Modelling the impact of cyclic hypoxia on cell-cycle regulation in cancer cells

Giulia Laura Celora

Worcester College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2022

Acknowledgements

I would like to thank the Engineering and Physical Sciences Research Council (EPSRC) and Cancer Research UK (CRUK) for funding this DPhil project and the Centre for Doctoral Training (DTC) for training me to face the challenges of research. All that I have achieved this year would have not been possible without the help, insights and support of my academic supervisors (the magic trio: Helen, Philip and Joe) who have taught me the invaluable skill of asking the right question at the right time. Equally I would like to thank my industrial supervisor Ester, who has helped me unravel the complexity of biology and trusted a clumsy mathematician like me into her lab. I would also like to thank my collaborators in and outside the Mathematical Institute: Sarah, Matt, Andreas, Barbara, Sam, Ben and Panos, for sharing with me their time, knowledge and ideas.

I would like to thank all of the people that made the past five years in Oxford a great and memorable part of my life. First, I need to acknowledge all the members of the Wolfson Centre for Mathematical Biology, and particularly Solveig and Sofia. I will miss our coffee breaks, lunches, dinners, cakes, happy hours and our attempts to become Tik-Tok stars (I still believe we have a chance; not sure about Prateek though). Then there is my Oxford family: Matteo, Johann, Julien, Prateek and Lea. I thank them for making 23 Walton Well Road my home for the past 4 years and not just my house; for bearing with me through the pandemics and supporting me in the up and down moments of my PhD. I thank le "capre", who have been my support and strength from high-school (or even kindergarten) onwards.

Finally, and most importantly, I want to thank my mother for her unwavering support and for choosing to fight, when many other people would have not. I am lucky if I inherited even half of her strength. And then, there is my sister. Without her advice (and stubbornness) I would not be in Oxford today.

Abstract

In vivo observations show that oxygen levels in tumours can fluctuate on fast timescales. As a result, cancer cells can be periodically exposed to pathologically low oxygen levels; a phenomenon known as *cyclic hypoxia*. Even though *in vitro* models of cyclic hypoxia exist, they fail to capture the complex and heterogeneous oxygenation dynamics of real tumours. Mathematical models can help to overcome current experimental limitations and, by doing so, offer new insights into the biology of cyclic hypoxia by predicting cell responses to a wide range of cyclic dynamics. Recent experimental evidence suggests that cyclic hypoxia alters the progression of cancer cells along the cell-cycle. In this thesis, we use a range of mathematical approaches to study cell-cycle dysregulation in cyclic hypoxia and its impact on the growth and survival of cell cultures.

Accordingly, we develop and analyse a series of mathematical models that can be used, alongside *in vitro* experiments, to increase understanding of cell responses to cyclic hypoxia. First, we present a deterministic cell-cycle model to describe the dynamics of a population of cells growing in different oxygen environments, in the absence of competition for space. Our model comprises a system of 6 linear differential equations with two time delays and describes the evolution of the number of cells in different cell-cycle states. The impact of hypoxia on cell-cycle progression is captured by introducing arrest (or checkpoint compartments) and allowing the rates at which cells transition between different cell-cycle states to depend on the oxygen environment they experience. We then develop a corresponding individual-based model to study the impact of cyclic hypoxia on cell survival in the presence and absence of treatment.

We apply the models we develop in two different ways. On the one hand, we use a combination of analytical and numerical approaches to identify cell-cycle control strategies that favour proliferation and/or cell survival in different oxygen environments. Specifically, we identify environmental regimes in which cells may, or may not, benefit from having defective cell-cycle checkpoints. Furthermore, we illustrate how our models can be used as a tool to complement and design *in vitro* experiment to uncover the biological mechanisms that drive cell responses to constant and cyclic hypoxia.

Contents

Chapter 1

Introduction

Despite the vast knowledge that has been acquired in the long-standing fight against cancer, this disease remains one of the leading causes of mortality worldwide and a major barrier to increasing average life-expectancy. According to the World Health Organization, 10 million people died of cancer in 2020 [135]. Even when detected at early stages, treatment outcomes differ markedly between patients, with one of the major causes of failure being intra-tumour heterogeneity.

In the past few decades, there has been a major shift in our understanding of cancer. The development of new technologies, such as single-cell sequencing, has showcased the complex composition of tumours, where cells with different genotypes and/or phenotypes can coexist in the same patch of tissue [19, 69, 97]. Such intra-tumour heterogeneity is currently a key limitation to the effectiveness of treatments [42] since it allows for phenotypes that are resistant to standard treatment strategies [7, 98]. These are selected for during treatment and can drive subsequent relapse. Understanding what drives intra-tumour heterogeneity is, therefore, an important first step for the design of more effective treatment strategies [7, 28, 98]. However, this is challenging as it combines genetic and epigenetic factors and, importantly, the external influence of the tumour micro-environment [19, 98].

In light of the research in intra-tumour heterogeneity, cancer is no longer viewed as a disease in which mutant cells acquire the ability to proliferate and avoid programmed cell-death. Rather, it is regarded as a complex ecosystem where malignant cells interact with neighbouring normal cells and their surrounding environment to favour malignant growth [63]. The tumour micro-environment is highly heterogeneous and this is reflected in the intra-tumour heterogeneity of cell populations. Indeed, cells can adapt to, and remodel, their local micro-environment to construct and maintain favourable ecological *niches*. Such local interactions between cancer cells and the tumour micro-environment arise from the interplay of molecular mechanisms

(*e.g.*, signalling pathways) and further affect the overall macro-scale dynamics. The complex multi-scale nature of *in vivo* cancer can be difficult to replicate in *in vitro* experiments. While *in vitro* experiments can help disentangle processes occurring at different temporal and spatial scales, they often fail to replicate the interactions that occur across multiple spatial and temporal scales. Mathematical models provide a means to integrate the acquired knowledge and increase understanding of the complexity underpinning the system. This is a crucial step in order to translate research into the clinic and move towards personalised care.

Spatio-temporal variability in tissue oxygenation is a characteristic feature of solid tumours [8, 57], where regions of abnormally low oxygen levels (i.e., *hypoxia*) form as a result of uncontrolled cell proliferation and abnormal vascular structures. Hypoxia is recognised as a major driver of tumour progression and the heterogeneity that correlates with tumour aggressiveness and therapeutic resistance. Adaptation of cancer cells to hypoxic stress is associated with several of the hallmarks of cancer, including genomic instability (i.e., higher mutation rate), metabolic reprogramming and apoptotic resistance [57, 63, 106]. *In vivo*, two forms of tumour hypoxia have been identified: diffusion-limited and perfusion-limited hypoxia. The former occurs as a result of distance from blood vessels and gives rise to regions exposed to constant (or chronic) hypoxia. In contrast, perfusion-limited hypoxia is due to perturbations in tumour blood flow and vascular remodelling, which result in transient and recurring exposure to hypoxia (i.e., cyclic hypoxia) [103]. While both forms of hypoxia contribute to disease progression and tumour radio-resistance, experimental evidence suggests that they do so by activating different biological mechanisms [4, 21]. While constant hypoxia has been widely studied, less is known about how fluctuations in oxygen levels affects cancer cell behaviour and, therefore, disease progression. This knowledge gap reflects the experimental difficulties involved in quantifying cyclic hypoxia *in vivo* and reproducing such conditions *in vitro*. Furthermore, the large number of degrees of freedom characterising cyclic hypoxia (such as periodicity and amplitude of oxygen levels) constitute a great challenge in the development of a general theory solely based on experiments.

In this thesis, we develop a new mathematical model that can be combined with *in vitro* models, to describe how different forms of hypoxia impacts cell progression along the mitotic cycle. We use our model to understand how cell-cycle dysregulation in cyclic hypoxia affects proliferation, survival and treatment responses of tumour cells. In doing so, we exploit a variety of mathematical techniques, including continuum, stochastic and data-driven modelling, depending on the question of interest and the

type of data available. Our goal is to build and validate robust mathematical models that explain experimental findings, aid in the design of novel experiments and help to bridge the gap between *in vitro* and *in vivo* responses to cyclic hypoxia.

In this chapter, we review the relevant biological and mathematical literature. In Section 1.1, we introduce the biology of cancer, with a focus on the cell-cycle (or mitotic cycle) and hypoxia. We explain how dysregulation in cell-cycle control mechanisms leads to uncontrolled proliferation of cancer cells and how hypoxia contributes to such dysregulation. In particular, we outline what is currently known about how hypoxia contributes to tumourogenesis and its role in altering important cellular functions, such as the cell-cycle and DNA damage repair. In doing so, hypoxia can manipulate how cancer cells respond to common treatments, such as radiotherapy (RT). This is discussed in Section 1.1.4, where we outline the radiobiological factors that determine cell sensitivity to RT.

Limitations in the study of tumours *in vitro* and the need for personalised cancer treatments have driven research in mathematical modelling of cancer. In Section 1.2, we review the relevant literature, focusing on standard approaches to model the cell-cycle and the relationship of these models to tumour growth laws. We also summarise the *in vitro* and *in vivo* experimental techniques used to study the cell-cycle, which have been used to validate existing phenomenological models. While cell-cycle models differ in their assumptions and formulations, mathematically, they can be grouped into the class of positive dynamical systems. In Section 1.2.3, we summarise some of the analytical and theoretical properties of these systems and known results about their asymptotic behaviour that will be used later in the thesis.

In contrast to theoretical physics, models of tumour growth are phenomenological, which implies that their parameters cannot be derived from known physical constants and cannot, in most cases, be measured directly. As such, for a model to have biological relevance it must be validated against experimental data. In Section 1.3, we review techniques that have been used to assess model identifiability and perform parameter inference, with a particular focus on the methods used throughout the thesis.

1.1 Cancer biology: "uncontrolled" proliferation

In eukaryotes, cell division (via the mitotic cycle) is regulated by a highly-conserved, complex network of feedback mechanisms that ensure the timely and accurate duplication and segregation of the genome. Cancer is a general term used to describe

a group of diseases in which cells divide continuously and excessively as a result of mutations that disrupt cell-cycle control [100]. Recent work has revealed that the common assumption that cancer cells undergo uncontrolled cell division is inaccurate and ignores the dependence of cancer cells on specific cell-cycle control mechanisms (or checkpoints) for survival. This novel paradigm has resulted in the design of treatment strategies [51, 108] that exploit differences in cell-cycle regulation between cancer and healthy cells. However, the use of these drugs in the clinic remains limited due to the lack of understanding of the biological mechanisms regulating the cell-cycle *in vivo* and how these are influenced by aspects of the tumour micro-environment (*e.g.*, hypoxia).

In this section, we review the mechanisms that are known to regulate the mitotic cell-cycle and explain how their disruption can lead to cancer (we refer to reviews in the literature [87, 100] for more details on the molecular mechanisms and pathways that orchestrate cell-cycle dynamics). We then summarise key experimental findings related to cell-cycle regulation in hypoxia and conclude by detailing open questions that we aim to address in this thesis.

1.1.1 The mitotic cell-cycle

As shown in Fig. 1.1, the mitotic cell-cycle is commonly divided into four phases: G1 (growth in preparation for DNA replication), S (DNA synthesis), G2 (growth and preparation for mitosis) and M (mitosis). As cells proceed along the cell-cycle there are two key decisions to take: whether to enter the cell-cycle (allowing initiation of DNA replication) and whether to proceed into mitosis. These decisions are regulated by complex interactions between cellular pathways and external stimuli. At the cellular scale, control mechanisms (or checkpoints) guarantee timely and accurate replication of the genome (in the S phase) and its correct segregation into two daughter cells (in the M phase). At the tissue scale, environmental cues, such as growth factors, nutrient levels and mechanical stress, can favour re-entry to, or arrest of, the mitotic cycle in order to maintain tissue homeostasis.

Cell-cycle checkpoints are essential for preventing the accumulation of errors and damage during cell division. As shown in Fig. 1.1b, checkpoints can be subdivided into three types [100]. The spindle assembly checkpoint is concerned with proper partitioning of the replicated DNA during the M phase. The DNA damage and replication stress checkpoints have evolved to prevent, and cope with, DNA damage, thus maintaining genetic stability; hence they are usually referred to more broadly as DNA damage response (DDR) [108]. The DNA damage checkpoint is activated by

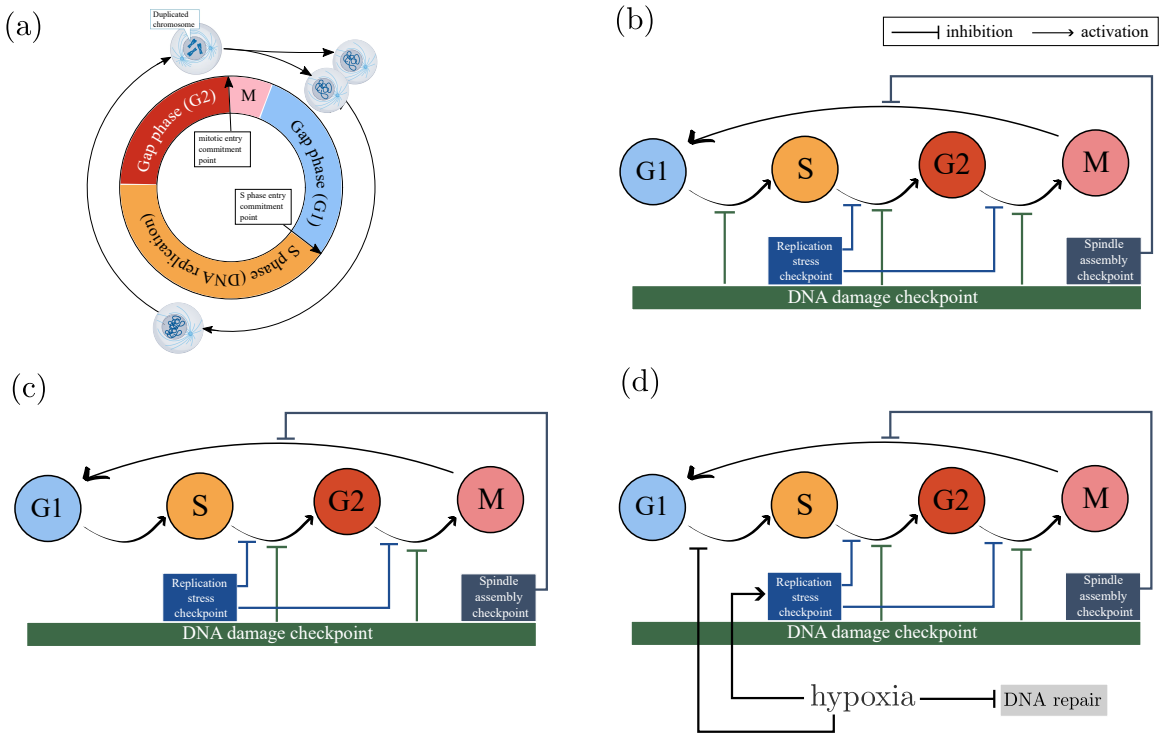


Figure 1.1: The mitotic cycle and its regulation: (a) schematic highlighting the standard phases of the cell-cycle and the two key events occurring: DNA synthesis and mitosis leading to the formation of two daughter cells. (b-d) Schematic illustrating the role of different checkpoints in delaying the timing of transitions between subsequent phases of the cell-cycle in (b) a normal cell; (c) a cancer cell and (d) a cancer cell exposed to hypoxia (pathologically low oxygen levels). More details on the role of different checkpoints can be found in the main text. The schematics are inspired by the work of Matthews et al. [100].

detection of DNA damage and involves different pathways depending on which phase of the cell-cycle a cell is in. Depending on the severity of the damage, the DDR can promote different cell fate decisions. Cells may transiently arrest, to allow time for DNA repair (resulting in delays in the duration of the cell-cycle), via inhibition of the G1/S and G2/M transitions and DNA replication. When DNA damage is irreparable, the DNA damage checkpoint promotes permanent cell-cycle arrest (*senescence*) or death via apoptosis. The replication stress checkpoint does not require the presence of DNA damage. It acts to prevent replication stress-induced DNA damage. As shown in Fig. 1.1b, this checkpoint is only activated during the S phase as a result of impaired DNA synthesis. DNA synthesis is initiated at precise locations along a chromosome called *replication forks* (RF), where the two strands of the DNA separate or "fork". Replication stress is defined as the accumulation of single-stranded DNA due to RFs

slowing down/stalling. Causes of fork stalling range from abnormal DNA topology to shortages of components required for DNA synthesis. The primary function of the replication stress checkpoint is to maintain the integrity of stalled forks and allow replication to resume once the cause of fork stalling is resolved. This is, however, a temporary solution and prolonged fork stalling can lead to severe DNA damage (*i.e.*, fork collapse). Exposure to replication stress during S phase impacts the latter G2/M transition. For mild levels of replication stress, delays in the initiation of mitosis are observed, while higher levels of replication stress can lead to replication catastrophe, which leads to either cell death or senescence [138].

1.1.2 Cell-cycle (dys-)regulation in cancer

As shown in Fig. 1.1c, in cancer cells, cell-cycle regulation is partially compromised when compared to healthy cells (Fig. 1.1b). The DDR is activated early during tumourigenesis as a barrier to oncogene activity and oncogenic stresses [9, 58]. Transition from a pre-cancerous tissue to cancer is often associated with cells losing one or more DNA damage checkpoint pathways in response to oncogenic stress. The most frequent mutations in cancer drive S phase entry and allow cells to escape apoptosis/cell-cycle arrest as the threshold for "severe" damage that dictates cell-fate decision increases. As a result, cancer cells continue to proliferate, acquiring higher levels of genetic instability (*i.e.*, increasing mutation frequency [106]). Since sustained proliferation exposes cancer cells to increased levels of replication stress, the integrity of the replication stress response is vital for cancer cell survival and, therefore, it is rarely dysfunctional in cancer. The rationale behind most cancer treatments, such as radiotherapy and chemotherapy, is to induce damage levels that are lethal in cancer cells but not in healthy cells. Inhibition of the DDR has proven effective for overcoming hypoxia-mediated resistance to radio- and chemo-therapy [51]. In particular, Prof. Hammond's lab and others have shown that cell-cycle regulation differs in hypoxic and normoxic conditions. They have also identified the molecular mechanisms connecting severely low oxygen levels, replication stress, activation of DDR and the pro-apoptotic p53 tumour suppressor [54, 52, 88, 110]. In Fig. 1.1d, we summarise their findings by schematically illustrating cell-cycle progression in hypoxia. However, these observations relate only to cell responses to acute/chronic hypoxic conditions; less is known about cell-cycle regulation in cyclic hypoxia. This motivates the work presented in the first two chapters of this thesis. In what follows we summarise what is currently known about cell-cycle regulation in different hypoxic environments. To

avoid confusion, we first give a quantitative definition of what we mean by hypoxia (and its different forms).

1.1.3 Cell-cycle (dys-)regulation in hypoxia

As a tumour develops, excessive cell proliferation beyond the capacity of the tissue leads to an imbalance between oxygen consumption and supply, resulting in pathologically low oxygen levels (*i.e.,* hypoxia). Adaptation of cancer cells to hypoxia is regulated by *hypoxia-inducible factors* (HIFs), which drive the activation of pathways associated with several of the hallmarks of cancer, such as angiogenesis, escape of cell death and cell-cycle dysregulation [4, 90, 81]. Stabilisation of HIF is associated with oxygen levels in the range $0 - 6\%\ O_2$ [4] and drives cells to switch their phenotype to decrease their energy requirements and facilitate tumour oxygenation (for example by release of angiogenic factors). When studying hypoxia, the specific oxygen concentration to which cells are exposed plays an important role in their response. This has led to a more refined classification of hypoxia: anoxia ($\approx 0\%\ O_2$), severe/radiobiological hypoxia ($\approx 0.1 - 1\%\ O_2$) and mild hypoxia ($\approx 1 - 2\%\ O_2$). In practice, tumour tolerance to shortages of oxygen depend on the tissue in which it emerges. As such, these ranges are merely indicative. Additional forms of hypoxia are needed to account for oxygen levels that vary over time. While constant hypoxia affects tumour regions at a significant distance from vessels, cyclic/intermittent hypoxia is observed both close to, and far from, blood vessels, with periods ranging from seconds to hours/days [4, 119]. Furthermore, there can be large variability in the amplitude and frequency of oxygen fluctuation within the same tumour. High frequency fluctuations are usually associated with vasomotor activity, while processes, such as vascular remodelling, can generate cycles with longer periods [103]. Recent work suggests that self-sustained fluctuations in the blood flow might be related to blood vessel network topology, which is known to be abnormal in tumours [17].

There are several challenges to be overcome when replicating hypoxic conditions using *in vitro* devices, such as hypoxic chambers [78]. Due to the large volume of air in these chambers, there is a limit to how low oxygen levels can drop and also how rapidly oxygen levels can change, precluding the study of anoxia and the high-frequency oxygen fluctuations that are observed *in vivo*. Another challenge relates to short exposure to ambient air (for example, when removing a culture dish from the chamber to make a measurement) which can have a severe impact on oxygen levels in the chamber, making it difficult to control the environment which cells experience [78]. Furthermore, while oxygen levels in the chamber may be spatially-homogeneous,

oxygen levels *in vivo* are often characterised by large oxygen gradients. Microfluidic devices can overcome some of the limitations of hypoxic chambers and achieve local spatio-temporal control of oxygen levels in different cultures (including monolayers and spheroids). However, microfluidic devices are still limited in how low oxygen levels can drop and also how fast oxygen levels can change in the device [18, 80]. Since this technology is still in development, only a few studies have used microfluidic devices to study hypoxia and most of the experimental results and data available (and used in this thesis) are from experiments in hypoxia chambers.

Motivated by available experimental data, in our work, we focus on the effect of periodic exposure to severe hypoxia. This may capture the oxygen dynamics in the pre-necrotic region of tumours where the most radio-resistant tumour cells are located [151]. In Fig. 1.1d, we illustrate how severe hypoxia (here simply denoted by hypoxia) affects cell-cycle transitions. *In vitro* experiments have shown that exposure to severe levels of hypoxia (specifically $< 0.5\%$ O_2), inhibits progress through S phase due to a rapid reduction in the rate of DNA synthesis [52, 114]. This has been attributed to impaired functioning of the enzyme ribonucleotide reductase (RNR) [52, 109], which mediates *de novo* production of *deoxynucleotide triphosphates* (dNTPs). Since dNTPs are the building blocks of DNA, the decrease in dNTP levels in severe hypoxia causes DNA synthesis to stall. In contrast, milder levels of hypoxia (1-2% O_2) do not impact dNTP levels [52]. The reduction in the rate of DNA synthesis activates the replication stress checkpoint that stabilises replication forks. As discussed above, the ability of cells to stabilise stalled forks is transient so that if severe conditions are prolonged ($>$ 12 hours) then cell death occurs [114] due to fork collapse [107]. As shown in Fig. 1.2, we use this 12-hour threshold to distinguish between acute ($<$ 12 hours) and chronic (>12 hours) constant hypoxia.

If oxygen levels are restored after acute hypoxia exposure, cells in the S phase can restart the cell-cycle although they may accumulate DNA damage, likely associated with the accumulation of reactive-oxygen species (ROS) during re-oxygenation [103]. Depending on the amount of damage sustained, activation of cell-cycle checkpoints causes cells to accumulate in the G2 phase and prevents damaged cells from entering mitosis [21, 59, 109]. Importantly, DNA damage repair is affected by hypoxia and inhibition of repair mechanisms also plays a role in the acute phase. The cellular decision to repress DNA repair has been interpreted as a cells initiating an energy-preservation programme [115]. Re-oxygenated cells risk accumulating DNA damage, which can not be efficiently repaired due to impaired DNA repair capabilities, resulting in higher levels of genomic instability [21]. This might be exacerbated

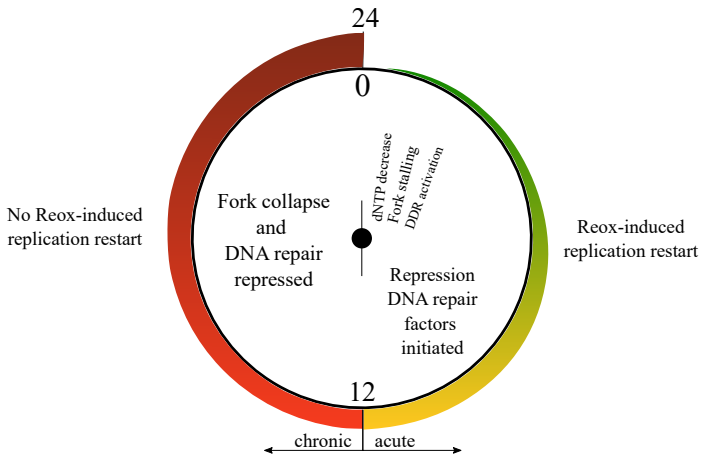


Figure 1.2: Schematic representation of how cell response changes with the duration of hypoxia ($< 0.1\%$ O_2). If exposure to hypoxia lasts for less than 12 hours, cells can restart DNA synthesis upon re-oxygenation (reox). If cells are re-oxygenated after 12 hours, most of them will be permanently damaged and unable to restart DNA synthesis. Image is adapted from [115].

in cells exposed periodically to hypoxia (*i.e.,* cyclic hypoxia) [5]. Increased genomic instability and the strong selective pressure in cyclic hypoxia may drive the emergence of more aggressive and resistant clones with impaired checkpoint activation. In this thesis, we use mathematical modelling to understand which cell-cycle regulation strategies might favour proliferation and survival of tumour cells in different cyclic hypoxia environments. In doing so, we can explore a wider range of cyclic hypoxia protocols than those that can be explored experimentally. From this point of view, we propose mathematical modelling as a tool to overcome the limitations of hypoxic chambers.

1.1.4 The relationship between cell-cycle, hypoxia and cell responses to radiotherapy

Radiotherapy (RT) is widely used to treat patients with solid tumours. RT is often combined with other treatments such as chemotherapy or immunotherapy and, also, surgery [147]. Radiotherapy acts by damaging biomolecules in tumour cells, such as proteins and DNA. Such damage can result in cells losing their clonogenic capacity (i.e., senescence) or in cell death via apoptosis or necrosis [147]. In the realm of radiobiology, it is well-known that different aspects of the tumour micro-environment play a critical role in determining cell response to radiotherapy [136]. For example, oxygen is a radio-sensitiser: it consolidates the fixation of DNA damage and, therefore, radio-induced cell killing. However, hypoxia can also indirectly affect a cell's ability to

resist treatment by altering its cell-cycle and activating DNA repair mechanisms [25]. Much less is known about how cyclic hypoxia affects the response of cells to RT. Nonetheless, there is compelling evidence that cells able to survive exposure to cyclic hypoxia do have increased radio-resistance [67, 146]. One explanation proposed in the literature is that cyclic hypoxia enables cells to improve their antioxidant capacity, which leads to lower levels of reactive ROS and, therefore, decrease RT-induced damage [67]. However, these studies do not consider how cell-cycle dysregulation due to cyclic hypoxia may affect RT outcomes. Given the complex interplay between cell-cycle regulation and fluctuating oxygen-levels, mathematical models can help to interpret and explain the results of experiments, by disentangling the effect of different mechanisms on the measured experimental outcome. This aspect will be discussed in more detail in Chapter 6.

1.2 Mathematical models of tumour growth

The simplest deterministic models of tumour growth describe how a tumour's volume $V = V(t)$ evolves over time, t. In this class of models, the most fundamental, and widely used, is the exponential growth model [84]:

$$\frac{dV}{dt} = \lambda V(t), \tag{1.1}$$

where the positive constant λ (also known as *Malthusian* parameter) indicates the net growth rate of the tumour volume. Eq. (1.1) is commonly used to describe situations where cells grow in a resource-rich environment where competition can be neglected (*e.g.*, *in vitro* growth of 2D monolayers). Under the assumption of constant cell density ρ [cell/volume], $V(t)$ can be related to the number of cancer cells $N = N(t) = V(t)\rho$. While Eq. (1.1) can be effective in describing the overall population level dynamics, it neglects information on the mechanisms that regulate increases in population size. In practice, the choice of whether to use such a simple phenomenological description will depend on the question of interest and the type of data available. From this point of view, it is yet to be understood the extent to which Eq. (1.1) represents an appropriate model of tumour growth under cyclic hypoxia. We investigate this question in Chapter 4, where we derive a growth law based on a more detailed description of cell-cycle regulation in cyclic hypoxia.

As shown in Fig. 1.3, tumour growth emerges from individual cancer cells following the mitotic cycle, which, as discussed in Section 1.1, is controlled by a complex network of signalling pathways that involve hundreds of macro-molecules. While a

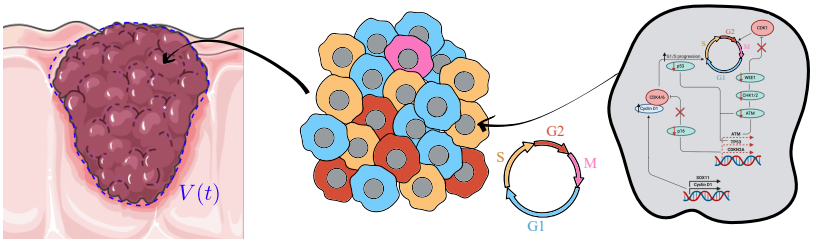


Figure 1.3: Schematic illustrating how cancer tissue growth emerges from cellular level mechanisms regulating cell progression along the mitotic cycle. The gene regulatory network on the right-hand side of the figure is adapted from Fig. 3 in [22].

complete description of tumour growth should start from the intracellular scale, a mathematical and computational approach that couples intracellular scale dynamics with the emergent macroscopic growth dynamics of tumours is still lacking. Instead, several mathematical formalisms have been proposed to investigate cell-cycle evolution at different spatial scales, in normal 'healthy' conditions, as well as perturbed environments, for example in the context of drug development. These formalisms encompass discrete [50, 140], continuous (either deterministic [1, 10, 12, 48, 94, 127, 126] or stochastic [3, 150]) and hybrid approaches [2, 128].

Mathematical models based on chemical reaction networks have been proposed to describe the complex pathway interactions that regulate cell-cycle progression in a single cell. Combining this modelling approach with the wealth of proteomic and, more recently, transcriptomic data has increased our understanding of the molecular and transcriptional mechanisms that regulate the cell-cycle in normal cells [111, 141] and which are dysregulated in cancer [1, 76]. Despite recent advances [44, 154], these models are yet to be upscaled to describe cellular decisions in a population level model due to the large number of proteins involved in such models (many more than the core network in Fig. 1.3).At the population-level, phenomenological models have been used instead to describe tumour growth. When interested in the population scale, population balance (PB) models are often used [10, 12, 48, 94, 126, 133]. These take the form of age- and/or phase-structured models, in which a structure variable is introduced to track progress through the cell-cycle. Being an abstract representation of the emergent behaviour that is driven by complex molecular mechanisms, such models are heavily dependent on experimental data for validation and parametrisation. Consequently, we start by summarising some of the most common and more recent experimental protocols used to investigate cell-cycle dynamics, since they have strongly influenced the mathematical techniques in the literature.

1.2.1 Measuring cell-cycle dynamics

The design of novel techniques to monitor cell-cycle dynamics has allowed the study of cell-cycle perturbations both in experimental model of tumours (cell-line cultures or mice) and tumour cells sampled from patients. Popular techniques for monitoring cell-cycle dynamics currently include flow cytometry and time-lapse microscopy using fluorescent ubiquitination-based cell-cycle indicator (FUCCI). Both methods can be used to indirectly identify the stage of the cell-cycle a cell is in, by measuring DNA content (flow cytometry) or the expression of cell-cycle related proteins (FUCCI). As shown in Fig. 1.4a, (standard) FUCCI labelling allows cells in the G1 phase and early S phase to be distinguished from cells in the late S/G2/M phase. The advantage of this approach is that it generates time- and spatially-resolved data that follow the fate of individual cells. However, since cells must be engineered prior to culture to express the fluorescent proteins, this method can not be applied to patient-derived cells.

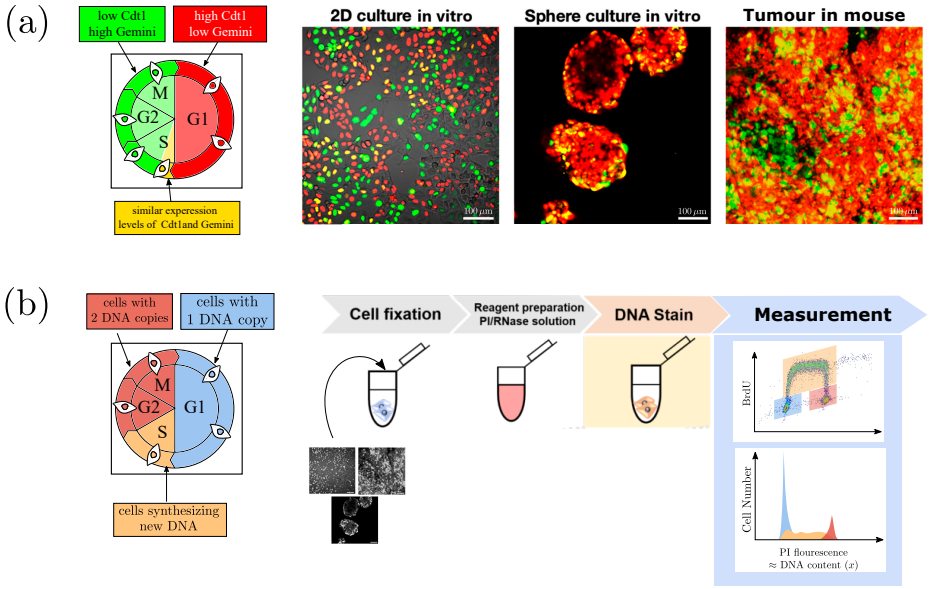


Figure 1.4: Experimental methods to estimate cell-cycle dynamics: (a) real time imaging with fluorescent ubiquitination-based cell-cycle indicator (FUCCI). Scale bar: 100 μm. (b) Flow cytometry based on DNA content.

Flow cytometry is a more established technique which can distinguish cells in different stages of the cycle, sorting them according to their DNA content (see Fig. 1.4b). While cells in the S phase are synthesising new DNA and, therefore, have a variable amount of DNA (from one to two copies), cells in the G1 and G2/M phases have exactly one and two copies of DNA, respectively. Although this method can fail to distinguish cells with similar DNA content (*i.e.*, cells in G1 and early S phases or, cells

in the G2/M and late S phases), bromodeoxyuridine (BrdU) labelling can overcome this limitation. Since cells in the S phase incorporate BrdU into newly synthesised DNA, it is possible to distinguish these cells by measuring BrdU uptake [139]. The bivariate distribution of cells over DNA content and BrdU labelling can therefore be used to estimate cell-cycle distributions which are usually presented as time series data for the evolution of the fractions of cells in the G1, S or G2/M phases. These estimates can be further refined by measuring the expression of proteins involved in cell-cycle regulation, such as cyclin-dependent kinases [36, 139]. The main limitation of this approach, in contrast to FUCCI, is that cells must first be harvested and then stained. While this allows patient-derived cells to be studied, any spatial information and single-cell dynamics are lost and only population-level metrics can be extrapolated. From this point of view, as we discuss in depth in Section 1.2, mathematical models of the cell-cycle have played an important role in relating flow cytometry data to individual cells progression along the cell-cycle. In contrast, quantitative methods to analyse and explain real-time FUCCI data are less well established.

1.2.2 Population balance models

Population balance models have been successfully used to interpret flow cytometry and FUCCI experiments. While models of this form may be formulated as deterministic or stochastic systems of either ordinary (ODEs) or partial (PDEs) differential equations (or a combination of the two), they are typically based on the principle of mass balance. The simplest deterministic cell-cycle model of this form is illustrated in Fig. 1.5a, where the population of cells at time t, $N = N(t)$, is subdivided into three subpopulation N_1, N_S and N_2 corresponding, respectively, to cells in the G1, S and G2/M phases of the mitotic cycle. Applying the principle of mass balance to each dependent variable and neglecting cell death, we obtain the following system of ODEs:

$$\frac{dN_1}{dt} = 2k_2N_2(t) - k_1N_1(t), \quad t > 0, \tag{1.2a}$$

$$\frac{dN_S}{dt} = k_1N_1(t) - k_SN_S(t), \quad t > 0, \tag{1.2b}$$

$$\frac{dN_2}{dt} = k_SN_S(t) - k_2N_2(t), \quad t > 0, \tag{1.2c}$$

which we close by imposing the initial conditions:

$$N_1(0) = \pi_1, \quad N_S(0) = \pi_S, \quad N_2(0) = \pi_2. \tag{1.2d}$$

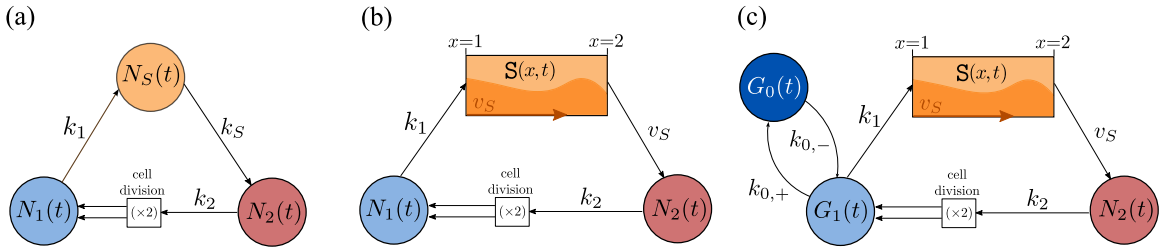


Figure 1.5: Population balance models of the cell-cycle. (a) Schematic of a simple cell-cycle model based which treats each cell-cycle phase as a well-mixed compartment. (b) A DNA-structured model of the cell-cycle that explicitly account for DNA synthesis by structuring cells in the S phase by their DNA content, x. (c) An extension of the model in panel (b) which accounts for quiescence in the G1 phase by introducing the G_0 compartment.

In Eqs. (1.2) the parameters $k_i \geq 0$ denote the constant rates of transition between compartments (*i.e.*, consecutive cell cycle phases), while $\pi_i \geq 0$ $(i = 1, S, 2)$ represent the initial number of cells in each compartment. Eqs. (1.2) can be compared directly with flow cytometry data (see Fig. 1.4b); when considering FUCCI data, the compartments may be redefined as representing florescent-green, yellow and red cells [144, 145]. While very useful in conditions of unperturbed growth, this model is no longer appropriate if cells are exposed to perturbations (such as hypoxia) that alter their progression through the cell-cycle. Such effects could be accounted for by making the transition rates k_i functions of time t. However, it is not always apparent what functional forms the coefficients should take. In this regard, introducing additional mechanisms into the model can be helpful. From our point of view, the main pitfall of an ODE modelling approach is that it does not capture DNA synthesis and, in particular limits in how fast cells can complete the cell-cycle. Imagine, for example, an experiment in which all cells are initially in the G1 phase. There will be a time delay between the beginning of the experiment and when the population size starts to increase because of the time required for cells to complete DNA replication (ranging between 10-20 hr in tumour cell-lines).

In a series of papers [10, 11, 12], Basse and coauthors developed cell-cycle models based on coupled PDEs in which cells are structured according to their DNA content. In these models, cells in the G1 and G2/M phases have constant DNA content (x), with $x = 1$ and $x = 2$, respectively, while x increases at a constant rate for cells in the S phase. The simplest model based on the framework proposed by Basse et al. is represented in Fig. 1.5b. This is identical to the previous ODE model defined by Eqs. (2.21) except for the way in which cells in the S phase are modelled. Cells

are assumed to enter the S phase with one copy of DNA ($x = 1$), to produce DNA at a constant rate $v_S > 0$ when they complete duplication (*i.e.,* $x = 2$) they enter the G2 phase. Referring to the schematic in Fig. 1.5b and applying the principle of mass balance, we obtain the following system of equations for the evolution of the dependent variables:

$$\frac{dN_1}{dt} = 2k_2N_2(t) - k_1N_1(t), \qquad t > 0, \tag{1.3a}$$

$$\frac{\partial \mathtt{S}}{\partial t} + \frac{\partial (v_S \mathtt{S})}{\partial x} = 0, \qquad t > 0,\ x \in (1,2), \tag{1.3b}$$

$$\frac{dN_2}{dt} = k_S \mathtt{S}(2,t) - k_2N_2(t), \qquad t > 0, \tag{1.3c}$$

with boundary conditions:

$$v_S \mathtt{S}(1,t) = k_1N_1, \tag{1.3d}$$

and initial conditions:

$$N_1(0) = \pi_1, \quad \mathtt{S}(x,0) = \mathtt{S}_0(x) \ \ x \in (1,2], \quad N_2(0) = \pi_2. \tag{1.3e}$$

In Eqs. (1.2) the parameters k_i, π_i $(i = 1, 2)$ and v_S are non-negative constants and $\mathtt{S}_0(x) > 0$ for all $x \in (1, 2]$. As we will discuss in Section 2.2.1.1, Eqs. (1.3) can be reduced to a system of delay differential equations if we use the method of characteristics to solve Eq. (1.3b). While Eqs. (1.2) and Eqs. (1.3) have the same number of unknown parameters, the latter requires information about the initial DNA distribution $\mathtt{S}_0(x)$. While this is readily available from flow cytometry experiments, it might be difficult to infer it from FUCCI data. In Chapter 2, we will start from Eqs. (1.2) and augment the model to account for cell-cycle dynamics in hypoxia.

Several theoretical studies have investigated the effect of chronic exposure to constant hypoxia on cell-cycle arrest [1, 23, 41, 104] with particular emphasis on hypoxia-induced G1 arrest. At the population level, quiescence is usually represented by introducing an additional phase to the standard cell-cycle: cell enter the so-called "G0" (or quiescent) phase [45, 61] prior to committing to DNA synthesis (i.e, prior to entering the S-phase) if local oxygen levels are low ($\approx 1\%\, O_2$). As shown in Fig. 1.5c, this mechanism can be introduced into our modelling framework by decomposing the N_1 compartment into two: the G_1 compartment accounts for cells that enter the S phase, while the G_0 compartment captures quiescent cells that have transiently withdrawn from the cell cycle. One might assume that the rates of transition into and out of G_0 depend on the oxygen concentration, $c = c(t)$, tp which cells are exposed to (*i.e.,*,

$k_{0\pm} = k_{0\pm}(c)$). However, as we will show in Chapter 3, cell response to acute/cyclic hypoxia is more complex, and can not be captured by this simple extension alone.

While DNA-structured models are useful for explaining flow cytometry data, they may not be the most effective framework for analysing FUCCI data. In this case, it may be preferable to construct a model that structures cells according to their "age" as well as their cell-cycle state (or phase). This has motivated the development of age-structured models [13, 32], in which the age of a cell in a certain phase of the cycle corresponds to the amount of time it has spent in that phase. Models combining both structure variables (i.e., DNA content and cell age) have also been proposed [34], but they are complex and difficult to validate against data. Cell-cycle specific properties, such as size [48], or protein expression levels [36], can also be included as additional structure variables. Population balance models can be extended to account for variability in compartment specific parameters, such as transition and/or death rates [48] and the lengths of cell-cycle phases [10], to capture the effect of different drugs on cell-cycle progression. Furthermore, spatially resolved population balance models have also been proposed to describe FUCCI data extracted from invasion assays [73, 144].

While deterministic population balance models provide effective descriptions of the dynamics of a large collection of cells, they are not appropriate to describe small populations of cells. To this end, re-framing population models in a stochastic and/or individual based (IB) framework may be more appropriate. As we will discuss in Chapter 5, these methods can be used to simulate *in silico* experiments which account for cell-to-cell variability by tracking individual cells' dynamics and fate. While different stochastic models of the cell-cycle have been proposed [3, 32, 33, 79, 153, 150], most are related, to some extent, to the *Erlang* model [47]. In this framework, the cell-cycle is described as a series of steps that proceed at some fundamental rate, which varies between cell-cycle phases. Recently, these approaches have been applied to single-cell tracking techniques based on fluorescent microscopy [32, 33, 79], to understand the temporal response of cell-cycle regulation after DNA damage [32] and identify differences in cell-cycle dynamics between normal and cancer cell lines [33]. One of the major limitations in the use of IB models is their computational costs which make it challenging to calibrate such models to experimental data. Developing alternative parameter inference approaches to overcome these limitations is an activate area of research, which encompasses *likelihood* free methods and machine learning approaches [46, 86]. While the calibration of IB models will not be addressed in this thesis, this aspect will need to be considered in the future to validate the framework and predictions from Chapters 5-6.

1.2.3 Cell-cycle models as positive dynamical systems

Deterministic population balance models follow into the class positive dynamical systems [49]. These are a subset of the most general class of *monotone dynamical systems*. Monotonicity indicates the property of a dynamical system of preserving order present in its initial conditions at all positive times. Monotone systems are usually referred to as positive systems since for any non-negative choice of initial conditions, solutions remain non-negative (to be understood here as each variable being non-negative at all times). A thorough description of the mathematical theory of monotone systems would require the introduction of several concepts from functional analysis and semigroup theory, which goes beyond the purpose of this thesis. We will here only give an intuitive explanation of the underlying theory and focus on those results that are applied in this thesis. For the reader interested in the topic of monotone non-linear systems and positive linear systems, we refer respectively to the monographs by Smith [129] and Farina and Rinaldi [49].

The term "*positive dynamical system*" refers to any type of dynamical system (*e.g.* deterministic/stochastic linear/non-linear systems of ODEs and/or PDEs with or without delays). Here we focus on the general class of linear non-autonomous dynamical systems of the form $u'(t) = \mathcal{A}(t)u(t)$, where $\mathcal{A}(t)$ is a family of bounded linear operators indexed by time and defined on a given Banach space $\mathbb{Z}$, with norm $\|\|_{\mathbb{Z}}$. We assume $\mathcal{A}$ is sufficiently smooth that its solution map $U(t,s)$ (defined so that $u(t) = U(t,s)u(s)$) is continuous and differentiable. In order to define a positive dynamical system, the space $\mathbb{Z}$ must be equipped with a defined positive cone C (for example, if $\mathbb{Z} = \mathbb{R}^n$, a common choice is $C = \mathbb{R}^n_+$). This leads naturally to the following definition of ordering: given $u_1, u_2 \in \mathbb{Z}$, $u_1 \leq u_2$ if $u_2 - u_1 \in C$, while $u_1 < u_2$ if $u_2 - u_1 \in C^+ = C \setminus \{0\}$, and $u_1 \ll u_2$ if $u_2 - u_1$ is an interior point of C. We then have that the system $u'(t) = \mathcal{A}(t)u(t)$ is positive if for all $u_s \in \mathbb{Z}$, $u_s \geq 0$ implies $U(t,s)u_s \geq 0$ for all $t \geq s$.

In this thesis, the solution $u(t)$ will be used to model the dynamics of a growing heterogeneous population of cells. Therefore, $u(t)$ will represent cell numbers or some other quantity estimating the size/mass of individuals that has physical meaning only if they take non-negative values. When the dynamical system describing the time evolution of u is derived applying the principle of mass balance (see discussion in Section 1.2), this guarantees that the equations preserve positivity of solutions. While positivity is an important property, it is not sufficient to characterise the asymptotic behaviour of the solution u. Instead, a stronger condition, *i.e.*, strong

positivity [49], is required. We note that strong positivity is not implied by the underlying physical properties of the system, *i.e.*, it is not guaranteed by applying the law of mass-balance. A linear dynamical system is said to be strongly positive if, for all $u_0 \in \mathbb{Z}$ and $u_0 > 0$, the solution $U(t,s)u_0 \gg 0$ (*i.e.*, all variables are eventually bounded away from zero). In what follows, we explain how we can prove the strong positivity of a simple cell-model taken from those presented in Section 1.1 and how strong positivity determines the model asymptotic dynamics.

Example 1.2.1. We now consider the model defined by Eqs. (1.3). As we will discuss in Section 2.2.1.1, the model can be rewritten as a delay differential equation (DDE) with constant delay, $\tau_S = v_S^{-1}$, coupled to an ODE:

$$N_1'(t) = -k_1 N_1(t) + 2k_2 N_2(t), \tag{1.4a}$$

$$N_2'(t) = k_1 N_1(t - \tau_S) - k_2 N_2(t), \tag{1.4b}$$

which we can shorten to $\boldsymbol{N}' = \mathcal{A}_0 \boldsymbol{N} + \mathcal{A}_{\tau_S} \boldsymbol{N}(t - \tau_S)$. A schematic of Eqs. (1.4) is presented in Fig. 1.6a. The natural space to define the dynamical system (1.4) is $\mathbb{Z} = \mathcal{C}([-\tau_S, 0], \mathbb{R}) \times \mathbb{R}$ with $u = (N_{1,t}, N_2)$ where $N_{1,t}(\theta) = N_1(t+\theta)$ with $\theta \in [-\tau_S, 0]$. Here, $\mathcal{C}$ is the space of continuous functions in the interval $[-\tau_S, 0]$. We equip $\mathbb{Z}$ with a maximum-norm so that $\|\boldsymbol{N}\|_{\mathbb{Z}} = \max\left\{ \max_{s\in[-\tau_S,0]} |N_{1,t}(s)|, |N_2| \right\}$, and the positive cone is defined as $\boldsymbol{N}$ such that $N_{1,t}$ are non-negative in the interval $[-\tau_S, 0]$ and $N_2 \geq 0$.

1. Eqs.(1.4) generate a positive dynamical system on $\mathbb{Z}$. Following [129, Chapter 5, Section 1], we have that Eqs. (1.4) define a positive dynamical system if the matrix $\mathcal{A} = \mathcal{A}_0 + \mathcal{A}_{\tau_S}$, where

$$\mathcal{A} = \begin{bmatrix} -k_1 & 2k_2 \\ k_1 & -k_2 \end{bmatrix}, \tag{1.5}$$

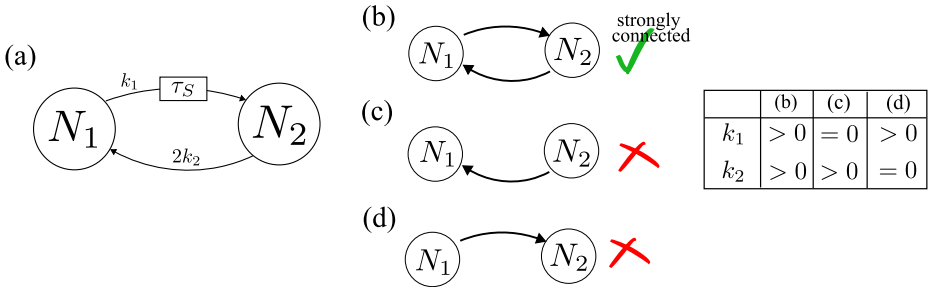
is Metzler (*i.e.*, it has all off-diagonal terms non-negative). This is indeed the case since the transition rates $k_{1,2}$ are, by definition, non-negative.

2. Eqs.(1.4) generate a strongly positive dynamical system on $\mathbb{Z}$. We now discuss the strong-positivity of Eqs. (1.4). Following standard results [129, Chapter 5], Eqs. (1.4) define a strongly positive dynamical system if the matrix $\mathcal{A}$ is irreducible. There are several ways to define and prove that a matrix is irreducible; here we will adopt an approach based on the theory of strongly connected graphs [14, Chapter 6]. We first construct the directed graph $\mathcal{G}_\mathcal{A}$ associated with matrix $\mathcal{A}$. Since the matrix $\mathcal{A}$ is constant, the resulting directed graph is time-independent. Based on Eq. (1.5), we construct the directed graph $\mathcal{G}_\mathcal{A}$ as follows:

(1) $\mathcal{G}_\mathcal{A}$ contains n_u nodes each corresponding to a dependent variable (in this case, N_1 and N_2);

(2) $\mathcal{G}_\mathcal{A}$ has an edge connecting the nodes i and j if $\mathcal{A}_{ij} > 0$.

The matrix $\mathcal{A}$ is irreducible if and only if the graph $\mathcal{G}_\mathcal{A}$ is strongly connected, *i.e.*, for any two nodes in the graph $i \neq j$ there is a path connecting them. For $\mathcal{A}$ in Eqs. (1.5), under the assumption $k_1 > 0$ and $k_2 > 0$, the graph $\mathcal{G}_\mathcal{A}$ is simple (see Fig. 1.6b) and it is apparent that this is strongly connected. We note that if any of the rates k_1 or k_2 is taken to be zero, then the corresponding graph $\mathcal{G}_\mathcal{A}$ would not be strongly connected (see Figs. 1.6c-1.6d) and, therefore, Eqs. (1.4) would just be positive but not strongly positive.



	(b)	(c)	(d)
k_1	> 0	$= 0$	> 0
k_2	> 0	> 0	$= 0$

Figure 1.6: Simple cell-cycle model with one constant delay. (a) Schematic of the model defined by Eqs. (1.4); (b)-(d) alternative graphs $\mathcal{G}_\mathcal{A}$ associated with the matrix $\mathcal{A}$ (see Eq. (1.5)) depending on different assumptions on model parameters.

In Chapter 4, we will see how the approach used in Example 1.2.1 generalises to more complex systems of autonomous and periodic differential equations with delays.

1.2.3.1 Dominant exponential solutions.

We now explain how strong positivity can be used to determine the asymptotic dynamics of a linear system of ordinary differential equations with delays (such as the system proposed in Example 1.2.1). To this end, we introduce the concept of dominant exponential solution for a solution map $U(t,s)$ [70] (see Definition 1.2.1).

Definition 1.2.1. *A solution map $(U(t,s))_{t\geq s}$ admits an exponentially dominant solution $\bar{u}(t) = e^{\lambda t}\xi$, where $\xi \in \mathbb{Z}$, $\xi \gg 0$ and $\lambda \in \mathbb{R}$. Furthermore, there exists a rank 1 projection Π_s such that for initial positive condition u_0 ($u_0 \in C^+$, where C is the positive cone on $\mathbb{Z}$), there exist two positive constants, $M > 0$ and $\epsilon > 0$:*

$$\|e^{-\lambda(t-s)}U(t,s)u_0 - c_s(u_0)\xi(s)\|_\mathbb{Z} \leq Me^{-\epsilon t}, \quad t \geq s. \tag{1.6}$$

where M and ϵ are positive constants, $c_s(u_0)$ is such that $\Pi_s u_0 = c_s(u_0)\xi(s)$.

In general, Definition 1.2.1 implies that, for any positive ($u_0 > 0$) initial conditions, on the long time scale all solutions are proportional to the same dominant exponential solution.

In general, any strongly positive autonomous system of ordinary differential equations with finite and constant delays, such as Example 1.2.1, admits a dominant solution. In particular, the existence of a dominant eigenvalue for these systems is linked to the spectral properties of the operator $\Psi = U(t+s, s)$ [53, Theorem 4.1]. This can be characterised using the *Krein–Rutman Theorem* [65, Theorem 7.2] for compact operators. Without going into the details (which can be found in [62]), we recall the concept of generalised eigenvalue and eigenspace pair, (μ, E_μ) for a linear operator, Ψ, where $\mu \in \mathbb{C}$ is a pole of order n_μ and $E_\mu = \mathrm{Ker}((\mu I - \Psi(t))^{n_\mu}) \subset \mathbb{Z}$. The Krein–Rutman Theorem guarantees that the dominant eigenvalue $\bar{\mu}$ of Ψ is real and simple so that $\bar{\mu} = e^{\lambda t}$ and that the corresponding eigenspace $E_{\bar{\mu}}$ has dimension one with basis $\{\xi\}$. Furthermore, it implies that $\xi \gg 0$ and that all other eigenvalues (μ) of $\Psi(t)$ are such that $|\mu| < \bar{\mu}$. Recalling Example 1.2.1, given that the delay τ_S is finite, this implies that $(\Psi(t))_{t\geq 0}$ is compact for any $t > \tau_S$ so that the Krein–Rutman Theorem applies. We conclude that Eqs. (1.4) admit a dominant exponential solution. Following the same argument as above, in Chapter 4 we will study the asymptotic behaviour of a $\mathcal{T}$-periodic, strongly positive cell-cycle model, consisting of ordinary differential equations with time-dependent delays. Briefly, in the periodic case, the existence of a dominant exponential solution is linked to the spectral properties of the period map $\Psi(s) = U(s+\mathcal{T}, s)$. Applying the Krein–Rutman Theorem to $\Psi(s)$ we can show that the dominant eigenvalue of the period map is real and simple. From [62, Theorem 1.1, Chapter 8], it follows that $U(t, s)$ admits a dominant exponential solution. While in the case of an autonomous dynamical system the eigenfunction $\xi \in \mathbb{Z}$ is constant, in $\mathcal{T}$-periodic systems ξ is $\mathcal{T}$-periodic, *i.e.*, $\xi(s) = \xi(s+\mathcal{T})$.

To summarise, we have shown how the theory of strongly positive dynamical system can be used to determine the asymptotic dynamics of a large class of population balance models; in particular, we find that the long time behaviour of the system is simpler than its transient dynamics, since it can be projected onto a lower-dimensional subspace in the state space of solutions ($\mathbb{Z}$).

1.3 Data driven modelling

When integrating mathematical models with *in vitro* data several challenges arise. First experimental data are often noisy and limited in quantity so that it can be diffi-

cult to identify (i.e., fitting) with sufficient certainty all unknown model parameters. *Practical identifiability* is concerned with the certainty with which model parameters can be estimated given noisy/incomplete data. *A priori* methods rely on Global Sensitivity Analysis (GSA), which estimates the relation between model outputs and inputs (*i.e.*, parameters). *A posteriori* approaches, instead, depend on the fitting technique being used. For maximum likelihood estimation, profile likelihood methods are usually adopted [118]; for Bayesian inference instead, identifiability can be directly investigated from the posterior distributions (see Section 1.3.2). In Chapter 3, we will use both GSA and Bayesian approaches to fit the cell-cycle model proposed in Chapter 2 to experimental data. In what follows, we briefly review the theory underlying these two methods and explain how they can be used to perform robust parameter inference.

1.3.1 Global Sensitivity Analysis

Sensitivity analysis (SA) is a popular quantitative method for investigating relationships between model inputs and outputs [117]. Consider a general mathematical model $\mathcal{M}$ which has n_Θ unknown parameters Θ_i ($i = 1, \dots, n_\Theta$). We denote by $\boldsymbol{\Theta} \in \mathbb{R}^{n_\Theta}$ the vector which contains unknown model inputs/parameters (which can include initial conditions if these are not known), and by $\boldsymbol{y} \in \mathbb{R}^{n_y}$ the model output. A mathematical model can generally be represented as a map between its model inputs (the parameters $\boldsymbol{\Theta}$) and outputs at a given time t:

$$\boldsymbol{y} = \mathcal{M}(t; \boldsymbol{\Theta}). \tag{1.7}$$

Sensitivity analysis is concerned with estimating how uncertainty in the input values $\boldsymbol{\Theta}$ propagates through the model output $\boldsymbol{y}$. Such information can be used for parameter screening (see Section 1.3.1.2) and parameter identifiability [117]. From this point of view, the more sensitive a measurable output is to a model input, the more easily we expect the input to be estimated. In contrast, parameters associated with small sensitivity can be changed without having a significant impact on the model outputs. These parameters are usually denoted as non-influential and they are fixed to a constant value, thus reducing the overall unknown model parameters.

Standard SA techniques can be split into two main categories: local and global methods [117]. *Local Sensitivity Analysis* (LSA) investigates the response of the model outputs to small (infinitesimal) variation in each parameter in a neighbourhood of specific point $\boldsymbol{\Theta}_0$ in parameter space. By contrast, *Global Sensitivity Analysis* (GSA) is computationally expensive but is well-suited to investigate non-linear interactions

between model parameters, and accounts for uncertainty in the prior knowledge of the parameter values. For each approach, several methods have been introduced depending on the specific aim of the analysis [117, 96]. Here we focus on *variance-based* approaches which are well-established; however, we note that promising alternative methods, such as derivative-based [82] and information-theoretic [93] sensitivity analysis, have been recently proposed.

1.3.1.1 Variance-Based Global Sensitivity Analysis (VGSA)

We now outline the mathematical framework underlining VGSA. In particular, we view the parameters $\boldsymbol{\Theta}$ as a random vector whose independent components take values in a compact subspace $\Omega = \Omega_1 \times \ldots \times \Omega_m \subset \mathbb{R}^{n_\Theta}$. We define the probability measure $\pi : \Omega \to [0, \infty)$ so that, given a subset $A \subset \mathbb{R}^{n_\Theta}$, the probability that $\boldsymbol{\Theta}$ takes a value in that subdomain is:

$$\mathbb{P}(\boldsymbol{\Theta} \in A) = \int_{A \cap \Omega} \pi(\boldsymbol{\Theta}) d\boldsymbol{\Theta} \leq 1. \tag{1.8}$$

Note that the probability distribution π captures our prior knowledge, which may be based on biological understanding or previously collected data. In the vocabulary of Bayesian inference (see Section 1.3.2), π is termed the prior distribution. When no prior information is available, the most common choice of π is a uniform distribution [96]. We will focus on this case, so that $\boldsymbol{\Theta}$ is uniformly distributed on the hyper-rectangle Ω, i.e $\boldsymbol{\Theta} \sim U(\Omega)$, where the components of the vector $\boldsymbol{\Theta}$ are assumed to be independent random variables. Without loss of generality, we assume Ω to be the hypercube $[0, 1]^{n_\theta}$ so that $\pi(\Theta) \equiv I_\Omega(\Theta)$, where $I_\Omega(\cdot)$ represents the indicator function. Starting from (1.7), we define the random variable $\boldsymbol{h}(\boldsymbol{\Theta}; t) = \mathcal{M}(t; \boldsymbol{\Theta})$, where we assume the function $\boldsymbol{h} : \Omega \to \mathbb{R}^{n_y}$ to be square-integrable (i.e., $\boldsymbol{h} \in \mathcal{L}^2(\Omega)$). We introduce the following decomposition:

$$\boldsymbol{h}(\boldsymbol{\Theta}; t) = \boldsymbol{h}_0 + \sum_{i=1}^{n_\Theta} \boldsymbol{h}_i(\Theta_i; t) + \ldots + \boldsymbol{h}_{12\ldots n_\Theta}(\Theta_1, \ldots, \Theta_{n_\Theta}; t) \tag{1.9}$$

where $\boldsymbol{h}_{i_1 \cdots i_s} : [0, 1]^s \to \mathbb{R}^n$, $1 \leq i_1 < \ldots < i_s \leq n_\Theta$, are also square integrable functions. If we impose the additional constraint on the expected value of the functions $\boldsymbol{h}_{i_1 \cdots i_s}$ with $i_k \in \{i_1, \ldots, i_s\}$, $\quad \mathbf{0} \in \mathbb{R}^n$, then the representation (1.9) is unique [132]. If we denote by $\mathcal{I}_s = \{i_1, \ldots, i_s\}$ the subset of the indices $\mathcal{I} = \{1, \ldots, n_\Theta\}$, then the *partial variances* can be defined as follows [132]:

$$\boldsymbol{V}_{\mathcal{I}_s}(t) = \int_{[0,1]^s} \boldsymbol{h}_{\mathcal{I}_s}(\boldsymbol{\Theta}_{\mathcal{I}_s}; t) \odot \boldsymbol{h}_{\mathcal{I}_s}(\boldsymbol{\Theta}_{\mathcal{I}_s}; t) \prod_{i \in \mathcal{I}_s} d\Theta_i = \mathrm{Var}[\boldsymbol{h}_{\mathcal{I}_s}](t), \tag{1.10a}$$

where $\odot$ is the Hadamard product (*i.e.*, element-wise multiplication). If we take the variance of both sides of Eq. (1.9) and divide by the left hand side, we obtain the following identity:

$$\mathbf{1}_{n\times 1} = \sum_i \boldsymbol{S}_i(t) + \sum_i^{n_\Theta}\sum_{j>i}^{n_\Theta} \boldsymbol{S}_{ij}(t) + \ldots + \boldsymbol{S}_{12\ldots n_\Theta}(t), \tag{1.11a}$$

$$\text{where } \boldsymbol{S}_{\mathbb{I}_s}(t) = \text{Var}\left[\boldsymbol{h}_{\mathbb{I}_s}\right](t) \oslash \text{Var}[\boldsymbol{h}](t), \tag{1.11b}$$

where $\oslash$ denotes the Hadamard division (*i.e.*, element-wise division). In Eqs. (1.11), the constant coefficients $\boldsymbol{S}_{\mathbb{I}_s} \in \mathbb{R}^n$ are called *Sobol indices*; they represent the fraction of the total variance associated with the variation of the subset of parameters $\{\Theta_{i_1}, \ldots, \Theta_{i_s}\}$. In other words, the Sobol indices capture the expected fraction of variance reduction in the output if the values of the parameters with indices in $\mathbb{I}_s$ were fixed [123]. The collection of all Sobol Indices gives a detailed picture of the different order of interactions between parameters, but their estimation becomes computationally expensive and intractable as the dimension of the parameter space n_Θ increases. To overcome these limitations, the *Total Sensitivity Index* (TSI) $\boldsymbol{S}_{T_i}$ has been introduced to provide information about all order interactions between parameters. The TSI is defined as [68]:

$$\boldsymbol{S}_{T_i}(t) = \boldsymbol{S}_i(t) + \ldots + \boldsymbol{S}_{12\ldots n_\Theta}(t) = \mathbb{E}[\text{Var}[\boldsymbol{h}|\boldsymbol{\Theta}_{\mathbb{I}^{-i}}]](t) \oslash \text{Var}[\boldsymbol{h}](t), \tag{1.12}$$

where $\boldsymbol{\Theta}_{\mathbb{I}^{-i}}$ indicates all components of the vector $\boldsymbol{\Theta}$ except for the i-th one. The TSI accounts for the contribution of the i-th parameter to the output, i.e. $\boldsymbol{S}_i$, plus the effect due to its non-linear interactions with all other parameters. In other words, this metric captures the 'total' effect of the parameter Θ_i, which includes the variance associated with Θ_i and the variance due to interactions of Θ_i the other parameters. The main advantage of $\boldsymbol{S}_{T_i}$, compared to computing higher order sensitivity indices, is the simplicity with which it can be estimated [123].

Several approaches have been proposed to numerically estimate Eq. (1.12). These approaches differ in the strategy used to efficiently sample the parameter space Ω. *Quasi-Monte Carlo* methods and the Fourier Amplitude Sensitivity Test (eFAST) [122] are the most commonly used. In this thesis, we will rely on Monte Carlo approaches to estimate higher order integrals by exploiting random sampling of the space Ω. Given a sequence of n_Θ-dimensional random vectors, $\{\boldsymbol{a}_i\}_{i=1}^{n_S}$ with $\boldsymbol{a}_i \in \Omega$, the Monte Carlo estimate of the integral of the function f in Ω is given by the quadrature rule [82, 131]:

$$\int_\Omega f(\boldsymbol{s})\, d\boldsymbol{s} \sim \frac{1}{n_S}\sum_{i=1}^{n_S} f(\boldsymbol{a}_i). \tag{1.13}$$

This approach can be extended to the use of *pseudorandom* sequences, where the sampling of the points $\boldsymbol{a}_i$ optimises the exploration of the domain Ω [82, 131]. Since the resulting methods are not completely random, they are termed *Quasi-Monte Carlo* methods. For a hypercube, i.e. $\Omega = [0,1]^{n_\Theta}$, a well-known sampling technique is the *Sobol* sequence [130]. As discussed in [82], this approach outperforms standard MC methods and yields a faster convergence rate. Details on how *Sobol* sequence sampling can be adapted to the computation of Sobol indices can be found in [96]. One of the key unknowns then setting up sensitivity analysis experiments is the proper number of samples n_S to use. We say that GSA has converged if its results do not change significantly if a larger number of samples n_S is used. By using QMC sampling, we can estimate the 95% confidence interval for each estimate of a sensitivity index and use the latter as a proxy for convergence by requiring it to be sufficiently small.

Having described how VGSA is performed, in what follows we describe how its outcomes can be used to reduce a complex model to a simpler formulation that is better suited to perform robust parameter fitting.

1.3.1.2 Parameter screening

The goal of parameter screening techniques is to classify model parameters depending on whether they influence model outputs. This information can be useful in reducing the complexity (*i.e.*, the number of unknown parameters) by indicating which parameters can be fixed without influencing the predictive power of the model (*i.e.*, the range of possible predicted outcomes). In Chapter 3, this procedure will allow us to decrease the number of unknown parameters in our cell-cycle model from 12 to 6.

To start with, it is necessary to define the conditions under which a parameter is classified as non-influential. This usually relies on a threshold-based approach: if the total sensitivity index is below a threshold, $\bar{S}$, then the parameter is classified as non-influential. When considering time dependent sensitivity analysis, this approach can be generalised by imposing a threshold on the maximum value of the total sensitivity index recorded over time and across multiple outputs: $\max_{t \in U_t} \|\boldsymbol{S}_T(t)\|_\infty < \bar{S}$, where U_t is the time interval along which the sensitivity indices are computed. However, there is no pre-defined way to choose an appropriate value for $\bar{S}$. This will usually depend on the problem or application of interest. In this thesis, we use GSA for *a priori* screening of those parameters that will not be identifiable given experimental data. In a perfect scenario, where experimental data are noise-free, we expect non-identifiable parameters to have zero sensitivity indices and we should fix $\bar{S} = 0$. In practice, experimental data are noisy so we fix $\bar{S} = 0.05 \ll 1$. Given that the

choice of $\bar{S}$ is arbitrary, we seek an approach to validate our screening procedure. Several statistical methods have been proposed (see [124] and references therein). Here we adopt an approach similar to that proposed in [113]. The main idea is to generate two parameter samples: the first is generated by sampling the entire parameter space, while the second fixes the non-influential parameters and samples only parameters identified as influential. Running the model forward for each parameter sample yields two output samples, named non-conditional ($\mathcal{Y}$) and conditional ($\mathcal{Y}|\boldsymbol{\theta}$), respectively. Testing the validity of the screening procedure is then reformulated as estimating whether the discrepancy between these two output samples is statistically significant. To this end, Pianosi *et al.* [113], adopt the Kolmogorov–Smirnov statistic (KS statistic) which compares the cumulative distribution functions of $\mathcal{Y}$ and $\mathcal{Y}|\boldsymbol{\theta}$ to determine whether the two samples are generated from the same distribution. This approach is only valid for one-dimensional output samples. For time series data and multiple outputs, as considered in this thesis, a different statistical test is required. We employ the analytic mean embeddings test first proposed by Chwialkowski *et al.* [37] and later optimised in [74]. The main idea is to approximate the probability distribution for $\mathcal{Y}$ and $\mathcal{Y}|\boldsymbol{\theta}$ via a class of real valued analytic functions and evaluate the distance between their embeddings (*i.e.*, their representation in the space of analytic functions) at a specified number of points J. The Hotelling's T-squared statistic, $\hat{\beta}_n$, is then used to decide whether or not to reject the null hypothesis (H_0), that is: $\mathcal{Y}$ and $\mathcal{Y}|\boldsymbol{\theta}$ are samples from the same probability distribution. If H_0 is true then, as the number of points in the sample increases, $\hat{\beta}$ approaches a chi-squared distribution with J degrees of freedom. Therefore, the null hypothesis is rejected if $\hat{\beta} > T_\alpha$, where the test threshold T_α is given by the $(1-\alpha)$-quantile of the chi-squared distribution with J degrees of freedom. The improvement proposed in [74] is based on optimising the J location at which the distance is evaluated so as to maximise the test power.

1.3.2 Bayesian parameter inference

Thus far, we have explained how GSA can be used to yield mathematical models that can be more effectively calibrated to data. We now review the parameter fitting techniques used in this thesis (specifically, in Chapter 3).

Bayesian inference methods can be used to generate probability distributions rather than point estimates of parameter values by accounting for the variability intrinsic in the experimental data. By quantifying the uncertainty in the estimates of model parameters, it also allows for direct assessment of parameter identifiability.

While not directly relevant for our applications, a major limitation of Bayesian inference methods is their high computational cost. This precludes their application to models with large numbers of unknown parameters, or models that are intrinsically computationally expensive to simulate (such as stochastic and/or agent based models).

1.3.2.1 The observation model

As already mentioned, Bayesian methods use a probabilistic approach by viewing the data as the realisation of a stochastic process that has to be specified. Let us now denote by $\boldsymbol{y}(t;\boldsymbol{\Theta},\mathcal{M}) \in \mathbb{R}^{n_y}$ the value of the model output at time t, as predicted by the model $\mathcal{M}$ for parameter values $\boldsymbol{\Theta} \in \mathbb{R}^{n_\Theta}$. We consider the experimental observations to be noisy realisations of the model solutions, $\boldsymbol{y}(t;\boldsymbol{\Theta},\mathcal{M})$. We denote by $\mathcal{E} = \{\tilde{\boldsymbol{y}}_m\}_{m=1}^{n_\mathcal{E}}$ the collection of experimental data taken at times t_m, and we make no special requirements on the distribution of these times. The likelihood function $\mathcal{L}_\mathcal{M}(\mathcal{E}|\boldsymbol{\Theta})$ represents the probability of observing the experimental data $\mathcal{E}$ under the model $\mathcal{M}$ with parameter values $\boldsymbol{\Theta}$. If we assume, as common, that the noise associated with each measurement is independent of other measurements, then:

$$\mathcal{L}_\mathcal{M}(\mathcal{E}|\boldsymbol{\Theta},\boldsymbol{\sigma}) = \prod_{m=1}^{n_\mathcal{E}} \phi\left(\tilde{\boldsymbol{y}}_m; \boldsymbol{y}(t_m;\boldsymbol{\Theta},\mathcal{M}),\boldsymbol{\sigma}\right), \tag{1.14}$$

where ϕ is a prescribed probability distribution, with mean $\boldsymbol{y}(t_m;\boldsymbol{\Theta},\mathcal{M})$. In Eq. (1.14), $\boldsymbol{\sigma}$ denotes parameters that characterise the function ϕ (commonly referred to as hyper-parameters). For example, a standard choice for the function ϕ is a multivariate normal probability distribution; then the hyper-parameters $\boldsymbol{\sigma}$ would be the vector formed by the entries of the square (upper-triangular) matrix $U \in \mathbb{R}^{n_y \times n_y}$ associated with the covariance matrix $\Sigma = U^T U$. If there is uncertainty in the properties of the noise, then the parameter $\boldsymbol{\sigma}$ can also be taken to be unknown. This choice introduces additional unknown parameters but allows for a full Bayesian treatment, where both the deterministic model, $\mathcal{M}$, and the observation model are estimated based on the data.

At the basis of Bayes approaches lies Bayes' Theorem (Eq. (1.15)). Given a model $\mathcal{M}$, any prior information about its parameter values is captured by the prior distribution, $\pi(\boldsymbol{\Theta},\boldsymbol{\sigma})$. Using Bayes' Theorem, we can update our prior distributions ($\pi(\boldsymbol{\Theta},\boldsymbol{\sigma})$) in light of experimental data as follows:

$$\pi_\mathcal{M}(\boldsymbol{\Theta},\boldsymbol{\sigma}|\mathcal{E}) = \frac{\pi(\boldsymbol{\Theta},\boldsymbol{\sigma})\mathcal{L}_\mathcal{M}(\mathcal{E}|\boldsymbol{\Theta},\boldsymbol{\sigma})}{\pi(\mathcal{E})}, \tag{1.15}$$

where $\pi(\mathcal{E})$ is a normalising factor, $\mathcal{L}_{\mathcal{M}}$ is the likelihood function and $\pi_{\mathcal{M}}(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E})$ is the posterior distribution for model $\mathcal{M}$. While for actual computation, the normalisation factor can be neglected, the likelihood function plays a key role as it captures assumptions about the stochastic process that generates the observed data.

1.3.2.2 Markov Chain Monte Carlo Methods

Except in very simple cases, due to the dimensions of the space of unknown parameters, it is not possible to compute Eq. (1.15) analytically. Instead, Markov Chain Monte Carlo (MCMC) methods are widely used to estimate $\pi_{\mathcal{M}}(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E})$. These methods are based on drawing values of the unknown parameters $(\boldsymbol{\theta}) = (\boldsymbol{\Theta}, \boldsymbol{\sigma})$ from approximate distributions and then correcting the draws to better approximate the target posterior distribution, $\pi_{\mathcal{M}}(\boldsymbol{\theta}|\mathcal{E})$. The output of this procedure is, therefore, a sequence of parameter values (or a chain) $\left\{\boldsymbol{\theta}^{(i)}\right\}_{i=1}^{n_c}$ where $n_c \gg 1$. As $n_c \to \infty$, the terms $\boldsymbol{\theta}^{(i)}$ are drawn from a distribution that better approximates the target posterior $\pi_{\mathcal{M}}(\boldsymbol{\theta}|\mathcal{E})$. Due to the asymptotic convergence of the methods, the first portion of a chain is disregarded and only the sub-chain $\left\{\boldsymbol{\theta}^{(i)}\right\}_{i=n_0}^{n_c}$, where n_0 is chosen arbitrarily such that $1 \ll n_0 < n_c$. Statistics that test for the convergence of MCMC methods can be applied to decide when to halt sampling (*i.e.*, when n_c is big enough); this will be discussed in Chapter 3. A common feature of the MCMC methods proposed in the literature is that the distribution used to draw a new parameter sample, T_{i+1} (also known as the transition distribution), has the *Markov property* (*i.e.*, it only depend on the last draw $\boldsymbol{\theta}^{(i)}$: $T_{i+1} = T_{i+1}(\boldsymbol{\theta}^{(i+1)}|\boldsymbol{\theta}^{(i)})$). In most methods, T_i depends on the iterate number i. What distinguishes existing MCMC methods is the choice of the transition probability T_i used to sample the parameter space. For a general review of MCMC methods, we refer to the standard textbooks by Gelman [56] and Lambert [85] and the documentation associated with the Python package PINTS (Probabilistic Inference on Noisy Time-Series) for Bayesian inference [38], which we use in this thesis. More details on the approach used for parameter estimates will be mentioned in Chapter 3.

1.3.3 Bayesian model selection

The aim of model selection techniques is to identify the (statistical) model that best captures the given experimental data amongst a set of candidate models. Classical approaches to model selection rely on quantitative metrics (or criteria) that allow comparison of fits to different models. The goal is to define a summary statistic that

estimates the predictive accuracy of a fitted model ($\mathcal{M}$), *i.e.*, how well the model would predict observations ($\tilde{\boldsymbol{y}}_{new}$) not included in the fitting procedure:

$$\begin{aligned}\text{prediction accuracy}(\mathcal{M}) &= \log\left[\pi_{\mathcal{M}}(\tilde{\boldsymbol{y}}_{new}|\mathcal{E})\right] \\ &= \log\left[\int_{\mathbb{R}^{n_\Theta+n_\sigma}} \mathcal{L}_{\mathcal{M}}(\tilde{\boldsymbol{y}}_{new}|\boldsymbol{\Theta},\boldsymbol{\sigma})\pi_{\mathcal{M}}(\boldsymbol{\Theta},\boldsymbol{\sigma}|\mathcal{E})d\boldsymbol{\Theta}d\boldsymbol{\sigma}\right],\end{aligned} \tag{1.16}$$

where $\pi_{\mathcal{M}}(\tilde{\boldsymbol{y}}_{new}|\mathcal{E})$ is known as the *posterior predictive distribution*. Several methods have been proposed to estimate Eq. (1.16), and they are usually tailored to the type of fitting approach used and the available data. In general, these methods can be split into two classes: information criteria and cross-validation approaches. The latter use a subset of the full data-set, $\mathcal{E}$, for fitting, and the remaining data (that are unseen) to estimate the prediction accuracy (see Eq. (1.16)). This approach can be problematic when considering time series data (as in this thesis), since it can be difficult to estimate model parameters when some data are withheld from the training set, for validation. Therefore, we focus on information criteria (IC) which estimate prediction accuracy using the same data employed for the fitting. This approach introduces biases into the estimate of model accuracy ($\mathtt{PA}_{IC}$) which are corrected for via a correction term (k_{IC}). Hence, all criterion take the form $\text{IC}(\mathcal{M}) = -2(\mathtt{PA}_{IC} - k_{IC})$, where the -2 term is by convention. When comparing models, we are interested in the relative value of the IC, and favour the model with the smallest IC. We note that the IC penalises more complex models because they are more likely to fit data well. More complex models tend to have more parameters which can lead to higher posterior uncertainty if the model is too complex for the data (i.e., it is over-fitted) [85]. When considering Bayesian model selection, the deviance information criterion (DIC) and Watanabe-Akaike information criterion (WAIC) are amongst the most common IC [56, 85]. Given a model $\mathcal{M}$, the DIC is defined as follows [85]:

$$\text{DIC}(\mathcal{M}) = 2k_{DIC}(\mathcal{M}) - 2\ln(\mathcal{L}_{\mathcal{M}}(\mathcal{E}|\hat{\boldsymbol{\Theta}})), \tag{1.17a}$$

$$\text{with } k_{DIC}(\mathcal{M}) = 2\text{Var}[\ln\mathcal{L}_{\mathcal{M}}|\mathcal{E}], \tag{1.17b}$$

where $\mathcal{L}_{\mathcal{M}}(\mathcal{E}|\hat{\boldsymbol{\Theta}},\hat{\boldsymbol{\sigma}})$ is the likelihood value estimated at the expected value of the unknown parameters (*i.e.,* $\hat{\boldsymbol{\Theta}}$ and $\hat{\boldsymbol{\sigma}}$). Its components are given by

$$\hat{\Theta}_i = \int_{\mathbb{R}^{n_\Theta+n_\sigma}} \Theta_i\,\pi_{\mathcal{M}}(\boldsymbol{\Theta},\boldsymbol{\sigma}|\mathcal{E})d\boldsymbol{\Theta}d\boldsymbol{\sigma}, \qquad i = 1,\ldots,n_\Theta, \tag{1.17c}$$

$$\hat{\sigma}_i = \int_{\mathbb{R}^{n_\Theta+n_\sigma}} \sigma_i\,\pi_{\mathcal{M}}(\boldsymbol{\Theta},\boldsymbol{\sigma}|\mathcal{E})d\boldsymbol{\Theta}d\boldsymbol{\sigma}, \qquad i = 1,\ldots,n_y. \tag{1.17d}$$

In Eq. (1.17b), Var is the variance of the log-likelihood, $\ln\mathcal{L}_{\mathcal{M}}$, here approximated via sampling from the estimated posterior. The WAIC improves on the DIC by

accounting for the uncertainty in the prediction of model accuracy, $PA_{\mathcal{M}}$, that is defined as follows:

$$\mathrm{WAIC}(\mathcal{M}) = 2k_{WAIC}(\mathcal{M}) - 2\ln(\hat{\mathcal{L}}_{\mathcal{M}}(\tilde{\boldsymbol{y}}_m)), \tag{1.18a}$$

$$\text{with } k_{WAIC}(\mathcal{M}) = \sum_{m=1}^{M} \mathrm{Var}[\ln \mathcal{L}_{\mathcal{M}}(\tilde{\boldsymbol{y}}_m)|\mathcal{E}], \tag{1.18b}$$

where $\hat{\mathcal{L}}_{\mathcal{M}}(\tilde{\boldsymbol{y}}_m)$ is the average likelihood to observe the data point $\tilde{\boldsymbol{y}}_m \in \mathcal{E}$:

$$\hat{\mathcal{L}}_{\mathcal{M}}(\tilde{\boldsymbol{y}}_m) = \int_{\mathbb{R}^{n_\Theta + n_\sigma}} \mathcal{L}_{\mathcal{M}}(\boldsymbol{y}_m|\boldsymbol{\Theta}, \boldsymbol{\sigma}) \pi_{\mathcal{M}}(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E}) d\boldsymbol{\Theta} d\boldsymbol{\sigma} \tag{1.18c}$$

Compared to the DIC, which uses point estimates to condition the log-likelihood, the WAIC averages over the posterior distribution. For this reason, the WAIC is preferred in a Bayesian context to estimate predictive accuracy; however, in the literature the DIC is more commonly adopted, since it is less computationally expensive to estimate. In Chapter 3, we use both approaches to obtain more reliable conclusions.

Information criteria can also be used to compare models with different likelihoods (*i.e.*, different observational models). However, this may require rescaling of the likelihood function. In particular, while $\mathcal{L}$ is invariant to changes in parametrization of $\boldsymbol{\Theta}$, it is not invariant under transformation of the data $\tilde{\boldsymbol{y}}_m$. Suppose, for example, that we compare two models which specify, respectively, the distribution of the observables $\tilde{\boldsymbol{y}}_m$ and their transformation $\boldsymbol{\Psi}(\boldsymbol{y}_m) \in \mathbb{R}^M$ (with Ψ being a non-linear bijective map). Then the proper way of defining the likelihood functions for the two observational models (denoted as $\mathcal{L}^{(1)}_{\mathcal{M}}$ and $\mathcal{L}^{(2)}_{\mathcal{M}}$) is:

$$\mathcal{L}^{(1)}_{\mathcal{M}}(\tilde{\boldsymbol{y}}_m|\boldsymbol{\theta}, \boldsymbol{\sigma}) = \mathcal{L}_{\mathcal{M}}(\tilde{\boldsymbol{y}}_m|\boldsymbol{\theta}, \boldsymbol{\sigma}), \tag{1.19a}$$

$$\mathcal{L}^{(2)}_{\mathcal{M}}(\tilde{\boldsymbol{y}}_m|\boldsymbol{\theta}, \boldsymbol{\sigma}) = \mathcal{L}_{\mathcal{M}}(\boldsymbol{\Psi}(\tilde{\boldsymbol{y}}_m)|\boldsymbol{\theta}, \boldsymbol{\sigma}) \det(J_{\boldsymbol{\Psi}}(\tilde{\boldsymbol{y}}_m)), \tag{1.19b}$$

where $J_{\boldsymbol{\Psi}}$ denotes the Jacobian of the transformation Ψ.

1.4 Thesis structure

The remainder of this thesis comprises 5 research chapters in which we investigate how different forms of hypoxia affect fundamental cellular mechanisms. In Chapters 2-4, we develop and analyse a deterministic mathematical model to interpret novel experimental data on the dysregulation of cell-cycle progression in different form of hypoxia [5]. Part of this work is published in [29]. In Chapters 5 and 6, we leverage a stochastic, individual based formulation of the model developed in the earlier part

of the thesis to investigate how different forms of hypoxia influence the survival of cancer cells in the presence and absence of treatment, specifically radiotherapy.

In **Chapter 2**, we develop a mechanistic model to describe the growth of a population of cancer cells structured according to their cell-cycle state. This approach allows us to relate mechanisms that impact cancer cells progression through the cell-cycle in hypoxia to the overall population level dynamics. We then perform numerical simulations to explore model predictions under different environmental conditions: normoxia, constant hypoxia and a range of cyclic hypoxia protocols. Our analysis reveals a rich set of behaviours and highlights the roles of specific cell-cycle control mechanisms in dictating how different oxygen environments affect tumour growth.

In **Chapter 3**, we combine the modelling framework presented in **Chapter 2** with experimental data from [5] for a colorectal RKO cancer cell-line to characterise cell-cycle dynamics during short-term exposure to cyclic hypoxia. We first use global sensitivity analysis to construct a reduced version of the model presented in **Chapter 2**. The reduced model arises from a separation of timescales associated with different mechanisms included in the original model. Using *Bayesian* inference methods to fit the proposed (reduced) model to data, we show this can explain the cell-cycle dynamics of RKO cells in both constant and cyclic hypoxia. We also use *Bayesian* model selection techniques to compare alternative mechanistic models that describe cell-cycle dynamics in hypoxia. In this way, we reveal the importance of memory effects, included in our model but previously neglected in the literature, in driving the experimentally observed cell-cycle dynamics in cyclic hypoxia. Furthermore, we show how our calibrated model can be used to inform the design of new experiments to yield novel insights on the cell-cycle dysregulation in cyclic hypoxia.

While experiments are limited in the time-scale that they can resolve, our model can be exploited to make predictions on the long term effects of hypoxia on the growth dynamics of cell cultures. In **Chapter 4**, we analyse the asymptotic behaviour of the model leveraging the theory of positive linear dynamical systems introduced in Section 1.2.3. This allows us to validate and extend the results obtained via numerical simulations in **Chapter 2**. In doing so, we derive information about the predicted growth dynamics in constant and cyclic hypoxia. Our analysis quantitatively relates cell-cycle control mechanisms to the ability of cancer cells to sustain growth in periodically hypoxic environments. In particular, we investigate environmental conditions which favour cells that maintain or lose checkpoint regulation.

While in **Chapters 2-4** we focus on the impact of cyclic hypoxia on tumour cell proliferation, in **Chapter 5 and 6**, we use our modelling framework to investigate

how dysregulation of the cell-cycle in hypoxia affects cell survival in the absence or presence of treatment. State-of-the-art experimental approaches to study cancer cell survival are based on clonogenic assays. In **Chapter 5**, we develop a computational framework to simulate clonogenic assays *in silico*, thus offering a new tool to study survival in a population of cells structured by their cell-cycle state. Our framework is based on the stochastic, individual-based version of the deterministic model proposed in **Chapter 2**. We start by comparing the predictions of the individual-based model with those of the deterministic model. By doing so, we illustrate why stochastic, in contrast to deterministic, models are better suited to study clonogenic assay experiments. We then use our framework to simulate the outcome of clonogenic assays performed in constant and in cyclic hypoxia. The results in **Chapter 5** suggest that a cancer cell ability to survive hypoxia depends on the cell-cycle regulation strategy it adopts. Interestingly, we find that, for a range of cyclic hypoxia conditions, mechanisms that offer cancer cell a proliferation advantage may simultaneously severely impair their chances of survival.

Based on experimental evidence, cyclic hypoxia can result in increased resistance to standard treatments, such as chemo- and radio-therapy, both *in vivo* and *in vitro*. It is well-known that cell-cycle dysregulation can have severe impact on cancer cell survival to treatment. In **Chapter 6**, we extend the computational framework developed in **Chapter 5** to study the extent to which cell-cycle dysregulation in cyclic hypoxia may impact cancer cell sensitivity to radiotherapy. We show how such an approach can be exploited to predict and design clonogenic assays that more accurately quantify changes in the radio-sensitivity of cancer cells pre-treated with cyclic hypoxia.

We conclude the thesis in **Chapter 7** where we present a summary of the main results and outline potential extensions of this thesis. Future directions include the development of a damage-structured model to study proliferation-survival trade-offs in cyclic hypoxia and extending our cell-cycle models to include spatial, as well as temporal, variation in oxygen levels.

Chapter 2

A mathematical model of cell-cycle progression in hypoxia

2.1 Overview

Despite evidence that oxygen levels in *in vivo* tumours are dynamic [4], most experimental and theoretical research has focused on constant hypoxia and relatively little is known about the effect that cyclic hypoxia has on tumour progression. Recently, thanks to the development of *in vitro* models of cyclic hypoxia, the number of experimental studies that investigate cyclic hypoxia has increased and we are now starting to understand the biological mechanisms that drive cell responses to fluctuating oxygen levels [4, 103, 125]. Interestingly, these preliminary studies show that responses of cells to cyclic and constant hypoxia are different. For example, the recent study by Bader et al. [5] has shown how cyclic and constant hypoxia differentially perturb cancer cell progression along the mitotic cell-cycle.

As discussed in Section 1.1.3, experimental evidence [5, 52, 114] suggests that the following two mechanisms are driving cell-cycle dynamics under different from of hypoxia: shortage in deoxynucleotides (dNTPs) – the building blocks of DNA, which leads to replication stress, and down-regulation of proteins associated with the DNA-damage response (DDR), which orchestrate cell repair and survival under replication stress. While it is possible to identify these mechanisms *in vitro*, the number of oxygen cycles that can be tested is limited by technical considerations and time constraints. In principle, infinitely many cyclic protocols could be designed by varying how long, and how often, cells are exposed to hypoxia. From this point of view, mathematical modelling can play a key role in exploring a wide range of environmental conditions, and generalising experimental findings that may be constrained to one cyclic protocol. However, there are no mathematical models in the literature that

can fully capture the available data on cell-cycle progression in fluctuating oxygen levels [5]. As discussed in Section 1.2, this is because most existing studies focus on constant (chronic) hypoxia, under which some of the mechanisms listed above are negligible.

In this chapter we develop and study a model of the cell-cycle that describes the growth of population of non-competing cells under different hypoxic environments (either constant or cyclic conditions). In Section 2.2, we present our 6-compartment cell-cycle model and detail how we account for the impact of transient and/or constant hypoxia on cell-cycle progression and cell viability. This model is a more general version of the one proposed in [29] to study cell-cycle dysregulation under acute (brief) exposure to constant and cyclic hypoxia. In Chapter 3, we will use model reduction techniques to show how the model analysed in [29] can be derived systematically from the model here proposed. In Sections 2.3 and 2.4, we use numerical simulation to explore model predictions in different oxygen environments: well-oxygenated conditions (Section 2.3), constant and cyclic hypoxia (Section 2.4). While normal cells have functional checkpoint, the growth advantage of cancer cells usually stems from having defective control of cell-cycle transitions. We use our model to study environmental conditions that confer a growth advantage to cells with one or more defective checkpoints. We conclude by summarising our findings in Section 2.5.

2.2 Model development

We are interested in describing the growth of cell cultures exposed to fluctuating oxygen environments, in the absence of competition for space. While this is consistent with the initial phase of cell culture growth, it does not accurately describe cell culture growth near confluence or *in vivo* conditions. We will discuss possible model extensions that address these limitations in Chapter 7, as part of future work. We model cell cultures as well-mixed populations and assume that all cells experience the same oxygen levels, $c = c(t)$, which are externally prescribed. As specified in Section 1.1, in this thesis, we define hypoxia as oxygen levels resulting in replication stress (*i.e.,* oxygen levels below the threshold $c_H \approx 1\%\, O_2$).

We model the impact of hypoxia on cell proliferation and viability by augmenting the simple 3-compartment DNA-structured model from Section 1.2.2 (see Eqs. (1.3)). There are two key mechanisms via which hypoxia affects cell-cycle progression (see Fig. 1.1d): (a) shortage in deoxyribonucleotides (dNTP), which are essential for the

onset and progression of DNA synthesis, and (b) down-regulation of proteins associated with the DNA-damage response (DDR), which orchestrate cell repair and survival under replication stress. To account for these effects, we introduce 3 additional model compartments to our simple cell-cycle model (*i.e.*, Eqs. (1.3)) and allow transition rates between compartments to evolve over time in response to oxygen levels (c) and the expressions level of repair proteins (m). While oxygen levels, $c(t)$, are externally prescribed, we explicitly describe the dynamics of $m(t) \geq 0$. In particular, based on experimental data [115], we assume $m(t)$ evolves according to a piecewise linear ODE:

$$\begin{aligned} \frac{dm}{dt} &= \begin{cases} -\mathcal{R}_M^- \left(m - M^-\right), & c(t) < c_H, \\ \mathcal{R}_M^+ \left(1 - m\right), & c(t) \geq c_H, \end{cases} \\ m(0) &= m_0. \end{aligned} \tag{2.1}$$

We acknowledge Eq. (2.1) is possibly the simplest choice for the dynamics of m, which is consistent with experimental data [115]. Alternative formulations could be introduced if required. In Eqs. (2.1), $\mathcal{R}_M^- > 0$, $\mathcal{R}_M^+ > 0$, $m_0 \in [0, 1]$ and $M^- \geq 0$ are constants. The rates $\mathcal{R}_M^+$ and $\mathcal{R}_M^-$ determine the timescales for re-equilibration of protein levels upon re-oxygenation and hypoxia respectively. In Eqs. (2.1), $m = 1$ corresponds to physiological expression levels and $m = M^-$ indicates pathological expression levels in hypoxia. In Section 2.2.7, we will discuss how most parameters in Eq. (2.1) can be obtained from existing experimental data.

Given the time-evolution of the two scalar fields c and m, our cell-cycle model consists of a 6-compartment (non-autonomous) partial differential equation (PDE) model. Fig. 2.1a presents a schematic of our cell-cycle model. Since this is quite complex, we decompose it into three sub-models corresponding to the 3 phases of the cell-cycle: the G1 (Fig. 2.1d), the S (Fig. 2.1b) and the G2/M (Fig. 2.1c) phases. We first introduce each sub-model independently, assuming that the influxes of cells in each cell-cycle phase, $I_{G1,S,G2}(t)$ (see dashed arrows in Fig. 2.1), are known. In Section 2.2.4, we connect the different sub-models by balancing fluxes in and out of consecutive cell-cycle phases. The governing equations for the full model are then summarised in Section 2.2.5.

2.2.1 A model of the S phase in hypoxia

As in Section 1.2, we structure cells in the S phase according to their DNA content, $x \in [1, 2]$ and denote by $\mathtt{S} = \mathtt{S}(x, t)$ the number of cells in the S phase with DNA content x at time t. In accordance to the schematic in Fig. 2.1b, the time-evolution

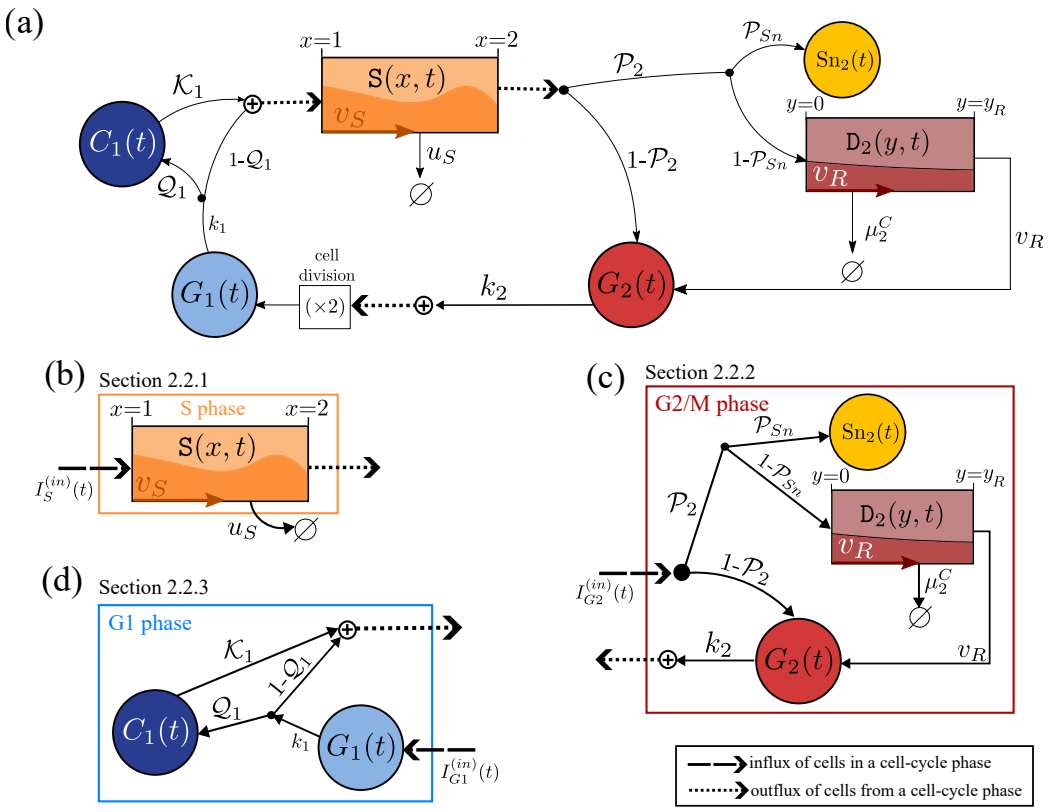


Figure 2.1: Schematic representation of our 6-compartment cell-cycle model. We decompose the full model (a) into three sub-models each corresponding to a different cell-cycle phase: G1 phase (see panel (d)), S phase (see panel (b)) and G2/M phase (see panel (c)). Cells start the cell-cycle in the G_1 compartment with one copy of DNA, they then transition to S and proceed to G2/M upon completion of DNA synthesis. At the end of G2/M phase cells divide giving rise to two daughter cells. For each sub-model (b)-(d) we indicate the section where we introduce governing equations. In Section 2.2.4, we link the different sub-models and then summarise the governing equations for the full model (a) in Section 2.2.5. In the schematic, circles and squares indicate well-mixed and structured compartments, respectively. Cells transition between compartments moving along arrows. At branching points, indicated by black dots, cell can follow different "fates" according to a given probability (for example cells that exit the G_1 compartment are redistributed with probability $\mathcal{Q}_1$ into C_1 and $1-\mathcal{Q}_1$ into $\mathtt{S}$). White dots represent convergence nodes. The out-flux from these nodes is obtained summing over the influxes. The empty set is used to indicate cell death.

of $\mathtt{S}(x,t)$ is governed by the following hyperbolic PDE:

$$\frac{\partial \mathtt{S}}{\partial t} + v_S(t)\frac{\partial \mathtt{S}}{\partial x} = -u_S(t)\mathtt{S}, \quad t > 0,\ x \in (1,2]. \tag{2.2a}$$

with boundary and initial conditions:

$$v_S(t)\mathtt{S}(1,t) = I_S^{(in)}(t), \tag{2.2b}$$

$$\mathtt{S}(x,0) = \mathtt{S}_0(x), \quad x \in (1,2]. \tag{2.2c}$$

In Eq. (2.2b), the function $I_S^{(in)}(t) \geq 0$ indicates the influx of cells in the S phase; its form will be given in Section 2.2.4. Since cell numbers can not be negative or infinite, we assume that $\mathtt{S}_0(x) \geq 0$ for all $x \in [1,2]$ and $\int_1^2 \mathtt{S}_0(x) < \infty$.

DNA synthesis in hypoxia. In Eq. (2.2a), we account for DNA synthesis by assuming that cells in the $\mathtt{S}$ compartment are advected along the x-axis at velocity $v_S(t) > 0$ [hr^{-1}] (i.e., they produce DNA at a rate $v_S(t)$). We model the impact of hypoxia on DNA synthesis by allowing v_S to evolve over time in response to the oxygen levels. Specifically, we propose the following piecewise linear ODE to describe how the rate of DNA synthesis v_S adapts to changes in the local oxygen levels $c(t)$:

$$\frac{dv_S}{dt} = \begin{cases} -\mathcal{R}_S^- \left(v_S - v_S^-\right), & c(t) < c_H, \\ \mathcal{R}_S^+ \left(v_S^+ - v_S\right), & c(t) \geq c_H, \end{cases} \tag{2.3a}$$

$$v_S(0) = v_{S,0}. \tag{2.3b}$$

In Eqs. (2.3), $\mathcal{R}_S^+$, $\mathcal{R}_S^-$, $v_{S,0}$, v_S^+ and v_S^- are non-negative constants. Here v_S^+ and v_S^- represent, respectively, the equilibrium velocities in well-oxygenated ($c > c_H$) and hypoxic ($c < c_H$) environments. Since DNA synthesis is slowed down in hypoxia, we consider $v_S^- < v_S^+$. It is straightforward to show that if $v_{S,0} \in [v_S^-, v_S^+]$, then $v_S(t) \in [v_S^-, v_S^+]$ at all times $t > 0$, so that v_S^+ and v_S^- can also be viewed, respectively, as the maximum and minimum rates of DNA synthesis. When oxygen levels drop below the critical value c_H, the rate of DNA synthesis decreases (i.e. $dv_S/dt \leq 0$ for $c < c_H$). As observed in [52], cells maintain a minimum level of DNA synthesis even after prolonged exposure to severe hypoxia ($c < c_H$). We therefore assume that $v_S^- > 0$, which guarantees that $\mathtt{S}(1,t)$ remains finite (see Eq. (2.2b)). When $c > c_H$, v_S increases (i.e., $dv_S/dt > 0$), towards its maximum value v_S^+.

Fork collapse. As discussed in Section 1.1.3, if hypoxia is prolonged (more than $\approx$ 12 hours [114]), reduced expression of proteins associated with DNA replication and DNA damage repair leads to the collapse of the replication forks and cell death [5, 114]. In Eq. (2.2a), this mechanism is captured by the term, $-u_S(m)\mathtt{S}$. We

assume that cells are sensitive to forks collapse if $m(t)$ drops below the threshold $M_{FC} \in (M^-, 1)$ and define the death rate u_S as:

$$u_S(m) = \mu_S \mathcal{H}(M_{FC} - m; \epsilon_{FC}), \tag{2.3c}$$

where $\mu_S > 0$ and $M_{FC} \in [M^-, 1)$ are constants. In Eq. (2.3c), $\mathcal{H}$ is a sigmoid function (here we adopt the hyperbolic tangent):

$$\mathcal{H}(\cdot; \epsilon) = \frac{1}{2}\left[\tanh\left(\frac{\cdot}{\epsilon}\right) + 1\right], \tag{2.4}$$

where ϵ controls the steepness of the sigmoid curve. In Eq. (2.3c), we fix $\epsilon_{FC} = 0.05$. In Eq. (2.3c), M_{FC} represents the expression levels of proteins for which the death rate u_S attains half of its maximum value, (*i.e.*, $\mu_S/2$). Based on Eq. (2.3c), the probability of a cell dying due to fork collapse decreases with m; since m decreases in hypoxia (see Eq. (2.1)) this implies that, the likelihood of a cell dying in the S phase increases with exposure to hypoxia.

2.2.1.1 Solving for $\mathtt{S}(x,t)$

We solve Eqs. (2.2) via the *method of characteristics*. As shown in Fig. 2.2, we divide the (x,t) plane into two regions: one in which the solution ($\mathtt{S}(x,t)$) is determined by the initial data, and one in which it is determined by the boundary data (i.e., Eq. (2.17c)).

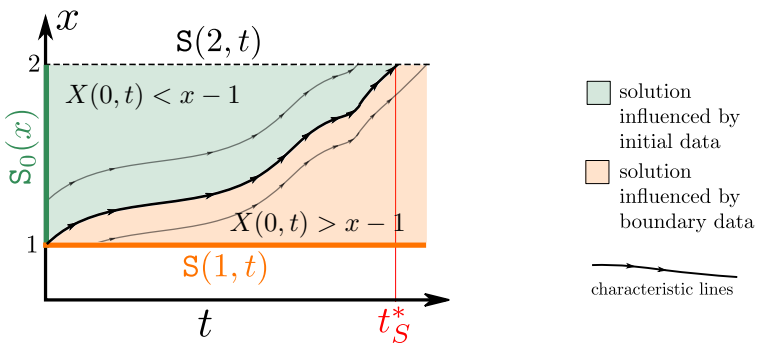


Figure 2.2: Schematic illustrating the characteristic curves defined by Eqs. (2.5)-(2.6) in the (t,x) plane (for the solution of $\mathtt{S}(x,t)$). In each case, the black bold curve divides the plane into two distinct regions: in the green region the solution ($\mathtt{S}(x,t)$) is determined by the initial data; in the orange region this is, instead, determined by the boundary data. In panel (a) the bold curve is defined by the condition $X(0,t) = \int_0^t v_S(\ell)d\ell = x - 1$ (see main text for definition $X(a,b)$).

We start by considering the region influenced by the initial data. In this region, we parametrise the characteristic curves via (τ_x, x_0), where $x_0 \in [1,2]$ corresponds

to the DNA content in a cell at time $t = 0$. This yields the following characteristic equations:

$$\begin{cases} \frac{dt}{d\tau_x} = 1, & t(0, x_0) = 0, \\ \frac{dx}{d\tau_x} = v_S(t(\tau_x, x_0)), & x(0, x_0) = x_0, \\ \frac{d\mathtt{S}}{d\tau_x} = -u_S(x(\tau_x, x_0), t(\tau_x, x_0))\mathtt{S}, & \mathtt{S}(0, x_0) = \mathtt{S}_0(x_0). \end{cases} \tag{2.5}$$

In the region influenced by the boundary data, we parametrise the characteristic curves via (τ_x,s), where $s > 0$ corresponds to the time at which a cell with DNA content x at time t entered the S phase. This yields the following characteristic system:

$$\begin{cases} \frac{dt}{d\tau_x} = 1, & t(0, s) = s, \\ \frac{dx}{d\tau_x} = v_S(t(\tau_x, x_0)), & x(0, s) = 1, \\ \frac{d\mathtt{S}}{d\tau_x} = -u_S(x(\tau_x, s), t(\tau_x, s))\mathtt{S}, & \mathtt{S}(0, s) = \frac{I_S^{(in)}(s)}{v_S(s)}. \end{cases} \tag{2.6}$$

Solving Eqs. (2.5)-(2.6), we obtain

$$\mathtt{S}(x, t) = \begin{cases} \mathtt{S}_0\left(x - X(0, t)\right) e^{-\int_0^t u_S(\ell)d\ell}, & X(0, t) \leq x - 1, \\ \frac{I_S^{(in)}(t-\tau_x(x,t))}{v_S(t-\tau_x(x,t))} e^{-\int_{t-\tau_x(x,t)}^t u_S(\ell)d\ell}, & X(0, t) > x - 1, \end{cases} \tag{2.7}$$

where $X(a, b) = \int_a^b v_S(\ell)d\ell$ and the delay function $\tau_x = \tau_x(x, t)$ satisfies

$$x - 1 = X(t - \tau_x(x, t), t), \qquad \text{if } X(0, t) > x - 1. \tag{2.8}$$

In our model, $\tau_x(x, t)$ indicates the amount of time a cell with DNA content x at time t has spent in the S compartment.

2.2.1.2 A reduced model of the S phase dynamics

Given Eqs. (2.7), we can derive an ODE model for the evolution of the total number of cells $S(t) = \int_1^2 \mathtt{S}(x, t)dx$. This is obtained by integrating Eqs. (2.2a) over $x \in [1, 2]$ and then using the explicit solution for S in Eq. (2.7). By doing so, we find that the time evolution of $S(t)$ is governed by the following differential equation with one time-dependent delay:

$$\frac{dS}{dt} = I_S^{(in)}(t) - I_S^{(out)}(t) - u_S(t)S(t), \qquad t > 0, \quad S(0) = \int_1^2 \mathtt{S}_0(x)dx, \tag{2.9a}$$

where:

$$I_S^{(out)}(t) = \begin{cases} v_S(t)\mathtt{S}_0\left(2 - \int_0^t v_S(\ell)d\ell\right) \exp\left\{-\int_0^t u_S(2 - X(\ell, t), \ell)d\ell\right\}, & t < t_S^*, \\ \frac{v_S(t)I_S^{(in)}(t-\tau_S(t))}{v_S(t-\tau_S(t))} \exp\left\{-\int_{t-\tau_S(t)}^t u_S(2 - X(\ell, t), \ell)d\ell\right\}, & t \geq t_S^*, \end{cases} \tag{2.9b}$$

where the scalar fields u_S and v_S are defined in Eqs. (2.3) and $t_S^* > 0$ is the unique time at which $\tau_S(t) = t$. In Eq. (2.9b), we have introduced the delay variable $\tau_S(t)$. This is implicitly defined by the integro-equation

$$1 = \int_{t-\tau_S(t)}^{t} v_S(\xi) d\xi, \tag{2.9c}$$

and indicates the time that a cell exiting the S phase at time t has taken to complete DNA duplication. In Eq. (2.3), prolonged or frequent exposure to hypoxia slows DNA replication and, therefore, increases the time τ_S that cells spend in the S phase (see Eq. (2.3)). Consequently, large values of τ_S implies that cells have been exposed to prolonged periods of replication stress that can lead to cell damage.

2.2.2 Modelling G2/M phase in hypoxia

As is standard in cell-cycle models [32], we combine the G2 and M phases into a single compartment under the assumption that the duration of the M phase (≈ 30 minutes) is much shorter than the G2 phase (≈ 5 hours).

To account for the impact of hypoxia on cells progression in the G2 phase, we split G2/M cells into three compartments: G_2, $\mathtt{D}_2$ and Sn_2. The variable $G_2 = G_2(t)$ counts the number of cells in the G2/M phase that are "undamaged" and as such can successfully undergo cell division. The variable $\mathrm{Sn}_2 = \mathrm{Sn}_2(t)$ indicates the number of cells in the G2 phase that are senescent. These cells have been severely damaged and, as result, have permanently withdrawn from the cell-cycle (*i.e.*, they will never exit the G2 phase). Senescence can be viewed as "cell death", in the sense that cells permanently lose their clonogenic capacity and, therefore, cease to be tumorigenic. However, senescence differs from apoptosis since cells may continue perform other functions. Cells in the $\mathtt{D}_2$ compartments are also damaged but have not entered senescence. Proper functioning of the DNA damage response (DDR) enables these cells to repair the damage accumulated and eventually re-enter the cell-cycle (see Section 1.1). We account for the repair of cells in the $\mathtt{D}_2$ compartment structuring this compartment by an internal, continuous variable $y \in [0, y_R]$. Here y is a non-dimensional measure of the extent to which cells have recovered since they entered the $\mathtt{D}_2$ compartment. We denote by $\mathtt{D}_2 = \mathtt{D}_2(y,t)$ the number of cells with repair level y at time t. Consistently with Fig. 2.1c, the evolutions of the cell numbers $G_2(t)$, $\mathrm{Sn}_2(t)$ and $\mathtt{D}_2(y,t)$ are governed by the following system of two ODEs coupled to an hyperbolic PDE:

$$\frac{\partial \mathtt{D}_2}{\partial t} + v_R(m)\frac{\partial \mathtt{D}_2}{\partial y} = -\mu_2^C\, \mathtt{D}_2, \quad t > 0,\ y \in (0, y_R], \tag{2.10a}$$

$$\frac{dG_2}{dt} = -k_2 G_2 + v_R(m)\mathtt{D}_2(t, y_R) + (1 - \mathcal{P}_2(\tau_S))I_{G2}^{(in)}(t), \quad t > 0, \tag{2.10b}$$

$$\frac{d\,\mathrm{Sn}_2}{dt} = \mathcal{P}_{Sn}(\tau_S)\mathcal{P}_2(\tau_S)I_{G2}^{(in)}(t), \quad t > 0. \tag{2.10c}$$

with boundary condition (for Eq. (2.10a)):

$$v_R(m)\mathtt{D}_2 = \mathcal{P}_2(\tau_S)(1 - \mathcal{P}_{Sn}(\tau_S))I_{G2}^{(in)}(t). \tag{2.10d}$$

Eqs. (2.10a)-(2.10c) are closed by imposing initial conditions:

$$\mathrm{Sn}_2(0) = \mathrm{Sn}_{2,0}, \quad G_2(0) = G_{2,0}, \quad \mathtt{D}_2(y, 0) = \mathtt{D}_{2,0}(y) \;\; y \in [0, y_R], \tag{2.10e}$$

where $G_{2,0}$ and $\mathrm{Sn}_{2,0}$ are constants and $\mathtt{D}_{2,0}(y)$ is a prescribed function. Since cell numbers can not be negative or infinite, we assume that $G_{2,0} \geq 0$, $\mathrm{Sn}_{2,0} \geq 0$, $\mathtt{D}_{2,0}(y) \geq 0$ for all $y \in [0, y_R]$ and $\int_0^{y_R} \mathtt{D}_2(y)dy < \infty$. In Eq. (2.10d), $I_{G2}^{(in)}(t) \geq 0$ indicates the flux of cells that enter the G2 phase at time t; its explicit form will be given in Section 2.2.4.

Repair dynamics. In Eq. (2.10a), the repair dynamics is modelled by the advection term: cells move along the y-axis at velocity $v_R(m) > 0$ [hr^{-1}]. As mentioned in Section 2.2.1, in hypoxia, repair mechanisms are impaired as cells lower translation of repair proteins, $m(t)$ (see Eq. (2.1)). Accordingly, we define the repair rate v_R as follows:

$$v_R(m) = \frac{v_R^+ m}{M_R + m}. \tag{2.11a}$$

In Eq. (2.11a) the constants $v_R^+ > 0$ and $M_R > 0$ represent respectively the maximum value that v_R can attain and the level of m at which v_R is equal to $v_R^+/2$. We note that this high-level description allows us to account for the effect of hypoxia on the repair dynamics.

In writing Eq. (2.10a), we assume that cells in the $\mathtt{D}_2$ compartment die at a constant rate μ_2^C [hr^{-1}]; for cell-lines with functional G2 checkpoint, we expect μ_2^C to be negligible ($\mu_2^C \approx 0$). In contrast, for cells with defective G2 checkpoint (for example due to a mutation), we expect $\mu_2^C > 0$. As these cells can not delay entrance to mitosis when damaged, they are likely to die mitotic catastrophe [101].

Cell division. We assume that only cells in the G_2 compartment can divide and that they do so at a constant rate $k_2 > 0$ [hr^{-1}] (see first term on the right hand side of Eq. (2.10b)).

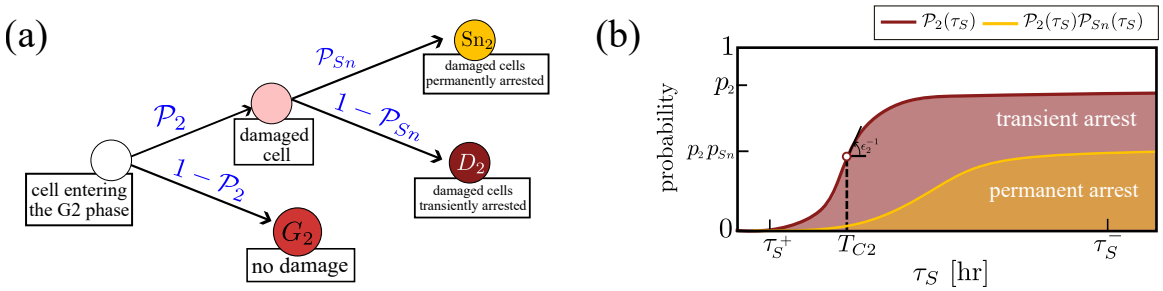


Figure 2.3: (a) Graph representing cell-fate decisions for cells entering the G2 phase. We indicate in blue the probability of cells following a given fate-path. (b) Plots of the probability $\mathcal{P}_2$ (see Eq. (2.11b)) of a cell entering the G2 phase being damaged and $\mathcal{P}_2\mathcal{P}_{Sn}$ (see Eq. (2.11b)-(2.11c)) corresponding to the probability of cells entering the G2 phase becoming senescent. As detailed in the text, the probabilities of cells having repairable/irreparable damage depends on the delay variable τ_S. As τ_S increases, cells become increasingly likely to accumulate damage and this being more severe.

The impact of hypoxia on the fate of cells entering the G2 phase. In writing Eqs. (2.10a)-(2.10d), we allow cells entering the G2/M phase to follow different fates. As shown in Fig. 2.3a, cells entering the G2 phase may become G_2, D_2 or Sn_2 cells with different probabilities. Specifically we assume that a fraction $\mathcal{P}_2 \in [0,1]$ will be damaged, while the remaining cells transit to the G_2 compartment (where they can successfully divide). Of the damaged cells, we assume a fraction $\mathcal{P}_{Sn} \in [0,1]$ become senescent while the remaining cells transit to the $\mathtt{D}_2$ compartment.

As discussed in Section 1.1.3, hypoxia slows down DNA replication. This results in pathologically long duration of the S phase ($\tau_S \gg 1$) during which cells experience high-levels of replication stress that can lead to cell damage. As shown in Fig. 2.3b, we account for this aspect assuming that the probability $\mathcal{P}_2(\tau_S)$ that cells exiting the S phase are damaged increases monotonically with τ_S. Based on Eq. (2.3), τ_S satisfies $1/v_S^+ \leq \tau_S \leq 1/v_S^-$. When oxygen levels are sufficiently high, (i.e. $c(t) > c_H$ for all t), neglecting an initial transient in the case $v_{0,S} \neq v_S^+$, we have that $\tau_S \approx 1/v_S^+ = \tau_S^+$. Conversely, when oxygen levels are constantly low, (i.e. $c(t) < c_H$ for all t), $\tau_S \approx 1/v_S^- = \tau_S^-$. We assume that the probability of cells being damaged in an oxygen-rich environment is negligible $\mathcal{P}_2(\tau_S^+) \approx 0$, while it is maximum when cells are exposed to prolonged constant hypoxia, $\mathcal{P}_2(\tau_S^-) \approx p_2 \in [0,1]$. We assume that $\mathcal{P}_2$ transition in between these two regimes by imposing:

$$\mathcal{P}_2(\tau_S) = p_2 \mathcal{H}(\tau_S - T_{C2}, \epsilon_2), \tag{2.11b}$$

where $\mathcal{H}$ is a sigmoid function (see Eq. (2.4)). In Eq. (2.11b), the parameter ϵ_2 controls the steepness of $\mathcal{H}$ while T_{C2} controls the location of its inflection point

where $\mathcal{P}_2(T_{C2}) = p_2/2$ (see Fig. 2.3b). We fix $\epsilon_2 = 0.1$ hr, while, based on the above discussion, we assume $\tau_S^+ \ll T_{C2} \ll \tau_S^-$. Of the damaged cells, a fraction $\mathcal{P}_{Sn}$ becomes senescent. This usually occurs when a cell detects irreparable DNA damage [5, 91]. As shown in Fig. 2.3, we assume that the probability of a cell becoming senescence increases with exposure to hypoxic stress (i.e., it increases with τ_S). Analogously to Eq. (2.11b), we also assume that the probability $\mathcal{P}_{Sn}$ follows a sigmoid profile

$$\mathcal{P}_{Sn}(\tau_S) = p_{Sn}\mathcal{H}(\tau_S - T_{Sn}, \epsilon_{Sn}). \tag{2.11c}$$

In Eq. (2.11c), we fix the parameter $\epsilon_{Sn} = 4.0$ hr, while $p_{Sn} \in [0, 1]$ and assume $T_{Sn} > T_{C2}$.

2.2.2.1 Solving for $\mathtt{D}_2(y, t)$

We solve Eq. (2.10a) following the same procedure outlined in Section 2.2.1.1 by using the *method of characteristics*. We find that $\mathtt{D}_2(y, t)$ is defined as follow:

$$\mathtt{D}_2(a,t) = \begin{cases} \mathtt{D}_{2,0}\left(y - Y(0,t)\right) e^{-\mu_2^C t}, & Y(0,t) \le y, \\ \dfrac{I_{D2}^{(in)}(t - \tau_y(y,t))}{v_R(t - \tau_y(y,t))} e^{-\mu_2^C \tau_y(y,t)}, & Y(0,t) > y, \end{cases} \tag{2.12}$$

where $I_{D2}^{(in)}(t) = I_{G2}^{(in)}(t)\mathcal{P}_2(\tau_S(t))(1 - \mathcal{P}_{Sn}(\tau_S(t)))$, and $Y(a, b) = \int_a^b v_R(\ell)d\ell$ and the function $\tau_y = \tau_y(y, t)$ satisfies

$$y = Y(t - \tau_y(y,t), t), \qquad \text{if } \; Y(0,t) \; > \; y. \tag{2.13}$$

Physically, $\tau_y(y, t)$ indicates the amount of time a cell with internal state y at time t has spent in the $\mathtt{D}_2$ compartment.

2.2.2.2 A reduced model of the G2 phase dynamics

Having solved for the variable $\mathtt{D}_2(y, t)$, we can derive a simplified version of Eqs. (2.10). We introduce the variable $D_2(t) = \int_0^{y_R} \mathtt{D}_2(y, t)dy$, which counts the total number of cells in the $\mathtt{D}_2$ compartment. We can derive an ODE for D_2 by integrating Eq. (2.10a) over the structure variable $y \in [0, y_R]$. Using Eqs. (2.10) and Eq. (2.12), we can write a system of governing equation for the evolution of the variables G_2, D_2 and Sn_2:

$$\frac{dG_2}{dt} = -k_2 G_2 + I_{D2}^{(out)}(t) + (1 - \mathcal{P}_2(t)) \; I_{G2}^{(in)}(t), \qquad t > 0, \tag{2.14a}$$

$$\frac{dD_2}{dt} = (1 - \mathcal{P}_{Sn}(t))\mathcal{P}_2(t) I_{G2}^{(in)}(t) - I_{D2}^{(out)}(t) - \mu_2^C D_2, \qquad t > 0, \tag{2.14b}$$

$$\frac{d\,\mathrm{Sn}_2}{dt} = \mathcal{P}_{Sn}(\tau_S)\mathcal{P}_2(\tau_S)I_{G2}^{(in)}, \qquad t > 0, \tag{2.14c}$$

where

$$I_{D2}^{(out)}(t) = \begin{cases} v_R(t)\mathtt{D}_{2,0}\left(y_R - \int_0^t v_R(\ell)d\ell\right)e^{-\mu_2^C t}, & t < t_R^*, \\ \frac{v_R(t)I_{G2}^{(in)}(t-\tau_R(t))}{v_R(t-\tau_R(t))}e^{-\mu_2^C \tau_R(t)}, & t \geq t_R^*, \end{cases} \tag{2.14d}$$

and with initial conditions:

$$G_2(0) = G_{2,0}, \quad \mathrm{Sn}_2 = \mathrm{Sn}_{2,0}, \quad D_2(t) = \int_0^{y_R} \mathtt{D}_{2,0}(y)dy. \tag{2.14e}$$

In Eqs. (2.14a)-(2.14d), the coefficients v_R, $\mathcal{P}_{Sn}$ and $\mathcal{P}_2$ are as defined in Eqs. (2.11) and t_R^* is the unique time at which $\tau_R(t) = t$. In Eq. (2.14d), we have introduced the delay variable $\tau_R(t)$, which satisfies the integro-equation

$$y_R = \int_{t-\tau_R(t)}^{t} v_R(\xi)d\xi. \tag{2.14f}$$

This indicates the delay between when a D_2 cell entered the G2 phase and the time t when it completes repair.

2.2.3 Modelling G1 phase in hypoxia

To account for the inhibition of the G1→S transition in hypoxia, we split cells in the G1 phase into two compartments: G_1 and C_1. Here, $C_1 = C_1(t)$ counts the number of cells in the G1 phase that have completed the G1 phase but have not yet initiated DNA synthesis due hypoxia-induced activation of the G1 checkpoint. The variable $G_1 = G_1(t)$ counts instead the number of cells in the G1 phase that are not arrested. Following the schematic in Fig. 2.1d, the evolutions of $G_1(t)$ and $C_1(t)$ are governed by the following system of two coupled ordinary differential equations:

$$\frac{dG_1}{dt} = I_{G1}^{(in)}(t) - k_1 G_1, \quad t > 0, \tag{2.15a}$$

$$\frac{dC_1}{dt} = \mathcal{Q}_1(c)k_1 G_1 - \mathcal{K}_1(c)C_1, \quad t > 0, \tag{2.15b}$$

with initial conditions:

$$G_1(0) = G_{1,0}, \quad C_1(0) = C_{1,0}, \tag{2.15c}$$

where $G_{1,0}$ and $C_{1,0}$ are prescribed non-negative constants. The term $I_{G1}^{(in)}(t) \geq 0$ in Eq. (2.15a) captures the influx of newborn cells into the G1 phase (this will be

specified in Section 2.2.4), while the positive constant k_1 [hr^{-1}] represents the rates at which cells leave G_1. We model the effect of hypoxia on the G1→S transition by assuming that transition rates into and out of the C_1 compartment (*i.e.*, $k_1\mathcal{Q}_1$ and $\mathcal{K}_1$) depend on current oxygen levels, c, via the following functional forms:

$$\mathcal{K}_1(c) = K_1\mathcal{H}(c - c_H; \epsilon_1), \tag{2.16a}$$

$$\mathcal{Q}_1(c) = q_1\,\mathcal{H}(c_H - c; \epsilon_1), \tag{2.16b}$$

where K_1 [hr^{-1}] and q_1 are positive constants with $q_1 \in [0, 1]$ and $c_H = 1\%\ O_2$ is the oxygen threshold introduced above. The function $\mathcal{H}$ is the sigmoid function given by Eq. (2.4) and $\epsilon_1 = 0.1\%\ O_2$. Based on Eqs. (2.16), when $c < c_H$, only a fraction $1 - q_1$ of the cells exiting the G_1 compartment proceeds to the S phase, while the remaining cells arrest in the C_1 compartment. In contrast, when $c > c_H$, cells completing the G1 phase directly transition to the S phase with no delays ($\mathcal{Q}_1(c) \approx 0$ when $c > c_H$) and arrested cells can re-enter the cycle at rate $\mathcal{K}_1$. Note that in Eq. (2.15) we have neglected cell death under the assumption that the viability of G1 cells is not affected by hypoxia. This would not be the case if we were trying to model anoxia (*i.e.*, complete lack of oxygen).

2.2.4 Coupling the different cell-cycle phases.

In Sections 2.2.1-2.2.3, we have introduced the 3 sub-models describing the time-evolution of cells in each cell-cycle phase. We now couple the dynamics in the different cell-cycle phases by specifying the form of the influxes $I_{G1,S,G2}(t) \geq 0$, into each cell-cycle phase. Referring to the schematic in Fig. 2.1 and applying the principle of mass balance, we obtain that the influxes $I_{G1,S,G2}(t)$ satisfy

$$I_{G1}^{(in)}(t) = 2k_2G_2, \tag{2.17a}$$

$$I_{S}^{(in)}(t) = [1 - \mathcal{Q}_1(c(t))]\,k_1G_1(t) + \mathcal{K}_1(c(t))C_1(t), \tag{2.17b}$$

$$I_{G2}^{(in)}(t) = v_S(t)\mathtt{S}(2, t). \tag{2.17c}$$

In Eq. (2.17a), the factor 2 on the right hand side arises because cell division produces two daughter cells. The right hand-sides of Eqs. (2.17b)-(2.17c) correspond respectively the flux of cells that exit the G1 and S phases.

2.2.5 Statement of the full cell-cycle model

We summarise the governing equations for our cell-cycle model. These are derived combining Eqs. (2.9), (2.14), and (2.15) via the definition of fluxes in Eqs. (2.17).

By doing so, we find that the time-evolution of the number of cells in each model compartment ($G_1(t)$, $C_1(t)$, $S(t)$, $G_2(t)$, $D_2(t)$ and $\mathrm{Sn}_2(t)$) is governed by the following system of coupled differential equations with two delays:

$$\frac{dG_1}{dt} = 2k_2G_2 - k_1G_1, \qquad t > 0, \tag{2.18a}$$

$$\frac{dC_1}{dt} = \mathcal{Q}_1(c)k_1G_1 - \mathcal{K}_1(c)C_1, \qquad t > 0, \tag{2.18b}$$

$$\frac{dS}{dt} = I_S^{(in)}(t) - I_S^{(out)}(t) - u_S(m)S, \qquad t > 0, \tag{2.18c}$$

$$\frac{dG_2}{dt} = -k_2G_2 + I_2^{(out)}(t) + \left[1 - \mathcal{P}_2(\tau_S)\right] I_S^{(out)}(t), \qquad t > 0, \tag{2.18d}$$

$$\frac{d\mathrm{Sn}_2}{dt} = \mathcal{P}_{Sn}(\tau_S)I_2^{(in)}(t), \qquad t > 0, \tag{2.18e}$$

$$\frac{dD_2}{dt} = \left[1 - \mathcal{P}_{Sn}(\tau_S)\right] I_2^{(in)}(t) - I_2^{(out)} - \mu_2^C D_2, \qquad t > 0, \tag{2.18f}$$

where the flux variables $I_S^{(out)}$, $I_S^{(in)}$, $I_2^{(in)}$ and $I_2^{(out)}$ are given by

$$I_S^{(out)}(t) = \begin{cases} v_S(t)\mathtt{S}_0 \left(2 - \int_0^t v_S(\ell)d\ell\right) \exp\left\{-\int_0^t \mu_S(\ell)d\ell\right\}, & t < t_S^*, \\ v_S(t) \exp\left\{-\int_{t-\tau_S(t)}^t \mu_S(\ell)d\ell\right\} \dfrac{I_S^{(in)}(t-\tau_S(t))}{v_S(t-\tau_S(t))}, & t \geq t_S^*, \end{cases} \tag{2.18g}$$

$$I_2^{(out)}(t) = \begin{cases} \mathtt{D}_{2,0} \left(y_R - \int_0^t v_R(\ell)d\ell\right) e^{-\mu_2^C t}, & t < t_R^*, \\ v_R(t)e^{-\mu_2^C \tau_R(t)} \left.\dfrac{I_2^{(in)}(\xi)[1-\mathcal{P}_{Sn}(\tau_S(\xi))]}{v_R(\xi)}\right|_{\xi=t-\tau_R(t)}, & t \geq t_R^*, \end{cases} \tag{2.18h}$$

$$I_S^{(in)}(t) = \left[1 - \mathcal{Q}_1(c(t))\right] k_1G_1(t) + \mathcal{K}_1(c(t))C_1(t), \tag{2.18i}$$

$$I_2^{(in)}(t) = \mathcal{P}_2(\tau_S(t))I_S^{(out)}(t). \tag{2.18j}$$

The delay functions τ_i ($i = S, R$) in Eqs. (2.18g)-(2.18j) are implicitly defined by the following integro-equations:

$$1 = \int_{t-\tau_S(t)}^t v_S(\xi)d\xi, \quad y_R = \int_{t-\tau_R(t)}^t v_R(\xi)d\xi, \ t > 0. \tag{2.18k}$$

while t_i^* ($i = S, R$) are defined by $\tau_i(t_i^*) = t_i^*$. The functions $\mathcal{Q}_1$, $\mathcal{K}_1$ are defined as in Section 2.2.3, u_S and v_S as in Section 2.2.1, and $\mathcal{P}_2$, $\mathcal{P}_{Sn}$, and v_R as in Section 2.2.2:

$$\mathcal{K}_1(c) = K_1\mathcal{H}(c - c_H; \epsilon_1), \tag{2.18l}$$

$$\mathcal{Q}_1(c) = q_1\,\mathcal{H}(c_H - c; \epsilon_1), \tag{2.18m}$$

$$u_S(m) = \mu_S\mathcal{H}(M_{FC} - m; \epsilon_{FC}), \tag{2.18n}$$

$$\mathcal{P}_2(\tau_S) = p_2\mathcal{H}(\tau_S - T_{C2}, \epsilon_2), \tag{2.18o}$$

$$\mathcal{P}_{Sn}(\tau_S) = p_{Sn}\mathcal{H}(\tau_S - T_{Sn}, \epsilon_{Sn}), \tag{2.18p}$$

$$\frac{dv_S}{dt} = \begin{cases} -\mathcal{R}_S^- \left(v_S - v_S^-\right), & c(t) < c_H, \\ \mathcal{R}_S^+ \left(v_S^+ - v_S\right), & c(t) \geq c_H, \end{cases} \tag{2.18q}$$

$$v_R(m) = \frac{v_R^+ m}{M_R + m}, \tag{2.18r}$$

where the expression levels of repair proteins evolves according to

$$\frac{dm}{dt} = \begin{cases} -\mathcal{R}_M^- \left(m - M^-\right), & c(t) < c_H, \\ \mathcal{R}_M^+ \left(1 - m\right), & c(t) \geq c_H. \end{cases} \tag{2.18s}$$

In Eqs. (2.18l)-(2.18s), the function $c = c(t)$ indicates the time-evolution of oxygen levels; this will take a different form depending on the environment we wish to simulate. Details will be given in Sections 2.3 and 2.4. The system is closed by imposing the following initial/history conditions:

$$\begin{gathered} G_1(0) = G_{1,0}, \quad G_2(0) = G_{2,0}, \quad C_1(0) = C_{1,0}, \quad \mathrm{Sn}_2 = \mathrm{Sn}_{2,0}, \\ S(0) = S_0 = \int_1^2 \mathtt{S}_0(\xi)d\xi, \quad D_2(0) = D_{2,0} = \int_0^{y_R} \mathtt{D}_{2,0}(\xi)d\xi, \\ v_S(0) = v_{S,0}(\xi), \quad \xi \in (-\tau_S(0), 0], \quad m(\xi) = m_0(\xi) \quad \xi \in [-\tau_R(0), 0], \end{gathered} \tag{2.18t}$$

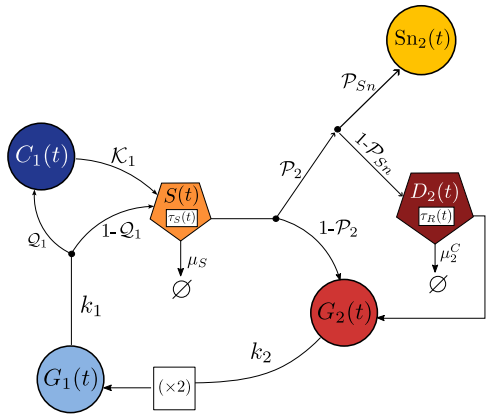
with $\mathtt{S}_0 : (1, 2] \to [0, \infty)$, $\mathtt{D}_{2,0} : (0, y_R] \to [0, \infty)$, $v_{S,0}(\xi) : [-\tau_S(0), 0] \to [v_S^-, v_S^+]$ and $m_0(\xi) : [-\tau_R(0), 0] \to [M^-, 1]$ are prescribed functions, while $G_{1,0} \geq 0$, $G_{2,0} \geq 0$, $C_{1,0} \geq 0$, $\mathrm{Sn}_{2,0} \geq 0$ are constants. Note that history conditions on the values of v_S and m are needed in order for the delays $\tau_S(t)$ and $\tau_R(t)$ (see Eq. (2.18k)) to be defined at all times $t \geq 0$. A schematic of the model defined by Eqs. (2.18) is presented in Fig. 2.4 alongside a summary of its variables. When comparing this schematic with that of the full model (Fig. 2.1), we note that the two structured compartments have been replaced by two well-mixed compartments and two, time-dependent delays: $\tau_S(t)$ and $\tau_R(t)$.

2.2.6 Relating the cell-cycle dynamics to population growth

Our cell-cycle model describes how variable oxygen environments perturb cell-cycle progression. We are interested in studying how dysregulation of the cell-cycle affects the growth of cell populations. To do so, we introduce the total number of cells, $N = N(t)$, which is defined by summing the number of cells in each model compartment:

$$N(t) = G_1(t) + C_1(t) + S(t) + G_2(t) + D_2(t) + \mathrm{Sn}_2(t). \tag{2.19a}$$

(a)



(b)

t	time [hr]
c	externally prescribed oxygen levels [% O_2]
m	expression level of repair proteins in cells
τ_S	delay between cells entering and exiting the S phase [hr]
τ_R	delay between damaged cells entering the G2 phase and them finishing repair [hr]
$G_1(t)$	number of cells in the G1 phase that are not arrested at time t
$C_1(t)$	number of cells that are arrested in the G1 checkpoint at time t
$S(t)$	number of cells in the S phase at time t
$G_2(t)$	number of cells in the G2/M phases that can divide at time t
$D_2(t)$	number of damaged cells transiently arrested in the G2 phase at time t
$\mathrm{Sn}_2(t)$	number of senescent cells in the G2 phase at time t

Figure 2.4: (a) Schematic representing our 6-compartment cell-cycle model when formulated as a system of differential equations with delays (see Eqs. (2.18)) after solving Eqs. (2.2a) and (2.10a) via the method of characteristics. As in Fig. 2.1, at branching points cell fluxes are redistributed with according to the weight indicated. (b) Table summarising model variables in Eqs. (2.18).

Differentiating Eq. (2.19a) with respect to time and using (2.18), we find that $N(t)$ satisfies:

$$\frac{dN}{dt} = k_2 G_2 - \mu_2^C D_2(t) - u_S(m) S(t), \tag{2.19b}$$

$$\text{with} \quad N(0) = \sum_{i=1}^{2} G_{i,0} + C_{1,0} + D_{2,0} + S_0 + \mathrm{Sn}_{2,0} := N_0. \tag{2.19c}$$

We can describe the composition of the population N via the cell fractions π_z which are defined by:

$$\pi_z(t) = \frac{z(t)}{N(t)}, \quad z \in \{G_1, C_1, S, G_2, D_2, \mathrm{Sn}_2\}. \tag{2.20}$$

These cell fractions correspond to the probability that a cell randomly chosen from the total population belongs to sub-population (*i.e.*, compartment) z.

2.2.7 Model parameters

Our model contains several unknown parameters, most of which will be cell-line (or tumour) specific. We estimate a subset of the model parameters from the literature, specialising them for the colorectal RKO cell-line, which has been extensively studied by our experimental collaborators. Parameter values are listed in Table 2.1 and are held fixed throughout the thesis. We acknowledge that fixing some of the parameters naturally limits our ability to explore model predictions. In particular, we here focus on cancer cells that have the same cell-cycle dynamics as RKO cells

in oxygen-rich environments but that differ in their response to fluctuating oxygen environments (which is the main focus of this thesis). As a result, we are still able to characterise the range of cell responses to different forms of hypoxia captured by our model. Fixing some of the model parameters will also facilitate model calibration to experimental data on cell-cycle dysregulation in RKO cancer cells, which will be performed in Chapter 3.

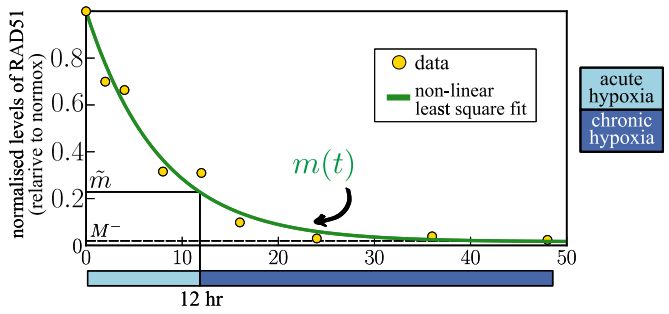
	Description	Typical value(s)
c_H	hypoxia threshold	$\approx 1\% \, O_2$
y_R	maximum level of repair for cells in the D_2 compartment	1
$\mathcal{R}_M^-$	rate of change of m when $c < c_H$	0.13 hr^{-1}
k_1	rate at which cells exit G_1	0.195 hr^{-1}
k_2	rate at which cells exit G_2	0.22 hr^{-1}
v_S^+	maximum velocity of DNA synthesis	0.083 hr^{-1}
$\tau_S^+ = 1/v_S^+$	duration of the S phase in oxygen-rich environments	12 hr
v_S^-	minimum velocity of DNA synthesis	0.005 hr^{-1}
$\tau_S^- = 1/v_S^-$	asymptotic duration of the S phase in constant hypoxia	200 hr
v_R^+	rate of damage repair when $c > c_H$	0.07 hr^{-1}
M^-	value of m in constant hypoxia $c < c_H$	0.0154
M_R	value of m for which repair velocity is inhibited	0.4
$\mathcal{R}_M^-$	rate of change of v_R when $c < c_H$	0.13 hr^{-1}
M_{FC}	value of m for which cells sensitise to fork collapse	0.22

Table 2.1: Summary of the parameters from Eqs. (2.18), that we can estimate from the literature. Typical values are given for the RKO cancer cell line. These parameters will be fixed throughout the thesis.

We estimate parameters associated with the evolution of repair protein expression levels, $m(t)$ (see Eq. (2.1)), using the data presented in Fig. 2.5, which describes the time-evolution of expression levels of the DNA repair protein RAD51 in constant hypoxia [115]. While in our model $m(t)$ corresponds to the expression of multiple different DNA repair proteins, we assume they behave similarly in response to hypoxia and use RAD51 dynamics as a prototypical response. In Fig. 2.5, the expression levels at each time point are normalised with respect to the expression levels in physiological oxygen-rich conditions (in our this correspond to $m = 1$). The data show a clear trend: the expression levels decrease monotonically as the period of exposure to hypoxia increases. The profile can be fitted to an exponential function (see the continuous curve in Fig. 2.5) justifying the functional form chosen for $m'(t)$ (see Eq. (2.1)). In

this case, model parameters can be inferred directly by interpolation of the data points with an exponential function ($m(t) = (1 - M_-)e^{-\mathcal{R}_M^- t} + M_-$). We could not find appropriate data to estimate $\mathcal{R}_M^+$, *i.e.*, the rate at which protein expression levels are restored upon reoxygenation.

(a)



(b)

parameters	values
$\mathcal{R}_M^-$	0.13 ± 0.015[1/hr]
M^-	1.54 ± 3.6 %

Figure 2.5: (a) Data from [115] on the time-evolution of expression levels of the DNA repair protein RAD51 in constant hypoxia (yellow dots). The green curve indicates the value of the function $m(t) = (1 - M^-)e^{\mathcal{R}_M^- t} + M^-$ (obtained by solving Eq. (2.1) with $m_0 = 1$ and $c(t) < c_H$ for all t) with M^- and $\mathcal{R}_M^-$ as specified in panel (b). (b) Estimates of the parameters M^- and $\mathcal{R}_M^-$ are obtained by fitting the function $m(t) = (1 - M^-)e^{\mathcal{R}_M^- t} + M^-$ to the experimental data using the `curve_fit` function in the *SciPy* library in Python. For each parameter we indicate the computed 67% confidence interval.

Using the data in Fig. 2.5, we find that the value of m after 12 hours in constant hypoxia drops to $\tilde{m} = 0.223$. As discussed in Section 1.1.3, experimentally it is observed that, when exposed to hypoxia for more than 12 hours, cells become sensitive to collapse of replication forks and repress repair mechanisms. Accordingly, we require that the death rate for cells in the S phase, $u_S(m)$ (see Eq. (2.18n)), is half of its maximum values when cells are exposed to constant hypoxia for 12 hours, *i.e.*, $u_S(\tilde{m}) = \mu_S/2$. By doing so, we obtain that $M_{FC} = \tilde{m} = 0.223$. To account for inhibition of repair in chronic hypoxia, we assume that the velocity of repair when $m = \tilde{m}$, $v_R(\tilde{m})$ (see Eq. (2.18r)), is half of its value at physiological levels, $v_R(\tilde{m}) = v_R(1)/2$. By doing so, we obtain $M_R = 0.40$. Finally, for the repair dynamics in physiological conditions, we use the reference time-scale of 20 hr [32] by setting the arbitrary threshold $y_R = 1$ and setting $v_R^+ = 0.07$ [hr^{-1}], so that $(v_R(1))^{-1} \approx 20$ hr; note that here the recovery level y does not have a clear physical analogue, so y_R can be chosen arbitrarily by suitably re-defining the value of v_R^+.

As we will discuss in Section 2.3.1, it is possible to derive equations that link cell-cycle model parameters (k_1, k_2, v_S^+) to quantities that can be directly measured via flow cytometry experiments on cell cultures grown in oxygen-rich conditions. Similar

	Description	Range	Default value	Units
$\mathcal{R}_M^+$	rate of change of m when $c > c_H$	$[10^{-2}, 2.0]$	0.4	$[\text{hr}^{-1}]$
$\mathcal{R}_S^+$	rate of change of v_S when $c > c_H$	$[10^{-2}, 2.0]$	0.25	$[\text{hr}^{-1}]$
$\mathcal{R}_S^-$	rate of change of v_S when $c < c_H$	$[10^{-2}, 2.0]$	0.5	$[\text{hr}^{-1}]$
K_1	maximum rate at which cells leave the C_1 compartment	$[0.01, 2.5]$	0.05	$[\text{hr}^{-1}]$
q_1	maximum probability of a cell arresting in the G1 phase	$[0.0, 1.0]$	0.9	
p_2	maximum probability of a cell entering the G2 phase being damaged	$[0.0, 1.0]$	0.8	
T_{C2}	value of τ_S for which $\mathcal{P}_2 = p_2/2$	$[13, 25]$	17	[hr]
p_{Sn}	maximum probability of a damaged cell entering the G2 phase becoming senescent	$[0.0, 1.0]$	0.1	
T_{Sn}	value of τ_S for which $\mathcal{P}_{Sn} = p_{Sn}/2$	$[20, 50]$	27	[hr]
μ_2^C	death rate in the D_2 compartment $[\text{hr}^{-1}]$	$[0.0, 0.1]$	0.0	$[\text{hr}^{-1}]$
μ_S	death rate due to fork collapse	$[0.0, 0.1]$	0.005	$[\text{hr}^{-1}]$

Table 2.2: Summary of the parameters that appear in Eqs. (2.18) that we could not directly estimate from the literature (prior to our work [29]). We specify physically realistic ranges of these parameters and indicate default value used in our simulations to model a cell with functional checkpoints (details given in the text).

results have been derived previously in the literature [10]. Using flow cytometry data from RKO cancer cell-line (details will be given in Section 2.3.1), we obtain the following estimates: $k_1 = 0.195$ $[\text{hr}^{-1}]$, $v_S^+ = 0.083$ [hr] and $k_2 = 0.22$ $[\text{hr}^{-1}]$. Based on experimental observation on DNA synthesis in chronic hypoxia [114], we expect v_S^- to be small. Since the limit $v_S^- \to 0$ is non-singular, it can be shown that, for $v_S^- \sim O(10^{-3})$, the solution is not sensitive to the precise value of v_S^- (results not shown). Therefore we fix $v_S^- = 0.005\,\text{hr}^{-1}$, which is sufficiently small to produce the desired qualitative behaviour.

Even after fixing the parameters in Table 2.1 as just described, our model still contains 11 unknown parameters (see Table 2.2). Based on our recent publication [29], in Chapter 3 we show how Bayesian inference can be used to estimate some of these unknown parameters. In this chapter, however, we explore in more generality the predictions of our cell-cycle model.

2.2.7.1 A cohort of *in silico* cell line.

As discussed in Section 1.1, cancer cells commonly present mutations that partially or fully impair regulation of cell-cycle transitions (or checkpoints). We are interested in using our model to characterise how defective checkpoints alter cell responses to different environmental conditions. To do so, we generate a cohort of *virtual* cell-lines which have either the G1 and S/G2 checkpoints mutated or both or none (see Table 2.3). In modelling our cell-lines, we assume they share all model parameters except for q_1, K_1, μ_2^C, p_{Sn} and $\mathcal{R}_S^+$. Shared parameters are fixed to the default values given in Tables 2.1 and 2.2 unless differently specified.

	G1 checkpoint	S/G2 checkpoints	Others
(+/+)	$q_1 = 0.9,\ K_1 = 0.05$	$\mathcal{R}_S^+ = 0.25,\ p_{Sn} = 0.1,\ \mu_2^C = 0.000$	Default values given in Tables 2.1 and 2.2
(+/−)		$\mathcal{R}_S^+ = 0.75,\ p_{Sn} = 0.9,\ \mu_2^C = 0.000$	
(+/x)		$\mathcal{R}_S^+ = 1.25,\ p_{Sn} = 0.1,\ \mu_2^C = 0.025$	
(−/+)	$q_1 = 0.9,\ K_1 = 1.00$	$\mathcal{R}_S^+ = 0.25,\ p_{Sn} = 0.1,\ \mu_2^C = 0.000$	
(−/−)		$\mathcal{R}_S^+ = 0.75,\ p_{Sn} = 0.9,\ \mu_2^C = 0.000$	
(−/x)		$\mathcal{R}_S^+ = 1.25,\ p_{Sn} = 0.1,\ \mu_2^C = 0.025$	
(x/+)	$q_1 = 0.2,\ K_1 = 1.00$	$\mathcal{R}_S^+ = 0.25,\ p_{Sn} = 0.1,\ \mu_2^C = 0.000$	
(x/−)		$\mathcal{R}_S^+ = 0.75,\ p_{Sn} = 0.9,\ \mu_2^C = 0.000$	
(x/x)		$\mathcal{R}_S^+ = 1.25,\ p_{Sn} = 0.1,\ \mu_2^C = 0.025$	

Table 2.3: List of the parameter values used for the numerical simulations. We generate a cohort of *in silico* cell lines by varying parameters q_1, K_1, $\mathcal{R}_S^+$, p_{Sn}. The last columns indicates where to find the values of the remaining model parameters. We identify parameter sets corresponding to cells with functional, partially-functional and defective G1 and G2/S checkpoints. For each cell-line we indicate first, mutations in the G1 checkpoint and then, mutations in the S/G2 checkpoints. For example, the notation (+/-) indicates a cell with functional G1 checkpoint while S/G2 checkpoints are only partially-functional.

Modelling mutations to the G1 checkpoint. We model cells with *functional* (+), *partially-functional* (-) and *defective* (x) G1 checkpoint by choosing different values for q_1 and $\mathcal{K}_1$ (see Table 2.3). We assume cells with a functional G1 checkpoint are likely to delay the G1→S transition when $c < c_H$ (high q_1) and they take a long time, after re-oxygenation, to re-enter the cell-cycle ($K_1 \ll k_1$). Cells with partially-functional (-) G1 checkpoint are assumed to arrest the G1→S transition when $c < c_H$ (high q_1) but they quickly re-enter the cell-cycle immediately after re-oxygenation

($K_1 \gg k_1$). Finally, cells with defective (x) G1 checkpoint lose the ability to delay the G1→S transition even when $c < c_H$ (low q_1 and $K_1 \gg k_1$).

Modelling mutations to the S/G2 checkpoints. We model cells with *functional* (+), *partially-functional* (-) and *defective* (x) S/G2 checkpoints by choosing different values for $\mathcal{R}_S^+$, p_{Sn} and μ_2^C (see Table 2.3). We assume that cells with functional S/G2 checkpoints have effective damage detection and repair mechanisms, which enables them to repair even more severe damage (low p_{Sn}, μ_2^C small), but also leads to prolonged delays in restarting DNA synthesis after replication stress (low $\mathcal{R}_S^+$). Cells with partially-functional (-) S/G2 checkpoints have less effective damage detection and repair mechanisms. As a result, they can repair only minor damage (μ_2^C small, p_{Sn} high) and experience intermediate delays in restarting DNA synthesis after replication stress (medium $\mathcal{R}_S^+$). Finally, we assume cells with defective (x) S/G2 checkpoints have lost the ability to detect (small p_{Sn}) and repair damage; as such, these cells attempt premature entry to mitosis (high μ_2^C) [101] and do not delay restart of DNA synthesis after replication stress (high $\mathcal{R}_S^+$).

2.3 Cell-cycle progression and growth in an oxygen-rich environment

We start by considering the case in which cells are exposed to a constant, oxygen-rich environment (i.e., $c(t) \equiv c_+ \gg c_H$). By assumption, all our *in silico* cell-lines have the same response to an oxygen-rich environment. We note that the results presented in this section have been published in [29].

Given Eqs. (2.3) and (2.1), and assuming $v_{S,0}(\xi) \equiv v_S^+$ and $m_0(\xi) \equiv 1.0$ for $\xi < 0$, we have that $v_i(t) \equiv v_i^+$ for $i = S, R$, and $\tau_S \equiv \tau_S^+ = 1/v_S^+$, while $\tau_R \equiv \tau_R^+ = y_R(1 + M_R)/v_R^+$. Furthermore, we have that $\mathcal{Q}_1 \approx 0$, $\mathcal{P}_2 \approx 0$. Assuming that the checkpoint compartments are initially empty (i.e., $C_{1,0} = D_{2,0} = \mathrm{Sn}_{2,0} = 0$) it follows that $C_1(t) = \mathrm{Sn}_2(t) = D(t) \equiv 0$ for all $t > 0$, so that there are no arrested cells. We conclude that when $c \equiv c_+ \gg c_H$, Eqs. (2.18) reduce to the following system:

$$\frac{dG_1}{dt} = 2k_2G_2 - k_1G_1, \quad t > 0, \tag{2.21a}$$

$$\frac{dS}{dt} = k_1G_1(t) - k_1G_1(t - \tau_S^+) \quad t > 0, \tag{2.21b}$$

$$\frac{dG_2}{dt} = -k_2G_2 + k_1G_1(t - \tau_S^+), \quad t > 0, \tag{2.21c}$$

subject to

$$G_1(\xi) = \begin{cases} G_{1,0}, & \theta = 0, \\ v_S^+ k_1^{-1} \mathtt{S}_0 \left(1 - \frac{\theta}{\tau_S^+}\right), & -\tau_S^+ < \theta < 0, \end{cases} \qquad S(0) = S_0 = \int_1^2 \mathtt{S}_0(x)dx, \quad G_2(0) = G_{2,0}. \tag{2.21d}$$

We note that Eqs. (2.21) are analogous to existing cell-cycle models, such as the model by Basse et al. [10, 11]. Since Eqs. (2.21) are linear, with a constant delay, they can be solved exactly via superposition of exponential functions $\boldsymbol{\xi}(\Lambda_i)e^{\Lambda_i t}$, where Λ_i $(i = 1, 2, \ldots)$ are the complex roots of the characteristic polynomial (see Eq. (2.25) in Section 2.3.1) and $\boldsymbol{\xi}(\Lambda_i)$ are the corresponding eigenvectors. In the case of DDEs, the characteristic polynomial is a transcendental equation, with an infinite number of roots and computing Λ_i is non-trivial. To investigate the transient dynamics, it is therefore more convenient to solve Eqs. (2.21) numerically. We use the Python *ddeint* package which integrates delay-differential equations. In Fig. 2.6 we present numerical solutions for two sets of initial conditions: cells are initially synchronised in either the G_1 (Figs. 2.6a-2.6c) or G_2 (Fig. 2.6d-2.6f) compartment. This corresponds to setting $\mathtt{S}_0 \equiv 0$ with $G_{1,0} = 1$ and $G_{2,0} = 0$ (top row), or $G_{2,0} = 1$ and $G_{1,0} = 0$ (bottom row).

Comparison of Figs. 2.6b and 2.6e shows that the evolution of the cell fractions π_z differs only up to time $t \approx 20$ hr; after this transient, the cell fractions evolve to constant values, denoted by π_z^+, which are independent of the initial conditions. By comparison, inspection of Figs. 2.6a and 2.6d shows that the evolution of G_1, S, G_2 and N differs, even at long times. For $t > 20$ hr, the variables exhibit similar qualitative behaviour, but values in Fig. 2.6d are higher than those in Fig. 2.6a. While cells synchronised in G_2 (see lower row in Fig. 2.6) start proliferating at the beginning of the simulations, cells initialised in G_1 (upper row) are delayed since they need to complete the S-phase before they can replicate. Once cells enter the G_2 compartment, the cell number N increases. Furthermore, for $t > 60$ hr, for both scenarios, $N(t)$, $G_1(t)$ and $G_2(t)$ increase exponentially at a constant rate. This is consistent with the results in Section 2.3.1, which show that, in the long term, the system converges to a regime of balanced exponential growth, with a population-growth rate λ_+ that is independent of the initial conditions.

The right-most panels in Fig. 2.6 show the evolution of the distribution $s(x, t)$

$$s(x,t) = \frac{\mathtt{S}(x,t)}{N(t)}, \quad 1 \le x \le 2, \tag{2.22}$$

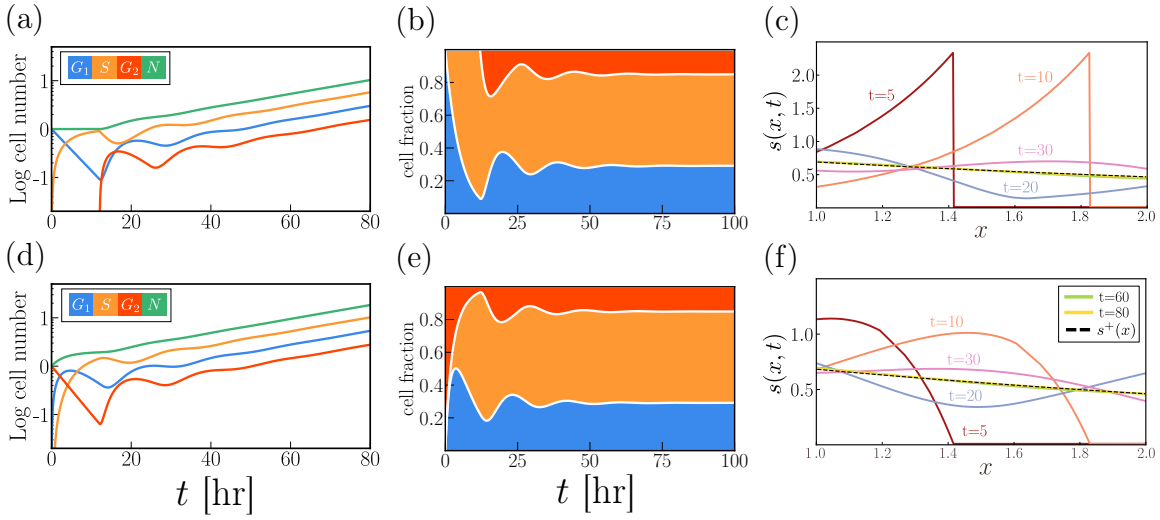


Figure 2.6: Numerical simulations of Eqs. (2.21) for two sets of initial conditions: (a-c) $\mathtt{S}_0 \equiv 0$, $G_{1,0} \equiv 0$ and $G_{2,0} = 1$; (d-f) $\mathtt{S}_0 \equiv 0$, $G_{1,0} \equiv 1$ and $G_{2,0} = 0$. For both simulations, the parameters k_1, k_2 and v_+ are set as in Table 2.1. (a) and (d) Time evolution of the cell numbers G_1, S, G_2 and N (see Eq. (2.19)) on a semi-logarithmic scale. (b) and (c) Cumulative plots of the cell fractions π_z with $z \in \{G_1, S, G_2\}$ (c) and (f) Profile of $s(x,t)$ (see Eq. (2.22)) at different time points (corresponding to the different coloured curves). We compare the long time behaviour with the analytically computed phase stationary solution, $s^+(x)$ (black dashed line).

where $\mathtt{S}(x,t)$ is defined by Eq. (2.7). The function $s(x,t)$ represents the probability that a cell in the S phase at time t has a given DNA content x (with $x \in [1,2]$). We note that in Fig. 2.6c, cells are initially highly synchronised in the S compartment, with the formation of a front that propagates at velocity v_S^+ (note the steep gradient in the profiles at time $t = 5$ and $t = 10$). This is because of the discontinuity in the initial data for G_1 (when $f_{1,0} = 1$ and $S_0 \equiv 0$, $G_1(\theta)$ in Eqs. (2.21) is discontinuous at $\theta = 0$). The discontinuity propagates along the x axis but then smooths out, due to de-synchronisation of cells in the G_1 and G_2 compartments. By contrast, in Fig. 2.6f, there is no discontinuity in the initial data for G_1 and, therefore, the profile for $s(x,t)$ is smoother. Despite these differences in the distributions at early times, $s(x,t)$ eventually evolves to the same stationary distribution (the curves for $t = 60$ and $t = 80$ are almost indistinguishable) and the time scales on which the system approaches the stationary distribution for the two initial conditions are comparable.

Following the notation introduced in [127], we term the asymptotic solution of the model a *phase stationary solution* (PSS) to indicate that, in this regime, the cell fractions π_{G1}, π_S and π_{G2}, and the distribution $s(x,t)$, remain constant in time. This is similar to predictions from other models in the literature [10, 12, 15, 40, 127] in the context of unperturbed growth; this regime is usually referred to as balanced,

or asynchronous, exponential growth [11, 16]. We note that our model does not account for resource competition and, therefore, only applies to the early phase of tumour formation or growth of cell cultures, before cells reach confluence (i.e., when competition for space can be neglected).

2.3.1 Balanced exponential growth and parameter estimates

We note from Eqs. (2.21) that the evolution of $S(t)$, the number of cells in the S phase, decouples from the equations for G_1 and G_2 so that we can consider the reduced model:

$$\frac{dG_1}{dt} = 2k_2G_2 - k_1G_1, \tag{2.23a}$$

$$\frac{dG_2}{dt} = -k_2G_2 + k_1G_1(t - \tau_S^+). \tag{2.23b}$$

Eqs. (2.23) form a system of linear delay differential equations with a constant (finite-size) delay; hence, their solutions can be written as a superposition of exponential functions, $\boldsymbol{\xi}(\Lambda_i)e^{\Lambda_i t}$, with eigenvalues $\Lambda_i \in \mathbb{C}$ that constitute the spectrum σ of Eq. (2.23) and corresponding eigenvector $\boldsymbol{\xi}_i(\Lambda_i) \in \mathbb{C}^{2\times 2}$. Further, the long time behaviour of the system will be dominated by the eigenvalue Λ_i with the largest real part, here denoted by λ_+. As shown in Section 1.2.3, Eqs. (2.23) define a strongly positive linear dynamical system which implies the dominant eigenvalue λ_+ is real and simple, and the corresponding eigenvector $\boldsymbol{\xi}(\lambda_+)$ has real and positive entries. The eigenvalues of Eqs. (2.23) are the roots of the characteristic equation:

$$\det(A(\Lambda) - \Lambda \mathrm{I}) = 0, \quad A(\Lambda) = \begin{bmatrix} -k_1 & 2k_2 \\ k_1 e^{-\Lambda\tau_+} & -k_2 \end{bmatrix} \tag{2.24}$$

where I is the identity matrix in $\mathbb{R}^{2\times 2}$. Evaluating the determinant explicitly we obtain the following transcendental equation for the eigenvalues Λ of Eqs. (2.23):

$$P_+(\Lambda) \equiv -(\Lambda + k_1)(\Lambda + k_2) + 2k_1k_2e^{-\Lambda\tau_S^+} = 0. \tag{2.25}$$

It can be shown that λ_+ is positive, so that the solution will eventually blow up, i.e. $\lim_{t\to\infty} G_1, G_2 = \infty$. Lemma 2.3.1 summarises the results on the asymptotic dynamics of Eqs. (2.23).

Lemma 2.3.1. *For any values of $k_1 > 0$, $k_2 > 0$ and $\tau_S^+ > 0$, the right-most root of $P_+(\Lambda)$, as defined by Eq. (2.25), is real and positive.*

Proof. As shown in Examples 1.2.1 in Section 1.2, Eqs. (2.23) define a strongly positive dynamical system with finite-size delays which implies that the dominant root of $P_+(\Lambda)$ is real and simple. This result can also be proven algebraically by manipulating the structure of P_+ (see [29]). We now prove that the latter must be positive, which is equivalent to showing that P_+ as at least one positive root. This is straightforward to prove since $P_+(\Lambda)$ is continuous and $P_+(0) > 0$, while $\lim_{\Lambda\to\infty} P_+(\Lambda) < 0$. Given that for $\Lambda > 0$, $dP_+/d\Lambda < 0$ (*i.e.*, P_+ is strictly monotonically decreasing), we have the stronger statement that λ_+ is the only real eigenvalue which is positive.

□

Based on Lemma 2.3.1, we know that for any choice of model parameters, under unperturbed growth, cells will eventually reach a regime in which the population grows exponentially at a constant rate λ_+. This is a common result of many cell-cycle models for *in vitro* systems and it is usually referred to as balanced or asynchronous exponential growth:

$$G_{1,2}(t) \approx \xi_{G_1,G_2} e^{\lambda_+ t}, \quad t \gg 1, \tag{2.26a}$$

where ξ_{G_1,G_2} are positive constants and the entries of the eigenvector $\boldsymbol{\xi}(\lambda_+)$. Having characterised the long time behaviour of $G_1(t)$ and $G_2(t)$, let us discuss what this implies for the other model variables, i.e. the distribution $\mathtt{S}(x,t)$ and the total number of cells, $N(t)$. Using Eq. (2.7) we find that, in the case of unperturbed growth and assuming $t > \tau_S^+$, the long time distribution of $\mathtt{S}(x,t)$ takes the form:

$$\mathtt{S}(x,t) \approx \underbrace{\tau_S^+ k_1 \xi_{G_1} e^{-\lambda_+ \tau_S^+ (x-1)}}_{\hat{s}(x)} e^{\lambda_+ t}, \quad t \gg 1, \tag{2.26b}$$

which can be written, by separating the DNA and time components, as $\mathtt{S}(x,t) \approx s^+(x)e^{\lambda_+ t}$. This implies that the number of cells in the S phase also grows exponentially, $S(t) = \xi_S e^{\lambda_+ t}$, where $\xi_S = \int_1^2 s^+(x)dx$. Combining this with Eq. (2.26a), we can compute the total number of cells and the cell fractions:

$$N(t) \approx (\xi_{G_1} + \xi_{G_2} + \xi_S) e^{\lambda_+ t} = \xi_N e^{\lambda_+ t}, \quad t \gg 1, \tag{2.26c}$$

$$\pi_z(t) \approx \pi_z^+ = \frac{\xi_z}{\xi_N}, \quad z \in \{G_1, S, G_2\}, \, t \gg 1. \tag{2.26d}$$

We find that the long time behaviour of the system is characterised by a stationary DNA-distribution, $s^+(x)$, of cells in the S phase, and constant cell fractions, π_z^+, which are independent of the chosen initial conditions. Following [149], we denote this regime as asynchronous exponential growth, as cells grow exponentially while the fraction of

cells in each phase of the cell-cycle remains constant. The distribution of cells in the S phase, this converges to the stationary distribution $s^+(x) = \tau_S^+ k_1 \pi_{G_1}^+ e^{-\lambda_+ \tau_S^+ (x-1)}$, which is monotonically decreasing in x (see black line in Figs. 2.6c and 2.6f). This indicates that, for $t \gg 1$, cells in the S phase are more likely to be starting DNA synthesis (with $x \approx 1$) than completing it (with $x \approx 2$). We note that the longer the duration of the S phase, or, equivalently, the smaller v_S^+, the steeper is the $s^+(x)$ curve and, therefore, the larger is the proportion of cells in the S phase that are concentrated around $x \approx 1$. In the limit where the DNA synthesis velocity is high, i.e. $\lambda_+ \tau_S^+ \to 0$, the distribution $s^+(x)$ flattens and cells are distributed more uniformly along the S phase.

2.3.2 Estimation of oxygen independent parameters

We now illustrate how we can use the phase stationary solution computed in Section 2.3.1 to estimate the model parameters k_1, k_2 and v_S^+ (or equivalently τ_S^+) associated with unperturbed progression along the mitotic cell-cycle.

We assume that cells have grown in an unperturbed environment for a sufficiently long time that they have reached the regime of balanced exponential growth. If we know the fraction of cells in each cell-cycle phase, π_j^+, and the growth rate, λ_+, of the population, we can uniquely identify the parameters k_1, k_2, τ_S^+ in the model. If we substitute the solution (2.26) into (2.21), then, upon re-writing $G_i \approx \pi_i^+ \xi_N e^{\lambda_+ t}$, we obtain the following algebraic system:

$$\lambda_+ \pi_{G_2}^+ = -k_2 \pi_{G_2}^+ + \pi_{G_1}^+ k_1 e^{-\tau_S^+ \lambda_+}, \tag{2.27a}$$

$$\lambda_+ \pi_{G_1}^+ = 2k_2 \pi_{G_2}^+ - k_1 \pi_{G_1}^+, \tag{2.27b}$$

$$\lambda_+ = k_2 \pi_{G_2}^+. \tag{2.27c}$$

From Eq. (2.27c) it is clear that k_2 is uniquely identified and positive. Substituting k_2 into Eq. (2.27b), we obtain an equation for k_1:

$$k_1 = \frac{2 - \pi_{G_1}^+}{\pi_{G_1}^+} \lambda_+, \tag{2.28}$$

where we note that $k_1 > 0$ since, by definition, $\pi_{G_1}^+ \leq 1$. Substituting the forms of k_1 and k_2 into Eq. (2.27a) we obtain an expression for τ_S^+:

$$\tau_S^+ = \frac{1}{\lambda_+} \ln \left(\frac{2 - \pi_{G_1}^+}{\pi_{G_2}^+ + 1} \right) \tag{2.29}$$

where the physical constraint $\pi^+_{G_1} + \pi^+_{G_2} \leq 1$ guarantees that the argument of the logarithm is always positive. Hence, given the measurements of $\pi^+_{G_1}$, $\pi^+_{G_2}$ and λ_+ we can uniquely identify the constant parameters in Eqs. (2.21). In practice, we can estimate the stationary values $\pi^+_{G_1}$ and $\pi^+_{G_2}$ using flow cytometry data. However, additional data are needed to determine the growth rate λ_+. This parameter can be related to the population doubling time, T_{doub} [10], which is known for most cell lines when cultured in standard media and in the absence of competition (i.e., low confluence). Indeed, it is straightforward to show that $T_{doub} = \lambda_+^{-1} \ln 2$.

	transition rates [1/hr]	average duration [hr]	cell fractions (from [5])
G1	$k_1 = 0.195$	$\tau_1 \approx k_1^{-1} \approx 5.14$	$\pi^+_{G_1} = 0.29$
S	$v_S^+ = 0.083$	$\tau_S^+ = v_+^{-1} \approx 12$	$\pi^+_S = 0.56$
G2/M	$k_2 = 0.22$	$\tau_2 \approx k_2^{-1} \approx 4.54$	$\pi^+_{G_2} = 0.15$

Table 2.4: Estimates of the cell-cycle parameters for the RKO cell line obtained using the phase stationary solution (PSS) and the experimental data from [5].

We use the values reported in [5] to estimate $\pi^+_{G_1}$ and $\pi^+_{G_2}$. For simplicity, we suppose that median values provide a good approximation to the 'true' cell fractions so that $\pi^+_{G_1} \sim 0.29$ and $\pi^+_{G_2} \sim 0.15$. We note also that the doubling time of RKO cells has been estimated to be about 21 hours [152] (correspond to $\lambda_+ \approx 0.033$ [1/hr]). The corresponding parameter estimates (obtained using Eq. (2.27)) are listed in Table 2.4 together with estimates of the duration of each cell cycle phase (given by the inverse of the rates of k_1, k_2 and v_S^+ [10]). We note that the RKO cell line is characterised by a particularly long S phase. By contrast, the durations of the G1 and G2/M phases are almost half the duration of the S phase, with G2/M being the shortest.

Having characterised the cell-cycle and growth dynamics when cancer cells are exposed to oxygen rich environment, we now use the model to explore how cell-cycle and growth dynamics are perturbed in hypoxia.

2.4 Characterising cell-cycle progression in different oxygen environments

In this section, we use numerical simulations of our model to investigate cell-cycle and growth dynamics in different hypoxic environments. In Chapter 4, we will use results on the asymptotic dynamics of positive linear systems with delays to study the long time behaviour of the solutions.

2.4.1 Predicting cell responses to constant hypoxia

Oxygen dynamics. We consider the case in which the cells are exposed to a constant, oxygen-deprived environment for $t \geq 0$ (i.e., by neglecting the first transient re-equilibration of the oxygen $c(t) \equiv c_- \ll c_H$ for $t > 0$).

Initial and history conditions. We assume that cells are initially in a regime of asynchronous exponential growth, which mimics cells being exposed to high oxygen levels for $t < 0$ ($c(t) \gg c_H$ for $t < 0$). We, therefore, initialise our simulations by assuming that the number of cells in each compartment is dictated by the phase stationary solution computed in Section 2.3.1. Without loss of generality, we fix $N(0) = 1$ corresponding to rescaling the number of cells in each compartment by the size of the population at time $t = 0$. Then, we have the following initial/history conditions:

$$\begin{aligned} &G_1(0) = \pi^+_{G_1}, \quad G_2(0) = \pi^+_{G_2}, \quad C_1(0) = \mathrm{Sn}_2(0) = 0, \\ &\mathcal{S}_0(x) = \tau^+_S k_1 \pi^+_{G_1} e^{-\lambda_+ \tau^+_S (x-1)}, \quad x \in (1, 2], \\ &v_S(\xi) = v^+_S, \quad m(\xi) = 1.0, \quad \xi < 0. \end{aligned} \tag{2.30}$$

We note that, for initial conditions (2.30), Eqs. (2.18) are non-autonomous, at least during the initial transient dynamics. For $t \gg 1$, transition rates between compartments and delays tend to constant values and the full model reduces to a system of linear delay-differential equations, with two constant delays (this model will be studied further in Chapter 4).

Set-up of the numerical simulations. We simulate the responses of our cohort of *in silico* cell-lines to constant hypoxia (see Section 2.2.7.1) by solving Eqs. (2.18) and (2.30) numerically using the Python package *ddeint* to integrate the delay-differential equations. In the simulations, parameter values are set as indicated in Table 2.3, except for the death rate of cells in the S phase, μ_S, that we allow to vary. We note that increasing μ_S corresponds to cells in the S phase being more sensitive to fork collapse in hypoxia. In hypoxia, cells with functional or partially-functional G1 arrest behave the same. We therefore report only results for cell-lines with functional and defective G1 checkpoint (see Section 2.2.7.1).

2.4.1.1 Dysregulation of cell culture growth in constant hypoxia

In Fig. 2.7, we compare the time evolution of the total population size, $N(t)$, when exposing our *in silico* cell-lines to constant hypoxia. Different panels correspond to a

different value of μ_S.

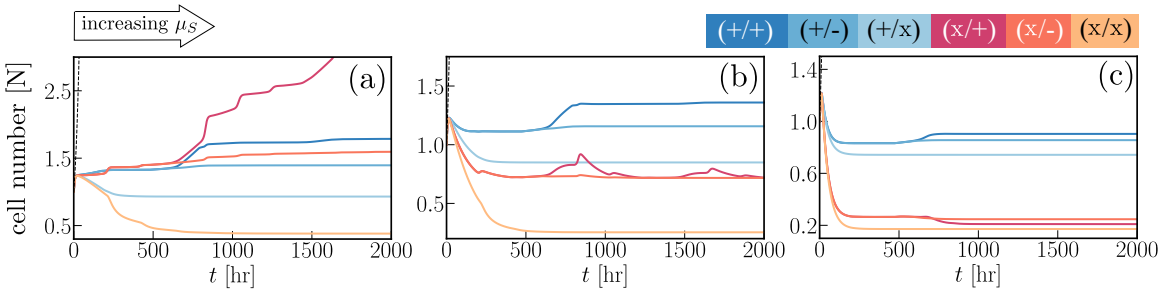


Figure 2.7: (a)-(c) Comparison of the evolution of the total cell number, $N(t)$, obtained by solving Eqs. (2.18) and (2.30), for different cell-lines (parameter values given in Table 2.3). The value of the death rate μ_S increases from left to right: (a) $\mu_S = 5 \times 10^{-4}$, (b) $\mu_S = 5 \times 10^{-3}$ and (c) $\mu_S = 2.5 \times 10^{-2}$. Note that the scaling of the y-axis varies amongst panels.

When μ_S is sufficiently small (see Fig. 2.7a), we observe significant variability in cell responses. For cell-lines (+/x) and (x/x), the total population size decreases over time due to cell-death in the D_2 compartment. This is more evident for (x/x) than (+/x) cells, implying that having a functional G1 checkpoint can protect cells from death. For (+,+), (+/-) and (x/-) cells, we observe complete growth inhibition at long times, with N converging to a constant value. In contrast, for (x/+) cells, which are characterised by functional S/G2 checkpoints but defective G1 checkpoint, the population size grows unbounded. As the value of μ_S increases (see Fig. 2.7b and 2.7c), there is less variability in cell responses. Specifically, for all cells, the total population size decreases towards a constant values. Nonetheless, based on the asymptotic dynamics of N, we find can divide cell-lines into two groups: while cell lines (+/+), (+/-) and (+/x) are characterised by a small loss of cells, cell death is much more pronounced for cell lines with defective G1 checkpoint (*i.e.*, (x/+), (x/-) and (x/x)). This is to be expected as the hypoxia-mediated G1 arrest has a preventive role, protecting cells from entering the S phase in disadvantageous conditions. We find instead that the S/G2 checkpoints have little influenced on cell responses. Interestingly, while (x/+) cells were the most "advantaged" for small values of μ_S, as μ_S increases they become one of cell populations that is most sensitive to cell-death in constant hypoxia. From a biological point of view, we expect μ_S to vary between cell-lines, depending on their sensitivity to fork collapse in constant (chronic) hypoxia. This suggests that having a functional G1 checkpoint is beneficial only if a cell entering the S phase is more likely to die prior to division than to proliferate (*i.e.*, large μ_S).

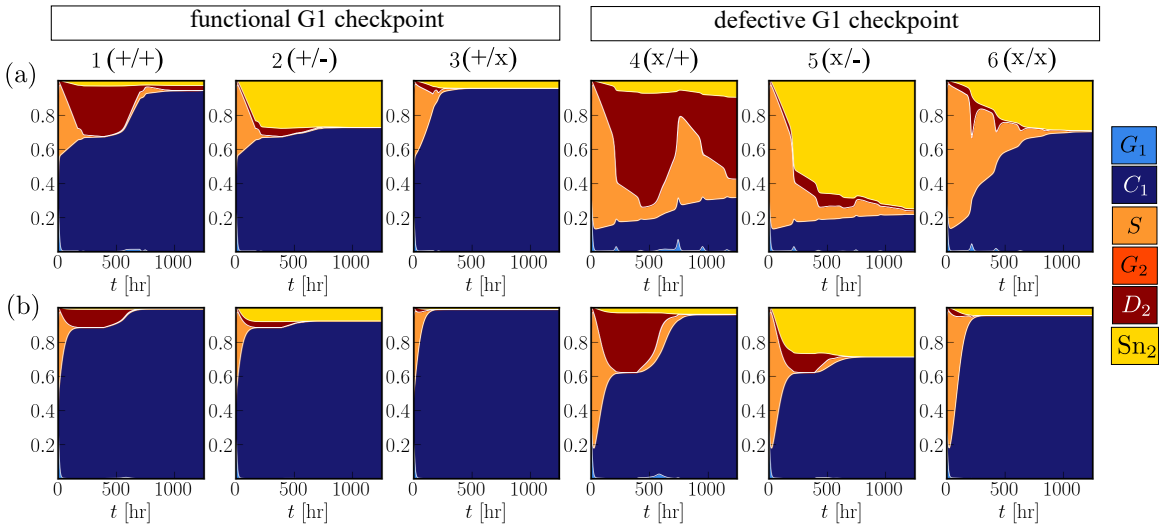


Figure 2.8: (a)-(b) Cumulative plot of the cell fractions, π_z $z \in \{G_1, C_1, S, G_2, D_2, \mathrm{Sn}_2\}$, indicating how the total population of cells, N, is distributed in between model compartments. These are estimates solving Eqs. (2.18) and (2.30) for the parameter sets listed in Table 2.3. The death rate μ_S is set to (a) $\mu_S = 5\times10^{-4}$ (as in Fig. 2.7a), (b) $\mu_S = 2.5 \times 10^{-2}$ (as in Fig. 2.7c).

2.4.1.2 Dysregulation of cell-cycle dynamics in constant hypoxia

In Fig. 2.8 we show how the composition of cell cultures (*i.e.,* how cells are distributed between the different model compartments) changes over time when these are exposed to constant hypoxia. On the top row, we compare the responses of our cohort of cell-lines for $\mu_S = 5\times10^{-4}$ (as in Fig. 2.7a), while on the bottom row we fix $\mu_S = 2.5\times10^{-2}$ (as in Fig. 2.7c).

Focusing on the top row first, we find that for all populations, except for the (x/+) population, eventually only consists of arrested cells (*i.e.*, C_1 and Sn_2 cells). We note that, in our model, under constant hypoxia, the C_1 compartment acts as a form of permanent arrest (since $\mathcal{K}_1 = 0$). For (+/+) and (+/x) populations, at the final time, almost 90% of cells are arrested in C_1, while for the (+/-) and (x/x) populations the asymptotic fraction of C_1 cells is lower ($\approx 60 - 70\%$). Interestingly, despite (+/-) and (x/x) populations have a similar asymptotic composition, their transient growth dynamics is significantly different (see Fig. 2.7a): while the number of (+/-) cells remains approximately constant due to cells arresting in the G1 phase, the number of (x/x) cells significantly decreases up to 500 hr after the beginning of our simulations due to death of D_2 cells in the G2 phase. Populations with similar internal structure can therefore arise from different transient dynamics. For the (x/+) population, the model predicts that a significant fraction of the population is actively

cycling (in particular a significant proportion of cells are in the S/G2 phase). This is reflected at the population level by an increase in the (x/+) population size (see pink curve in Fig. 2.7a). Interestingly, for the first $t \approx 600$ hours, the dynamics of the (x/+) and (+/+) populations are almost identical (see Fig. 2.7a), although their internal compositions are different. Similar-population-level dynamics can therefore be associated with populations with significantly different internal compositions. This suggests that data on cell numbers alone might not be always sufficient to distinguish cells with different cell-cycle progression in hypoxia.

As μ_S increases (see Fig. 2.8b), the model predicts less variability in the evolution of the internal composition: all populations eventually only consists arrested cells are present. In general, the fraction of cells arrested in the C_1 compartment is larger in Fig. 2.8b than in Fig. 2.8a, while the fraction of Sn_2 cells is smaller. This can be readily explained as follows: by increasing μ_S, cells are likely to die in the S phase before entering the G2 phase. While the long time behaviour seems to be almost independent of the whether cells have a functional or defective G1 checkpoint, this affects the initial phase (first 100 hours). For cell-lines with functional G1 checkpoint, the fraction of cells in C_1 (π_{C_1}) increases monotonically throughout the simulation. In contrast, in cell-lines with defective G1 checkpoint the increase in π_{C_1} is delayed as cells tend to first accumulate in the S compartment. Unlike before, the increase in π_{C_1} is not driven by hypoxia-mediated G1 arrest but rather by the large number of cells that die in the S phase.

Overall, our numerical simulations highlight that our cell-cycle model captures a variety of responses to constant hypoxia that may be representative of different cell-lines. Furthermore, our results suggest that for cells unable to cope with replication stress (high μ_S), having functional checkpoints is advantageous as it protects them from death. In contrast, for cell lines that are more tolerant to hypoxia-induced replication stress (small μ_S), bypassing G1 arrest can be beneficial by allowing for slow, but sustained, population growth (see cell-line 4 in Fig. 2.7a). However, this is true if cells have functional S/G2 checkpoints.

2.4.2 Predicting cell responses to cyclic hypoxia

We now use our model to investigate dysregulation of the cell-cycle and growth dynamics in cyclic (*i.e., periodic*) hypoxia.

Oxygen dynamics. Compared to the case of constant hypoxia, we now have more degrees of freedom, which includes the unknown model parameters as well as the oxygen dynamics. For simplicity, we restrict attention to periodic oxygen fluctuations, similar to those that are commonly used experimentally. The dynamics of the oxygen levels, $c = c(t)$, at time $t > 0$ are given by:

$$\frac{dc}{dt} = \begin{cases} \mathcal{R}_{ox}(c_+ - c), & \mathcal{T}_- < \mathrm{mod}(t, \mathcal{T}_- + \mathcal{T}_+) < \mathcal{T}_- + \mathcal{T}_+, \\ \mathcal{R}_{ox}(c_- - c), & 0 \ < \mathrm{mod}(t, \mathcal{T}_- + \mathcal{T}_+) < \mathcal{T}_-, \end{cases} \tag{2.31}$$

where $\mathcal{R}_{ox}$, $\mathcal{T}_\pm$ and $c_\pm$ are a positive constant and $\mathrm{mod}(a, b)$ is the modulo operator which returns the remainder of dividing a by b. We here assume that oxygen levels equilibrate fast (on the timescale of minutes) by taking $\mathcal{R}_{ox} = 10$ [1/hr] so that, the oxygen levels $c(t)$ fluctuate with period $\mathcal{T} = \mathcal{T}_- + \mathcal{T}_+$ between $c_+ > c_H$ and $c_- < c_H$. Furthermore, $\mathcal{T}_-$ and $\mathcal{T}_+$ are representative of, respectively, the times that cells spend in hypoxia and in a re-oxygenated environment in each oxygen cycle. Without loss of generality we consider $c_+ = 2.1\% O_2$ and $c_- = 0.1\% O_2$ following the protocol adopted by our experimental collaborator (more details on the experimental set up are given in Section 3.5).

Initial and history conditions. As for the case of constant hypoxia considered in Section 2.4.1, we assume that the cells are originally in a regime of exponential growth so that we start again our simulation with the initial conditions given by Eqs. (2.30).

Set-up of the numerical simulations. We simulate the responses of our cohort of *in silico* cell-lines to cyclic hypoxia environments (see Section 2.2.7.1) solving Eqs. (2.18) and (2.30) numerically using the Python package *ddeint* to integrate the delay-differential equations. Parameter values used in the simulations are as given in Table 2.3. We simulate different cyclic hypoxia environments solving Eq. (2.31) for a range of choices of $\mathcal{T}_+$ [hr] and $\mathcal{T}_-$ [hr].

2.4.2.1 Dysregulation of cell culture growth in cyclic hypoxia

In Fig. 2.9, we present the predicted dynamics of the total population size, $N(t)$, when exposing cell cultures to different cyclic hypoxia environments. During the early transient ($t \approx 30$ hours), the model predicts all populations experience growth arrest; while most cells behave similarly, mutations associated with the S/G2 checkpoints (specifically the value of $\mathcal{R}_S^+$) can lead to differences in the growth dynamics. In

particular, we find that cells with defective S/G2 checkpoints initially benefit from being able to quickly restart DNA synthesis upon re-oxygenation.

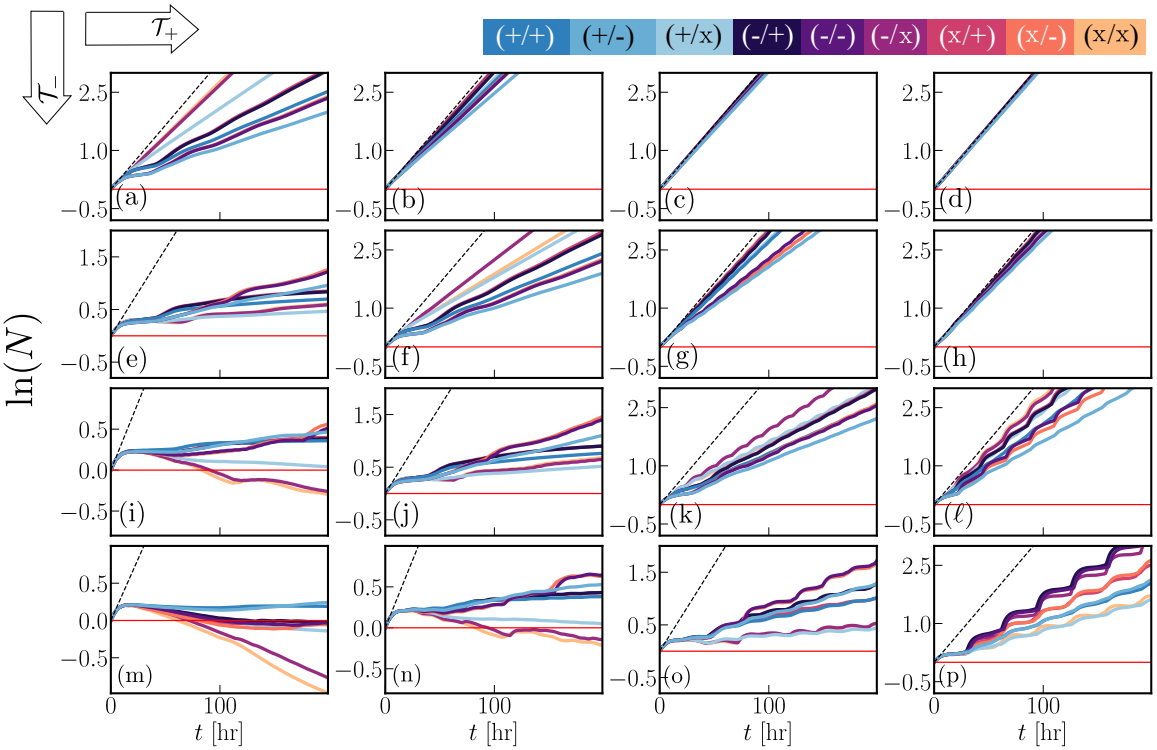


Figure 2.9: (a)-(p) Time evolution of the total number, N, in cell cultures grown in cyclic hypoxia obtained by simulating Eqs. (2.18) and (2.30)-(2.31). Each colour indicates simulation for a different choice of parameter values (or cell-line); parameter values are listed in Table 2.3. From left to right, we increase $\mathcal{T}_+$, *i.e.*, the period in-between exposure of cells to hypoxia: (first column) $\mathcal{T}_+ = 0.5$ hr, (second column) $\mathcal{T}_+ = 2$ hr, (third column) $\mathcal{T}_+ = 8$ hr and (forth column) $\mathcal{T}_+ = 16$ hr. From top to bottom, we increase $\mathcal{T}_-$: (first row) $\mathcal{T}_- = 0.5$ hr, (second row) $\mathcal{T}_- = 2$ hr, (third row) $\mathcal{T}_- = 5$ hr and (fourth row) $\mathcal{T}_- = 16$ hr. The red horizontal curve indicated $N = 1$ [cell], while the dotted black line corresponds to the evolution of N if cells were to be maintained in oxygen-rich conditions instead of being exposed to hypoxia, *i.e.*, $N = e^{\lambda_+ t}$ [cell].

Cell responses are more variable when considering the long time dynamics. On the one hand, we find that responses of the same cell-line vary when grown in different cyclic hypoxia conditions. Furthermore, the same oxygen protocol can drive a range of responses in different cell-lines. For example, in Fig. 2.9n, while the (-/-) population grows slowly, the size of the (+/+) population is approximately constant and the number of (x/x) cells decays. Nonetheless, we can identify general trends.

By increasing $\mathcal{T}_-$, cells are exposed to increasingly toxic environmental conditions and population growth slows down. In contrast, increasing $\mathcal{T}_+$ decreases toxicity,

leading to higher population growth rates. Once cells are allowed sufficient time to recover in between brief periods of hypoxia (*i.e.*, $\mathcal{T}_+$ is sufficiently high and $\mathcal{T}_-$ sufficiently low), the evolution of N is similar to that of cells grown in oxygen-rich conditions (see the black dotted line). In this regime, we find little difference between cells with functional and those with dysfunctional checkpoints. For intermediate durations of hypoxia $\mathcal{T}_-$ and $\mathcal{T}_+ \approx \mathcal{T}_-$ the model predicts a decrease in the growth rate for all cell populations (see Figs. 2.9e and Figs. 2.9i) as a result of larger delays in the G1→S and S→G2 transitions. However, while some populations continue to grow in cyclic hypoxia, other appear to arrest or even decay in size. Interestingly, while growth arrest characterises the dynamics of several cell-lines, the biological mechanisms that drive arrest are different: for (+/-) cells, homeostasis occurs because cells withdraw from the cell-cycle (see Fig. 2.9i); in contrast, for (x/x) cells, homeostasis is achieved via a balance between cell death and proliferation (see Fig. 2.9e). If we increase the period of exposure to hypoxia, $\mathcal{T}_-$, even further (see Fig. 2.9m), cells become sensitive to fork collapse and, for most cell-lines, this results in large numbers of cells dying.

(a)

$\mathcal{T}_-$ \ $\mathcal{T}_+$	0.5 hr	2 hr	8 hr	16 hr
0.5 hr	(x/x)	(x/x)	(x/x)	(x/x)
2 hr	(x/-)	(-/x)	(x/x)	(x/x)
5 hr	(x/-)	(x/-)	(-/x)	(x/x)
16 hr	(+/-)	(-/-)	(-/-)	(-/+)

(b)

$\mathcal{T}_-$ \ $\mathcal{T}_+$	0.5 hr	2 hr	8 hr	16 hr
0.5 hr	(+/-)	(+/-)	(+/-)	(+/-)
2 hr	(+/x)	(+/-)	(+/-)	(+/-)
5 hr	(x/x)	(+/x)	(+/-)	(+/-)
16 hr	(x/x)	(x/x)	(+/x)	(+/x)

Figure 2.10: (a) For each panel in Fig. 2.9, we indicate the cell population with the largest number of cells at the end of the simulations; (b) same as (a) but we indicate the cell population with the lowest number of cells. We use the same colour-scheme as in Fig. 2.9 to indicate different cell-lines.

The tables in Fig. 2.10 indicate which cell-lines reach, respectively, the largest (Fig. 2.10a) and the smallest (Fig. 2.10b) population size at the end of our simulations. When toxicity is high (i.e., $\mathcal{T}_- \gg 1$), having defective checkpoints is disadvantageous. This is analogous to what we observed when exposing cell-lines to constant hypoxia. By contrast, in mildly toxic environments, the opposite holds and cells benefit from lacking checkpoint control mechanisms. In transitioning between these two extremes, cells that have defective S/G2 checkpoints become increasingly disadvantaged as toxicity increases. Less intuitive is the role of mutations in the G1 checkpoint. Overall, cells seem to benefit from bypassing G1 arrest in hypoxia, but there are intermediate levels of toxicity for which cells with partially-functioning G1

checkpoint give rise to larger populations than cells with defective G1 checkpoint. Interesting, we find that the effect of a new mutation in the G1 (S/G2) checkpoint(s) depends on whether the S/G2 (G1) checkpoints are (is) functioning. For example, (x/-) cells are better off in a toxic environment (see row 3, column 1 of Fig. 2.10a). We conclude that having a defective G1 checkpoint is favourable. However, this is not true for cells with defective S/G2 checkpoints (see row 3, column 1 of Fig. 2.10b). Indeed, a double (x/x) mutant is worst off when grown under the same environmental conditions. It is, therefore, non-trivial to predict how evolution might shape the cell-cycle control strategy adopted by cells exposed to cyclic hypoxia.

2.4.2.2 Dysregulation of cell-cycle dynamics in cyclic hypoxia

We now investigate the time-evolution of the internal composition of cell populations grown under different forms of cyclic hypoxia. In Fig. 2.11, we present the evolution of the fraction of cells in each model compartment for the (+/-) population. In Figs. 2.11a-2.11d, we fix the duration of the hypoxia phase to $\mathcal{T}_- = 2$ hours, while in Figs. 2.11e-2.11h, $\mathcal{T}_- = 16$ hours. Moving along each row, the period of recovery between phases of hypoxia, $\mathcal{T}_+$, increases from $\mathcal{T}_+ = 0.5$ to $\mathcal{T}_+ = 16$ hours.

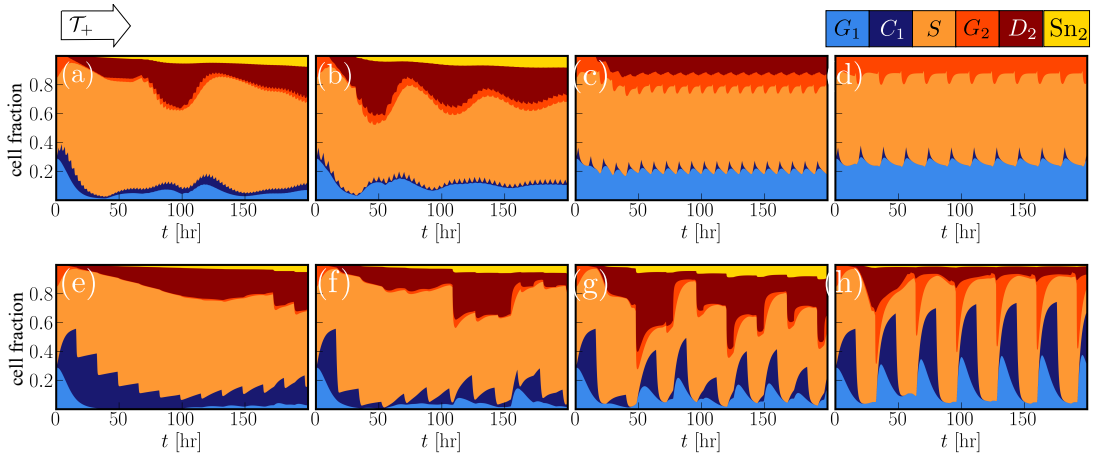


Figure 2.11: (a)-(h) Time-evolution of the fraction of cells in each model compartment. Plots are obtained solving Eqs. (2.18) and (2.30)-(2.31) for the parameter set (-,+) in Table 2.3. From left to right, we increase $\mathcal{T}_+$ (see Eq. (2.31)): (a) and (e) $\mathcal{T}_+ = 0.5$; (b) and (f) $\mathcal{T}_+ = 2$; (e) and (g) $\mathcal{T}_+ = 8$; (d) and (h) $\mathcal{T}_+ = 16$ hours. (a)-(d) We fix $\mathcal{T}_- = 2$ (as in Figs. 2.9e-2.9h); (e)-(h) $\mathcal{T}_- = 16$ hours (bottom row and as in Figs. 2.9m-2.9p).

Focusing on the top row, for all values of $\mathcal{T}_+$, the acute response to cyclic hypoxia leads to a transient increase in the fraction of cells in the S phase (due to the reduced rate of DNA synthesis). For small values of $\mathcal{T}_+$ (see Figs. 2.11a-2.11b), the fraction

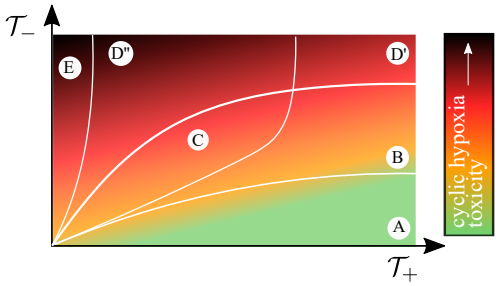
of cells in the G_1 and G_2 compartments rapidly decreases and by time $t = 50$ hours these two compartments are effectively empty. After approximately $t = 20$ hours, as cells complete the S phase, we see accumulation of damaged cells in the G2 phase (see increase in the fraction of D_2 and Sn_2 cells). As D_2 cells complete repair, at about $t = 80$ hours, the number of cells in G_1 starts to increase; this corresponds to the re-start of proliferation observed in Fig. 2.9e (see dark-blue curve). The increased duration of the cell-cycle, due to delays in the G1→S, S→G2 and G2→M transitions, drives low frequency damped fluctuations in the cell fractions. While we expect the latter to damp down with time, high frequency variations in the cell fractions are expected to persist as they are due to the instantaneous changes in the oxygen levels. As we increase $\mathcal{T}_+$ (see Fig. 2.11b), the evolution of the cell-cycle distribution is qualitatively similar. Nonetheless, since reoxygenation periods are longer, cells spend less time in the S and D_2 compartments so that the low frequency fluctuations in the cell-cycle composition occur on faster time-scales than in Fig. 2.11a. As we increase $\mathcal{T}_+$ to be much larger than $\mathcal{T}_-$ (as in Figs. 2.11c-2.11d), the impact of cyclic hypoxia on the cell-cycle dynamics changes. First, in Fig. 2.11c, we observe no accumulation of senescent cells and a significant increase in the number of cells in the G_2 compartment. By further increasing $\mathcal{T}_+$, we do not observe accumulation of damaged D_2 cells but we can still detect a significant change in the cell-cycle distribution due primarily to the variable duration of the S phase. The dynamics become more complex when we increase $\mathcal{T}_-$ to 16 hours. In general, we find more rapid and significant changes in the internal composition of the cell population. Even for large value of $\mathcal{T}_+$, we now observe accumulation of D_2 (damaged) cells in the G2 phase (see in Figs. 2.11h) and C_1 cells in the G1 phase. In this regime, we find that cells tend to naturally synchronise in their progression along the cell-cycle with time intervals during which the population consists almost entirely of cells in the S phase. These was not observed when considering shorter exposure to hypoxia (compare the top and bottom row of Fig. 2.11) for which cells that to be more evenly distribution across model compartments.

2.5 Discussion

In this chapter, we presented a novel mathematical model to describe the growth of a population of cells structured by their cell-cycle state when exposed to different oxygen environments. The key mechanisms captured in our model are: regulation of different cell-cycle transitions in hypoxia, as well as death due to fork collapse and

premature mitosis. In Sections 2.3 and 2.4, we used numerical simulations to explore the range of behaviours exhibited by the model under different oxygen conditions: a physiological, oxygen-rich environment, and constant or cyclic exposure to hypoxia as observed in pathological tissues. In physiological conditions, the model reduces significantly in complexity and exhibits behaviour similar to those of existing cell-cycle models: cell cultures asymptote to a regime of asynchronous exponential growth. In conditions of constant hypoxia, we find that population growth is significantly slowed down. Nonetheless, the model exhibits a variety of responses, including exponential growth, decay and confluence. In contrast to existing models that capture the impact of hypoxia on cell viability, our model includes a delay in cell death based on experimental observations [114]. While such a delay does not affect the long-term growth dynamics of cell populations in constant hypoxia, it does play a significant role when exposure to hypoxia is periodic.

The most interesting application of our model is to conditions of cyclic hypoxia, which have not been explored in previous studies. Here we focus on cyclic conditions where oxygen fluctuations are periodic with period $\mathcal{T} = \mathcal{T}_+ + \mathcal{T}_-$. In our simulations, cells are repeatedly exposed to hypoxia for $\mathcal{T}_-$ hours and then re-oxygenated for $\mathcal{T}_+$ hours. Our findings are summarised in the schematic in Fig. 2.12 where we divide the $(\mathcal{T}_+, \mathcal{T}_-)$-space into regions that yield qualitatively different cellular responses. We note that this a qualitative, rather than quantitative, representation of the findings from Section 2.4.2. From Fig. 2.12, it is apparent that cyclic hypoxia encompasses



	delay S\G2 transition	delay G2\M transition	fork collapse
A	✓	✗	✗
B	✓	✓/✗	✗
C	✓	✓	✗
D'	✓	✓/✗	✓/✗
D"	✓	✓	✓/✗
E	✓	✓	✓

Figure 2.12: Schematic for how the response of cells depends on the properties of the cyclic hypoxia protocol considered. Here, toxicity describes the extent to which cellular functions (proliferation and viability) are compromised as a result of cells being exposed to cyclic hypoxia. The table indicates which biological mechanisms are predicted to play a role for the different regions identified in the $(\mathcal{T}_+, \mathcal{T}_-)$ parameter space. The check marker indicates that the mechanism is activated, the cross that this is never activated, both check and cross markers imply that activation is not continuous but fluctuates.

multiple environmental conditions that trigger different cellular responses and variable

degrees of toxicity. Here the term toxicity refers to the extent to which physiological cell functions, such as proliferation and cell viability, are compromised as a result of cells being exposed to cyclic hypoxia. For sufficiently short periods of hypoxia, the model predicts a reduction in the population growth rates due to delays in the S→G2 transition caused by the reduced rates of DNA synthesis during hypoxia (see region A in Fig. 2.12). As cells are more frequently exposed to longer periods of hypoxia, progress along the cell-cycle is more significantly perturbed due to the accumulation of a large number of damaged cells in the G2 phase. In region C of Fig. 2.12, cells in the S phase are exposed to high levels of replication stress independently of when they initiate DNA synthesis. As a result, most cells entering the G2 phase are damaged and, as such, have to delay transition to mitosis (*i.e.*, $\mathcal{P}_2(t) \approx p_2$). In contrast, for environmental conditions in region B, the amount of replication stress cells experience in the S phase is highly dependent on when cells initiate DNA synthesis. As a result, $\mathcal{P}_2(t)$ fluctuates between 0 and its maximum value p_2, so that damaged cells accumulate in the G2 phase only periodically. Finally, if exposure to hypoxia is prolonged, fork collapse leads to cell death in S phase; cell sensitivity to fork collapse can be either continuous (E) or periodic (D′-D″) when $\mathcal{T}_+$ is sufficiently large. Recalling the results in Fig. 2.10, we find that having defective cell-cycle checkpoints is not advantageous in all environmental conditions (in the $(\mathcal{T}_+, \mathcal{T}_-)$ plane). While in region A, cells with both G1 and S/G2 checkpoints defective have a growth advantage over cells with functioning checkpoints, having defective checkpoints is deleterious when cells are grown in more toxic environments (such as regions D). Overall, our results suggest that the idea of cyclic hypoxia as a single phenomena, as commonly used in the literature, may be inaccurate; rather this term encompasses several phenomena.

In this chapter, we explored cell responses to different cyclic hypoxia protocols. However, these are only a few examples from the range of possible environmental conditions cells might be exposed to *in vivo*. Since numerical simulations do not allow a thorough exploration of the $(\mathcal{T}_-, \mathcal{T}_+)$-space, in Chapter 4 we will investigate the asymptotic dynamics predicted by our model. In addition to providing a more complete picture of the range of possible cell responses to cyclic hypoxia, this analysis will validate the predictions of the numerical simulations presented in this chapter. However, before studying more extensively the predictions of our model, in Chapter 3, we are going to validate our model against flow cytometry data on cell-cycle dynamics in hypoxia [5] and investigate why existing models, unlike ours, fail to capture experimental data.

Chapter 3

Predicting cell-cycle dynamics in *in vitro* hypoxia experiments

3.1 Overview

In Chapter 2, we introduced a mathematical model to describe the cell-cycle and growth dynamics of a population of cells exposed to different forms of hypoxia. The development of this model was necessary because existing cell-cycle model could not account for observations from recent experiments (see Section 1.1 and references therein). In this chapter, we use our modelling framework to explain cell-cycle progression in *in vitro* hypoxia experiments, including brief (*i.e.*, acute) exposure to constant and cyclic hypoxia. We also show how our model can be used to design new experiments in order to better understand the biological mechanisms that contribute to cell-cycle control in hypoxia. In doing so, we use the range of data-driven modelling techniques introduced in Section 1.3 to construct an identifiable, and minimal, model that can be calibrated to the data available.

We start by simplifying the model from Chapter 2 to describe responses to acute hypoxia using global sensitivity analysis (see Section 1.3.1). Starting from the simplified model, we then introduce a class of models ($\mathcal{S}_{\mathcal{M}}$) of different degrees of complexity by incrementally adding biological mechanisms that are believed to play a contribute to cell-cycle dysregulation in hypoxia (*e.g.*, delays in the cell-cycle transitions). In doing so, we aim to identify the minimal model that can explain experimental data on cell-cycle dysregulation in acute hypoxia [5]. In order to compare the different models, we use *Bayesian model selection*: we first calibrate our class of models against experimental data by using *Bayesian inference* methods and then identify the "best" model structure based on the *deviance information criteria*. This analysis reveals that both the checkpoint compartments are necessary to describe the experimentally

observed dynamics and suggests that memory effects are a characteristic feature of cell-cycle progression in cyclic, but not chronic, hypoxia. Our findings explain why previous studies, which aimed to describe cell-cycle responses to constant hypoxia, neglected such memory effects. We conclude in Section 3.5, by using the selected and calibrated model to predict cell responses to different oxygen environments. We then explain how these results could inform the design of new experiments to validate the model and/or improve the accuracy of the parameter estimates. This chapter is largely based on the publication [29]:
G. L. Celora, S. B. Bader, E. M. Hammond, P. K. Maini, J. Pitt-Francis, H. M. Byrne, *"A DNA-structured mathematical model of cell-cycle progression in cyclic hypoxia"*, J. Theor. Biol. (545), 2022.

3.1.1 Experimental data

This chapter is inspired by experimental data showing how the cell-cycle of RKO (colorectal cancer) cells changes when they are cultured *in vitro* as 2D monolayers and exposed to fluctuating oxygen levels [5]. As shown in Fig. 3.1, the cells are cultured in chambers where the oxygen levels $c = c(t)$ (where t is time) are carefully controlled and assumed to be spatially homogeneous. At prescribed time points, a subset of the cells is analysed using flow cytometry to estimate the fractions of cells in the G1, S and G2/M phases of the cell-cycle; we denote these fractions by F_1, F_S and F_2, respectively. We refer to the time-evolution of F_1, F_S and F_2 as the observed cell-cycle dynamics. Since each measurement requires cells to be harvested, we assume that the measurement errors are independent. As cell death is not observed in the experiments, we have that by definition $F_1 + F_S + F_2 = 1$, so only two of the three cell fractions are needed to fully characterise the cell-cycle dynamics.

The model developed in Chapter 2 decomposes cell-cycle phases into multiple compartments to account for the delays in cell-cycle transitions due to hypoxia. This is achieved by decomposing the G1 and G2/M phases of the cell-cycle are decomposed into two ($\mathcal{Z}_1 = \{G_1, C_1\}$) and four ($\mathcal{Z}_2 = \{G_2, R_2, D_2, \mathrm{Sn}_2\}$) compartments, respectively. As the experimental set-up does not allow direct measurement of all model variables, this leads to the question of whether the model is practically identifiable, based on flow cytometry data alone. We can relate model outputs to experimental data by defining the fractions $f_1(t)$, $f_S(t)$ and $f_2(t)$:

$$f_i(t) = \sum_{z \in \mathcal{Z}_i} \pi_z(t), \quad i \in \{1, S, 2\}\,, \tag{3.1}$$

where $\{\pi_z\}_{z\in\mathcal{Z}}$ (see Eq. (2.20)) describe the fraction of cells in each model compartment; the set of all model variables, $\mathcal{Z}$, is subdivided into the three subsets $\mathcal{Z}_1$ and $\mathcal{Z}_2$ (as defined above) and $\mathcal{Z}_S = \{S\}$, where the classification is based on DNA content (in line with the flow cytometry output).

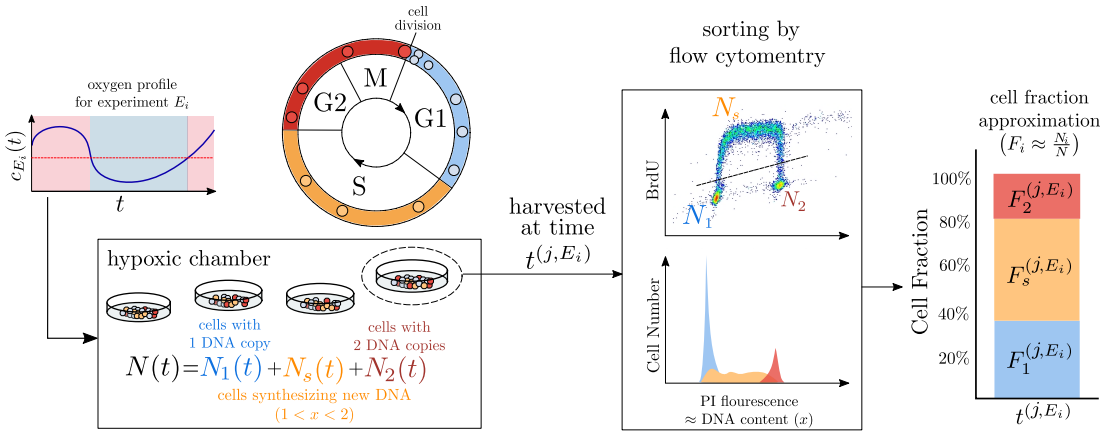


Figure 3.1: Schematic representation of flow cytometric data. Cells are cultured in a hypoxic chamber where they are exposed to a given oxygen environment. At n_{E_i} prescribed time points ($t^{(j,E_i)}$, $j = 1,\ldots,n_{E_i}$) during experiment E_i, estimates of the cell-cycle distribution are obtained by harvesting a subpopulation of cells (from the total number of cells, N, in the chamber) and using flow cytometry to sort them according to their DNA content (PI fluorescence intensity) and BrdU uptake. This enables to estimate the fractions F_1, F_s and F_2 of cells in the subpopulation with one, intermediate (i.e. in between one and two), and two copies of DNA, respectively.

As shown in Fig. 3.2, two experimental protocols are tested. At time $t = 0$, cells are exposed to either constant hypoxia ($c(t) \equiv c_- \approx 0.1\%\, O_2$, $t > 0$), or periodic cycles of hypoxia (2 hr at $c = c_-$ and 2 hr at $c_+ = 2\%\, O_2$). Prior to both experiments, the cells are cultured in normoxia ($21\%\, O_2$) so that measurements at time $t = 0$ contain information on the cell-cycle distribution in unperturbed physiological conditions. As discussed in Section 2.3.1, in an oxygen-rich environment, and in the absence of competition, cells are typically in a regime of balanced exponential growth for which the values of the cell fractions, f_m, are stationary (i.e., they do not change over time). Hence, a single set of measurements should be sufficient to fully characterise the cell-cycle dynamics under these environmental conditions. We, therefore, divide the data into three different sets: E_0 (normoxia), E_1 (constant hypoxia) and E_2 (cyclic hypoxia). We note that there is no significant change in the cell-cycle dynamics when oxygen levels decreases from 21% to 2% O_2. This is consistent with our modelling assumption that oxygen levels affect the cell-cycle dynamics only if they drop below a critical threshold $c_H = 1\%\ O_2$. The data from experiment E_0 have been

used in Section 2.3.2 to estimate model parameters (*i.e.*, k_1, k_2 and v_S^+) describing unperturbed cell-cycle dynamics in oxygen-rich environments.

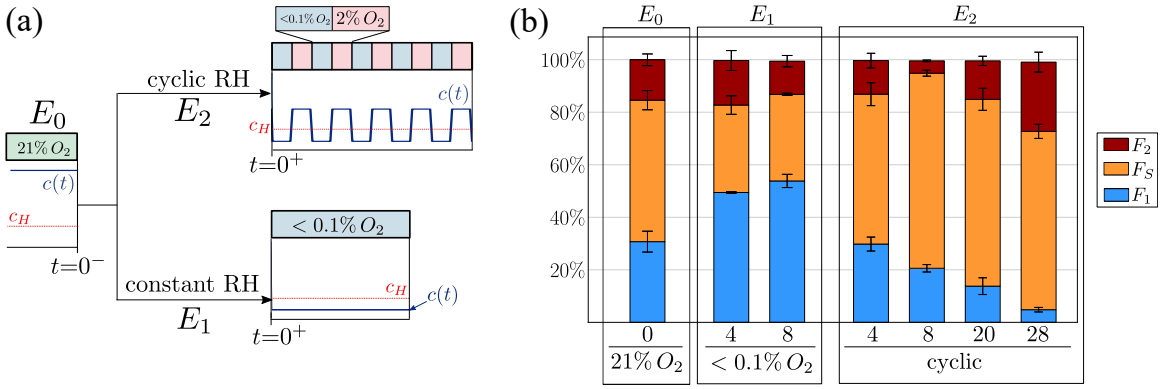


Figure 3.2: (a) Data are partitioned into three sets corresponding to the three experiments: normoxia (E_0), constant hypoxia (E_1) and cyclic hypoxia (E_2). We plot the evolution of the oxygen levels, $c(t)$ (see blue curve), and compare it to the threshold c_H at which DNA synthesis is impaired (see red horizontal line). (b) Histogram summarising the data from [5]. At each time point, the histograms illustrate the mean cell fractions estimated from multiple (between two and four) measurements. Error bars indicate one standard deviation in the estimated values. The complete data sets are available in [29].

The histograms in Fig. 3.2 summarise the data from [5], and were obtained by averaging over two to four independent replicates of each experiment. The data show that both constant (E_1) and cyclic (E_2) hypoxia redistribute the cells along the cycle. For experiment E_1, cells tend to accumulate in the G1 phase (note the increase in the fraction F_1), whereas in experiment E_2, the fraction F_1 decreases. In this case, cells initially concentrate in the S phase (note the increase in the fraction F_S) and accumulate in the G2/M phase at later times (see increase in F_2 at time $t = 28$ hr). In the remainder of this chapter, we will use inference techniques to calibrate and validate the model from Chapter 2 against the time-series data from experiments E_1 and E_2.

3.2 A model for cell-cycle dysregulation under short-term exposure to hypoxia

Given the discrepancy between the complexity of our model and the data available, we first address the question of model identifiability. The model from Chapter 2 describes both the short (or acute) and long-time effect of exposure to different forms of hypoxia.

The available data, however, only focus on the short-time response of cells. This raises the question of which mechanisms amongst those included in the full model are relevant to the short time-scale, and which can be neglected. We recall that our model contains 11 unknown parameters: $\mathbf{\Theta} = [K_1, q_1, \mathcal{R}_S^+, \mathcal{R}_S^-, p_2, T_{C2}, p_{Sn}, T_{Se}, \mathcal{R}_M^+, \mu_2^C, \mu_S]$. The other model parameters are set to the default values in Table 2.2. We use global sensitivity analysis to identify which parameters, and therefore, mechanisms, affect the model outputs when simulating experiments E_1 and E_2.

3.2.1 Numerical set up for the parameter screening

For the global sensitivity analysis we use the variance-based approach described in Section 1.3.1. This requires specification of prior distributions for each parameter. Following a standard approach, for all model parameters, except μ_S, μ_2^C and T_{Se}, we use a uniform distribution over the biologically relevant ranges indicated in Table 2.2. Given that the ranges for μ_S and μ_2^C span several orders of magnitude, we assume that their logarithm is uniformly distributed in the range $[-6, -1]$. For T_{Sn}, the physical constraint $T_{Sn} > T_{C2}$ precludes the use of a standard uniform distribution. Instead, we rewrite T_{Sn} as $T_{Sn} = T_{C2} + \Delta T_{Sn}$ and assume that ΔT_{Sn} is uniformly distributed in $[0, 30]$. We then compute the total sensitivity index by sampling the parameter space as described in Section 1.3.1, yielding $\kappa = 7168$ parameter sets $\mathbf{\Theta}^{(k)}$, $k = 1, \ldots, \kappa$.

For each parameter set $\mathbf{\Theta}^{(k)}$ we simulate Eqs. (2.18) in the interval $U_t = [0, t_f]$ using Eqs. (2.30) from Section 2.4.1 as initial conditions. We do this to capture the fact that, before the start of each experiment, cells are maintained in a regime of asynchronous exponential growth. We extract from each simulation the cell fractions $\{f_i(t)\}_{i=1,S,2}$ (defined by Eq. 3.1) which correspond to the physical quantities measured during each experiment. To replicate the environmental conditions in the chamber during the experiments from [5], we use the following functional form for the oxygen levels $c = c(t)$ at time $t \in U_t$:

$$c(t) = c_+ + (c_+ - c_-) \sum_{i=1} \left[H_\epsilon(t - t_i^{(R)}) - H_\epsilon(t - t_i^{(H)}) \right], \tag{3.2}$$

where H_ϵ is defined in Eq. (2.4) with $\epsilon = 0.01$, $t_i^{(H)}$ and $t_i^{(R)}$ are the times at which oxygen levels decrease and increase across the threshold $c = c_H$, respectively. We first simulate the model by fixing $t_i^{(R)} = 4(i-1) + 2$ and $t_i^{(H)} = 4(i-1)$, $c_- \approx 0.1\% \, O_2$ and $c_+ = 2\% \, O_2$ to reproduce the 2hr+2hr cycle corresponding to experiment E_2 in Fig. 3.2 (final time $t_f = 28$ hours). This is similar to using Eqs. (2.31) with $\mathcal{T}_+ = 2$

and $\mathcal{T} = 4$. However, for consistency with [29], we use Eq. (3.13) instead. We then simulate the model with the same initial conditions, but setting $t_1^{(H)} = 0$ and $t_1^{(R)} = \infty$, to obtain the chronic hypoxia (E_1) conditions with $c < c_H$ for $t \in U_t$ (final time $t_f = 8$ hours). Using the simulated curves and the Python library `SALib` [72], we compute the time-dependent total sensitivity index $\boldsymbol{S}_T(t) \in \mathbb{R}^3$ for each parameter with respect to each cell fraction defined in Eq. (3.1). We use the maximum of the infinity norm of the total, time-dependent, sensitivity index ($S_T^\infty = \max_{t \in U_t} \|\boldsymbol{S}_T(t)\|_\infty$) as a metric to perform parameter screening, using the threshold $S_T^\infty \leq \bar{S} = 0.05$ to identify non-influential parameters.

Based on the parameter screening procedure, we propose a new model where the non-influential parameters are set to a default value $\boldsymbol{\theta}_0$. We sample randomly 2000 points from a uniform distribution to generate $\kappa = 2000$ parameter sets $\boldsymbol{\Theta}^{(k)}$, $k = 1, \ldots, \kappa$. We then simulate the full model to produce κ solution paths for both experimental conditions, E_1 and E_2. From the latter, we extrapolate the values of the three cell fractions f_i ($i = 1, S, 2$) at prescribed times. We arrange the observations, for both cyclic and chronic hypoxia experiments, into a single vector $\boldsymbol{y}^{(k)} \in \mathbb{R}^{3 \times n_t}$, where n_t is the total number of sampled times amongst the two experimental conditions. This leads to the first sample of observations, $Y(t) = \left\{\boldsymbol{y}^{(k)}\right\}_{k=1}^{\kappa}$, generated from the original model. We then generate a new (independent) sample of parameter sets $\boldsymbol{\Theta}^{(k)}|\boldsymbol{\theta}_0$, where influential parameter are sampled from the associated uniform distribution while non-influential parameters are set to the prescribed default value $\boldsymbol{\theta}_0$. We use the sample $\boldsymbol{\Theta}^{(k)}|\boldsymbol{\theta}_0$ to generate κ realisations of the model that are conditioned to the default values of the non-influential parameters, $\boldsymbol{\theta}_0$. As before, we extrapolate the predicted values for the three cell fractions f_i ($i = 1, S, 2$) to produce a new sample $Y_{|\boldsymbol{\theta}_0}(t) = \left\{\boldsymbol{y}_{|\boldsymbol{\theta}_0}^{(k)}\right\}_{k=1}^{\kappa}$, where $\boldsymbol{y}_{|\boldsymbol{\theta}_0}^{(k)} \in \mathbb{R}^{3n_t}$. We then use the mean embeddings test (see Section 1.3.1.2) to decide whether there is statistical evidence that Y and $Y_{|\boldsymbol{\theta}_0}$ are generated by different distributions. We use a significance level of $\alpha = 0.001$ and a total of $J = 5$ evaluation points, so that we reject the null-hypothesis (*i.e.*, Y and $Y_{|\boldsymbol{\theta}_0}$ are generated from the same distribution), if the test statistics are above the 0.999-quantile of the $\chi^2(5)$ distribution, where χ^2 is the chi-squared distribution. Here the choice of a small value of α, as suggested by [124], is guided by the main goal of the test, which is to validate the parameter screening. Knowing that the two samples come from different models, we expect deviation, yet we want such deviations to be sufficiently small. To compute the mean embeddings test statistics, we rely on the available code from [74].

3.2.2 Results of the model reduction

The results of the parameter screening procedure are illustrated in Fig. 3.3. We identify as influential a subset of the model parameters; namely $[K_1, q_1, \mathcal{R}_S^+, \mathcal{R}_S^-, p_2, T_{C2}]$. The remaining parameters, instead, have sensitivity indices well below the threshold $\bar{S} = 0.05$.

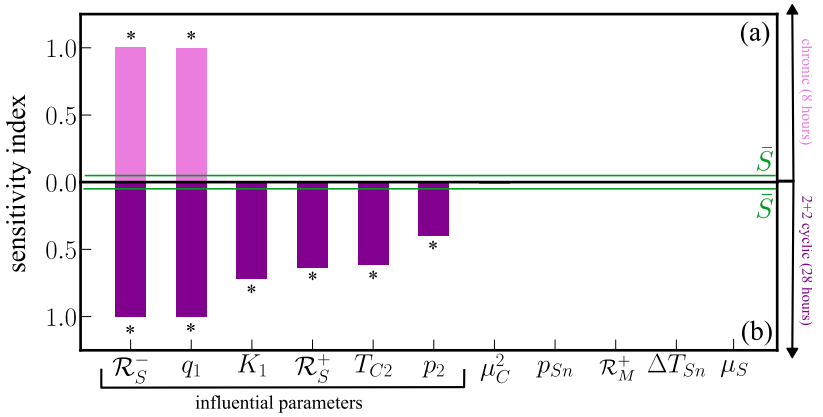


Figure 3.3: (a) Values of S_T^∞ obtained from simulating the constant hypoxia experiment (E_1). (b) Values of S_T^∞ obtained from simulating the cyclic hypoxia experiment, E_2. Details on the numerical procedure used to estimate S_T^∞ are given in Section 3.2.1. For each experimental condition, influential parameters (*i.e.*, parameters for which $S_T^\infty > \bar{S} = 0.05$) are indicated with a (*). Overall we identify parameters as influential if they are influential in at least one of the two experimental set-ups considered. Details on the numerics for estimating Sobol indices are given in Section 3.2.1.

The validity of the parameter screening is confirmed by adopting the mean embeddings two-sample testing procedure [74] described in Section 1.3.1.2. In doing so, we set non-influential parameters to the following arbitrary default values:

$$\mu_S = 0.025\,\mathrm{hr}^{-1},\ \mathcal{R}_M^+ = 0.4\,\mathrm{hr}^{-1},\ \Delta T_{Sn} = 0\,\mathrm{hr},\ p_{Sn} = 0.1,\ \mu_2^C = 0.0\,\mathrm{hr}^{-1}. \tag{3.3}$$

We find that the two-sample test validates the parameter screening. The null hypothesis, *i.e.*, that the conditional and unconditional models are equal (see Section 1.3.1.2 for more details) can not be rejected (p-value=0.76 $\gg$ 0.001), suggesting that the difference between the conditional and unconditional models is not statistically significant. We conclude that the behaviour of the full model can be captured by setting non-influential parameter to the values listed in Eq. (3.3), at least for the time-frame and conditions tested experimentally. We acknowledge that the choice of default values of the parameters given in Eq. (3.3) is arbitrary. However, the sensitivity analysis indicates that the specific values of parameters chosen should not influence predictions on the short term response of cells to cyclic hypoxia. This should not, therefore,

significantly influence the analysis we perform in this chapter, where we focus on cell responses to short periods of exposure to constant and cyclic hypoxia.

When setting the non-influential parameters to the values in Eq. (3.3), we find that entire mechanisms captured by the cell-cycle model presented in Chapter 2 can be neglected. Given the value of $\mathcal{R}_M^+$ in Eq. (3.3) and taking $\mathcal{R}_M^-$ and M_- from Table 2.1, we can solve for the variable $m(t)$ (see Eq. (2.1)) and the delay $\tau_R(t)$ (see Eq. (2.18k)). We find the expression levels of repair proteins, $m(t) > M_{FC}$ for all times both when simulating experiment E_1 and E_2. Consequently, $u_S(m) \approx 0$ (see Eq. (2.3c)) regardless of μ_S being positive. This is consistent with *in vitro* experiments during which no cell-death is observed. Furthermore, we find that $\tau_R(t) < t$ at all times when simulating experiments E_1 and E_2. This implies that, on the time scale of the experiments, there are no cells leaving the D_2 compartment (see Eq. 2.18h). We conclude that while the experimental data available may contain information on the number of cells that enter either the D_2 or Sn_2 compartments (see Fig. 3.4a), the data are not sufficient to distinguish between transiently (D_2) and permanently (Sn_2) arrested cells since neither of them re-enter the cell-cycle before the end of the experiment. For this reason, we propose the reduced model in Fig. 3.4b where we combine Sn_2 and D_2 cells into a single well-mixed compartment, namely C_2. While cells enter the C_2 compartment with probability $\mathcal{P}_2$ after completing the S phase (which depends on the values of parameters p_2 and T_{C2}; see Eq. (2.11b)), they can not exit it. Furthermore, we neglect cell death in C_2 in line with the observation that no cell death is observed during the *in vitro* experiments. We denote the reduced model $\mathcal{M}$. The governing equations for $\mathcal{M}$ are obtained from Eqs. (2.18]) setting $u_S(t) \equiv 0$, $\mu_2^C = 0$ and $\mathcal{P}_{Sn} \equiv 1$ and switching the name of the Sn_2 compartment to C_2. As a result we obtain a system of 5 delay-differential equations with a single time-dependent delay. This contrasts with our original model (see Fig. 3.4a) of 7 delay-differential equations with two time-dependent delays. Furthermore, while the original model has 11 unknown parameters, the reduced model contains only those 6 parameters that have been identified as influential from the parameter screening analysis (*i.e.*, $\mathbf{\Theta} = [K_1, q_1, \mathcal{R}_S^+, \mathcal{R}_S^-, p_2, T_{C2}]$). When using the reduced model to make predictions on experimental outcomes, it is important to recall its limitations. To this end, in Section 3.5 focus on using the reduced model to make predictions in the realms of its validity. We will consider oxygen protocols where fork collapse is also negligible and focus on analysing model predictions that are independent of the fate decision of cells arresting in the G2 compartment.

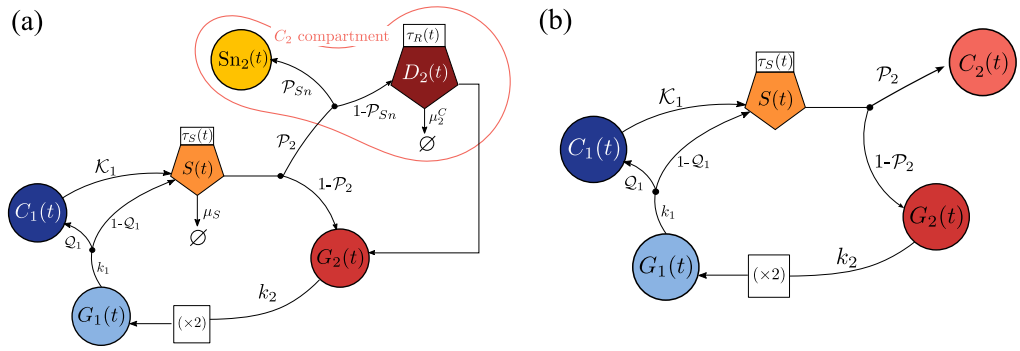


Figure 3.4: Comparison between the model proposed in Chapter 2 (a) and reduced model obtained based on the result of the sensitivity analysis (Fig 3.4b). We see that the main difference is in the description of the G2 checkpoint: while the original model contained two compartments (D_2 and Sn_2) the reduced model combines these three compartment into a unique well-mixed compartment named C_2. Furthermore, three edges have been deleted: the one associated with death due to fork collapse in the S phase ($u_S(t) \approx 0$), the one related to death in the D_2 compartment due to damaged cells attempting mitosis ($\mu_2^C D_2 \approx 0$) and the one accounting for cell transitioning from the D_2 to the G_2 compartment.

3.3 Fitting model $\mathcal{M}$ to flow cytometry data

In this section, we discuss the calibration of the proposed model, $\mathcal{M}$, to the data available. The model $\mathcal{M}$ is simulated by solving Eqs. (2.18) (setting $\mathcal{P}_{Sn}(\tau_S) \equiv 1$, $u_S(m) \equiv 0$ and $\mu_2^C \equiv 0$ and renaming the Sn_2 compartment C_2) with initial conditions (2.30). Since we have no information of the total cell numbers, we assume without loss of generality that the initial population size is 1. As explained in [29], this does not affect the model calibration since we are only interested in fitting our linear model to data on cell fractions. For the fitting, we adopt the Bayesian inference approach outlined in Section 1.3.2. This requires us to specify prior distributions for the model parameters and the likelihood function.

3.3.1 The observation model and the inference procedure

Given the small amount of data available, we calibrate the model by pooling the data from the experiments in constant hypoxia (E_1) and cyclic hypoxia (E_2) into a single dataset $\mathcal{E}$. We therefore postpone model validation until more data will be available. In Fig. 3.2, for each time point, we reported the mean and standard deviation for multiple repeats of the experiment; however, for the estimation of unknown parameters we consider each individual experimental measurement (the complete data set can be found in [29]). We denote by $F_z^{(i,j)}$ ($z \in \{1, S, 2\}$) the i-th measurement performed

at time $t^{(i,j)}$ during experiment E_j, with $j \in \{1,2\}$. Given that F_1, F_s and F_2 are not independent (recall $F_S = 1 - F_1 - F_2$), only two of the three values are needed for the fitting; without loss of generality we choose F_1 and F_2. We collect the data in the set $\mathcal{E} = \left\{(t^{(i,1)}, F_1^{(i,1)}, F_2^{(i,1)})\right\}_{i=1}^{L_1} \cup \left\{(t^{(i,2)}, F_1^{(i,2)}, F_2^{(i,2)})\right\}_{i=1}^{L_2}$, where $L_1 = 4$ and $L_2 = 14$ are the number of measurements in sets E_1 and E_2, respectively.

We denote by $f_z(t; \boldsymbol{\Theta}, \mathcal{M}, E_j)$ (with $z \in \{1,2\}$) the values of the cell fractions at time t for the experiment E_j predicted by the reduced model, $\mathcal{M}$ for parameter values $\boldsymbol{\Theta}$. We here consider the, experimental observations $F_{1,2}^{(i,j)}$ to be noisy realisations of the model solutions, $f_{1,2}(t^{(i,j)}; \boldsymbol{\Theta}, \mathcal{M}, E_j)$. We assume that the observation errors are independent, additive and normally distributed with zero mean and variance $\sigma_{1,2}^2$. Furthermore, we consider $\sigma_{1,2}$ to be constant in time and to be the same for both constant (E_1) and cyclic (E_2) experiments. Instead of specifying values of the variances σ_1 and σ_2, we view them as unknown parameters that are learnt from the data. Since the cell fractions are normalised, we assume that σ_1 and σ_2 are uniformly distributed in $[0,1]$ (*i.e.*, $\pi_{\mathcal{M}}(\boldsymbol{\sigma}) \sim \mathcal{U}(0,1) \times \mathcal{U}(0,1)$). Based on the assumptions outlined above, $\mathcal{L}_{\mathcal{M}}$ is given by:

$$\mathcal{L}_{\mathcal{M}}(\mathcal{E}|\boldsymbol{\Theta}, \boldsymbol{\sigma}) = \prod_{z=1}^{2} \prod_{j=1}^{2} \prod_{i=1}^{L_j} \phi\left(F_z^{(i,j)}; f_z(t^{(i,j)}; \boldsymbol{\Theta}, \mathcal{M}, E_j), \sigma_z^2\right), \tag{3.4}$$

where $\boldsymbol{\sigma} = [\sigma_1, \sigma_2]$, and ϕ is the normal probability density with mean $f_z(t; \boldsymbol{\Theta}, \mathcal{M}, E_j)$ and variance σ_z^2.

In the absence of prior information on the values of the model parameters $\boldsymbol{\Theta} = [K_1, q_1, \mathcal{R}_S^+, \mathcal{R}_S^-, p_2, T_{C2}]$, we assume that each component of the vector $\boldsymbol{\Theta}$ (i.e., Θ_i for $i = 1, \ldots, \kappa$) is uniformly distributed on the intervals $[\Theta_i^{min}, \Theta_i^{max}]$, where the extremes of the intervals are taken from Table 2.2; hence $\pi_{\mathcal{M}}(\boldsymbol{\Theta})$ is given by:

$$\pi_{\mathcal{M}}(\boldsymbol{\Theta}) = \prod_{i=1}^{\kappa} \left(\Theta_i^{max} - \Theta_i^{min}\right)^{-1} I_{[\Theta_i^{min}, \Theta_i^{max}]}(\Theta_i), \tag{3.5}$$

where $I_A(\cdot)$ is the indicator function for the general set A. Using Bayes' Theorem, we can update our prior distributions ($\pi_{\mathcal{M}}, (\boldsymbol{\Theta})$, $\pi(\boldsymbol{\sigma})$) in light of the available experimental data:

$$\pi_{\mathcal{M}}(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E}) = \frac{\pi_{\mathcal{M}}(\boldsymbol{\Theta})\pi(\boldsymbol{\sigma})\mathcal{L}(\mathcal{E}|\boldsymbol{\Theta}, \boldsymbol{\sigma})}{p(\mathcal{E})}. \tag{3.6}$$

As discussed in Section 1.3.2, the size of the parameter space does not allow us to estimate exactly the posterior distribution. Instead, we estimate $\pi_{\mathcal{M}}(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E})$ using Markov Chain Monte Carlo Methods implemented the Python package PINTS (Probabilistic Inference on Noisy Time-Series) [38]. Specifically, as suggested by Johnstone

et al. [75], prior to starting our MCMC routine, we compute a good initial guess by maximizing the likelihood function $\mathcal{L}_{\mathcal{M}}$ (see Eq. (1.14)). Based on Eq. (1.15) and the choice of a uniform prior, maximising the posterior is equivalent to maximising the likelihood function. Sampling random initial guesses from π_{pr}, we solve the optimisation problem for the log-likelihood using the `CMA-ES` algorithm [38, 64]. We then use the output of the optimization routine to initiate the MCMC simulation (we compute, in total, three chains). We sample from the posterior distribution by using `HaarioBardenet MCMC`, which is a Metropolis-Hastings algorithm with adaptive covariance, where the first 8000 iterations are performed without adaptation (as suggested by [75]). We compute up to 30000 iterations for each chain and discard the first 10000 as "warm-up". As in [39], we assess the convergence of the MCMC chains by estimating $\hat{R}$ [85, Chapter 13], and we accept the sampled posterior if $\hat{R} < 1.05$ (note that $\hat{R} \to 1$ as the algorithm converges).

3.3.1.1 Posterior predictions for the cell-cycle dynamics

We assess the predictions of the calibrated model on the cell-cycle distribution $\boldsymbol{F} = [F_1, F_2]$ via the predictive posterior distribution, $\pi_{\mathcal{M}}(\boldsymbol{F}|\mathcal{E})$, which we first introduced in Section 1.3.2 to discuss model selection and, we recall, is defined as:

$$\pi_{\mathcal{M}}(\boldsymbol{F}|\mathcal{E}) = \int_{\mathbb{R}^{n_\Theta + n_\sigma}} \mathcal{L}_{\mathcal{M}}(\boldsymbol{F}|\boldsymbol{\Theta}, \boldsymbol{\sigma}) \pi_{\mathcal{M}}(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E}) d\boldsymbol{\Theta} d\boldsymbol{\sigma}, \tag{3.7}$$

In general, it is not be possible to estimate the integral in Eq. (3.7). We instead use sampling techniques using the chains generated via MCMC. In more detail, we sample κ parameter sets, $\left\{(\boldsymbol{\Theta}^{(i)}, \boldsymbol{\sigma}^{(i)})\right\}_{i=1}^{\kappa}$, from the estimated posterior $\pi_{\mathcal{M}}(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E})$ (*i.e.*, randomly choosing from the "converged" Markov chains generated in the calibration step as described above). Predictive posterior estimates for the outputs of the model are obtained by sampling 800 parameter sets, $(\boldsymbol{\Theta}^{(i)}, \boldsymbol{\sigma}^{(i)})$, from $\pi(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E})$. For each set, we run the model forwards to generate 800 predictive curves for each phase fraction $f_{1,2}^{(i)}(t)$, with $i = 1, \ldots, 800$. This gives posterior distributions for the model prediction of cell fractions. To obtain the posterior prediction for the *measured* cell fractions $F_{1,2}^{(i)}(t)$, we add to the simulated cell fractions, $f_1(t; \boldsymbol{\Theta}^{(i)}, \mathcal{M}, E_j)$ and $f_2(t; \boldsymbol{\Theta}^{(i)}, \mathcal{M}, E_j)$, the corresponding measurement errors $\boldsymbol{e}(t) \sim \mathcal{N}(0, \boldsymbol{\sigma}^{(i)})$. We then estimate $F_S^{(i)}$ as $F_S^{(i)} = 1 - F_1^{(i)} - F_2^{(i)}$ and $f_S^{(i)}$ as $f_S^{(i)} = 1 - f_1^{(i)} - f_2^{(i)}$. At each time point, we compute the mean of the 800 predictive curves (for either the "true" or measured fractions) and the corresponding 68%- and 95%-confidence intervals. To extrapolate point estimates of the model parameters, we use the mean of the posterior

distribution. Additional results on the estimated posterior distributions can be found in Appendix A.2.

3.3.2 Results of the fitting

Looking at the posterior predictions in Fig. 3.5, we see that the reduced model $\mathcal{M}$ captures the experimental data experiments E_1 and E_2, with all experimental data points (see dots in the plots) falling in the 95%-confidence interval of the posterior predictions for the measured fractions, F_1, F_S and F_2.

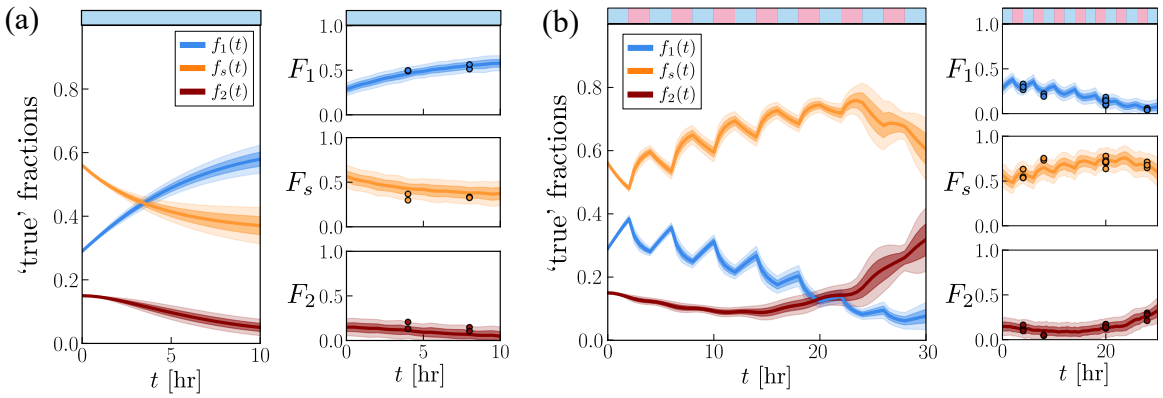


Figure 3.5: Posterior prediction distribution of the selected model $\mathcal{M}$ for (a) constant and (b) cyclic hypoxia. For both scenarios we plot the posterior estimates for model outputs (or 'true' fractions), $f_i(t)$, and the predicted measured fractions $F_i(t)$ with $i \in \{1, s, 2\}$; the predicted evolution of F_m is compared to the experimental data (dots). For each model output we plot the expected value, while the dark and light shaded areas indicate the 68%- and 95%- confidence intervals, respectively. The top panels show the evolution of the oxygen tension in the two experiments (light blue – hypoxia; light pink – physiological oxygen levels).

3.3.2.1 Parameter Identifiability

We end our discussion on model calibration by briefly considering practical identifiability of the unknown parameters. Following [43, 66], we define a parameter as practically identifiable if we can constrain its value to a reasonably small region of the parameter space, *i.e.*, the posterior $\pi(\Theta_i|\mathcal{E})$ has a single well-defined peak.

Looking at the profile of the marginal posterior distributions for the different parameters (see last Fig. 3.6), we see that most marginal posteriors have a bell-like shape, with a unique, well-defined, maximum for most parameters. However, this is not the case for parameter K_1 (see first panel in Fig. 3.6); its posterior distribution tends to flatten for large K_1 ($K_1 > 1$). This indicates greater uncertainty in the

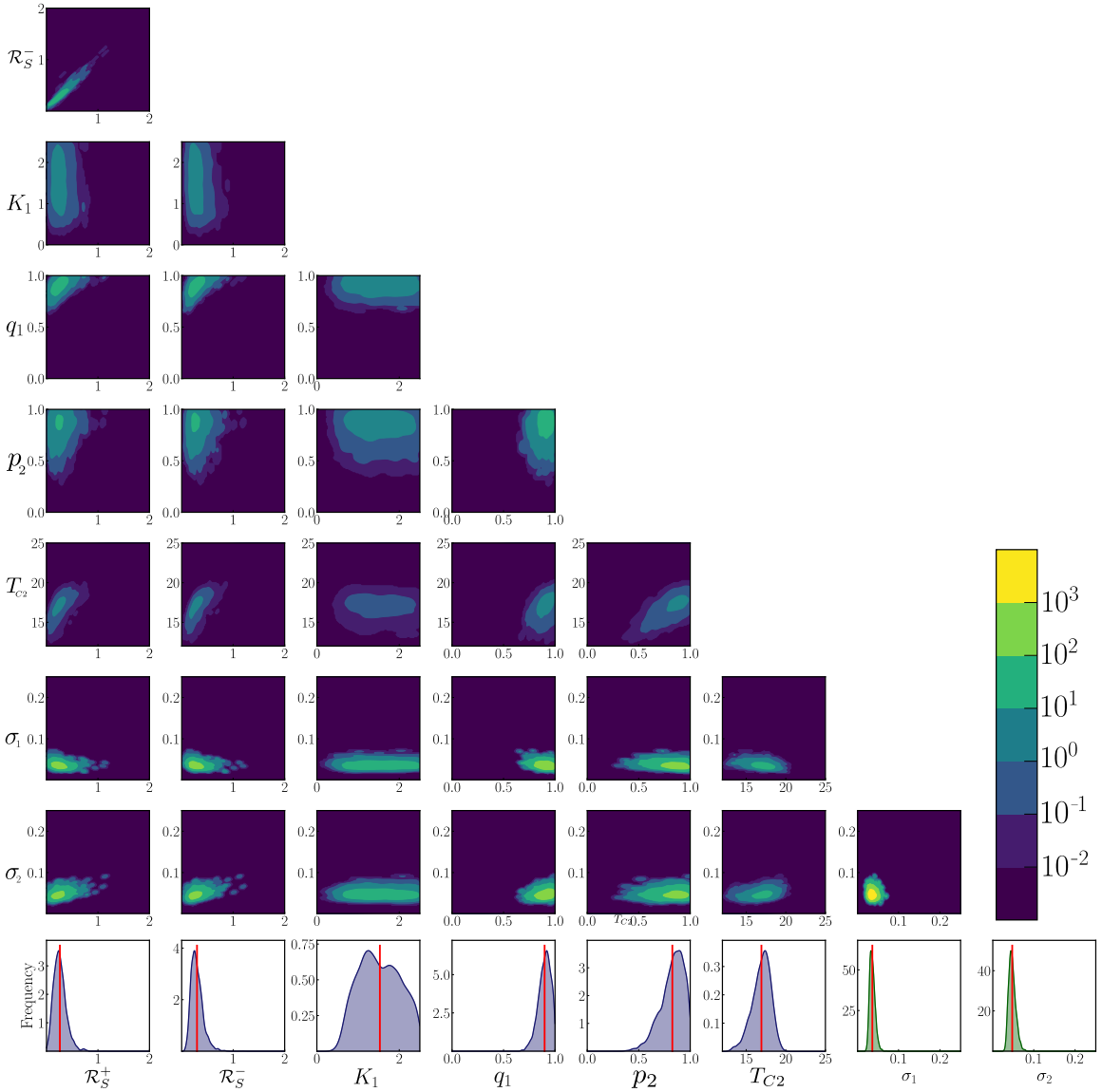


Figure 3.6: Estimated joint (surface plots) and marginal (last row) posterior distributions for the model parameters $\Theta \in \left\{\mathcal{R}_S^+, \mathcal{R}_S^-, K_1, q_1, p_2, T_{C2}\right\}$ and the error parameters σ_i for $i = 1, 2$. The distributions were obtained from the MCMC samples generated as discussed in the text above. In the surface plots for the joint distributions, the yellow areas correspond to high posterior probability and the blue areas correspond to low probability (see colour-bar on the right). We see that for most parameters the posterior distributions have a well-defined peak. This is not the case for K_1 (see third panel in the last row) where the much larger support and the non-defined peak of the posterior is a sign of practical non-identifiability. In the last row, the red vertical lines indicate the mean of the marginal posterior distribution as reported in Table A.2, where additional summary statistics extrapolated from the marginal distribution are also given. Note that in the last row the y-axes are scaled differently.

estimation of this parameter for which we can identify only a lower, but not an upper, bound. Furthermore, joint posterior distributions show that there is no correlation between K_1 and other model parameters. We conclude that the non-identifiability of K_1 does not stem from it being correlated with other model parameters and, therefore, we look for alternative explanations. From Eq. (2.16a), we note that K_1 is only relevant in data from experiment E_2 (i.e., when oxygen levels c are raised above the threshold c_H). If we consider, for example, the first time at which this happens (*i.e.*, time $t_1^{(R)} = 2hr$), then, over the period $t_1^{(R)} = 2hr < t < t_2^{(H)} = 4hr$, $\mathcal{Q}_1(t) \approx 0$ and the evolution of $C_1(t)$ can be determined by solving Eq. (2.18b) to obtain:

$$C_1(t) = C_1(t_1^{(R)})e^{-K_1(t-t_1^{(R)})}, \quad t \in (t_1^{(R)}, t_2^{(H)}). \tag{3.8}$$

Therefore, when the first measurement is taken, $C_1(t_2^{(H)}) = C_1(t_1^{(R)})e^{-2K_1}$. If $K_1 > 1$ then $e^{-2K_1} \ll 1$ and compartment C_1 rapidly empties after re-oxygenation. Consequently, unless C_1 is very large, $C_1(t_2^{(H)}) \approx 0$, regardless of the value of $K_1 > 1$. In order to resolve this fast time-scale we would need to collect experimental data at an earlier time point, say t, for which $t - t_1^{(R)} < K_1^{-1}$; alternatively, we could choose an oxygen cycle for which a larger number of cells accumulate in the C_1 compartment. This could be achieved by prolonging the period for which the cells are exposed to severe hypoxia (i.e., $c < c_H$).

We use *in silico* data to support our claim. In particular, we fix the value of the parameters and simulate the model considering a (4.5,7.5)-cyclic hypoxia protocol, where cells are exposed to hypoxia for 4.5 hours and then to physiological oxygen levels for 7.5 hours before being exposed to hypoxia again; this is simulated *in silico* by setting $t_i^{(R)} = 12(i-1) + 4.5$ and $t_i^{(H)} = 12(i-1)$ in Eqs. (3.2). We record the cell fractions f_i $(i = 1, 2)$ at times $t = [4, 9, 12, 20, 28]$. For each time point, we sample the random variable $F_{1,2}$ three times by adding Gaussian noise with mean zero and variance $\hat{\boldsymbol{\sigma}}$ to $f_{1,2}$. This procedure leads to the synthetic dataset $\tilde{\mathcal{E}}$, where we do not measure the system immediately after re-oxygenation, but one hour afterwards, when, based on the above discussion (see text after Eq. (3.8)), we expect the value of K_1 will be most relevant. We then repeat the MCMC procedure outlined in Section 3.3.1 to estimate the posterior $\pi_{\mathcal{M}}(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E}, \tilde{\mathcal{E}})$, which combines information on both the experimental and *in silico* data. To do so, we replace the uninformative uniform prior with the posterior probability distribution obtained (illustrated in Fig. 3.6). In particular, we adopt the *sequential learning* approach used in [24]. This approach is based on the application of Bayes' rule under the assumption that the two data sets

$\mathcal{E}$ and $\tilde{\mathcal{E}}$ are independent:

$$\pi_{\mathcal{M}}(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E}, \tilde{\mathcal{E}}) = \frac{\mathcal{L}_{\mathcal{M}}(\tilde{\mathcal{E}}, \mathcal{E}|\boldsymbol{\Theta}, \boldsymbol{\sigma})\pi_{\mathcal{M}}(\Theta)\pi(\boldsymbol{\sigma})}{\pi(\tilde{\mathcal{E}}, \mathcal{E})} = \underbrace{\pi_{\mathcal{M}}(\boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathcal{E})}_{\text{updated prior}} \frac{\mathcal{L}_{\mathcal{M}}(\tilde{\mathcal{E}}|\boldsymbol{\Theta}, \boldsymbol{\sigma})}{\pi(\tilde{\mathcal{E}})}. \quad (3.9)$$

In writing Eq. (3.9) we have used the assumption of independence enables us to write $\pi(\tilde{\mathcal{E}}, \mathcal{E}) = \pi(\mathcal{E})\pi(\tilde{\mathcal{E}})$ and $\mathcal{L}_{\mathcal{M}}(\tilde{\mathcal{E}}, \mathcal{E}|\boldsymbol{\Theta}, \boldsymbol{\sigma}) = \mathcal{L}_{\mathcal{M}}(\tilde{\mathcal{E}}|\boldsymbol{\Theta}, \boldsymbol{\sigma})\mathcal{L}_{\mathcal{M}}(\mathcal{E}|\boldsymbol{\Theta}, \boldsymbol{\sigma})$. Importantly, in writing Eq. (3.9), we are implicitly assuming that the model parameters $\boldsymbol{\Theta}$, and the noise parameters $\boldsymbol{\sigma}$, have the same distribution under all experimental conditions considered. This is indeed the case for our set up where the noise distribution for the *in silico* data is known. To compute the updated prior from the estimated posterior we use the kernel density estimation method implemented in the `scipy.stats` library in python.

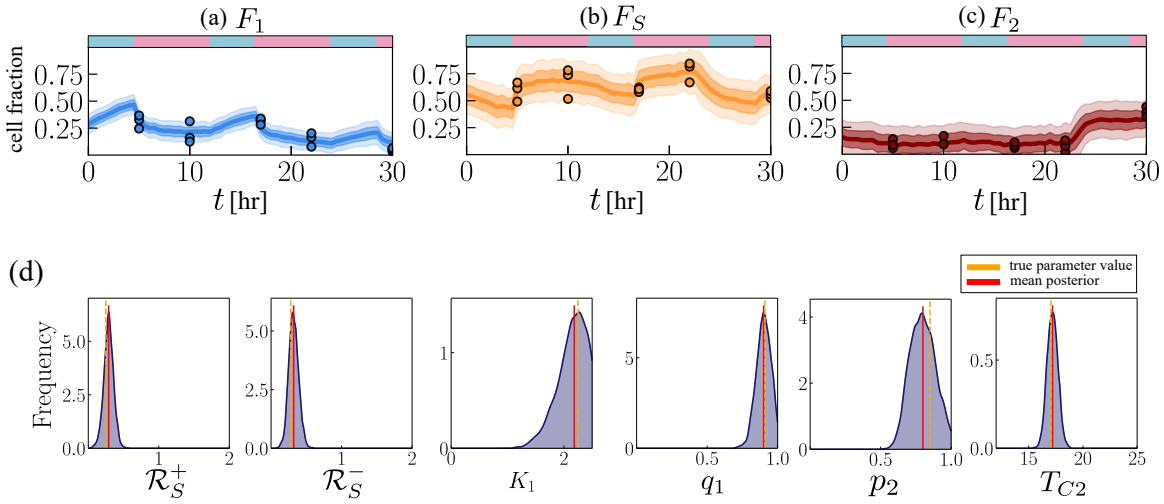


Figure 3.7: Testing identifiability of K_1 using *in silico* data. The data (dots in panels (a)-(c)) are generated by adding noise (with variance $\boldsymbol{\sigma} = [0.04, 0.05]$) to simulations of the proposed model with parameter values $\boldsymbol{\Theta} = [2.25, 0.25, 0.28, 0.85, 0.91, 17.05]$. Panels (a)-(c) compare the estimated predictive posterior distribution for the cell fractions F_i ($i \in \{1, S, 2\}$) and the *in silico* data. Panel (d) illustrates marginal posterior distributions for the model parameters. The vertical red lines indicate the posterior mean for each parameter and compares it to the true value (orange dashed line).

The results of fitting our model to $\mathcal{E}$ and $\tilde{\mathcal{E}}$ are illustrated in Fig. 3.7. As shown in Fig. 3.7a-3.7c, the posterior predictive distribution recapitulates the data for the time-evolution of all cell fractions. We can also retrieve the correct values for the model parameters. Plotting the marginal posterior distribution in Fig. 3.7d, shows that the true parameter values lie in the support of the estimated distributions and

they are well-approximated by the mean of the posterior. Furthermore, in contrast to Fig. 3.6, the posterior distribution for K_1 has the standard bell-like signature of practically identifiable parameters. The results from the *in silico* data therefore support our claim that the failure to identify K_1 from the experimental data might be due to suboptimal experimental design rather than a structural weakness of our model.

3.3.3 The role of the observation model

As mentioned in Section 1.3.2, Bayesian inference is concerned with fitting a stochastic model to data. Besides our modelling choices on the deterministic dynamics underlying the true system dynamics, there is therefore the question of how the choice of the likelihood function might influence the inference procedure. Given that the stochastic process underlying the experimental data is not known, we want to test the robustness of our conclusions and parameter estimates to changes in the likelihood function. In the previous sections, we have taken possibly the simplest choice of an additive Gaussian noise. However, this presents some drawbacks. The values of the experimental data (by definition of cell fractions) lies in the range $[0, 1]$. This condition is satisfied for the posterior predictive distribution for the model outputs, but not by our predictions on the observed cell fractions as the latter have support that extends to the whole real line. In practice, the estimated values for the standard deviations (see Fig. 3.6) are such that the probability of observing a value outside $[0, 1]$ is small, but not exactly zero.

For completeness, we test alternative observational models that are guaranteed to satisfy the physical constraints on the model variables. We first apply the following transformation:

$$\mathbf{\Psi}(F_1, F_2) = \begin{bmatrix} F_1 + F_2 \\ \frac{F_1}{F_1+F_2} \end{bmatrix}. \tag{3.10}$$

In Eq. (3.10), the entries of the vector Ψ can be interpreted as probabilities: $\Psi_1 = F_1 + F_2 \in [0, 1]$ indicates the probability that a cell is not in the S phase, while $\Psi_2 = F_1/(F_1 + F_2) \in [0, 1]$ is the probability that a cell is in the G1 phase conditioned on the knowledge that this is not in the S phase. We then assume that $\mathbf{\Psi}$ is a vector of independent random variables following a *Beta distribution*, $\Psi_i \sim \mathrm{Be}(\alpha_i, \beta_i)$. Here $\alpha_i > 0$ and $\beta_i > 0$ are the shape parameters. The Beta distribution is often used to model probabilities; for example, statistical models based on the Beta distribution are popular in the realm of population genetics [6]. Following [6], we rewrite the shape parameters α and β as $\alpha = \mu\kappa$ and $\beta = (1 - \mu)\kappa$, so that μ is the mean value of the

distribution. Then, given μ and κ, the probability distribution of a random variable $\psi \sim Be(\alpha(\mu,\kappa),\beta(\mu,\kappa))$ reads:

$$\phi_{Be}(\psi;\mu_i,\kappa_i) = \frac{\psi^{\alpha(\mu_i,\kappa_i)-1}(1-\psi)^{\beta(\mu_i,\kappa_i)-1}}{B(\alpha(\mu_i,\kappa_i),\beta(\mu_i,\kappa_i))}, \quad \psi \in [0,1], \tag{3.11}$$

where $B(a,b)$ is a normalization constant that can be expressed in terms of a product of Gamma functions. We therefore propose an alternative likelihood function:

$$\mathcal{L}^{(B)}_{\mathcal{M}}(\mathcal{E}|\boldsymbol{\Theta},\boldsymbol{\kappa}) = \prod_{j=1}^{2}\prod_{i=1}^{L_j}\left(\prod_{m=1}^{2}\phi_{Be}\left(\Psi_m(\boldsymbol{F}^{(i,j)});\Psi_m(\boldsymbol{f}^{(i,j)}_{\boldsymbol{\Theta}}),\kappa_m\right)\right)\left|\det J_{\boldsymbol{\Psi}}(\boldsymbol{F}^{(i,j)})\right|, \tag{3.12}$$

where $\boldsymbol{f}^{(i,j)}_{\boldsymbol{\Theta}} = \boldsymbol{f}(t^{(i,j)};\boldsymbol{\Theta},\mathcal{M},E_j)$. Note that Eq. (3.12) is similar to Eq. (3.4) except that the Gaussian kernel is now replaced by the Beta kernel which is applied to the transformed output $\boldsymbol{\Psi}$. Furthermore, we have introduced a multiplicative term corresponding to the determinant of the Jacobian of the transformation $\boldsymbol{\Psi}$. We refer to Section 1.3.2 to understand why this term is included. We repeat the fitting procedure by using the definition of likelihood in Eq. (3.12) and find similar estimates for model parameters with similar shapes for the marginal distribution (see results in Appendix A.1). We also use Bayesian model selection to compare the two observational models (see Section 1.3.3) and find that the two models have similar performance. We conclude that our estimates are robust to changes in the error model and, for the regime considered (where probabilities do not approach extremum values), the Gaussian error model applies.

3.4 Testing the validity of our modelling assumptions: a class of cell-cycle models

We started by deriving a reduced version of the model presented in Chapter 2 that we claim explains the available experimental data. In Section 3.3.1 we showed that the proposed model can indeed be fitted to the data and all parameters are, or can be, identified. However, the question remains of whether all mechanisms included in the model are needed to explain the experimental data. Furthermore, a distinctive feature of our model is that there is some correlation between the durations of the S and G2 phases; this introduces a memory effect which is not usually included in models of the cell-cycle in hypoxic environments. At this point, our modelling choices are grounded in what is known from the biological literature, and the qualitative experimental

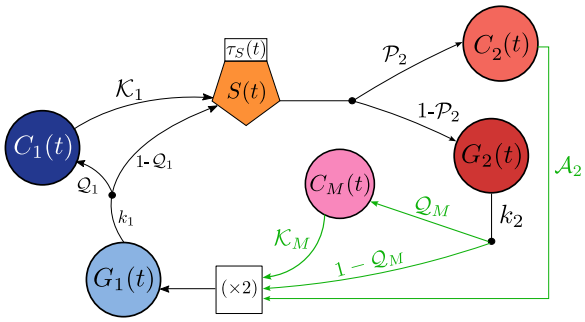


Figure 3.8: Schematic of the enlarged model used for the data-driven hypothesis testing. Here all compartments are as in Fig. 3.4b except for the introduction of the novel checkpoint compartment C_M. In green we have highlighted the arrows and coefficients in the diagram that are not present in the proposed model (Fig. 3.4b).

observations of our collaborators. Here, we seek a more rigorous and quantitative approach to justify our assumptions using data-driven modelling techniques.

As discussed in Section 1.2, it is, instead, more common to assume only an instantaneous activation of cell-cycle arrest in response to low oxygen levels. We therefore explore in more detail whether this instantaneous-response hypothesis could also describe the experimental data to the same accuracy as our proposed model, $\mathcal{M}$ (see schematic in Fig. 3.4b), for the short term effect of hypoxia on the cell-cycle dynamics. To this end, we augment the model $\mathcal{M}$ to address the following two questions: 1) Could the accumulation of cells in the G2/M phase be explained by hypoxia inhibiting the M→ G1 transition? 2) Is the time scale on which cells in the G2 checkpoint recover really much slower than the duration of the experiments?

To address these questions, we construct a new model, that accounts for all biological mechanisms included in our model $\mathcal{M}$ (see schematic in Fig. 3.4b) and two additional mechanisms that allow us to answer the questions raised above. As shown in Fig. 3.8, this includes a novel checkpoint compartment C_M that captures the possibility of hypoxia-mediated inhibition of the M→G1 transition (*i.e.*, addresses question 1 above). We model C_M analogously to the C_1 compartment, with rates of transition into and out of the compartment denoted by $\mathcal{Q}_M = q_M \mathcal{H}_{\epsilon_c}(c_H - c(t))$ and $\mathcal{K}_M = K_M \mathcal{H}_{\epsilon_c}(c(t) - c_H)$, respectively. To address question 2, we allow cells to exit the C_2 compartment at a rate $\mathcal{A}_2(t) = \alpha_2 \mathcal{H}_{\epsilon_c}(c(t) - c_H)$. The model equations corresponding to the schematic in Fig. 3.8 can be derived by applying the principle of mass balance to obtain:

$$\frac{dG_1}{dt} = 2(k_2(1 - \mathcal{Q}_M(t))G_2 + \mathcal{A}_2(t)C_2 + \mathcal{K}_M(t)C_M) - k_1 G_1, \qquad t > 0, \tag{3.13a}$$

$$\frac{dC_1}{dt} = \mathcal{Q}_1(t)k_1 G_1 - \mathcal{K}_1(t)C_1, \qquad t > 0, \tag{3.13b}$$

$$\frac{dS}{dt} = I_S^{(in)}(t) - I_S^{(out)}(t), \qquad t > 0, \tag{3.13c}$$

$$\frac{dG_2}{dt} = -k_2 G_2 + (1 - \mathcal{P}_2(\tau_S))\ I_S^{(out)}(t - \tau_S(t), t), \qquad t > 0, \tag{3.13d}$$

$$\frac{dC_2}{dt} = \mathcal{P}_2(\tau_S) I_S^{(out)}(t) - \mathcal{A}_2(t) C_2, \qquad t > 0, \tag{3.13e}$$

$$\frac{dC_M}{dt} = \mathcal{Q}_M(t) k_2 G_2 - \mathcal{K}_M(t) C_M, \qquad t > 0, \tag{3.13f}$$

$$\frac{dv_S}{dt} = \begin{cases} -\mathcal{R}_S^- \left(v - v_S^-\right), & c(t) < c_H, \\ \mathcal{R}_S^+ \left(v_S^+ - v\right), & c(t) \geq c_H, \end{cases}, \qquad t > 0, \tag{3.13g}$$

wherein

$$1 = \int_{t-\tau_S(t)}^{t} v_S(\xi) d\xi, \tag{3.13h}$$

$$I_S^{(in)}(t) = (1 - \mathcal{Q}_1(t)) k_1 G_1(t) + \mathcal{K}_1(t) C_1(t), \tag{3.13i}$$

$$I_S^{(out)}(t) = v_S(t) \left. \frac{I_S^{(in)}(s)}{v_S(s)} \right|_{s=t-\tau_S(t)}, \tag{3.13j}$$

$$\mathcal{K}_i(t) = K_i\ H_\epsilon(c(t) - c_H), \quad i \in \{1, M\}, \tag{3.13k}$$

$$\mathcal{A}_2(t) = \alpha_2\ H_\epsilon(c(t) - c_H), \tag{3.13l}$$

$$Q_i(t) = q_i\ H_\epsilon(c_H - c(t)), \quad i \in \{1, M\} \tag{3.13m}$$

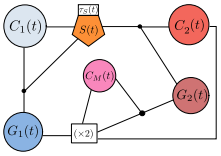
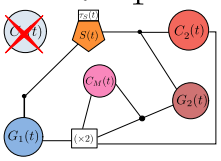
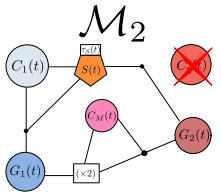
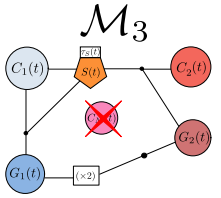
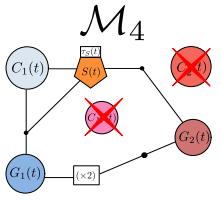
$$\mathcal{P}_2(\tau_S) = p_2\ H_{\epsilon_p}(\tau_S - T_{C2}). \tag{3.13n}$$

and

$$\begin{aligned} v(\xi) &= v_S^+, \quad G_1(\xi) = \pi_{G_1}^+ e^{\lambda_+ \xi}, && -\tau_S(0) \leq \xi \leq 0, \\ S(0) &= \pi_S^+, \quad G_2(0) = \pi_{G_2}^+, \quad C_i(0) = 0 && i \in \{1, 2, M\}. \end{aligned} \tag{3.13o}$$

We denote by $\mathcal{M}_0$ the model given by Eqs. (3.13). The reduced model $\mathcal{M}$ is obtained from Eqs. (3.13) neglecting the elements highlighted in green. Starting from the latter we create a class of models, $\mathcal{S}_\mathcal{M}$, by reducing its complexity. A list of these models are presented in Table 3.1. While the list of possible models is not exhaustive, it is sufficient to test the hypotheses introduced at the beginning of this section. While all models reduce to Eqs. (2.21) under normoxia, they differ in their response to hypoxia. The alternative models are derived from the model $\mathcal{M}_0$, by fixing either $C_1 \equiv 0$ ($\mathcal{M}_1$), $C_2 \equiv 0$ ($\mathcal{M}_2$) or/and $C_M \equiv 0$ ($\mathcal{M}_3$ and $\mathcal{M}_4$). In this list, our proposed model $\mathcal{M}$ is $\mathcal{M}_3$, which we already know can capture the experimental data. We note that the number of unknown parameters associated with model $\mathcal{M}_k$ decreases as k increases (i.e., as the model complexity reduces).

Table 3.1: Schematics showing the cell-cycle control mechanisms included in each model $\mathcal{M}$ belonging to the class of models $\mathcal{S}_{\mathcal{M}}$. When a mechanism is included in the model, we list the parameters associated with it. The model proposed in Section 3.2 ($\mathcal{M}$) is $\mathcal{M}_3$.

	delay transition				$\boldsymbol{\Theta} \in \mathbb{R}^k$
	G1→ S	S→ G2	G2→ M	M→ G1	
$\mathcal{M}_0$ 	True $(K_1,\ q_1)$	True $(\mathcal{R}_S^+,\ \mathcal{R}_S^-)$	True $(T_{C2},\ p_2,\ \alpha_2)$	True $(K_M,\ q_M)$	$k = 9$
$\mathcal{M}_1$ 	False	True $(\mathcal{R}_S^+,\ \mathcal{R}_S^-)$	True $(T_{C2},\ p_2\ ,\alpha_2)$	True $(K_M,\ q_M)$	$k = 7$
$\mathcal{M}_2$ 	True $(K_1,\ q_1)$	True $(\mathcal{R}_S^+,\ \mathcal{R}_S^-)$	False	True $(K_M,\ q_M)$	$k = 6$
$\mathcal{M}_3$ 	True $(K_1,\ q_1)$	True $(\mathcal{R}_S^+,\ \mathcal{R}_S^-)$	True $(T_{C2},\ p_2)$	False	$k = 6$
$\mathcal{M}_4$ 	True $(K_1,\ q_1)$	True $(\mathcal{R}_S^+,\ \mathcal{R}_S^-)$	False	False	$k = 4$

3.4.1 Comparison of different cell-cycle models

We start with a qualitatively comparison of the family of models $\mathcal{S}_{\mathcal{M}}$, using predictions from point estimate fits for both experiments E_1 (constant hypoxia) and E_2 (cyclic hypoxia). Point estimate for the parameter values are obtained by calibrating each model to the experimental data in Fig. 3.2, following the procedure described in Section 3.3.1, and extrapolating the mean of the posterior distribution (parameter values given in Appendix A.2). The results are illustrated in Figs. 3.9-3.10, where predictions from the most complex model $\mathcal{M}_0$ serve as a reference (see white dotted line in the plots).

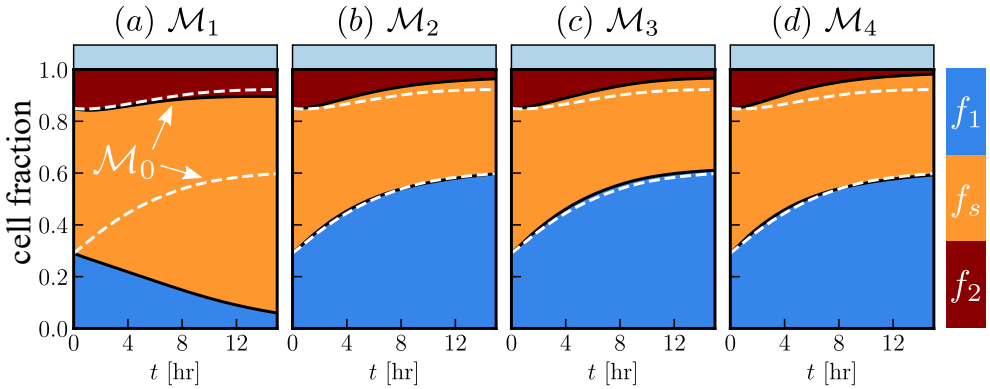


Figure 3.9: Comparison of the cell-cycle dynamics predicted by solving numerically the models $\mathcal{M}_1$-$\mathcal{M}_4$ listed in Table 3.1 for constant hypoxia (scenario E_1). We here use the best fit obtained for each model. The white dotted lines in the panels indicate the evolution predicted by model $\mathcal{M}_0$, used here as a reference. The parameter values of k_1, $v_{\pm}$ and k_2 are as in Table 2.1, while the oxygen-dependent parameters $\boldsymbol{\Theta}$ are taken to be the estimated mean values reported in Table A.2.

In Fig. 3.9, we compare model predictions when cells are exposed to constant hypoxia (E_1). In this case all models, except $\mathcal{M}_1$, predict that f_1 increases, in line with the experimental observations from [5]. In contrast, model $\mathcal{M}_1$ (Fig. 3.9c) is incompatible with the experimental data, predicting a monotonic decrease in $f_1(t)$ and cell accumulation in the S phase (observe the large increase in $f_s(t)$ over time). This suggests that under constant hypoxia, the C_1 checkpoint plays a key role. However, since the simpler model $\mathcal{M}_4$ predicts the same trends as those observed for model $\mathcal{M}_0$, we deduce that both C_2 checkpoints (i.e., C_2 and C_M) can be neglected. By contrast, neglecting the C_1 checkpoint would not significantly impact predictions under cyclic hypoxia. Indeed, Fig. 3.10a shows that model $\mathcal{M}_1$ is in qualitative agreement with model $\mathcal{M}_0$. While $\mathcal{M}_1$ does not capture the rapid fluctuations in $f_1(t)$ predicted by $\mathcal{M}_0$, the overall trend is the same, with $f_1(t)$ decreasing after each cycle. Focusing on

Figs. 3.10b and 3.10d, we see good agreement with $\mathcal{M}_0$ until $t \approx 10$ hours. At later times, both models $\mathcal{M}_2$ and $\mathcal{M}_4$ underestimate the values of f_2 and fail to capture the late accumulation of cells in the G2/M phase (see Fig. 3.2). Interestingly, the predictions of model $\mathcal{M}_2$ are similar to those of model $\mathcal{M}_4$ (where both C_2 checkpoints are neglected). This is because, when the model is fitted to the data, the probability q_M of cells arresting in C_M is estimated to be small ($q_M = 0.08$). This suggests that including a C_M checkpoint that is mediated by instantaneous oxygen levels would not help to explain the experimental observations. Given that model $\mathcal{M}_3$ matches the expected behaviour, we conclude that the C_2 checkpoint is needed to explain the cell-cycle dynamics observed under cyclic hypoxia conditions. This reinforces our claim that memory effects are likely to be important for driving the cell-cycle dynamics under cyclic hypoxia. By contrast, memory effects can be neglected under conditions of acute constant hypoxia.

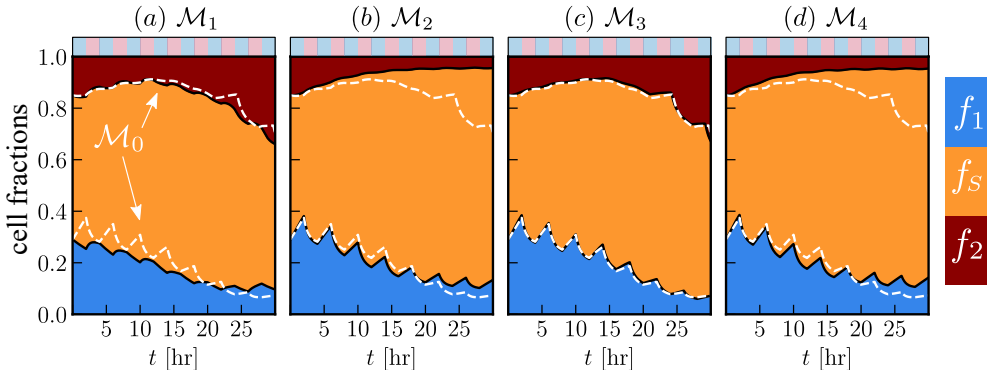


Figure 3.10: Comparison of the cell-cycle dynamics predicted by solving numerically the models $\mathcal{M}_1$-$\mathcal{M}_4$ listed in Table 3.1 for cyclic RH (scenario E_2). The white dotted lines in the panels indicate the evolution predicted by model $\mathcal{M}_0$, used here as a reference. The value of parameters k_1, $v_S^{\pm}$ and k_2 are as in Table 2.1, while the oxygen-dependent parameters $\boldsymbol{\Theta}$ are taken to be the estimated mean values reported in Table A.2 in Appendix A.2.

The results from Figs. 3.9 and 3.10 highlight the inability of models $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_4$ to recapitulate the experimental data. Further, the predictions from $\mathcal{M}_0$ and the much simpler model, $\mathcal{M}_3$, are similar, which suggests that the compartment C_M does not significantly influence the dynamics and that, in conditions of acute hypoxia, cells in the C_2 checkpoint might not re-enter the cell-cycle. Discriminating between models $\mathcal{M}_0$ and $\mathcal{M}_3$ is, therefore, not straightforward and requires consideration of other metrics, in addition to how well the model fits the experimental data. To do this we implement Bayesian model selection using the approach described in Section 1.3.3.

Recall that here the main idea is to estimate the predictive accuracy of the model based on commonly used information criteria (ICs), specifically the DIC and WAIC. The most accurate model is then defined as the one with the lowest value of the information criterion considered. The estimated values of the information DICs and WAICs for the models $\mathcal{M} \in S_{\mathcal{M}}$ are reported in Table 3.2. We see that the two criteria yield similar results, suggesting our conclusions are robust to changes in the information criterion considered.

Based on these estimates, model $\mathcal{M}_0$ is the "best" model in the class $\mathcal{S}_{\mathcal{M}}$. However, it just slightly outperforms the simpler model $\mathcal{M}_3$. As expected from the results presented in Figs. 3.9 and 3.10, $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_4$ exhibit poor performance. In this case, the difference in the estimated ICs is rooted in the estimates of the model predictive accuracy, PA, which is markedly reduced suggesting that these three models fail to fit the experimental data. We conclude that including the C_1 and C_2 compartments (*i.e.* cell-cycle arrest in the G1 and G2 phases) is necessary to reproduce the experimentally observed cell-cycle dynamics. On the contrary, the C_M compartment does not seem to play a significant role in explaining the experimental data.

Table 3.2: Comparison based on the deviation information criterion (DIC) and widely-applicable information criterion (WAIC) of the models in class $\mathcal{S}_{\mathcal{M}}$ that were introduced in Table 3.1. Both information criteria are defined as $-2(\mathtt{PA} - k)$, where PA measures the predictive accuracy of a model, and the correction term k (known as effective number of parameters) accounts for overfitting. We refer to Section 1.3.2 for more details. The two columns ΔDIC and ΔWAIC indicate the relative DIC and WAIC scores with respect to model $\mathcal{M}_0$, i.e., ΔDIC=DIC($\mathcal{M}$)-DIC($\mathcal{M}_0$) and ΔWAIC=WAIC($\mathcal{M}$)-WAIC($\mathcal{M}_0$).

	k_{DIC}	$\mathtt{PA}_{DIC}$	DIC	ΔDIC	k_{WAIC}	$\mathtt{PA}_{WAIC}$	WAIC	ΔWAIC
$\mathcal{M}_0$	5.87	67.98	-124.22	0	3.48	67.51	-128.07	0
$\mathcal{M}_1$	6.31	41.82	-71.01	49.21	4.13	42.17	-76.57	51.5
$\mathcal{M}_2$	4.74	46.14	-82.79	37.43	3.08	46.20	-86.25	41.82
$\mathcal{M}_3$	5.85	65.99	-120.27	3.95	3.49	66.00	-125.02	3.05
$\mathcal{M}_4$	4.25	45.84	-83.18	37.04	3.01	46.07	-86.12	41.95

We see that model $\mathcal{M}_3$, which neglects the compartment C_M, is only slightly worse than $\mathcal{M}_0$ according to both the DIC and the WAIC. The two models yield similar predictive accuracy (PA) as well as "effective number of parameters" (k). This suggests that the additional complexity due to the introduction of the C_M compartment plays only a minor role in explaining the cell-cycle dynamics reported experimentally. In the spirit of adopting the simplest model, we continue employing model $\mathcal{M}_3$ (or

equivalently $\mathcal{M}$) in the remainder of this thesis. Additional experiments and data are needed to determine whether arrest of cells in G2/M due to low oxygen levels (as captured by the compartment C_M) is biologically relevant or might be an artefact of measurement errors and the small number of data points currently available. This interpretation is confirmed by the profile of the marginal posterior for q_M (see Fig. A.2 in Appendix A), whose mean value is approximately $q_M = 0.16$ indicating a very small probability of cells arresting in C_M. This is in contrast to the value of q_1 for which the posterior distribution peaks at $q_1 \approx 1$. We conclude that the instantaneous response of cells to low oxygen levels is to arrest in the G1 phase while they still proceed to mitosis. Similarly, we find that the distribution of α_2 is skewed towards $\alpha_2 = 0$ suggesting that a negligible number of cells are predicted to leave the C_2 compartment. In other words, this suggests that the time-scale required for cells arrested in G2 to transition to mitosis is longer than that resolved experimentally, as we assumed in deriving the reduced model $\mathcal{M}_3$ (or $\mathcal{M}$) in Section 3.2.

Table 3.3: Comparison of the models in class $\mathcal{S}_{\mathcal{M}}$ (see Table 3.1) under different environmental conditions based on the widely-applicable information criterion (WAIC). The WAIC is evaluated by splitting the dataset $\mathcal{E}$ into two: $\mathcal{E}_1$ and $\mathcal{E}_2$ corresponding to data from the experiments E_1 (constant hypoxia) and E_2 (cyclic hypoxia), respectively.

	k_{WAIC}		$\mathtt{PA}_{WAIC}$		WAIC		ΔWAIC	
	E_1	E_2	E_1	E_2	E_1	E_2	E_1	E_2
$\mathcal{M}_0$	1.17	2.34	13.75	53.74	-25.17	-102.80	0	0
$\mathcal{M}_1$	1.71	2.21	3.40	38.69	-3.38	-72.96	21.79	29.84
$\mathcal{M}_2$	0.49	2.45	11.50	34.67	-22.02	-64.45	3.15	38.35
$\mathcal{M}_3$	1.20	2.29	12.32	53.68	-22.24	-102.77	2.93	0.03
$\mathcal{M}_4$	0.51	2.50	11.47	34.59	-21.92	-64.19	3.25	38.61

In computing the information criteria scores in Table 3.2, we used all of the data available, *i.e.*, data from both constant and cyclic hypoxia. As suggested by Figs. 3.9 and 3.10, the models perform more or less alike depending on the environmental conditions considered in the experiment. We re-compute the information criteria by splitting the full data set into two, depending on the hypoxic conditions under which they were measured (see Table 3.3). Since the DIC and WAIC yield similar results, we focus only on the latter. When stratifying the WAIC according to experimental conditions, we find that differences between model performances are significantly dependent on which subset of data is used. In cyclic hypoxia (experiment E_2 in

Table 3.3), the difference in WAIC for model $\mathcal{M}_0$ and $\mathcal{M}_3$ is negligible (basically zero). A small advantage in using model $\mathcal{M}_0$ over $\mathcal{M}_3$ is evident only in constant hypoxia (experiment E_1). However, the difference is again small and not significant. Furthermore, when comparing models $\mathcal{M}_2$ and $\mathcal{M}_3$, the latter performs better, suggesting that even in cyclic hypoxia it may be more plausible that inhibition of cell division in hypoxia results from cells delaying the G2→M transition due to replication stress rather than hypoxia mediated arrest of mitosis (as modelled by the C_M compartment). In cyclic conditions, there is a clear separation of models $\mathcal{M}_0$ and $\mathcal{M}_3$ from the others. In contrast under constant hypoxia, all models have a similar performance except for model $\mathcal{M}_1$. This observation is in agreement with the output of our global sensitivity analysis (see Fig. 3.3) and the simulations in Fig. 3.9. Taken together, this evidence suggests that in constant hypoxia the cell-cycle dynamics are dominated by prolonged inhibition of the G1→S transition, while other mechanisms can be neglected without much loss in performance. It is therefore not surprising that, in the absence of data on cell-cycle dynamics in cyclic hypoxia, cell-cycle models have neglected other mechanisms, in particular the impact of replication stress on the G2→M transition.

3.5 Experimental Design

To conclude, we show how the calibrated model from Section 3.3.2 (*i.e.*, $\mathcal{M}$) can be exploited to inform experimental design. We are particularly interested in designs that can help test some of the model predictions, and improve parameter estimates.

One of the most interesting findings of Chapter 2 is that, according to our model, cyclic hypoxia, as currently defined in the literature, groups together environmental conditions that can drive significantly different cellular responses. For example, when exposed to the (2,2)-cyclic protocols tested experimentally, cells arrest in the G2 checkpoint. However, with cells longer periods of re-oxygenation, we expect cyclic hypoxia to drive delays in the S→G2 transition but not significant delays in the G2→M transition. In other words, the model predicts that a correlation between the duration of the S and G2/M phases of the cell-cycle is not always observed. This is illustrated in Fig. 3.11, where we use point estimates of the inferred parameters to simulate cell response to different cyclic hypoxia protocols. While the duration of the hypoxic phase is set to 2 hr, we allow the time of reoxygenation, $\mathcal{T}_+$, to increase. Looking at Fig. 3.11a, we see a transition in the population growth dynamics as $\mathcal{T}_+$ increases above 5. For small $\mathcal{T}_+$, the model predicts growth inhibition (with N flattening) while

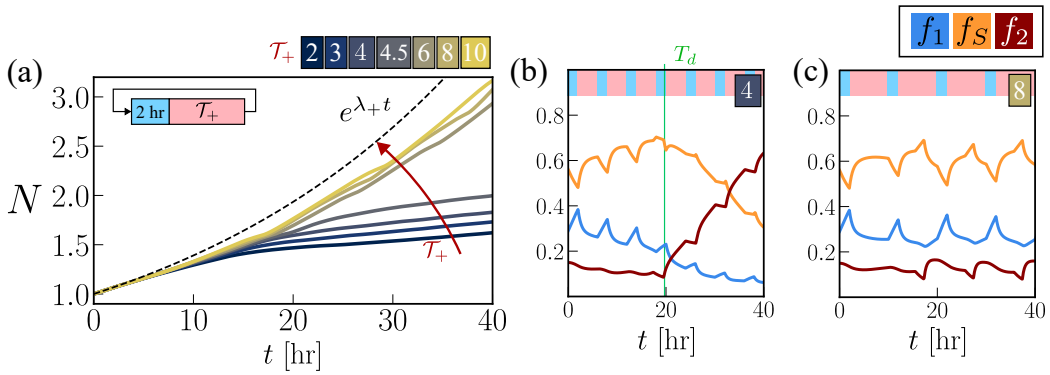


Figure 3.11: Simulation of the cell-cycle dynamics under different cyclic hypoxia conditions. As illustrated in the schematic in panel (a), we consider oxygen cycles where the duration of the hypoxia phase is fixed at 2 hr while we consider variable intervals $\mathcal{T}_+$ between subsequent exposure to hypoxia. (a) Evolution of the population size $N(t)$ for different oxygen cycles as predicted by Eqs. (2.18). The colour of the curves maps to the value of $\mathcal{T}_+$ as indicated in the legend. The black curve ($e^{\lambda_+ t}$) indicates the expected growth dynamics in the case of an oxygen-rich environment (see Section 2.3.1). (b)-(c) Evolution of the phase fraction f_i (see solid lines) with $i \in \{1, S, 2\}$ as predicted by the model for three different cyclic protocols: (b) (2,4)-cyclic hypoxia, *i.e.*, 2 hr in hypoxia with a 4 hr break; (c) (2,8)-cyclic hypoxia, *i.e.*, 2 hr in hypoxia with an 8 hr break. The green vertical line indicates the time, T_d, at which the fraction of cells in the C_2 compartment first rises above the threshold ($\bar{\pi} = 0.05$); when no vertical line is indicated it implies that, for the duration of the simulation, the threshold is never crossed. Rates k_1, k_2 and $v_S^{\pm}$ are set to the default values in Table 2.1, while the unknown parameter values $\boldsymbol{\Theta}$ are the mean posterior estimate obtained by fitting the model to the data (see row $\mathcal{M}_3$ in Table A.2 in Appendix A.2).

for $\mathcal{T}_+ \geq 6$ hr, the population size continue to increase albeit at a lower rate than it would in an oxygen-rich environment. The transition between proliferation inhibition and milder reduction in proliferation rates is dictated by the activation of the G2 checkpoint (*i.e.*, transition of cells in the C2 compartment). This is illustrated in Figs. 3.11b-3.11c, where we illustrate the evolution of the cell-cycle distribution, for different values of $\mathcal{T}_+$. Where proliferation is inhibited (see Figs. 3.11b), the model predicts that a large number of cells will accumulate in the G2 phase. In contrast, when the proliferation is reduced (but not suppressed), the model predicts fluctuations in the fraction of cells in the G2 phase, but no sustained accumulation. To show that the difference in these two regimes is due to cells transitioning into the C_2 compartment, we compute the time T_d at which the fraction of cells in $C_2(t)$ first rises above the threshold $\bar{\pi} = 0.05$, which implies a non-negligible fraction of the population is delayed while attempting to undergo the G2→M transition. In Fig. 3.11b, we see that T_d is finite and coincides with the time at which f_2 starts increasing (despite

some fluctuation). On the contrary, when $\mathcal{T}_+$ is sufficiently large and proliferation is not inhibited, as in Fig. 3.11c, $C_2(t)$ never rises above the value $\bar{\pi}N(t)$ and, therefore, $T_d = \infty$.

The results presented in Fig. 3.11 rely on point estimates of the model parameters. However, one of the advantages of Bayesian inference is that we can capture uncertainty in parameter estimates. In Fig. 3.12 we illustrate how model predictions for the cell-cycle dynamics change when we account for uncertainty (*i.e.*, we illustrate posterior predictions for the cell fractions). In particular, we consider the (2,8)-cyclic hypoxia protocol (2 hours in hypoxia with an 8 hr break) as in Fig. 3.11c. In the absence of uncertainty, we predicted a systematic increase in the fraction of cells in the S phase with no activation of the C_2 compartment. Once uncertainty is taken into account, we find large variability in model predictions at longer times ($t > 25$ hr). Indeed the 68%-confidence interval encompasses the possibility of $f_s(t)$ either increasing or decreasing compared to its initial value, when $t \gg 1$; further, activation of C_2 can not be ruled out as indicated by the value of $f_2(50)$ ranging between $0.1 \lessapprox f_2(50) \lessapprox 0.4$ (see Fig. 3.12c). Despite the large uncertainty in the value of $f_2(t)$, the confidence intervals for $f_1(t)$ and $f_S(t)$ remain reasonably small. The uncertainty in the cell-cycle distribution is reflected in the predictions for the number of cells $N(t)$ (see Fig. 3.12d) where the value of N at the final time, $t = 50$ hr, ranges between $3.5 < N(50) < 4.2$. We note that the uncertainty in model predictions increases over time even if we do not account for uncertainty in the fate of cells entering the C_2 compartment. Based on the results in Fig. 3.12, we can use our calibrated model to predict the cell-cycle dynamics on the short time scale (≈ 25 hr); the large uncertainty at long times prevents us from making reliable predictions on the cell-cycle distribution in this time frame (for this specific cyclic hypoxia protocol). However, the predicted uncertainty is still informative when considering experimental design. In this case, the objective is twofold. On the one hand, we seek to identify experiments that would facilitate model validation. From this point of view, we look for instances where the uncertainty in our predictions is small (*e.g.*, on the short time scale for experiments in Fig. 3.12). On the other hand, where possible, we look for experiments that can improve the accuracy of estimates of model parameters. Unlike before, this requires scenarios where uncertainty in the model predictions is large and, therefore, new measurements can help to refine parameter estimates (such as for the long time dynamics in Fig. 3.12).

In line with the discussion above, we explore a wider range of cyclic hypoxia protocols and classify them according to whether they would be informative for the

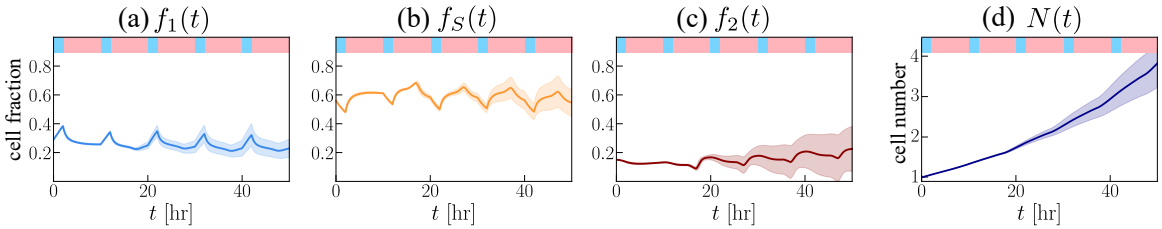


Figure 3.12: Posterior estimates for the time evolution of the cell-cycle distribution (*i.e.*, $f_i(t)$ with $i \in \{1, s, 2\}$) and the total cell number N for the same cyclic protocol as in Fig. 3.11c: 2 hr in hypoxia with a break of 8 hr. We plot the mean estimates as well as the 68%-confidence intervals in the predictions highlighted by the shaded region.

fitting or for model validation. In doing so, we use the variable T_d to distinguish between the scenarios of cyclic hypoxia only mildly or severely affecting the cell-cycle dynamics (see Fig. 3.11). When accounting for uncertainty, the variable T_d becomes random and instead of point estimates we obtain a distribution. For the simulation in Fig. 3.12, the posterior cumulative distribution of T_p, $\pi_{\mathcal{M}}(T_d \leq x|\mathcal{E})$, is illustrated in Fig. 3.13a. Here we consider the support to be between 0 and $t_f = 50$ hr, that is, the duration of the simulations. By evaluating $p_d = \pi_{\mathcal{M}}(T_d \leq t_f|\mathcal{E})$, we can estimate that there is approximately a 20% chance that the number of cells arrested in the C_2 compartment will grow to constitute more than 5% of the total cancer cell population. In other words, there is large uncertainty on how cells would respond to the oxygen protocol considered. We can repeat the same calculation for a range of protocols.

In Fig. 3.13b we illustrate how the probability p_d varies as we change both the duration of the hypoxic phase, $\mathcal{T}_-$, and the interval between subsequent periods of hypoxia $\mathcal{T}_+$ (see schematic on the top of Fig. 3.13b). We find that, for sufficiently small $\mathcal{T}_+$ (values below the dark purple curve in Fig. 3.13b) the model predicts, with high certainty, activation of the G2 checkpoint ($p_d > 0.95$). On the contrary, for sufficiently high $\mathcal{T}_+$ (values above the yellow line in Fig. 3.13b), the model predicts, with high certainty, that the G2 checkpoint will not be activated ($p_d < 0.05$). Despite leading to opposite conclusions, what these experimental designs have in common is the high-level of certainty with which the model predicts their outcome. We, therefore, classify these as designs that aid model validation (see Fig. 3.13c). For intermediate values of $\mathcal{T}_+$, model predictions are, instead, more uncertain, $0.05 \leq p_d \leq 0.95$, implying that the model can not predict the response of cells with certainty. As discussed above, this type of experimental design can be useful in refining parameter estimates due to the large uncertainty associated with model predictions (see Fig. 3.13c).

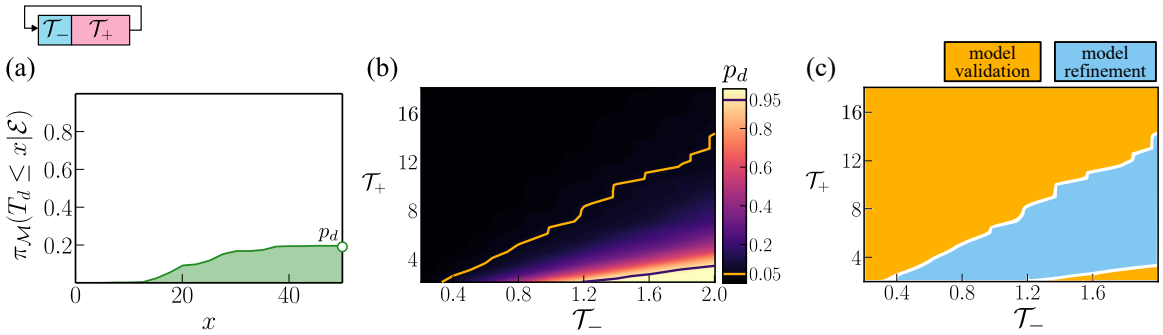


Figure 3.13: Posterior estimates for the time T_d at which the number of cells in the C_2 compartment first rises above 5% of the total population. (a) Cumulative posterior distribution for the numerical simulation presented in Fig. 3.12; here p_d is used to indicate the value of the cumulative distribution at $x = t_f$, *i.e.*, $p_d = \pi_M(T_d \leq t_f|\mathcal{E})$ (b) estimated value of p_d (defined as indicated in panel (a)) for a variety of oxygen protocols: $\mathcal{T}_-$ hours in hypoxia with a break of $\mathcal{T}_+$ hours in physiological oxygen conditions (see schematic on top of panel (a)). We indicate on the colourmap the contour lines $\pi_M(T_d \leq x|\mathcal{E}) = 0.05$ and $\pi_M(T_d \leq x|\mathcal{E}) = 0.95$. (c) Using the contour lines identified in panel (b), we subdivide the space of experimental designs $(\mathcal{T}_-, \mathcal{T}_+)$ into two regions: in yellow we identify experimental designs that can be used to validate our model predictions and assumptions; in light blue instead we indicate experimental designs that can help in the refinement of our parameter estimates.

In practice, there are limitations on the regions of the experimental design space $(\mathcal{T}_+, \mathcal{T}_-)$ that can be tested. Experimentally, it might be difficult to replicate oxygen dynamics where oxygen levels change on a faster time scale than two hours due to the time required for oxygen levels to equilibrate *in vitro* [116]. This, however, is not a limitation for our model. From this point of view, with the current parameter estimates, the model already makes some definite predictions on the response of RKO cells to short periods of hypoxia. We indeed find, with high certainty, that infrequent ($\mathcal{T}_+ > 2$ hr) and brief ($\mathcal{T}_- < 0.4$ hr) periods of hypoxia would not substantially perturb the cell-cycle dynamics and would only result in a slight increase in the duration of the S phase, but no delays in the subsequent G2 phase. This illustrates how, despite the remaining uncertainty in parameter estimates, the model can still retain predictive power and unravel cell responses to environmental conditions that are observed in tumours but that, with the available technology, could not be otherwise investigated experimentally.

3.6 Discussion

In Chapter 2 we developed a new model to describe the effect of cell-cycle dysregulation in hypoxia on the growth dynamics of a population of cancer cells. While based on biological evidence, the model is phenomenological and, therefore, requires validation against experimental data. We here relied on recently published data from [5] to calibrate and validate the model. These resolve the short term effects of constant and cyclic hypoxia on cell-cycle dynamics in the RKO cancer cell-line. Unfortunately, no data are currently available that quantify the long-term effect of cyclic hypoxia on cell proliferation and viability, which can be predicted by our model. In light of the gap between the complexity of our model and the data available, we first used model reduction techniques, based on sensitivity indices, to tailor our model to the quality of the data available. We found that an adequate model, $\mathcal{M}$, can be obtained from the original one by neglecting the long-term fate decisions of cells arresting in the G2 phase. This reduces the number of unknown model parameters from 11 down to 6. We then showed how Bayesian inference techniques can be used to calibrate our model and that the predictions of the calibrated model agree well with the available experimental data. We also show that the calibration procedure is not sensitive to assumptions on the likelihood function used to estimate model parameters.

As often in biology, multiple mechanisms have been proposed to explain cell responses to hypoxia. It is, therefore, not apparent which mechanisms should be included in a model and which are, instead, negligible. In order to justify the assumptions used to construct our model, in Section 3.4 we used data-driven modelling approaches to test and compare the role of different mechanisms driving cell-cycle dynamics in hypoxic conditions. To do so we generated a class of candidate models, $\mathcal{S}_{\mathcal{M}}$, based on an augmented version of our original model $\mathcal{M}$. By applying Bayesian model selection techniques we were able to systematically show that the inclusion of the C_1 and C_2 compartments is necessary to capture the experimentally observed cell-cycle dynamics. In contrast, the addition of the checkpoint compartment C_M, omitted in our model, can be neglected without significantly impacting the quantitative agreement between the model predictions and the data (particularly those predictions related to cell-cycle dynamics in cyclic hypoxia). These results suggest that memory effects, instead of an instantaneous response to oxygen levels, drive cell arrest in G2/M under acute exposure to cyclic hypoxia.

To conclude, we have shown how our modelling framework can be used to predict cell-cycle dynamics and inform the design of *in vitro* experiments aimed both at

refining our parameter estimates, as well as validating predictions of the calibrated model. While experiments are limited in the time-frame they can resolve (on the order of days), we expect tumour cells *in vivo* to experience hypoxia for much longer times. From this point of view, mathematical modelling can offer unique insights into cell responses that have not been investigated experimentally. In the next chapter,we will show how the original model proposed in Chapter 2 can be applied to study the long-term effect of different hypoxic environments on the growth dynamics of tumours. Since growth curves are measurable, our predictions can be tested experimentally and, therefore, used to verify the validity or reveal the limitations, of our modelling framework.

Chapter 4

Predicting the asymptotic growth dynamics of cell cultures in hypoxia

4.1 Overview

In the previous chapters, we introduced a new mathematical model to describe the cell-cycle dynamics of a population of cancer cells and showed how it can be used to predict the short-term cell-cycle dynamics of the RKO cancer cell-line when grown in different hypoxic environments. Having validated our model against experimental data, we now use it to investigate how long term exposure to different forms of hypoxia impact the growth dynamics and the tumourogenic capacity (*i.e.*, their ability to form a growing mass) of cells cultured *in vitro*. In Chapter 2, we leveraged numerical simulations to explore the short and long term cell-cycle dynamics under a range of environmental conditions. These preliminary results already suggest that the model can capture a rich range of dynamics, from exponentially growing populations to complete growth arrest. In this chapter, we leverage the theory of positive dynamical systems presented in Section 1.2.3, to investigate the model predictions on the long term effect of hypoxia on the growth dynamics of cancer cells that have different cell-cycle control strategies. In doing so, we gain insights on which cell-cycle control mechanisms allow population growth under hypoxia and which, instead, inhibit growth.

In Section 4.2, we revise the concept of asynchronous exponential growth that we first introduced Section 2.3.1 as a property of a population of cells growing in resource-rich environments. Here, we want to link the biological interpretation of asynchronous exponential growth with the asymptotic dynamics of strongly positive linear systems and generalise its definition to situations in which populations grow in periodically changing environments. In Section 4.3, we use the model to study the

asymptotic growth dynamics of a population of cancer cells exposed to constant hypoxia. Under these conditions, the cell-cycle model proposed in Chapter 2 reduces to a system of autonomous delay-differential equations which differs from the model derived in Section 2.3.1 for a constant, oxygen-rich environment since it accounts for the checkpoint compartments. While the model equations yield a strongly positive linear dynamical system, the system does not always have the property of asynchronous exponential growth. We derive analytical conditions on model parameters that guarantee the property of asynchronous exponential growth. These results yield biological insights into the cell-cycle control mechanisms that drive growth arrest in cancer cell populations exposed to hypoxia.

In Section 4.4, we exploit the asymptotic dynamics of our model to study the long term effect of cyclic hypoxia. Under these conditions, the cell-cycle model proposed in Chapter 2 reduces to a system of non-autonomous periodic delay-differential equations. Leveraging again the theory of strongly positive linear dynamical systems, we show that, under appropriate assumptions on the model parameters, the model has the property of asynchronous exponential growth. However, in contrast to the case of constant hypoxia, it is not possible to derive explicit analytical conditions to determine whether the system display the property of asymptotic exponential growth. We instead rely on a numerical method based on Galerkin-approximations. After briefly introducing this approach in Section 4.4.1, in Section 4.4.2 we focus on presenting and interpreting our results. Interestingly we find that predictions of asynchronous exponential growth, depend not only on how a cell regulates its cell-cycle in hypoxia, but also on the specific cyclic hypoxia protocol considered. From this point of view, our results echo the discussion from Section 2.5 in Chapter 2 and suggest that it is not possible to determine generally the effect of cyclic hypoxia on a population of cells; instead, it is more appropriate to discuss the impact of a specific subset of oxygen protocols on a cell's ability to proliferate. In Section 4.4.3, we combine the results from Section 4.4.2 and simple heuristic arguments to obtain a cell-specific classification of cyclic hypoxia that maps oxygen dynamics to toxicity. We conclude by summarising our findings in Section 4.5.

4.2 Asynchronous exponential growth

The concept of asynchronous exponential growth was first introduced by cell biologists to describe the growth of cell populations [149]. As discussed in Section 2.3.1, the regime of asynchronous exponential growth describes a situation where the total

number of cells, N, grows exponentially at a constant rate ω (*i.e.*, $N \propto e^{\omega t}$) and the distribution of cells in the different phases of the cell-cycle tends to a constant profile defining what we have called the *phase stationary solution*, that is, the solution is independent of how the cells were initially distributed along the cell-cycle. Mathematicians have formalised the concept of asynchronous exponential growth as a property of the solution of linear autonomous dynamical systems whose solutions converge asymptotically to phase stationary solutions [149]. For an autonomous system the property of asynchronous exponential growth is equivalent to requiring the existence of an exponentially dominant solution with positive growth rate λ (see Section 1.2.1). However, the concept of asynchronous exponential growth, unlike the one of dominant solution, does not apply to periodic system, for which phase stationary solution do not exists. Here, we are interested in extending the concept of asynchronous exponential growth to account for the growth dynamics of cells exposed to environments that change periodically over time.

In the literature, it is common to quantify population growth dynamics (or proliferation capacity) via the per capita growth rate ω, which is defined as the infinitesimal relative increase in the number of individuals in a population:

$$\omega = \frac{1}{N}\frac{dN}{dt}. \tag{4.1}$$

While we use the term "growth rate" to denote $\omega(t)$, this can be either positive or negative. As such, $\omega(t)$ gives an indication of whether the population is increasing or decreasing in size and how fast it is evolving. For a population growing exponentially, $\omega(t)$ is constant and corresponds to the so-called *Malthusian* parameter ω [149]. However, growth (in absence of competition) of a structured population in a periodically changing environment is better described mathematically by a non-autonomous, periodic, positive linear system, whose asymptotic dynamics do not result in $\omega(t)$ being constant, but rather to $\omega(t)$ fluctuating over time. This suggests that, in dynamic environments, $\omega(t)$ might not be a suitable metric to quantify how a periodically-changing environment affects a population's overall growth dynamics.

To illustrate this point we consider the simple example presented in Fig. 4.1. Here, a population of cells grows in a periodically changing environment and we assume that its size evolves according to the following law: $N(t) = (2 + \cos(2\pi t))e^{\lambda t}$. The evolution of N is plotted in Fig. 4.1a. While the total cell number, N, fluctuates over time, there is an apparent positive trend that suggests the population would eventually grow unbounded. This feature, however, is not readily obvious if we consider the time-evolution of the (instantaneous) growth rate, $\omega(t)$, as the latter fluctuates

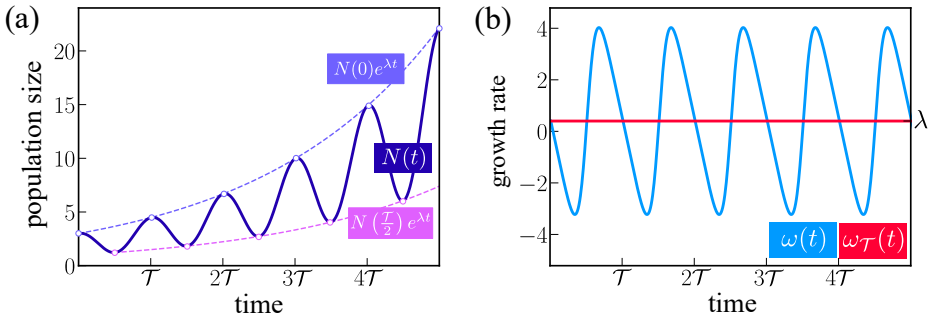


Figure 4.1: Describing growth in a periodic environment. (a) Evolution of the total population size as prescribed by the function $N(t) = (2+\cos(2\pi t))e^{\lambda t}$ with $\lambda = 0.3\mathcal{T}$. (b) Plot comparing the time-evolution of the growth rate $\omega(t)$ (see Eq. (4.1)) and the generalised Malthusian parameter $\omega_{\mathcal{T}}(t)$ (see Definition 4.2.1) for a population following the growth plotted in panel (a).

between positive and negative values (see blue curve in Fig. 4.1b). In this case, the instantaneous growth rate captures information about fluctuations in the population size (*i.e.*, the fast time scale), but it fails to characterise the long term dynamics of the population (*i.e.*, the slow time scale). There has been increasing interest in the literature to extending standard concepts such as the Malthusian parameter to describe the evolution of heterogeneous populations in fluctuating environments (see [70] and references therein). Unlike in constant environments, a more natural way of studying the evolution of N in periodically changing conditions (which fluctuate with period $\mathcal{T}$) is to look at the overall change in the population over $\mathcal{T}$. To do so, we can define a discrete map $\Phi(t) : [0, \infty) \to [0, \infty)$ which maps $N(t)$ into $N(t + \mathcal{T})$. In our simple example the map is independent of t and given by $\Phi = e^{\lambda \mathcal{T}}$. In Fig. 4.1a, the discrete map is represented by the white dots, which can be interpolated via a growing exponential function $e^{\lambda t}$ (see dashed curve). This discrete description of the time-evolution naturally leads to the alternative definition of growth rate.

Definition 4.2.1. *Let us consider a population with size $N = N(t)$ growing in an environment that fluctuates with period $\mathcal{T}$. We define its growth rate as:*

$$\omega_{\mathcal{T}}(t) = \frac{1}{\mathcal{T}} \ln\left(\frac{N(t+\mathcal{T})}{N(t)}\right). \tag{4.2}$$

As shown in Fig. 4.1b for the example discussed above, the growth rate $\omega_{\mathcal{T}}$ is constant over time and captures the overall positive trend in the evolution of N. In general, we will refer to cell populations whose dynamics follow the growth law $N(t) = CN(t+\mathcal{T})$ ($C > 0$) as being in a regime of quasi-periodic exponential growth. It is therefore natural to introduce the concept of asynchronous quasi-periodic exponential growth. Mathematically, a non-autonomous $\mathcal{T}$-periodic linear system will have the property of

exponential growth if it admits an exponentially dominant solution. From a biological point of view, this regime corresponds to a state where the total number of cells undergoes quasi-periodic exponential growth, with period $\mathcal{T}$, whilst the distribution of cells in different states of the cell-cycle changes periodically, with its average over intervals of the form $[t, t + n\mathcal{T}]$ $(n \in \mathbb{N})$ being constant.

In the next sections, we analyse how different forms of hypoxia impair the ability of cancer cell cultures to give rise to exponentially growing cell populations. From a mathematical point of view, this analysis will be equivalent to determining whether the parameters associated with a specific cell behaviour allow the model to exhibit asynchronous growth (either exponential or quasi-periodic exponential).

4.3 Asymptotic growth dynamics in constant hypoxia

Before diving into the analysis, we first recall some of the results from the numerical simulations presented in Section 2.4.1. These results suggested that our model exhibits three different "growth" regimes: sustained growth (cell-line 4 in Fig. 2.7a); growth inhibition driven by arrest (cell-line 1 in all panels of Fig. 2.7) and growth inhibition driven by cell death (cell-line 6 in Fig. 2.7b). Based on the per capita population growth rate $\omega(t)$ (see Eq. (4.1)), we can discriminate between these different scenarios as follows: $\omega(t) \to \omega > 0$ (sustained growth); $\omega(t) \to 0^+$ (growth inhibition driven by arrest) and $\omega(t) \to 0^-$ (growth inhibition driven by death). In this section, we investigate how the different mechanisms included in the model (i.e., groups of parameters) interact to determine the asymptotic population growth dynamics.

The model presented in Chapter 2 contains several dependent variables. Some are continuous fields that depend uniquely on the oxygen levels and determine how cells behave. These include: the internal level of proteins (m); the velocities at which cells can synthesise DNA (v_S), and cell recovery (v_R); the delays associated with the duration of the S phase ($\tau_S(t)$) and the time required for a cell to recover in the G2 checkpoint ($\tau_R(t)$). When considering constant hypoxia, it is straightforward to show that, after a brief transient, these variables equilibrate to new constant values:

$$m = M^-, \quad v_{s,R} = v_{S,R}^-, \quad \tau_{S,R} = \tau_S^- \tag{4.3}$$

$$\text{where } v_R^- = v_R^+ \frac{M^-}{M_R + M^-}, \ \tau_S^- = \frac{1}{v_S^-}, \ \tau_R^- = \frac{y_R}{v_R^-}. \tag{4.4}$$

All the parameters in Eq. (4.4) are defined in Table 2.1. The remaining dependent variables describe how the number of cells in each model compartment evolves over time. We subdivide these variables into three groups: $\boldsymbol{X}(t) = [G_1(t), G_2(t)]$, $\boldsymbol{Y}(t) = [S(t), D_2(t)]$ and $\boldsymbol{A}(t) = [C_1(t), \mathrm{Sn}_2(t)]$. The reason for such a subdivision will become clear in what follows. By substituting Eqs. (4.4) into Eqs. (2.18), we find that, for $t \gg 1$, the dynamics of $\boldsymbol{X}$ can be approximated by the following system of coupled ordinary and delay differential equations:

$$\frac{dG_1}{dt} = 2k_2 G_2 - k_1 G_1, \tag{4.5a}$$

$$\begin{aligned}\frac{dG_2}{dt} = &-k_2 G_2 + k_1(1-p_2)(1-q_1)\exp\left[-\mu_S \tau_S^-\right] G_1(t-\tau_S^-) \\ &+ (1-p_{Sn})p_2 k_1 (1-q_1) e^{-\mu_S \tau_S^- - \mu_2^C \tau_R^-} G_1(t-\tau_{S+R}^-),\end{aligned} \tag{4.5b}$$

while $\boldsymbol{A}$ evolves according to:

$$\frac{dC_1}{dt} = q_1 k_1 G_1, \tag{4.6a}$$

$$\frac{d\mathrm{Sn}_2}{dt} = k_1 p_{Sn} p_2 (1-q_1)\exp\left[-\mu_S \tau_S^-\right] G_1(t-\tau_S). \tag{4.6b}$$

If t is sufficiently large (specifically $t > \tau_R^- + \tau_S^-$), we can solve exactly for the $\boldsymbol{Y}$ variables:

$$S(t) = (1-q_1)k_1 \int_{t-\tau_S^-}^{t} G_1(\xi) e^{-\mu_S(t-\xi)} d\xi, \tag{4.7a}$$

$$D_2(t) = (1-q_1)k_1 e^{-\mu_S \tau_S^-} p_2 (1-p_{Sn}) \int_{t-\tau_R^-}^{t} G_1(\xi - \tau_S^-) e^{-\mu_2^C (t-\xi)} d\xi. \tag{4.7b}$$

In Eqs. (4.5)-(4.7), we define $\tau_{S+R}^- = \tau_S^- + \tau_R^-$, while p_2, p_{Sn} and μ_S are defined as in Table 2.1. From Eqs. (4.5)-(4.6), it is clear that the variables $\boldsymbol{X}$ represent compartments that the cells can enter and leave under constant hypoxia, while variables $\boldsymbol{A}$ relate to compartments that cells can enter but not leave. Variables $\boldsymbol{Y}$, instead, are associated with compartments where cells remain for a finite, prescribed amount of time.

In contrast to normoxic conditions, when the model is specialised for constant hypoxia (i.e., Eqs. (4.5))-(4.7), it admits non-trivial steady states. More precisely, it admits infinitely many stationary solutions that can be parametrised as follows:

$$\hat{\boldsymbol{X}} = (0,0), \quad \hat{\boldsymbol{Y}} = (0,0), \quad \hat{\boldsymbol{A}} = (N_{C1}, N_{Sn}), \tag{4.8}$$

where N_{C1}, N_{Sn} are arbitrary non-negative constants. The only non-trivial solutions allowed by the model are, therefore, solutions in which all cells arrest from the mitotic

cycle either by transitioning into the C_1 checkpoint, or by becoming senescent (*i.e.*, transitioning into Sn_2). Since the model is linear, we have that solutions either converge to the equilibrium point in Eqs. (4.8), or they grow unbounded.

As we will show in what follows, the asymptotic dynamics of the system can be characterised by analysing at the subsystem associated with $\boldsymbol{X}$ (*i.e.*, Eqs. (4.5)). For proper choices of the model parameters (as detailed in Theorem 4.3.1), Eqs. (4.5) yield a strongly positive system. Based on Definition 1.2.1, this system admits an exponentially dominant solution of the form $\boldsymbol{\xi_x} e^{\lambda_- t}$, where λ_- is a simple (with multiplicity one) real eigenvalue with a strictly positive eigenvector $\boldsymbol{\xi_X} > 0$, which indicates the fraction of cells in each compartment G_1 and G_2 at long times. We identify three possible scenarios:

- if $\lambda_- > 0$, then the system admits the property of asynchronous population growth. Eventually, all model variables, and the total population size, N, converge to a regime of exponential (or geometric) growth at rate λ_-, (*i.e.*, $\omega(t) \to \lambda_-$);
- if $\lambda_- = 0$, then the population eventually reaches a regime of linear (or algebraic) growth, (*i.e.*, $N \propto t$ as $t \to \infty$);
- if $\lambda_- < 0$, then the system evolves towards a stationary solution as defined by Eq. (4.8), (*i.e.*, $\omega(t) \to 0$ as $t \to \infty$).

We compute the eigenvalue λ_- by deriving the characteristic polynomial P_- for Eqs. (4.5) by seeking for solutions of the form $\boldsymbol{X} = \boldsymbol{\xi_X}(\Lambda) e^{\Lambda t}$. After some algebra, we find that P_- satisfies

$$P_-(\Lambda) = 2k_1k_2(1-q_1)e^{-(\Lambda+\mu_S)\tau_S^-}\tilde{P}_-(\Lambda) - (\Lambda+k_1)(\Lambda+k_2), \tag{4.9a}$$

$$\text{where } \tilde{P}_-(\Lambda) = (1-p_2) + (1-p_{Sn})p_2 e^{-(\Lambda+\mu_2^C)\tau_R^-}. \tag{4.9b}$$

Using the expressions given in Eqs. (4.9), we can characterise the sign of λ_- and, therefore, whether cell cultures maintain their growth when exposed to hypoxia. Our findings are summarised in Theorem 4.3.1 and reveal that the long term effect of hypoxia depends on the value of the parameter

$$\chi = [(1-p_2) + (1-p_{Sn})p_2 e^{-\mu_2^C \tau_R^-}](1-q_1)e^{-\mu_S \tau_S^-}. \tag{4.10}$$

Specifically, if χ is greater than the critical value $\chi_c = 1/2$, then the cells eventually converge to a regime of balanced exponential growth at a rate λ_-, which is less than

the proliferation rate in oxygen-rich conditions (λ_+). If $\chi \leq \chi_c$, then cell proliferation is eventually inhibited. This condition has a clear physical interpretation when we consider the meaning of each factor in Eq. (4.10). The term in the square brackets represents the probability that a cell successfully completes the G2→M transition, either without delays (with probability $1 - p_2$) or by first allowing time for repair (second term in the sum in the square brackets). The factors $(1-q_1)$ and $e^{-\mu_S \tau_S^-}$ represent, respectively, the probability that a cell successfully completes the G1→S and S→G2 transitions. Since all transitions are independent events, χ corresponds to the overall probability that a cell successfully completes one mitotic cycle. The condition $\chi > 1/2$ implies that, for a cell population to grow in hypoxia, the probability that a cell in the colony successfully completes one round of the mitotic cell-cycle must be greater than the probability it will fail.

Theorem 4.3.1. *For any value of $k_1 > 0$, $k_2 > 0$, $0 < q_1 < 1$ and $0 < p_2 < 1$, $0 < p_{Sn} < 1$, $\mu_S, \mu_2^C \geq 0$ and $\tau_S^- > \tau_S^+ > 0$, the right-most root of $P_-(\Lambda) = 0$, denoted by λ_-, as defined by Eq. (4.9), is real. Let χ be defined as in Eq. (4.10). Then we have the following behaviour:*

1. *if $\chi > \frac{1}{2}$, then $\lambda_- > 0$. Additionally, we have that $\lambda_- < \lambda_+$, where λ_+ is the right-most root of Eq. (2.25).;*

2. *if $\chi \leq \frac{1}{2}$, then $\lambda_- \leq 0$.*

Proof. First, we prove that, under the conditions on the parameter values listed above, the system defining $\boldsymbol{X}$ is a strongly positive dynamical system [129].

1. Eqs. (4.5) generate a positive dynamical system. Following Example 1.2.1 in Section 1.2.3, we construct the matrix $\mathcal{A}$ associated with Eqs. (4.5a)-(4.5b) when delays are neglect:

$$\mathcal{A} = \begin{bmatrix} -k_1 & 2k_2 \\ k_1\chi & -k_2 \end{bmatrix}, \tag{4.11}$$

where χ is as in Eq. (4.10). Given the conditions on the model parameters in the statement of Theorem 4.3.1, $\chi > 0$, $k_1 > 0$ and $k_2 > 0$ so that it follows that the matrix $\mathcal{A}$ is positive (and therefore Metzler), and, hence, that Eqs. (4.5a)-(4.5b) define a positive dynamical system.

1. Eqs. (4.5) generate a strongly positive dynamical system. To prove irreducibility, we follow the same procedure as in Example 1.2.1 and construct the directed graph $\mathcal{G}_\mathcal{A}$ associated with the matrix $\mathcal{A}$. From Fig. 4.2, it is apparent that the graph is strongly connected. In at most 1 steps, it is possible to reach any node

in the graph from any other node. Given that the graph $\mathcal{G}_{\mathcal{A}}$ is strongly connected, we conclude that the matrix $\mathcal{A}$ is irreducible and, so, Eqs. (4.5a)-(4.5b) yield a strongly positive dynamical system. This implies that, the dominant eigenvalue of Eqs. (4.5a)-

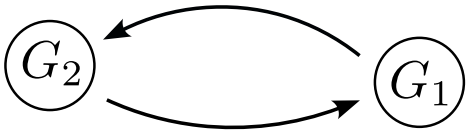


Figure 4.2: Directed graph, $\mathcal{G}_{\mathcal{A}}$, constructed from the matrix $\mathcal{A}$ defined by Eq. (4.11) following the procedure described in Section 1.2.3.

(4.5b), and therefore, the dominant root of the polynomial P_- will be simple and real. To investigate the sign of λ_-, we follow arguments similar to those used in the proof of Lemma 2.3.1 and look at the sign of $P_-(0) = k_1 k_2(2\chi - 1)$, where χ is defined by Eq. (4.10).

1. If $\chi > \frac{1}{2}$, then $P_-(0) > 0$. Given that $\lim_{\Lambda\to\infty} P_-(\Lambda) = -\infty$ for $\Lambda > 0$ and P_- is a continuous function of Λ, we conclude that there exists at least one positive real root. Given that $\tau_S^- > \tau_S^+$, we also have that λ_- is the only positive root. Let us now consider the right-most root λ_+ for the polynomial P_+ (see Eq. (2.25)). We have that $(\lambda_+ + k_1)(\lambda_+ + k_2) = 2k_1 k_2 e^{-\lambda_+ \tau_S^+}$, where $\tau_S^+ < \tau_S^-$ by assumption. Evaluating $P_-(\lambda_+)$ and, after some algebra, we obtain $P_-(\lambda_+) < 0$. Given that $P_-(\lambda_+) < P_-(\lambda_-)$ and $P_-(\Lambda)$ is monotonically decreasing for positive Λ, this implies that $\lambda_+ > \lambda_-$.

2. Suppose now that $\chi \leq 1/2$ so that $P_-(0) \leq 0$. Given that $P_-(\Lambda) \to -\infty$ (where $'$ denotes the derivative) as $\Lambda \to +\infty$ and $P_-'(\Lambda) < 0$ for $\Lambda > 0$, there can not be a positive real root of $P_-(\Lambda)$. Therefore we have that the largest *real* eigenvalue for the polynomial P_- is negative.

□

While Theorem 4.3.1 characterises the long time dynamics of the subset of variables $\boldsymbol{X}$, we must now consider the asymptotic population growth dynamics (*i.e.*, the dynamics of the total cell numbers, $N(t)$). This requires us to solve for the remaining unknowns (*i.e.*, $\boldsymbol{Y}$ and $\boldsymbol{A}$). Based on Theorem 4.3.1, we find that the solution of Eqs. (4.5a)-(4.5b) eventually converges to:

$$X(t) \approx \xi_X e^{\lambda_- t}, \quad \text{for } t \gg 1, \text{ with } X \in \{G_1, G_2\}, \tag{4.12a}$$

where the constants ξ_X are subject to the constraint $(\lambda_- + k_1)\xi_{G_1} = 2k_2\xi_{G_2}$. Furthermore, using Eqs. (4.6a)-(4.7), we obtain that:

$$S(t) \approx \frac{(1-q_1)k_1[1-e^{-(\mu_S+\lambda_-)\tau_S^-}]}{\mu_S+\lambda_-}\,\xi_{G_1}e^{\lambda_- t} \equiv \xi_S e^{\lambda_- t}, \tag{4.12b}$$

$$D_2(t) \approx \frac{(1-q_1)k_1p_2(1-p_{Sn})e^{-(\mu_S+\lambda_-)\tau_S^-}[1-e^{-(\mu_2^C+\lambda_-)\tau_R^-}]}{\mu_2^C+\lambda_-}\,\xi_{G_1}e^{\lambda_- t} \equiv \xi_{D_2}e^{\lambda_- t}, \tag{4.12c}$$

$$C_1(t) \approx N_{C1} + \begin{cases} \frac{q_1k_1\xi_{G_1}}{\lambda_-}e^{\lambda_- t}, & \lambda_- \neq 0, \\ q_1k_1\xi_{G_1}t, & \lambda_- = 0, \end{cases} \tag{4.12d}$$

$$\mathrm{Sn}_2(t) \approx N_{Sn} + \begin{cases} \frac{(1-q_1)p_2k_1p_{Sn}\xi_{G_1}e^{-(\mu_S+\lambda_-)\tau_S^-}}{\lambda_-}e^{\lambda_- t}, & \lambda_- \neq 0, \\ (1-q_1)p_2k_1p_{Sn}\xi_{G_1}e^{-\mu_S\tau_S^-}t, & \lambda_- = 0, \end{cases} \tag{4.12e}$$

where, as previously, N_{C1} and N_{Sn} are non-negative constants. Given Eqs. (4.12), we now have all the information needed to characterise the system long time behaviour. Our findings are summarised in Proposition 4.3.1.

Proposition 4.3.1. *Under the assumptions on the model parameters stated in Theorem 4.3.1, we find that:*

- *if $\lambda_- > 0$, then Eqs. (4.6)-(4.7) have the property of balanced exponential growth. Eventually, for all initial conditions, the population size increases exponentially at rate λ_-,* (i.e., $\lim_{t\to\infty} N = \xi_N e^{\lambda_- t}$)*;*
- *if $\lambda_- = 0$, then the population eventually reaches a regime of linear growth,* (i.e. $\lim_{t\to\infty} N = \xi_N t$)*;*
- *if $\lambda_- < 0$, then the system evolves towards a stationary solution, as defined by Eq. (4.8), with $\lim_{t\to\infty} N = \bar{N} = N_{C1} + N_{Sn} \geq 0$. If $\bar{N} = N_{C1} + N_{Sn} > 0$, we find that the asymptotic behaviour of the instantaneous population growth $\omega(t)$ (see Definition 4.2.1), depends on the value of the constant γ where:*

$$\gamma = \frac{\lambda_- + k_1}{2} - \frac{\mu_S k_1(1-q_1)[1-e^{-(\lambda_-+\mu_S)\tau_S^-}]}{\lambda_-+\mu_S} - \frac{\mu_2^C p_2(1-p_{Sn})k_1(1-q_1)e^{-(\lambda_-+\mu_S)\tau_S^-}[1-e^{-(\lambda_-+\mu_2^C)\tau_R^-}]}{\lambda_-+\mu_2^C}. \tag{4.13}$$

If $\gamma > 0$, then growth inhibition is driven by cell arrest ($\lim_{t\to\infty}\omega(t) = 0^+$), while if $\gamma < 0$, then growth inhibition is driven by cell death ($\lim_{t\to\infty}\omega(t) = 0^-$).

Proof. Recalling that the total cell number, N, is given by the sum of the number of cells in all compartments (see Eq. (2.19)), its asymptotic growth dynamics as a function of λ_- follows straightforwardly from Eqs. (4.12).

We now focus on the case $\lambda_- < 0$, to characterise how ω asymptotically approaches 0. Using the definition of dN/dt (see Eq. (2.19)) and Eqs. (4.12), we find that:

$$\frac{dN}{dt} \approx \gamma \xi_{G_1} e^{\lambda_- t}, \quad t \gg 1, \tag{4.14}$$

where γ is defined as in Eq. (4.13). When $\lambda_- < 0$, $dN/dt \to 0$ and $N \to \bar{N} = N_{C1} + N_{Sn}$. In the critical case, $\bar{N} = N_{C1} + N_{Sn} = 0$, $N(t)$ approaches 0 at rate λ_-. Instead, when $\bar{N} > 0$, then the asymptotic behaviour of the population growth rate, ω, is given by:

$$\omega(t) \approx \frac{\gamma}{N_{C1} + N_{Sn}} \xi_{G_1} e^{\lambda_- t}. \tag{4.15}$$

Looking at Eq. (4.15), it is apparent that the sign of $\omega(t)$ is determined by the sign of γ. □

Proposition 4.3.1 provides a complete picture of the possible outcomes predicted by the model. Neglecting the critical scenario $\lambda_- = 0$, we recover the three scenarios observed in the numerical simulations: exponential growth ($\lim_{t\to\infty} \omega(t) = \lambda_- > 0$); growth inhibition driven by arrest ($\lim_{t\to\infty} \omega(t) = 0^+$); growth inhibition driven by death ($\lim_{t\to\infty} \omega(t) = 0^-$).

4.3.1 The impact of defective checkpoints on the growth of cell cultures in constant hypoxia

Using the results from Theorem 4.3.1 and Proposition 4.3.1, we can determine how the long time growth dynamics of cell-cultures depend on the values of model parameters (*i.e.*, the extent to which cells have functioning checkpoint regulation of the cell-cycle in hypoxia). The results are illustrated in Fig. 4.3. For different values of the parameters μ_S and q_1, we compute the curves $\chi(\mu_2^C, p_{Sn}) = 1/2$ (see light blue curves) and $\gamma(\mu_2^C, p_{Sn}) = 0$ (see dark blue curves), which partition the (μ_2^C, p_{Sn}) plane into three regions corresponding to the three asymptotic behaviours we identified in Section 4.3. Following the same notation introduced in Chapter 2, different points in the (μ_2^C, p_{Sn}) plane correspond to cell-lines with a different regulation of the G2→M transition in hypoxia. Large values of μ_2^C and small values of p_{Sn} correspond to cells with a defective G2 checkpoint that leads to cell death. By contrast, large values of p_{Sn} and small values of μ_2^C correspond to cells with a partially-functional G2 checkpoint.

Finally, small values of both p_{Sn} and μ_2^C correspond to cells with a functional G2 checkpoint. Figs. 4.3a-4.3d indicate responses of cells with a defective G1 checkpoint (i.e., $q_1 = 0.2$ as for cell-lines 4-6 in Section 2.4.1), while Figs. 4.3e-4.3h study the responses of cells with a functional G1 checkpoint. Moving from left to right along the same row, the value of μ_S increases. The coloured squares indicate the location in (μ_2^C, p_{Sn})-space of the cell-lines (*i.e.*, parameter sets) used to generate Fig. 2.7 in Chapter 2. This shows that the predictions of the asymptotic analysis are in good agreement with the full dynamical simulations. In particular, Figs. 4.3a (cell-lines 4-6 in in Section 2.4.1) and 4.3e (cell-lines 1-3 in Section 2.4.1) are to be compared to the results in Fig. 2.7a, while Figs. 4.3d (cell-lines 4-6 in Section 2.4.1) and 4.3h with the results in Fig. 2.7c (cell-lines 1-3 in Section 2.4.1).

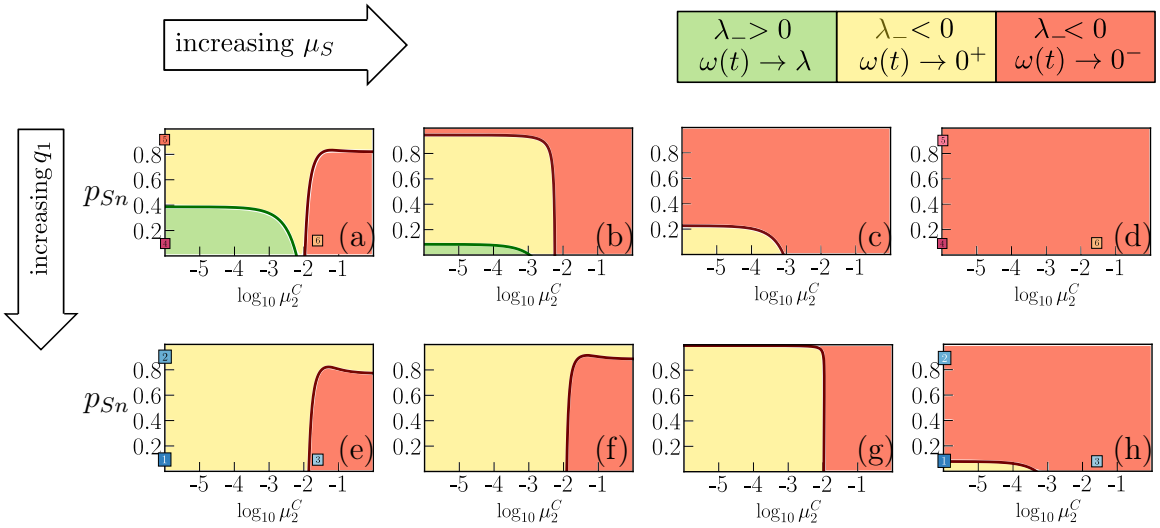


Figure 4.3: Diagrams representing the long time behaviour of the system in the (μ_2^C, p_{Sn}) parameter space for different values of q_1 and μ_S. Moving from left to right, μ_S takes the values $\mu_S = 5\times10^{-4}$, $\mu_S = 2\times10^{-3}$, $\mu_S = 3.6\times10^{-3}$ and $\mu_S = 2.5\times10^{-2}$; while moving from top to bottom, q_1 takes the values $q_1 = 0.2$ (a-d) and $q_1 = 0.9$ (e-h). The green curve is defined by $\lambda(\mu_2^C, p_{Sn}) = 0$ while the dark red curve corresponds to $\gamma(\mu_2^C, p_{Sn}) = 0$ (see Eq. (4.13). The parameter values corresponding to the *in silico* cell-line in Table 2.3 are also indicated using the same colour scheme as in Fig. 2.7.

As expected, the size of the dark red region (corresponding to growth inhibition driven by cell death) increases with μ_S and, for sufficiently large $\mu_S > \mu_S^{(cr)}$, it is the only possible model outcome (see panel (d)). Interestingly, the area of the red region increases in size much faster when increasing μ_S for smaller values of q_1. This is particularly evident when looking at the third column of Fig. 4.3. In this case, given the same values of μ_S, μ_2^C and p_{Sn}, cells with a functional G1 checkpoint (i.e.,

panels (e)-(d)) are less likely to die than cells with a defective G1 checkpoint. In other words, even those cell-lines that have a functional G2 checkpoint (small μ_S and μ_2^C) require effective activation of the G1 checkpoint to prevent cell death at long times. This is consistent with the numerical simulations: proper activation of the G1 checkpoint can protect cells from cell death, which is advantageous in cells that are highly sensitive to hypoxia (*i.e.*, large μ_S). The opposite holds for cells that are less sensitive to hypoxia (i.e., low μ_S). Indeed, escape from the G1 checkpoint allows for sustained proliferation and overall population growth (see green region in Figs. 4.3a and 4.3b), which is not possible if, instead, cells arrest in the G1 phase (*i.e.*, arrest in the C_1 compartment) with high probability (see Fig. 4.3d and 4.3e).

The results in Fig. 4.3 are particularly interesting as they highlight how cells with different sensitivity to hypoxia (*i.e.*, different μ_S) may benefit from different levels of regulation of cell-cycle transitions in hypoxia. Here, as is common in the literature on evolutionary game theory and ecology, benefit (or fitness) can be identified with higher population growth rate (i.e., λ_-). Another interesting model prediction is that homeostasis, (*i.e.*, constant cell numbers), can emerge from different biological mechanisms: either a large rate of cell death or a large probability of cell arrest. While senescent cells lose the ability to proliferate permanently, cells that arrest in the G1 checkpoint (*i.e.*, C_1 compartment) maintain their ability to replicate via the mitotic cell-cycle upon re-oxygenation. The implication of this fact will be studied in the next section where we discuss the long term impact of cyclic hypoxia.

4.4 Asymptotic growth dynamics in cyclic hypoxia

Having discussed the long term effect of constant hypoxia, we now consider the long term effect of cyclic hypoxia. Based on the numerical simulations presented in Section 2.4.2, we expect the asymptotic dynamics predicted by the model will vary greatly with the choice of model parameters associated with cell regulation of the mitotic cycle, and with the cyclic hypoxia protocol considered. From a mathematical point of view, characterising the long term behaviour of Eqs. (2.18) under fluctuating, rather than constant, oxygen levels is more challenging. This is due to the time dependent evolution of the model delays, $\tau_S(t)$ and $\tau_R(t)$, and the transition coefficients between different model compartments. Nevertheless, we will follow the same approach as in the previous section, relying on the theory of strongly positive dynamical system.

For the oxygen dynamics, $c(t)$, we assume that the oxygen levels rapidly switch between their maximum (c_+) and minimum (c_-) values crossing the hypoxic threshold

c_H, *i.e.* $c_- < c_H < c_+$. More specifically, we take $c : [0, \infty) \to [c_-, c_+]$ to be a continuous, $\mathcal{T}$-periodic function (*i.e.*, $c(t) = c(t + \mathcal{T})$) that crosses the hypoxic threshold at times $t_i^{(R)} = \mathcal{T}_- + i\mathcal{T}$ and $t_i^{(H)} = (i+1)\mathcal{T}_+$ and $\mathcal{T} = \mathcal{T}_+ + \mathcal{T}_-$. Here $\mathcal{T}_-$ will correspond to the duration of the hypoxic phase and $\mathcal{T}_+$ to the duration of the re-oxygenation phase between consecutive periods of hypoxia. We refer to these environmental conditions as the $(\mathcal{T}_-, \mathcal{T}_+)$-cyclic protocol. Here we will be interested in comparing with the simulation presented in Chapter 2. Therefore, we use the same formulation as given by Eq. (2.31) in Section 2.4.2. We approximate the solution of Eq. (2.31) under the assumption that $\mathcal{R}_{ox} \gg 1$, *i.e.*, that the oxygen levels equilibrate quickly between the prescribed equilibrium values:

$$c(t) = \begin{cases} (c_+ - c_-)e^{-\mathcal{R}_{ox}\xi(t)} + c_-, & 0 \le \xi(t) < \mathcal{T}_-, \\ (c_- - c_+)e^{-\mathcal{R}_{ox}\xi(t)} + c_+, & \mathcal{T}_- \le \xi(t) < \mathcal{T}, \end{cases} \tag{4.16}$$

where $c_+ = 2.1\%\, O_2$ and $c_- = 0.1\%\, O_2$, and $\xi(t) = t + t_H \mod \mathcal{T}$ where $t_H = \mathcal{R}_{ox}^{-1} \ln((c_+ - c_-)/(c_+ - c_H))$ with $c_H = 1\%\, O_2$. As in Section 2.4.2, we assume that $\mathcal{R}_{ox} \gg 1$, so that t_H is just a small correction; however, this term guarantees that the hypoxia phase ($c < c_H$) starts exactly at time $t = 0$, which will simplify several of the computations in the next sections. As mentioned in Section 2.4.2, under the assumption that $\mathcal{R}_{ox} \gg 1$, the parameters $\mathcal{T}_-$ and $\mathcal{T}_+ = \mathcal{T} - \mathcal{T}_-$ represent good estimates of the time cells spend in hypoxia ($c < c_H$) and in reoxygenation ($c > c_H$), respectively.

Given the periodic nature of the oxygen levels, most the independent variable/coefficient in the model (see Eq. (2.18)), acquire the same periodicity property including: the internal level of proteins ($m(t)$); the velocities at which cells synthesise DNA ($v_S(t)$) and recover from damage/stress ($v_R(t)$); the delays associated with the duration of the S phase ($\tau_S(t)$) and repair in the G2 checkpoint ($\tau_R(t)$). Based on Eqs. (3.2), it is straightforward to show that all these variables converge towards a limit cycle $\hat{\boldsymbol{U}}(t)$ that has the same period, $\mathcal{T}$, as the oxygen fluctuations:

$$\hat{\boldsymbol{U}}(t) = [m^\dagger(t), v_S^\dagger(t), v_R^\dagger(t), \tau_S^\dagger(t), \tau_R^\dagger(t)]. \tag{4.17}$$

For the protein levels $m^\dagger(t)$ and the velocities $v_{S,R}^\dagger(t)$, we can derive an explicit equation, whereas the delays $\tau_{S,R}^\dagger(t)$ are only defined implicitly. These will be discussed in more detail in Section 4.4.3. For now, what is of interest is that the delays $\tau_{S,R}^\dagger$ are continuous and $\mathcal{T}$-periodic functions. As in Section 4.3, we group our variables into three subsets: $\boldsymbol{X}(t) = [G_1(t), C_1(t), G_2(t)]$, $\boldsymbol{Y} = [S(t), D_2(t)]$ and $\boldsymbol{A}(t) = [\mathrm{Sn}_2(t)]$. As shown in the schematic of Fig. 4.4, the variables in $\boldsymbol{X}$ correspond to compartments

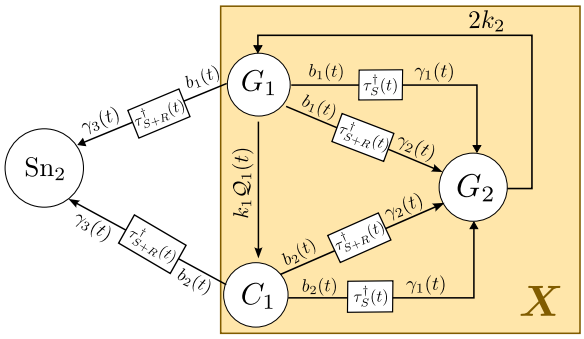


Figure 4.4: Scheme of the model for long time exposure to cyclic hypoxia. Arrows indicate fluxes in between model compartments occurring at the rate indicated. The boxes indicate delay operators.

that have both an influx and an outflux of cells, while those in $\boldsymbol{Y}$ do not. Unlike in constant hypoxia, arrest in the C_1 compartment is reversible upon re-oxygenation. Hence this variable is now included in $\boldsymbol{X}$. The variables $\boldsymbol{Y}$ instead have finite memory and their asymptotic solution can be expressed in terms of other model variables (as in the constant hypoxia case).

By substituting Eqs. (4.17) into Eqs. (2.18), we find that the evolution of $\boldsymbol{X}$ is determined by the following coupled ordinary and delay differential equations with $\mathcal{T}$-periodic coefficients:

$$\begin{aligned}\frac{d\boldsymbol{X}}{dt} = A(t)\boldsymbol{X}(t) &+ \gamma_1(t)\boldsymbol{b}^T(t-\tau_S^\dagger(t))\boldsymbol{P}^T\boldsymbol{X}(t-\tau_S^\dagger(t))\boldsymbol{e}_3 \\ &+ \gamma_2(t)\boldsymbol{b}^T(t-\tau_{S+R}^\dagger(t))\boldsymbol{P}^T\boldsymbol{X}(t-\tau_{S+R}^\dagger(t))\boldsymbol{e}_3,\end{aligned} \qquad t > s \gg 1. \tag{4.18a}$$

In Eqs. (4.18a), $A \in \mathbb{R}^{3\times 3}$ and $\boldsymbol{b}(t) = [b_1(t), b_2(t)] \in \mathbb{R}^{2\times 1}$ are $\mathcal{T}$-periodic matrices that link to the coefficients of the model from Chapter 2 (see Eqs. (2.18l)-(2.18s)) as follows:

$$A(t) = \begin{bmatrix} -k_1 & 0 & 2k_2 \\ k_1\mathcal{Q}_1(t) & -\mathcal{K}_1(t) & 0 \\ 0 & 0 & -k_2 \end{bmatrix}, \quad \boldsymbol{b}(t) = \begin{bmatrix} k_1(1-\mathcal{Q}_1(t)) \\ \mathcal{K}_1(t) \end{bmatrix} \tag{4.18b}$$

$$\gamma_1(t) = \frac{v_S^\dagger(t)}{v_S^\dagger(t-\tau_S^\dagger(t))} e^{-d_S(t)}(1-p_2^\dagger(t)), \tag{4.18c}$$

$$\gamma_2(t) = v_R^\dagger(t)e^{-d_R(t)}\left[\frac{e^{-d_S(\xi)}p_2^\dagger(\xi)(1-p_{Sn}^\dagger(\xi))v_S^\dagger(\xi)}{v_R^\dagger(\xi)v_S^\dagger(t-\tau_S^\dagger(t))}\right]_{\xi=t-\tau_R^\dagger(t)}. \tag{4.18d}$$

In Eqs. (4.18b)-(4.18d), $d_S(t) = \int_{t-\tau_S^\dagger(t)}^t u_S(\xi)d\xi$, $d_R(t) = \int_{t-\tau_R^\dagger(t)}^t \mu_2^C d\xi$, $p_i^\dagger(t) = \mathcal{P}_i(\tau_S^\dagger(t))$ with $i = 2, Sn$. Here $d_{S,R}$ are related to cell death, while $p_2^\dagger$ indicates the time-dependent probability that a cell exiting the S phase at time t enters the G2

checkpoint, and $p^\dagger_{Sn}$ the fraction of these cells that become senescent. In Eqs. (4.18a), the operator T indicates the transpose of a matrix, $\boldsymbol{P} = [\boldsymbol{e}_1\ \boldsymbol{e}_2] \in \mathbb{R}^{4\times 2}$ with $\boldsymbol{e}_i$ being the i-th unit vector of the canonical base in $\mathbb{R}^4$. Furthermore, the time-dependent delays $\tau^\dagger_S(t)$, $\tau^\dagger_R(t)$ and $\tau^\dagger_{S+R}(t) = \tau^\dagger_R(t) - \tau^\dagger_S(t - \tau^\dagger_R(t))$ are known $\mathcal{T}$-periodic continuously differentiable functions that satisfy the condition $d\tau^\dagger_j/dt(t) < 1$ for all t and $j \in \{S, R, S+R\}$. We note that, by definition, we have that $\tau^\dagger_{S+R} > \tau^\dagger_S$. This follows from the fact that $d\tau^\dagger_S/dt < 1$ which implies that the function $\tau^\dagger_S$ is a contraction, *i.e.*, $|\tau^\dagger_S(\xi_1) - \tau^\dagger_S(\xi_2)| \leq L|\xi_1 - \xi_2|$ with $0 < L < 1$ for all $\xi_1, \xi_2 \in \mathbb{R}$. This implies that $\tau^\dagger_S(t) - \tau^\dagger_{S+R} = \tau^\dagger_S(t) - \tau^\dagger_S(t - \tau^\dagger_R(t)) - \tau^\dagger_R(t) < L\tau^\dagger_R(t) - \tau^\dagger_R(t) < 0$ for all times, so that $\tau^\dagger_{S+R} > \tau^\dagger_S$. The evolution of Sn_2 again decouples and satisfies the equation:

$$\frac{d\mathrm{Sn}_2}{dt} = \gamma_3(t)\boldsymbol{b}^T(t - \tau^\dagger_S(t))\boldsymbol{P}^T\boldsymbol{X}(t - \tau^\dagger_S(t)), \tag{4.18e}$$

where $\gamma_3(t) = v^\dagger_S(t)e^{-d_S(t)}p^\dagger_{Sn}(t)p^\dagger_2(t)$, while $\boldsymbol{P}$ and $\boldsymbol{b}$ are as above. Once solved for $\boldsymbol{X}$, we can compute $\boldsymbol{Y}$ by using the following integral equations:

$$S(t) = \int_{t-\tau_S(t)}^{t} \boldsymbol{b}^T(\xi)\boldsymbol{P}^T\boldsymbol{X}(\xi)e^{-(D_S(t)-D_S(\xi))}d\xi, \tag{4.19a}$$

$$D_2(t) = \int_{t-\tau^\dagger_R(t)}^{t} \left[p^\dagger_2(1-p^\dagger_{Sn})\right]_{\tau^\dagger_S(\xi)} \boldsymbol{b}^T(\xi - \tau^\dagger_S(\xi))\boldsymbol{P}^T\boldsymbol{X}(\xi - \tau^\dagger_S(\xi))e^{-d_S(\xi)-(D_R(t)-D_R(\xi))}d\xi, \tag{4.19b}$$

where $D_i : \mathbb{R} \to [0, \infty)$ is such that $d_i(t) = D_i(t) - D_i(t - \tau^\dagger_i(t))$ with $i = S, R$. Here we have written the equations in a slightly more general form; besides helping in the analysis, avoiding specifying the functional form of the time-dependent coefficients in Eqs. (4.18a) allows us to find the general conditions they have to satisfy to yield asynchronous growth.

As in the case of constant hypoxia, in cyclic conditions, the model admits infinitely many equilibrium points that can be parametrised as follows:

$$\hat{\boldsymbol{X}} = (0, 0, 0), \quad \hat{\boldsymbol{Y}} = (0, 0), \quad \hat{\mathrm{Sn}}_2 = N_{Sn}, \tag{4.20}$$

where N_{Sn} is an arbitrary non-negative constant. Based on Eq. (4.20), we see that in cyclic hypoxia, the only non-trivial solutions that the model admits are those for which only senescent cell survive (*i.e.*, no proliferative cells remain). As in the previous section, there are two scenarios to consider: either $\boldsymbol{X} \to \infty$ and so do all other variables (this will be the case of asynchronous growth), or $\boldsymbol{X} \to \boldsymbol{0}$ and $\boldsymbol{Y} \to \boldsymbol{0}$ while Sn_2 converges to a finite value, $N_{Sn} \geq 0$. In other words, the long time behaviour of the system can be characterised by simply looking at the dynamics of $\boldsymbol{X}$, *i.e.*, Eqs. (4.18a).

Without going into the details, based on the assumptions listed above on the regularity of model coefficients and delays, the general theory of retarded functional differential equations guarantees that the model is well-posed [62, 143, 142] and the standard theory of positive dynamical systems introduced in Section 1.2.3 can be applied. While Eqs. (4.18a) are linear, their solution can not be expressed as the superposition of the eigenfunctions $e^{\Lambda t}$ since the equations are not autonomous. Nonetheless, as stated in Lemma 4.4.1, upon a few additional assumptions on the properties of the coefficients in Eqs. (4.18a), we can characterise the asymptotic behaviour of solutions to Eqs. (4.18a). Since the proof of Lemma 4.4.1 is slightly more involved than for previous examples discussed in this thesis, we relegate it to Appendix B.1. This combines the theory of strongly positive linear systems and results from [62, Chapter 8] on the existence of a Floquet representation for periodic differential equations with delays.

Lemma 4.4.1. *Suppose that the following conditions on the coefficients in Eqs. (4.18a)-(4.18d) are satisfied:*

$$k_1, k_2 > 0, \quad \mathcal{Q}_1 : \mathbb{R} \to [0, q_1], \quad \mathcal{K}_1 : \mathbb{R} \to [0, \infty), \quad p^{\dagger}_{2,Sn} : \mathbb{R} \to [0, p_{2,Sn}], \tag{4.21}$$

where $q_1 \in (0,1)$, $p_{2,Sn} \in (0,1)$ are as in Theorem 4.3.1. Suppose further that

$$\max_{t\in[0,\mathcal{T})} \mathcal{Q}_1(t) > 0, \quad \max_{t\in[0,\mathcal{T})} \mathcal{K}_1(t - \tau_S^{\dagger}(t)) > 0, \tag{4.22}$$

Then, for any positive function $\boldsymbol{X}(t)$ satisfying Eq. (4.18a), there exists $c > 0$, such that $\lim_{t\to\infty} \boldsymbol{X}(t) = ce^{\lambda_{\dagger} t}\boldsymbol{\xi}_{\boldsymbol{X}}(t)$, where $\lambda_{\dagger} \in \mathbb{R}$ and $\boldsymbol{\xi}_{\boldsymbol{X}}(t) \gg 0$ is a $\mathcal{T}$-periodic function (i.e., $\boldsymbol{\xi}_{\boldsymbol{X}}(t) = \boldsymbol{\xi}_{\boldsymbol{X}}(t + \mathcal{T})$).

Proof. The proof can be found in Appendix B.1. □

As shown in Lemma 4.4.1, conditions on the model coefficients are needed to characterise the long term behaviour of the system. Most of these conditions follow naturally from the biological system we are modelling, (*e.g.*, non-negativity of parameter values). However, the interpretation of conditions (4.22) is less apparent straightforward. We recall that conditions (4.22) are necessary to prove that the dynamical system is strongly positive, but are not necessary for it to be positive (see proof of Lemma 4.4.1 in Appendix B.1). From a biological point of view, these imply that there is always some level of heterogeneity in how cells from the same population (*i.e.*, cell-line) respond to cyclic hypoxia. Given that our model satisfies the conditions in Lemma 4.4.1 applying Definition 1.2.1, we conclude that the reduced

system for the time-evolution of the variables $\boldsymbol{X}$ has a dominant exponential solution. Using the terminology introduced in Section 4.2, Eqs. (4.18a) have the property of asynchronous quasi-periodic exponential growth when $\lambda_\dagger > 0$. This, however, does not readily imply that the population will follow quasi-periodic exponential growth dynamics, as this will depend also on the asymptotic dynamics of the variables $\boldsymbol{Y}$ and $\boldsymbol{A}$. Our findings are summarised in Theorem 4.4.1, where we leverage the results from Lemma 4.4.1 to study the asymptotic behaviour of the population size $N(t)$. Analogous to the constant hypoxia case, we find that if Eqs. (4.18a) have the property of asynchronous quasi-periodic exponential growth, then the whole system has this same property; in contrast, if the trivial steady state is asymptotically stable ($\lambda_\dagger < 0$) for Eqs. (4.18a) then the growth of the cell culture is arrested.

Theorem 4.4.1. *Under the assumptions of Lemma 4.4.1, and denoting by $\lambda_\dagger$ the dominant Floquet exponent for Eqs. (4.18a), we find that:*

- *if $\lambda_\dagger > 0$, then the population eventually converges to a regime of asynchronous quasi-periodic exponential growth at rate* (i.e., $\lim_{t\to\infty} N(t) = e^{\lambda_\dagger t}\xi_N(t)$ *where* $\xi_N(t) = \xi(t+\mathcal{T})$, *and* $\lim_{t\to\infty} \omega_{\mathcal{T}}(t) = \lambda_\dagger$, *where* $\omega_{\mathcal{T}}(t)$ *defined by Eq. (4.2));*
- *if $\lambda_\dagger = 0$, then the population eventually follows an arithmetic quasi-periodic growth law (*i.e., $\lim_{t\to\infty} N(t) = \bar{N}(t)$, *where* $\bar{N}(t)$ *satisfies* $\bar{N}(t+\mathcal{T}) = \bar{N}(t) + C$*);*
- *if $\lambda_\dagger < 0$, then the system evolves towards a stationary solution,* (i.e., $\lim_{t\to\infty} N = N_{Sn} \geq 0$ *and* $\lim_{t\to\infty} \omega_{\mathcal{T}}(t) = 0$, *where* $\omega_{\mathcal{T}}(t)$ *defined by Eq. (4.2)).*

Proof. When $\lambda_\dagger > 0$, then from Lemma 4.4.1 we know that $\lim_{t\to\infty} \boldsymbol{X}(t) = ce^{\lambda_\dagger t}\boldsymbol{\xi}_{\boldsymbol{X}}(t)$ where $\boldsymbol{\xi}_{\boldsymbol{X}}(t) = \boldsymbol{\xi}_{\boldsymbol{X}}(t+\mathcal{T})$. It follows that also $\lim_{t\to\infty} S(t) = e^{\lambda_\dagger t}\xi_S(t)$, $\lim_{t\to\infty} D_2 = e^{\lambda_\dagger t}\xi_{D_2}(t)$, $\lim_{t\to\infty} \mathrm{Sn}_2 = e^{\lambda_\dagger t}\xi_{\mathrm{Sn}_2}(t)$, where $\xi_i(t)$ with $i \in \{S, D_2, \mathrm{Sn}_2\}$ are $\mathcal{T}$-periodic functions. In other words, all model variables eventually grow according to a quasi-periodic geometric growth law $z(t+\mathcal{T}) = Cz(t)$, where $C = e^{\lambda_\dagger \mathcal{T}} > 0$. Consequently, also their sum (*i.e.*, the total population size, $N(t)$), follows the same asymptotic trend of asynchronous (quasi-periodic) exponential growth.

When $\lambda_\dagger = 0$, Lemma 4.4.1 implies that $\lim_{t\to\infty} \boldsymbol{X}(t) = c\boldsymbol{\xi}_{\boldsymbol{X}}(t)$, where $\boldsymbol{\xi}_{\boldsymbol{X}}(t)$ is a $\mathcal{T}$-periodic function. Using Eq. (4.19), we find that the same holds for the variables $\boldsymbol{Y}(t)$: $\lim_{t\to\infty} S(t) = \xi_S(t)$, $\lim_{t\to\infty} D_2(t) = \xi_{D_2}(t)$, where $\boldsymbol{\xi}_{S,D_2}(t)$ are $\mathcal{T}$-periodic function. Integrating Eq. (4.18e), we find that, for $t \gg 1$, after each period $\mathcal{T}$, Sn_2 increases by the same amount $\Delta_{\mathrm{Sn}_2} \geq 0$, *i.e.*, $\mathrm{Sn}_2(t+\mathcal{T}) = \Delta_{\mathrm{Sn}_2} + \mathrm{Sn}_2(t)$. Estimating $N(t)$ by taking the sum of all model variables, we obtain that $N(t+\mathcal{T}) = C + N(t)$ for $t \gg 1$, where $C = \Delta_{\mathrm{Sn}_2} \geq 0$.

When $\lambda_\dagger < 0$, the steady states in Eq. (4.20) are stable. Lemma 4.4.1 implies that the variables $\boldsymbol{X}$ eventually monotonically decrease toward zero exponentially fast. Again, using Eq. (4.19), we find that the same holds for the variables $\boldsymbol{Y}$. Integrating Eq. (4.18e), we find that $\mathrm{Sn}_2(t+\mathcal{T}) = \mathrm{Sn}_2(t) + O(e^{\lambda_\dagger t})$, which implies $\mathrm{Sn}_2(t)$ monotonically increases towards a constant value that we denote $N_{Sn} \geq 0$. Consequently, we have that $\lim_{t\to\infty} N = N_{Sn}$. □

4.4.1 Numerically approximating the Malthusian parameter

We now estimate the dominant exponent $\lambda_\dagger$ (or Malthusian parameter) to determine under which environmental conditions and choice of model parameters, cells are predicted to maintain their proliferation capacity (*i.e.*, $\lambda_\dagger > 0$), when periodically exposed to hypoxia. Unlike in constant environments, under general cyclic conditions we can not derive a characteristic polynomial that implicitly defines the exponent $\lambda_\dagger$. However, the latter can be approximated numerically. Due to their relevance in engineering applications, several methods have been proposed to approximate the spectrum of DDE systems with time dependent delays, *e.g.*, semi- or full-discretization methods [71]. Here, we use the method developed by Vyasarayani [77, 120, 121], which allow us to derive a finite dimensional approximation of Eqs. (4.18a) via Galerkin approximations. This approach involves three steps. First, Eqs. (4.18a) are converted into a system of partial differential equations, with time periodic coefficients and boundary conditions by introducing the variable $\boldsymbol{x}(t,s) = \boldsymbol{X}(t + s\tau^\dagger_{S+R}(t))$, which satisfies the system:

$$\frac{\partial \boldsymbol{x}}{\partial t} = \frac{1 + \tau^\dagger_{S+R}(t)'s}{\tau^\dagger_{S+R}(t)} \frac{\partial \boldsymbol{x}}{\partial s}, \tag{4.23a}$$

with boundary conditions:

$$\begin{aligned} \left.\frac{\partial \boldsymbol{x}}{\partial t}\right|_{(t,0)} - A(t)\boldsymbol{x}(t,0) - \gamma_1(t)\boldsymbol{b}(t-\tau^\dagger_S(t))\boldsymbol{e}_3\boldsymbol{P}^T\boldsymbol{x}\left(t, -\frac{\tau^\dagger_S(t)}{\tau^\dagger_{S+R}(t)}\right) \\ - \gamma_2(t)\boldsymbol{b}(t-\tau^\dagger_{R+S}(t))\boldsymbol{P}^T\boldsymbol{x}(t,-1)\boldsymbol{e}_3 = 0. \end{aligned} \tag{4.23b}$$

The solution of Eqs. (4.23) is then approximated via a Galerkin approximation that uses as basis functions the Legendre polynomials (η_i):

$$\boldsymbol{x}(t,s) \approx \sum_{i=1}^{n_\eta} \boldsymbol{\beta}_i(t)\eta_i(s) = \boldsymbol{\eta}(s)^T B(t), \quad \boldsymbol{\beta}_i \in \mathbb{R}^4, \quad B \in \mathbb{R}^{n_\eta \times 4}, \tag{4.24}$$

where n_η is an arbitrary chosen constant that determines the accuracy of the approximation. Note that we can retrieve the original variable $\boldsymbol{X} = \boldsymbol{x}(t,0) \approx \sum_i^{n_\eta} \beta_i$. Substituting Eq. (4.24) into Eqs. (4.23) and manipulating the equations, we obtain a system of $4n_\eta$ ordinary differential equations with $\mathcal{T}$-periodic delays in $4n_\eta$ variables of the form:

$$\boldsymbol{B}' = \mathcal{A}_{n_\eta}(t)\boldsymbol{B}, \tag{4.25}$$

where the vector $\boldsymbol{B}(t)$ contains all entries of the matrix $B(t)$ (details of the derivation of Eq. (4.25) are given in Appendix B.2). Given that Eqs. (4.25) are finite dimensional, the standard Floquet theory for differential equations applies. We can compute the Floquet exponents of Eqs. (4.25) in terms of the Floquet transition matrix $U_{n_\eta}(t,s) \in \mathbb{R}^{4n_\eta \times 4n_\eta}$ defined as:

$$\frac{dU_{n_\eta}}{dt} = \mathcal{A}_{n_\eta}(t)U_{n_\eta}, \quad t > s, \tag{4.26a}$$

$$U_{n_\eta}(s,s) = \mathbb{I}_{4n_\eta \times 4n_\eta}, \tag{4.26b}$$

where $\mathbb{I}$ is the identity matrix. In particular, if Λ is a Floquet multiplier of Eqs. (4.25), then $\gamma = e^{\Lambda \mathcal{T}}$ is an eigenvalue of the *period-map* $\psi_{n_\eta}(s) = U_{n_\eta}(s+\mathcal{T}, s)$. It is well-known that the eigenvalues of $\psi_{n_\eta}(s)$ are independent of s, so, without loss of generality, we choose $s = 0$ and we introduce the monodromy operator $\Psi_{n_\eta} = \psi_{n_\eta}(0)$, which is computed solving Eqs. (4.26) in the interval $t \in [0, \mathcal{T}]$. We estimate $\lambda_\dagger \approx \mathcal{T}^{-1} \ln |\gamma_{n_\eta}|$, where γ_{n_η} is the dominant eigenvalue of matrix Ψ_{n_η}. A summary of the procedure used to estimate $\lambda_\dagger$ is given in Algorithm 1. We note that, although we have proved that the dominant Floquet exponent $\lambda_\dagger$ for Eqs. (4.18a) is real (see discussion Section 1.2.3.1) so that $e^{\lambda_\dagger \mathcal{T}}$ is real and positive, we can not conclude the same for the dominant Floquet exponent associated with the finite-dimensional system given by Eqs. (4.25). While investigating the properties of the approximated system goes beyond the scope of this thesis, in all of the examples considered, γ_{n_η} is found to be a positive number, suggesting that the approximated operator Ψ_{n_η} might inherit the property of positivity from Ψ. In Appendix B.3 we show that as we increase n_η, the estimated value of $\lambda_\dagger$ using Galerkin approximation converges to a fixed value that agrees with the growth rate approximated via the full numerical simulation of the model. In what follows, we fix $n_\eta = 15$, which yields estimates of $\lambda_\dagger$ in agreement with the full numerical simulations to three significant figures.

Given a predefined uniform discretization of $(\mathcal{T}_+, \mathcal{T}_-)$-space with $n_+ \times n_-$ squares of size $\Delta\mathcal{T}$ and a choice of model parameters, we obtain an approximation of the map $(\mathcal{T}_-, \mathcal{T}_+) \to \lambda_\dagger(\mathcal{T}_-, \mathcal{T}_+)$ by using the procedure outlined in Algorithm 1. As

Algorithm 1 Estimate the growth rate $\lambda_\dagger$

function ODESOLVER($F(t,x)$, x_0, $[a,b]$)
    # solves the problem $x'(t) = F(t,x)$ with $x(a) = x_0$
    based on `DifferentialEquations.jl` package to solve ode problems in `Julia`
return $x(\xi)$ with $\xi \in [a,b]$
end function
function EIGS(A)
    # computes the dominant eigenvalue of the matrix A
    based on `LinearAlgebra.jl` package in `Julia`
end function
function GENERATOR(t, n_η, $\boldsymbol{\Theta}$)
    # returns the approximated operator $\mathcal{A}_{n_\eta}(t)$ (see Eqs. (4.25)) given the values of model parameters $\boldsymbol{\Theta}$
end function
Input: $\mathcal{T}_+$,$\mathcal{T}_-$, n_η, $\boldsymbol{\Theta}$ (model parameters), ODESOLVER, EIGS, GENERATOR
Output: $\lambda_\dagger$
procedure SOLVE $\lambda_\dagger$
    $\mathcal{T} = \mathcal{T}_+ + \mathcal{T}_-$
    $\Psi_{n_\eta} = \mathbf{0}_{4n_\eta \times 4n_\eta}$
    for $i = 1, \ldots 4n_\eta$ **do**
        $\mathcal{A}_{n_\eta}(\cdot)$ = GENERATOR$(\cdot, n_\eta, \boldsymbol{\Theta})$
        $\boldsymbol{B}$ =ODESOLVER$(\mathcal{A}_{n_\eta}(\cdot)x,\ \boldsymbol{e}_i,\ (0,\mathcal{T}])$
        $\Psi_{n_\eta}[:,i] = \boldsymbol{B}(\mathcal{T})$.
    end for
    γ_{n_η} =EIGS(Ψ)
    $\lambda_\dagger = \frac{\ln|\gamma_{n_\eta}|}{\mathcal{T}}$
end procedure

in Chapter 2 (see Section 2.2.7.1), we compare model predictions for our cohort of virtual cell-lines that vary in how their cell-cycle checkpoints function. Parameter values corresponding to the different cell-lines are listed in Table 2.3.

We recall that cells are predicted to reach a regime of asynchronous (quasi-periodic) exponential growth in cyclic hypoxia only if $\lambda_\dagger(\mathcal{T}_-, \mathcal{T}_+) > 0$. In this regime, larger values of $\lambda_\dagger$ correspond to cell cultures that, asymptotically, grow at faster rates. Therefore we say that a cell-line in our virtual cohort has a growth advantage in a given $(\mathcal{T}_-, \mathcal{T}_+)$-cyclic environment if it has the largest growth rate $\lambda_\dagger(\mathcal{T}_-, \mathcal{T}_+)$. If $\lambda_\dagger(\mathcal{T}_-, \mathcal{T}_+) < 0$, then the model predicts that the cells will eventually lose ability to proliferate and the culture size will reach a homeostatic value. In this regime, smaller values of $\lambda_\dagger$ imply that cells in the culture arrest growth more quickly, *i.e.*, $\omega_\mathcal{T}(t) \to 0$ at a faster rate ($\omega_\mathcal{T}(t)$ defined as in Eq. (4.2)). However, in contrast to the

case of constant hypoxia, we do not know whether the population size N is decaying ($\omega_{\mathcal{T}} \to 0^-$) or growing ($\omega_{\mathcal{T}} \to 0^+$) towards it homeostatic size.

4.4.2 The impact of defective checkpoint on the growth of cell cultures in cyclic hypoxia

The estimated values of $\lambda_\dagger$ for our cohort of *in silico* cell-lines are illustrated in Fig. 4.5. In producing Fig. 4.5, we have set $\mu_S = 0.025$, which implies that, all the cell-lines considered are predicted to arrest their growth in conditions of constant hypoxia (see Fig. 4.3). In an oxygen-rich environment, all cell-lines are instead predict to eventually grow exponentially at the same growth-rate $\lambda_+ \approx 0.033$ [1/hr] (see Section 2.3.2). We find that their behaviours can be substantially different when they are exposed to cyclic hypoxia.

As expected from the numerical simulations presented in Section 2.4.2, the estimated asymptotic growth dynamics are highly heterogeneous with respect to the responses of the different cell-lines and the different cyclic-hypoxia protocols. In all panels, we find that, when $\mathcal{T}_+ \gg \mathcal{T}_-$, the growth rate approaches its upper limit, λ_+, corresponding to population growth rate in oxygen-rich conditions. Hence the model predicts that the growth dynamics of all cell-lines are unperturbed when they are seldom exposed to brief periods of hypoxia. When considering cyclic hypoxia protocols for which $\mathcal{T}_+ > \mathcal{T}_-$, all cell-lines maintain the property of asynchronous exponential growth. However, differences start to appear when we compare the responses of cell-lines with different S→M transition in hypoxia. Cells with functional S/G2 checkpoints (see third column) experience a sharp decay in their population growth rate, which is predicted to be independent of whether they have a functional G1 checkpoint. In contrast, cells that have defective S/G2 checkpoints (see first column), do not exhibit a sharp decay in their proliferation rate. This suggests that in this regime of cyclic hypoxia, having a partially-functional S/G2 checkpoint gives cells a growth advantage. When comparing panels in the first column, differences in the G1 checkpoint can also have a significant impact on the predicted growth dynamics; specifically, cells with a partially-functional G1 checkpoint but defective S/G2 checkpoints (see panel (d)) are the most proliferative, with a proliferation rate that remains close to the upper limit λ_+. When $\mathcal{T}_- \sim \mathcal{T}_+$, cyclic hypoxia causes a severe reduction in the growth rate of all cell-lines. In such scenarios, cells with partially-functional, and not those with defective, S/G2 checkpoints give rise to the fastest proliferating populations.

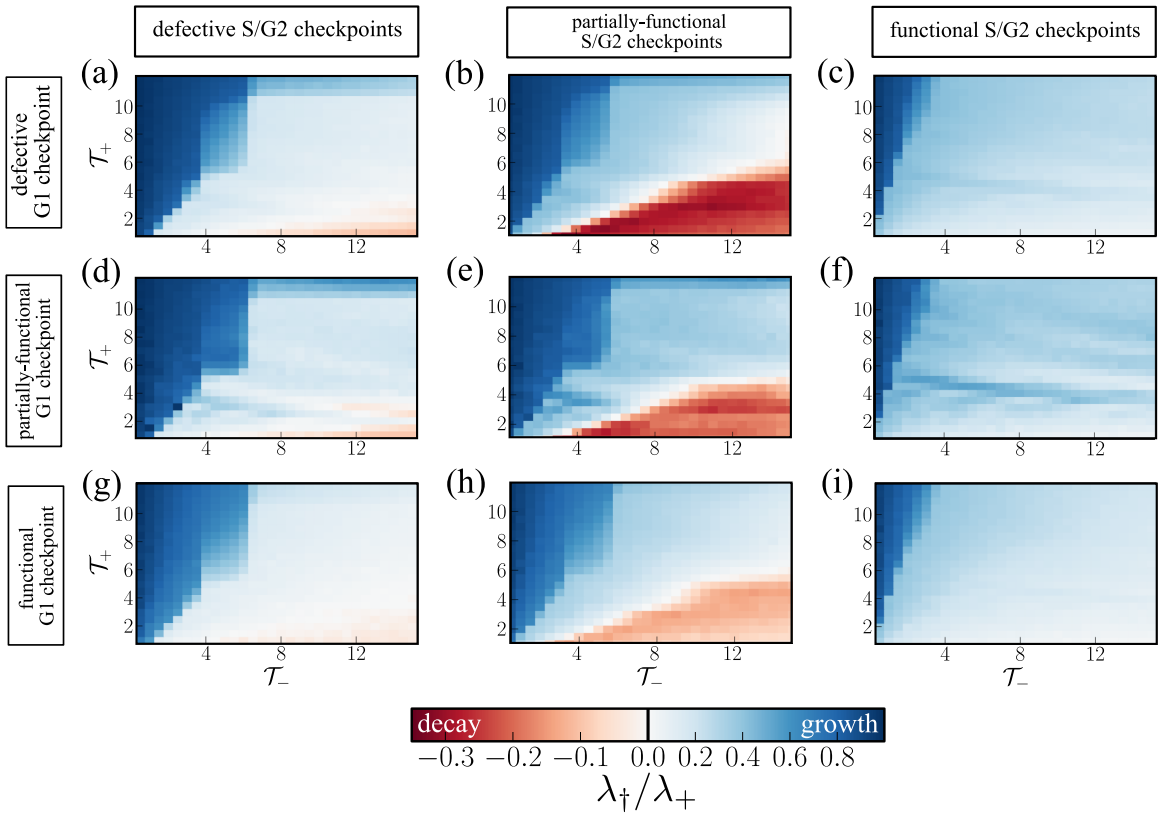


Figure 4.5: (a)-(i) Heat-maps illustrating how the asymptotic growth rate $\lambda_\dagger$ depends on the specifics of the oxygen protocol used ($\mathcal{T}_+$, $\mathcal{T}_-$). Growth rates are rescaled by $\lambda_+ \approx 0.033$ [1/hr], which is the growth rate in an oxygen rich environment. Details on the numerical method used to estimate $\lambda_\dagger$ are given in Section 4.4.1. Each row presents estimates of $\lambda_\dagger$ for cells with defective (left column), partially-functional (centred column) and functional (right column) S/G2 checkpoints. Each column presents estimates of $\lambda_\dagger$ for cells with defective (first row), partially-functional (middle row) and functional (bottom row) G1 checkpoints. Parameter values are as specified in Table 2.3 (except $\mu_S = 0.025$).

When considering cyclic protocols for which $\mathcal{T}_- \gg \mathcal{T}_+$ (*i.e.*, the bottom right corner of the ($\mathcal{T}_-$,$\mathcal{T}_+$) space), we find that the ability of cells to maintain their proliferative capacity is strongly dependent on the cell-cycle control mechanisms they adopt. For example, cyclic protocols with large $\mathcal{T}_-$ and small $\mathcal{T}_+$ result in complete proliferation arrest (*i.e.*, $\lambda_\dagger < 0$) in cells that have either partially-functional or defective S/G2 checkpoints (first column), but not in cells with functional S/G2 checkpoints (middle column). This is in line with the results from Chapter 2, where we found that while having defective S/G2 checkpoints was favourable in cyclic protocols where exposure to hypoxia was very short, having functional S/G2 checkpoints was beneficial for protocols with longer periods of hypoxia.

Overall, the results presented in this section suggest that the toxicity of cyclic hypoxia (*i.e*, the extent to which it impairs cell proliferation) varies between cell-lines and also depends on the specific dynamics of cyclic hypoxia; specifically how often, and for how long, cells are exposed to hypoxia. In general, the $(\mathcal{T}_-, \mathcal{T}_+)$-space can be partitioned into regions in which cyclic hypoxia has significantly different impacts on cell culture growth: low toxicity (dark blue regions –growth dynamics are unperturbed compared to an oxygen-rich environment); medium toxicity (light blue regions –growth is severely impaired); high toxicity (red regions –complete growth arrest). In the next section, we will explore the biological mechanisms that determine the boundaries of these regions. We will derive simple relations between model parameters and investigate the extent to which cyclic hypoxia is toxic for a cell with specific cell-cycle control mechanisms.

4.4.3 Cyclic hypoxia as multiple environmental conditions

In the previous section, we employed numerical techniques to explore how cyclic hypoxia affects the proliferative capacity of cells. While the results are insightful, the complexity of the equations precludes an intuitive understanding of how the different biological mechanisms included in the model contribute to the cells asymptotic growth dynamics under cyclic hypoxia. In this section, we derive more intuitive explanations to understand relationship between the oxygen dynamics (*i.e.*, the values of $\mathcal{T}_+$ and $\mathcal{T}_-$) and the extent to which this impairs the proliferative capacity of different cell-lines.

Recall that our model accounts for five key mechanisms: (1) delay in the S→G2 transition (associated with the velocity $v_S^\dagger$); (2) delay of the G1→S transition (associated with the parameters $\mathcal{Q}_1$ and $\mathcal{K}_1$); (3) delay of the G2→M transition (associated with $\mathcal{P}_2$); (4) premature senescence (associated with the transition rate $\mathcal{P}_{Sn}$ and $\mathcal{P}_2$); (5) fork collapse (associated with the coefficient u_S, see Eq. (2.3c), and the level of repair proteins $m^\dagger$). While, under our modelling assumptions, the dynamics of mechanisms (1) and (2) are directly affected by oxygen-levels, this is not the case for mechanisms (3)-(5) which depend on the current and past oxygen environments cells are and have been exposed to. From this point view, mechanisms (1)-(2) are always activated in cyclic hypoxia but they vary smoothly with $\mathcal{T}_-$ and $\mathcal{T}_+$ in the extent to which they affect proliferation. In particular, the larger $\mathcal{T}_-$, the longer the delays in the S→G2 and G1→S transitions, while the opposite holds when increasing $\mathcal{T}_+$. Mechanisms (3)-(5) have less straightforward impact on cell proliferation due to the

highly non-linear relation between the oxygen levels, $c(t)$, and the values of model coefficients associated with these mechanisms.

We first focus on the effect of cyclic hypoxia on cell arrest in the G2 phase by deriving estimates for the dynamics of the probability $p_2^\dagger(t)$. We then consider the extent to which cyclic hypoxia affects the viability of cells in the G2 phase by looking at the dynamics of the probability $p_{2\to X}^\dagger(t)$ that a cell never completes the G2 phase. Next we discuss how different cyclic hypoxia protocols impact fork stability, or equivalently, the ability of cells to restart DNA synthesis after exposure to hypoxia. This requires us to study the asymptotic evolution of the metric $T_{FC}(t)$ which measures the extent to which cells in the S phase are sensitive to fork collapse.

4.4.3.1 Cell-fate in the G2 phase

According to our model, delays in the G2→M transition are due to cells needing time to repair damage. When this is not possible, damaged cells lose their viability and never exit the G2 phase. We introduce the following two metrics to estimate the extent to which these two mechanisms influence cell responses to cyclic hypoxia:

$$p_2^\dagger(t) = \mathcal{P}_2(\tau_S^\dagger(t)) = \mathbb{P}\begin{pmatrix}\text{a cell exiting the S phase at time } t\\ \text{arrests in G2}\end{pmatrix}, \tag{4.27a}$$

$$\begin{aligned} p_{2\to X}^\dagger(t) &= \mathcal{P}_2(\tau_S^\dagger(t))\mathcal{P}_{Sn}(\tau_S^\dagger(t)) + \mathcal{P}_2(\tau_S^\dagger(t))[1-\mathcal{P}_{Sn}(\tau_S^\dagger(t))]e^{-\mu_2^C\tau_R^\dagger(t)} \\ &= \mathbb{P}\begin{pmatrix}\text{a cell exiting the S phase at time } t\\ \text{will lose its viability}\end{pmatrix}. \end{aligned} \tag{4.27b}$$

We note that Eq. (4.27b) is analogous to the first factor in the definition of the critical parameter χ (see Eq. (4.10)) that determined whether the model presented exhibits asynchronous exponential growth in constant hypoxia.

In order to study the dynamics of $p_2^\dagger$ and $p_{2\to X}^\dagger$, we first derive an expression for the asymptotic delays $\tau_S^\dagger$ and $\tau_R^\dagger$ in cyclic hypoxia characterising, respectively, the dynamics of DNA synthesis and damage/stress recovery.

DNA synthesis. Based on the prescribed evolution of the oxygen levels in Eq. (4.16), we can easily show that the velocity v_S asymptotes to a limit-cycle which is synchronised with the oxygen fluctuations. Starting by characterising DNA synthesis, we have that

$$v_S^\dagger(t) = \begin{cases} v_S^- - (v_S^- - \hat{V}_S^+)e^{-\mathcal{R}_S^-\xi(t)}, & 0 \le \xi(t) < \mathcal{T}_- \\ v_S^+ - (v_S^+ - \hat{V}_S^-)e^{-\mathcal{R}_S^+(\xi(t)-\mathcal{T}_-)}, & \mathcal{T}_- \le \xi(t) < \mathcal{T} \end{cases}, \tag{4.28a}$$

where $\xi(t) = t \mod \mathcal{T}$, $\hat{V}_S^-$ and $\hat{V}_S^+$ respectively indicate the minimum and maximum value that $v_S^\dagger$ takes during a cycle. These values are attained, respectively, at the end and beginning of the hypoxic phase and are defined as follows:

$$\hat{V}_S^- = \frac{\gamma_S^+ v_S^-(\gamma_S^- - 1) + v_S^+(\gamma_S^+ - 1)}{\gamma_S^+\gamma_S^- - 1},\ \hat{V}_S^+ = \frac{v_S^-(\gamma_S^- - 1) + v_S^+\gamma_S^-(\gamma_S^+ - 1)}{\gamma_S^+\gamma_S^- - 1}. \quad (4.28b)$$

In Eqs.(4.28), $\gamma_S^+ = e^{\mathcal{R}_S^+\mathcal{T}_+} > 1$, $\gamma_S^- = e^{\mathcal{R}_S^-\mathcal{T}_-} > 1$. Note that since $v_S^+ > v_S^-$ we always have that $\hat{v}_S^+ > \hat{v}_S^- > 0$. Starting from Eqs. (4.28), we can compute the average velocity of DNA synthesis over an oxygen cycle (of period $\mathcal{T}$), which is defined as:

$$\bar{v}_S^\dagger = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} v_S^\dagger(\xi) d\xi. \quad (4.29)$$

Eq. (4.29) allows us to decompose the asymptotic dynamics of the duration of the S phase, $\tau_S^\dagger(t)$, into two contributions:

$$\tau_S^\dagger(t) = \mathcal{T} n_S + \tau_x\left(\delta_x, \xi(t)\right), \quad n_S = \left\lfloor \frac{1}{\mathcal{T}\bar{v}_S^\dagger} \right\rfloor, \quad (4.30a)$$

$$\delta_x = 1 - \bar{v}_S^\dagger n_S \mathcal{T}, \quad (4.30b)$$

where ξ is again defined as $\xi(t) = t \mod \mathcal{T}$ and $n_S \in \mathbb{N}_0$. In Eq. (4.30a), the first term indicates the constant part of $\tau_S^\dagger$ while the second term $0 \leq \tau_x < \mathcal{T}$ is implicitly defined by Eq. (2.8) and represents the variable component of the delay, which is not a multiple of the period $\mathcal{T}$. Generally, the delay $\tau_S^\dagger(t)$ will vary greatly over a period $\mathcal{T}$, however, there are critical choices of the cyclic protocols for which $\bar{v}_S^\dagger$ is a multiple of the period $\mathcal{T}$; as a result δ_x is zero and the asymptotic delay becomes constant and an exact multiple of the period ($\tau_S^\dagger(t) \equiv n_S\mathcal{T}$).

Recovery dynamics in the G2 checkpoint. In order to study $\tau_R^\dagger(t)$, we must first solve for the asymptotic behaviour of $m(t) \to m^\dagger(t)$. This follows the same behaviour as $v_S^\dagger$ (see Eq. (4.28) upon appropriate update of the parameters ($\mathcal{R}_S^{+,-} \to \mathcal{R}_M^{+,-}$, $v_S^+ \to 1$ and $v_S^- \to M^-$)). Given an expression for $m^\dagger(t)$, it is straightforward to derive $v_R^\dagger$ using Eq. (2.11a), which can be used to estimate the average velocity ($\bar{v}_R^\dagger$); since $v_R^\dagger$ is not an exponential function, like $v_S^\dagger$, but a sigmoid, the algebra is slightly more complicated but we can still obtain an explicit formula for the mean repair velocity. The delay $\tau_R^\dagger$ is then defined analogously to $\tau_S^\dagger$ (see Eq. (4.30)) as

$$\tau_R^\dagger(t) = \mathcal{T} n_R + \tau_y\left(\delta_y, \xi(t)\right), \quad n_R = \left\lfloor \frac{y_R}{\bar{v}_R^\dagger \mathcal{T}} \right\rfloor \quad (4.31a)$$

$$\delta_y = y_R - \bar{v}_R^{\dagger} n_R \mathcal{T}. \tag{4.31b}$$

As the velocity $v_R^{\dagger}$ in cyclic hypoxia is always in the range (v_R^-, v_R^+) (where v_R^- is defined by Eq. (4.4)), the delay will always be in the range $\tau_R^- < \tau_R^{\dagger}(t) < \tau_R^+$. Again there are critical choices of $\mathcal{T}_+$ and $\mathcal{T}_-$ for which $\tau_R^{\dagger}(t)$ is constant and an exact multiple of the period of the fluctuation. For a general choice of parameters, it is unlikely that both $\tau_S^{\dagger}$ and $\tau_R^{\dagger}$ will be synchronised with the oxygen fluctuations.

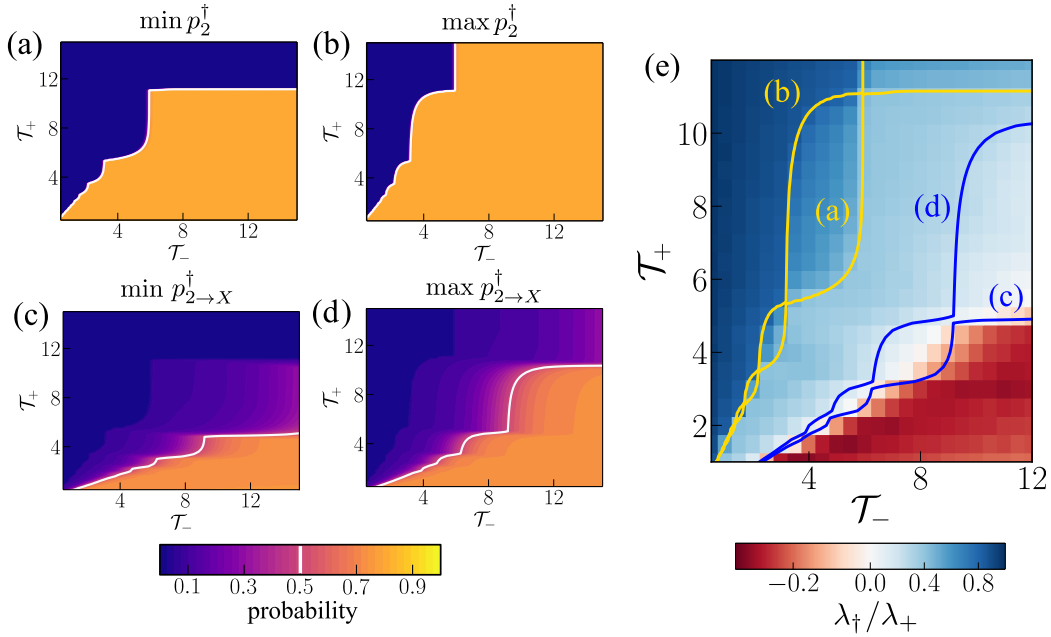


Figure 4.6: (a)-(b) We plot how the minimum (a) and maximum (b) probability of arrest $p_2^{\dagger}(t)$ (see Eq. (4.27a)) vary with the prescribed oxygen dynamics (*i.e.*, $\mathcal{T}_+$ and $\mathcal{T}_-$). (c)-(d) We plot how the minimum (c) and maximum (d) probability of premature senescence $p_{2\to X}^{\dagger}(t)$ (see Eq. (4.27b)) vary with $\mathcal{T}_+$ and $\mathcal{T}_-$. (a)-(d) The white curves indicate the level set 0.5. (e) The same plot as Fig. 4.5c where the yellow and blue curves correspond, respectively, to the isoclines from panels (a)-(b) and (c)-(d). All figures are generated using the same values of the parameters corresponding to the cell-line (x/-) in Table 2.3.

Results. While we can estimate the mean velocity, $\bar{v}_{R,S}^{\dagger}$, explicitly, we can only derive an implicit definition for $\tau_{R,S}^{\dagger}$ and, therefore, also for the probabilities $p_z^{\dagger}$ with $z \in \{2, 2 \to X\}$. In Fig. 4.6 we illustrate the predicted minimum and maximum values of $p_z^{\dagger}$ as a function of the cyclic hypoxia protocol considered for a cell-line with partially-functional S/G2 checkpoints. In Appendix B.3 we comment on how estimates of $\tau_{R,S}^{\dagger}$ are obtained to produce Fig. 4.6. In Figs. 4.6a-4.6d, we indicate the level sets corresponding to the plotted probabilities being one half.

Focusing first on Figs. 4.6a-4.6b, we show how the minimum and maximum of the function $p_2^\dagger(t)$ changes with $\mathcal{T}_-$ and $\mathcal{T}_+$. When $\min p_2^\dagger > 0.5$, we expect prolonged delays in the G2→M transition, as cells are always more likely to be damaged than being undamaged regardless of when they enter the G2 phase. By contrast, if $\max p_2^\dagger < 0.5$, cells are always less likely to arrest than not arrest in the G2 phase as the level of replication stress they experienced in the S phase is mild. Between these two extrema, $p_2^\dagger$ fluctuates between regimes where cells $p_2^\dagger \approx 0$ and $p_2^\dagger \approx p_2$ as the level of replication stress caused by cyclic hypoxia largely depends on when cells initiate DNA synthesis. For sufficiently large $\mathcal{T}_-$, the white curve in Fig. 4.6a ($\min p_2^\dagger = 0.5$) asymptotes horizontally to a constant value of $\mathcal{T}_+^{(cr)}$. This implies that regardless of the length of the hypoxia phase, if $\mathcal{T}_+ > \mathcal{T}_+^{(cr)}$, then there will a be period where cells exiting the S phase will be more likely not to arrest in G2 phase. Conversely, focusing on the level set indicated in Fig. 4.6b ($\max p_2^\dagger = 0.5$), this vertically asymptotes to a critical value $\mathcal{T}_-^{(cr)}$. This implies that when $\mathcal{T}_- > \mathcal{T}_-^{(cr)}$ there are always periods where cells are highly-likely to arrest in the G2 phase. When superimposing the two level sets in Figs. 4.6a-4.6b on the recorded values of $\lambda_\dagger$ for the (x/-) cell-line (see Fig. 4.6e), we see that these almost exactly define the boundaries of the transition from the regimes $\lambda_\dagger \approx \lambda_+$ and $0 < \lambda_c yc \ll \lambda_+$.

Focusing now on Figs. 4.6c-4.6d, we can characterise the extent to which cells are likely to become senescent due to exposure to hypoxia. To this end, values of $\min p_{2\to X}^\dagger > 0.5$ imply cells are more likely to lose their clonogenic capacity upon exiting the S phase than to divide. In contrast, if $\max p_{2\to X}^\dagger < 0.5$, cells are more likely to divide than to lose their proliferative capacity due to cyclic hypoxia induced senescence. Between these two extreme regimes, the risk of cells losing their proliferative capacity is highly dependent on when cells exit the S phase. This indicates that the fate of cells is largely dependent on when they transition into the S phase, as this will affect the amount of damage they accumulate due to cyclic hypoxia. As before, we can superimpose the level sets in Figs. 4.6c-4.6d on the estimated rates $\lambda_\dagger$ for the same (x/-) cell line (see Fig. 4.6e). By doing so, we find that the curve $\min p_{2\to X}^\dagger = 0.5$ delimits the region of the $(\mathcal{T}_-, \mathcal{T}_+)$-space where $\lambda_\dagger$ is significantly negative corresponding to cyclic hypoxia environments that are highly toxic for cells. The region enclosed between the two level sets indicates, instead, where the transition between positive (cells growing exponentially) and negative (growth arrest) values of $\lambda_\dagger$ occurs.

As we will discuss in Section 4.5 the level sets in Fig. 4.6 depend on the properties of the cell-line under consideration (*i.e.*, choice of parameter values), in particular

whether or not they have function S/G2 checkpoint.

Sensitivity to fork collapse We now discuss an approach to quantify the extent to which cells in the S phase are sensitive to fork collapse in different cyclic hypoxia conditions. To this end, we recall that fork stability is lost when the expression levels of repair proteins, $m^{\dagger}(t)$, drop below the threshold M_{FC} (see Eq. (2.3c)). Therefore we propose the following metric to estimate the impact of fork collapse on the asymptotic growth dynamics of cells in cyclic hypoxia:

$$\mathcal{T}_{FC} = \int_0^{\mathcal{T}} I_{[0,M_{FC}]}(m^{\dagger}(t))\, dt, \tag{4.32}$$

where $I_A(\cdot)$ is again the indicator function for the set A and the evolution of $m^{\dagger}(t)$ is as defined in page 126 (see text above Eq. (4.31). Here $\mathcal{T}_{FC}$ measures the fraction of the period $\mathcal{T}$ associated with the oxygen fluctuations where cells are sensitive to fork collapse.

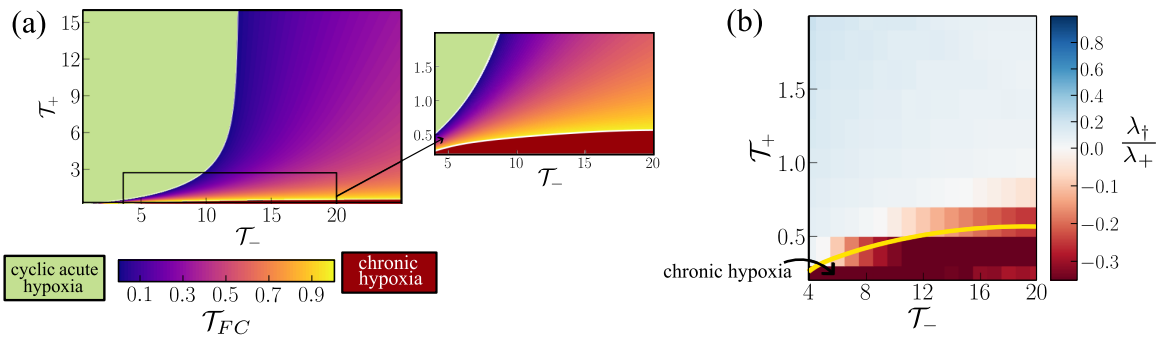


Figure 4.7: (a) We plot the time cells are sensitive to fork collapse, $\mathcal{T}_{FC}$ (see Eq. (4.32)), per unit period, $\mathcal{T}$, in the $(\mathcal{T}_-, \mathcal{T}_+)$ space. Parameter values are as in Table 2.1. We use $\mathcal{T}_{FC}$ to split the space $(\mathcal{T}_-, \mathcal{T}_+)$ into three regions corresponding to significantly different cellular responses: chronic hypoxia ($\mathcal{T}_{FC} = \mathcal{T}$); cyclic acute hypoxia ($\mathcal{T}_{FC} = 0$); cyclic chronic hypoxia ($0 < \mathcal{T}_{FC} = \mathcal{T}$). The small inset zooms in the region of chronic hypoxia. (b) We superimpose the boundary of the region $\mathcal{T}_{FC} = 1$ on the estimated value of $\lambda_\dagger$ for the (x/+) cell-line (we use the same parameters as Fig. 4.5b but different range of values for $\mathcal{T}_+$ and $\mathcal{T}_-$).

In Fig. 4.7a, we show how $\mathcal{T}_{FC}$ depends on the oxygen protocol considered. Since all cell-lines have the same response to fork collapse, $\mathcal{T}_{FC}$ follows the same profile for all of them. As expected, $\mathcal{T}_{FC}$ increases with the duration of the hypoxia phase $\mathcal{T}_-$ and decreases as the duration of the re-oxygenation phase, $\mathcal{T}_+$, increases. Based on the value of $\mathcal{T}_{FC}$, we find a natural subdivision of the $(\mathcal{T}_-, \mathcal{T}_+)$ space into regions where cyclic hypoxia results in cells having different sensitivity fork collapse. In

the red regions, where $\mathcal{T}_+ \ll \mathcal{T}_-$, even though cells are re-oxygenated, they are not able to recover and they remain sensitive to fork collapse. Under these environmental conditions, cell responses are predicted to be similar to those predicted under constant hypoxia: almost complete inhibition of DNA synthesis and a large rate of cell death in the S phase. On the contrary, when $\mathcal{T}_+ \gg \mathcal{T}_-$, cells in the S phase are never sensitive to fork collapse and, despite being exposed to replication stress, they maintain their ability to restart DNA synthesis. Between these two extreme regimes, cells alternate between highly and mildly toxic environments (during which cells in the S phase are sensitive to fork collapse). We denote the latter regime as cyclic chronic hypoxia to indicate that periodically cells behave as if they were exposed to chronic hypoxia. This differs from constant hypoxia in which the oxygen levels remain below the hypoxic threshold c_H at all times. In Fig. 4.7b, we superimpose the level set $\mathcal{T}_{FC} = 0.5$ in Fig. 4.7a the predicted value of $\lambda_\dagger$ for (x/+) cells, *i.e.*, cells that have defective G1 checkpoint but have fully functional S/G2 checkpoints (same parameter values as Fig. 4.5i). We note that, since the ranges of cyclic protocols used to generate Fig. 4.5b did not include any example of chronic hypoxia, in Fig. 4.7b we have considered a very different range of values of $\mathcal{T}_+$ and $\mathcal{T}_-$. Here we chose a cell-line with functional regulation of the S→M transition since, for this cell-line death is primarily driven by fork collapse, since functioning of the S and G2 checkpoints prevents lose of cell-viability due to other form of damage. We find the chronic hypoxia region almost exactly predicts cyclic hypoxia protocols that yield a negative value of $\lambda_\dagger$ indicating full growth arrest.

4.5 Discussion

In this chapter we leveraged the theory of strongly positive dynamical systems and asymptotic properties of systems of differential equations with delays to characterise the long term impact of hypoxia on the growth dynamics of cell populations. In particular, we investigated conditions under which cell cultures maintain their ability to generate exponentially growing "tumours" in different hypoxic environments. In Section 4.2, we reviewed the concept of asynchronous exponential growth and how it is used to characterise the growth dynamics of structured populations in constant environments via definition of the growth rate (or *Malthusian* parameter). We also showed how this concept extends naturally to periodic environments by revising the definition of the *Malthusian* parameter.

In Section 4.3 we characterised the long term effect of hypoxia on the predicted growth dynamics. We found that the ability of cells to proliferate is dominated by the sensitivity of cells in the S phase to fork collapse, as captured by the value of the death rate μ_S. Only in regimes where μ_S is very small, can cells maintain their ability to proliferate in hypoxia, but this will be dictated by how they regulate the S→M transition.

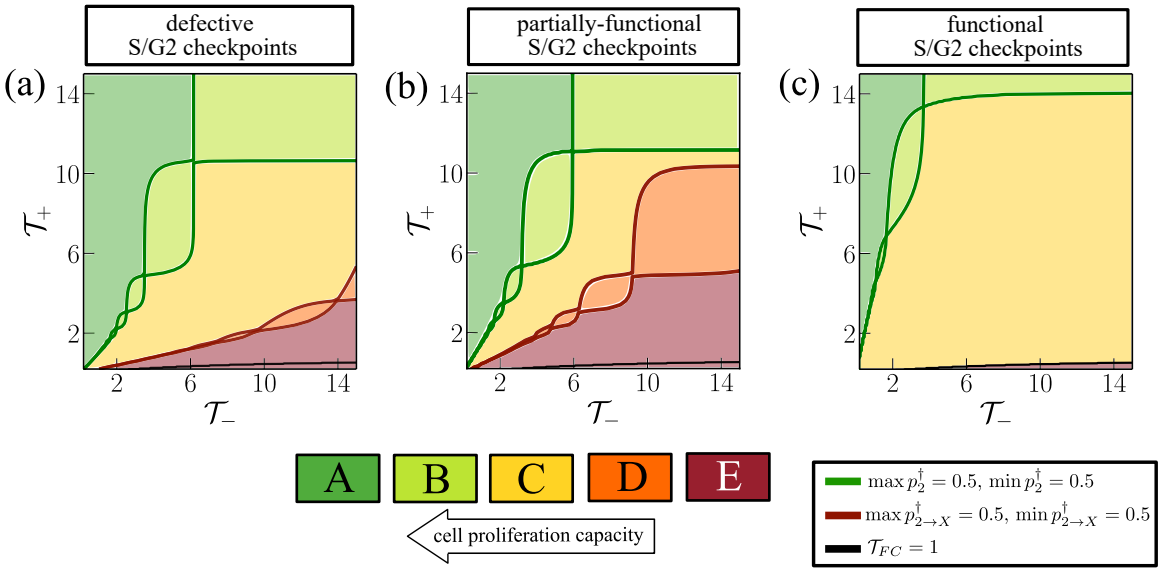


Figure 4.8: Plots summarising the possible cyclic hypoxia regimes predicted by our model for with (a) defective, (b) partially-functional and (c) functional S/G2 checkpoints. Parameter values are the same as in Chapter 2. The green curves indicate the level sets $\max p_2^\dagger = 0.5$ and $\min p_2^\dagger = 0.5$ (see Fig. 4.6) where $p_2^\dagger$ is defined as in Eq. (4.27a). The red curves indicate the level sets $\max p_{2\to X}^\dagger = 0.5$ and $\min p_{2\to X}^\dagger = 0.5$ (see Fig. 4.6) where $p_{2\to X}^\dagger$ is defined as in Eq. (4.27b). The black curve, instead, indicates the level sets $\mathcal{T}_{FC} = 1$, where $\mathcal{T}_{FC}$ is defined by Eq. (4.32).

Even when considering cell populations that behave similarly in constant hypoxia, we find that their long term response to cyclic hypoxia can be highly heterogeneous. This is because the mechanisms that determine the proliferative capacity of cells in constant and cyclic hypoxia are different. Unlike in constant hypoxia, we must rely on numerical methods to study the asymptotic predictions of the model in cyclic hypoxia. In general, we find that the ability of cells to maintain their proliferative capacity is largely impacted by whether their S/G2 checkpoints are functional, while the regulation of the dynamics of the G1 checkpoint has less of an impact, and is only evident for specific cyclic hypoxia protocols. Guided by the numerical results, in Section 4.4.3 we propose simple metric that allow us to readily relate model parameters (*i.e.*, how a

cell regulates its progression along the mitotic cell-cyclic) to the asymptotic behaviour predicted by the model (*i.e.*, cell culture growth dynamics in cyclic hypoxia). Using the results from Section 4.4.3, we partitioned the $(\mathcal{T}_-, \mathcal{T}_+)$-space into regions where cyclic hypoxia has similar impact on the growth of cell cultures. In Fig. 4.8 we summarise our results, by illustrating how the classification of the cyclic hypoxia space varies for cells with functional (Fig. 4.8a), partially-functional (Fig. 4.8b) and defective (Fig. 4.8c) S/G2 checkpoints. We can identify three main regions A, C and E, which identify cyclic hypoxia protocols that, on the long term, result in: unperturbed proliferation of cells (low toxicity), severe impairment of the proliferative capacity of cells (medium toxicity) and complete growth arrest (high toxicity). While the transition from low to medium toxicity is dictated by the activation of the G2 checkpoint, the cause of complete growth arrest (transition from C to E) varies between cells. For cells that have defective S/G2 checkpoints (left panel), loss of proliferative capacity is due to dysfunctional repair dynamics. For cells with partially-functional S/G2 checkpoints, which might be interpreted as cells from a healthy tissue, the boundary of the region E is dictated by the extent to which cells entering the G2 phase become senescent. In contrast, cells with functional S/G2 checkpoints (right panels) maintain their proliferation capacity for most of the cyclic hypoxia conditions shown, except in the very bottom-right corner of the diagram, where growth arrest is driven primarily by fork collapse. Between these three extreme conditions, we find that two intermediate regions might exist: B and D. Here we find that the proliferation capacity of cells is more significantly affected by whether or not cells have a functional G1 checkpoint. This is because the time at which cells initiate DNA synthesis dictate the extent to which cells experience replication stress in the S phase. Fig. 4.8 offers an easy framework to determine response of different cell-lines to the same cyclic hypoxia protocol, which requires computation of the Malthusian parameter $\lambda_\dagger$ (as done in Section 4.4.1) only when cells experience the same level of toxicity.

An interesting implication of our analysis, which is apparent from Figs. 4.8, is that there is no cell-cycle regulation strategy that enables a cell to have a general advantage (defined as higher proliferation rate) in cyclic hypoxia; rather, distinct strategies confer cells a growth advantage only under particular cyclic hypoxia protocols. So far, we have focused on characterising cell-cycle dysregulation in cyclic hypoxia and how this impacts the long term growth dynamics of cell cultures in different environments. However, proliferation is just one of the many traits, such as survival and migration, that characterise the ability of cancer cells to drive disease progression [20]. In the

next chapters we will use our model to investigate how cyclic hypoxia affects cell survival, in the presence and absence of treatment.

Chapter 5

An individual-based model of cell-cycle dysregulation in hypoxia

5.1 Overview

In Chapter 2 we introduced a deterministic model of the mitotic cell-cycle that describes how the number of cells in each cell-cycle phase evolves over time (*i.e.*, cell-cycle dynamics) in different hypoxic environments. As we saw in Chapter 3, this framework can be used to interpret flow cytometry data which show how different forms of hypoxia alter the cell-cycle dynamics of cancer cells in the short term. In Chapter 4, we used the model to predict the long term effects of cyclic hypoxia. However, it remains to validate some of the modelling assumptions about how dysregulation of the cell-cycle in cyclic hypoxia affects cell viability. A standard experimental approach to quantify the impact of a damaging agent on cell viability is to use clonogenic assays. In this chapter, we develop a cell-cycle model that can be used to simulate the outcome of clonogenic assays. The utility of our framework goes beyond the study of cell viability in hypoxia: as we will discuss in Chapter 6, it can also be used to investigate the mechanisms by which hypoxia affects the response of cell responses to treatment, specifically radiotherapy.

As discussed in Section 1.2, a population level description of cell-cycle dynamics can not be readily applied to all types of experimental data. It is, therefore, important to tailor the mathematical framework to the type of experimental data available for model validation and calibration. In Fig. 5.1, we present a schematic of a standard clonogenic assay that may be used to study cell survival in hypoxia. Unperturbed, asynchronously dividing cells are initially plated in a nutrient-rich environment and allowed to attach to a new culture plate. The cells are plated individually, so that if all cells survive, then the number of colonies observed at the end of the experiment in

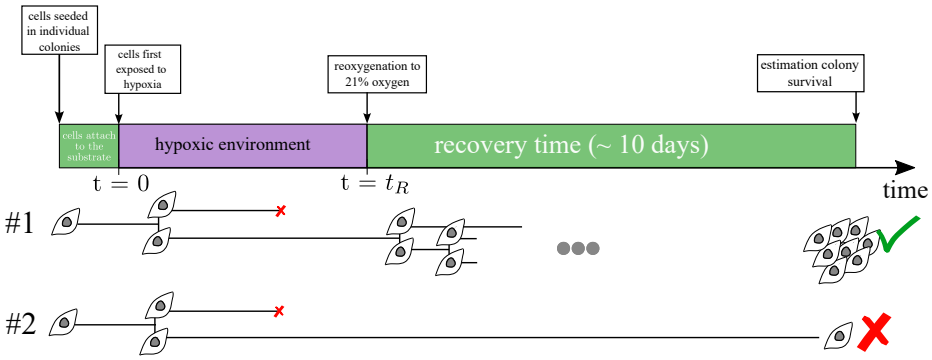


Figure 5.1: Schematic of the experimental set-up of a clonogenic assay experiments to assess cells viability after exposure to a hypoxic environment (could be either constant or cyclic hypoxia). We present the evolution of two representative lineages during the experiment. Lineage #1 gives rise to a detectable colony while lineage #2 is classified as non-viable at the end of treatment. The crosses for interrupted lineages indicate cell death.

the plate should equal the number of cells initially seeded. In practice, some cells will not survive plating so that, even when cells are allowed to grow without exposure to any damaging agent, the number of colonies formed at the end of the experiment will likely differ from the number of cells that were initially seeded. To avoid incorporating this bias into estimates of cell survival due to exposure to hypoxia, survival fractions are normalised with respect to a negative control (*i.e.*, a clonogenic assay where plated cells are not exposed to hypoxia). After the initial plating phase, cells are exposed to hypoxia for t_R hours. The cells are then re-oxygenated and allowed to grow for up to 10 days. The survival fraction is estimated by counting the number of colonies in the plate. In practice, only colonies that are sufficiently large to be visible are counted (these colonies usually consist of at least 50 cells). All this considered, the measured survival fraction, SF, is estimated to be:

$$\mathrm{SF} = \frac{n_f^{(H)}}{n_0^{(H)}} \times \frac{n_0^{(C)}}{n_f^{(C)}}, \tag{5.1}$$

where $n_0^{(C)}$ and $n_f^{(C)}$ represent, respectively, the number of cells initially seeded and the number of colonies measured at the end of the control experiment, while $n_0^{(H)}$ and $n_f^{(H)}$ represent, respectively, the number of cells initially seeded and the number measured at the end of the experiment when the cells are exposed to hypoxia. Thus, clonogenic assays quantify the ability of a single cell, that has been successfully plated, to form a detectable colony (or lineage) following exposure to hypoxia. As shown in Fig. 5.1, a lineage may be disrupted by cell death. Alternatively, cells may survive

but experience prolonged delays in the mitotic cycle due to senescence (see lineage #2 in Fig. 5.1).

Since clonogenic assays follow the evolution of individual lineages, a population level description may not be suitable. This is apparent when looking at the time-evolution of lineage #1 in Fig. 5.1. Although one cell in the lineage dies, its death does not influence the final outcome since another cell in the lineage survives and gives rise to a detectable colony at the end of treatment. A model that can describe clonogenic assays should retain information about the composition of different lineages rather than the total number of viable cells in the population (as is typically done in deterministic models). For this reason, in this chapter, we introduce a stochastic, spatially well-mixed, individual based model (IBM) of the cell-cycle, where the rules used to update each agent are based on the deterministic model presented in Chapter 2. After introducing the stochastic model in Section 5.2, we show that its predictions for the cell-cycle dynamics in different oxygen environments agree with those of the deterministic model from Chapter 2, at least in the regime with a large number of cells. Discrepancies are instead observed at low population numbers where stochastic effects are significant and can induce bifurcations in the predicted model outcomes. This aspect is investigated in Section 5.3.1, where we revisit some of the results from Chapter 2 in light of small populations. Interestingly, we find that when cell numbers are small, the model predicts a trade-off between proliferation and survival in hypoxic conditions. Such a trade-off is known to contribute to the evolution and genetic diversity of tumours, which is known to be a driver of tumour resistance to treatment. After analysing our new computational model, in Section 5.4, we show how it can be used to simulate the outcomes of clonogenic assays in different hypoxic environments. We show that the model can capture a range of responses which may be representative of the heterogeneity in survival responses to hypoxia of different cell-lines. We also discuss the role of intrinsic variability in cells with the same phenotype (*i.e.*, regulation of the cell-cycle) due to the differential effect of hypoxia on cells in different stages of the cell-cycle. We conclude in Section 5.5, by summarising the key findings from this chapter and their biological implications for how cyclic hypoxia might affect treatment efficacy. This aspect will be explored in more detail in Chapter 6, where we will show how our stochastic computational framework can be used to investigate how fluctuating oxygen environments may affect cell survival to radiotherapy.

5.2 Stochastic individual-based (IB) model

Following the model presented in Section 2.2, we consider a population of cells in a well-mixed system. Cells are exposed to a spatially-homogeneous, but dynamic, oxygen environment. Oxygen levels, $c = c(t)$, are externally prescribed. As before, we define hypoxia as oxygen levels that fall below the threshold c_H (see Table 2.1). Since cell numbers during clonogenic assay experiments remain low, we neglect any competition for space and nutrients. The model could be extended to include spatial effects and to directly test such assumptions. However, we postpone this discussion to Chapter 7, as part of future work.

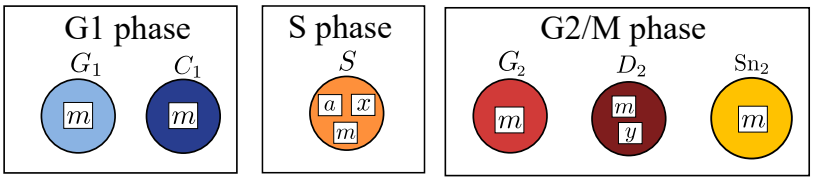


Figure 5.2: Illustration of the subpopulations present in our IB model of the cell-cycle. Internal variables associated to each cell-cycle state are indicate in the white boxes.

As illustrated in Fig. 5.2 and following the same modelling assumption as in Chapter 2, we divide the cell population into 6 sub-populations depending on the cell-cycle state they are in: $\mathcal{Z} = \{G_1, C_1, S, G_2, D_2, \mathrm{Sn}_2\}$. We recall that G_1 and C_1 states correspond to cells in G1 phase; where G_1 cells are actively cycling and C_1 cells are transiently arrested due to hypoxia. Cells in states G_2, D_2 or Sn_2 are in the G2/M phase; the former have the ability to successfully divide, while D_2 and Sn_2 cells can not divide due to damage. D_2 cells are transiently arrested and can re-enter the cell-cycle whereas Sn_2 (or senescent) cells are permanently arrested. We refer to Section 2.2 for more detail. We denote via $Z(t)$ (with $Z \in \mathcal{Z}$) the total number of cells in subpopulation Z at time t. We group sub-populations depending on the cell-cycle phase to which they belong: $\mathcal{Z} = \mathcal{Z}_1 \cup \mathcal{Z}_S \cup \mathcal{Z}_2$, where $\mathcal{Z}_1 = \{G_1, C_1\}$, $\mathcal{Z}_S = \{S\}$ and $\mathcal{Z}_2 = \{G_2, D_2, \mathrm{Sn}_2\}$. We then denote by $N_k(t)$ the number of cells in subset $\mathcal{Z}_k$ (with $k \in \{1, S, 2\}$) at time t and, by $N(t) = N_1(t) + N_S(t) + N_2(t)$, the total number of cells in the population at time t.

Depending on their cell-cycle state, cells can be characterised by one or more internal variables. All cells are structured according to the variable $m \geq 0$ representing the levels of expression of proteins associated to DNA damage response of the j-th agent; this is normalised so that $m = 1$ corresponds to physiological levels. The variable m is analogous to the global field m in the deterministic model presented in Chapter 2

(see Eq. (2.1)), but instead of using a global variable, the IBM framework enables us to follow its evolution in each cell. The expression level of proteins is assumed to be inherited during cell division and upon transition between phases of the mitotic cycle. Agents of type S have two additional continuous internal variables: x and a. Here $x \geq 1$ is a measure of the DNA content of the j-th cell in the S subpopulation at time t, while $a \geq 0$ plays the role of an internal clock capturing the time a cell has spent in the S phase. To this end, the variable a can be viewed as the analogue of the delay variable, τ_S, in the deterministic model from Chapter 2. Cells transitioning to the S phase are initialised with $x = 1$ and $a = 0$ and they exit once $x \geq 2$. We structure cells of type D_2 by the internal variable $y \geq 0$ representing their level of repair. Cells entering the state D_2 are initialised with $y = 0$ which increases over time as cells repair. As for the deterministic model presented in Section 2.2, D_2 cells successfully re-enter the cell-cycle once $y \geq y_H$, where y_H is a prescribed arbitrary threshold (see Table 2.1).

5.2.1 Modelling cell-cycle progression

Cells can update their cell-cycle state, divide or die with probabilities that may depend upon the value of their internal state variables and/or the oxygen levels, c, based on similar modelling assumptions to those introduced in Chapter 2. Given a sufficiently small time step Δt, during the time interval $[t, t + \Delta t)$, the following rules are used to update the cell-cycle state of each agent:

(R1) Cells in state C_1 transition to the S phase or remain in state C_1 with probabilities $\mathcal{K}_1(c)\Delta t$ and $1 - \mathcal{K}_1(c)\Delta t$, respectively. Here, $\mathcal{K}_1(c)$ is as defined in Chapter 2: $\mathcal{K}_1(c) = K_1\mathcal{H}(c_H - c; \epsilon_1)$, and $\mathcal{H}$ is again the sigmoid function (see Eq. (2.4)), ϵ_1 controls its steepness and $K_1 > 0$ is as defined in Table 2.1.

(R2) Cells in state G_1 remain in state G_1 with probability $1 - k_1\Delta t$, whereas the become of type C_1 or S with probabilities $\mathcal{Q}_1(c)k_1\Delta t$ and $(1 - \mathcal{Q}_1(c))k_1\Delta t$, respectively. Here $\mathcal{Q}_1$ is defined as in Chapter 2: $\mathcal{Q}_1(c) = q_1\mathcal{H}(c - c_H; \epsilon_1)$, where $q_1 \in [0, 1]$ is as defined in Table 2.1.

(R3) Cells in state G_2 divide giving rise to two G_1 daughter cells with probability $k_2\Delta t$ or remain in state G_2 with probability $1 - k_2\Delta t$.

(R4) Cells in state D_2 with internal state variable $y \geq y_H > 0$ re-enter the cell-cycle transitioning to state G_2. Cells in state D_2 with internal state variable $y < y_H$

die with probability $\mu_2^C \Delta t$ as a result of attempting mitotic division prior to successfully completing repair.

(R5) Cells of type S with internal protein expression levels m die with probability $u_S(m)\Delta t$, where u_S is defined by Eq. (2.3c) in Chapter 2: $u_S(m) = \mu_S \mathcal{H}(M_{FC} - m; \epsilon_{FC})$, where μ_S and M_{FC} are as in Table 2.1. If a cell survives and its DNA content $x \geq 2$, it transitions to either state G_2, D_2 or Sn_2 with probabilities that depends on the internal variable a: $1 - \mathcal{P}_2(a)$, $(1 - \mathcal{P}_{Sn}(a))\mathcal{P}_2(a)$ and $\mathcal{P}_{Sn}(a)\mathcal{P}(a)$, respectively. Here, $\mathcal{P}_2$ and $\mathcal{P}_{Sn}$ are defined as in Eq. (2.11b):

$$\mathcal{P}_2(a) = p_2 \mathcal{H}(a - T_{C2}, \epsilon_p), \quad \mathcal{P}_{Sn}(a) = p_{Sn} \mathcal{H}(a - T_{Sn}; \epsilon_{Sn}), \tag{5.2a}$$

where p_2, p_{Sn}, T_{C2} and T_{Sn} are constant model parameters defined in Table 2.1.

5.2.2 Modelling the evolution of a cell internal state

We account for the impact of hypoxia on DNA synthesis, protein expression levels and repair dynamics by allowing cells to update their internal state variables. Again, we follow similar modelling assumptions to the ones introduced in Chapter 2. Given a sufficiently small time step Δt and oxygen levels $c = c(t)$, after having updated their cell-cycle state (according to the rules outlined in Section 5.2.1), cells update their internal variables according to the following rules:

(R6) Cells in state D_2 recover from the stress/damage that they have accumulated at a rate that depends on the expression levels of repair proteins, m. In particular, the internal variable y increase of an amount Δy:

$$\Delta y = \frac{v_R^+ m}{M_R + m} \xi_y \Delta t, \tag{5.2b}$$

where v_R^+ and M_R are defined as in Table 2.1 and ξ_y is an exponentially distributed random variable with mean 1, *i.e.* $\xi_y \sim \mathrm{Exp}(1)$. Here ξ_y accounts for intrinsic variability in cell response due to phenomena not explicitly captured in the model. We choose an exponential distribution so that, in each cell, y is a monotonically non-decreasing function of time.

(R7) Cells of type S increase their DNA content, $x(t)$, of an amount Δx:

$$\Delta x = \xi_x \Delta t v_S(t), \tag{5.2c}$$

where the rate $v_S(t)$ depends on the oxygen levels according to the following piecewise linear ODE:

$$\frac{dv_S}{dt} = \mathcal{R}_S(c(t))\left(v_S^{(eq)}(c(t)) - v_S\right), \quad t > t_0, \tag{5.2da}$$

$$v_S(t_0) = v_{S,0}, \tag{5.2db}$$

where:

$$\mathcal{R}_S(c) = \begin{cases} \mathcal{R}_S^+, & c > c_H, \\ \mathcal{R}_S^-, & c < c_H, \end{cases} \quad , \quad v_S^{(eq)}(c) = \begin{cases} v_S^+, & c > c_H, \\ v_S^-, & c < c_H, \end{cases} \quad , \tag{5.2dc}$$

and $\mathcal{R}_S^{+/-}$ and $v_S^{+/-}$ are constant parameters as defined in Chapter 2 (see Table 2.1). In Eq. (5.2c) we have that $\xi_x \sim \text{Exp}(1)$. As before, we choose an exponential random variable to capture heterogeneity and guarantee that the DNA level in a cell is monotonically increasing. As we will see, by choosing an exponential distribution, we will recover, in constant oxygen environments, the standard Erlang distribution used in stochastic models of the cell-cycle [33, 32]. The internal clock a of each cell of type S is also updated at each time step by Δt, *i.e.*, $a(t + \Delta t) = a(t) + \Delta t$.

(R8) All cells change the expression level of DNA protein m by an amount Δm:

$$\Delta m = \xi_m \Delta t v_m(m, c), \quad j = 1, \ldots, N(t), \tag{5.2e}$$

where ξ_m is a normally distributed random variable with variance σ^2, (*i.e.*, $\xi_m \sim \mathcal{N}(0, \sigma^2)$), which captures stochastic fluctuations in expression levels of proteins associated with the noise intrinsic in gene regulatory networks [137]. The average rate of change of m follows the deterministic model $v_m(m, c) = \mathcal{R}_M(c)\left(M^{(eq)}(c) - m\right)$ where

$$\mathcal{R}_M(c) = \begin{cases} \mathcal{R}_M^+, & c > c_H, \\ \mathcal{R}_M^-, & c < c_H, \end{cases} \quad , \quad M^{(eq)}(c) = \begin{cases} 1.0, & c > c_H, \\ M^-, & c < c_H, \end{cases} \quad , \tag{5.2f}$$

and $\mathcal{R}_M^{+/-}$ and M^- are constant parameters as defined in Table 2.1.

5.2.3 Computational implementation

Numerical simulations of the IB model are performed in Python. An outline of the procedure followed in the simulation is presented in Figure 5.3. To simulate the oxygen dynamics, $c(t)$, we adopt the approach used in Chapter 2.

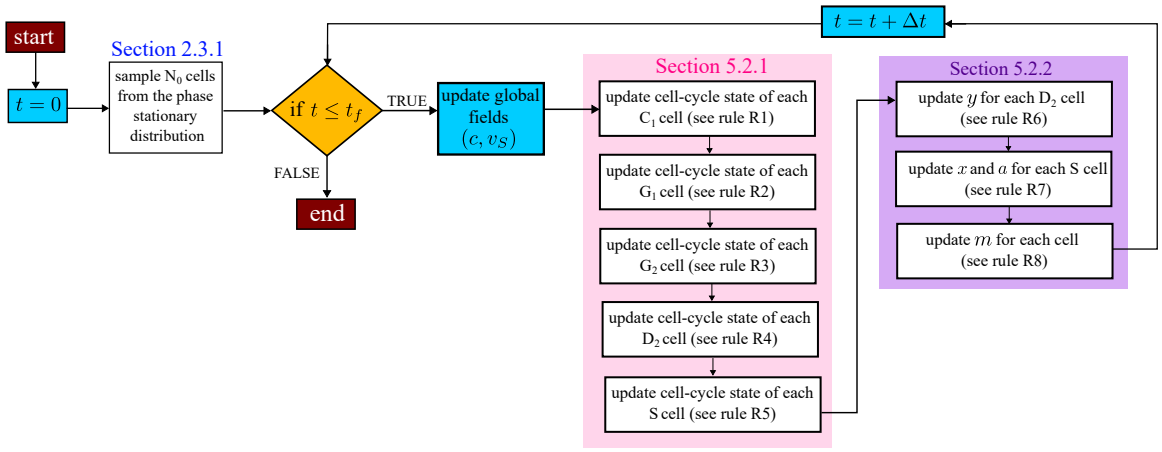


Figure 5.3: Flowchart illustrating the procedure underlying the computational implementation of the stochastic IB model. There are two main subroutines identified by the shaded areas: pink – update of cell-cycle state (details given in Section 5.2.1)) and purple – update of internal variables (details given in Section 5.2.2).

Model parameters The time step Δt is set to $\Delta t = 0.02$, while all other parameter values are as described in Section 2.2.7. As in Chapters 2 and 4, we generate a cohort of *in silico* cell-lines that differ in how they regulate cell-cycle transitions in hypoxia. We identify parameter sets corresponding to cells with functional (+), partially-functional (-) or defective (x) G1 and S/G2 checkpoints. We recall that functional checkpoints result in cells delaying cell-cycle transitions (*i.e.*, arresting) in hypoxia whereas defective checkpoints lead to cells attempting proliferation even in hypoxia. Detailed on how we model different cell-lines are given in Section 2.2.7.1. Parameter values corresponding to the different cell-lines are listed in Table 2.3; we here only change the value of the death rate in the S phase (from $\mu_S = 0.005$ to $\mu_S = 0.025$ for all cell-lines) and the death rate of D_2 cells with defective S/G2 checkpoints (from $\mu_2^C = 0.025$ to $\mu_2^C = 0.05$).

Initial conditions For consistency with the previous chapters and *in vitro* experimental conditions, cells are assumed to be initially distributed according to phase stationary solution estimated in Section 2.3.1: cells are initially in either state G_1, S or G_2 respectively with probabilities $\pi_{G_1}^+$, π_S^+ and $\pi_{G_2}^+$, where $\pi_{G_1}^+ + \pi_S^+ + \pi_{G_2}^+ = 1$. Values of π_{G_1,S,G_2}^+ are listed in Table 2.4. Furthermore, cells in the S phase have DNA distribution $\hat{s}/\pi_S^+$, where $\hat{s}$ is defined in Eq. (2.26b). All cells are initialised with $m = 1$, corresponding to them having physiological expression level of repair proteins. The initial cell-cycle state of a cell is selected by generating a random number θ from a uniform distribution on $[0, 1]$; if $\theta \leq \pi_S^+$ or $\theta - \pi_S^+ < \hat{\pi}_1$ the cell is given

cell-cycle state S and G_2, respectively; otherwise, it is initialised as a G_2 cell. For cells of type S the internal variable x (DNA content) is selected as $x = 1 + \hat{s}^{-1}(\theta)$ (defined as above) and, consistently, $a = (x-1)\tau_S^+$.

5.3 Comparing the deterministic and IB models

We start by comparing results generated from the stochastic and deterministic (continuous) models, for the three environmental conditions analysed in Chapter 2: oxygen-rich, constant hypoxia and cyclic hypoxia. In particular, for the same set of parameter values and initial conditions, we perform 100 repetitions of the stochastic model which is initialised with 150 cells. For each simulation, we output the time-evolution of the total number of cells in each phase of the mitotic cycle, *i.e.*, $N_m(t)$ with $m \in \{1, S, 2\}$. As shown in Fig. 5.4, the ensemble average predictions

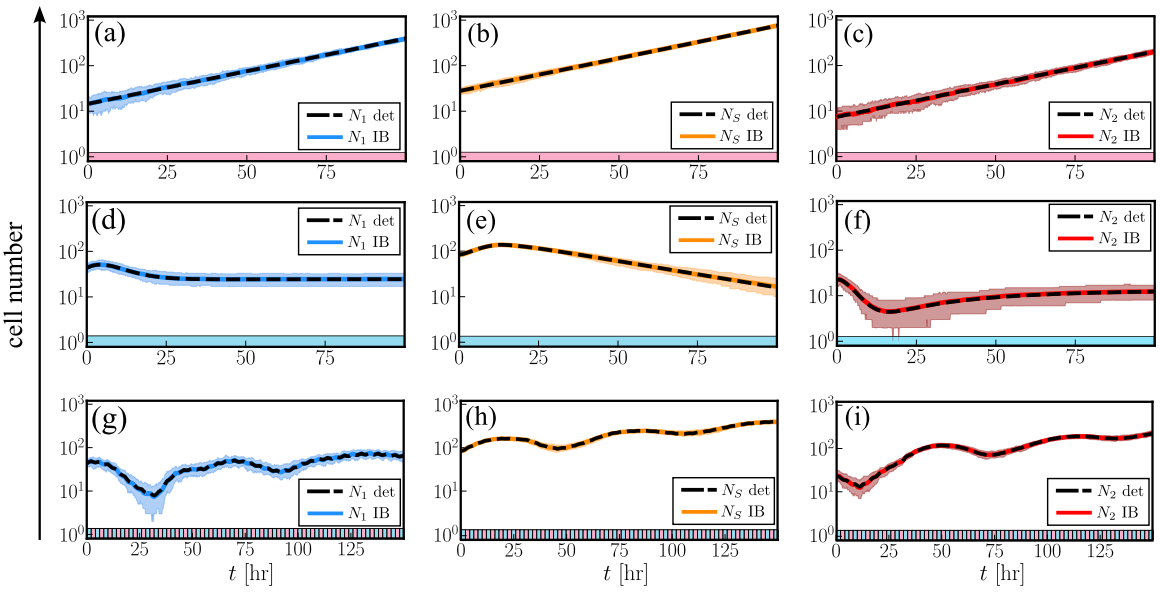


Figure 5.4: We plot the evolution of the total number of cells in each phase of the cell-cycle, N_m with $m \in \{1, S, 2\}$, as predicted by the IB (see continuous, coloured lines) and deterministic (det) models (see dashed, black line) when exposing (+,+) cells to three different oxygen environments (a-c) normoxia; (d-f) constant hypoxia; (g-i) (2,2)-cyclic hypoxia. The results from the deterministic (det) model are obtained solving Eqs. (2.18). The results from the IB model correspond to the average over 100 realisation and the related 99%-confidence interval is displayed by the coloured areas surrounding the curves. Cell numbers are on a log-scale. All simulation are initialised with 150 cells. Parameter values for the deterministic and IB model are the same and set as indicated in Section 5.2.3.

of the IB model (see continuous, coloured lines) are in good agreement with the re-

sults from the deterministic model (see black, dashed lines), for all oxygen protocols considered. Furthermore, the confidence intervals on the predictions of the IB model (indicated by the shaded areas) are narrow. This indicates that realisations of the IB model all yield the same qualitative behaviour, which agrees with the predictions of the deterministic model. In Section 5.3.1 we will explore what happens when we initialise the model with a much smaller number of cells. While the deterministic model only captures perturbations of the overall cell-cycle dynamics in hypoxia, by following changes in cell-number, the predictions of the stochastic model yield additional information by following the cell-cycle progress of individual cells. For example, the IBM framework can be used make predictions about how hypoxia perturbs the timings of cell-cycle transitions between subsequent phases of the cell-cycle (results not shown due to space limitations).

5.3.1 Predicting survival for small cell numbers

At the start of this section, we showed that when cell numbers are large, there is good agreement between predictions of the stochastic and deterministic models. Discrepancies are expected, however, when considering populations consisting of only a few cells. In this case, stochastic fluctuations may have a significant effect on the long time dynamics of the stochastic model. In particular, we observe parameter regimes where extinction and escape of an initially small population of cancer cells are both possible outcomes of the stochastic model, a property that the (linear) deterministic model from Chapter 2 does not exhibit. When considering oxygen-rich environments, the probability of cell death is effectively zero. In this case, realisations of the stochastic model always converge to the same regime of asynchronous exponential growth predicted by the deterministic model (see Section 2.3.1). In contrast, when cells are exposed to hypoxia (either continuously or periodically), the lack of nutrients can impair cell viability (by inducing cell death and/or senescence). Depending on the probability of a cell surviving hypoxia, persistence and extinction of a population of cells may be observed in realisations of the stochastic model initialised with N cells sampled from the same initial distribution and with the same values of model parameters. In the context of tumour development, a small number of abnormal cells characterises disease initiation, the fixation of new mutant clones, and post-treatment growth, when the survival of a few cells can be sufficient to drive disease recurrence. For these reasons, clonogenic assays have been designed to estimate the effect of toxic

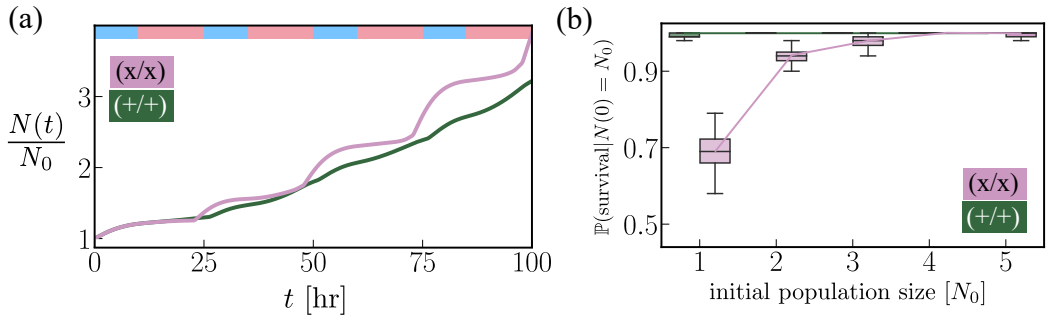


Figure 5.5: (a) We illustrate the predictions of the deterministic model for the evolution of the population size in a (10,15)-cyclic hypoxia protocol. (b) Estimates of probability of survival for the two cell-lines, extrapolated from simulations of the stochastic model, as a function of the initial size of the population, N_0. Parameters values are as specified in Section 5.2.3.

agents on the fate of single-cells (by following lineages) instead of measuring the effect of treatment on the overall population level growth dynamics of *in vitro* tumours. This distinction will become more apparent when we look at a specific example.

We consider the growth of cultures of two cell-lines that have different cell-cycle regulation mechanisms. Following Chapter 2, the notation (x/x) indicates mutant cells with both the G1 and S/G2 checkpoints defective, while for (+/+) cells all checkpoints are functional. In Fig. 5.5a we use the deterministic model to predict the growth dynamics of these two *in vitro* cell cultures during exposure to a (10,15)-cyclic hypoxia protocol and $N(0) = 1$ cells. Over time, both cultures evolve to a regime of quasi-periodic geometric growth (see Section 4.2) with the population consisting of (x/x) cells growing at a faster rate than the (+/+) cell culture. This suggest that (x/x) cells have a growth-advantage over (+/+) cells in (10,15)-cyclic hypoxia.

When simulating multiple realisations of the same experiment with our stochastic model, we find that the ensemble mean growth dynamics over several realisations agrees with the predictions from the deterministic model. Nonetheless, when comparing the predicted growth dynamics between realisations we observe divergent outcomes. For example, if we start the simulation from a single (x/x) cell, then we find instances where the cell gives rise to a growing tumour (*i.e.*, escape scenario), but also simulations in which no tumour forms (*i.e.*, extinction scenario). From this point of view, it is more appropriate to describe the predictions of the stochastic model in terms of a survival probability (or tumour initiation capacity) rather than looking at its ensemble growth dynamics.

To estimate the survival probability for a given cell culture, we first simulate 600 lineages starting from a single cancer cell randomly sampled from an exponentially

growing population in an oxygen-rich environment and following the procedure in Fig. 5.3. From time $t = 0$, the cells are exposed to a prescribed hypoxic environment. At the end of the simulations, we classify a lineage as surviving if it has the ability to proliferate (*i.e.*, it consists of at least one viable and non-senescent cell), or extinct. To estimate the survival probability of a culture which initially consists of N_0 cells, we generate κ cultures, each consisting of N_0 lineages randomly sampled from the 600 previously generated. This procedure is valid since lineages proceed independently (if competition were included then this procedure would not be possible). The probability of survival of a colony is then estimated as the ratio between the number of cultures that give rise to a tumour (*i.e.*, they consist of at least one surviving lineage) and κ (*i.e.*, the number of simulated cultures). Given that the survival probability is estimated from stochastic simulations, it is itself a random variable. To estimate its distribution we employ a standard bootstrap method with resampling.

Using the procedure above, we estimate the probability of survival of cell cultures initiated from either (x/x) or (+/+) cells as we vary the initial culture size (N_0). Results are shown in Fig. 5.5b. As expected, the survival probability increases with N_0 and in the limit $N_0 \gg 1$ it asymptotes to a Dirac-delta function centred at 1 (*i.e.*, survival is certain). For (+/+) cell cultures, the probability of survival is approximately one, even when the cultures are initialised from a single cell ($N_0 = 1$). In contrast, the model predicts that (x/x) cell cultures have a much lower probability to survive cyclic hypoxia for small N_0. In particular, we find that for $N_0 = 1$, the probability of survival reduces to $\approx 75\%$. To summarise, we find that while, on average, (x/x) cell cultures grow faster than (+/+) cell cultures under (10,15)-cyclic hypoxia, they are also more likely to go extinct than (+/+) cell cultures. These findings suggest that, while cells with impaired activation of checkpoints can lead to tumours with larger growth rates in cyclic hypoxia, they also have decreased chances of surviving periodic exposure to hypoxia. In contrast, cells with more stringent cell-cycle control mechanisms yield tumours that grow more slowly under cyclic hypoxia, but have a higher chance of surviving. Such trade-offs between proliferation and survival have been investigated previously in the mathematical oncology literature, in particular in relation to the idea of the cost of resistance [20]. However, to the best of our knowledge, it has not been discussed in the context of cyclic hypoxia.

As mentioned in earlier chapters of this thesis, the term cyclic hypoxia relates to a range of environmental conditions. We can combine predictions from the deterministic and stochastic models to investigate whether cells are exposed to trade-offs between proliferation and survival generally in fluctuating oxygen environments or only under

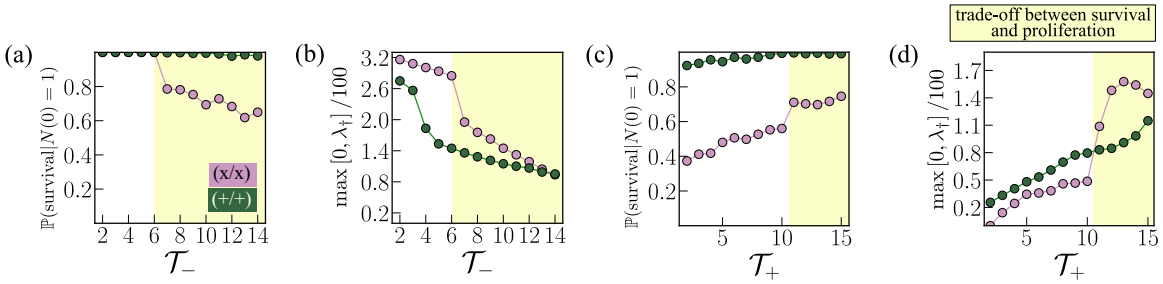


Figure 5.6: (a)-(b) Estimates of the (a) survival probability and (b) asymptotic growth rate in (10,$\mathcal{T}_-$)-cyclic hypoxia for the two cell-lines considered in Fig. 5.5. (c)-(d) Estimates of the (c) survival probability and (d) asymptotic growth rate in ($\mathcal{T}_-$,15)-cyclic hypoxia for the two cell-lines considered in Fig. 5.5. Here survival is estimated from stochastic simulations of the model (we here report the mean survival), while the rate $\lambda_\dagger$ is estimated as discussed in Section 4.4.1. The shaded area indicates cyclic conditions for which the cell-line with the largest growth rate is not the one with the largest survival probability.

specific cyclic hypoxia protocols. To estimate the proliferation capacity of cancer cells under cyclic hypoxia, we use the approach presented in Chapter 4, where we used the predictions of the deterministic model to estimate the asymptotic growth rate, $\omega_{\mathcal{T}}^{\dagger} = \max\{0, \lambda_\dagger\}$. Survival instead is estimated as above. In Fig. 5.6 we show how the trade-off between survival and proliferation is not predicted to be a general feature of cyclic hypoxia, but rather a property of specific cyclic hypoxia environments. In Figs. 5.6a-5.6b we illustrate how the estimated survival and proliferation capacity of our two cell-lines change across the ($\mathcal{T}_-$,15)–class of cyclic hypoxia protocols. A trade-off between proliferation and survival is predicted only for conditions where the period of exposure to hypoxia, $\mathcal{T}_-$, is longer than the critical value $\mathcal{T}_- \approx 6$ hours. As $\mathcal{T}_-$ increases further, the predicted growth rate of the two populations converges, suggesting that the trade-off exists only for intermediate values of $\mathcal{T}_-$. In Figs. 5.6c-5.6d we focus on another class of cyclic hypoxia protocols: (10,$\mathcal{T}_+$)–cyclic hypoxia. For small $\mathcal{T}_+$, the model predicts there is no trade-off between survival and proliferation. However, this regime differs from the one observed for small $\mathcal{T}_-$. In this case, it is the (+/+) cells that have an advantage over (x/x) cells, both in terms of survival and proliferation; in contrast, for small $\mathcal{T}_-$, both cell-lines have the same survival probability, but that (x/x) cells have a proliferation advantage. When considering large $\mathcal{T}_+$, the difference between the growth rates, $\omega_{\mathcal{T}}^{\dagger}$, of the two cell-lines decreases, while the difference in survival probability asymptotes to a constant, non-zero value. This suggests that, for large $\mathcal{T}_+$, the extent of the trade-off between survival and proliferation is negligible as (+/+) will have a clear advantage in term of survival over (x/x) cells while both cell-lines have similar proliferation capacity. We conclude

that, under very toxic (large $\mathcal{T}_-$ and small $\mathcal{T}_+$) and mildly toxic (large $\mathcal{T}_+$ and small $\mathcal{T}_-$) cyclic hypoxia conditions, survival and proliferation capacity are not mutually exclusive. In very toxic environments, cells are overall better off by keeping intact their cell-cycle control mechanisms in order to favour survival. In contrast, in mildly toxic environments, cells are better off losing cell-cycle control mechanisms so as to obtain a growth advantage. In intermediate regimes, it is unclear what strategy cells may adopt due to the existing trade-off between survival and proliferation.

While providing interesting insights on the role of cell-cycle control strategies in different cyclic hypoxia environments, the above example illustrates how we can use our stochastic computational model to estimate cell survival (*i.e.*, viability) under hypoxia. However, if we want to make useful predictions on cell survival in hypoxia, we must use our model to estimate quantities that can be measured experimentally. To this end, in the remainder of this chapter, we use our modelling framework to simulate *in silico* experiments involving clonogenic assays.

5.4 A model of clonogenic assays in hypoxia

As mentioned in the introduction, the primary goal of this chapter is to develop a computational model to predict the outcome of clonogenic assays. Here our focus is on estimating survival in hypoxia, but our approach can be used more widely to quantify the survival probability of cell populations exposed to damaging agents whose impact depends on a cells' cell-cycle state. In this section, we use our framework to investigate how survival under hypoxia may vary amongst cells with different cell-cycle control mechanisms.

5.4.1 *In silico* simulation of clonogenic assays in hypoxia

The standard protocol for clonogenic assays is illustrated in Fig. 5.1. We aim to replicate closely the different steps in the experiment using our stochastic model. We initialise our simulations by seeding N_0 (here, $N_0 = 150$) cells randomly sampled from the phase stationary distribution characterising asynchronous exponential growth. To identify lineages, we mark each initially seeded cell with a unique tag, which is inherited by all its daughter cells upon division. We expose the cells to either constant or periodic hypoxia for t_R hours. Then the cells are re-oxygenated and allowed to grow unperturbed for another 10 days. At the end of the 10 days of unperturbed growth, the surviving lineages/colonies are counted; here a lineage is classified as viable if it contains at least 50 cells. Given N_f viable lineages, the survival fraction is estimated

by the ratio SF $= N_f/N_0$. We note that this expression is slightly different from the formula used experimentally (see Eq. (5.1)); this is because in our computational model the survival fraction for the controlled experiment is always exactly 1, since external factors affecting cell viability are neglected.

Estimating survival curves We compute survival for increasing discrete values of t_R, *i.e.,* increasing the time for which cells are exposed to longer periods constant or cyclic hypoxia. Continuous survival curves are obtained by linear interpolation of the discrete values of SF(t_R). Given that SF(t_R) is estimated via realisations of the stochastic model, it is itself a random function. We therefore repeat the procedure described above 100 times to estimate the distribution of the random variable SF. When results are presented, we report the mean and inter-quantile range (IQR) over 100 realisation of the survival curves, SF(t_R).

5.4.2 The role of checkpoints in cell survival under hypoxia

We start by investigating cell survival under constant hypoxia. The results presented in Fig. 5.7 compare the estimated survival curves for our cohort of *in silico* cell-lines. In each panel, we compare cells with the same checkpoint control of the S→M transition, but variable activation of the G1 checkpoint. The red bar above each plot indicates the experimental conditions under which cells in the S phase are sensitive to fork collapse. This allows us to determine to what extent fork collapse affects the result of the *in silico* clonogenic assays. When discussing the results, we will refer to lineage survival, rather than cell survival, since this is what clonogenic assay are designed to measure. In particular, we discuss survival of (a, b)-lineages, where $a, b \in \{+,\text{-},\text{x}\}$.

In all panels, we find that the pink curve is always lower than the other. This implies that having a defective G1 checkpoint reduces lineages survival regardless to whether the S/G2 checkpoints are functional or not (compare pink line - defective G1 checkpoint - with the green and purple lines). Furthermore, in each panel, the green and purple lines are indistinguishable. We conclude that there is no significant difference in lineages consisting of cells with partially-functional (-) or functional (+) G1 checkpoint. Moving on to comparing the predictions of survival in different panels, we find significant differences in the qualitative shape of the survival curves. While lineages with defective S/G2 checkpoints (see Fig. 5.7a) are insensitive to hypoxia for small t_R, if they are exposed to hypoxia for more than ≈ 6 hours then their chance of survival is severely decreased. We note that the steep decrease in the survival fraction

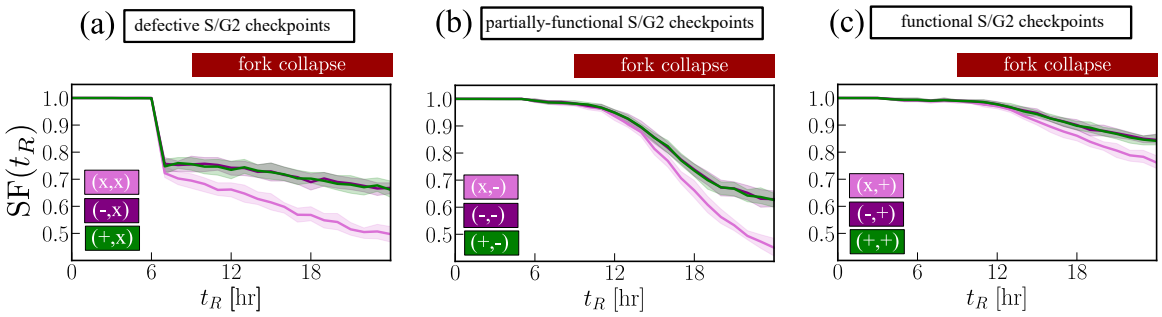


Figure 5.7: (a)-(c) Survival curves SF(t_R) for cells grown in constant hypoxia. Different panels compare cell-lines with (a) defective, (b) partially-functional and (c) functional S/G2 checkpoints. Curves of the same colour correspond to cell-lines with the same G1 checkpoint dynamics: (pink) defective, (indigo) partially-functional, (green) functional. On the top of each plot, the red bar indicates experiments for which cells are sensitive to fork collapse. Details on how survival curves are computed are given in Section 5.4.1. Parameters values are as specified in Section 5.2.3.

at $t_R = 6$ hours can not be linked to fork collapse since this is activated only for $t_R > 10$. Instead, loss of viability is associated with D_2 cells attempting mitosis prior to completing repair and dying as a result. For $t_R > 6$ hours, the chances of survival decrease almost linearly with t_R, with the gradient being the greatest for cells that have defective G1 checkpoint (*i.e.*, (x/x) cells). Cell lineages with partially-functional S/G2 checkpoints (see Fig. 5.7b) have a good chance of survival up to approximately 12 hours in hypoxia. For longer exposure times, survival decreases non-linearly trend unlike the survival curves from the other panels which follow an approximately linear trend for $t_R \gg 1$. For these lineages (*i.e.*, (x,-), (-,-) and (+,-)) reduced survival is due to both fork collapse and premature senescence (*i.e.*, accumulation of Sn_2 cells). Finally, cell lineages with functional activation of the S/G2 checkpoints (see Fig. 5.7c) have a greater chance of surviving hypoxia. In this case, almost all lineages survive if they are exposed to hypoxia for less than approximately 12 hours. For longer periods of hypoxia, as cells in the S phase become sensitive to fork collapse, the survival probability decreases almost linearly. Summarising the results presented in Fig. 5.7, lineage sensitivity to constant hypoxia depends on how long cells are exposed to constant hypoxia and the extent to which they can suppress cell-cycle transitions during hypoxia. If the duration of exposure to hypoxia is sufficiently long, then survival is maximal for lineage that have functional G1 checkpoint.

We now investigate cell survival in cyclic hypoxia. As mentioned throughout this thesis, cyclic hypoxia captures a variety of environmental conditions that can drive significantly different responses in cells. Here we focus on one example: a (4,2)-cyclic

hypoxia protocol, (*i.e.,* 4 hours in hypoxia every 6 hours). Based on the analysis in Section 4.4.3, our model predicts that, this protocol, in contrast to constant hypoxia, will not induce fork collapse regardless of how long the cells are exposed to cyclic hypoxia. The predicted survival curves for our cohort of *in silico* cell-lines are presented

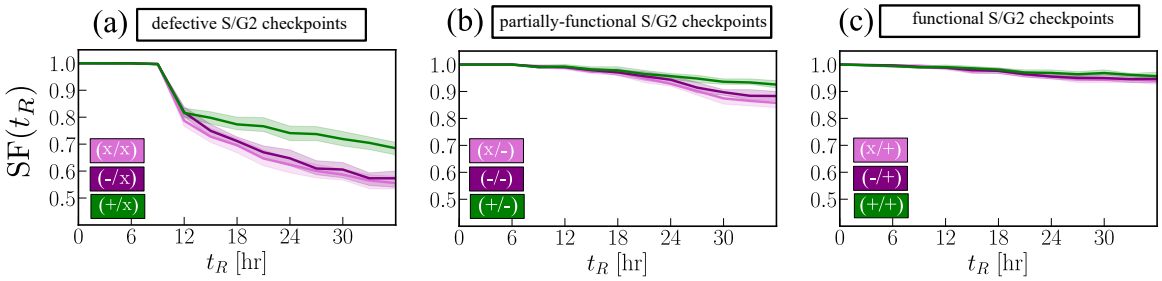


Figure 5.8: (a)-(b) Survival curves (SF(t_R)) for cells exposed to (4,2)-cyclic hypoxia. Panels and colour scheme are organised as in Fig. 5.7. Details on how survival curves are computed using our IB model are given in Section 5.4.1.

in Fig. 5.8. As in Fig. 5.7, each panel compares cells with the same regulation of the S→M transition, but variable activation of the G1 checkpoint (different probability of delaying the G1→S transition in hypoxia). Interestingly, lineages formed by cells with partially-functional (-) and defective (x) G1 checkpoint (compare pink and purple curves in any of the panels) are equally sensitive to cyclic hypoxia; this is despite the fact that the two cell-lines have significantly different sensitivity to constant hypoxia. In contrast, cell lineages with a functional (+) G1 checkpoint are more likely to survive prolonged exposure to cyclic hypoxia (see green curves). When comparing cell-lines with different abilities to control the S→M transition (*i.e.*, curves from different panels but same colour), we find that the survival curves follow a similar trend to those observed under constant hypoxia (see Fig. 5.7) except for lineages with partially-functional S/G2 checkpoints (see Fig. 5.8b) whose survival curves in (4,2)-cyclic hypoxia are now almost indistinguishable from the one of cells with functional Fig. 5.8c). In Fig. 5.8a (as in Fig. 5.7a), we observe a steep decrease in cell survival on exposure to cyclic hypoxia for longer times ($t_R > 8$ hours), which is analogous to the steep decrease in survival when the same lineages where exposed to $t_R > 6$ hours in constant hypoxia. As in constant hypoxia, functional S/G2 checkpoints leads to increased survival (Fig. 5.8c), while compromised activation of these checkpoint controls (as in Fig. 5.8c) results in a substantial fraction of cells losing their viability if they are exposed to cyclic hypoxia for more than 10 hours.

Overall, we find that lineages with a functional G1 checkpoint have higher survival probabilities under constant and (4,2)-cyclic hypoxia. This is consistent with the re-

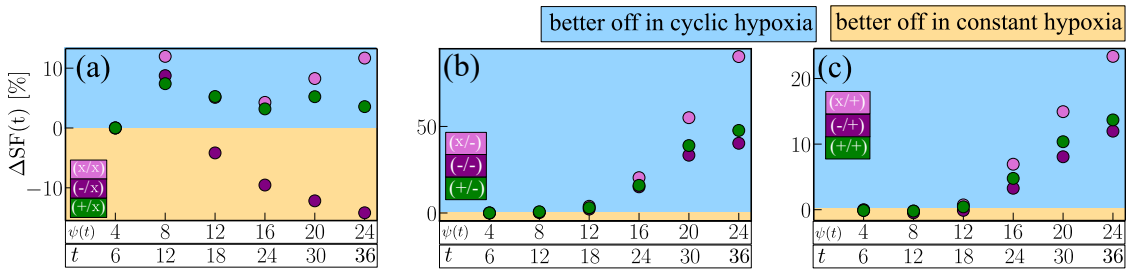


Figure 5.9: (a)-(c) Relative survival, ΔSF (see Eq. (5.1)) in constant (Figs. 5.7) and cyclic (Figs. 5.8) hypoxia for cell-lines with: (green) functional, (purple) partially-functional, (pink) defective G1 checkpoint. Results for cells with (a) defective, (b) partially-functional and (c) functional S/G2 checkpoints.

sults in Section 5.3.1, where we found that a functioning G1 checkpoint prevented extinction of tumours of small size under exposure to cyclic hypoxia. However, the extent to which lineage survival increases if cells have a functional G1 checkpoint depends whether or not they have functional S/G2 checkpoints. This is particularly evident when simulating clonogenic assays in (4,2)-cyclic hypoxia. If the S/G2 checkpoints are functional (see Fig. 5.8c), then the benefit of having a functional G1 checkpoint is minimal; in contrast, if a cell has defective S/G2 checkpoints (see for example Fig. 5.8a), having a functioning G1 checkpoint (see (+/x) curve) can significantly improve its lineages' chances of survival.

To conclude, we compare the results presented in Figs. 5.7 and 5.8. In doing so, we simulations of the IB model in which cells are exposed to hypoxia for the same overall duration. For example, an experiment where cells are exposed to 2 cycles of the (4,2)-cyclic protocol ($t_R = 12$) will be compared to an assay where cells are exposed continuously to hypoxia for 8 hours ($t_R = 8$). Consider the mapping $\psi : [0, \infty) \to [0, \infty)$ that returns the overall time cells have spent in hypoxia after t hours of the (4,2)-cyclic protocol. We introduce the following metric which compares survival in cyclic and constant hypoxia:

$$\Delta\mathrm{SF}(t) = \frac{\mathrm{SF}^{(2)}(t) - \mathrm{SF}^{(1)}(\psi(t))}{\mathrm{SF}^{(1)}(\psi(t))}, \tag{5.1}$$

where $\mathrm{SF}^{(1)}(\xi)$ and $\mathrm{SF}^{(2)}(\xi)$ represent the mean survival estimated, respectively, under constant and cyclic hypoxia with $t_R = \xi$. In Eq. (5.1), $\Delta\mathrm{SF} > 0$ implies that lineages have a better chance of surviving in (4,2)-cyclic hypoxia rather than constant hypoxia (and conversely if $\Delta\mathrm{SF} < 0$). The estimated values of ΔSF for our cohort of cell-lines are illustrated in Fig. 5.9. As in the previous figures we distinguish lineages according to their regulation of the G1→S transition by using different colours, while

lineages consisting of cells with different S→M regulation are in different panels. From Fig. 5.9, it is apparent that the general question– "*are cells more sensitive to constant or cyclic hypoxia?*" –is not well-posed. Indeed, we find that the sign of $\Delta\mathrm{SF}(t)$ may vary when comparing the response of different lineages and, also, when considering the value of $\Delta\mathrm{SF}(t)$ for the same cell-line but at different times (see, for example, the results for (-/x) cells in Fig. 5.9a). For cells with functional or partially-functional S/G2 checkpoints, $\Delta\mathrm{SF}(t)$ is almost always positive and increases with t_R (see Figs. 5.9b-5.9c) regardless of how a cell-line regulates the G1→S transition. This suggests that cells with a functional or partially-functional S/G2 checkpoints are more likely to withstand (*i.e.*, it has a higher chance to survive) exposure to periodic, rather than, constant hypoxia (given same time of exposure). This is not always the case for lineages that have defective S/G2 checkpoints (see Fig. 5.9a). In this case the sign of $\Delta\mathrm{SF}$ depends upon the duration t of exposure to hypoxia, as well as on the cells ability to regulate the G1→S transition (compare different colours in the same panel). Interestingly, we find that (-,x) lineages are more likely to withstand prolonged exposure to chronic rather than cyclic hypoxia. This is because in the case of a partially-functioning G1 checkpoint, C_1 cells quickly initiate DNA synthesis after re-oxygenation, which is deleterious in cyclic hypoxia but does not play a role under constant hypoxia.

5.4.3 Cell-cycle dependent survival in hypoxia

So far, we have focused on how different choices of model parameters (*i.e.*, considering different cell-lines) yield different predictions on cell survival under hypoxia. We now focus on intra-specific heterogeneity (*i.e.*, capturing the viability in survival of cells from the same population). In our models, intra-specific heterogeneity results from the fact that hypoxia affects cells differently depending on which phase of the cell-cycle they are in. This can significantly impact the outcomes of clonogenic assays where cells are synchronised in a specific phase of the cell-cycle prior to plating. While experimental technique exists to do so, perfect synchronisation might be difficult to achieve in practice. In our computational model, we can simply extrapolate information on survival conditioned on the state cells are in when initially seeded (for our model that is either G_1 or S or G_2). Since we are interested in charactering heterogeneity in the survival of cells belonging to the same population, we here focus on the (-/x) cell-line (see purple curves in Figs. 5.7-5.8); however the results can vary greatly depending on the cell-line and the type of hypoxic environment (results not shown). As in the previous section, when presenting our results, we will refer to the

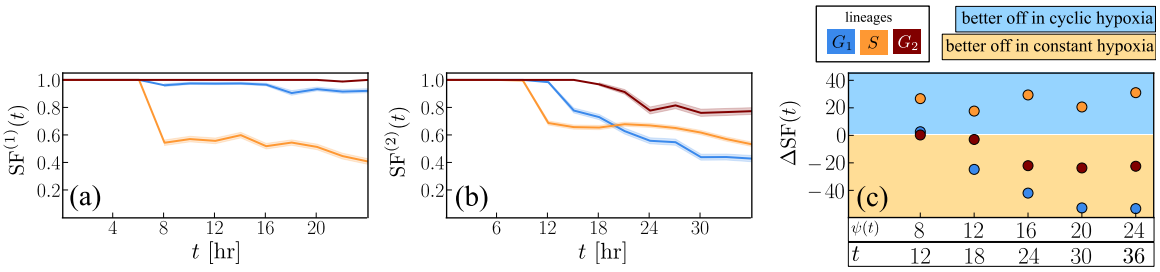


Figure 5.10: (a-b) Survival curves in constant (a) and (4,2)-cyclic hypoxia (b) for lineages where the parental cells are sampled from different phases of the cell-cycle. The data in (a) and (b) are respectively obtained from the simulations that yield the purple curve in Figs. 5.7a and 5.8a. In (c) we plot the estimated value of ΔSF (see Eq. (5.1)) for lineages initiated by cells in different phases of the cell-cycle.

survival of lineages instead of cells. In particular we will discuss the survival of G_1-, S- and G_2-lineages that correspond to lineages initiated, respectively, by a G_1, S and G_2 cell. We neglect the C_1 and D_2 lineages since our working assumption is that cells are initially sampled from a population growing in the regime of unperturbed exponential growth.

The results for synchronised (-/x) cells are presented in Fig. 5.10. In Fig. 5.10a we show survival curves in constant hypoxia. The model predicts that the G_1 and G_2 lineages are effectively insensitive to hypoxia while the S lineages have a significantly lower chance of survival. This is because S cells inevitably experience replication stress as they are synthesising DNA when exposed to hypoxia. In contrast, activation of the G1 checkpoints protects G_1 and G_2 cells from initiating DNA synthesis during hypoxia and forces them to wait until reoxygenation. When looking at survival in cyclic hypoxia (see Fig. 5.10b), the model predicts significant differences. While cyclic hypoxia also impacts the survival of S-lineages, it does so to a lower extent than constant hypoxia. In contrast, cyclic hypoxia reduces the survival of the G_1- and G_2-lineages, which are predicted to be sensitive to cyclic hypoxia but not to constant hypoxia. This is consistent with the results from the previous section which suggested that a poorly functional G1 checkpoint is not sufficient to protect cells from cyclic hypoxia. Indeed, the ease with which cells can initiate DNA synthesis by transitioning to the S phase is also important. The difference between the sensitivity of cells synchronised in the G1 and G2/M phases also has an intuitive explanation. Based on our modelling assumptions, cells initially in the G_2 phase are certain to proliferate at least once, even if they are exposed to hypoxia. Consequently, the survival probability of a G_2-lineage is effectively equivalent to the probability that at least one of two G_1-lineages will survive, which is significantly lower than the

probability of a single G_1-lineage surviving. Overall, our results illustrate that, even within the same population, inherent heterogeneity in the distribution of cells along the mitotic cycle can impact the outcome of clonogenic assays. Therefore, whenever quantifying cell survival under hypoxia, it is important to take into consideration the initial cell-cycle distribution of the cell population used.

In Fig. 5.10c we compare the value of $\Delta\mathrm{SF}(t)$ (defined in Eq. (5.1)) for lineages initiated by cells in different cell-cycle phases. We recall that if $\Delta\mathrm{SF}(t) > 0$ then cells are more likely to survive exposure to hypoxia if this is periodic rather than constant, while the opposite holds when $\Delta\mathrm{SF}(t) < 0$. Our model predicts that, for all three types of lineages, the sign of $\Delta\mathrm{SF}(t)$ is independent of t. In particular, we find that cells in the S-lineage are consistently more sensitive to constant hypoxia (and conversely for cells synchronised in G1 or G2/M). Looking back at the profile of $\Delta\mathrm{SF}$ when considering an asynchronous population (purple curve), it is now apparent that the sharp change in its sign between $t = 8$ and $t = 12$ hours is due to intra-specific heterogeneity between lineages caused by the asynchronous cell-cycle distribution of cells from the same cell-line and initial culture.

5.5 Discussion

While in Chapter 4 we studied how cell-cycle dysregulation in hypoxia affects cell proliferation and the overall growth dynamics cell cultures, in this chapter the focus has been on characterising the impact of cell-cycle dysregulation in hypoxia on cell survival. To do so, we developed a computational model to simulate clonogenic assays that enables us to study the effect of different forms of hypoxia on cancer cell survival. Our computational framework is based on a stochastic, individual based model (IBM) that is derived from the deterministic model proposed in Chapter 2. As such, it captures the cell-cycle dependent effect of hypoxia on cell viability and proliferation. While the deterministic model was suitable for studying population-level cell-cycle dynamics, its stochastic counterpart enables us to characterise individual cells.

While the ensemble average of the stochastic model is well approximated by the deterministic model, single realisations of the stochastic model can yield significantly different predictions on the long-term growth dynamics of cell cultures initially consisting of only a few cells. From this point of view, a stochastic framework, in contrast to a deterministic description, is more suitable for making predictions about cell survival in hypoxia, where randomness can play a major role (see the example presented

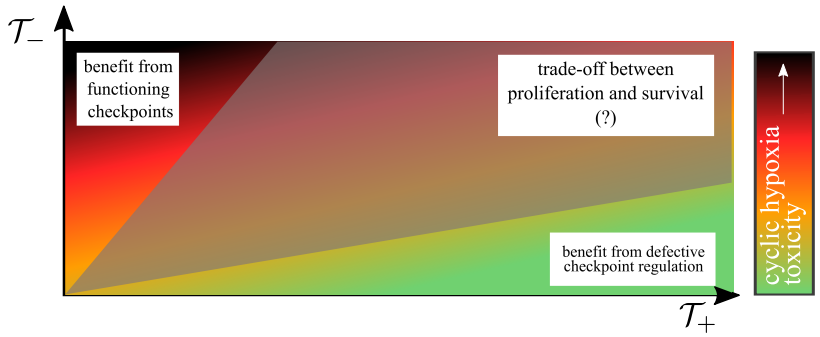


Figure 5.11: Schematic illustrating under which cyclic hypoxia environment cells might benefit from having functioning or defective checkpoints. As in Fig. 2.12, toxicity is defined as the level to which proliferation of cancer cells is impaired under cyclic hypoxia. The grey area indicates the intermediate regime between very toxic and mildly toxic conditions where it is unclear what the strategy is favourable for cells.

in Section 5.3.1). Interestingly, we found that the stochastic model predicts the existence of a trade-off between fast proliferation (at the population level) and survival (at the single-cell level) in cells exposed to certain cyclic hypoxia environments. More specifically, defective cell-cycle control mechanisms (or checkpoints) may yield faster growing populations in a cyclic hypoxia environment; however, this can be at the cost of significantly lowering cell survival. Our findings are illustrated in Fig. 5.11, where we update the diagram first presented in Chapter 2 (see Fig. 2.12) to combine information on proliferation and survival in cyclic hypoxia. While in severely toxic cyclic hypoxia environments cells are better off maintain at least partially functional cell-cycle checkpoints, the opposite is true in mildly toxic environments where defective checkpoints enables cells to grow faster. However, for intermediate regimes, where the model predicts a trade-off between survival and proliferation, it is unclear what cell-cycle control strategy would favour tumour initiation and progression. In particular, we find that cells with functional checkpoints are more likely to survive these intermediate conditions than cells that have defective cell-cycle checkpoints. However, when cells with defective cell-cycle checkpoints manage to survive and reach a critical population size (so that stochasticity is negligible), they yield tumour populations which proliferate faster. Overall, our findings suggest that cyclic hypoxia may contribute to intra-tumour heterogeneity by favouring the emergence of cell lineages with cell-cycle regulation mechanisms that are not selected for in region of normoxia and constant hypoxia.

Having investigated the predictions of the stochastic model, in Section 5.4 we used the IB model to develop a computational framework to simulate clonogenic as-

says. By generating an *in silico* cohort of cell-lines (as in Chapter 2), we used our framework to study survival curves for populations with different cell-cycle control mechanisms in different hypoxic environments: constant and cyclic hypoxia (specifically a (4,2)-cyclic hypoxia protocol). The model predicts significant heterogeneity in the estimated survival curves amongst our cohort of *in silico* cell-lines. Furthermore, in Section 5.4.3 we used our model to investigate intra-specific heterogeneity in the survival of cells with the same cell-cycle regulation mechanisms. We found that survival may significantly depend on how cells are initially distributed along the cell-cycle and that cells that are likely to survive exposure to constant hypoxia can still be highly sensitive to cyclic hypoxia because of the differential effects the two environmental conditions have on checkpoint regulation of cell-cycle transitions.

Here, we have focused on using our computational model to simulate clonogenic assays in hypoxia. A similar approach could be used to investigate the effect of exposure to cell damaging agents (or environmental factors) whose toxicity depends on the cell-cycle phase cells are in. In the next chapter, we will illustrate how this framework can be extended to simulate the effect of radiotherapy (RT) on cell survival and how our computational model can be used to understand the impact of cell-cycle dysregulation in hypoxia on the sensitivity of cancer cells to RT.

Chapter 6

Predicting radiotherapy responses of cells grown in cyclic hypoxia

6.1 Overview

In the previous chapter we introduced a computational approach to simulate clonogenic assays and showed how this can be used to study cell survival in hypoxia. In this chapter we use our framework to study how cell-cycle dysregulation in cyclic hypoxia affects cancer cell responses to radiotherapy. As mentioned in Section 1.1.4, there is *in vitro* and *in vivo* evidence that exposure to cyclic hypoxia can reduce RT efficacy. However, the mechanisms by which this occurs are still poorly understood. In Chapters 2 and 3, we combined numerical simulations with *in vitro* data to show how cyclic hypoxia perturbs cell-cycle progression. Our goal in this chapter is to augment the computational framework from Chapter 5 to study the extent to which dysregulation of the cell-cycle affects the sensitivity to RT of cells grown in cyclic hypoxia.

It is well known that sensitivity to RT depends on the cell-cycle state a cell is in. Since the late 1960s, experimentalists have investigated how sensitivity to RT changes as cells progress along the cell-cycle [112]. While the relation between cell-cycle and RT is likely to be cell-line and tumour specific, some general trends have been identified. In particular, cells are most RT-sensitive during the M phase and most RT-resistant during the S phase (in particular towards its end) [112]. Furthermore, there is evidence that checkpoint activation can protect cells from RT-mediated death [35, 95]. While arrested cells have the ability of repairing *potentially lethal damage* (PLD) [95] post irradiation, PLD leads to cell-death in cells that are actively cycling.

In Section 6.2 we explain how we augment the computational model proposed in Section 5.4 to simulate the impact of RT on the outcome of clonogenic assays.

In Section 6.3, we use our modelling framework to inform experimental designs that can properly quantify the extent to which cyclic hypoxia affects cell sensitivity to RT. In Section 6.4 we use our approach to determine the differential impact of cell-cycle dysregulation on RT outcomes for *in silico* cell-lines with and without PLD repair. We conclude by summarising our findings in Section 6.5.

6.2 Simulating clonogenic assays with RT

We model a standard clonogenic assay to estimate the impact of cyclic hypoxia on RT-survival as performed by our collaborators. As shown in Fig. 6.1, two independent experiments are run: a positive control where cells are irradiated after exposure to cyclic hypoxia, and a negative control where cells are only exposed to cyclic hypoxia. While the positive control has information on cell survival to both cyclic hypoxia and RT, the negative control is used to measure survival to cyclic hypoxia alone. Combining information from these two experiment, our collaborators aim to estimate the probability of a cell surviving RT given that it has survived pre-treatment in cyclic hypoxia.

In both experiments, N_0 cells from an asynchronously growing population are initially seeded in a new culture plate at time $t = 0$. Cells are cultured in a spatially well-mixed environment and below confluence so that competition for space and nutrient is negligible. Oxygen levels, $c = c(t)$, in the culture are externally prescribed. After being seeded cells are exposed to a given oxygen environment for t_R hours after which they are re-oxygenated to air oxygen levels (*i.e.*, $c = 21\%\ O_2$) and allowed to grow unperturbed for 10 days. Cells in the positive control experiment are irradiated immediately after re-oxygenation ($t = t_R^+$). After 10 days post re-oxygenation, our collaborators estimate RT-survival in cyclic hypoxia, SF_{RT}, as the ratio between the fraction of viable colonies (or lineages) in the positive ($n_f^{(+)}/N_0$) and negative ($n_f^{(-)}/N_0$) controls

$$\mathrm{SF}_{RT}(t_R) = \underbrace{\left(\frac{n_f^{(+)}(t_R)}{N_0}\right)}_{\substack{\approx \text{ probability a lineage survives} \\ \text{RT and cyclic hypoxia}}} \Big/ \underbrace{\left(\frac{n_f^{(-)}(t_R)}{N_0}\right)}_{\substack{\approx \text{ probability a lineage survives} \\ \text{cyclic hypoxia}}} . \tag{6.1}$$

We note that Eq. (6.1) can be viewed as an application of Bayes' Theorem (see Section 1.3.2) for conditional probabilities where SF_{RT} estimates the probability of a lineage that survives exposure to cyclic hypoxia for t_R hours to survive RT. In

writing Eq. (6.1), it is assumed that, for the whole duration of the experiment, each initially seeded cell gives rise to an independent colony (*i.e.*, we neglect the possibility of colonies merging or fragmenting).

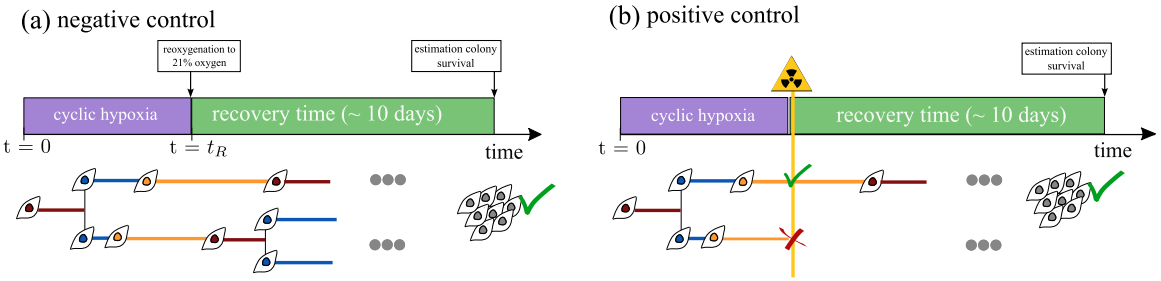


Figure 6.1: Schematic illustrating the standard protocol used by our collaborators to estimate the impact of cyclic hypoxia on RT-survival. This comprises two experiments: (a) negative control – cells are not irradiated; (b) positive control – cells are irradiated. Survival is then estimated as the ratio between the survival fractions from experiment (b) and from experiment (a).

6.2.1 Modelling effect of RT on lineage survival

As in Chapter 5, lineages consist of cells that can exists in any of 6 cell-cycle states: G_1, C_1, S, G_2, D_2 and Sn_2. Cells update their state according to the procedure summarised in Fig. 6.2. This follows the IB model presented in Chapter 5 (see Fig. 5.3) with an additional subroutine to capture the impact of radiotherapy (details in the next section). Consistently with experiments, at the final time $t_f = t_R + 10$, a lineage is classified as viable if it consists of at least 50 cells.

6.2.1.1 Cell-cycle dependent RT-survival

For simplicity, we assume that irradiation alters a cell viability instantaneously. As in Chapter 5, a cell is defined as viable if it has the potential to proliferate. Since senescent cells (*i.e.*, Sn_2 cells) have permanently withdrawn from the cell-cycle, they are considered as non-viable and they are unperturbed by RT. Considering a cell which is viable at the time of irradiation, $t = t_R^+$, we model the impact of a single dose of radiotherapy as a Bernoulli process with probability dependent on a cell cell-cycle state; specifically, either a cell in cell-cycle state z ($z \in \mathcal{Z}_v = \{G_1, C_1, S, G_2, D_2\}$) remains viable with probability $S_{RT}(z)$, or it loses the ability to proliferate with probability $1 - S_{RT}(z)$. If a cell loses viability due to RT, it is simply removed from the lineage. In general, the survival probability S_{RT} will depend on several biological and environmental factors, such as the dose of RT and the oxygen levels at the

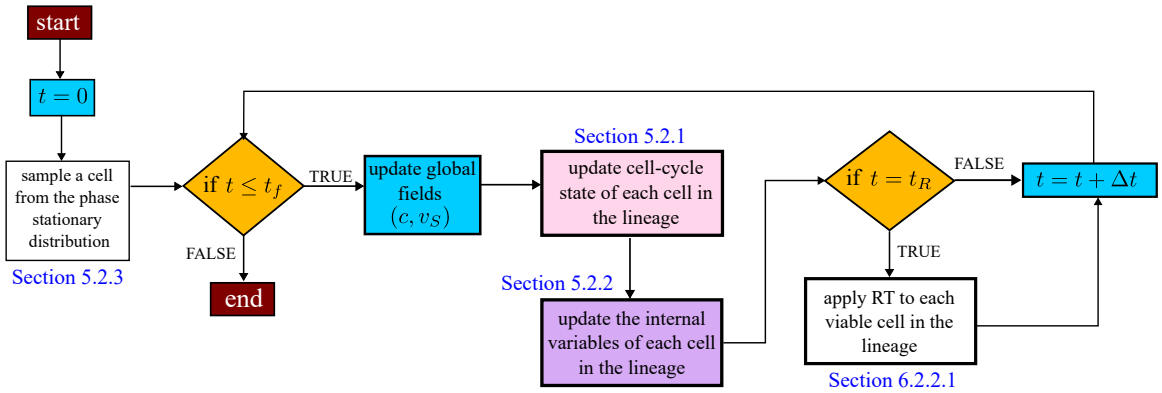


Figure 6.2: Flowchart illustrating the procedure underlying the computational implementation of the stochastic IB model for each lineage. We indicate the section of the thesis where details of the implementation of each subroutine can be found. Here, the constants t_R, t_f and the functions $c(t)$ (oxygen levels) and $v_S(t)$ (average DNA synthesis) are inputs to the model.

moment of irradiation. For simplicity, we will keep these factors constant across our simulations.

We capture heterogeneity in cell responses to RT by splitting our cell population into sensitive ($\mathtt{s}$) and resistant ($\mathtt{r}$) cells. Our approach to model heterogeneous responses to RT is naive. However, our model could be easily extended to include more than two cell sub-populations to capture more refined differences in RT-sensitivity of cells as they move along the mitotic-cycle. In the absence of detailed information, we have opted for the simplest approach. We denote the probabilities of survival of sensitive and resistant cells after being irradiated by $\mathbb{P}_{\mathtt{s}}$ and $\mathbb{P}_{\mathtt{r}}$, where, by definition $\mathbb{P}_{\mathtt{s}} < \mathbb{P}_{\mathtt{r}}$. We then associate to each type of viable cell in our model, $z \in \mathcal{Z}_v$, either the label $\mathtt{s}$ (sensitive) or $\mathtt{r}$ (resistant); this is formalised by introducing the map $\varphi : \mathcal{Z}_v \to \{\mathtt{s}, \mathtt{r}\}$. Then, given the cell-cycle state z, its probability of survival is $S_{RT}(z) = \mathbb{P}_{\varphi(z)}$.

Based on findings from the literature (see Section 6.1 and references therein), we explore two alternative hypotheses of how the cell-cycle affects RT efficacy:

a. Without potentially lethal damage repair. Cells in the S phase are resistant to RT, while all other cells are sensitive: *i.e.*, $\varphi(S) = \mathtt{r}$ and $\varphi(z) = \mathtt{s}$ for $z \in \{G_1, C_1, G_2, D_2\}$.

b. With potentially lethal damage repair. Cells in the S phase and in checkpoint states (*i.e.*, C_1 and D_2 cells) are resistant to RT, while other cells are sensitive: *i.e.*, $\varphi(z) = \mathtt{r}$ for $z \in \{C_1, S, D_2\}$ and $\varphi(z) = \mathtt{s}$ for $z \in \{G_1, G_2\}$.

As we will see, these two different hypothesis can yield different conclusions on how cyclic hypoxia contributes to RT resistance by perturbation cell progression along the cell-cycle.

6.2.2 Alternative estimates of survival

Following the experimental protocols used by our collaborators, cells are seeded before exposure to cyclic hypoxia and RT, and therefore survival fractions are associated with lineages rather than individual cells. In order to estimate the impact of cyclic hypoxia on cell sensitivity we introduce a second survival metric $\mathrm{SF}^{\diamond}_{RT}$. This is estimated as follows:

$$\mathrm{SF}^{\diamond}_{RT}(t_R) = \left(\frac{n_f^{(+,\diamond)}(t_R)}{n_R^{(+,\diamond)}(t_R)}\right) \Big/ \left(\frac{n_f^{(-,\diamond)}(t_R)}{n_R^{(-,\diamond)}(t_R)}\right). \tag{6.2}$$

In Eq. (6.2), $n_R^{(+,\diamond)}$ and $n_R^{(-,\diamond)}$ correspond to the total number of cells in the culture at time $t = t_R$ (immediately before applying RT) in the positive and negative control experiments. The variables $n_f^{(+,\diamond)}$ and $n_f^{(-,\diamond)}$ correspond instead to the number cells (of the $n_R^{(+,\diamond)}$ and $n_R^{(-,\diamond)}$ cells present in the culture at time $t = t_R$), that have a total of ≥ 50 offspring at the end of the experiment ($t_f = t_R + 10$ days). Note that, by definition, cell survival is always less than or equal to lineage survival, *i.e.*, $\mathrm{SF}_{RT}(t_R) > \mathrm{SF}^{\diamond}_{RT}(t_R)$. While $\mathrm{SF}^{\diamond}_{RT}$ can be easily estimated computationally, it is more difficult to measure it experimentally. Practically, this would require an experimentalist to run a clonogenic assay where cells are initially grown for t_R hours in cyclic hypoxia, then plated in the assay and irradiated. However, based on discussions with our experimental collaborators, it is not advisable to follow this alternative protocol experimentally as it would not allow them to compare different experimental conditions (for example when testing different cyclic hypoxia protocols, *i.e.*, t_R), since seeded cells would have to be sampled from different solutions. From this point of view, our computational model can provide novel insights by connecting the experimentally measurable quantity, SF_{RT}, to the more accurate metric of survival, $\mathrm{SF}^{\diamond}_{RT}$, that may be not directly assessable in the lab.

6.2.3 Numerical implementation

We implement our IB model of clonogenic assays in Python and use it to estimate lineage and cell RT-survival for cultures grown in cyclic hypoxia for variable durations, t_R, prior to irradiation. Oxygen levels in the time interval $[0, t_R]$ fluctuate between $c_- = 0.1\%\ O_2 < c_H$ and $c_+ = 2.1\%\ O_2 > c_H$ according to Eq. (2.31); for $t \in [t_R, t_f]$,

where $t_f = t_R + (10 \text{ days})$, c is simply set to $c = 21\% O_2 \gg c_H$. As in other thesis chapters, we denote by $(\mathcal{T}_-, \mathcal{T}_+)$-cyclic hypoxia a protocol in which cells are repeatedly exposed to hypoxia (*i.e.*, $c < c_H$) for $\mathcal{T}_-$ hours followed by re-oxygenation (*i.e.*, $c = c_+ > c_H$) for $\mathcal{T}_+$ hours. We set the RT survival probabilities $\mathbb{P}_{\mathtt{s}}$ and $\mathbb{P}_{\mathtt{r}}$ to arbitrary values: $\mathbb{P}_{\mathtt{s}} = 0.1$ and $\mathbb{P}_{\mathtt{r}} = 0.5$. This is a reasonable assumption since in all our simulations cells are re-oxygenated to $21\% O_2$ before being irradiated and exposed to the same RT dose. In order to compare experiments where cells are irradiated at different oxygen levels and/or exposed to different doses, a more refined model for $\mathbb{P}_{\mathtt{s},\mathtt{r}}$ is needed to account for the impact of these two factors on cell survival [102]. We simulate responses to RT for cells with functional checkpoints, *i.e.*, (+/+) cells, by considering parameter values in Table 2.3, except for the death rate of cells in the S phase which we set to $\mu_S = 0.025$ (as in Chapter 5). Initial conditions are as given in Section 5.2.3.

Since SF_{RT} is estimated via stochastic simulations, it is itself a random variable. We therefore estimate information on its mean and distribution by generating 100 realisations of the same clonogenic assay. For each realisation, we simulate the positive and negative control experiments by generating $N_0 = 2500$ lineages following the procedure summarised in Fig. 6.2. Note that for the negative control RT is not applied.

6.3 Comparing RT-survival in lineages and cells pre-treated with cyclic hypoxia

In Fig. 6.3, we report cell and lineage RT-survival estimates (denote by SF_{RT} and $\text{SF}^{\diamond}_{RT}$, respectively) obtained when growing the same cell-line for t_R hours to two different cyclic hypoxia protocols: either (2,4)-cyclic hypoxia (Fig. 6.3a) or (4,2)-cyclic hypoxia (Fig. 6.3b). This example elucidates when and why we expect differences between cell and lineage RT-survival in cyclic hypoxia.

The difference between SF_{RT} and $\text{SF}^{\diamond}_{RT}$ is apparent from both panels in Fig. 6.3. Focusing on panel (a), we find that $\text{SF}^{\diamond}_{RT}(t_R)$ and $\text{SF}_{RT}(t_R)$ follow an opposite trend. While $\text{SF}^{\diamond}_{RT}(t_R)$ slightly decreases with t_R (*i.e.*, cells grown for longer in cyclic hypoxia are more sensitive to RT), $\text{SF}_{RT}(t_R)$ significantly increases with t_R. Since (2,4)-cyclic hypoxia is mildly toxic for cells, these continue proliferating almost unperturbed in the period $t = [0, t_R]$. As a result, the average size of a lineage at the time of irradiation ($t = t_R^+$) increases with t_R and this compensates for the fact that individual cells in the lineage are more sensitive to RT the longer they are exposed to (2,4)-cyclic hypoxia.

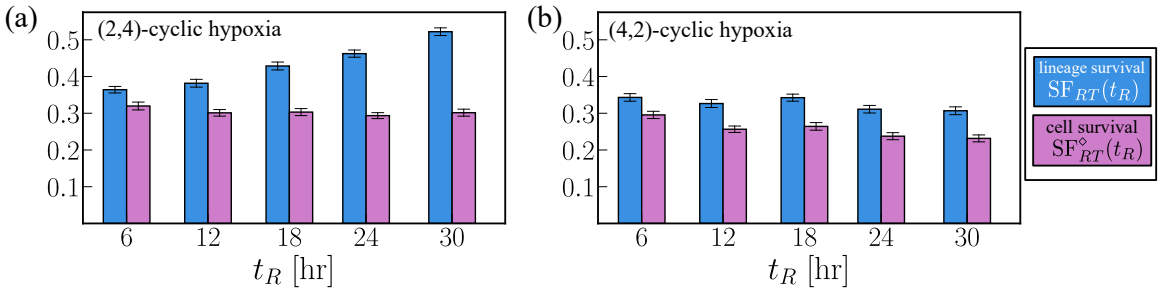


Figure 6.3: (a) Comparison of RT-survival metrics, SF_{RT} (Eq. (6.1)) and $SF^{\diamond}_{RT}$ (Eq. (6.2)), for cell cultures grown for 6, 12, 18, 24 or 30 hours in (2,4)-cyclic hypoxia. Survival metrics are estimated as detailed in Section 6.2.1 neglecting potentially lethal damage repair (*i.e.*, hypothesis A in Section 6.2.1.1). The bar-plot indicates the mean over 100 realisation of the clonogenic assays, while the error bar indicates the related variances. Parameters values are taken as defined in Section 5.2.3. (b) Same as (a) but cells are grown for 6, 12, 18, 24 or 30 hours in (4,2)-cyclic hypoxia prior to irradiation.

While (2,4)-cyclic hypoxia is mildly toxic for cells, (4,2)-cyclic hypoxia greatly reduces cell proliferation. In this case, the average size of a lineage remains approximately constant as t_R increases. As a result, while cell and lineage survivals are slightly different, they both decrease as cells are pre-treated with cyclic hypoxia for longer periods. We therefore conclude that, lineage survival is not always a good metric to estimate the impact of toxic agents (such as cyclic hypoxia) on cell sensitivity to RT; particularly in those experiments where cells are exposed to low levels of toxicity and proliferation is not halted.

6.4 The impact of the potentially lethal damage repair on RT-survival

Having presented our general framework, we use it to investigate how potentially lethal damage repair affects sensitivity to RT in cells grown in cyclic hypoxia.

In Fig. 6.4, we compare model predictions of cell-survival, $SF^{\diamond}_{RT}$, when growing cells with functional checkpoints in (4,2)-cyclic hypoxia for different durations ($t_R = [0, 6, 12, 18, 24, 30, 36]$ hours) prior to being irradiated. In Fig. 6.4a (as in Fig. 6.3b) predictions of survival are obtained assuming that only cells in the S phase are radioresistant (see hypothesis A in Section 6.2.1.1). Overall, we find that exposure to cyclic hypoxia slightly increases cell sensitivity to RT. This is because of activation of the G1 checkpoint when cells are grown in cyclic hypoxia. As cells arrest in C_1 checkpoint (where they are sensitive to RT), the fraction of resistant S

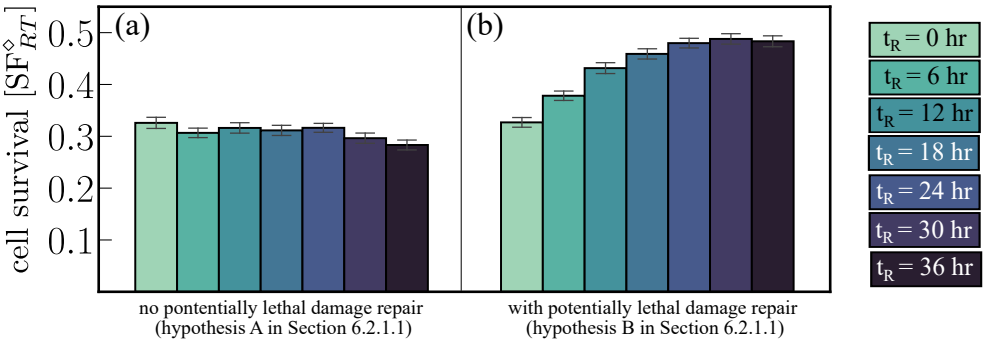


Figure 6.4: (a) Predicted survival to RT, $\mathrm{SF}^{\diamond}_{RT}$ (see Equation (6.2)), for cells grown in (4,2)-cyclic hypoxia for $t_R = [0, 6, 12, 18, 24, 30, 36]$ hours when neglecting potentially lethal damage repair (see hypothesis A in Section 6.2.1.1). The bar-plot indicates the mean value of $\mathrm{SF}^{\diamond}_{RT}$ over 100 realisation of the clonogenic assays, while the error bar indicates the related variances. Details on the set up of the simulations and parameter values are given in in Section 6.2.3. (b) Same as (a) but accounting for potentially lethal damage repair in checkpoint compartments (see hypothesis B in Section 6.2.1.1).

cells in the population decreases. At the population level, this redistributions of cells along the cell-cycle result in increased sensitivity to RT. However, this is not true when accounting for potentially lethal damage repair (see Fig. 6.4b). In this case, prolonged exposure to cyclic hypoxia decreases cell sensitivity to RT. This is because, when modelling potentially lethal damage repair, C_1 cells are also assumed to be resistant to RT. In this case, the accumulation of C_1 cells in cyclic hypoxia increases, instead of decreasing, cell RT-survival. We therefore conclude that activation of cell-cycle checkpoints in cyclic hypoxia can protect cells that allow for potentially lethal damage repair from RT-mediated death.

6.5 Discussion

Experimental evidence suggests that cells grown in cyclic hypoxia acquire resistance to radiotherapy (RT) both *in vivo* [146] and *in vitro* [4, 67, 83]. However, little is known about the biological mechanisms by which cyclic hypoxia increases cell resistance to treatment. As discussed in Section 1.1.4, cell-cycle dysregulation is known to impact RT efficacy. In this chapter, we investigated how cell-cycle dysregulation in cyclic hypoxia affects cells sensitivity to RT. To do so, in Section 6.2, we augmented the computational model from Chapter 5 to simulate the outcome of clonogenic assays in which cells that grown in cyclic hypoxia are irradiated. A key feature of our model is that RT-killing is assumed to depend on a cell cell-cycle state. In Section 6.3,

we used our framework to show how and under which conditions, standard *in vitro* clonogenic assays used by our collaborators may fail to quantify the impact of cyclic hypoxia on cells' ability to survive RT. From this point of view, our computational model can be applied alongside experimental data to better understand the impact of cyclic hypoxia on RT outcomes.

It has been observed that checkpoint activation can protect cells from RT-induced death enabling them to repair potentially lethal damage. In Section 6.4, we used our computational framework to study how potentially lethal damage repair impacts RT-survival in cells grown in cyclic hypoxia. When we assumed checkpoint activation shields cells from RT-induced death, we found that cells are more resistant to RT the longer they are exposed to cyclic hypoxia. In contrast, when assuming that checkpoint activation does not affect cells radio-sensitivity, cell-cycle dysregulation in cyclic hypoxia increased cell sensitivity to RT. Overall, our results suggest that cell-cycle dysregulation in cyclic hypoxia can impact RT-efficacy. However, more work and experimental data are needed to predict the extent to which cell-cycle dysregulation plays a role in RT responses of cancer cells exposed to cyclic hypoxia *in vivo*.

Chapter 7

Conclusions and Future Work

Recent advances in imaging techniques are allowing for a more detailed analysis of the spatio-temporal dynamics of oxygen in *in vivo* mouse models. This reveals that cyclic hypoxia is present in several types of solid tumours [4, 99]. Nonetheless, little is known on how cyclic hypoxia contributes to tumourogenesis and tumour resistance to treatment. Limitations in the ability to generate *in vitro* models of cyclic hypoxia are currently hindering progress in understanding this phenomenon. While some of these challenges can be overcome by technological advances, the level of heterogeneity in how oxygen levels fluctuate in different tumours, or even different regions of the same tumour, prevents us from designing an *in vitro* model of cyclic hypoxia that can properly account for such variability. In this thesis we have shown how combining mathematical and *in vitro* models can overcome these limitations and, by doing so, offer unique insights on how different hypoxic environments affect cancer cell functions and response to radio-therapy. In what follows, we summarise the results of Chapters 2–6 and then outline ideas for future work.

Our focus has been on developing a novel mathematical model of cell-cycle dysregulation in cancer cells cultured in different hypoxic environments and on investigating how this impacts on other important cellular functions; specifically: cell culture growth and survival in the presence and absence of radiotherapy. In Chapter 2 we developed a deterministic structured PDE model of the cell-cycle that describes how the number of cells in different cell-cycle states evolves over time. We showed how the model can be reduced to a system of non-autonomous differential equations with 2 time-dependent delays using the method of characteristics. Leveraging numerical methods, we used the model to simulate the growth and cell-cycle dynamics of cancer cell cultures exposed to different hypoxic environments. In doing so, we also illustrated how cell-cycle regulation mechanisms affect response to hypoxia. In Chapter 3 we calibrated our model using flow cytometry data describing how hypoxia affects the

cell-cycle dynamics of RKO colorectal cancer cells [5]. In addition to showing that our cell-cycle model can more accurately describe the available experimental data than existing models, we illustrated how it can also be used to aid experimental design.

Having validated our modelling framework, in Chapters 4–6 we used it to investigate the impact of cell-cycle dysregulation on the growth and survival of cell culture exposed to different forms of hypoxia. In Chapter 4 we studied the long-term effect of cyclic hypoxia on cell culture growth dynamics by leveraging the theory of linear positive dynamical systems. Our results illustrate how the growth dynamics of cell cultures in cyclic hypoxia depends on how the extent to which cancer cells maintain cell-cycle checkpoint regulation. When cells are exposed to hypoxia infrequently and for brief periods (low toxicity), they benefit from complete loss of checkpoint regulation. In contrast, in highly toxic environments (frequent exposure to long periods of hypoxia), as in constant hypoxia, checkpoint integrity is advantageous, allowing cell cultures to continue growing, even though they do so at a much slower rate. In intermediate regimes, we find that partial loss of the checkpoint regulation yields cell cultures with higher growth rate. While illustrating the complexity of possible cell responses to cyclic hypoxia, our findings show how the heterogeneous time-evolution of oxygen levels in different tumour regions might drive phenotypic diversity in tumours by selecting for cells with different cell-cycle control mechanisms.

In Chapter 5 we investigated how cell-cycle dysregulation impacts cancer cell survival in different hypoxic environments. To do so, we recast our cell-cycle model in a stochastic, individual-based framework, that allows us to follow cell-cycle progression and fate of individual cells and their progeny (*i.e.*, lineages). Our results illustrate how, when cells are exposed to hypoxia infrequently and for brief periods (low toxicity), their risk of losing viability is negligible. However, for intermediate to highly toxic cyclic hypoxia environments, cancer cell survival can be significantly reduced unless cells maintain functional cell-cycle checkpoint regulation. Combining these results with those from Chapter 4 we found that, while proliferation-driven cancer cells thrive in slightly-toxic cyclic hypoxia, survival-driven cancer cells have an advantage in highly-toxic hypoxic conditions. For intermediate regimes, however, it is unclear what strategy would most favour cells as they experience trade-offs between proliferation and survival. In the second part of Chapter 5 we showed how our IBM model can be used to simulate the impact of hypoxia on clonogenic assays that are commonly used to estimate cancer cell survival under exposure to toxic agents/environments. In Chapter 6 we augmented our computational model of clonogenic assays to study the impact of cell-cycle dysregulation in cyclic hypoxia on cancer cell radio-resistance.

While experimental data are not yet available for this model, we showed how it could be combined with data from *in vitro* clonogenic assays to quantify the extent to which cell-cycle dysregulation can account for increased cancer cell radio-resistance. Our preliminary results highlight that cell-cycle redistribution during cyclic hypoxia is likely to impact cancer cell responses to radiotherapy; as such, the emergence of cells in tumours that adopt different cell-cycle regulation strategies might relate to the emergence of more radio-resistance phenotypes.

7.1 Future directions

There are several possibilities to extend the work presented in this thesis. Due to limited space, here we list some of them.

7.1.1 Experimental validation

While in Chapter 3 we validated our model against flow cytometry data describing cell-cycle dysregulation in RKO cells grown in different forms of hypoxia, we expect large variability amongst different cell-lines. To this end, it would be interesting to test our model against flow cytometry data from cell-lines that are known to have different cell-cycle checkpoint regulation. Collecting flow cytometry data from a wide range of cancer cell-lines would help also validating some of our theoretical predictions from Chapter 2 and 4 on the relation between the functionality of checkpoints and cell response to hypoxia. Another natural extension of the work is to couple the computational framework developed in Chapters 5 and 6 to simulate clonogenic assays with experimental data. While the cell-cycle is known to affect survival under several treatments, quantitative estimates of these effects are lacking. From this point of view, the computational framework developed in Chapters 5 and 6 could help to design experiments that would improve our current understanding and quantitatively test assumptions, such as the relevance of potential damage repair effect, on how RT-sensitivity depends on the cell-cycle state of cancer cells at the moment of irradiation.

7.1.2 To proliferate or to repair: a cell dilemma

The results presented in Chapters 2, 4 and 5 indicate that cyclic hypoxia can impose cells trade-offs between proliferation and survival depending on the extent to which cyclic hypoxia damages cancer cells. Prompted by this finding, it would be interesting to use structured population models to investigate the relationship between

damage dynamics and the extent of survival-proliferation trade-offs cell experience in different hypoxic environments. In our cell-cycle models, the role of damage has been included indirectly by relating delays in the S→G2 transition to the extend to which cells have been damaged by replication and reoxygenation stresses. While this simplifying assumption allowed us to develop a model that could be easily fitted to the experimental data available (see Chapter 3), a more refined model is needed to investigate how damage dynamics in hypoxia relate to cancer cells experiencing proliferation and survival trade-offs. To this end, we propose a novel deterministic structured-population model where cancer cells are structured according to their damage level, $y \geq 0$, which mediates cell proliferation, death and repair. To start with, for the sake of simplicity, details of the cell-cycle dynamics might be neglected and re-introduced in a later iteration of the model.

We assume that cells can exists in either of two states, P_0 and P_1, and denote by $u_0 = u_0(y,t)$ and $u_1 = u_1(y,t)$ respectively the number of cells of type P_0 and P_1 with damage level y at time t. The time-evolution of the distributions u_0 and u_1 is then described by the following system of coupled advection-reaction PDEs defined for time $t > 0$ and $y \in [0, \infty)$:

$$\frac{\partial u_0}{\partial t} + \frac{\partial V_0(y,t)u_0}{\partial y} = \lambda_0(y,t)u_0 - \mu_0(y)u_0 - \alpha_{0,1}(y)u_0 + \alpha_{1,0}(y)u_1, \tag{7.1a}$$

$$\frac{\partial u_1}{\partial t} + \frac{\partial V_1(y,t)u_1}{\partial y} = \alpha_{0,1}(y)u_0 - \alpha_{1,0}(y)u_1, \tag{7.1b}$$

which is closed by imposing proper boundary and initial conditions. In writing Eq. (7.1a), we assume that P_0 cells proliferate and die respectively at a rates $\lambda_0(y)$ and $\mu(y)$, dependent on a cell internal level of damage y. The internal damage level of P_0 cells changes at a rate $V_0 = V_0(y,t)$, which we model as the sum of two contributions: $V_0 = v_0(y) + \delta v(t)$. Here $v_0(y) \geq 0$ captures the natural rate at which cells accumulate damage during the cell-cycle, whereas $\delta v(t) \geq 0$ captures damage due to a toxic agent (such as hypoxia) to which cells are exposed to. In writing Eq. (7.1b), we assume that P_1 cells are quiescent; as such they do not proliferate or die but they can repair. Specifically, we may assume they can change their damage level at a rate $V_1 = V_1(y,t)$. Again we decompose V_1 into two terms $V_1 = -v_1(y) + \delta v(t)$, where $\delta v(t)$, as above captures the effect of a toxic agent. In the absence of toxic agents, P_1 cells have the ability to repair damage and as such we take $v_1(y) \geq 0$. We model cell plasticity by allowing cells to switch between the P_0 and P_1 states. This is captured by the blue terms in Eqs. (7.1). We assume cells switch from being of type P_0 to type P_1 and backwards at rates $\alpha_{0,1}(y)$ and $\alpha_{1,0}(y)$, which can be taken to dependent on

damage variable y. As mentioned above, the effect of hypoxia (or any other damaging agent) can be captured by the factor $\delta v(t)$ in the definition of V_0 and V_1 given above.

On the one hand, this modelling framework would allow us to theoretically explore how different forms of hypoxia impact the growth dynamics of populations of phenotypically-different cancer cells (similarly to the analysis presented in Chapter 4). For example, we envision comparing populations with different transition rates, $\alpha_{0,1}(y)$ and $\alpha_{1,0}(y)$ to model responses of survival-driven cells ($\alpha_{1,0} \gg \alpha_{0,1}$) and proliferation-driven cells ($\alpha_{1,0} \ll \alpha_{0,1}$) to different oxygen environments. On the other hand, this model could provide novel insight on the dynamics of damage accumulation for a given cell phenotype in different hypoxic environments. From this point of view, the model could be employed to study the relationship between cyclic hypoxia and genetic instability (*i.e.*, higher mutation rate). While our experimental collaborators have shown that (2,2)-cyclic, more than constant, hypoxia contributes to genetic instability in RKO cell-line [5], it is not clear how their findings generalise to different forms of hypoxia. Eqs. (7.1) can help us understand how damage accumulates in cells exposed to different hypoxia environments.

While in the first attempt, a single population might be considered, following a similar approach to [134], it would be interesting to include competition in the model and consider co-cultures of cells with different phenotypes, such as proliferation- and survival-driven cells, to investigate how evolution affects the culture composition in different hypoxic environments.

7.1.3 Spatial modelling

In this thesis, we have focused on developing mathematical models to study cell-cycle dysregulation in 2D cell cultures. While cell cultures are commonly used experimentally to study cancer biology, it is more apparent that the latter are not a realistic *in vivo* model of tumours. This is particularly true when considering models of hypoxia: while oxygen levels are spatially-homogeneous in 2D culture, *in vivo* they are both temporally and spatially-heterogeneous. As a result, we expect regions within the same tumour to be exposed to different modes of oxygen fluctuation, which can lead to the development of different niches which may favour specific phenotypes. For example, in the work I have done in collaboration with my supervisor Prof. Helen Byrne and Prof. Panos Kevrekidis [30], we have shown how spatial heterogeneity in oxygen levels can have a significant impact intra-tumour heterogeneity and a tumour growth dynamics. Interestingly, we found that the spatially-resolved model from [30]

presented dynamics that can not be captured by well-mixed models of 2D cell cultures (investigated in our previous publication [31]).

An interesting extension of the cell-cycle model developed in Chapter 2 would be to account for both spatially- and temporally-heterogeneous oxygen environments. Following [23, 104, 105], we might start by modelling multicellular tumour spheroids for which spatially-resolved cell-cycle data can be obtained using FUCCI (see Section 1.2).

A standard mathematical approach to describe tumour spheroid growth is based on the widely applied model by Greenspan [60]. While the simplicity of this model, that allows for analytical tractability, is a positive feature in several contexts [105], this limits its application to study spheroid growths in fluctuating oxygen environments. This is because Greenspan's framework assumes that cell behaviour depends only on the current oxygen levels that cells experience. However, as highlighted in Chapter 3, when considering fluctuating oxygen levels, cells do not always respond instantaneously to oxygen levels. To this end, single- or multiphase approaches might be a more effective way to introduce spatial effects in our cell-cycle model [27, 55, 73, 89]. Within this approach, a tissue is decomposed into distinct phases (*e.g.*, tumour cells, blood vessels, extracellular fluid) and each phase might consist of different sub-populations. Mass and momentum balance are applied to each phase and the system is closed by making constitutive assumptions about their properties. The resulting models typically comprise mixed systems of PDEs on a growing domain, with physically motivated spatial fluxes. While in our well-mixed scenario global fields are used to describe the time-evolution of expression levels of proteins, m, and of the rate of DNA synthesis v_S, this approach would not be possible once spatial effects are included. These would instead have to be treated either as structure variables (if cell movement is included), or as local fields that depednd both on space and time. Given the complexity and computational cost of such a model, an individual-based modelling framework might be more appropriate. From this point of view, the stochastic model developed in Chapter 5 could be augmented to account for space using an overlapping sphere model, such as in [26], or a less computationally-expensive approach where cells are modelled as point clouds as in [104]

Regardless of the approach used to build a spatially-resolved model of the cell-cycle dysregulation in tumour spheroids exposed to fluctuating oxygen environments, we expect such models will yield important new insights into how cyclic hypoxia affects tumour development and intra-tumour heterogeneity. Another extension would be to

include radiotherapy to understand how spatio-temporal variation in tumour oxygen levels impact their response to treatment.

7.2 Final Remarks

We developed and analysed a new mathematical model of the cell-cycle that accounts for delays in cell-cycle transition in temporally evolving oxygen environments. Furthermore, we used this framework to investigate how cell-cycle dysregulation in cyclic hypoxia impacts cell-culture growth dynamics and survival in the presence and absence of treatment (radiotherapy). In doing so, we leveraged a deterministic and individual-based versions of the model showing how they can be used to simulate different *in vitro* experiments. While we have focused on the application of this framework to cyclic hypoxia, our modelling approach could be extended to describe cell-cycle dysregulation in cancer cells due to periodic exposure to other damaging agents. This, for example, is relevant to study treatment strategies that apply drugs following an on/off basis. We believe our approach offers a useful tool to help in the design and interpretation of *in vitro* experiments, as well as their translation to *in vivo* tumours, where oxygen levels are highly temporally- and spatially-heterogeneous.

Appendix A

Calibration to flow cytometry data

A.1 Comparison of different observation models

Here we present additional results from the fitting procedure described in Section 3.3.3 in the main text. In Fig. A.1, we show the estimated posterior distributions obtained from fitting to experimental data (see Fig. 3.2) model $\mathcal{M}$ in Section 3.2 combined with the Beta observation model introduced in Section 3.3.3. When we compare the marginal posterior distributions for the Beta (see bottom row Fig. A.1) and Gaussian (see top row Fig. A.1) observation models, we find that the two agree. We compare the performance of the two models using Bayesian model selection (detail in Section 1.3.3). The values of the DIC and the WAIC score for the two observation models are stated in Table A.1. While the smallest DIC score corresponds to the Beta noise model, the smallest WAIC score is recorded for the Gaussian noise model. Nonetheless, the differences in the DIC and WAIC scores between the two models are so small that, regardless of which information criteria we adopt, we conclude that the two models are equally good at describing the available data.

	$\mathcal{M}$ + Gaussian noise	$\mathcal{M}$ + Beta noise
k_{DIC}	5.85	4.25
$\texttt{PA}_{DIC}$	65.99	65.78
DIC	-120.27	-123.06
k_{WAIC}	3.49	3.32
$\texttt{PA}_{WAIC}$	66.00	65.66
WAIC	-125.02	-124.67

Table A.1: Comparison based on the deviation information criterion (DIC) and widely-applicable information criterion (WAIC) of two observation models (see Section 3.3.3 for more details). Both information criteria are defined as $-2(\texttt{PA} - k)$, where PA measures the predictive accuracy of a model, and the correction term k (known as effective number of parameters) accounts for overfitting.

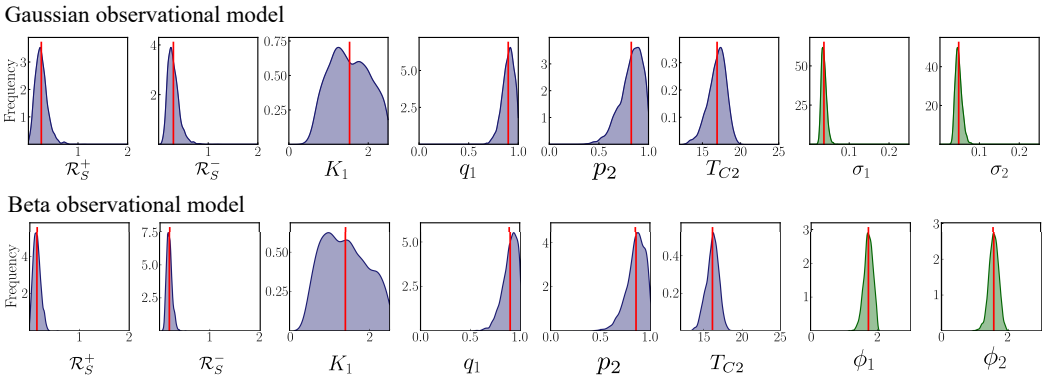


Figure A.1: Marginal posterior distributions for model $\mathcal{M}$ with the Gaussian (top) and Beta (lower) observation model (see Section 3.3.3). The distributions were obtained by using the MCMC samples generated as discussed in Section 3.3.1. The red vertical lines indicate the mean of the marginal posterior distribution as reported in Table A.2, where additional summary statistics extrapolated from the marginal distribution are also given. The hyper-parameters $\phi_{1,2}$ characterise the shape of the Beta distribution (*i.e.*, the error model) as described in Section 3.3.3. Note that in the last row the y-axes are scaled differently.

A.2 Comparison of different mechanistic models

In Fig. A.2, we show the estimated posterior distributions for our class of models $\mathcal{S}_{\mathcal{M}}$ (see Table 3.1). Summary statistics of the marginal posterior distributions are listed in Table A.2. For the point estimates of parameter values in §3.4.1, we use the mean of the marginal posterior distributions (see Table A.2).

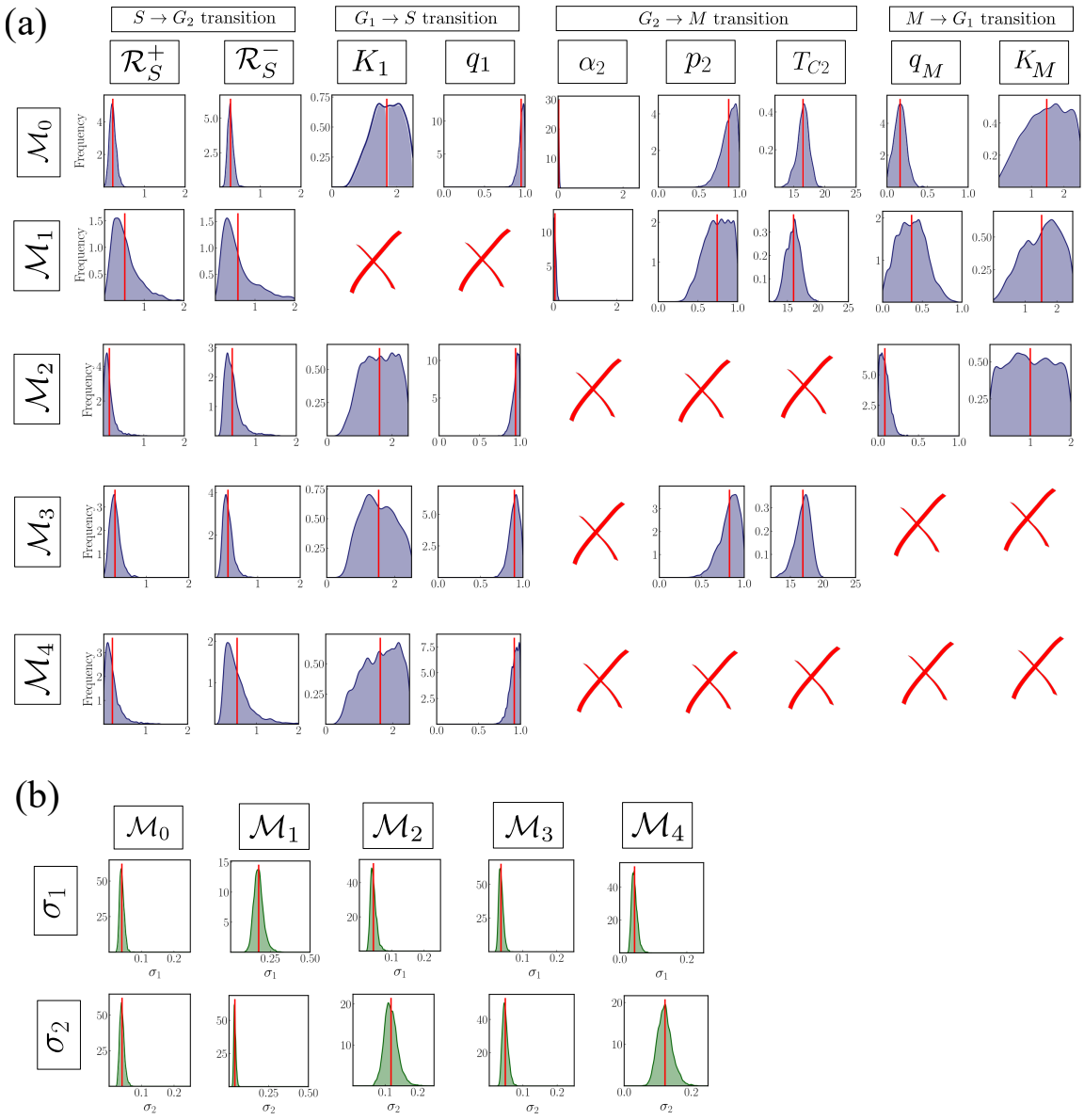


Figure A.2: Marginal posterior distributions for model parameters (a) and hyper-parameters (b) in models $\mathcal{M}_0 - \mathcal{M}_4$ from Table 3.1. In panel (a), we divide model parameters into groups depending on the biological mechanisms they relate to. Red crosses indicate that a parameter is not present in the model. The distributions were obtained by using the MCMC samples generated as discussed in the text above. The red vertical line indicates the mean of the marginal posterior distribution as reported in Table A.2, where additional summary statistics extrapolated from the marginal distribution are also given. Note that the y-axes are scaled differently.

Table A.2: Summary statistics for the marginal posterior distributions for the family of models $\mathcal{S}_{\mathcal{M}}$ listed in Table 3.1 in Section 3.4. Here we report the means and standard deviations together with the quantiles (Q_i) of the marginal posterior distributions. The final column shows the value of $\hat{R}$ used to estimate the convergence of the MCMC algorithm, where convergence corresponds to $\hat{R} \leq 1.05$.

Model		**mean**	std	$Q_{25\%}$	$Q_{50\%}$	$Q_{75\%}$	$\hat{R}$
$\mathcal{M}_0$	$\mathcal{R}_S^+$	0.22	0.08	0.16	0.21	0.26	1.01
	$\mathcal{R}_S^-$	0.26	0.07	0.21	0.26	0.30	1.02
	K_1	1.69	0.47	1.34	1.70	2.08	1.01
	q_1	0.95	0.04	0.93	0.96	0.98	1.02
	p_2	0.87	0.09	0.81	0.89	0.94	1.02
	T_{C2}	16.47	0.92	15.91	16.52	17.11	1.01
	α_2	0.02	0.01	0.01	0.02	0.03	1.03
	q_M	0.16	0.07	0.11	0.16	0.21	1.01
	K_M	1.48	0.64	1.00	1.53	2.03	1.01
	σ_1	0.04	0.01	0.04	0.04	0.04	1.02
	σ_2	0.04	0.01	0.04	0.04	0.04	1.01
$\mathcal{M}_1$	$\mathcal{R}_S^+$	0.51	0.30	0.29	0.45	0.65	1.01
	$\mathcal{R}_S^-$	0.55	0.35	0.29	0.46	0.72	1.02
	p_2	0.74	0.16	0.62	0.74	0.87	1.01
	T_{C2}	15.92	1.17	15.08	15.96	16.74	1.01
	α_2	0.05	0.03	0.02	0.05	0.07	1.04
	q_M	0.37	0.17	0.24	0.36	0.49	1.02
	K_M	1.47	0.59	1.01	1.50	1.94	1.01
	σ_1	0.18	0.03	0.16	0.17	0.20	1.02
	σ_2	0.04	0.01	0.03	0.04	0.04	1.02
$\mathcal{M}_2$	$\mathcal{R}_S^+$	0.15	0.14	0.06	0.12	0.19	1.05
	$\mathcal{R}_S^-$	0.42	0.21	0.28	0.38	0.49	1.04
	K_1	1.61	0.51	1.20	1.63	2.04	1.01
	q_1	0.89	0.06	0.86	0.90	0.94	1.01
	q_M	0.08	0.06	0.04	0.08	0.12	1.02
	K_M	1.00	0.57	0.53	0.98	1.48	1.01
	σ_1	0.04	0.01	0.04	0.04	0.05	1.02
	σ_2	0.12	0.02	0.10	0.12	0.13	1.01
$\mathcal{M}_3$	$\mathcal{R}_S^+$	0.26	0.12	0.17	0.25	0.33	1.00
	$\mathcal{R}_S^-$	0.30	0.13	0.22	0.28	0.58	1.00
	K_1	1.53	0.50	1.13	1.50	1.94	1.00
	q_1	0.90	0.06	0.86	0.91	0.95	1.00
	p_2	0.83	0.11	0.76	0.85	0.91	1.00
	T_{C2}	16.93	1.18	16.21	17.05	17.78	1.00
	σ_1	0.04	0.01	0.03	0.04	0.04	1.00
	σ_2	0.05	0.01	0.04	0.05	0.05	1.00
$\mathcal{M}_4$	$\mathcal{R}_S^+$	0.21	0.19	0.08	0.16	0.27	1.04
	$\mathcal{R}_S^-$	0.53	0.31	0.32	0.45	0.66	1.03
	K_1	1.63	0.53	1.22	1.66	2.08	1.01
	q_1	0.93	0.05	0.89	0.93	0.97	1.01
	σ_1	0.04	0.01	0.04	0.04	0.05	1.01
	σ_2	0.12	0.02	0.11	0.12	0.14	1.01

Appendix B

Analytical and numerical set up to study the asymptotic growth dynamics in cyclic hypoxia

B.1 Proof of Lemma 4.4.1

In this section, we provide details on the proof of Lemma 4.4.1 from Chapter 4 by using the results from Section 1.2.3 on the spectral property of strongly positive operators.

B.1.1 Preliminaries.

Solutions of Eq. (4.18a) are of the form $\boldsymbol{x}_t(\theta) = [G_t^1(\theta), C_t^1(\theta), G_2(t)]$ and belong to the space $\mathbb{X} = \mathcal{C}([-r, 0], \mathbb{R}^2) \times \mathbb{R}$, where $r = \max_t \tau_{S+R}^{\dagger}(t)$ and the functions G_t^1, C_t^1 are respectively defined as $G_t^1(\theta) = G_1(t+\theta)$ and $C_t^1(\theta) = C_1(t+\theta)$ where $\theta \in [-r, 0)$. As mentioned in Section 4.4, Eqs. (4.18a) are well-posed on $\mathbb{X}$ so that we can apply the standard theory of positive dynamical systems introduced in Section 1.2.3. As in Section 1.2.3, we introduce the solution map $(U(t,s))_{t\geq s}$ associated with Eqs. (4.18a): $\boldsymbol{x}(t) = U(t,s)\boldsymbol{x}(s)$.

B.1.2 An intermediate result.

As in Example 1.2.1, we start by proving some properties on the matrix $\mathcal{A}(t)$, associated with formulating Eqs. (4.18a) in the absence of delays.

Lemma B.1.1. *Suppose that the following conditions on the coefficients matrix $\mathcal{A}(t)$ are satisfied:*

$$k_1, k_2 > 0, \quad \mathcal{Q}_1 : \mathbb{R} \to [0, q_1], \quad \mathcal{K}_1 : \mathbb{R} \to [0, \infty), \quad p_{2,Sn}^{\dagger} : \mathbb{R} \to [0, p_{2,Sn}^{\dagger}], \tag{B.1}$$

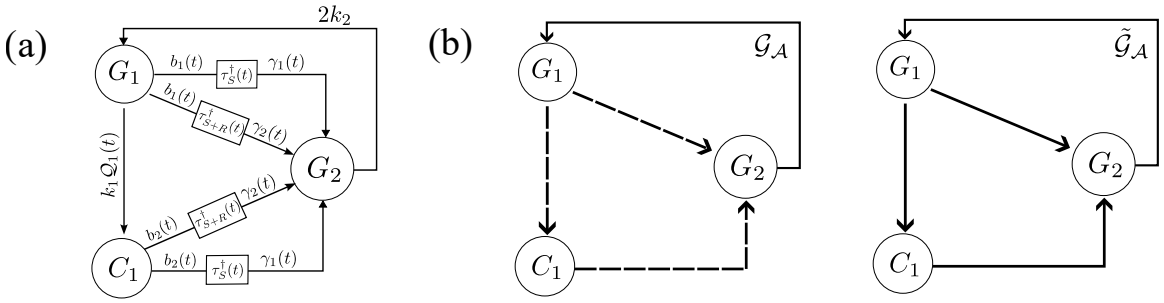


Figure B.1: (a) Model schematic associated with the variables $\boldsymbol{X}$ (taken from Fig. 4.4. (b) Directed time-dependent graphs, $\mathcal{G}_\mathcal{A}$, associated with the matrix $\mathcal{A}(t)$ defined by Eq. (B.3) when conditions (B.1)-(B.2) are satisfied. We indicate with continuous arrows the edges that are present at all times, while the dashed edges are those that are not necessarily present at all times. In (b), $\tilde{\mathcal{G}}_A$ is the static directed graph obtained from $\mathcal{G}_A$ when neglecting time-ordering. The procedures followed to generate $\mathcal{G}_A$ and $\tilde{\mathcal{G}}_A$ are outlined in the main text.

*where $q_1 \in (0,1)$, $p_{2,Sn} \in (0,1)$ are as in Theorem 4.3.1. Then $\mathcal{A}(t)$ is non-negative (*i.e.*, Metzler). Furthermore, if*

$$\max_{t\in[0,\mathcal{T})} \mathcal{Q}_1(t) > 0, \qquad \max_{t\in[0,\mathcal{T})} \mathcal{K}_1(t-\tau_S^\dagger(t)) > 0, \tag{B.2}$$

then $\mathcal{A}(t)$ is also irreducible.

Proof. **1. The matrix $\mathcal{A}$ is Metzler.** We recall that $\mathcal{A}(t)$ is defined as

$$\mathcal{A}(t) = \begin{bmatrix} -k_1 & 0 & 2k_2 \\ k_1\mathcal{Q}_1(t) & -\mathcal{K}_1(t) & 0 \\ \mathcal{A}_{31}(t) & \mathcal{A}_{31}(t) & -k_2 \end{bmatrix}. \tag{B.3}$$

where $\mathcal{A}_{31}(t) = b_1(t-\tau_S^\dagger(t))\gamma_1(t)+b_1(t-\tau_{S+R}^\dagger(t))\gamma_2(t)$ and $\mathcal{A}_{32}(t) = b_2(t-\tau_S^\pm(t))\gamma_1(t)+b_2(t-\tau_{S+R}^\dagger(t))\gamma_2(t)$. The matrix $\mathcal{A}(t)$ is Metzler if all its off-diagonal entries are non-negative. As conditions (B.1) imply that b_1, b_2, γ_1 and γ_2 are all non-negative functions, it is apparent that $\mathcal{A}$ is Metzler. We, therefore, conclude that Eqs. (4.18a) yield a positive dynamical system.

2. The matrix $\mathcal{A}$ is irreducible. To prove that $\mathcal{A}(t)$ is irreducible we use an approach similar to that used in Example 1.2.1 where we leverage the theory of strongly connected graphs. We start by constructing the directed graph, $\mathcal{G}_\mathcal{A}(t)$, associated with the matrix $\mathcal{A}$ following the procedure detailed in Section 1.2.3: each variable corresponds to a node, and an edge from node i to node j ($i \neq j$) exists when the corresponding entry in the matrix $\mathcal{A}$, (*i.e.*, $\mathcal{A}_{ij}$), is positive. The resulting graph $\mathcal{G}_\mathcal{A}$ is represented in Fig. B.1. One key difference from the other examples discussed in this

thesis is that the $\mathcal{G}_{\mathcal{A}}$ is time-dependent [148]. Based on the conditions (B.1)-(B.2), we have that all edges illustrated in Fig. B.1 exist, at least at one point in time, but not necessarily at all times. In Fig. B.1, we indicate with continuous arrows those edges that are permanent, and with dashed arrows those that are time-dependent. The concept of strongly connected nodes can be extended to time-varying graphs upon proper definition of what it means for two nodes to be connected. Without going into the details (which can be found in [148] and citations within), we say that the node i is connected to j if there exists a time-respecting path that allows movement from node i to node j ($i \neq j$) in finite time. Consequently, a strongly connected time-varying graph is one in which every node is temporally connected to every other node. Given that the matrix $\mathcal{A}(t)$ is periodic, if any edge exists in $\mathcal{G}_{\mathcal{A}}$ at a given time t_1 then it will exist (at least) at all times $t = t_1 + n\mathcal{T}$ with $n \in \mathbb{Z}$. This implies that time ordering can be neglected if we are only interested in the connectivity. Consider the static graph, $\tilde{\mathcal{G}}_{\mathcal{A}}$, derived by including all edges that exist in $\mathcal{G}_{\mathcal{A}}$ and neglecting time-ordering. If there is a path of length n_{ij} connecting nodes i and j in $\tilde{\mathcal{G}}_{\mathcal{A}}$, then there is a path in $\mathcal{G}_{\mathcal{A}}$ of duration less or equal to $n_{ij}\mathcal{T}$ connecting nodes i and j; hence i and j are temporally connected in $\mathcal{G}_{\mathcal{A}}(t)$. We conclude that, if the corresponding static graph $\tilde{\mathcal{G}}_{\mathcal{A}}$ is strongly connected, then the time-dependent graph $\mathcal{G}_{\mathcal{A}}$ is also strongly connected.

Looking at Fig. B.1b, it is apparent that $\tilde{\mathcal{G}}_A$ is a strongly connected graph. We therefore conclude that, when conditions (B.1)(B.2) are both satisfied the time-dependent graph $\mathcal{G}_A$ is strongly connected. This implies that, under these conditions, the matrix $\mathcal{A}(t)$ is both positive and irreducible. □

B.1.3 Final result

The results in Lemma B.1.1 are sufficient to show that the dynamical system without delays, $u' = \mathcal{A}(t)u$, is strongly positive. However, this only implies that $(U(t,s))_{t\geq s}$ is positive (not strongly positive) on $\mathbb{X}$ and that $(U(t,s))_{t\geq s}$ is strongly positive on $\mathbb{X}_s = \mathcal{C}([-\tau^{\dagger}_{R+S}(s), 0], \mathbb{R}^2) \times \mathbb{R}$. This however is sufficient to prove Lemma B.1.2.

Lemma B.1.2. *Suppose that coefficients in Eqs. (4.18a)-(4.18d) satisfy the conditions in Lemma B.1.1. Then, for any positive function $\boldsymbol{X}(t)$ satisfying Eq. (4.18a), there exists $c > 0$, such that $\lim_{t\to\infty} \boldsymbol{X}(t) = ce^{\lambda t}\boldsymbol{\xi}_{\boldsymbol{X}}(t)$, where $\lambda \in \mathbb{R}$ and $\boldsymbol{\xi}_{\boldsymbol{X}}(t) \gg 0$ is a $\mathcal{T}$-periodic function* (i.e., $\boldsymbol{\xi}_{\boldsymbol{X}}(t) = \boldsymbol{\xi}_{\boldsymbol{X}}(t+\mathcal{T})$).

Proof. From Lemma B.1.1, we have that $(U(t,s))_{t\geq s}$ is strongly positive on $\mathbb{X}_s$. As standard in the study of periodic systems, we introduce the period map $\Psi_s = U(t+$

$s, s)$. Then, since $(U(t,s))_{t\geq s}$ is strongly positive on $\mathbb{X}_s$ and Eqs. (4.18a) have finite delays, we have that Ψ_s is a strongly positive and compact operator on $\mathbb{X}_s$. By the Krein–Rutman Theorem [65, Theorem 7.2], we have that the operator Ψ_s, defined on $\mathbb{X}_s$, has a dominant real and simple eigenvalue $\mu = e^{\lambda}$ and its eigenfunction $\boldsymbol{\Xi}_s(\theta) \gg 0$ ($\boldsymbol{\Xi}_s(\theta) \in \mathbb{X}_s$) is uniquely defined up to a constant. Importantly, as shown in [92], the spectrum of Ψ_s is independent of the choice of s and whether Ψ_s is defined on $\mathbb{X}_s$ or $\mathbb{X}$. Then, since our model admits a Floquet representation (see [62, Theorem 1.1, Chapter 8]), we have that for any choice of initial time s and initial solution $\boldsymbol{x}_s(\theta) \in \mathbb{X}_s$, $\lim_{n\to\infty} \Psi_s^n \boldsymbol{x}_s(\theta) = e^{\lambda n \mathcal{T}} c_s(\boldsymbol{x}_s)\boldsymbol{\Xi}_s(\theta)$, where $\theta \in [-\tau^{\dagger}_{R+S}(s), 0]$, $c_s(\boldsymbol{x}_s)$ is a constant depending on the initial condition $\boldsymbol{x}_s$ and the initial time s and, the eigenfunctions $\boldsymbol{\Xi}_s(\theta)$ are $\mathcal{T}$-periodic in s, *i.e.*, $\boldsymbol{\Xi}_s(\theta) = \boldsymbol{\Xi}_{s+\mathcal{T}}(\theta)$. Furthermore, consider $\bar{s}$ such that $\tau^{\dagger}_{S+R}(\bar{s}) = r$, then it can be shown that (by uniqueness of the eigenfunctions $\boldsymbol{\Xi}_s$ – up to a normalisation factor) for any t such that $t \in [\bar{s}, \bar{s} + \mathcal{T}]$, $(U(t,\bar{s})\boldsymbol{\Xi}_{\bar{s}})(\theta) = \tilde{\boldsymbol{\Xi}}_t(\theta)$, where $\tilde{\boldsymbol{\Xi}}_t(\theta) = \boldsymbol{\Xi}_t(\theta)$ on $\theta \in [-\tau^{\dagger}_{S+R}(t), 0]$. We therefore have that $\lim_{t\to\infty}(U(t,s)\boldsymbol{x}_s)(\theta) = e^{\lambda t} c_s(\boldsymbol{x}_s) e^{-\delta[t,\bar{s}]\mathcal{T}} \tilde{\boldsymbol{\Xi}}_{\delta[t,s]}(\theta)$, where $\delta[t,s] = (t-s) \mod \mathcal{T}$. Since $\tau^{\dagger}_{S+R} > 0$, this implies that $\lim_{t\to\infty}(U(t,s)\boldsymbol{x}_s)(0) = e^{\lambda t} c_s(\boldsymbol{x}_s) e^{-\delta[t,s]\mathcal{T}} \boldsymbol{\Xi}_{\delta[t,s]}(0)$. We define now the function $\boldsymbol{\xi}(t) = e^{-\delta[t,\bar{s}]\mathcal{T}} \tilde{\boldsymbol{\Xi}}_{\delta[t,\bar{s}]}(0)$, then this is a strictly positive (in the sense specified in Section 1.2.3 – since $\boldsymbol{\Xi}_s \gg 0$) and $\mathcal{T}$-periodic function in $\mathbb{R}^3$. Then for any function $\boldsymbol{X}(t)$ satisfying Eq. (4.18a), this can be written as $\boldsymbol{X}(t) = (U(t,\bar{s})\boldsymbol{x}_{\bar{s}})(0)$, where $\boldsymbol{x}_{\bar{s}}(\theta) = [G^1_{\bar{s}}(\theta), C^1_{\bar{s}}(\theta), G_2(\bar{s})]$, so that $\lim_{t\to\infty} \boldsymbol{X}(t) = ce^{\lambda t}\boldsymbol{\xi}(t)$, where the constant $c = c_{\bar{s}}(\boldsymbol{x}_{\bar{s}}) > 0$. As $\boldsymbol{\xi}$ is a positive (uniquely defined) function and $\lambda \in \mathbb{R}$ is also uniquely defined. This completes our proof. □

B.2 Galerkin approximation for time-dependent delay differential equations

In this section we detail the analytical steps that, starting from the time-dependent PDE, Eqs. (4.23), lead to the discrete time-periodic dynamical system, Eqs.(4.25) based on the Galerkin approximation proposed by Vyasarayani et al. in a series of papers [77, 120, 121]. Recal from Section 4.4.1 that we can re-write our system of $\mathcal{T}$-periodic equations with time-delays in the form:

$$\frac{\partial \boldsymbol{x}}{\partial t} = \tau^{\dagger}_{S+R}(t)^{-1}\frac{\partial \boldsymbol{x}}{\partial s} + \frac{\tau^{\dagger}_{S+R}(t)'s}{\tau^{\dagger}_{S+R}(t)}\frac{\partial \boldsymbol{x}}{\partial s}, \tag{B.4a}$$

with boundary conditions:

$$\left.\frac{\partial \boldsymbol{x}}{\partial t}\right|_{(t,0)} - A(t)\boldsymbol{x} - \gamma_1(t)\boldsymbol{b}(t-\tau_S^\dagger(t))\boldsymbol{P}^T\boldsymbol{x}\left(t, \frac{\tau_S^\dagger(t)}{\tau_{S+R}^\dagger(t)}\right)\boldsymbol{e}_3 \\ - \gamma_2(t)\boldsymbol{b}(t-\tau_{R+S}^\dagger(t))\boldsymbol{P}^T\boldsymbol{x}(t,1)\boldsymbol{e}_3 = 0. \tag{B.4b}$$

The solution of Eqs. (4.23) is then approximated via a Galerkin approximation :

$$\boldsymbol{x}(t,s) \approx \sum_{i=1}^{n_\eta} \boldsymbol{\beta}_i(t)\eta_i(s) = \boldsymbol{\eta}(s)^T B(t), \quad \boldsymbol{\beta}_i \in \mathbb{R}^4, \quad B \in \mathbb{R}^{n_\eta \times 4}, \tag{B.5}$$

where n_η is an arbitrary chosen constant that determines the accuracy of the approximation and $\eta_i{}_{i=1}^{\infty}$ are a given set of orthogonal basis functions:

$$\int_{-1}^{0} \eta_i(s)\eta_j(s) = \begin{cases} 0, & i \neq j, \\ > 0, & i = j. \end{cases} \tag{B.6}$$

By substituting Eq. (B.5) into Eqs. (B.4), we obtain:

$$\boldsymbol{\eta}(s)^T \frac{dB(t)}{dt} = \tau_{S+R}^\dagger(t)^{-1} \frac{\partial \boldsymbol{\eta}(s)^T}{\partial s} B(t) + \frac{\tau_{S+R}^\dagger(t)'s}{\tau_{S+R}^\dagger(t)} \frac{\partial \boldsymbol{\eta}^T(s)}{\partial s} B(t), \tag{B.7a}$$

with boundary conditions:

$$\boldsymbol{\eta}^T(0)\frac{dB}{dt} - A(t)\boldsymbol{\eta}^T(0)B(t) - \gamma_1(t)\boldsymbol{b}(t-\tau_S^\dagger(t))\boldsymbol{P}^T\boldsymbol{\eta}\left(-\frac{\tau_S^\dagger(t)}{\tau_{S+R}^\dagger(t)}\right)^T B\,(t)\,\boldsymbol{e}_3 \\ - \gamma_2(t)\boldsymbol{b}(t-\tau_{R+S}^\dagger(t))\boldsymbol{P}^T\boldsymbol{\eta}^T(-1)B(t)\boldsymbol{e}_3 = 0. \tag{B.7b}$$

We now focus on Eq. (B.7a) and introduce the notation $B_i(t)$ $(i = 1, 2, 3)$ to indicate the i-th column of the matrix B so that $\boldsymbol{\eta}(s)B_i(t) \approx x_i(t,s)$ where x_i is the i-th component of the vector function $\boldsymbol{x}$. We multiply Eqs. (B.7a) for each basis function $\boldsymbol{\eta}_j$, $j = 1, \ldots, n_\eta$ and integrate in the domain $[-1, 0]$ to obtain:

$$\mathtt{W}^{(0)}\frac{dB_i(t)}{dt} = \frac{1}{\tau_{S+R}^\dagger(t)}\mathtt{W}^{(1)}B_i(t) + \frac{\tau_{S+R}^\dagger(t)'}{\tau_{S+R}^\dagger(t)}\mathtt{W}^{(2)}B_i(t), \quad i = 1, \ldots, 4, \tag{B.8a}$$

where the matrices $\mathcal{W} \in \mathbb{R}^{n_\eta \times n_\eta}$ are defined as

$$\mathtt{W}_{ij}^{(0)} = \int_{-1}^{0} \eta_i(s)\eta_j(s)ds, \quad \mathtt{W}_{ij}^{(1)} = \int_{-1}^{0} \eta_i(s)\eta_j'(s)ds, \quad \mathtt{W}_{ij}^{(2)} = \int_{-1}^{0} s\eta_i(s)\eta_j'(s)ds. \tag{B.8b}$$

We can now combine Eqs. (B.8) and Eq. (B.7b) to obtain an overdetermined ODE system:

$$\mathcal{W}^{(0)}\frac{\partial \boldsymbol{B}(t)}{\partial t} = \mathcal{W}^{(1)}(t)\boldsymbol{B}(t), \tag{B.9a}$$

where $\boldsymbol{B} = [B_1, B_2, B_3] \in \mathbb{R}^{3n_\eta}$ and the matrices $\mathcal{W} \in \mathbb{R}^{3(n_\eta+1)\times 3n_\eta}$ are given by:

$$\mathcal{W}^{(0)} = \begin{bmatrix} \mathtt{W}^{(0)} & 0 & 0 \\ 0 & \mathtt{W}^{(0)} & 0 \\ 0 & 0 & \mathtt{W}^{(0)} \\ \mathbb{I} \otimes \boldsymbol{\eta}^T(0) & & \end{bmatrix}, \quad \mathcal{W}^{(1)}(t) = \begin{bmatrix} \tilde{\mathcal{W}}^{(1)}(t) & 0 & 0 \\ 0 & \tilde{\mathcal{W}}^{(1)}(t) & 0 \\ 0 & 0 & \tilde{\mathcal{W}}^{(1)}(t) \\ & \omega & \end{bmatrix}, \tag{B.9b}$$

$$\tilde{\mathcal{W}}^{(1)}(t) = \frac{1}{\tau^\dagger_{S+R}(t)}\mathtt{W}^{(1)} + \frac{\tau^\dagger_{S+R}(t)'}{\tau^\dagger_{S+R}(t)}\mathtt{W}^{(2)}, \tag{B.9c}$$

$$\begin{aligned} \omega = A(t) \otimes \boldsymbol{\eta}^T(0) + (\gamma_1(t)\boldsymbol{b}(t - \tau^\dagger_S(t))\boldsymbol{P}^T) \otimes \boldsymbol{\eta} \left(-\frac{\tau^\dagger_S(t)}{\tau^\dagger_{S+R}(t)} \right)^T \\ + \gamma_2(t)[\boldsymbol{b}(t - \tau^\dagger_{R+S}(t))\boldsymbol{P}^T] \otimes \boldsymbol{\eta}^T(-1), \end{aligned} \tag{B.9d}$$

where $\otimes$ indicates the tensor product. Since Eq. (B.9d) is overdetermined, we can not expect a solution to exist. Following [77], we instead find the solution to the problem:

$$\frac{d\boldsymbol{B}(t)}{dt} = \underbrace{(\mathcal{W}^{(0)})^\dagger \mathcal{W}^{(1)}(t)}_{\mathcal{A}_{n_\eta}} \boldsymbol{B}(t), \tag{B.10}$$

where $(\mathcal{W}^{(0)})^\dagger$ indicates the Moore-Penrose inverse matrix of $\mathcal{W}^{(0)}$, which is commonly adopted to find least-squares solution to linear, overdetermined regression models. We note that other approaches, such as the spectral tau method or Lagrange multipliers [120, 121], can be used to solve the overdetermined system. Having derived the final form of the operator $\mathcal{A}_{n_\eta}$, we now provide more details on how we compute the matrices $\mathcal{W}^{(0)}$ and $\mathcal{W}^{(1)}$.

B.3 Numerical implementation and convergence

Eq. (B.10) holds for a general choice of Galerkin approximation. Following [77], we consider as basis functions $\boldsymbol{\eta}$, the shifted (on the domain $[-1, 0]$) Legendre polynomials. Under this assumption the matrices $\mathtt{W}^{(0)}$, $\mathtt{W}^{(1)}$ and $\mathtt{W}^{(2)}$ take the simple form [121]:

$$\mathtt{W}^{(0)}_{ij} = \begin{cases} 0, & i \neq j, \\ \frac{1}{2i-1}, & i = j \end{cases}, \quad \mathtt{W}^{(1)}_{ij} = \begin{cases} 2, & i < j, \ (i+j) \bmod 2 = 1, \\ 0, & \text{otherwise}, \end{cases}, \tag{B.11a}$$

$$\mathtt{W}^{(2)}_{ij} = \begin{cases} 0, & j > i, \\ (-1)^{i+j}, & i \leq j, \\ \frac{i-1}{2i-1}, & i = j, \end{cases}, \tag{B.11b}$$

where $i = 1 \ldots, n_\eta$. In order to construct the matrix $\mathcal{A}_{n_\eta}$ we are left with specifying the form of the delays $\tau^\dagger_S$ and $\tau^\dagger_R$ and the time-dependent parameters in the model,

specifically $D_S(t) = \int_{-\tau_S^\dagger(t_H)}^{t} u_S(s)ds$, $\mathcal{Q}_1(t)$ and $\mathcal{K}_1(t)$. For the delays, we compute $\tau_{S,R}^\dagger(0)$ by solving Eqs.(4.30) and (4.31a) using the `least square` function part of the *scipy* library in Python. We then compute $\tau_{S,R}^\dagger(t)$ with $(0, \mathcal{T}]$, solving the non-linear system of ODEs:

$$\frac{d\tau_{S,R}^\dagger}{dt} = 1 - \frac{v_{S,R}^\dagger(t)}{v_{S,R}^\dagger(t - \tau_{S,R}^\dagger)}, \quad 0 < t \leq \mathcal{T} \tag{B.12a}$$

$$\tau_{S,R}^\dagger(0) = \text{known value}, \tag{B.12b}$$

where $v_{S,R}^\dagger$ are known functions ($v_S^\dagger$ is defined by Eq. (4.28) and $v_R^\dagger$ is computed as described in Section 4.4.3.1 – page 126). We then solve for the death term $D_S(t)$ by solving the system:

$$\frac{dD_S}{dt} = u_S(m^\dagger(t)), \quad -\tau_{S+R}(0) < t \leq \mathcal{T} \tag{B.12c}$$

$$\tau_{S,R}^\dagger(-\tau_{S+R}(0)) = 0, \tag{B.12d}$$

where $m^\dagger(t)$ is known and computed as described in Section 4.4.3.1 (page 126). Parameters $\mathcal{Q}_1$, $\mathcal{K}_1$ are defined as in Eqs. (2.16a), where the oxygen levels are prescribed by Eqs. (3.2).

In Fig. B.2a we plot the estimates of $\lambda_\dagger$ obtained with the Galerkin approximation as a function of n_η. We find that the method quickly asymptotes to an approximately constant value by $n_\eta = 5$. When comparing with the estimates of the growth rate obtained via dynamical simulations (see Fig. B.2b), we find that the two agree to within 3 significant figures. Given the fast convergence of the Galerkin approximation method, we set $n_\eta = 15$ when computing the results presented in Section 4.4.2.

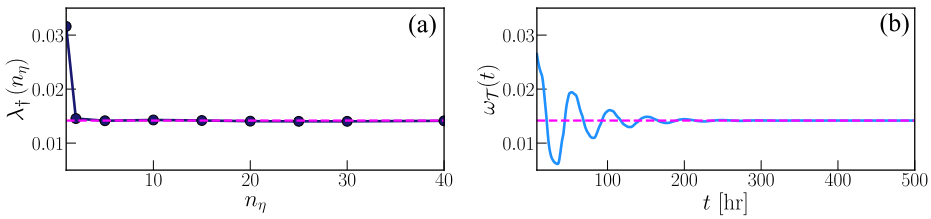


Figure B.2: Comparison of growth rate estimates. (a) Plot of the Galerkin approximation for the population growth rate $\lambda_\dagger$, in a (2,6)-cyclic hypoxia environment as a function of the dimension n_η of the approximated monodromy operator; (b) Evolution of the growth rate $\omega_\mathcal{T}(t)$ (see Definition 4.2.1) for a full dynamical simulation of the model (Eqs. (2.18)) under (2,6)-cyclic hypoxia. In both panels the magenta dotted line indicates the final value of $\omega_\mathcal{T}$ recorded in the simulations.

Bibliography

[1] T Alarcón, H M Byrne, and P K Maini. A mathematical model of the effects of hypoxia on the cell-cycle of normal and cancer cells. *J. Theor. Biol.*, 229:395–411, 2004.

[2] R. Alfieri, E. Bartocci, E. Merelli, and L. Milanesi. Modeling the cell cycle: From deterministic models to hybrid systems. *BioSystems*, 105(1):34–40, 2011.

[3] A. Altinok, D. Gonze, F. Lévi, and A. Goldbeter. An automaton model for the cell cycle. *Interface Focus*, 1(1):36–47, 2011.

[4] S. B. Bader et al. Review cyclic hypoxia: An update on its characteristics, methods to measure it and biological implications in cancer. *Cancers*, 13(1):1–20, 2021.

[5] S. B. Bader, T. S. Ma, et al. Replication catastrophe induced by cyclic hypoxia leads to increased APOBEC3B activity. *Nucleic Acids Research*, 49(13):7492–7506, 2021.

[6] D. J. Balding and R. A. Nichols. A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity. *Genetica*, 96(1-2):3–12, 1995.

[7] M. Baliu-Piqué et al. Breast cancer heterogeneity and response to novel therapeutics. *Cancers*, 12(11):1–16, 2020.

[8] H. Barker et al. The tumour microenvironment after radiotherapy: Mechanisms of resistance and recurrence. *Nat. Rev. Cancer*, 15:409–25, 2015.

[9] J. Bartkova et al. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature*, 434(7035):864–870, 2005.

[10] B. Basse et al. A mathematical model for analysis of the cell cycle in cell lines derived from human tumors. *J. Math. Biol.*, 47(4):295–312, 2003.

[11] B. Basse et al. Modelling cell death in human tumour cell lines exposed to the anticancer drug paclitaxel. *J. Math. Biol.*, 49(4):329–357, 2004.

[12] B. Basse et al. Modelling the flow of cytometric data obtained from unperturbed human tumour cell lines: Parameter fitting and comparison. *Bull Math Biol*, 67(4):815–830, 2005.

[13] B. Basse and P. Ubezio. A generalised age- and phase-structured model of human tumour cell populations both unperturbed and exposed to a range of cancer therapies. *Bull Math Biol*, 69(5):1673–1690, 2007.

[14] A. Bátkai et al. *Positive operator semigroups: from finite to infinite dimensions*, volume 257 of *Operator theory, advances and applications*. Birkhauser Cham, 2017.

[15] R. E. Begg et al. The steady-states of a multi-compartment, age-size distribution model of cell-growth. *Eur. J. Appl. Math.*, 19(4):435–458, 2008.

[16] G. I. Bell. Cell growth and division: III. conditions for balanced exponential growth in a mathematical model. *Biophys. J.*, 8(4):431–444, 1968.

[17] Y. Ben-Ami et al. Structural features of microvascular networks trigger blood-flow oscillations. *Bull Math Biol*, 84:1–36, 2022.

[18] I. Berger Fridman et al. High throughput microfluidic system with multiple oxygen levels for the study of hypoxia in tumor spheroids. *Biofabrication*, 13(3):035037, 2021.

[19] J. R. M. Black and N. McGranahan. Genetic and non-genetic clonal diversity in cancer evolution. *Nat. Rev. Cancer*, 21:1–14, 2021.

[20] A. M. Boddy et al. Life history trade-offs in tumors. *Curr. Pathobiol. Rep.*, 6(4):201, 2018.

[21] R. G. Bristow and R. P. Hill. Hypoxia and metabolism: Hypoxia, DNA repair and genetic instability. *Nat. Rev. Cancer*, 8(3):180–192, 2008.

[22] P. J. Bröckelmann et al. Targeting DNA Repair, Cell Cycle, and Tumor Microenvironment in B Cell Lymphoma. *Cells*, 9(10):2287, 2020.

[23] A. P. Browning et al. Quantitative analysis of tumour spheroid structure. *eLife*, 10:e73020, 2021.

[24] A. P. Browning, P. Haridas, and M. J. Simpson. A bayesian sequential learning framework to parameterise continuum models of melanoma invasion into human skin. *Bull Math Biol*, 81(3):676–698, 2019.

[25] A. M. Buckley et al. Targeting hallmarks of cancer to enhance radiosensitivity in gastrointestinal cancers. *Nat Rev Gastroenterol Hepatol*, 17(5):298–313, 2020.

[26] J. A. Bull et al. Mathematical modelling reveals cellular dynamics within tumour spheroids. *PLoS Comput. Biol.*, 16(8):e1007961, 2020.

[27] H. Byrne and L. Preziosi. Modelling solid tumour growth using the theory of mixtures. *Mathematical Medicine and Biology*, 20(4):341–366, 2003.

[28] S. R. Cajal et al. Clinical implications of intratumor heterogeneity: challenges and opportunities. *J. Mol. Med.*, 98(2):161, 2020.

[29] G. L. Celora et al. A DNA-structured mathematical model of cell-cycle progression in cyclic hypoxia. *J. Theor. Biol.*, 545:111104, 2022.

[30] G. L. Celora et al. Spatio-temporal modelling of phenotypic heterogeneity in tumour tissues and its impact on radiotherapy treatment. *J. Theor. Biol.*, page 111248, 2022.

[31] G.L. Celora et al. Phenotypic variation modulates the growth dynamics and response to radiotherapy of solid tumours under normoxia and hypoxia. *J. Theor. Biol.*, 527, 2021.

[32] H. X. Chao et al. Orchestration of DNA damage checkpoint dynamics across the human cell cycle. *Cell Systems*, 5(5):445–459.e5, 2017.

[33] H. X. Chao et al. Evidence that the human cell cycle is a series of uncoupled, memoryless phases. *Mol. Syst. Biol.*, 15(3):e8604, 2019.

[34] S. J. Chapman et al. A nonlinear model of age and size-structured populations with applications to cell cycles. *ANZIAM Journal*, 49(2):151–169, 2008.

[35] M. T. Chew et al. Potential lethal damage repair in glioblastoma cells irradiated with ion beams of various types and levels of linear energy transfer. *J. Radiat. Res.*, 60(1):59–68, 2019.

[36] G. Chiorino and M. Lupi. Variability in the timing of G1/S transition. *Math. Biosci.*, 177-178:85–101, 2002.

[37] K. P Chwialkowski et al. Fast two-sample testing with analytic representations of probability measures. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015.

[38] M. Clerx et al. Probabilistic inference on noisy time series (PINTS). *J. Open Res. Softw.*, 7(1):23, 2019.

[39] J. Collis et al. Bayesian calibration, validation and uncertainty quantification for predictive modelling of tumour growth: A tutorial. *Bull Math Biol*, 79(4):939–974, 2017.

[40] J. J. Crivelli et al. A mathematical model for cell cycle-specific cancer virotherapy. *J Biol Dyn*, 6(sup1):104–120, 2012.

[41] A. Csikász-Nagy. Computational systems biology of the cell cycle. *Brief. Bioinformatics*, 10(4):424–434, 2009.

[42] I. Dagogo-Jack and A. T. Shaw. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol*, 15(2):81–94, 2018.

[43] A. C. Daly et al. Inference-based assessment of parameter identifiability in nonlinear biological models. *J R Soc Interface*, 15(144):20180318, 2018.

[44] J. de Boeck et al. A modular approach for modeling the cell cycle based on functional response curves. *PLoS Comput. Biol.*, 17(8):e1009008, 2021.

[45] A. Ducrot et al. An in vitro cell population dynamics model incorporating cell size, quiescence, and contact inhibition. *Math Models Methods Appl Sci*, 21(SUPPL. 1):871–892, 2011.

[46] J. Dyer et al. Black-box bayesian inference for economic agent-based models. *arXiv:10.48550*, 2022.

[47] A. K. Erlang. The theory of probabilities and telephone conversations. *Nyt. Tidsskr. Mat. Ser. B*, 20:33–39, 1909.

[48] S. Fadda et al. A novel population balance model to investigate the kinetics of in vitro cell proliferation: Part I. model development. *Biotechnol. Bioeng.*, 109(3):772–781, 2012.

[49] L. Farina and S Rinaldi. *Positive linear systems : theory and applications*. Pure and applied mathematics. Wiley, New York ; Chichester, 2000.

[50] A. Fauré et al. Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics*, 22(14):124–131, 2006.

[51] E. Fokas et al. Targeting ATR in DNA damage response and cancer therapeutics. *Cancer Treatment Reviews*, 40(1):109–117, 2014.

[52] I. P. Foskolou et al. Ribonucleotide Reductase Requires Subunit Switching in Hypoxia to Maintain DNA Replication. *Molecular Cell*, 66(2):206–220, 2017.

[53] M. V.S. Frasson and S. M. Verduyn Lunel. Large Time Behaviour of Linear Functional Differential Equations. *Integral Equations and Operator Theory 2003 47:1*, 47(1):91–121, 2003.

[54] R. A. Freiberg et al. DNA Damage during Reoxygenation Elicits a Chk2-Dependent Checkpoint Response. *Mol. Cell. Biol.*, 26(5):1598, 2006.

[55] H. B. Frieboes et al. Three-dimensional multispecies nonlinear tumor growth-II: Tumor invasion and angiogenesis. *J. Theor. Biol.*, 264(4):1254–1278, 2010.

[56] A. Gelman et al. *Bayesian Data Analysis*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, 2013.

[57] J.S. Good and K.J. Harrington. The hallmarks of cancer and the radiation oncologist: Updating the 5rs of radiobiology. *Clin. Oncol.*, 25(10):569 – 577, 2013.

[58] V. G. Gorgoulis et al. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature*, 434(7035):907–913, 2005.

[59] T. Goto et al. Visualizing cell-cycle kinetics after hypoxia/reoxygenation in HeLa cells expressing fluorescent ubiquitination-based cell cycle indicator (Fucci). *Exp. Cell Res.*, 339(2):389–396, 2015.

[60] H. P. Greenspan. Models for the Growth of a Solid Tumor by Diffusion. *Studies in Applied Mathematics*, 51(4):317–340, 1972.

[61] M. Gyllenberg and G. F. Webb. A nonlinear structured population model of tumor growth with quiescence. *J. Math. Biol.*, 28(6):671–694, 1990.

[62] J. K. Hale. *Theory of functional differential equations*. Applied mathematical sciences (Springer-Verlag New York Inc.) ; 3. Springer-Verlag, New York, 2d ed edition, 1977.

[63] D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 100(1):57 – 70, 2000.

[64] N. Hansen, S. D. Müller, and P. Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evol Comput*, 11(1):1–18, 2003.

[65] P. Hess. *Periodic-parabolic Boundary Value Problems and Positivity*. Pitman research notes in mathematics series. Longman Scientific & Technical, 1991.

[66] E. Hines et al. Determination of parameter identifiability in nonlinear biophysical models: A bayesian approach. *J. Gen. Physiol.*, 143:401–416, 2014.

[67] J. Hlouschek et al. Targeting SLC25A10 alleviates improved antioxidant capacity and associated radioresistance of cancer cells induced by chronic-cycling hypoxia. *Cancer letters*, 439:24–38, 2018.

[68] T. Homma and A. Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliab. Eng. Syst*, 52(1):1–17, 1996.

[69] C. G. Hubert and J. D. Lathia. Seeing the GBM diversity spectrum. *Nature Cancer*, 2(2):135–137, 2021.

[70] H. Inaba. The Malthusian parameter and R_0 for heterogeneous populations in periodic environments. *MBE*, 9(2):313–346, 2012.

[71] T. Insperger and G. Stépán. Semi-discretization for time-delay systems: Stability and engineering applications. *Applied Mathematical Sciences (Switzerland)*, 178:i–174, 2011.

[72] T. Iwanaga et al. Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses. *SESMO*, 4:18155, 2022.

[73] W. Jin et al. Mathematical Model of Tumour Spheroid Experiments with Real-Time Cell Cycle Imaging. *Bull Math Biol*, 83(5):44, 2021.

[74] W. Jitkrittum et al. Interpretable distribution features with maximum testing power, 2016.

[75] R. H. Johnstone et al. Uncertainty and variability in models of the cardiac action potential: Can we build trustworthy models? *J. Mol. Cell. Cardiol.*, 96:49–62, 2016.

[76] Y. Jung et al. Crosstalk between Plk1, p53, cell cycle, and G2/M DNA damage checkpoint regulation in cancer: computational modeling and analysis. *NPJ Syst. Biol. Appl.*, 7(1):1–13, 2021.

[77] S. S. Kandala et al. Pole placement for delay differential equations with time-periodic delays using Galerkin approximations. *J. Comput. Nonlinear Dyn.*, 16(9), 2021.

[78] M. H. Kim et al. Engineering tools for regulating hypoxia in tumour models. *J Cell Mol Med*, 25(16):7581–7592, 2021.

[79] J. J. Klowss et al. A stochastic mathematical model of 4D tumour spheroids with real-time fluorescent cell cycle labelling. *J R Soc Interface*, 19(189), 2022.

[80] R. Koens et al. Microfluidic platform for three-dimensional cell culture under spatiotemporal heterogeneity of oxygen tension. *APL Bioeng.*, 4(1):016106, 2020.

[81] C. Koumenis and B. G. Wouters. "Translatin" tumor hypoxia: Unfolded protein response (UPR)-dependent and UPR-independent pathways. *MCR*, 4(7):423–436, 2006.

[82] S. Kucherenko et al. Monte Carlo evaluation of derivative-based global sensitivity measures. *Reliab. Eng. Syst*, 94(7):1135–1148, 2009.

[83] R. Kumareswaran et al. Cyclic hypoxia does not alter RAD51 expression or PARP inhibitor cell kill in tumor cells. *Radiother Oncol*, 116(3):388–391, 2015.

[84] N. G. Laleh et al. Classical mathematical models for prediction of response to chemotherapy and immunotherapy. *PLoS Comput. Biol.*, 18(2):e1009822, 2022.

[85] B. Lambert. *A Student's Guide to Bayesian Statistics*. SAGE Publications, 2018.

[86] B. Lambert et al. Bayesian inference of agent-based models: a tool for studying kidney branching morphogenesis. *J. Math. Biol.*, 76(7):1673–1697, 2018.

[87] B. Lemmens and A. Lindqvist. DNA replication and mitotic entry: A brake model for cell cycle progression. *J. Cell Biol.*, 218(12):3892–3902, 11 2019.

[88] K. B. Leszczynska et al. Preclinical testing of an Atr inhibitor demonstrates improved response to standard therapies for esophageal cancer. *Radiother Oncol*, 121(2):232–238, 2016.

[89] T. D. Lewin et al. A three phase model to investigate the effects of dead material on the growth of avascular tumours. *MMNP*, 15, 2020.

[90] C. Li et al. Exploring hypoxic biology to improve radiotherapy outcomes. *Expert Rev. Mol. Med.*, 24, 2022.

[91] Lossaint, G. and others. Reciprocal regulation of p21 and Chk1 controls the cyclin D1-RB pathway to mediate senescence onset after G2 arrest. *J. Cell. Sci.*, 135(8), 2022.

[92] Yijun Lou and Xiao Qiang Zhao. A theoretical approach to understanding population dynamics with seasonal developmental durations. *J. Nonlinear Sci.*, 27(2):573–603, 2017.

[93] N. Lüdtke et al. Information-theoretic sensitivity analysis: a general method for credit assignment in complex networks. *J R Soc Interface*, 5(19):223–235, 2008.

[94] S. Maad Sasane. An age structured cell cycle model with crowding. *J. Math. Anal. Appl.*, 444(1):768–803, 2016.

[95] M. J. Marchese et al. Potentially lethal damage repair in human cells. *Radiother Oncol*, 9(1):57–65, 1987.

[96] S. Marino et al. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.*, 254(1):178–196, 2008.

[97] A. Marusyk et al. Intra-tumour heterogeneity: A looking glass for cancer? *Nat. Rev. Cancer*, 12(5):323–334, 2012.

[98] A. Marusyk et al. Intratumor heterogeneity: The rosetta stone of therapy resistance. *Cancer Cell*, 37(4):471–484, 2020.

[99] S. Matsumoto et al. Imaging cycling tumor hypoxia. *AACR*, 70(24):10019–10023, 2010.

[100] H. K. Matthews et al. Cell cycle control in cancer. *Nat. Rev. Mol. Cell Biol.*, 23(1):74–88, 2021.

[101] M. M. Mc Gee. Targeting the mitotic catastrophe signaling pathway in cancer. *Mediators Inflamm.*, 2015, 2015.

[102] Stephen Joseph McMahon. The linear quadratic model: Usage, interpretation and challenges. *Phys. Med. Biol.*, 64(1):01TR01, 2019.

[103] C. Michiels, C. Tellier, and O. Feron. Cycling hypoxia: A key feature of the tumor microenvironment. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1866(1):76 – 86, 2016.

[104] R. J. Murphy et al. Designing and interpreting 4D tumour spheroid experiments. *Commun. Biol.*, 5(1):1–11, 2022.

[105] R. J. Murphy et al. Growth and adaptation mechanisms of tumour spheroids with time-dependent oxygen availability. *bioRxiv*, 2022.

[106] S. Negrini et al. Genomic instability an evolving hallmark of cancer. *Nat. Rev. Mol. Cell Biol.*, 11(3):220–228, 2010.

[107] N. Ng et al. Challenges to DNA replication in hypoxic conditions. *The FEBS Journal*, 285(9):1563–1571, 2018.

[108] M. J. O'Connor. Targeting the DNA Damage Response in Cancer. *Molecular Cell*, 60(4):547–560, 2015.

[109] M. Olcina, P. S. Lecane, and E. M. Hammond. Targeting hypoxic cells through the DNA damage response. *Clin. Cancer Res.*, 16(23):5624–5629, 2010.

[110] M. M. Olcina et al. Replication stress and chromatin context link ATM activation to a role in DNA replication. *Molecular Cell*, 52(5):758–766, 2013.

[111] N. Pandey and P. K. Vinod. Mathematical modelling of reversible transition between quiescence and proliferation. *PLOS ONE*, 13(6):e0198420, 2018.

[112] T. M. Pawlik and K. Keyomarsi. Role of cell cycle in mediating sensitivity to radiotherapy. *IJROBP*, 59(4):928–942, 2004.

[113] F. Pianosi and T. Wagener. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environ Model Softw*, 67:1–11, 2015.

[114] I. M. Pires et al. Effects of acute versus chronic hypoxia on DNA damage responses and genomic instability. *Cancer Research*, 70(3):925–935, 2010.

[115] I. M. Pires et al. Exposure to acute hypoxia induces a transient DNA damage response which includes Chk1 and TLK1. *Cell Cycle*, 9(13):2502, 2010.

[116] T. L. Place et al. Limitations of oxygen delivery to cells in culture: An underappreciated problem in basic and translational research. *Free Radical Biology and Medicine*, 113:311–322, 2017.

[117] G. Qian and A. Mahdi. Sensitivity analysis methods in the biomedical sciences. *Math. Biosci.*, 323:108306, 2020.

[118] A. Raue et al. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, 2009.

[119] A. Ron et al. Volumetric optoacoustic imaging unveils high-resolution patterns of acute and cyclic hypoxia in a murine model of breast cancer. *Cancer Research*, 79(18):4767–4775, 2019.

[120] A. Sadath and C. P. Vyasarayani. Galerkin approximations for stability of delay differential Equations with time periodic coefficients. *J. Comput. Nonlinear Dyn.*, 10(2), 2015.

[121] A. Sadath and C. P. Vyasarayani. Galerkin approximations for stability of delay differential equations with time periodic delays. *J. Comput. Nonlinear Dyn.*, 10(6), 2015.

[122] A. Saltelli et al. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics*, 41(1):39–56, 1999.

[123] A. Saltelli et al. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Comput. Phys. Commun.*, 181(2):259–270, 2010.

[124] F. Sarrazin et al. Global sensitivity analysis of environmental models: Convergence and validation. *Environ Model Softw*, 79:135–152, 2016.

[125] K. Saxena and M. K. Jolly. Acute vs. Chronic vs. cyclic hypoxia: Their differential dynamics, molecular mechanisms, and effects on tumor progression. *Biomolecules*, 9(8), 2019.

[126] E. Sherer et al. Identification of age-structured models: Cell cycle phase transitions. *Biotechnol. Bioeng.*, 99(4):960–974, 2008.

[127] K. Simms, N. Bean, and A. Koerber. A Mathematical Model of Cell Cycle Progression Applied to the MCF-7 Breast Cancer Cell Line. *Bull Math Biol*, 74(3):736–767, 2012.

[128] R. Singhania et al. A Hybrid Model of Mammalian Cell Cycle Regulation. *PLoS Comput. Biol.*, 7(2), 2011.

[129] H. L. Smith. *Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems*. Mathematical surveys and monographs. American Mathematical Society, Providence, R.I., 1995.

[130] I. M. Sobol. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Comput. Math. & Math. Phys.* , 7(4):86–112, 1967.

[131] I. M. Sobol. On quasi-Monte Carlo integrations. *Math Comput Simul*, 47:103–112, 1998.

[132] I. M. Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math Comput Simul*, 55(1-3):271–280, 2001.

[133] L. Spinelli et al. Modelling the balance between quiescence and cell death in normal and tumour cell populations. *Math. Biosci.*, 202(2):349–370, 2006.

[134] M. A.R. Strobl et al. Turnover Modulates the Need for a Cost of Resistance in Adaptive Therapy. *Cancer Research*, 81(4):1135–1147, 2021.

[135] H. Sung et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J. Clin.*, 71(3):209–249, 2021.

[136] T. Suwa, M. Kobayashi, J. Nam, and H. Harada. Tumor microenvironment and radioresistance. *Exp. Mol. Med.*, 53(6):1029–1035, 2021.

[137] M. Thattai and A. Van Oudenaarden. Intrinsic noise in gene regulatory networks. *PNAS*, 98(15):8614–8619, 2001.

[138] L. Toledo, K. J. Neelsen, and J. Lukas. Molecular Cell Perspective Replication Catastrophe: When a Checkpoint Fails because of Exhaustion. *Molecular Cell*, 66:735–749, 2017.

[139] P. Ubezio. Unraveling the complexity of cell cycle effects of anticancer drugs in cell populations. *Discrete Contin. Dyn. Syst. - B*, 4(1):323–335, 2004.

[140] P. Ubezio et al. Quantitative assessment of the complex dynamics of G1, S, and G2-M checkpoint activities. *Cancer Research*, 69(12):5234–5240, 2009.

[141] A. Verdugo et al. Molecular mechanisms creating bistable switches at cell cycle transitions. *Open Biology*, 3(MAR), 2012.

[142] E. I. Verriest. Robust stability of time varying systems with unknown bounded delays. *Proc. IEEE Conf. Decis. Control*, 1:417–422, 1994.

[143] E. I. Verriest. Inconsistencies in systems with time-varying delays and their resolution. *IMA J. Math. Control Inform.*, 28(2):147–162, 2011.

[144] S. T. Vittadello et al. Mathematical models for cell migration with real-time cell cycle dynamics. *Biophys. J.*, 114(5):1241–1253, 2018.

[145] S. T. Vittadello et al. Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchronization in cell proliferation. *J R Soc Interface*, 16(157):20190382, 2019.

[146] B. J. Wadsworth et al. Transiently hypoxic tumour cell turnover and radiation sensitivity in human tumour xenografts. *Br. J. Cancer*, 126(11):1616–1626, 2022.

[147] H. Wang et al. Cancer radiosensitizers. *TIPS*, 39(1):24–48, 2018.

[148] Y. Wang et al. Time-dependent graphs: Definitions, applications, and algorithms. *Data Sci. Eng.*, 4(4):352–366, 2019.

[149] G. F. Webb. An operator-theoretic formulation of asynchronous exponential growth. *Trans Am Math Soc*, 303(2):751, 1987.

[150] T. S. Weber et al. Quantifying the length and variance of the eukaryotic cell cycle phases by a stochastic model and dual nucleoside pulse labelling. *PLoS Comput. Biol.*, 10(7), 2014.

[151] C. M. West and F. Slevin. Tumour Hypoxia. *Clin. Oncol.*, 31(9):595–599, 2019.

[152] F. Witzel et al. Analysis of impedance-based cellular growth assays. *Bioinformatics*, 31(16):2705–2712, 2015.

[153] C. A. Yates et al. A multi-stage representation of cell proliferation as a markov process. *Bull Math Biol*, 79(12):2905–2928, 2017.

[154] A. Zinovyev et al. Modeling progression of single cell populations through the cell cycle as a sequence of switches. *Front. Mol. Biosci.*, 8:1340, 2022.