi.1007178.

[64] Van Zanten H., An Introduction to Stochastic Processes in Continuous Time, lecture notes, unpublished, November 8, 2004.

[65] Vandermeer J., How Populations Grow: The Exponential and Logistic Equations, *Nature Education Knowledge*, 3(10):15, 2010.

[66] Wu S., S. Powers, W. Zhu, et al., Substantial contribution of extrinsic risk factors to cancer development, *Nature*, 529:43-47, 2016, DOI:10.1038/nature16166

[67] https://www.cancerresearch.org/immunotherapy/why-immunotherapy

[68] Pros and Cons of Immunotherapy, *WebMD Medical Reference*, 2021, https://www.webmd.com/cancer/immunotherapy-risks-benefits.

[69] What types of mutation are there?, *Yourgenome*, 2017, https://www.yourgenome.org/facts/what-types-of-mutation-are-there.