Cellular decision-making models in yeast

Aldo Estefano Encarnación Segura

A thesis presented for the degree of

Doctor of Philosophy

Supervised by Professor James A R Marshall, Dr Thomas Bose and

Dr Andreagiovanni Reina

Department of Computer Science

University of Sheffield

United Kingdom

September 2020

Abstract

Decision-making is ubiquitous throughout all levels of biological complexity, from social insect colonies to individual cells and multi-cellular organisms. The study of decision-making by different fields has suggested that there are shared underlying principles and decision-making mechanisms that can be used to describe the behaviour of any given biological system, regardless of its specific nature. The relatively recent application of decision theory to the study of cellular systems has provided great insights into the nature of different cellular processes.

In this thesis, I aim to explore and describe the sugar consumption dynamics observed in a yeast culture growing in a binary-sugar mixture. In order to study this cellular process, I develop several mathematical models to describe the sugar consumption behaviour of yeast. The models are influenced by the work of Pais et al. (2013) on house-hunting honeybee swarms. I use experimental data gathered from yeast cultures grown in maltose, galactose, and a mixture of both sugars to validate and parameterise the models. I show that with a single parameterisation, the models are capable of replicating the metabolic and biomass experimental data of yeast growing in single sugar, as well as binary-sugar mixtures.

Additionally, the models are studied by means of bifurcation and dynamical systems analysis. As pointed out by Aidelberg et al. (2014), microorganisms growing in a media with two different sugars present one of three different consumption strategies: 1) simultaneous consumption of both sugars present in the media, 2) exclusive consumption of one of the two sugars available, and 3) no consumption. I show that

the models developed present features such as decision-deadlock, deadlock-breaking bifurcations, and deadlock-restoring bifurcations, which give rise to these consumption strategies. We also show that the transition between these regimes depends on the value of key parameters of our models.

Declaration

I declare that this thesis is original work, apart from the exceptions indicated by a reference in the text, and that it has not been submitted to any other university for examination.

Acknowledgements

I want to express my enormous gratitude to my supervisory team. I owe an outstanding debt to Dr Thomas Bose, who during the first three years of my PhD helped me tremendously by displaying immense patience and support and sharing his expertise. Likewise, Dr Andreagiovanni Reina gave me great support and provided me with expert advice, specifically during my last year. Finally, I want to thank Professor James Marshall, who, as my primary supervisor, was crucial to the completion of this thesis and the work described in it. His generosity and care are only matched by his expertise and judicious advice. I will forever be thankful for the responsibility they decided to share and commitment they partook in by taking me as a student. It has been a very enriching experience.

Additionally, I am grateful to all members of the BET Lab group. Our meetings, conversations and advice were always enriching and valuable.

I want to express my gratitude to the National Council for Science and Technology (CONACYT) for the financial support they provided me with during my first three years. Likewise, I thank the European Research Council (ERC) for the funding on my last year.

Finally, I want to thank my parents, Bernardo and Xochitl. None of this could have happened without you and your love and support. I am eternally in grateful. To my sister Miranda, my extended family and my friends, both the old and the new, I can only say that these four years were that much more enjoyable because of you.

Statement of contribution

All of the experimental data, as well as the development of the *platereader* Python library were developed by the Swain Lab independently of this project. I did not contribute to either one of these tasks. The extent of our collaboration included two short visits to the Swain Lab in the University of Edinburgh (the first visit had a duration of one day whilst the second visit lasted for ten days). During the first one day long visit, we discussed this project in its early stages and received feedback which helped calibrate the direction in which we would continue to design our models. During the second visit I was introduced to their experimental data and the plate data analysing software *platereader*, which as I have previously mentioned, was developed by them. Subsequently, we continued to communicate with them through online meetings in which they helped by providing insightful feedback. Their data are used in this project with their full consent with the condition that the credit be properly stated. The data sets used in this project and specified in table 4.1 are publicly available at Bandeira and Swain (2020).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this introductory chapter I discuss the study of decision-making within the context of cellular behaviour, I focus primarily on the behaviour regarding carbon catabolism and selection of carbon sources, which is the point of interest of this thesis. I describe the advance that have being made in regards of understanding the sugar preferences observed in microorganisms. Additionally, I discuss how the scientific community has tackled the challenge of studying biochemical networks, as well as the difference between the study of the behaviour of a microbial population, contrasted with the behaviour displayed by the individuals within that population. Additionally, I describe how stochasticity has being studied in microbial and cellular systems. Finally, I state the objectives of this thesis as well as describing the specific system we are interested in studying, and the analyses we performed on the mathematical models we developed.

1.1 Cellular decision-making

Cellular biology and microbiology are fields in which decision-making has been thoroughly studied. Cellular decision-making has always been present in one way or another since the first studies about carbon consumption in microbes emerged. As

we have already mentioned, in recent years, research efforts made by dissimilar fields such as neuroscience, psychology, behavioural biology and microbiology have shown remarkable parallels in the mechanisms and interaction patterns employed by different biological systems in order to reach a decision. The application of concepts from other fields within a cellular or microbiology context has provided great insights into cellular decision-making (Perkins and Swain, 2009; Tyson et al., 2003; Rhee et al., 2012). Decision-making in cellular behaviour has been studied primarily in cellular differentiation, bacterial motility, and nutrient selection. In what follows we will talk mostly about decision-making in nutrient selection.

One of the most important decisions a microbe must make is what carbon source consumption strategy to adopt in response to the specific nutrients available to it. In the specific case of a binary-sugar mixture, the microbe has three potential metabolic profiles it could display: 1) simultaneous consumption of both sugars, 2) preferential consumption of either one of the sugars, or 3) no consumption of either sugar (Aidelberg et al., 2014). The question of which nutrients are preferentially consumed by a given organism has been a main subject of study and observation for microbiologists for quite a long time (Monod, 1949; D'amore et al., 1989). Although the quest to explain the underlying mechanisms through which those decisions are implemented by the organism is addressed through the study of metabolic pathways and biochemical interactions, the tools from decision-making theory can also be applied to explore the subject.

Concepts such as value-sensitivity and inhibition are particularly useful in the context of cellular behaviour since they are inherent properties of their regulatory systems (Horák, 2013). Aidelberg et al. (2014) have recently established a preference hierarchy of carbon sources in *E. coli* based on binary-sugar mixtures experiments. Preferential consumption of nutrients has been observed in microorganisms since the beginnings of the field (Monod, 1949). It is usually accepted that cells prioritise the consumption of whichever of the available sugars provides the greatest growth advantage. Nutrients that provide a greater growth advantage have a higher value

to the cell. Hence they will be preferred than those that give a lower growth advantage. For most microorganisms, it is a well established fact that their preferred carbon source is glucose, which they will consume exclusively regardless of which other sugars are also available (Horák, 2013). This preference, and the overarching existence of a consumption hierarchy, not only underscores the notion of value-based decisions being implemented at the cellular level, but also highlights how inhibition plays a role within the cellular decision-making dynamics, just as it does within social insect decision-making (Pais et al., 2013). In addition, if the value of sugar i is associated with its extracellular concentration and sugar i was the alternative being exploited by the microorganism, then this value would smoothly decrease over time as a result of the cellular system consuming sugar i, and thereby decreasing its extracellular concentration. Depending on the implementation, this feature could entail dynamical phenomena such as hysteresis loops to emerge in the system as the value of the alternatives is diminishes, as shown by Pais et al. (2013); Pomerening et al. (2003) and Laurent and Kellershohn (1999). The carbon catabolite repression, also known by its more casual name "glucose effect", refers to a global inhibitory effect that glucose metabolism exerts upon the consumption of other sugars (Horák, 2013). This means that, whenever glucose is present and is being consumed by the cell, an entire network of regulatory pathways will trigger into action to stop the uptake and consumption of any other sugars. As glucose is depleted, the tight control will also diminish, allowing the cell to consume other sugars (Horák, 2013). Although the name might imply otherwise, there is evidence that non-glucose sugars can trigger the global repressive effect, although it is not yet entirely clear which sugars can be considered to trigger this inhibitory effect, nor to what extent or under which conditions (Lodi et al., 1991; Herrero et al., 1985; Chambers et al., 2004; Horak et al., 2002).

The study of complex biochemical networks is a great challenge in biology. However, as Alon (2007c) points out, the biological complexity observed in metabolic systems can be simplified to be studied and analysed in a way that allows researchers to draw

insights from the biological machinery. This perspective bears a close resemblance to the concept of heuristics and how they are built. Hutchinson and Gigerenzer (2005) describe heuristics as robust strategies, which consist in building blocks that make use of the different capacities inherent to the organism. This is reflected by the observation made by Tyson et al. (2003); diagrams of the biochemical interactions in molecular regulatory mechanisms bear close resemblance to electrical circuitry. This similarity is found also in the way the regulatory systems are built. The complexity of the cellular regulatory network emerges from simpler modules connected to one another. Work done by Alon (2007a) highlights the circuit-like nature of molecular biology. The simple recurring patterns of interaction or *Network motifs* are found in a plethora of different organisms, and in all them they perform the same information processing operation (Alon, 2007b). Alon (2007a) discusses how network motifs are likely the result of convergent evolution, being independently discovered due to the important functions they perform. Additionally, they are found in a number of other biological networks.

Over time, the regulatory networks that control the metabolic cellular response to external stimuli optimise their responses to the expected environmental conditions, even when this forces them to follow sub-optimal strategies when conditions change. A way in which cellular populations tackle environmental variability is through phenotypic heterogeneity. Phenotypic heterogeneity is a form of diversity that presents itself in genetically identical cellular populations. Individuals within these populations display different phenotypes independently of environmental changes. As Ackermann (2015) describes in his review, phenotypic heterogeneity is common, particularly when implementing a sensing strategy is sub-optimal due to the environmental fluctuations. This strategy is a bet-hedging strategy. More specifically, it is a stochastic switching strategy, which allows the population to tackle the problem of uncertainty about their environment by adopting diverse survival or feeding strategies. Diversifying the phenotype within a population is beneficial because it allows some individuals to survive through dramatic environmental fluctuations.

Therefore, phenotype heterogeneity ensures that the genotype survives in a variable environment. Investing all resources into a single good strategy under a specific set of current environmental conditions is the optimal move in a stable state. However, if the environment changes, the group will go through a period of adaptability whilst it changes and implements a new strategy befitting the new environment.

As stated by Ackermann (2015), populations will not invest all their resources into a single optimal strategy, they will rather diversify their strategies. That way, if the environment suddenly changes, there is already a strategy put in place that ensures that some individuals of the population will survive the new environmental conditions. This diversification occurs at the individual level within these populations. Additionally, and broadly speaking, one could categorise the different strategies exhibited by the population as either optimistic or pessimistic. McNamara et al. (2011) show that depending on the nature of the environmental fluctuations, spatial or temporal, natural selection will favour either optimistic or pessimistic behaviour respectively. What is notable, is that both optimistic and pessimistic behaviour drive organisms to behave in a different manner that what would be optimal for their given environment. Optimistic individuals will behave as if environmental conditions are better than they actually are, and vice versa. This is relevant because acting in such an apparent irrational way, under certain conditions, proves to yield long term benefits for the genotype of the population by increasing its fitness, at the cost of short term measures such as the number of surviving offspring. Interestingly, this behaviour fits with Olofsson et al. (2009) definition of bet-hedging: lowering fitness variance in order to maximise long-term fitness. This phenotypic heterogeneity is accomplished through stochastic switching, which is randomly switching strategies throughout the entire population, meaning that some individual cells will display one strategy, whereas other individual will display a different one. Stochastic switching has been especially studied within the microbial context. Kussell and Leibler (2005) show that, in principle, adopting a stochastic switching strategy can be more beneficial to the survival of the population than relying on environmental sensing when the

environment changes infrequently. Their simulations highlight the relation between switching rates and environmental change statistics: the switching rates are proportional to the probability of the environment to change and inversely proportional to the average duration of each environmental state. They also point out that if the environment changes too quickly, sensing becomes an optimal strategy over stochastic switching. These ideas have been expanded by subsequent studies. Salathé et al. (2009) study how stochastic switching evolves in microbial populations and how the switching rate evolves in accordance with the environmental fitness landscape, meaning, how costly it is for the organisms to not be well adapted to the environmental conditions. Notably, they found that when the cost of being maladapted to an environment is low (small selective pressure), stochastic switching rates not only decrease but they are not proportional to the environmental rate of change. Another conclusion of their work is that when the cost of being maladapted to the environmental conditions varies with the environmental state, stochastic switching might not evolve at all. Other studies examine the underlying genetic mechanisms that facilitate stochastic switching. As Møller et al. (2013) point out, several environmental parameters can change at once, which presents a great challenge to microbial populations. In their work they propose that in order to tackle such a hurdle, organisms such as *S. cerevisiae* implement what is called a *hub-switch*, which is a genetic or epigenetic switch that triggers several phenotype changes at once. Instead of relying on different genetic or epigenetic events to occur, which would increase the complexity of adaptation to new environmental conditions, a single event that is tied to several phenotypical changes would increase the robustness of the stochastic switching bet-hedging strategy.

Perkins and Swain (2009) discuss how a stochastic switching strategy is present in cellular decision-making as an optimal bet-hedging strategy in situations where elements of the environment are unknown, which is plausible given that a cell's sensing accuracy is not perfect. In addition, without such stochastic switching it might be impossible to obtain any phenotypic heterogeneity within a cellular population,

which is the basis of diversified bet-hedging. A population of cells is phenotypically heterogeneous when cells with the same genetic content express a different phenotype, in other words, activate different genes. This gives rise to phenotypic sub-populations. A relevant instance of heterogeneity presents itself when a cellular culture grows in a medium with two different sugars. Traditional approaches to this classical scenario state that bacteria will perceive the two sugars available to them and proceed to consume their preferred substrate. The sugar consumption is reflected by their exponential growth. Once they consume the first sugar to completion, they will experience a period of slow, if not completely arrested growth rate, during which the bacteria reshape their metabolism in order to consume the remaining sugar by degrading specific enzymes and synthesising new ones. Once they have re-adapted to the second nutrient they experience a second phase of exponential growth. This phenomena is called diauxic growth (Hogg, 2005). This view of diauxic growth assumes that all individual cells go through the same process of reshaping their genetic expression to match the environmental changes. However, recent studies show that a thorough analysis of the individual cellular phenotypes reveals a diversified bet-hedging strategy taking place (Solopova et al., 2013; Koirala et al., 2016). The analysis shows that one sub-population displays the phenotype necessary to consume one sugar whilst a second sub-group presents an alternative phenotype that matches the other nutrient present in the medium. This realisation modifies the conventional view of the diauxic growth. These results suggest that the period of arrested growth is not due to a population-wide readjustment, but rather to one of the phenotypic sub-populations stopping their growth altogether, whilst an already existing sub-population ramps up their growth due to the favourable environmental conditions (Solopova et al., 2013).

The importance of single-cell analysis is highlighted by the fact that, as Boulineau et al. (2013) point out, our knowledge regarding cellular growth behaviour has largely come from bulk experiments that provide information from the entire population leaving single-cell behaviour unclear. Ozbudak et al. (2004a) emphasise that

inferring single-cell behaviour from population techniques can entail misleading conclusions. In their own work analysing bistability in the lac operon, population average showed a gradual activation of the system, whilst individual cells are either fully induced or not at all. Other authors have found similar differences between population dynamics and single-cell analysis underscoring the need to analyse single-cell behaviour (Solopova et al., 2013; Koirala et al., 2016).

Phenotypic heterogeneity can also permit the existence of a division of labour within a cellular population, which can allow it to grow faster (Ackermann, 2015). This particular instance of phenotypic heterogeneity can be asymmetrical, meaning that one cell type displays a phenotype that benefits other cell types, without reciprocity from the benefited cell sub-populations. It can be conceptualised as either a "cheating" strategy, as some cells that do not contribute any resources to an activity, accrue almost if not all of the benefit derived from the work of others, or an altruistic one, as some cells sacrifice their own individual fitness by exhibiting a behaviour that benefits the entire population (Perkins and Swain, 2009; Veening et al., 2008). A perhaps more unambiguously benign instance of division of labour is when it allows a population to overcome the incompatibility between different cellular processes, both of which can be necessary for the population to thrive. This is the case with filamentous cyanobacteria. In order for them to fixate nitrogen, individual cells need to produce nitrogenase, an enzyme which turns nitrogen into ammonia. However, this enzyme is dismantled in the presence of oxygen, which is released during photosynthesis. The strategy that cyanobacteria evolved in order to tackle this conundrum is to segregate these processes into different sub-populations, one that carries on with photosynthesis, and a second specialist group called heterocyst that provides the appropriate oxygen-free environment required for nitrogenase (Adams, 2000).

The study of cellular behaviour from the standpoint of decision-making has, and is still providing great insights about the functioning of the complex metabolic networks that allow cells to make decisions. The analysis of cellular decision-making

with tools and concepts originated in neuroscience and psychology, provides an approach that not only allows researchers to overcome the biological complexity of the cell, but also reinforces the existing notion that the same decision-making mechanisms are shared throughout nature, regardless of the specific biological system.

1.2 Objectives of this thesis

The aim of this project is to identify the underlying principles that guide cells to make decisions from the perspective of decision-making theory. I will focus on the mechanism through which cells decide which strategy to adopt when they are presented with different sugar alternatives; whether to consume them sequentially, in which case they must decide which carbon source to consume preferentially before the other, and under what conditions can cells consume nutrients simultaneously.

The scenario we are interested in studying with this project is the sugar consumption behaviour displayed by *S. cerevisiae*, or baker's yeast. Specifically, we will analyse the behaviour of a yeast culture growing in a medium containing two different sugars, in which both maltose and galactose are available for consumption. We will analyse this system from the value-based decision-making standpoint. In order to do this, we will develop a mathematical model that will be able to replicate experimental data, and which can be analysed as a dynamical system under steady-state conditions. As we have mentioned in this chapter, microorganisms have consumption preferences regarding the sugars they consume (Alon, 2007b). Aidelberg et al. (2014) show that, through metabolic experiments and analysis, different carbon sources can be ranked in a hierarchy of preference, which is specific to the organism. In principle, a carbon source is preferred over another as a function of the growth advantage it provides to the organism. Through the process of evolution, microorganisms have configured their metabolism and sensing mechanisms in such a way as to facilitate the catabolism of their preferred sugars over the non-preferred carbon sources.

This fact is most clearly observed with glucose. As glucose is the preferred carbon source of many organisms, including yeast, entire families of glucose transporters are dedicated to the internalisation of it. The preference for glucose is so robust that in order to increase the efficiency of the glucose transport, yeast synthesises different individual transporters depending on the concentration of glucose detected in the environment. Glucose metabolism also triggers a whole network of regulatory pathways that prevent other sugars from being consumed, ensuring the preferential consumption of glucose (Horák, 2013). These metabolic manifestations of the preference for glucose are a consequence of the growth advantage that microorganisms perceive from consuming glucose. Our objective is to capture the "value" that each sugar has to the organism, through the parameterisation of our mathematical model. The parameterisation is made by fitting the model to the experimental data.

Through dynamical systems analysis, we aim to elucidate the decision-making dynamics that the model is capable of. The results gathered from this approach, primarily through the discovery of bifurcations, permit a qualitative interpretation of the choices available to the system. We expect that through mathematical modelling we can conduct a broader analyses which will allow us to study the expected dynamics in equal-value and non-equal value decision-making scenarios. In this regard, this work is motivated primarily by the studies done in house-hunting honeybee swarms by Pais et al. (2013). We take their approach as a starting point for our analysis.

Our design philosophy aims for simplicity whenever possible, symmetry and non-specificity. The model should be as minimal as possible. The goal is to be able to explain the behaviour that we observe in the experimental data with as little mathematical complexity as possible, whilst still retaining accuracy. Clearly, this is a trade-off, and as we make the model simpler to facilitate manipulation and analysis, which will become a necessity with the dynamical systems analysis, we lose accuracy in the experimental data fit. There must be a balance between the ease of analysis and manipulation, and the model's capacity to replicate the experimental

data. The model must also be symmetric. In this context, symmetry means that, in a binary decision context, both alternatives are described by either equations or sets of equations that are structurally identical and only differ in their parameterisation. In more detail, the system we are interested in is the binary decision between two carbon sources. Most carbon sources are catabolised through distinct metabolic pathways. For instance, the two sugar tested in the experiments we are using to conduct our modelling efforts, have used maltose and galactose. Taking into account membrane transporters, regulatory factors and enzymes, maltose requires three proteins to be transformed into glucose, which then enters glycolysis. Galactose on the other hand, makes use of eight different proteins before it is transformed into glucose-6-phosphate, which is the substrate for the second step of the glycolytic pathway (Horák, 2013). A straightforward approach used in systems biology would be to do a simulation of the entire system, specifying both of these metabolic pathways, taking each biochemical interaction into account and giving it its own mathematical representation (Palsson, 2011). However, this approach would result in a larger system that would steer away from two of our goals: 1) develop and test a model with a decision-making like structure, and 2) we are aiming to develop a model that is as simple as possible. Hence, all the enzymatic and proteic interactions will be lumped up into a single and generalised equation as other authors such as Narang and Pilyugin (2007); Narang (1998) and Chu and Barnes (2016) have done. We expect the different parameterisations of both equations to be enough to produce the behaviour we want to simulate from the experimental data. From a decision-making perspective, the symmetric structure is important to place our model within the context of other decision-making model structures (notably, the canonical decision-making models described by Bogacz et al. (2006)), which are symmetrical in nature. Finally, our model must be non-specific. We aim to construct a model that captures the broad dynamics of the cellular carbon consumption system. Regardless of the specific sugars used in the experiments, the model must be versatile enough that through parameterisation alone, a new decision scenario between other different carbon sources could be modelled. In order to do this, we attempt

not to describe any particular metabolic pathway or substrate-specific mechanism. We aim to describe a general representation of the overall biochemical interactions that allow cells to consume nutrients regardless of their specific chemical identity.

In closing, with this thesis, we hope to find an answer to the following questions: We know from the literature that microorganisms have a hierarchy of nutrients they prefer to feed on (Aidelberg et al., 2014). There is also ample knowledge and evidence regarding yeast's metabolism, and it's regulation. With this in mind, and from the perspective of decision-making theory, what general mechanisms of yeast's metabolism and regulation play a vital role in the system's behaviour and nutrient preferences. By this, I do not mean specific regulatory pathways of signal molecules, but rather, broad and general mechanisms such as inhibition or value sensitivity. We anticipate that this thesis will draw parallels between cellular decision-making dynamics, and the more general principles from decision-making theory originated in other fields of research. We have placed constraints on the structure and the complexity with which we will design our model. With such limitations in place, how accurately can our model fit to experimental data? What details in the experimental data are not-captured by our model? And, what elements of our model are essential to fit the data? Lastly, in the context of decision-making, in a binary decision problem, the agent chooses one of the alternatives, and it is rewarded accordingly. Is this dynamic of unambiguously choosing one of the alternatives over the other present in the cellular system we intend or are the system's preferences expressed differently?

Chapter 2

Literature review

In this chapter I introduce decision-making as an overarching field of study that aims to uncover the underlying mechanisms that allow organisms or groups of organisms to make decisions about their environment and themselves. We go over the biological significance of the study of decision-making as well as some of the strategies that organisms use to facilitate and streamline their decisions in an always fluctuating environment, of which no certain information is available to them. I discuss several different approaches to the study of decision-making: the study of perceptual decision-making in the field of neuroscience and psychology, collective decision-making, nutritional decision-making and value-sensitive decision-making. Going into further detail, I discuss the relevance that value-sensitivity has in the study of naturalistic decisions. Next, I cover the canonical models of neurologically and psychologically-inspired decision-making models, as well as how decision-making has been studied in other systems, such as insect groups and slime moulds, remarking on the significance of multistability within the decision-making dynamics of a system. Additionally, I present a relevant discussion of yeast metabolism, more specifically of the carbon regulation network that allows yeast to control its carbon source consumption, as well as enact decisions regarding its sugar uptake. Subsequently, I go over the main different modelling approaches that have been employed

in order to study and analyse metabolic dynamics.

2.1 Decision-making

The premise of an agent (human or otherwise) having to choose between different alternatives, each one prompting their own set of associated rewards and costs as well as the cost of not making a choice accruing over time, presents a dilemma that has occupied both biologists and philosophers for a very long period of time. A classical example of such scenarios is the Buridan's ass paradox. In this traditional layout, an ass stands at equal distance between water and hay whilst being equally hungry and thirsty. Some other versions have the ass in between two bundles of hay. Philosophers from ancient times such as Al-Ghazali and Aristotle, coming all the way up to our modern times and including the eponymous french philosopher Buridan, have ruminated over the fate of the hesitant animal, or man in some cases (Aristotle, 2006; Chislenko, 2016). The classic outcome of this paradox is that, unable to choose between to equally valued resources, the ass would forever linger in between both options until it dies of both thirst and hunger. We can imagine that, despite of the conclusion presented by this scenario, the ass would not have gone either hungry or thirsty for long. The animal could just choose the pail of water and satiate its thirst, then eat its way through the pile of hay, or vice versa. Just as easily it could alternate between the hay and the water.

The study of decision-making processes, such as the one exemplified by the Buridan's ass paradox, as a field of research aims to understand how agents make their decisions by identifying the underlying mechanisms that allow these processes to take place, as well as to identify the optimal strategies that an agent can take in a given situation in order to maximise their reward, or minimise their loses. Notably, research efforts made by dissimilar fields such as neuroscience, psychology and behavioural biology have shown remarkable parallels in the mechanisms and

interaction patterns employed by different biological systems in order to reach a decision. These similarities have allowed concepts from neuroscience and psychology, originally envisioned to describe the primate brain decision-making, to be applied in the study of other wildly different biological systems such as social insect colonies foraging behaviour, or nest selection for instance (Marshall et al., 2009; Reina et al., 2018).

Decision-making has usually been looked at from the perspective of classical utility theory which dictates that the objective of any decision-maker is to maximise their utility, which means that they are assumed to be rational. However, experimental data from human subjects has proven that humans can be quite irrational decision-makers (Hanks and Summerfield, 2017). As Sasaki and Pratt (2018) point out, in order to be an ideal decision-maker, a subject would always calculate the utility of each option before making a decision. However, this is not what is observed in human subjects. Instead, humans use heuristics and rules of thumb in order to facilitate their decision-making. The implementation of these mechanisms can explain the departure from rationality observed in humans and animals (Hanks and Summerfield, 2017).

In the following subsections we will cover some of the strategies that organisms use in order to facilitate their decision-making processes, generalise rules and deal with the effects of stochasticity. These strategies are heuristics, bet-hedging and stochastic switching, which can be considered as a bet-hedging sub-strategy.

2.1.1 Heuristics

Hutchinson and Gigerenzer (2005) explain that rules of thumb and heuristics are prevalent mechanisms used by organisms to facilitate the decision-making process. They help organisms make decisions with limited amounts of information and prevent overspecialisation in a single environmental scenario. As Tversky and Kahne-

man (1974) mention, heuristics exist in great part to compensate for the degree of uncertainty in the information, as well as in its value, used to make a decision. In humans at least, it has the consequence of reducing the number of complex probability-assessing tasks a subject must do in order to make a choice because they generalise a particular rule so it can be easily applied in different similar scenarios. Microorganisms also seem to employ this approach to expedite their decision-making. *E. coli* cultures have been shown to use practical heuristics in order to calibrate their metabolic response to a specific environment through their regulatory network. This approach that permitted optimal growth rate in some environmental conditions also caused sub-optimal growth under different circumstances (Towbin et al., 2017).

Whether or not an organism uses heuristics or rules of thumb, its decision-making process is related to the generalist-specialist dichotomy in ecology. Generalist strategies provide an advantage to organisms that need to quickly adapt to rapidly and ever-changing environmental conditions, whilst specialist strategies work better when the environment is highly predictable and stable (Henke-Von Der Malsburg and Fichtel, 2018). Organisms can either adapt rapidly to change and sub-optimally benefit in any given environment, or they can optimally take advantage of a specific environmental condition whilst sacrificing adaptability (New et al., 2014).

Strategies that make use of heuristics are usually considered to be generalist strategies. As Hutchinson and Gigerenzer (2005) point out, the implementation of heuristics might loose its advantage in static environments where generalisation is not needed. Heuristics seemed to have evolved as a response to change. In static environments where conditions change at a lower rate compared to more dynamic environments, other strategies which are specific to the current and static environmental conditions might be optimal. This observation reflects the arguments presented by Arkes and Ayton (1999). Their work with human subjects shows that, following a generalist approach by making use of heuristics, can make the subjects fall into sunk-cost fallacies and sub-optimal decision-making, as it sacrifices the capacity to change a decision when it is no longer cost-efficient.

2.1.2 Bet-hedging

The natural environment in which an organism lives changes constantly, both in space and time. The information an organism has about its environment and its changes is not perfect but rather uncertain, noisy and incomplete. Organisms implement bet-hedging in an effort to tackle this uncertainty by displaying behaviours or phenotypes that insure favourable outcomes regardless of the changes the environment presents. The technical definition of bet-hedging is a strategy that reduces arithmetic mean fitness in order to reduce variance in fitness (Ripa et al., 2009; Slatkin, 1974). Olofsson et al. (2009) defines it as lowering fitness variance between years to maximise long-term fitness. This is particularly relevant in changing and unpredictable environmental conditions. Specialisation in a specific set of environmental conditions means that the organism would have high fitness in those specific conditions. This also means that it would do badly if those conditions were ever to change, resulting in high variance over time. In contrast, a generalist would have a much lower fitness variance regardless of the environmental conditions because they are better suited to survive in changing environmental conditions. They never do as well as the specialist, nor as bad in terms of fitness. This concept, which is studied mainly in evolution and ecology, is similar to the risk aversion principle taken from utility theory. As described by Philippi and Seger (1989), the risk aversion principle states that the associated cost of negatively deviating from the mean is larger that the benefit associated to a positive deviation of the same amount. In biological terms, this principle says that what organisms tend to do is avoid risks of death at all costs, even if it means sacrificing an opportunity to thrive in the given environment.

This principle is better appreciated by analysing the three sub-strategies that exist within the bet-hedging purview. According to Olofsson et al. (2009), these are: 1) conservative, 2) diversified and 3) coin flipping. Seger and Brockmann (1987) go into more detail regarding the different bet-hedging definitions, listing up to 5, al-

though they focus their attention on their analogues for conservative and diversified or probabilistic diversification which is the name they give this strategy. Conservative bet-hedging refers to an organism always using the same low investment and low risk strategy. The diversified strategy is what we have described in the previous paragraph, investment in several strategies simultaneously. The third strategy, coin flipping or stochastic switching, is the random switch between different strategies by different individuals (Olofsson et al., 2009). These three different strategies can be exemplified with this perennial conundrum: With limited resources, should an organism produce many small eggs or a few larger (with more chances of individual survivability) ones? Mirroring our three different bet-hedging strategies, the question can be rearranged as follows: Should the organism in question a) always invest in a few large eggs (conservative), b) diversify the egg size each time according to a fixed distribution (diversified), or c) randomly switch the egg size each cycle (coin flipping)? A combination of all three is also a possibility.

Cooper and Kaplan (1982) have a seminal article examining the stochastic switching as a survival strategy. In their work, they show that adopting a coin flipping strategy can outcompete adopting pure strategies, especially in an environment that presents unpredictable temporal fluctuations. In more recent studies that explore the effects of spatial and temporal fluctuations in expanding populations, it is shown that environments that have a fast rate of change are concomitant with the population adopting a single optimal strategy, whereas a phenotypically diverse population can be optimal in slower changing environments (Villa Martín et al., 2019).

2.1.3 Perceptual decision-making in psychophysics

In the preceding section we discussed how decision-making has been a subject of ample study in ecology, evolutionary and behavioural biology, and behavioural psychology. Within the field of psychology and neuroscience, there has been a clear focus on the study and mathematical analysis of perceptual decisions. In this sec-

tion we will cover this as it pertains to naturalistic decisions. We will overview the assumptions being made in the traditional zero-one perceptual decision-making approach as well as the solutions offered by value-sensitive decision-making.

Within the field of decision-making, the concept of perceptual decisions receives a great amount of attention, in no small part due to the advances of psychology and neuroscience. Perceptual decision-making describes the process of observation, analysis and categorisation of information perceived by the senses (Hanks and Summerfield, 2017). As we have discussed in the previous section, these processes are subject to inherent biases and incomplete information.

From an economic viewpoint, the study of perceptual decision reflects on the gains and loses of each decision. The outcome of most decisions made by any organisms are in essence a balancing act between the increasing cost of the decision-making process and the potential gains from making the right choice. This approach to the analysis of decision-making usually places a great deal of attention on the question of optimality. It also implies that, in a two-alternative scenario, there is a correct alternative and incorrect one; choosing the right option in a timely manner accrues rewards, whilst choosing the wrong option incurs costs to the decision-maker. Alternatively, taking too long to choose also represents a cost, independently of the alternative chosen. As Bogacz et al. (2006) mention, one of the simplest scenarios that is considered within the optimality framework is one called the two alternative forced choice (TAFC). The TAFC is a simplified scenario of real life situations in which an agent must make a decision between two alternatives, with uncertain information about the options and with limited time to make the decision.

The decision-making process described by TAFC has three assumptions listed by Bogacz et al. (2006): 1) the decision-maker accumulates evidence on both alternatives, 2) once it has accumulated enough evidence in favour of one of the alternatives, the decision is taken, and 3) the process is stochastic. As Bogacz et al. (2006) and Pirrone et al. (2014) point out, there is a body of behavioural and neurological

evidence that fits within the TAFC framework (Ratcliff and Rouder, 2000; Usher and McClelland, 2001; Ratcliff and L. Smith, 2004). In order to describe and study these processes, many mathematical models motivated by the neurological findings have been produced. The simplest of them is the drift diffusion model (DDM), a very well studied and characterised standard decision-making model. The DDM works precisely within the assumptions listed above, and it has also been shown to describe the optimal mechanism to solve TAFC decision-making (Ratcliff, 1978; Bogacz et al., 2006). Bogacz et al. (2006) present a very complete review of different neurologically-motivated models used to study perceptual decisions and how they can be made to approximate the DDM under specific conditions. I discuss and provide a brief summary of the models discussed by Bogacz et al. (2006) in section 2.2.1.

Analysing decision-making from the framework of optimality theory necessitates to consideration of the effect of the cost of time and decision accuracy. The time to make a decision is limited and the faster the decision is made, the smaller the cost is, whilst the risk of remaining in an undecided state is kept at a minimum. In terms of accuracy, the agent is pressured into choosing the correct option as often as possible, or to make as few errors as possible, in order to maximise the reward. Both of these constraints are presumed to play strong evolutionary roles in the optimisation of different decision-making related mechanisms. Decision-time and error-rate are usually a point of great interest when studying and analysing decision-making models (Bogacz et al., 2006; Pirrone et al., 2014). This raises the conundrum commonly known as the speed-accuracy trade-off: agents that prioritise faster decision-making, inevitably make more mistakes, and agents that aim for more accurate decisions must invest more time gathering evidence and making the decision (Pirrone et al., 2014). In models such as the DDM, this trade-off is managed by the value of the decision thresholds (Bogacz et al., 2006).

2.1.4 Nutritional decision-making

Going back to the Buridan's ass paradox, another solution to the problem is offered by the geometric framework approach. The geometric framework describes an n-dimensional nutrient space that spans n-axes, each one representing a particular nutrient or food-item. In such a nutritional space the current nutritional state of the animal, as well as its nutritional target can be represented by two different points. Hence, the nutritional deficit can be defined as the distance between these two points. In said nutrient space, the animals or agents try to overcome their deficits by making decisions that take them closer and closer to their target (Bose et al., 2017). Originally proposed by Simpson and Raubenheimer (1993), the geometric framework is usually implemented in a way that each axis represents one nutrient type (protein, carbohydrates, lipids, etc). The animal has to choose a sequence of decisions that will take it to its nutritional target by potentially consuming different available food items each with different nutrient compositions.

The cost associated with nutritional decisions are a defining factor that shapes the animal's foraging strategy. If there are no costs associated with switching from one food source to another and the food sources are considered to be always available, there is no unique optimal strategy. However, if one considers the costs of switching between two food items, or the fact that food availability might be in jeopardy, different strategies arise. By considering the possibility of interruption, due to food items becoming unavailable, Houston et al. (2011) have shown that a *switching line* that leads directly to the nutritional target emerges. In the context of a two-dimensional nutrient space, corresponding to two different nutrients, on one side of this line, the subject must consume exclusively food A, on the other side, food B. The strategy that emerges then for a subject that has two different food options and two nutritional deficits, is to reach the switching line, and once on it, both nutrients alternatives are consumed simultaneously at a ratio, in an attempt to be as close to a straight path between the current state and the nutritional target as

possible. Additionally, Marshall et al. (2015) show the effect that the cost associated with switching between the nutrient alternatives and cross-inhibition has on such models. Their study demonstrates that by including cross-inhibition as part of the system's dynamics, the performance of the system improves. The performance is measured by calculating the penalty associated with the remaining deficits once a simulation is interrupted.

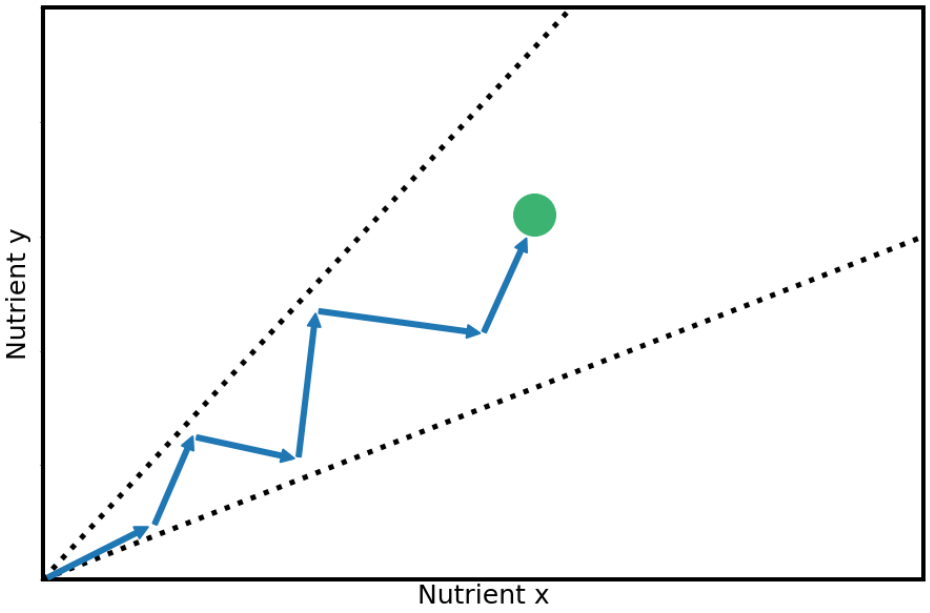


Figure 2.1: Nutrient space for two nutrients x and y. The two dotted lines represent two different food items with different ratios of both nutrients. By switching between the two foods in sequential feeding sessions (indicated by the blue arrows), the foraging animal aims to reach the nutritional target represented by the green circle from its initial position at the origin. An alternative representation measures the minimisation of a nutritional deficit, rather than nutrient intake. In such cases, the movement of the animal targets the origin of the plane, where the nutrient deficit is eliminated.

Bose et al. (2019b) show that such a foraging strategy can be produced with other modelling approaches such as neural decision-making models. Specifically a model derived from the pooled inhibition model presented by Wang (2002). The pooled inhibition model describes three different neural populations, two that integrate evidence for one out of two alternatives, and a third neural population that inhibits the activity of the former two populations. Bose et al. (2019b) show that oscillating

dynamics is much more efficient at reducing both deficits, and with the correct parameterisation this behaviour can be produced with the pooled inhibition model.

The relevance of the geometric framework is not limited to the study of ecology. Recent studies have explored how this approach might be used to make advances in human health and nutrition, as well as poultry farming (Simpson et al., 2017; Cowieson, 2014).

2.1.5 Collective decision-making

Collective decision-making studies how groups of individuals make decisions with no leadership (Bose et al., 2017). The study of the mechanisms through which individuals interact with one another so that the group makes a decision is at the core of collective decision-making. Different alternatives can be supported by different fractions of the population. However, the group achieves consensus by collectively choosing one of the alternatives unambiguously. This is analogous to what is observed in neural systems (Kao et al., 2014). The interaction between individuals with simple rules gives rise to the dynamics of the group. Collective behaviour can give rise to the *superorganism*, a concept that describes a highly interdependent group that acquires the features that would normally be ascribed to the individual, such as group-cognition. Sasaki and Pratt (2018) argue that the concept of the superorganism is most appropriately used with groups in which the fitness of the individual largely depends on the success of the entire group. It is often observed that the decision-making capabilities of superorganisms are superior than those of the individuals within the group (Masuda et al., 2015). Couzin (2009) has done a thorough analysis of collective decision-making, and goes into great detail about how groups can outperform individual decision-making performance and overcome uncertainty and noise. This, however, is not always the case. It has been shown that when the difference in quality between two different nest-sites is large, ant colonies perform worse at making accurate decisions than individuals. The reverse however is

also true, when the difference is small (which is considered to be a difficult decision), the colony does outperform the individual (Sasaki and Pratt, 2018).

Reaching consensus is one of the most relevant challenges of a social group. Failing to reach a single decision as a group can have extremely detrimental consequences for the group and its members. In the case of house-hunting honeybees, failing to reach consensus regarding where to establish a new nest can have potentially devastating consequences for the survivability of the swarm. Lindauer (1957) documents an instance in which an unbreakable deadlock within a honeybee group led the swarm to build its nest in a low-quality location that resulted in the colony being unable to survive the German winter. Inability to reach consensus can also presents the risk of fragmentation the group (Couzin and Krause, 2003). Since some features such as the risk of predation and the accuracy of decisions greatly depend on the number of individuals within the group, fragmentation of the group is an undesirable outcome (Partridge, 1982; Sumpter et al., 2008).

To avoid this, groups performing collective decision-making must manage and overcome the inherent internal conflict and non-compatible preferences in order to reach consensus. The resolution to such obstacles is not always reducible to the majority rule. In fact, as Couzin et al. (2011) highlight, opinionated minorities can drive the group completely. However, groups with uninformed individuals, i.e. individuals with very weak opinions, insulate the group from minority control, restoring the control of the group to the majority.

Another interesting concept is the capacity of groups to learn collectively. Kao et al. (2014) have studied the implication about the way both individuals and the collective learn about their environment, and how this affects the decisions of the collective. Their modelling work showed that the optimal collective decision-making performance was achieved when the individuals learned to associate environmental cues to reward within the social context of their group, as opposed to individuals learning by themselves, which performed poorly.

When talking about collective decision-making, a prescient question arises: How does a group with no leader reach a decision? Probably the most well-known mechanism that facilitates decision-making within a group is majority voting. Considering a binary decision problem, majority voting means that a decision will be taken in favour of the option supported by more than half of the members of the group. The majority rule is considered by some authors as a collective decision-making standard. Hastie and Kameda (2005) point out that in human societies, both ancient and modern, majority voting as a decision mechanism is prevalent.

The majority voting rule is closely related to the Condorcet's jury theorem. Simply stated, the Condorcet's jury theorem argues that in binary problems, the accuracy of a homogeneous group of individuals will approach 1 as the number of individuals increases, if the individuals themselves are good decision-makers (i.e. the individual probability of being correct is larger than 0.5). In contrast, if the individual accuracy is poor (less than 0.5), then the group's accuracy will approach 0. As Marshall et al. (2019a) mention, the Condorcet's jury theorem permeates the way we think about collective decisions. One key interpretation from this theorem is that majority rule is considered to be the best collective decision mechanism a group can have. Marshall et al. (2019a) reach two important conclusions: 1) in many situations, the Condorcet's jury theorem does not hold true, and 2), usually, a majority voting rule is sub-optimal to a sub- or super-majority *quorum* decision rule. In contrast to what the Condorcet's jury theorem predicts, they show that by replacing the decision rule from majority voting to a properly chosen sub- of super-majority quorum decision, a group's accuracy will approach 1, even if the individual accuracy is less than 0.5. Their results highlight the importance of utilising quorums as a collective decision-making mechanism, as it can optimise collective decision-making by improving a group's accuracy beyond what a majority voting rule can provide.

Quorums and *quorum sensing* are widely studied mechanisms in decision-making. Quorum sensing refers to the how individuals within a group start an action only when the number of members of the group committed to the same course of action

reaches a certain threshold (Masuda et al., 2015). Quorum sensing presents itself in a cornucopia of systems and its a core concept in collective behaviour and consensus decision-making. Quorum sensing is present in bacterial populations in the form of individual bacteria communicating with each other through signal molecules. In this way, they can coordinate behaviours that take place at the population level (Balázsi et al., 2011). Quorum sensing in bacteria seems to be so vital that *V. harveyi* appears to have evolved a quorum sensing circuit that reduces noise and individuality when population level behaviour is taking place (Long et al., 2009). Insect colonies are prime systems in which quorum sensing has been documented. Seeley and Kirk Visscher (2004) have documented quorum sensing taking place in house-hunting honeybees swarms just as they decide a new location in which to establish a new hive, and Pratt et al. (2002) have reported quorum sensing during the colony emigration process in *Leptothorax albipennis* ants.

2.1.6 Information theory

In nature, organisms must make decisions in order to increase their chances of survival and reproduction. Choosing the best alternative as often as possible is imperative. However, organisms face a plethora of obstacles that make this task challenging. All decisions any organism makes in nature will be made based on uncertain and incomplete information about an environment that fluctuates across time and space, with limited resources (including the time available to make the decision), and facing competition from other organisms. Despite these obstacles, organisms try to increase the efficiency of their decision-making processes based on the information, limited as it is, that they can gather in order to adapt to the environmental conditions they find themselves in.

A useful approach to study these problems has come from information theory. As Bowsher and Swain (2014) mention, there is a challenge in understanding how cells respond to a fluctuating environment with biochemical mechanisms that are inher-

ently stochastic. They also discuss how, information theory and more specifically, mutual information, can aid in quantifying the influence of stochasticity in these systems. In this context, gene regulation can be viewed as a channel that receives an input x and provides (or maps them onto) an output y. Stochasticity in this channel means that for a given value of x, there is not a single value y, but rather the output is drawn from a probability distribution (Tkačik and Walczak, 2011). If we also consider the inputs to be drawn from a probability distribution, then an important question to ask and challenge to solve is how strongly are the inputs and the outputs dependent on each other. Mutual information is a branch of information theory that has proved to be helpful in this regard. Generally speaking, mutual information relates to the ability to infer a signal from an output, or the amount of information that can be obtained regarding one random variable by observing a second random variable. According to Tkačik and Bialek (2016), a stochastic channel also implies that there is a limit to the amount of information that can be successfully transmitted. In conjunction with mutual information, rate-distortion theory, another branch of information theory which is used to calculate the amount of information necessary to achieve a desired level of accuracy, has been successfully used to study decision-making problems such as chemotaxis, and gradient sensing in bacteria (Andrews and Iglesias, 2007).

2.1.7 Value-sensitive decision-making

In the previous section we have discussed the traditional approach to the study of perceptual decisions, which optimises the reward of each decision based on the speed-accuracy trade-off. One of the drawbacks of this approach to the study of decision-making is that the agents are assumed to optimise a zero-one loss function. In other words, the conceptualisation of decision accuracy implies that only one alternative is correct and the rest are incorrect options.

Bogacz et al. (2006) mention that many scenarios can be simplified to a decision between two alternatives in which the decision-maker has to prioritise either accuracy or speed. This set of conditions referred to as the two-alternative forced-choice (TAFC) task, has prompted a vast number of formal modelling efforts as well as neuro-physiological studies, the results of which support models such as the Drift Diffusion Model (Ratcliff and McKoon, 2008; Usher and McClelland, 2001). Studying decision-making within the TAFC framework has been useful in a plethora of different fields, from ecology to economics, neuroscience and psychology. However, under different circumstances, this approach can be quite limiting in the analysis of naturalistic decisions. As Pirrone et al. (2014) explain, the mechanisms behind this type of decision-making are very well known, in large part because they match the laboratory experimental conditions, where subjects are rewarded only when they make the correct choice. They argue that, although some naturalistic decisions do match this zero-one loss function, in most naturalistic choices there is no correct-incorrect dichotomy, rather alternatives are likely to have different values, and the decision-maker is rewarded according to the value of the chosen alternative, such as choosing a food item to be consumed (Pirrone et al., 2014; Pais et al., 2013). This feature is known as magnitude-sensitivity or value-sensitivity. In relation to decision times, Bose et al. (2019a) mention that value-sensitivity is characterised by decreasing decision times in response to increasing values of 1) the difference between the value of the alternatives, and 2) the average value of the alternatives. This feature has been observed in humans (Pirrone et al., 2018b), monkeys (Pirrone et al., 2018a) and slime moulds (Dussutour et al., 2019). In value-based decisions, and unlike the perceptual decision-making, regardless of the option selected, the agent will always receive a reward, the value of which will depend on the alternative chosen. This is the difference between perceptual discriminatory decisions and value-based decisions (Sugrue et al., 2005).

As Bose et al. (2017) mention in their review, the dynamics found in value-based decision-making have led to the modification of the speed-accuracy trade-off concept,

used in the perceptual discrimination decision-making, into the speed-value trade-off. This concept encapsulates the new dilemma faced by agents performing a value-based decision. Since the decision-maker has to prioritise either decision time (speed) or obtaining highest possible reward (value). In this context, value can be interpreted differently depending on the context of the decision and it is by no means a fixed concept. In perceptual decisions, value can indicate the magnitude of a signal, and in value-sensitive decisions, value can represent a reward, for instance, the perceived energetic content of a food item or the suitability of a potential location for a new shelter. The reward is what the organism actually gets, for instance, the energy contribution of a food item.

Equal-value alternative decision-making

In the study of value-based decisions, an interesting case to analyse is that of a decision between two options of equal value. The aforementioned Buridan's ass paradox describes an instance of equal-value alternative decision-making. An ass at equal distance of two piles of hay that are indistinguishable from each other. In the original outcome of this dilemma, the ass dies of hunger unable to chose one pile of hay over the other because they are equal, both in distance and in the perceived quality. Without knowing much about the mathematics behind value-based decision-making, or the philosophy of intentionality and free will, inherently we doubt that an animal would simply starve when presented with two equally nutritious food items because it could not decide which one is the best. As Chislenko (2016) mentions, there are three main philosophical answers to the Buridan'ass dilemma: 1) denying the possibility that equal-value decisions exist, 2) "biting the bullet" and accepting that the ass must starve, and 3) choosing randomly in order to break the decision deadlock. Randomly choosing one of the two alternatives as a solution to the Buridan's ass was proposed by Rescher (1960), as a solution to problems of choice in the absence of preference as he calls them. This solution to the philosophical reading of the Buridan's ass problem mirrors the findings of the

mathematical analysis of equal-value alternative decisions. Specifically, they are analogous to the results yielded by the modelling efforts that describe the decision faced by house-hunting honeybee swarms. Honeybee swarms selecting a location in which to build their new nest are composed of scout-bees whose job is to explore and evaluate different potential locations. If two alternatives have the same value (in this context value means the suitability of the location to be a nest, determined by the protection it offers from the environment and predators, as well as cavity volume and other structural factors), the swarm faces the same predicament as the Buridan's ass: how to make a choice without preference, or in other words, how to make a decisions if the estimated nest locations have identical value. According to Seeley et al. (2012), and very much like neural decision-making systems, honeybee nest-site scouts not only accumulate evidence in favour of each potential nest-site, but they also deliver stop signals to scouts committed to different sites as a form of cross-inhibition. Cross-inhibition is an important feature of some neural decision-making models, and it refers to the mechanism by which each neural population that integrates evidence for one of the alternatives, inhibits the activation of the competing neural populations (Bogacz et al., 2006). In honeybees, cross-inhibition takes the form of scouts directing inhibition signals (stop-signals which are basically headbutts) at scouts committed to competing sites, whenever they promote their site to uncommitted bees via the waggle dance. In cellular metabolism, inhibition takes the literal form of regulatory metabolic pathways that inhibit the activity of target genes in response to different external and internal chemical signals. Seeley et al. (2012) show that models that lack cross-inhibition will not break their decision deadlock (the ass starves to death unable to decide), whereas models that implement cross-inhibition between the nest-site scout populations, successfully break deadlock by choosing one of the two alternatives at random.

In contrast, in equal-value alternative problems, speed-accuracy decision-making models such as the DDM predict a state of deadlock. As Pirrone et al. (2018b) mention, optimising accuracy and disregarding the absolute value of the alternatives

makes equal- and near-equal-value alternatives decisions, particularly hard within such an approach. Because these models do not take into account the average or absolute magnitude of the alternatives, but rather just focus on the difference between them, as the difference in value diminishes, the system risks a state of deadlock. Additionally, the mean value of the alternatives does not affect the decision times in such systems. Equal-and-high or equal-and-low cases are treated in the same way by value-insensitive systems. As long as the difference in value is the same, high- and low-mean value alternatives will have the same decision times.

Work by Tajima et al. (2016, 2019) on optimal value-based decision-making for binary and multichoice problems puts forward the argument that drift diffusion models present the optimal mechanism to describe such decisions, regardless of the number of alternatives. However, as pointed out in a commentary made by Marshall (2019), their results hinge on the assumption that the time invested in taking the decision has a linear cost, which is subtracted from the rewards gained by making the decision. Cost can refer to the cost of delaying a decision or the investment of resources related to the accumulation of evidence. In the cellular context, a cost would be anything that negatively affects growth rate. Time being a linear cost runs contrary to the reality of most naturalistic decisions, in which costs are considered to increase with the passage of time, rather than remain constant. This has been shown experimentally. Steverson et al. (2019) tested the prediction of models including both a linear (or additive as it is called in their paper) and a multiplicative cost by contrasting them with experimental data. They show that the experimental results they obtained support the hypothesis of a multiplicative cost of time passed. Pirrone et al. (2018a) also found similar results in their experiments with both human and monkey subjects. Additionally, Steverson et al. (2019) also highlight another limitation of approaching value-sensitive decisions with a drift diffusion approach. DDM makes predictions based on the difference in value between the two alternatives, not taking into account the absolute magnitude of them. This does not fit with their results, which show that high-valued alternatives decrease decision

times. This fact only emphasises the limitation of this approach in predicting some naturalistic decisions in which the mean value plays an important role in determining the dynamics of the system. Pirrone et al. (2018a) mentions that value-insensitive models such as the DDM can tackle this disadvantage by implementing additional mechanisms such as collapsing boundaries or asymmetric inhibition. However, these methods do not confer the models with magnitude-sensitivity.

The value-sensitivity approach on the other hand, facilitates the breakage of decision deadlocks because it optimises value of the potential reward to be gained, rather than accuracy within a zero-one loss function problem. The rational of value-sensitivity states that in equal-value cases, a deadlock state can be broken depending on the average value of the alternatives. Equal-and-high value cases have shorter decision times than equal-and-low value cases. This can be easily understood if one imagines an agent choosing between two equally low value food items. In this case, it might be worth it to simply wait in case a better third option presents itself, as shown by Pais et al. (2013). On the other hand, if the subject is faced by two equally high alternatives, it is more beneficial to choose as fast as possible.

Returning to the nest hunting honeybee swarm, varying the strength of the cross-inhibition in this model also affects the dynamics of the system. As Pais et al. (2013) and Reina et al. (2017) show, a swarm that finds itself in a state of deadlock can break out of it by increasing the strength of cross-inhibition over time. This will eventually allow them to randomly choose one of the alternatives. Because this system is value-sensitive, the average value of the options also affects the decision-making dynamics of the system. As Bose et al. (2017) point out, if a honeybee swarm faces a choice between two potential nest locations of equal but low value, then the system will be remain in a state of deadlock. Neither option is good enough to overcome the potential benefit of waiting for a better third option, which is the sensible choice. Experiments with slime mould by Dussutour et al. (2019) also provide proof of another aspect of value sensitivity. They observe that when a slime mould has two food sources at equal distance and of equal quality, then the average

value of the food patches determines the decision time. As the average value of the patches increased, the time it took the slime mould to reach either one of the two food sources diminished.

2.2 Decision-making systems in biology

Decision-making has been studied and analysed in a plethora of different biological systems, from neurons to insect colonies and cells. The broad range of fields in which decision-making principles take place highlights how ubiquitous they are in all of biology. In this section we briefly touch upon the different fields and systems in which decision-making has being studied or applied.

2.2.1 Neuroscience

Decision-making study in neuroscience have focused on producing models of the neural circuits responsible for assessing the alternatives available and make the decision through the activation of different population of neurons. The models have not only been derived from neurological experimental data, but also psycho-physical experiments with human subjects.

Bogacz et al. (2006) present a review of the canonical decision-making models in neuroscience. In what follows we give a brief description of these models based upon the classification presented by Bogacz.

DDM

One of the first and most used formal decision-making models is the drift diffusion model (DDM). Originally proposed by Ratcliff (1978), the DDM has been shown to be very successful in describing a vast quantity of data from different experiments

(Ratcliff and Rouder, 2000; Usher and McClelland, 2001; Ratcliff and L. Smith, 2004; Ratcliff, 1978). Notably, when fitting to experiments such as signal detection, lexical decision and recognition memory, the DDM renders the best fit compared to other neurologically inspired models Ratcliff and L. Smith (2004). As Bogacz et al. (2006) and Marshall et al. (2009) point out, the DDM is a continuous case of the sequential probability ratio test (SPRT), which is discrete in nature. The SPRT has been proven to be the optimal decision-making mechanism for TAFC, as it minimises the decision time for any given error rate (Wald and Wolfowitz, 1948). This means that the SPRT will reach a decision the fastest for any desired accuracy. This property is also present in the DDM. Hence, the DDM is optimal for TAFC decision-making tasks. The DDM describes the integration of the difference in evidence between two different options or hypothesis.

The DDM describes how the difference in evidence supporting either one of the two alternatives varies over time. This accumulation process continues until the difference of evidence reaches a certain threshold, which leads to a decision in favour of the alternative whose boundary was just crossed. The process of evidence integration is described as a continuous random walk process. Equation (2.1) shows what is known as the pure DDM,

$$dx = A\,dt + c\,dW, \tag{2.1}$$

where x is the position of a particle in a time/evidence line where the position of the particle represents the difference between evidence supporting the correct and incorrect alternatives over time. x is also referred to as an integrator unit. An unbiased process starts at $x(0) = 0$. The change in the position of x, or the change in the difference in evidence between both alternatives dx, is determined by the term $A\,dt$, which is the drift that describes the accumulation of evidence in favour of the correct option. The sign of the constant A correlates to which one of the alternatives is the correct one. If A is positive then option 1 is correct, whereas if

A is negative, the correct option is option 2. The term $c\,dW$ is the white noise that affects the process of evidence accumulation. The decision is then reached when the position of x crosses a decision boundary, indicating that the difference between the evidence of both alternatives is large enough to have crossed a threshold in favour of one options, allowing the decision to be taken. Figure 2.2 shows examples of unbiased decision-making processes simulated with the DDM.

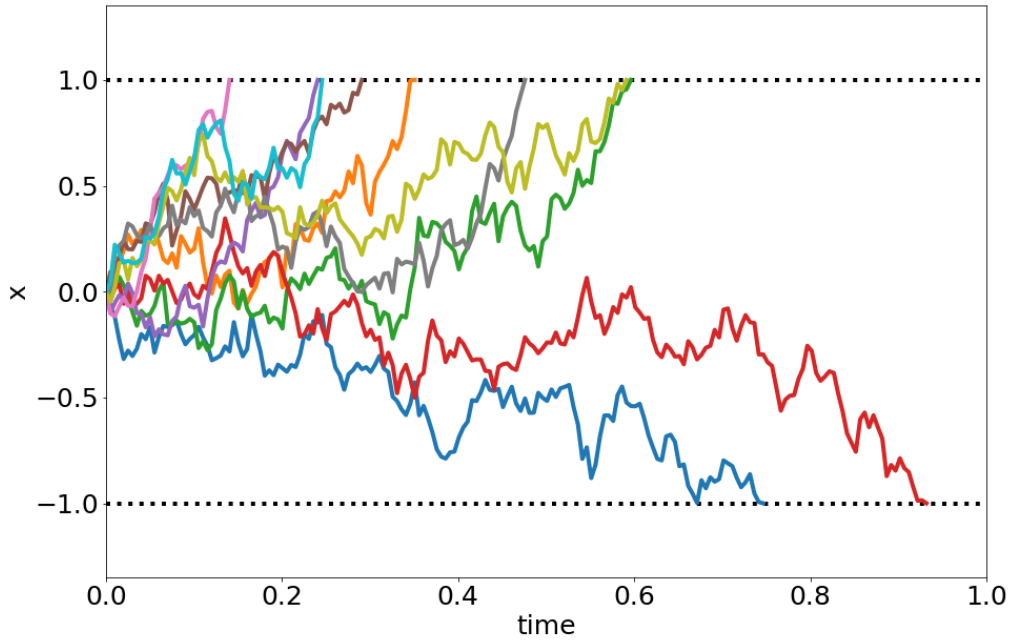


Figure 2.2: Examples of evidence accumulation with the drift diffusion model (DDM). The x is the difference between the amounts of evidence accumulated supporting the two options. Each one of these ten time series corresponds to one decision process taking place. The parameters used are: $A = 1$, $c = 1$, $x(0) = 0$ and with threshold $= 1$. The dotted lines at $x = 1$ and $x = -1$ are the decision thresholds for the correct and incorrect options, respectively. Once a path crosses one of these boundaries the decision is taken in favour of that option.

Race model

The race model presented by Bogacz et al. (2006) is a continuous-time variant of the model described originally by Vickers (1970). The DDM consists in one unit that integrates all the evidence for both options. The race model is two-dimensional, therefore, two units integrate evidence for each of the alternatives independent from each other. The race model is effectively two drift diffusion equations

$$
\begin{aligned}
dy_1 &= I_1\,dt + c\,dW_1 \\
dy_2 &= I_2\,dt + c\,dW_2.
\end{aligned}
\tag{2.2}
$$

In contrast to the DDM (2.1), the race model accumulates absolute evidence rather than the difference between them. Both integrator units y_i represent the evidence accumulated in favour of their corresponding alternative. The $I_i\,dt$ terms modulate the rate at which each units accumulate evidence. Each independent evidence accumulation process is subjected to white noise, set by the term $c\,dW_i$. Two different decision conditions can be applied: a temporal one or an evidence threshold. A temporal decision condition means that at a specific time, the alternative with the higher y_i value is selected. In a scenario with no time constraints, an evidence threshold value exists and whichever alternative reaches this value first, is selected. Unlike the following models, the race model is not neuro-scientifically motivated. However, other race-type decision-making mechanisms have been described in *Bacillus subtilis* cells (Kuchina et al., 2011).

Mutual inhibition

This model consists in four different units that represent four different neuron populations: for each alternative, a neuron group provides evidence in support of their cognate option, and a second neuron group accumulates the evidence provided by the first group. The integrator units in this model inhibit each other's activity directly. The mutual inhibition model reads

$$
\begin{aligned}
dy_1 &= (-k\,y_1 - w\,y_2 + I_1)\,dt + c\,dW_1 \\
dy_2 &= (-k\,y_2 - w\,y_1 + I_2)\,dt + c\,dW_2.
\end{aligned}
\tag{2.3}
$$

The integrator units y_i have an inherent decay of activity, or leak, that takes place at a rate determined by k. The integrator units accumulate evidence at a rate de-

termined by the activity of the evidence-provider units I_i. The integrator units also inhibit each other's activity with strength w. Again, as in the previously described models, the term $c\,dW_i$ sets the contribution of white noise. A decision is made when either unit crosses a threshold value.

Feed-forward inhibition

This model is somewhat similar to the mutual inhibition model (2.3). The difference here is that the inhibition, suffered by the integrator neurons, comes from the input units providing the evidence. The feed-forward inhibition reads

$$\begin{aligned} dy_1 &= I_1\,dt + c\,dW_1 - u(I_2\,dt + c\,dW_2) \\ dy_2 &= I_2\,dt + c\,dW_2 - u(I_1\,dt + c\,dW_1). \end{aligned} \tag{2.4}$$

In this model, there is no decay of activity and the input units perform two tasks: provide evidence for their corresponding accumulators, and inhibit the activity of the competing accumulator. The strength of the cross-inhibition is mediated by parameter u. The decision is taken as soon as one of the two integrator units crosses a threshold value.

Pooled inhibition

This model was originally proposed by Wang (2002). This model adds a fifth unit to the four-unit structure implemented by the mutual and the feed-forward models. The sole purpose of this fifth unit is to inhibit the two integrator neuron populations. The model is written as follows

$$
\begin{aligned}
dy_1 &= (-k\,y_1 - w\,y_3 + \upsilon\,y_1 + I_1)\,dt + c\,dW_1 \\
dy_2 &= (-k\,y_1 - w\,y_3 + \upsilon\,y_2 + I_1)\,dt + c\,dW_1 \\
dy_3 &= -k_{inh}\,y_3 + w'(y_1 + y_2)\,dt.
\end{aligned}
\tag{2.5}
$$

Like the mutual inhibition model (2.3), the activity of integrator units decays over time at a rate mediated by parameter k, and they integrate evidence at a rate defined by the input units I_i. The effect of inhibition comes from the inhibitory unit y_3, not from each other, and its strength is modulated by parameter w. Additionally, the accumulators in this model have a self-excitatory feedback loop that helps to maintain the activity of the accumulator. The inhibitory unit y_3 also exhibits activity decay at rate k_{inh}. Both integrator populations of neurons y_1 and y_2 excite the inhibitory group of neurons represented by unit y_3. The activation of this interaction is weighted by w'. Again, this model reaches a decision whenever the activity of one of the integrator units reaches a determined threshold.

The models presented by Bogacz et al. (2006) can be considered the canonical decision-making model architectures. All of the models are motivated by experimental data from neurological studies as well as psychophysical experiments done in humans, primates and other animals such as rats (Hanks and Summerfield, 2017). Following their development, these models have continued to be used in psychology and neuroscience. Moreover, the architectures and underlying structures of the models can be seen paralleled in the study of other biological systems, such as slime mould (Zabzina et al., 2014), insect swarm behaviour (Pais et al., 2013) and cellular behaviour (Chu and Barnes, 2016; Nené et al., 2012).

2.2.2 Decision-making in insect colonies

The study of decision-making is particularly prevalent in the study of social insect behaviour. In colonies of such social species, individuals are highly related to one

another and work with a high degree of cooperation within the group. This means that the individual interests are aligned with the interests of the entire group.

Quintessential examples of collective decision-making in insect groups are honeybees swarms and ant colonies. They present a notable instance of the concept of the superorganism and group-cognition. As Sasaki and Pratt (2018) mentions, the knowledge of the colony does not depend on knowledgeable individuals within the colony. In fact, individuals only have partial information. It is the interaction between the individuals that allows group-cognition to emerge. This is particularly apparent in the case of foraging colonies in which most individuals only know one of the potential food sources, yet the colonies "know" all of them, and therefore, they can "choose" how to allocate their individuals to the different food options available.

The process through which honeybee swarm chooses a location for their new nest has been the subject of thorough study. Seeley et al. (2012) describe the nest-site selection mechanism in detail: when a honeybee swarm leaves their hive to create a new colony, scout bees leave the group to find potential nest locations. Then, they come back to the group and try to recruit uncommitted honeybees to the nest-site they themselves are committed to through the "waggle dance", a mechanism that allows scouts to recruit uncommitted bees to their site by communicating the direction and distance of the site. The scouts also deliver "stop signals" to dancing scouts promoting different sites than the one they are committed to. This cross-inhibition between the scouts allows the swarm to reach a decision. A scout-bee receiving stop signals reduce the length of its waggle dance. Once enough bees are committed to the same site, the swarm collectively makes the decision and relocates to the winning option to build a new nest. Pais et al. (2013) present a model that explores the dynamics of this process for a two-alternative decision. Reina et al. (2017) expands on the work done by Seeley et al. (2012) in order to produce a N-nest-site selection model for cases with more than two options. It is worth noting that both insect colonies and neural systems seem to preform decision-making through the accumulation of evidence by different population with cross-inhibition

occurring in between them. Marshall et al. (2009) describe the parallels that exist between decision-making in primate brains and insect colonies. Their work shows how analysis techniques used on neural models can also be used to study decision-making processes in insect groups.

Decision-making in ants colonies has been equally extensively studied. Particularly, the behaviour of foraging ants has been illuminating. Ants follow a trail of pheromones left by previous ants, from the nest to the food source. The quantity of pheromones in the trail is in proportion with the quality of the food source. Modelling efforts have successfully modelled foraging ants by accounting for the competition effect that exists between the trails of pheromones. The trails compete for ants in terms of how attractive the trail is for the foraging individuals, which is in proportion the the quality of the food source. Unlike the case with house-hunting honeybee swarms, ants do not necessarily need to chose one food source. Nicolis and Deneubourg (1999) work shows that, under certain conditions, an ant colony can exploit multiple food sources simultaneously. Selection of new nest-sites have also being a subject of great interest in studies on ant colonies. Masuda et al. (2015) explore how a response-threshold model, that includes variability among ants on their level of acceptance of a new nest site, delivers results that are supported by data. Similarly, Sasaki and Pratt (2018) analyse nest selection by ants through the use of neural models, namely, the drift diffusion model, further cementing the existence of underlying decision-making principles that are applicable to all decision-making systems.

2.2.3 Other biological systems

Foraging is a fundamental task that every individual or group of individuals must succeed at, in order to survive. Decision-making in the context of choosing between sources of food has not only been explored in insect colonies, but also in slime mould. Zabzina et al. (2014) present a detailed modelling work of the slime mould *Physarum*

polycephalum foraging for food and having to choose between two identical sources of food. Their model describes collective decision-making that reaches a decision through quorum sensing. Their modelling efforts, which are confirmed by their experimental data, present a scenario in which the slime mould has three courses of action available to itself: it can either consume none of the options, consume one of the two, or consume both of them simultaneously. Similarly, Vogel et al. (2018), also found the slime mould *P. polycephalum* to be capable of multistability. They find that when the alternatives are of low quality, the slime mould would generally exploit them both. However, at high average value, the symmetry breaks and the slime mould makes a decision for one of the two options presented. The slime mould proves itself to be multistable at medium quality values. In this experiment, Vogel et al. (2018) report that at this value, half of time, the slime mould exploit both options, whilst the other half the commit to only one. Both of these articles deal with the concept of symmetry-breaking, which in this context means that a system in a homogeneous environment, transitions into asymmetric behaviour. In other words, the system makes a decision in a two equal-alternatives scenario.

Multistability, or bistability, are features of some systems which indicate that the system, under certain conditions, can find itself in any of multiple different biological states (two in the case of bistability). This property is intrinsic to many systems, including the slime mould in the equal alternative scenario, and very present in cellular decision-making. Some bistable systems present the property known as hysteresis, which implies that the current state of the system, depends on where the system started. This can be conceptualised as a switch. For instance, a switch-like response curve that presents bistability and hysteresis, can be in either its "on" or "off" state depending on the value of the trigger signal. However, the trigger signal threshold value at which the switch from on-to-off takes place, is not the same value at which the switch from off-to-on occurs (Ferrell, 2002; Laurent and Kellershohn, 1999; Ozbudak et al., 2004a). Bistability can also present itself as irreversible switches in which once the switch happens, there is no coming back. The switch-

ing behaviours obtained through bistability are used by cells in apoptosis, cellular differentiation and cell cycle progression (De La Cruz et al., 2015). The maturation of *Xenopus* oocytes (frog eggs), which is decided through the complex interaction between different molecular signals, is such a system. In fact, the maturation occurs thanks to two different molecular cascades that are irreversible switches. *Xenopus* oocytes has also being used to study mitosis or cell division, which occurs as a competition between two different types of enzymes with opposite effects, kinases and phosphatases (Domingo-Sananes et al., 2011). Apoptosis, the process of cellular death, might also be another process that is a one-way, or irreversible, switch, as Tyson et al. (2003) points out. However, new studies have indicated that this might not be the case, and that under certain conditions cells can turn off apoptosis even if the process had already begun. This is particularly prevalent in cancer cells, which presents the challenge of cancer cells that can overcome the effects of chemotherapy (Tang et al., 2009). Non-cancer cells are known to be able to reverse the process of apoptosis, at least at the very early stages (Elmore, 2007; Geske et al., 2001). Ferrell (2002) mentions that in order for a system to present bistability, it is necessary that a positive feedback loop (or the equivalent) exist. It is also necessary that, within the feedback circuit, some form of non-linearity be included.

2.2.4 Swarm robotics

The study of collective decision-making has also been implemented in the field of robotics, specifically swarm robotics and produced significant insights into biologically inspired models (Talamali et al., 2019, 2020; Mitri et al., 2013). Swarm robotics can be particularly insightful and helpful to the research of decision-making whenever the physical environment and the interactions that the individuals from the group have with it are a major factor in the decision-making process. Such components, especially if they are very specific, can be sometimes hard to model, which is why physical experiments with robotic swarms can be quite an effective tool to

study biological systems (Bose et al., 2017).

2.3 Biological background of yeast metabolism

The dynamics of carbon consumption in microorganisms is an essential sub-field of interest within microbiology. Seminal early studies by Monod (1949) as well as Novick and Weiner (1957) were amongst the first to describe nutrient consumption dynamics in microorganisms and establish some of the first concepts that have later been expanded upon throughout the following decades. One such concept that has received much attention is carbon source consumption. More specifically, the behaviour an organism presents when placed in a medium with two or more available carbon sources. As Aidelberg et al. (2014) mention, it is well established that when a population of cells finds itself under such conditions, it can either consume both sugars simultaneously, or consume the sugars sequentially by first depleting the preferred carbon source. Both Beisel and Afroz (2015) and Koirala et al. (2016), point out that, single cells within that population can find themselves in one of three different states: 1) they can be completely un-induced or uncommitted to either one of the sugars, 2) they can be induced to one of the two sugars, or 3) both simultaneously. One of the first examples of sequential preferential consumption was presented by Monod (1949) by showing that, when *E. coli* is presented with a glucose-lactose mixture, the bacteria consumes glucose first, their preferred carbon source, and once depleted it moves on to lactose. Over the following decades several studies have built upon those findings and have shown that carbon sources preferences in microorganisms can be organised in hierarchies based on the growth rate they provide. Aidelberg et al. (2014) presented a sugar consumption hierarchy for *E. coli* based on their analysis of gene expression patterns. When presented with two alternatives, the metabolic genes which correspond to the preferred sugar activate in order to build up the molecular machinery required for its consumption. The high levels of genetic activity detected reflect the commitment of the cell to a high

valued alternative. This commitment is reinforced by the lack of genetic activity present in the genes of non-preferred sugars.

For most organisms glucose is the preferred carbon source and their regulatory networks are usually tuned to respond with great specificity to its presence and metabolism (Stülke and Hillen, 1999). For *Saccharomyces cerevisiae*, commonly known as baker's yeast, the general order of consumption is relatively well documented: glucose is the preferred sugar and it is consumed first over any other available alternative. Once glucose is depleted, any other fermentable sugars are then consumed over any non-fermentable carbon sources, which are consumed last (Schweizer and Dickinson, 1999; Broach, 2012). The carbon catabolism dynamics observed in yeast, as well as the resulting hierarchy, are explained by the interaction of a number of biochemical pathways of cross-regulation by which the catabolism of the preferred sugar represses the uptake and metabolism of less valuable sugars (Broach, 2012; Gancedo, 1998; Horák, 2013; Wang et al., 2015). This order of consumption is maintained by carbon catabolite repression (CCR). This feature makes reference to the universally observed phenomenon that, in the presence of their preferred carbon sources, microorganisms will repress the utilisation pathways of non-preferred carbon sources in order to ensure preferential selective consumption (Görke and Stülke, 2008). As Lengeler (2001) explains, it is the presence of a preferred and rapidly metabolisable sugar such as glucose or fructose that triggers the inhibition of the catabolic pathways of non-preferred carbon sources (Chambers et al., 2004). Stülke and Hillen (1999) and Görke and Stülke (2008) provide comprehensive reviews in which they examine the different mechanisms used by different microorganisms to regulate their carbon metabolism. As Brückner and Titgemeyer (2002) point out, bacterial CCR not only ensures preferential consumption but also limits catabolism according to the cellular metabolic capacities. Brückner and Titgemeyer (2002) also describe different regulatory mechanisms, some of which are sugar specific, whilst others act on a global scale affecting a vast number of metabolic pathways, even those cognate to the triggering carbon source. For instance, *E. coli* growing in lac-

tose will trigger inhibitory mechanisms that autoregulate the expression of the lac operon, the DNA unit responsible for lactose uptake and initial metabolism, thereby reducing the activity of lactose consumption, presumably, in order to not exceed the metabolic capacity of the cell (Brückner and Titgemeyer, 2002; Hogema et al., 1999). In *S. cerevisiae*, CCR pathways also exert a level of autoregulation in order to reduce the transition time once glucose is depleted (Kaniak et al., 2004). Petit et al. (2000) explore how the kinase Hxk2p, which is one of the enzymes that catalyses the first step of glycolysis by phosphorylating the glucose molecule, autoregulates the high affinity glucose transporter. Similarly, maltose and galactose have also been documented regulating their own catabolism (Jiang et al., 2000; Horak et al., 2002).

CCR in yeast, and in *S. cerevisiae* in particular, has been thoroughly studied (Gancedo, 1998; Santangelo, 2006; Carlson, 1999; Horák, 2013; Johnston et al., 1994; Broach, 2012; Schüller, 2003). As is the case with *E. coli*, CCR refers to a rather complex array of varied mechanisms which are often redundant and act at different levels of metabolic activity in response to different signals (Wang et al., 2004; Kaniak et al., 2004; Lane et al., 2018; Vega et al., 2016; Zaman et al., 2009; Novak et al., 2004). As is the case with most microorganisms, including yeast, glucose is also the preferred sugar. Therefore, CCR, or the glucose effect as it is often called, is mainly triggered by the presence of glucose. When glucose is detected, a multitude of mechanisms both repressive and inducing are triggered (Horák, 2013; Jiang et al., 1997; Herrero et al., 1985; Kayikci and Nielsen, 2015; Johnston et al., 1994; Chubukov et al., 2014; Bendrioua et al., 2014; Ozcan and Johnston, 1999). In fact, close to 40% of the entire genome is affected within minutes after introducing glucose into the environment (Broach, 2012; Wang et al., 2004). Reviews by Carlson (1999); Zaman et al. (2009); Gancedo (2008) and Novak et al. (2004) provide great insight into the details of this transcriptional restructuring and highlight the inherent redundancy of the regulatory network. At the level of genetic expression, easily fermentable sugars, usually glucose, silence the genetic expression of less valuable carbon sources, both fermentable and non-fermentable (Schüller, 2003).

For instance, in a glucose/galactose mixture, galactose metabolism is repressed by the Snf1p/Mig1p pathway, the main glucose repression pathway in yeast (Horák, 2013). Johnston et al. (1994) show that the Snf1p/Mig1p pathway is responsible for the 1000-fold inhibition that the galactose metabolism endures when yeast grows on a glucose rich medium. Papamichos-Chronakis et al. (2004) provide an in-depth explanation of how the Snf1p/Mig1p pathway interacts with other regulatory elements in yeast in order to produce its inhibitory effect. The Snf1p/Mig1p pathway is by no means the only regulatory pathway glucose makes use of. Zaman et al. (2009) have a detailed breakdown of the signalling pathways by which glucose represses genetic transcription. The most important among them will be discussed in the text below.

CCR also encompasses mechanisms at the post-transcriptional level that can increase the degradation rate of specific mRNAs as well as affect the synthesis rate of specific proteins. Cereghino and Scheffler (1996) not only explain this process thoroughly, but their findings indicate that for some genes the rate of mRNA turnover might be the most important element of the regulatory process. Andrade et al. (2005) present similar work that shows how lactose uptake is repressed by glucose-triggered mRNA decay. Work done by Federoff et al. (1983) and DeJuan and Lagunas (1986) and others shows how glucose induces mRNA decay as part of its repression of maltose and galactose uptake respectively (Klein, 1997; Hu et al., 2000). At the post-translational level, CCR induces the internalisation of membrane transporters of non-preferred carbon sources, as well as their inactivation by proteolysis, which can occur in the citosol or within a vacuole (Brondijk et al., 2001; Horak and Wolf, 1997). This process is usually called catabolite inactivation. Horak et al. (2002) and Jiang et al. (2000) describe how the metabolic activity of Hxk2p, the enzyme responsible for catalysing the first step in the glycolytic pathway, triggers GAL2p and FBPase proteolysis. As Medintz et al. (1996) mention, transport inhibition is not only the result of the membrane transporters being degraded but also in some cases it is preceded by a phosphorylation that rapidly inactivates the transporter. Novak et al. (2004) also point out that in the case of maltose transporter

Mal61p, the catabolic inactivation seems to respond to both glucose sensing and glucose transport. It is not uncommon to have all different levels of repression acting at once. Glucose-triggered inhibition at all levels of both maltose and galactose metabolisms has been well documented (Novak et al., 2004; Johnston et al., 1994; DeJuan and Lagunas, 1986).

A great deal of attention has been placed on the identity of the glucose signal that triggers CCR into action. Although it has not been elucidated in its entirety, some progress has been made. Carlson (1999) gives an overview on the subject by going through some of the main candidates. A growing amount of evidence points to glucose phosphorylation during the first step of glycolysis as an essential part of the regulatory glucose signal. Specifically, the kinase Hxk2p, which catalyses the first step of the glycolytic pathway, and the product of that reaction, D-glucose 6-phosphate (G6P), have both been thoroughly analysed as potential "trigger molecules" (Meijer et al., 1998; Ahuatzi et al., 2006). Jiang et al. (2000) mentions that early stages of catabolism, specifically the first step of glucose phosphorylation, seem to be required to activate the inhibition effect. Lane et al. (2018) findings strengthen this point by suggesting that G6P might regulate Snf1p, which in turn serves as a global repressor to numerous non-glucose carbon source metabolic pathways (Teusink et al., 1998). Studies by Cereghino and Scheffler (1996), on the other hand, show that glucose phosphorilation, not G6P nor F6P, which is a fructose metabolic intermediate, is enough to trigger post-transcriptional regulation of Ip proteins which are required for succinate metabolism. Other authors have further suggested that Hxk2p plays a central role in yeast CCR. Work done by Andrade et al. (2005) has pointed out that mutants without the Hxk2 gene see their glucose-triggered repression alleviated. Vega et al. (2016) also support the notion that Hxk2p is heavily involved in the repression apparatus, as well as being an internal glucose sensor. An important element to note is that not every pathway that makes up the CCR phenomenon responds to the same signal. Work done by Kaniak et al. (2004) shows that the regulatory pathways Snf3p/Rgt2p and Snf1p/Mig1p, although intertwined with one

another, might respond to different signals. Snf3p/Rgt2p is a glucose sensing mechanism, primarily responsible for activating the synthesis of concentration-specific glucose transporters, as well as the activation of various transcription factors that play a role in CCR (Horák, 2013). They also point out that the cell might benefit from responding to different glucose signals whilst having a coordinated activation of different regulatory pathways. Snf3p/Rgt2p shows activity in response to high glucose concentration, whereas the Snf1p/Mig1p pathway, considered a master regulator, responds to glucose metabolism through its interaction with Hxk2p (Vega et al., 2016). Meijer et al. (1998) have suggested that the Snf3p/Rgt2p sensing mechanism has the capacity to sense intracellular glucose concentration, and it is this that triggers some regulatory pathways. Their results also rule out glucose flux through the membrane as a potential trigger of the inhibitory network. It is rather glucose concentration that activates glucose repression. In addition to this, some studies on glucose-induced galactose inhibition showed that the inhibitory mechanism is triggered by the galactose/glucose ratio, rather than by any absolute concentration (Escalante-Chong et al., 2015; Nguyen-Huu et al., 2015).

Whilst the Snf3p/Rgt2p and Snf1/Mig1p pathways have been characterised as the two main regulatory pathways involved in carbon regulation, the cAMP/PKA or RAS-cAMP pathway, considered to be the third main carbon regulatory pathway, has also been analysed in great depth (Rolland et al., 2002; Tamaki, 2007; Horák, 2013). In contrast to both Snf3p/Rgt2p and Snf1/Mig1p, which exert a rather specific control of the carbon metabolism in yeast, the control that cAMP/PKA pathway has over the yeast metabolism is much broader, affecting many other cellular functions, not only carbon metabolism. As Rolland et al. (2002) mention in their review, cAMP/PKA pathway in yeast is essential to regulate overall metabolic functions, stress response and proliferation. Studies in *E. coli* metabolism regulation support this finding. You et al. (2013) show that in *E. coli*, the function of the cAMP signalling pathway is focused in the global redistribution of proteomic resources. They conclude that the cAMP pathway is not involved in regulating

carbon metabolism. These findings are mirrored in yeast by several authors who have found that the cAMP/PKA pathway is not essential to achieve glucose-induced repression, indicating that at the very least there are cAMP-independent pathways that redundantly inhibit non-glucose carbon source metabolism (Wang et al., 2004; Horak et al., 2002; Eraso and Gancedo, 1984). Diagrams of the regulatory pathways discussed in this section can be found in references Horák (2013); Rolland et al. (2002); Kayikci and Nielsen (2015).

Compared to glucose, the repressive potential of non-glucose carbon sources has received scant examination. Although the study of bacterial systems has unveiled a hierarchy of consumption within non-glucose carbon sources, the mechanisms by which this happens are not yet fully understood (Beisel and Afroz, 2015; Aidelberg et al., 2014; Chubukov et al., 2014). Likewise, studies in *Saccharomyces cerevisiae* suggest that a hierarchy of consumption still exists between non-glucose carbon sources, both fermentable and non-fermentable (Gancedo, 1998; Jiang et al., 2000; Brondijk et al., 2001; Horak et al., 2002; Simpson-Lavy and Kupiec, 2019). Although it is known that the CCR network is not glucose exclusive, and in fact other sugars such as fructose, mannose and galactose have long been known to elicit a repressive effect, the extent to and conditions under which non-glucose sugars trigger a repressive effect is not yet clear. (Sierkstra et al., 1993; Cereghino and Scheffler, 1996; Andrade et al., 2005; D'amore et al., 1989; Hong et al., 2011; Simpson-Lavy et al., 2017). Sucrose, for instance, can interact with upstream elements of the cAMP/PKA pathway, namely Gpr1p, with higher affinity than either glucose or fructose (Lemaire et al., 2004). In particular, galactose has been well documented for its repressive capacity. Both Lodi et al. (1991) and Herrero et al. (1985) show that galactose not only regulates its own metabolism, but can also inhibit other metabolic pathways including the expression of respiration related enzymes. In some cases, these sugars can repress just as much as glucose. Other authors, however, have been unable to find galactose-induced inhibition (Chambers et al., 2004). Horak et al. (2002) reports that, under specific circumstances and in non-wild type strains of

Saccharomyces cerevisiae, galactose does not induce FBPase degradation, which has been reported previously by Herrero et al. (1985), albeit in different experimental conditions. In the same experiments, Horak et al. (2002) shows that maltose represses galactose metabolism, which contrasts heavily with maltose's reputation as a non-repressive sugar. Other studies present strong evidence for maltose-induced repression in different yeast species (Eraso and Gancedo, 1984; Gancedo, 1998). Nevado et al. (1993) document that galactose's ability to repress glucose, fructose and mannose depends greatly on the strain and the growing conditions of the cell population. Metabolic pathways involved in the consumption of non-fermentable carbon sources are also repressed by non-glucose sugars (Herrero et al., 1985; Chambers et al., 2004). Non-fermentable carbon sources exert some degree of inhibition as well. Belinchon and Gancedo (2003) show that non-fermentable carbon sources can repress gluconeogenic enzymes. Similarly, acetate has been shown to exert some repressive effect over ethanol metabolism (Simpson-Lavy and Kupiec, 2019).

Yeast is capable of both respiration and fermentation. Which one it uses depends on the carbon source available. Yeast presents what is known as the Crabtree effect. The Crabtree effect is the fermentation of sugars when oxygen is available and the sugar concentration is high. A condition under which one could expect respiration to occur, especially if one considers that respiration is much more energetically beneficial than fermentation. As Pfeiffer and Morley (2014) point out, there are several evolutionary reasons behind this seemingly energetically counterintuitive behaviour. One explanation is that fermentation might be a better option in order to defend a sugar rich environment since through fermentation yeast produces ethanol which is toxic to other microorganisms. In addition, ethanol can be further consumed by yeast. A second point that Van Voorhies et al. (2012) also highlight, is that although respiration is more efficient, it is slower, whilst fermentation produces less energy at a faster rate, which might give yeast a selective advantage. On this note, Bruggeman et al. (2020) offer a very substantial article on how microorganisms are selected to maximise their immediate growth rate. In this context, seemingly counterintuitive

phenomena such as the Crabtree effect can be explained as strategies that maximise the growth rate of the microorganism.

The importance of single-cell analysis is highlighted by the fact that, as Boulineau et al. (2013) point out, our knowledge regarding cellular growth behaviour has largely come from bulk experiments that provide information from the entire population, leaving single-cell behaviour unclear. Ozbudak et al. (2004b) emphasise that inferring single-cell behaviour from population techniques can entail misleading conclusions. In their work, which focuses on the analysis of bistability in the lac operon, the results from the population average show a gradual activation of the system, whilst individual cells are either fully induced or not at all. Other authors have found similar differences between population dynamics and single-cell analysis, underscoring the importance of differentiating population from single-cell behaviour (Koirala et al., 2016; Solopova et al., 2013; Beisel and Afroz, 2015).

2.4 Previously developed models

Various different theoretical approaches have been used in the past to model and study the behaviour of microorganisms. **High-throughput modelling** approaches such as flux balance analysis (FBA) which use stoichiometric matrices as their input allow the modelling of entire metabolic networks comprising a vast number of metabolic reactions at once (Palsson, 2011; Varma and Palsson, 1994). Such models are typically used to study a particular metabolic steady-state. They use data on metabolic pathways to build their matrices and produce accurate and testable predictions (Daran-Lapujade et al., 2004b). Usually, an objective function, usually growth rate maximisation, is used to find the metabolic state that optimises this function (Nilsson and Nielsen, 2016). Further developments of this approach into dynamical FBA or mass action stoichiometric simulation (MASS) models have extended the applicability and predictive power of stoichiometric approach (Jamshidi

and Palsson, 2010; Mahadevan et al., 2002). Another established modelling strategy is the use of **ordinary differential equations (ODEs)**. This approach has proven to be versatile as it allows the study of individual gene expression, as it has been done in order to analyse the lac operon on numerous occasions (Ozbudak et al., 2004b; Santillán and Mackey, 2004; Yildirim et al., 2004; Noel et al., 2009), the reconstruction of dynamic metabolic networks (Palsson, 2011), and the modelling of entire cellular processes that encompass copious amounts of genes and metabolic pathways into a small number of ODEs (Narang and Pilyugin, 2007; Kuchina et al., 2011). This modelling approach has also led to the use of bifurcation and dynamical systems analysis, which in recent years has provided insights into cellular behaviour (Nené et al., 2012; Zabzina et al., 2014). The work of Nikolov et al. (2010) is an example of an overarching modelling approach that integrates ODE modelling with sensitivity and bifurcation analysis to provide predictive simulations. Similarly, Neelamegham and Liu (2011) provide an example of how modelling and experiments are being integrated in the field of glycobiology.

A significant distinction within the different modelling approaches is that between deterministic and stochastic strategies. A deterministic modelling approach assumes that there are no random fluctuations. The output of such a system is determined by the parameters of the model and the initial conditions. This means that with the same input, this type of model will always produce the same output. On the other hand, a stochastic modelling approach assumes that there is noise, both internal and external, that affects the behaviour of the system. Stochastic models will produce different outputs every time they are run, even with the same initial conditions. It can be argued that, as Perkins and Swain (2009) mention, since life at the cellular level is intrinsically stochastic, a stochastic modelling strategy is always better suited to simulate cellular systems. However, the inclusion of stochasticity adds more complexity to the simulation task, a resource investment that might not always be necessary. El-Samad and Khammash (2010) also emphasise the importance of accounting for stochasticity when modelling cellular systems. Some established

stochastic modelling approaches they mention are the master equation, which accounts for the probabilistic nature of cellular reactions, and the Langevin equation, which usually describes the stochasticity from external sources. As Pais et al. (2013) mention, when one uses stochastic differential equations, such as the Langevin equation, the source of noise is external to the system, whereas the source of noise in the master equation is internal, and dependent on the size of the system. The master equation, however, can prove to be difficult to analyse and solve. Therefore, the usual approach is to numerically approximate the solution to the master equation, via the Gillespie algorithm (Gillespie, 1976). The Gillespie algorithm simulates stochastic chemical reactions systems, with the probability of any given reaction to take place as a function of the concentration of its cognate substrates. Chu and Barnes (2016) present an approach to model a stochastic biochemical network using the Gillespie algorithm. The model consists of coupled chemical reaction equations that are then simulated using a version of the Gillespie algorithm which, in their case, was modified to allow growth and cell division to be modelled, since the original is suited to simulate reactions within a single cell. A plethora of different techniques and mathematical approaches are being use to simulate stochastic systems. Desponds et al. (2019) use a Markov chain, a different probabilistic approach, to study the decision process that takes places in a fly embryo. Similarly, Nguyen-Huu et al. (2015) use a delayed stochastic simulation algorithm (dSSA) to study the response of a minimum galactose metabolism model to changes in its environment. Another useful approach to study stochastic systems is information theory (section 2.1.6). Concepts such as mutual information have been widely used in the context of gene expression and cellular decision-making to address the inherent stochasticity of these systems (Bowsher and Swain, 2014; Tkačik and Walczak, 2011; Tkačik and Bialek, 2016).

Building models that can be used to study and analyse a particular system has become a necessity. However, constructing models and designing specific computational tools in order to analyse a particular system can prove to be labour intensive

as well as technically complex. In the past couple of decades, there have been efforts to the creation of modelling tools that can address the gap between the textual descriptions of biological systems, and the mathematical models that describe those systems as well as the computational tools required to analyse them.

Tools that offer a few specific analytic functionalities are widely available. Tools such as MATCONT and Dynamica, for Matlab and Mathematica, respectively, as well as XXP-AUT offer very robust dynamical systems analysis and bifurcation analysis (Dhooge et al., 2003; Beer, 2016; Ermentrout, 2002). GillesPy, StochSS and StochKit2 provide modelling software that focuses on performing stochastic simulations based on the Gillespie stochastic simulation algorithm (Abel et al., 2017; Drawert et al., 2016; Sanft et al., 2011). More generalised modelling tools that aim to encapsulate a vast array of analytic functionalities also exist. For instance, MuMoT is a highly accessible tool that allows the user to describe a system as a set of chemical reactions, as well as to perform a wide array of analyses such as dynamic systems analysis, bifurcation analysis and stochastic simulations (Marshall et al., 2019b). The emergence and continuous development of such computational tools allows researchers, particularly those that might not have the technical background, to analyse and study their systems with mathematical rigour in an efficient manner.

Most mathematical models use and describe recurring patterns of interaction between different molecular elements. These recurring regulatory molecular circuits have coined the name of *network motifs*. As Alon (2007a) points out in his review, the chief idea behind the network motifs is that each one of them carries out a specific function. Network motifs can be classified based on the information process they perform. Other authors, such as Schaerli et al. (2014), show that the variety between different motifs can be redundant since some motifs show the same dynamical profile or can be reduced to other more simple motifs without loosing their dynamic behaviour. Ferrell (2002); Laurent and Kellershohn (1999) and other authors provide a lot of detail about how different motifs elicit certain bifurcations and other different dynamical properties as well as their biological implications such

as bistability (Ozbudak et al., 2004b; De La Cruz et al., 2015). Tyson et al. (2003) review a variety of different mathematical terms and examples for specific dynamical properties. For examples of articles that go in-depth on specific motifs see (Hart et al., 2013, 2012; Yi et al., 2016; Briat et al., 2016; Shoval et al., 2010).

Chapter 3

Models developed

In this chapter I describe the models that were developed in order to reproduce experimental data and to analyse the decision-making dynamics through dynamical systems analysis. I touch upon some relevant models from literature that have informed our modelling efforts. I continue by presenting the models that we developed, discussing how suitable they are for the purposes of this project. These models follow several of the canonical decision-making structures described by Bogacz et al. (2006) that were discussed in Chapter 2. Additionally I also describe the process of preliminary evaluation through which we screened and selected our model. Furthermore, I provide a small summary of a particular feature of the experimental data in order to better explain the reasoning behind some of our design choices. A full description of the experimental data is provided in the Chapter 4. I finalise by providing an in-depth description and breakdown of the global inhibition model, discussing the notable features we have used to represent key aspects of yeast metabolism, as well as the four different variants of the model that have been analysed in Chapters 4 and 5.

3.1 Relevant models from literature

In this small section I will go over a few models from literature, or just some features of them, that are particularly relevant to our design efforts, and particularly, point out some of the elements that have been incorporated into the design of our own models.

3.1.1 Models by Narang et al. (1997) on *E. coli*

Throughout the years, Atul Narang has developed several models use to study different aspects of *E. coli* growth and metabolism (Narang and Pilyugin, 2007; Narang, 1998; Narang et al., 1997). Specifically, the model by Narang et al. (1997) presents some similarities to ours in the way it models the bacterial system. We have incorporated some elements from one of them, primarily the inhibition term they use in their model, which reads

$$\begin{aligned}
\frac{dc}{dt} &= V_g\, c\, \frac{p}{Km_p + p} \\
\frac{dp}{dt} &= \sum_{j=1}^{2} Y\, V_x\, \frac{x_j}{Km_x + x_j} - V_g\, \frac{p}{Km_p + p} - V_g\, p\, \frac{p}{Km_p + p} \\
\frac{ds_i}{dt} &= -V_s\, e_i\, c\, \frac{s_i}{K_s + s_i} \\
\frac{dx_i}{dt} &= V_s\, e_i\, \frac{s_i}{K_s + s_i} - V_x\, \frac{x_i}{Km_x + x_i} - V_g\, x_i\, \frac{p}{Km_p + p} \\
\frac{de_i}{dt} &= V_e\, \frac{x_i}{(K_e\,(1 + \frac{p}{K_{inh}})) + x_i} + B_e - K_{dcy}\, e_i - V_g\, e_i\, \frac{p}{Km_p + p},
\end{aligned} \tag{3.1}$$

where c is biomass, p is the concentration of precursors, s_i and x_i are the extra- and intracellular concentrations of substrate, respectively, and e_i is the concentration of enzyme or the "lumped" system of inducible enzymes associated with the internalisation and catabolism their cognate sugars. Narang's models and ours bear close resemblance in their structure, however, there are some key differences that are worth pointing out. The last terms of equations dp/dt, dx_i/dt and de_i/dt of

model (3.1) describes the rate of dilution of these cellular components within the cell. In our approach we have decided not to include this element. This decision was made based on the preliminary evaluation explained in section 3.2.3. In addition, we think that without these dilution terms, we could better analyse and evaluate the contribution of inhibition and alternative's value to the decision-making process. Moreover, in our final model (3.24) we have included a the decay term to the biomass equation (B). Because our model couples biomass to the metabolic activity levels through a common resource term, to some extent it could stand for the effect depicted in model (3.1).

In the early stages of our analyses we explored the original as well as modified versions of model (3.1) and other Narang models present in other studies (Narang and Pilyugin, 2007; Narang, 1998). The initial fit to the experimental data was subpar compared with our models, hence, we decided to move forward with the model we developed.

3.1.2 Model by Zabzina et al. (2014)

An interesting example of collective decision-making models is provided by the work of Zabzina et al. (2014). They modelled the foraging dynamics of slime mould *Physarum polycephalum*. A feature that I will highlight from their two-dimensional model is the system size term they use to limit the growth of the foraging branches, since it is very similar to the term we use to represent the competition for a common resource in our own model (3.24). Their model reads

$$\begin{aligned} \frac{dX_1}{dt} &= \phi\,(M - X_1 - X_2)\frac{X_1^2}{k^2 + X_1^2} - v\,X_1 \\ \frac{dX_2}{dt} &= \phi\,(M - X_1 - X_2)\frac{X_2^2}{k^2 + X_2^2} - v\,X_2. \end{aligned} \tag{3.2}$$

The term $(M - X_1 - X_2)$ represents the uncommitted units of the system, and it

is the the common resource that both X_1 and X_2 need in order to increase. M is simply the system size (number of individuals), which decreases as units X_1 and X_2 increase. This in turn reduces the overall output of the entire induction term, thereby slowing the growth of units X_1 and X_2.

3.1.3 Model by Chu and Barnes (2016)

The model presented by Chu and Barnes (2016), which has a similar structural approach as our models, focuses on permease synthesis as a way of analysing the trade-offs that exist between fast adaptation to new environments and rapid growth. The reaction that produces permease P_i with substrate E_i, which is the internal nutrient i, has the following rate

$$\text{leak}_i + a_i \frac{E_i^2}{K_i^2 + E_i^2} \left(1 - \frac{(P_1 + P_2)^2}{K_L^2 + (P_1 + P_2)^2}\right). \tag{3.3}$$

A single aspect of this expression that is of particular interest to us is the way that the current amount of total permease, $(P_1 + P_2)$ limits its own synthesis rate. In this Hill-like equation, the parameter K_L is the saturation capacity of the membrane. As P_1 and P_2 are synthesised the membrane has less available space to accommodate more permeases, which in turn has the effect of diminishing the rate of production.

3.2 Models developed

In this section I present the models that were developed and tested. We considered three different model structures or architectures, which are taken from Bogacz et al. (2006): 1) Mutual inhibition, 2) Feed-forward inhibition, and 3) Global inhibition. I discuss their structure, as well as the limitations. Furthermore, I discuss the developing process for the global inhibition model, as well as describe the relevant

features that make it the most suitable model to study the system we are interested in.

3.2.1 Mutual inhibition model

The mutual inhibition model describes the concentration of two sugar alternatives, as a system of two ordinary differential equations.The model reads

$$\begin{aligned} \frac{dx_1}{dt} &= I_1\, x_1 - \alpha_1\, f(x_1) - \beta_1\, g_1(x_1, x_2) \\ \frac{dx_2}{dt} &= I_2\, x_2 - \alpha_2\, f(x_2) - \beta_2\, g_2(x_1, x_2). \end{aligned} \tag{3.4}$$

In this model we considered three different elements that drive the decision process: the inducing or input term $I_i\, x_i$ represents all the mechanisms that drive the metabolism and/or intake of sugar i, which ultimately drive the system to choose one alternative over the other i. In this particular model, the input term is a first order reaction, the strength of which is modulated by parameter I_i. The rate β_i determines the strength of the inhibition that the competing alternatives exert on each other. Finally, α_i is the rate of decay or consumption of sugar i. In this system, the feedback is captured by the $I_i\, x_i$ term, whilst non-linearity is introduced through the equations $f_i(x_i)$ and $g_i(x_1, x_2)$, which read

$$f_i(x_i) = \frac{x_i^n}{K_i^n + x_i^n} \tag{3.5}$$

$$g_i(x_1, x_2) = \frac{x_1^n\, x_2^n}{(K_1^n + x_1^n)\,(K_2^n + x_2^n)}. \tag{3.6}$$

Some features we are interested in obtaining and analysing with our models are bistability and super-critical pitchfork bifurcations, since they are very powerful tools used to interpret and analyse the decision-making behavioural profile of a

system. These dynamical properties have been shown to present themselves in several models of biological systems such as the decision-making model of house-hunting honeybee swarms, which is the work that has motivated this project (Pais et al., 2013). Similarly, Zabzina et al. (2014) have analysed the decision-making behaviour of slime moulds through dynamical system analysis and the further study of the deadlock-breaking bifurcations of their system.

We find that model (3.4) presents some critical hindrances: 1) In terms of dynamical systems analysis, we are interested in exploring the multistable nature of decision-making systems. However, this model presents little behaviour of interest, since we found no bifurcations in our analysis. Additionally, the bi- and multistable nature of cellular and genetic expression models has been very well documented (Laurent and Kellershohn, 1999; Ozbudak et al., 2004b). Some of these studies have unveiled the mathematical elements required for these systems to exhibit multistability. Particularly, Ferrell (2002) explicitly point out that positive feedback, or autocatalysis, and non-linearity within the feedback motif are essential for bistable systems to emerge. The work of Zabzina et al. (2014) provides a neat example of this. In their model (equations (3.2)), the equations that track the commitment to either alternative consist of a linear decay term and nonlinear positive feedback term. Regardless of its simplicity, the dynamical systems analysis they performed revealed that their model displayed quite complex behaviour in the form of bistability. Note that model (3.4) lacks any non-linearity in its positive feedback terms $I_i\, x_i$, in contrast to the negative feedback terms $f_i(x_i)$ and $g_i(x_1, x_2)$, which are non-linear in nature. 2) The x_i equations represent sugar concentration rather than metabolic activity. Although both are related to one another, metabolic activity is a much more immediate measurement of the system's commitment to either one of the sugar alternatives. Additionally, the experimental data we are interested in replicating describes metabolic activity rather than sugar concentration. We conclude that model (3.4) is unsuitable for our purposes

The Hill function

It is worth discussing the Hill equation, as it is a key element of all our models. Equations (3.5) and (3.6) are Hill functions. These expressions describe biological processes which are activated or repressed by a given substrate, usually a transcription factor or a repressor. Hill equations are widely used to represent gene response curves. The canonical Hill equation for an activator reads

$$f(x) = \frac{V\,x^n}{K^n + x^n}, \tag{3.7}$$

where three parameters of the Hill equation are K, V and n. Parameter K is called the activation coefficient or the dissociation constant. K determines the inflexion point of our curve and it represents the substrate concentration at which our reaction reaches half of its maximum value (Alon, 2007a). The value of this constant has an inverse relationship with the affinity of the reaction for its substrate. The second parameter, V, is the maximum expression level. High levels of expression can only be achieved when the activator x has a much higher concentration than the activation coefficient K. The third parameter is the Hill coefficient n. As Alon (2007a) states, this parameter determines the steepness of the curve. The larger it is the steepest it is. Originally, the Hill coefficient was used to estimate a number of ligands that a specific receptor needed in order to produce a functional response. However, Weiss (1997) comments that the Hill coefficient is not capable of providing such details and instead it should be interpreted as an interaction coefficient that reflects the level of cooperativity in receptor-ligand interactions. In systems biology it represents the sensitivity of a reaction for its substrate or the level of positive binding between receptor and substrate, which determines the steepness of the curve (Palsson, 2011). Traditionally, the Hill coefficient value used to describe biological processes is 1, especially for the microbial growth curves, which results in the classic Monod hyperbolic curve (El-Samad and Khammash, 2010; Weiss, 1997). However, in all models with the exception of model (3.4), we use a Hill coefficient

value of 2 by default, which produces a sigmoidal curve. Our decision is motivated by the analyses of Ozbudak et al. (2004b), whose results show that a mathematical approximation with a Hill coefficient value of ≈ 2 is the value that best approximates their experimental data for the lac operon in *Escherichia coli* (Yildirim et al., 2004; Narang et al., 1997; Yagil and Yagil, 1971). Chu and Barnes (2016) demonstrated that this value was the lowest integer to produce bifurcations with their genetic models. Similarly, we also found a richer dynamical behaviour with a coefficient value of 2, as well as a better fit to our experimental data.

3.2.2 Feed-forward inhibition model

The feed-forward inhibition model describes a four dimensional system of ordinary differential equations. This model describes the concentration of two sugar alternatives, and the activity of their corresponding metabolisms. The model reads

$$\begin{aligned}
\frac{dn_1}{dt} &= -h_1\, p_1\, \frac{n_1^2}{(k_{n1}^2 + n_1^2)} \\
\frac{dn_2}{dt} &= -h_2\, p_2\, \frac{n_2^2}{(k_{n2}^2 + n_2^2)} \\
\frac{ds_1}{dt} &= h_1\, p_1\, \frac{n_1^2}{(k_{n1}^2 + n_1^2)} - \gamma_{s_1}\, s_1 \\
\frac{ds_2}{dt} &= h_2\, p_2\, \frac{n_2^2}{(k_{n2}^2 + n_2^2)} - \gamma_{s_2}\, s_2 \\
\frac{dp_1}{dt} &= a_1\, p_1 - \gamma_{p_1}\, p_1 + b_1\, \frac{s_1^2}{(k_{s1}^2 + s_1^2)}\, \frac{k_{r1}^2}{(s_2^2 + k_{r1}^2)} \\
\frac{dp_2}{dt} &= a_2\, p_2 - \gamma_{p_2}\, p_2 + b_2\, \frac{s_2^2}{(k_{s2}^2 + s_2^2)}\, \frac{k_{r2}^2}{(s_1^2 + k_{r2}^2)}.
\end{aligned} \tag{3.8}$$

In this model, n_i and s_i are the extracellular and intracellular concentrations of sugar i respectively, whilst p_i is the metabolic activity level associated with the metabolism of sugar i. Both equations make use of the Hill equation structure (equation (3.7)) for their induction term, as well as the repressor dynamics in case of the p_i equations. Pragmatically, the metabolic activity levels p_i can be represented by the levels of a key enzyme in the catabolic pathway of the corresponding carbon

source. In this case, we originally considered p_i to represent the levels of active sugar-specific membrane transporter molecules. The intracellular sugar concentration s_i increases proportionally to the cognate metabolic activity p_i which internalises the extracellular nutrient n_i at a maximum rate h_i and with a dissociation constant k_{ni}. The γ_{s_i} parameter describes the rate of sugar consumption, primarily through its catabolism, although this term does include any form of degradation such as dilution.

The a_i parameter is the basal rate of p_i synthesis. Chu and Barnes (2016) call this term leak_i (eq. (3.3)). Parameter a_i is the promoter's expression independent from the effect of either the activating sugar and the repressing sugar. As a brief observation on this point, a_i parameter should be independent from p_i similar to how the leak_i is independent from P_i in model (3.3). The γ_{pi} parameter is the rate of transporter degradation. The last term of the transporter concentration equation is the induction term. It is comprised of two interactions with opposing effects that modify the protein's synthesis rate: the activating effect of the corresponding sugar, and the repressing effect of the competing sugar alternative. The b_i parameter in the expression is the maximum activity rate. Both k_{si} and k_{ri} are dissociation constants to the cognate activating sugar and the repressing competing sugar respectively. They modulate the affinity with which the sugar substrates exert their effect over the expression of p_i, be it an activating or repressing effect. Unlike the mutual-inhibition models described in Chapter 2, the source of inhibition is the input units, not the activity units. Translated to our model, the source of inhibition is s_i rather than p_i.

In contrast with model (3.4), the feed-forward inhibition model (3.8) the induction term and the inhibitory term are integrated as one single term. From the biological perspective, this change better reflects the reality of the molecular biology of genetic regulation. Having an induction term separated from the repressive/silencing terms does not reflect the reality of the carbon catabolism regulation in yeast. In simple terms, a gene's expressed is controlled by a regulatory sequence. This regulatory se-

quence interacts with other molecules that can have either an inducing or a silencing effect in terms of genetic activity. When it interacts with an inducer or activator, the gene's expression is induced (the gene is activated), hence the synthesis of whichever protein that gene codes for increases. On the other hand, if a repressing or silencing molecule interacts with the regulatory sequence, the expression of this gene is reduced, or silenced.

3.2.3 Global inhibition model development

We moved from the two-dimensional mutual inhibition model (3.4) to the feed-forward inhibition model (3.8). The two-dimensional model is far to simple to be a relevant representation of the system. Secondly, the limited results of the initial dynamical systems analysis called for a redesign of the model's autocatalytic terms, from linear to non-linear in order to facilitate the emergence of multistable behaviour (Ferrell, 2002). Additionally, the literature shows that yeast regulates its selection of carbon sources through a complex array of regulatory pathways that seem to affect, with varying levels of intensity, the transportation and metabolism of all different sugars. Similarly, different sugars seem to be able to present a repressing effect at different levels, although there is conflicting evidence about which sugars are considered or to be repressive and to what extent. More details about yeast's metabolism are explored in much further detail in Chapter 2. The feed-forward inhibition model (3.8) represents a scenario in which the sugars in question repress each other directly and specifically. Biologically this would imply that each sugar has a specific regulatory pathway that then acts upon other sugars in a targeted manner. However, the literature suggests that the same network of regulatory pathways affects the metabolic pathways of all sugars. The literature also suggests that this regulatory network can be triggered by various sugars with different degrees of intensity (Lodi et al., 1991; Herrero et al., 1985; Chambers et al., 2004; Horak et al., 2002). Notably, glucose is the most repressive sugar in yeast, since it is also its

preferred substrate (Horák, 2013; Stülke and Hillen, 1999; Chambers et al., 2004).

We decided to explore a different model structure: the pooled inhibition model, which is taken from Bogacz et al. (2006). I discuss this model in more detail in section 2.2.1. In this model, units y_1 and y_2 represent groups of neurons accumulating evidence in favour of their respective alternatives. The activity of these units also activates a third group of neurons, y_3, which is in charge of exerting an inhibitory effect on both decision units y_1 and y_2. It is important to note that both decision units activate the inhibitory unit, which then inhibits them as can be observed in the figure 3.1.

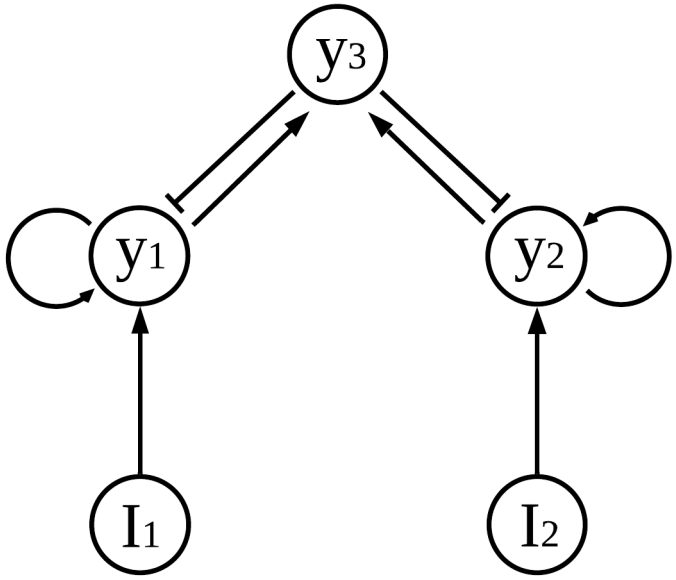


Figure 3.1: Canonical pooled inhibition decision-making model architecture (Bogacz et al., 2006). This model characterises itself for its feedback loops and the inhibition-specific unit y_3.

This structure makes it much more suitable to represent yeast's regulatory network. It also facilitates the analysis of biomass dynamics. This is an important addition, since the experimental data we have worked with includes metabolic data as well as biomass data.

In the following subsections I provide a small summary of a specific aspect of the experimental data that is of particular relevance to our model design. A complete

discussion of the experimental data is provided in Chapter 4. I also provide a brief description of the model development process, which is specific to the global inhibition model.

Grouping pattern in the binary-sugar mixture experimental data

As a compact summary, the experimental data we worked with are time series that display the biomass and the metabolic activity levels of specific maltose and galactose catabolic enzymes in yeast cultures growing over a period of ~ 25 hours in a medium with: 1) maltose, 2) galactose, and 3) both maltose and galactose. In each of these regimes, several yeast cultures were grown simultaneously in media with different concentrations of the sugars. We describe the experimental data in much greater detail in Chapter 4.

An notable feature observed in sugar mixture data is that the metabolic activity presents a grouping pattern. In the experimental results shown in figure 3.2, nine different conditions were tested by pairing up three different concentrations of both maltose and galactose. The fluorescence curves are organised in three groups, each one of them consisting of three curves. The order of the three groups is dependent on the sugar concentration: the higher the sugar concentration, the higher activity level of its own metabolism. The order of the curves within the group, however, appears to correspond to the concentration of the competing sugar: the higher the concentration of the competing sugar; the higher the activity levels of the metabolism of the competing sugar.

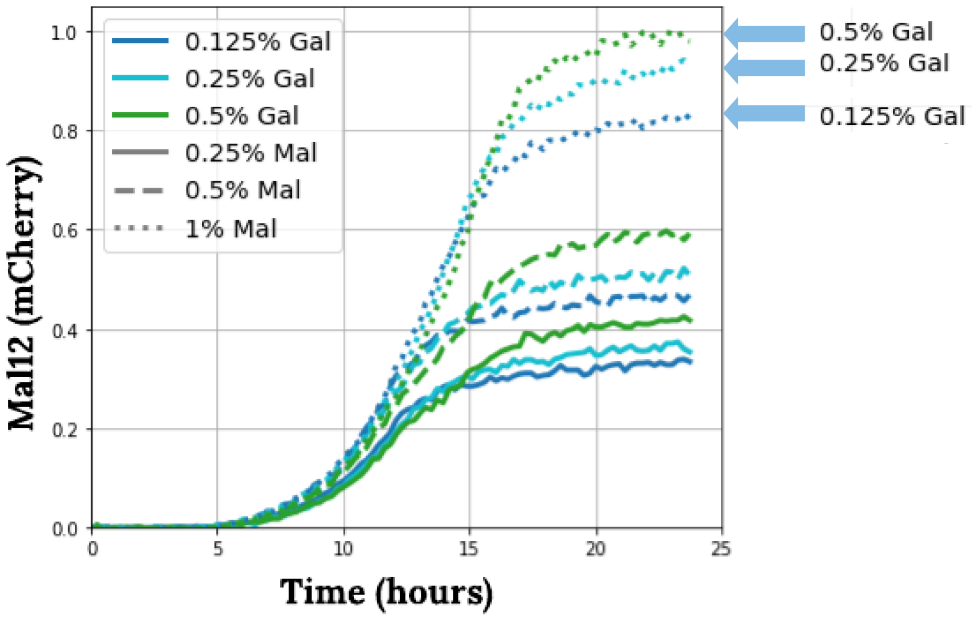


Figure 3.2: Mal12 levels as a function of maltose and galactose concentration in yeast grown over approximately 24 hours. Mal12 is a key enzyme in the maltose catabolic pathway. It splits the maltose molecule into two glucose molecules, which can be further processed. Its activity is traced by tagging with a fluorophore, which in this case is mCherry. This data corresponds to the binary-sugar mixture experiments from 25 March 2017 PR1 (Table A.28).

Preliminary visual evaluation through manual parameter tuning

In order to develop a global inhibition model we followed a simple process for model development which is described in figure 3.3. We began by developing several preliminary candidate models by varying overall structure, interactions between the different variables and specific mathematical terms. All of them, however, shared the pooled inhibition model structure (Fig. 3.1) as the blueprint for their design. Along with the analysis of the preliminary models, we also analysed models from Narang et al. (1997) and Narang (1998), both their original version, in which the Hill coefficient of the Hill-like functions present in the model is = 1, as well as a modified version in which the Hill coefficient is = 2.

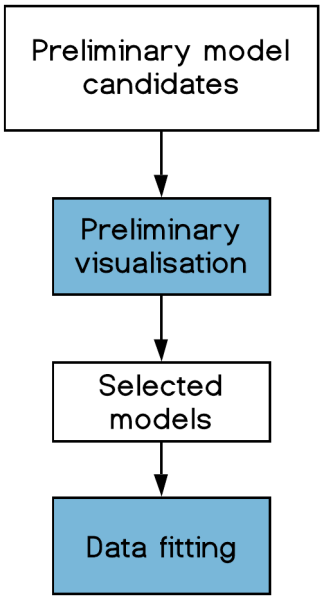


Figure 3.3: Flow diagram of the steps followed to select a model for data fitting.

The preliminary candidate models were then parameterised with parameters taken and calculated from literature. However, the graphical output obtained with such parameterisation is unsatisfactory and it does not resemble the experimental results our models are meant to simulate (the experimental results can be observed in section 4.1.2. Specifically, the experimental results consist of curves of, roughly, sigmoidal shape, whereas the models parameterised with values taken from the literature produced approximately straight lines at $y = 0$ over the same time range as the experimental results. Moreover, taking this parameterisation as a starting point for the fitting process (described in detail in section 4.2) yielded subpar results when compared to the results produced by adjusting the initial parameters. The adjustment can be described as a visual manual fit that consisted on varying parameter values and visually comparing the model's output to the experimental results. The objective of this was to take the model's output from a straight line to a curve resembling a sigmoidal curve, which parallels the experimental output much better than the straight lines produced by the un-adjusted parameterisation. Another prime concern was to explore if the model could replicate the grouping pattern discussed in the previous section. We found that if the model could not reproduce

such behaviour at this stage through visual manual fit, then it would not reproduce it even after the fitting process had taken place. Furthermore, I performed initial dynamical systems analysis on some key parameters that we were interested in, in order to asses the models' dynamic behaviour. The model we chose out of this process followed the structure observed in figure 3.4.

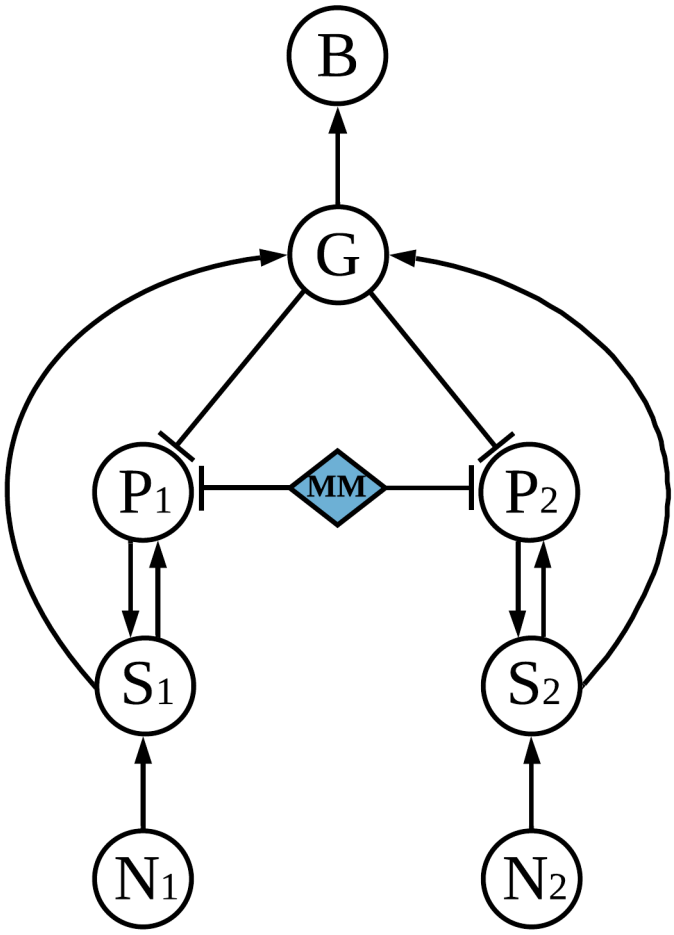


Figure 3.4: Architecture of the selected model.

Global inhibition model for single sugar data

The model yielded by the process detailed in section 3.2.3 is a system of eight ordinary differential equations that reads

$$
\begin{aligned}
\frac{dB}{dt} &= V_B \frac{G^2}{K_G^2 + G^2} - Dcy_B B \\
\frac{dG}{dt} &= Y_1 V_{G1} S_1 + Y_2 V_{G2} S_2 - V_B G \frac{G^2}{K_G^2 + G^2} \\
\frac{dN_i}{dt} &= -V_{Si} P_i \frac{N_i^2}{K_{Si}^2 + N_i^2} \\
\frac{dS_i}{dt} &= V_{Si} P_i \frac{N_i^2}{K_{Si}^2 + N_i^2} - V_{Gi} S_i \\
\frac{dP_i}{dt} &= V_{Pi} A_i - Dcy_{Pi} P_i + V_{Pi} (M - P_1 - P_2) f_i(S_i, G),
\end{aligned}
\tag{3.9}
$$

where N_i is the extracellular carbon source i ($i \in \{1, 2\}$), S_i is the intracellular carbon source i, P_i is the sugar i metabolic pathway activity levels. It can also be thought of as the concentration of a critical enzyme in the catabolic pathway of sugar i, which can then be taken as a approximation of the overall activity of the entire metabolic pathway. This approach is parallel to the one used in the experimental design that produced the data we use to parameterised our models (described in section 4.1.1. State variable G is the global inhibition signal of unspecified biochemical identity but assumed to be related to the accumulation of biomass precursors, similar to the P equation in (3.1) (the non-specificity of the molecule represented by G is explained is discussed in detail in section 2.3). Finally, B is biomass. The induction term $f_i(S_i, G)$ in the P_i equations can take two forms depending on the modelling of the inhibition: 1) It can be represented as a Hill function with substrate G acting as a repressor

$$
f_i(S_i, G) = \frac{S_i^2}{K_{Pi}^2 + S_i^2} \frac{K_{inhi}^2}{G^2 + K_{inhi}^2},
\tag{3.10}
$$

or 2) it can be integrated into the activating Hill function with substrate S_i, by negatively affecting the affinity of the substrate:

$$f_i(S_i, G) = \frac{S_i^2}{\left(K_{Pi}\left(1 + \frac{G}{K_{inhi}}\right)\right)^2 + S_i^2}. \tag{3.11}$$

Both induction terms considered yield similar results regarding fitting to the experimental data. However, they have different implications biologically as well as in the dynamic systems analyses results. The first term (3.10), describes a process in which S_i acts as a inducer molecule and G as a repressor molecule. In this term, the activation of P_i can be though of as a process in which both molecules, activator and repressor, exert their effects over P_i's synthesis independently from each other. In contrast, the second term (3.11), which takes the inhibition term from the model presented by Narang et al. (1997), integrates the inhibitory effect of G and the inducing effect of S_i into the same term. This term implies that instead of G and S_i exerting their effects independently, G inhibits the expression of P_i by reducing the affinity of S_i. In other words, G increments the value of the effective affinity constant.

In terms of fitting the model to the experimental data sets, model (3.9) satisfactorily fits single sugar data sets. However, it is incapable of properly reproducing the metabolic activity curves of the binary-sugar mixture data sets, due to the grouping pattern's complexity. Although the curves are properly separated into groups by the concentration of their cognate sugar, the order of the curves within the groups is inversely proportional to the competing sugar concentration. This pattern is the exact opposite to that observed in the experimental data in figure 3.2. Sugar mixture data suggests that there is positive interaction between the two metabolic branches. The concentration levels of either maltose or galactose's metabolic enzyme seem to respond positively to the concentration of not only their corresponding sugar, but also the competing sugar, to a minor degree. Model 3.9 does not have any interaction between the two metabolic branches, as can be seen in both the equations of the model, and its structure (Fig. 3.4).

The last stages of models development, which I describe in the following subsections, are focused on modelling the positive cross-interaction between the two branches that is observable in the data, and the selection of a common resource term.

Cross-interaction between the branches

With the goal of accounting for the grouping pattern observed in the binary-sugar mixture data (fig. 3.2), whilst retaining the model's capacity to fit to single sugar data, several modifications to model (3.9) were considered. I performed a preliminary manual fit of the model iterations to the sugar mixture data.

Positive cross-interaction between the branches was considered to be present at two levels: Sugar internalisation and metabolic activation. Biologically, the former represents the idea that the membrane transporters that allow sugar molecules access to the inside of the cell for their subsequent metabolisation, are not specific, and both sugars can make use of both membrane transporters, albeit with different affinity constants and internalisation rates. Cross-interaction at the metabolic activation level implies that both sugars, once inside the cell, can activate both metabolic pathways to a different extent. We assume that sugar i is the main substrate of its own metabolism, and a secondary substrate for the metabolism of other sugars.

Implementation of cross-interaction at the sugar internalisation level is accomplished by modifying both S_i equations as follows:

$$\begin{aligned}\frac{dS_1}{dt} &= V_{S1}\,P_1\,\frac{N_1^2}{K_{S1}^2+N_1^2} + V_{S2}\,P_2\,\frac{N_1^2}{K_{SI1}^2+N_1^2} - V_{G1}\,S_1\\ \frac{dS_2}{dt} &= V_{S2}\,P_2\,\frac{N_2^2}{K_{S2}^2+N_2^2} + V_{S1}\,P_1\,\frac{N_2^2}{K_{SI2}^2+N_2^2} - V_{G2}\,S_2.\end{aligned} \tag{3.12}$$

Based on the results from preliminary analysis, we concluded that the metabolic activation type of cross-interaction was sufficient to render the grouping pattern observed in the experimental data. At the metabolic activation level, two different

terms were considered. Whereas the induction terms $f_i(S_i, G)$ (3.10) and (3.11) have one activator, which is their corresponding sugar, and one inhibitor, which is the global inhibitor G, the modified induction terms $f_i(S_1, S_2, G)$ are functions of both sugars, which act as activators, and of G which acts as an inhibitor. Both of these terms (3.13) and (3.14) continue to use the Hill function (3.7) as basis for their structure. The nature of the competing sugar activation is what differentiates the two induction terms: In the term $RFO_1(S_1, S_2, G)$ (3.13), the activating effect of the competing sugar S_2 is a first order reaction, which means that the reaction depends on the concentration of only S_2. In the second induction term we use, $RSO_1(S_1, S_2, G)$ (3.14), the activating effect of competing sugar S_2 depends on the concentration of both S_1 and S_2, making it a second order reaction. Both inducing terms read

$$RFO_1 = \frac{(S_1 + S_2\, I_1)^2}{(S_1 + S_2\, I_1)^2 + (K_{P1}\,(1 + \dfrac{G}{K_{inh1}}))^2} \tag{3.13}$$

$$RSO_1 = \frac{(S_1\,(1 + S_2\, I_1))^2}{(S_1\,(1 + S_2\, I_1))^2 + (K_{P1}\,(1 + \dfrac{G}{K_{inh1}}))^2}. \tag{3.14}$$

The new parameter I_i is introduced in both of these terms to mediate the activating influence that the competing sugar has. Similarly to the inducing terms with no cross-interaction (3.10) and (3.11), the inhibitory effect of G in both RFO_i and RSO_i terms can be modelled as a Hill function, or directly integrated into the affinity constant, similar to the term used by Narang et al. (1997). The equations displayed here are shown with the latter approach. With this approach, the effective affinity is inversely proportional to the amount of G. Parameter K_{Pi} is the affinity constant that P_i has for its substrates, whereas parameter K_{inhi} moderates the inhibitory effect that global inhibition signal G exerts. The higher the concentration of G is, the lower the affinity of the metabolic pathway for its substrates is, which in turn lowers the levels of P_i.

Equation (3.15) shows how the RSO_1 term looks like with Hill-like inhibition. This approach describes G as a direct repressor of metabolic activity, independently of its substrate and its affinity constant.

$$RSO_1 = \frac{(S_1\,(1 + S_2\,I_1))^2}{(S_1\,(1 + S_2\,I_1))^2 + K_{P1}^2}\,\frac{K_{inh1}^2}{K_{inh1}^2 + G^2}. \tag{3.15}$$

The P_i equations represent the lumped sum of metabolic pathways that are related to the internalisation and consumption of their corresponding sugars, we believe that the integrated inhibition captures the array of inhibitory mechanisms in yeast discussed in Chapter 2 better than the direct inhibition of the Hill-like term. Additionally, both approaches were equally successful in fitting to both the single sugar and the binary-sugar mixture experimental data in our preliminary analysis. However, they showed significant differences in the initial dynamical systems analysis. The integrated inhibition of equations (3.13) and (3.14) provided a much richer dynamical profile, which is why we chose this approach to model inhibition.

Common resource term

Allocation of finite resources to a given task implies a reduction of the available amount of the same resources to carry out other processes. The idea to use a competition for finite resources is present in various modelling efforts. Chu and Barnes (2016) use the surface saturation capacity for permeases as a finite resource that limits the synthesis of new permeases in their system. Modelling the foraging of slime mould *Physarum polycephalum*, Zabzina et al. (2014) make use of a term that, in their case, keeps track of the number of uncommitted units within the system. Only by recruiting from this pool of uncommitted units, can either pool of committed units to either option continue to grow toward the food alternative they have committed to. Similarly, Pais et al. (2013) also uses the fraction of uncommitted bees to modulate the growth of sub-populations of honeybees committed to a potential

nest site in which the swarm can establish itself by building a new nest.

In a cellular context, some common resources that constrain cellular processes, as described by Weiße et al. (2015), are finite levels of energy, finite levels of ribosomes for protein synthesis, and finite levels of cellular mass. In their review, Chubukov et al. (2014) point out how the TOR pathway in yeast might mediate translation rates by controlling ribosome biogenesis in response to amino acid concentration. Towbin et al. (2017) also point out that resources such as ribosomes increase linearly with growth rate. Furthermore, the study of cell economics shows that growth rate maximisation is constrained by the trade-offs of synthesising specific proteins and enzymes, citosolic space as well as the capacity of the ribosomes (Molenaar et al., 2009). On that point, Bosdriesz et al. (2015) points out that growth rate is dependent on ribosome concentration, and how efficiently ribosomes are being used. Additionally, they mention that because protein concentration changes minimally, the synthesis of certain proteins comes at the expense of others. The induction terms of the P_i equations in model (3.9) are multiplied by a common resource linear term. We considered and tested four different variations of this term, which read

$$\begin{aligned}\frac{dP_1}{dt} &= V_{P1}\,A_1 - Dcy_{P1}\,P_1 \\ &\quad + V_{P1}\,(M - P_1 - P_2)\,f_i(S_1, S_2, G)\end{aligned} \tag{3.16}$$

$$\begin{aligned}\frac{dP_1}{dt} &= V_{P1}\,A_1 - Dcy_{P1}\,P_1 \\ &\quad + V_{P1}\,(M\,B - P_1 - P_2)\,f_i(S_1, S_2, G)\end{aligned} \tag{3.17}$$

$$\begin{aligned}\frac{dP_1}{dt} &= V_{P1}\,A_1 - Dcy_{P1}\,P_1 \\ &\quad + V_{P1}\,(M\,G - P_1 - P_2)\,f_i(S_1, S_2, G)\end{aligned} \tag{3.18}$$

$$\frac{dP_1}{dt} = V_{P1}\, A_1 - Dcy_{P1}\, P_1 + V_{P1}\,(1 - \frac{(P_1 + P_2)^2}{(M\,B)^2 + (P_1 + P_2)^2})\, f_i(S_1, S_2, G). \tag{3.19}$$

With each of these equations, the units of parameter M are modified to accommodate the changes made to the common resources term. Model (3.9) used common resource term (3.16) for the analysis performed. The non-linear common resource term (3.19) is taken from the model by Chu and Barnes (2016), whereas the linear terms are slight modifications of the term used by Zabzina et al. (2014). The effects of the common resource term were tested with RFO_i (3.13) as induction term. The results of fitting the model to experimental data both single sugar and binary-sugar mixture showed a great degree of homogeneity, particularly between terms (3.17) and (3.18). I decided to choose term (3.17) to be used in the final model. In this term, the increase of biomass B entails a proportional increase of the finite resource amount. A metabolic pathway would inevitably necessitate a resource which is finite and in demand by the catabolic pathways of competing carbon sources, as well as other metabolic processes. Therefore, this common resource term plays an important role in the overall dynamics of the system.

Cellular growth functions

We use two different equations to describe the rate of biomass growth. The first equation makes use of a Hill function (Monod, 1949). This equation reads

$$\frac{dB}{dt} = V_B\, \frac{G^2}{K_G^2 + G^2} - Dcy_B\, B. \tag{3.20}$$

At first glance the apparent zero order growth kinetics might be cause for objection. We argue that, through the structure of our model (Figure 3.5b), biomass B levels indirectly affect its own growth rate. B grows by using G as its substrate, the levels of which increase proportionally to the concentration of intracellular sugar S_i.

Subsequently, the S_i levels increase proportionally with P_i. P_i activity is repressed in three different ways: 1) its own decay, 2) the inhibition effect of the global inhibition signal G and 3) the constraint imposed by the common resource term, and more specifically by parameter M, and biomass level B. The chosen common resource term (3.17) shows that the common resource constraint increases proportionally with biomass levels B. Increasing levels of B alleviate the common resource constraint. The larger the common resource $M * B$ value is, the larger the P_i values are. Larger P_i values increase S_i, which in turn increase G. Ultimately, increasing biomass B induces its own growth by increasing the amount of common resource required by both metabolic pathways. This process, called global feedback, is properly described as the coupling of the cellular growth of individual cells with all gene expression regardless of the status of specific signalling pathways (Shahrezaei and Marguerat, 2015). Kargi (2009) presented a similar approach in their work reducing Monod kinetics to the logistic equation, effectively describing batch growth kinetics in terms of the limiting substrate in lieu of the current biomass value. Secondly, we have used the logistic growth equation as activation term using $G * K_G$ as an indicator of the the carrying capacity of the system (Weisstein, 2019). This equation reads

$$\frac{dB}{dt} = V_B\, B\, \frac{(G\, K_G - B)}{G\, K_G} - Dcy_B\, B^2. \tag{3.21}$$

Now, although the logistic equation itself already accounts for deaths within the population (Weisstein, 2019) we found better results when we included an explicit decay term. The second order kinetics used here are reminiscent of models used to describe migration dynamics in insect populations (Murray, 2002).

It can be observed that in both equations, the two terms differ in the degree of the state variable B. We had much better success fitting the OD data when the decay term had a higher degree than the growth term. If both terms had the same degree, the equation would not fit the data properly. We decided to try two combinations. The Hill function growth equation had a growth term with a zero degree and a decay

term with degree one. The logistic growth equation's growth term has a degree one, whilst its decay term's degree is two. The logistic growth equation does include the state variable B within the fraction part of the term, but the observation still stands between the decay and the $V_B\,B$ part of the growth term. Again, we found that this arrangement was necessary to achieve better fits. This arrangement, however, does change the units of equation B. In particular, the B equation with the Hill function becomes unitless, which is unorthodox. We decided to keep this equation and test it regardless. The units of the B equation with the logistic function are conventional, as can be seen in table 3.1.

It is also important to note that both the B equation and the decay term of the equation G change in unison. Particularly, the decay term $f_{dcy.G}(G, B)$ is very similar to the activating term in B, $f_{ind.B}(G, B)$. However, it is not identical, rather, the non-linear term serves as the rate of the reaction, which is also a function of a different state variable depending on the equation that the term finds itself into, B or G. The two equations we use for biomass, the Hill like function and linear decay equation (3.20), and the logistic growth with quadratic decay equation (3.21), have their analogues in equation G, which read as follows

$$f_{ind.B}(G,B) = V_B\,\frac{G^2}{K_G^2+G^2} \qquad \rightarrow \qquad f_{dcy.G}(G,B) = V_B\,G\,\frac{G^2}{K_G^2+G^2} \tag{3.22}$$

$$f_{ind.B}(G,B) = V_B\,B\,\frac{(G\,K_G-B)}{G\,K_G} \qquad \rightarrow \qquad f_{dcy.G}(G,B) = V_B\,G\,\frac{(G\,K_G-B)}{G\,K_G}. \tag{3.23}$$

Table 3.1: State variables units. State variable B has to different units depending on the $f_{ind.B}(G, B)$ function used.

Variable	Units
N_i	gr/gdw
S_i	gr/gdw
P_i	gprot/gdw
G	gr/gdw
B	unitless (eq. (3.20)) and gdw/L (eq. (3.21))

3.2.4 Global inhibition model with cross-interaction

I have taken model (3.9), which is successful in reproducing single sugar data, and modified its induction terms to account for the grouping pattern observed in the binary-sugar mixture data. The global inhibition model with cross-interaction reads

$$\begin{aligned}
\frac{dB}{dt} &= f_{ind.B}(G, B) - f_{dcy.B}(B) \\
\frac{dG}{dt} &= Y_1\, V_{G1}\, S_1 + Y_2\, V_{G2}\, S_2 - f_{dcy.G}(G, B) \\
\frac{dN_i}{dt} &= -V_{Si}\, P_i\, \frac{N_i^2}{K_{Si}^2 + N_i^2} \\
\frac{dS_i}{dt} &= V_{Si}\, P_i\, \frac{N_i^2}{K_{Si}^2 + N_i^2} - Y_i\, V_{Gi}\, S_i - V_B\, S_i \\
\frac{dP_i}{dt} &= V_{Pi}\, A_i - Dcy_{Pi}\, P_i + V_{Pi}\, (M\, B - P_1 - P_2)\, f_i(S_1, S_2, G),
\end{aligned} \tag{3.24}$$

where $f_i(S_1, S_2, G))$ is the induction term of P_i and can take two forms which are described in section 3.2.3: RFO_i (3.13) or RSO_i (3.14). This model (3.24) and (3.9) are virtually the same. The only differences between them are the P_i induction terms, the decay of equations S_i, which is slightly modified, and the common resource term. Additionally, biomass equation B, as well as the decay term for global inhibition unit G, can take two forms, which are introduced and discussed in subsection 3.2.3. The units of each state variable are shown in table 3.1.

This model structure bears close resemblance to the pooled inhibition model, one of the canonical decision-making model architectures used in neuroscience (Bogacz et al., 2006). The pooled inhibition model in neuroscience was derived originally

by Wang (2002) based on neurophysiological observations. It describes two neural populations, y_1 and y_2, each one of them accumulates evidence in favour of one alternative. Each decision unit has a self-activating feedback loop in order to maintain their activity levels. They represent individual neurons being activated within the population. The way both populations compete is by activating a third neural group which inhibits both decision neuron groups, y_1 and y_2. Figure 3.5 displays the structure of model (3.24) as well as that of the pooled inhibition model, highlighting the close dynamic resemblance. It is worth noting that other structures such as the race, feed forward and mutual inhibition models can also be used by cellular networks (Narang, 1998; Narang and Pilyugin, 2007; Ozbudak et al., 2004b; Kuchina et al., 2011).

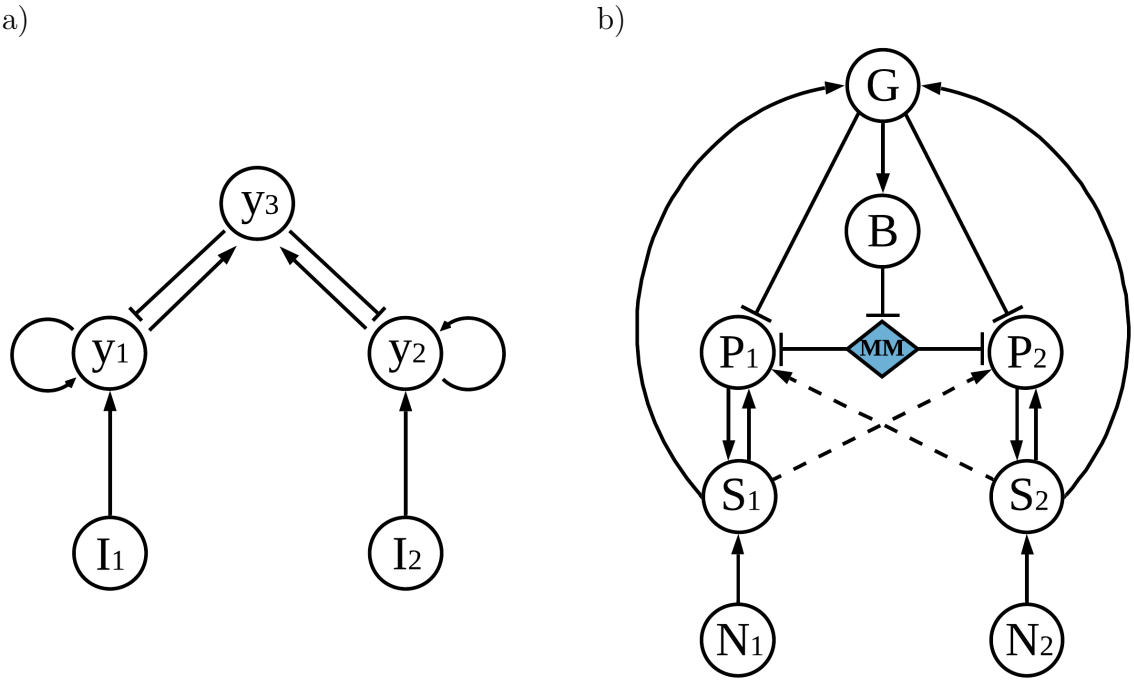


Figure 3.5: Model architecture comparison. a) The pooled inhibition decision-making model used in neuroscience (Wang, 2002). The response units y_1 and y_2 are activated by the input units I_i and by themselves through a feedback loop. The response units also activate the inhibitory unit y_3 which in turn represses the levels of both response units. b) Model 3.24. Although our model has more state variables and more complex interactions between them, the overall dynamics of the pooled inhibition model are maintained.

Lastly, it is important to note that, although our model comprises different processes that proceed at their respective timescales, and that time-delays can be present between them, our model does not include delays nor timescale separations between

the various state variables. Work from Santillán and Mackey (2004); Nguyen-Huu et al. (2015) and Nikolov et al. (2010), do implement time-delays in the modelling of biological systems. In particular, Nikolov et al. (2010) present a detailed analysis of the catabolite repression of the lac operon that emphasises the time-delay between the processes their system encompasses. In the interest of simplicity, we omit such elements from our models and their analyses. We expect that this allows us to study better the direct interactions between state variables and features such as inhibition.

3.3 Conclusions

The aim of the model we have developed is to study the cellular decision-making process between two carbon sources. Our goal is to generate a single model capable of reproducing experimental data from both single sugar and binary-sugar mixture regimes. It is important to note that the models are, by design, generic and symmetric, which, at least in theory, should allow the use of such models in similar cellular decision-making scenarios regardless of the metabolic pathways in question or the microorganism. In order to do this, we have developed several models following three different models architectures: 1) Mutual inhibition, 2) Feed-forward inhibition, and 3) Pooled inhibition. The global inhibition model (3.24), which has the pooled inhibition blueprint, is the most successful model in reproducing experimental data from both single sugar and binary-sugar mixture experiments, according to our preliminary analysis. Additionally, the pooled inhibition architecture represents the yeast's carbon catabolism regulation better than the other two structures considered. In order to represent the binary-sugar mixtures, two induction terms are considered plausible: RFO_i (3.13), and RSO_i (3.14). Likewise, two different B equations are also considered: Hill-like cellular growth (3.20) and logistic growth (3.21). This yields four model iterations that we fit to the experimental data in Chapters 4, and analyse as dynamical systems in Chapter 5.

Chapter 4

Fitting experimental data

In this chapter I present the experimental data used in this project which was provided by Lucia Bandeira and Peter Swain from the Swain Lab from the Department of Synthetic Biology at the University of Edinburgh. I explain how it was analysed and what data sets were used from it. Subsequently, I detail the process through which the models introduced in the previous chapter were parameterised, as well as the specific error metrics used to evaluate each fit. Finally, I present the relevant results.

4.1 Experimental realisation

Experimentally, a plethora of technologies allows researchers to study closely both single cell and population behaviour. Studying cell cultures in flasks and chemostat growing conditions is a common experimental approach. Chemostats are bioreactors that allow continuous cultures to remain in steady-state, by maintaining the internal conditions constant through a continuous supply and overflow of content (Vazquez, 2018). They have proven to be very useful in studying metabolism and its regulation due to the level of control over individual parameters such as temperature, pH, cell dilution rate, growth rate, as well as the concentrations of individual

nutrients (Daran-Lapujade et al., 2004a). Experiments on single-cells have been of crucial significance in revealing cellular behaviour, as was shown in the results from the seminal work of Novick and Weiner (1957) on enzyme induction phenomena in the late 50's. Single-cell analysis was originally done by multiple and subsequent dilutions, however it has been refined throughout the decades thanks to the emergence of new technologies including flow cytometry and microfluidic devices, both of which allow a high level of versatility as well as provide high detail of single-cell measurements in a short time (Ormerod, 2008; Reece et al., 2016). Microfluidics is the science and technologies that allow the control of small amounts of fluid (in the range of $10^{-9}L$) (Farré et al., 2012). Flow cytometry refers to a system in which cell are made to flow passed a point where, one by one, they are measured. As (Ormerod, 2008) points out, in principle, the measurement could be any of a number of different types of analyses, and flow cytometry refers to the equipment that allows for this single-cell measurement to be made. For examples of flow cytometry and microfluidics experiments see references (New et al., 2014; Wang et al., 2015; Venturelli et al., 2015) and (Boulineau et al., 2013).

Next, I present a description of the experimental data in this project. I first will discuss the experimental setup, as well as provide some detail about the fluorescent tags used in the experiments. Subsequently, I will describe the two types of experimental data sets that were utilised: single sugar and binary-sugar mixture data sets. I will briefly touch upon the most significant aspects of the *platereader* Python library developed by Swain Lab which was used in order to analyse and transform the raw experimental data.

The data used in this project were provided by Lucia Bandeira and Peter Swain from the Swain Lab in the Department of Synthetic Biology at the University of Edinburgh.

4.1.1 Methods

The experiments consisted in growing yeast cultures for approximately 25 hours with medium containing different sugars at varying concentrations. The experiments can be categorised into single sugar and binary-sugar mixture regimes; under single sugar conditions the cultures were only fed with either galactose or maltose at different concentrations, whereas under a sugar mixture regime sugars were fed simultaneously with both maltose and galactose at different concentrations (both experimental regimes included a glucose run as control, and sugar mixture experiments also included single sugar batches as control, both of which are not shown in the figures presented). Two types of data were gathered from these experiments: the fluorescence levels of both tags (mCherry and GFP) that indicate the relative levels of enzymatic concentration (which in turn we assume to be proportional to the overall activity of the metabolic pathway the tagged enzymes belong to), as well as absorbance levels, which correlate with the amount of cellular mass in the sample. Table 4.1 provides more details about the experimental layout. Fluorescence levels are reported in R.F.U. (relative fluorescence units), whereas absorbance (in the graphs indicated as OD, which stands for optical density) data are reported in A.U. (arbitrary units). In each single sugar experiment there were three replicates per condition, meaning that each sugar concentration was measured three times in parallel with three independent yeast cultures in three different wells on a 96 well plate. The sugar mixture experiments had only two replicates. All of these experiments are batch cultures, meaning that once the media and the inoculum are set, they are let to incubate and no more media nor nutrients are added, and no contents are removed (Hogg, 2005). It is important to remark that the data gathered from these experiments inform about the state of the population, not the individual cells that are part of the culture.

Table 4.1: The specific experimental data that were used in our analyses have been separated into single sugar and sugar mixtures. The four single sugar experiments are divided into two sets: two experiments with maltose and two with galactose. The binary-sugar mixture experiments are divided into two sets of experiments which differ in the concentration of galactose used. For these two sets of experiments the maltose concentrations are identical. They are: 0.25%, 0.5%, 1%, 1.5%, 2%, 3% and 4%. the galactose concentrations used are: 0.125%, 0.25%, 0.5% for one set of experiments, and: 0.75%, 1%, 1.5% and 2% for the second set of binary-sugar mixture experiments. These two sets are respectively labelled PR1 and PR2, which refers to the physical plate reader used to read the samples.

Type of experiment	**Date**	**Data used**	**Strain**	**Maltose concentration**	**Galactose concentration**
Single sugar	16 Mar 2017	mCherry, OD	680 (Mal12:mCherry)	0.25 - 4%	-
Single sugar	16 Mar 2017	mCherry, OD	710 (Mal12:mCherry, Gal10:GFP)	0.25 - 4%	-
Single sugar	10 Apr 2017	GFP, OD	708 (Gal10:GFP)	-	0.125 - 2%
Single sugar	10 Apr 2017	GFP, OD	710 (Mal12:mCherry, Gal10:GFP)	-	0.125 - 2%
Sugar mixture (PR1)	25 Mar 2017	mCherry, GFP, OD	710 (Mal12:mCherry, Gal10:GFP)	0.125 - 1%	0.125 - 0.5%
Sugar mixture (PR1)	26 Mar 2017	mCherry, GFP, OD	710 (Mal12:mCherry, Gal10:GFP)	0.125 - 1%	0.125 - 0.5%
Sugar mixture (PR2)	25 Mar 2017	mCherry, GFP, OD	710 (Mal12:mCherry, Gal10:GFP)	0.125 - 1%	0.75 - 2%
Sugar mixture (PR2)	26 Mar 2017	mCherry, GFP, OD	710 (Mal12:mCherry, Gal10:GFP)	0.125 - 1%	0.75 - 2%

In the specific experiments that we used, the yeast cultures had enzymes Mal12p and Gal10p tagged with mCherry and GFP respectively in order to measure their fluorescence levels. The fluorescence levels are assumed to be proportional to the concentration of the protein they are coupled with. Enzyme concentration is in turn assumed to be representative of the metabolic activity related to the consumption of the protein's cognate sugar.

It must be mentioned that there is a caveat to this assumption, which is that fluorophore maturation dynamics represent an extra layer of complexity. This means that the recorded fluorescence observations might differ from the actual protein dynamics. Once a fluorophore is expressed by a cell, it need to undergo (depending on the specific fluorescent molecule) some further steps before it is functional, and its fluorescence properly measurable (Milo and Phillips, 2015). This process is called maturation and each fluorophore has a different maturation time, meaning that there is a discrepancy between the time at which the fluorophore was synthesised in the cell, and the time at which its signal is detectable. Work by Macdonald et al. (2013) has demonstrated that in *E.coli* EGFP (Enhanced GFP) and mCherry have maturation times of 15 ± 3.5 minutes, and 155 ± 10 minutes respectively at room temperature (20 C). Khmelinskii et al. (2012) have reported maturation times for sfGFP (super folded GFP, a variant of the wild type GFP that has a shorter maturation time due to its fast folding nature) and mCherry in *Saccharomyces cerevisiae* of 5.63 ± 0.82 minutes for sfGFP, and both 16.91 ± 1.23 and 30.30 ± 1.88 minutes for mCherry. Milo and Phillips (2015) provide a comprehensive table comparing the different maturation times of several fluorophores measured in different experimental conditions. The pattern that can be recognised is that, in general, mCherry has a higher maturation time compared to GFP and most of its variants. Milo and Phillips (2015) judiciously advise that implementing a separation of timescales between the fluorophore dynamics, and the process of interest would be a useful modification in the pursuit of accuracy. We did not implement this separation of timescales. However, we expect that since the half-lives of both tagged enzymes are

in the order of several hours, the timescale differences, as well as their effect on our results can be considered to be trivial.

Both enzymes are essential to the metabolism of their respective sugars. Mal12p breaks maltose into two glucose molecules whereas Gal10p is the first enzyme of the Leloir pathway that catalyses galactose's transformation into glucose 6-phosphate which can be further metabolised through the glycolytic pathway (Vanoni et al., 1989; Timson, 2007). As mentioned, optical density (OD) was also measured, as this is a standard technique used to assess the biomass of a cell colony (Sonnleitner et al., 1992; Mytilinaios et al., 2012). As mentioned before, I used the Python package *platereader* developed by the Swain Lab to correct for the media fluorescence in the case of mCherry and GFP readings, as well as correcting the non-linearities of OD measurements. Additionally, the software tool also provided the first and second derivative of each data curve (Swain et al., 2016; Lichten et al., 2014).

The corrections performed on the fluorescent data take all replicates from each condition and output a single corrected time series. This is why none of the figures depicting fluorescence data from individual experiments have an error region. On the other hand, the non-linearities present in the raw absorbance data can be corrected on each individual well. Consequently, the OD values observed in individual experiments are the mean values taken from the corrected values of each individual well for each sugar concentration (three for single sugar experiments, and two for the sugar mixture experiments).

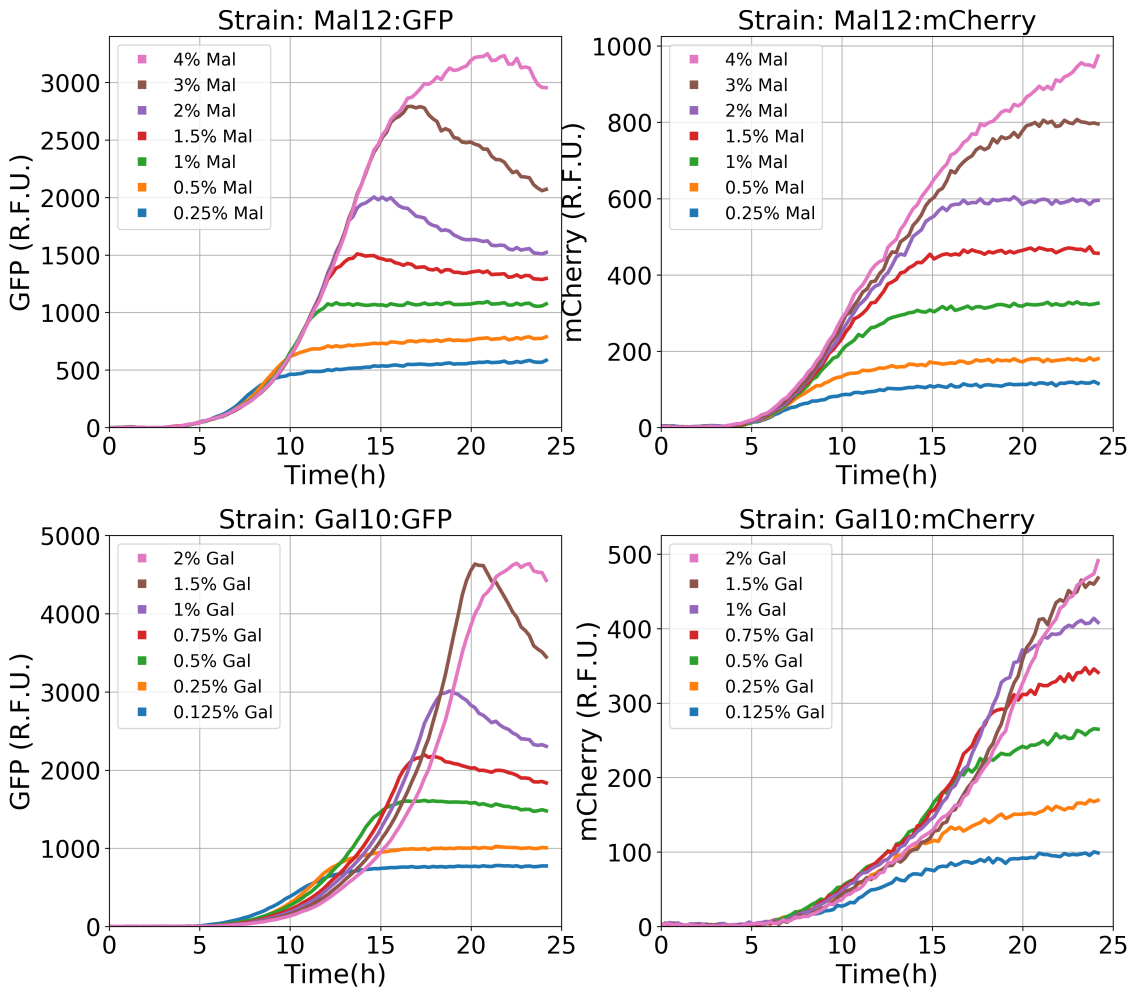


Figure 4.1: Fluorescent data of the two fluorescent molecules GFP and mCherry observed in the same experimental conditions. **Top**: Fluorescent data of yeast cultures grown in maltose. **Bottom**: Fluorescent data of yeast cultures grown in galactose.

The same experiment performed on the same day, read by the same equipment can deliver noticeably different time series depending on the molecule used to tag the enzyme of interest. The different dynamics displayed by the two fluorophores can be observed in figure 4.1. Two of observations to be made here are: 1) The values of fluorescence reported vary drastically between mCherry and GFP. In the maltose experiment the maximum fluorescence values are $\approx$ 3400 and $\approx$ 1000 R.F.U.s for GFP and mCherry respectively, whereas in the galactose experiments these values are $\approx$ 4700 and $\approx$ 500 R.F.U.s for GFP and mCherry respectively. This is also observable when comparing the maltose activity curves (Fig. 4.3) with the galactose curves (Fig. 4.4) from different experiments. It is also readily apparent that the relationship between the fluorescence values from mCherry and GFP is inconsistent

across experimental conditions. The GFP (R.F.U.) / mCherry (R.F.U.) ration in the maltose experiment is close to 3, whilst in the galactose experiments, the ratio is much closer to 10. 2) The signal from GFP appears to decay much faster than that of mCherry. The different dynamics from the same enzyme would suggest that this effect is much more related to the degradation of the tagging molecules.

4.1.2 Results

As indicated in table 4.1, single sugar experiments consist of seven different concentrations of their respective sugar, whilst in the binary-sugar mixture experiments, the yeast culture grows in nine or twelve different conditions (PR1 and PR2 respectively). These conditions are combinations of different concentrations of both maltose and galactose.

Figures 4.3 to 4.7 offer a visualisation of the experimental data. It can be noted that the curves in each graph have three recognisable stages that parallel the archetypal stages found in the microbial growth curve: 1) A lag phase that starts at $t = 0$ in which the curve remains mostly flat, 2) the exponential phase characterised by exponential increase and 3) the stationary phase in which the growth reaches a plateau (Hogg, 2005). Figure 4.2 shows these three section of the microbial growth curve. This pattern is observed in both the fluorescent data and the optical density data. It is worth noting that, in none of the optical density curves, which measures the cellular mass, do we observe the fourth stage: death. On the other hand, the fluorescent data does show signal decay. This effect is particularly noticeable in the GFP fluorescent data, regardless of the enzyme being tagged (Mal12p or Gal10p), which suggests that the decay of the signal is related to the fluorophore's dynamics rather than an indication of proteolysis (Fig. 4.1).

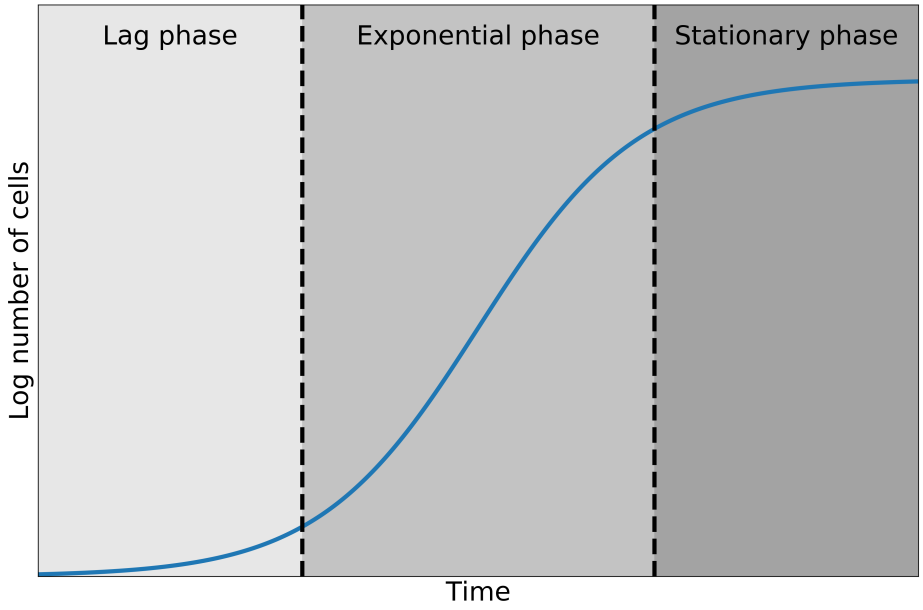


Figure 4.2: Segmentation of the microbial growth curve. This curve covers only the first three phases, it does not show the death phase.

Figures 4.3 and 4.4 show the experimental data corresponding to yeast cultures grown in maltose and galactose, respectively. Changes between different curves within the stationary and exponential growth phases are subtle, and in the case of OD in maltose grown cultures, there is nearly a complete overlap. The most noticeable change in all three data sets occurs at their respective stationary phases, which is when the sugar has been depleted. This stage is remarkably flat in both OD and mCherry (Mal12p), showing little decay. GFP (Gal10p) data, however, shows more complex dynamics compared to mCherry. It can be noted that the inflexion points of each curve change noticeably as a function of the galactose concentration. As galactose concentration increases the decay present in the stationary phase increases as well, and the second inflexion point becomes the maximum point of the curve, a property which is not present in the mCherry data set.

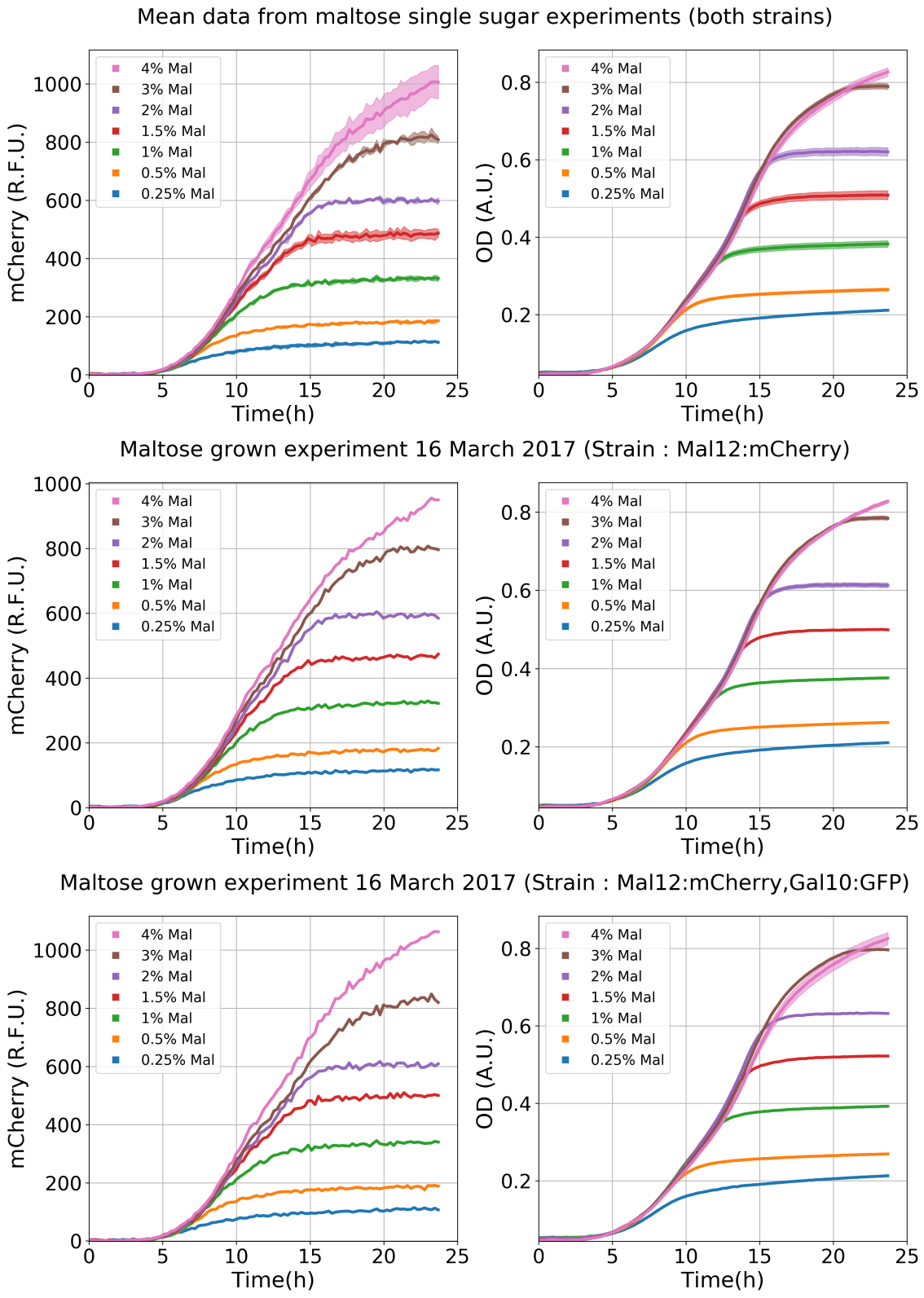


Figure 4.3: Fluorescent and absorbance data from yeast cultures grown in maltose. The graphs on the left correspond to the mCherry fluorescent measurement, and the graphs on the right show the absorbance values that indicate the amount of cellular mass. **Top**: The mean values from the experiments in two different strains, 680(Mal12:mCherry) and 710(Mal12:mCherry, Gal10:GFP), with an error region corresponding to one standard deviation. **Middle and Bottom**: The experimental data from the two different strains, respectively.

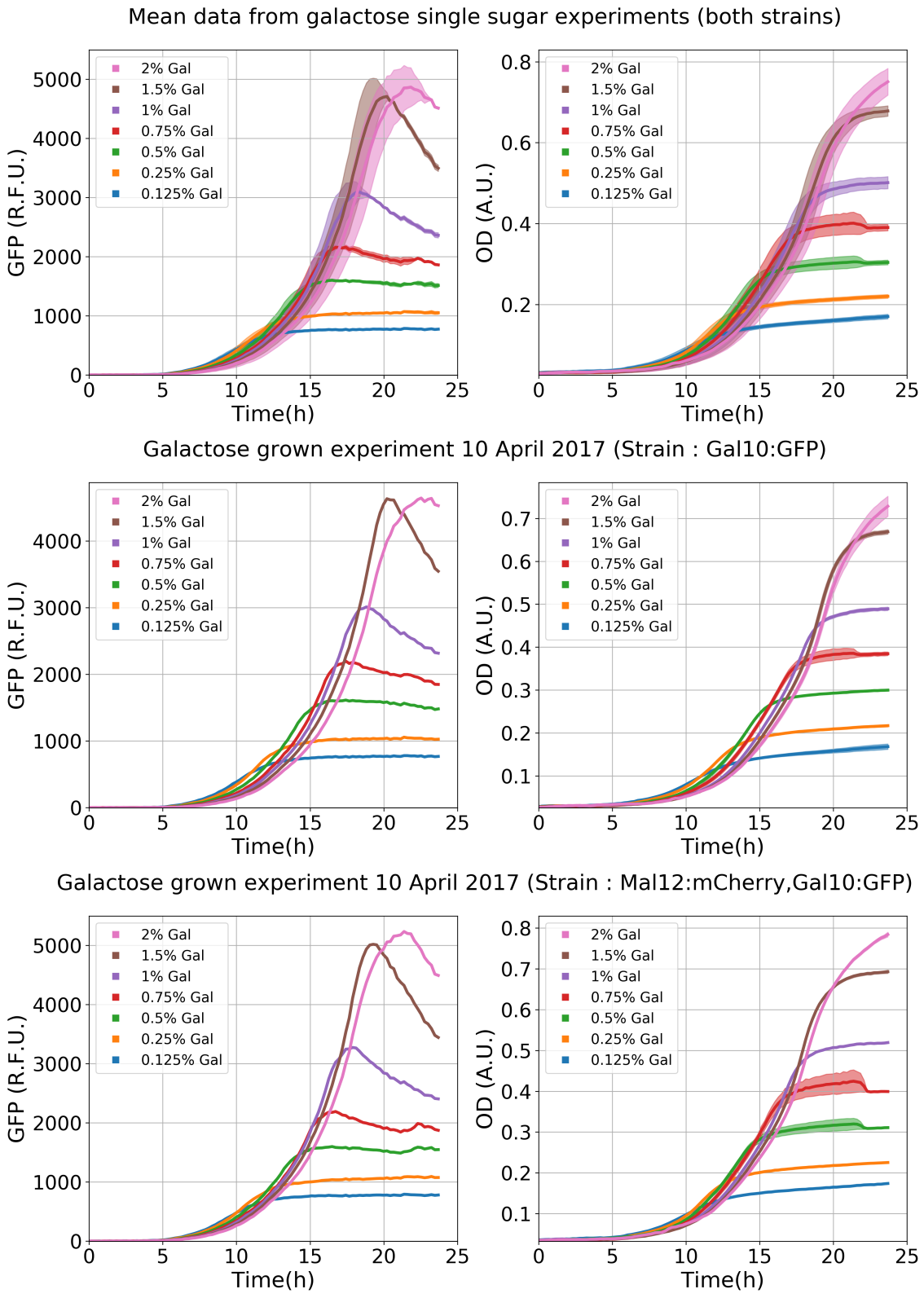


Figure 4.4: Fluorescent and absorbance data from yeast cultures grown in galactose. Similarly to figure 4.3, the graphs on the left correspond to the GFP fluorescent measurement, and the graphs on the right show the absorbance values that indicate the amount of cellular mass. **Top**: The mean values from the experiments in two different strains, 708 (Gal10:GFP) and 710(Mal12:mCherry, Gal10:GFP), with an error region corresponding to one standard deviation. **Middle and Bottom**: The experimental data from the two different strains, respectively.

Notably, the curves for a 1.5% galactose concentration (figure 4.4) present a spike (the second inflexion point) that surpasses the maximum point of the 2% galactose curve. Such behavioural properties are present in other experiments in which Gal10p was tagged with mCherry, hence, we do not presume this effect is exclusively caused by the dynamics of the fluorophore, rather we think it is more likely a consequence of the Gal10p dynamics. Figures 4.5, 4.6 and 4.7 show the experimental data for the sugar mixture experiments. The labels *PR1* and *PR2* used to differentiate between the two data sets refer to the plate readers used for the measurements of each experiment. Between them the only difference is the galactose concentration used, as indicated in table 4.1.

In these experiments, several concentrations of both maltose and galactose are available to the yeast cultures. PR1 data consists of three different concentrations of both maltose and galactose, whereas PR2 has three maltose concentrations and four galactose concentrations, generating nine and twelve different experimental conditions respectively. PR1 fluorescent data is organised in three groups with three curves per group, the order of which is contingent on the concentration of the cognate sugar. The order within the groups, however, appears to correspond to the competing sugar concentration. PR1 shows a great degree of overlap over the stationary and exponential phases, separating in the stationary phase (the grouping effect and the overlap are more evident in mCherry). PR1 and PR2 differ only in the galactose concentrations used, which elicit noticeable differences. By comparing the mCherry curves from both experimental data sets PR1 and PR2, it can be noticed that curves from the PR2 data set take longer to reach their stationary phase, and in their exponential phase the slope of every curve increases as the concentration of galactose increases, a feature which is not present in the PR1 data set. An interpretation for this is that the higher galactose concentrations in PR2 beget a slower synthesis rate for Mal12p without affecting the pattern observed in the stationary phase of PR1 (higher galactose concentration equals a higher Mal12p level).

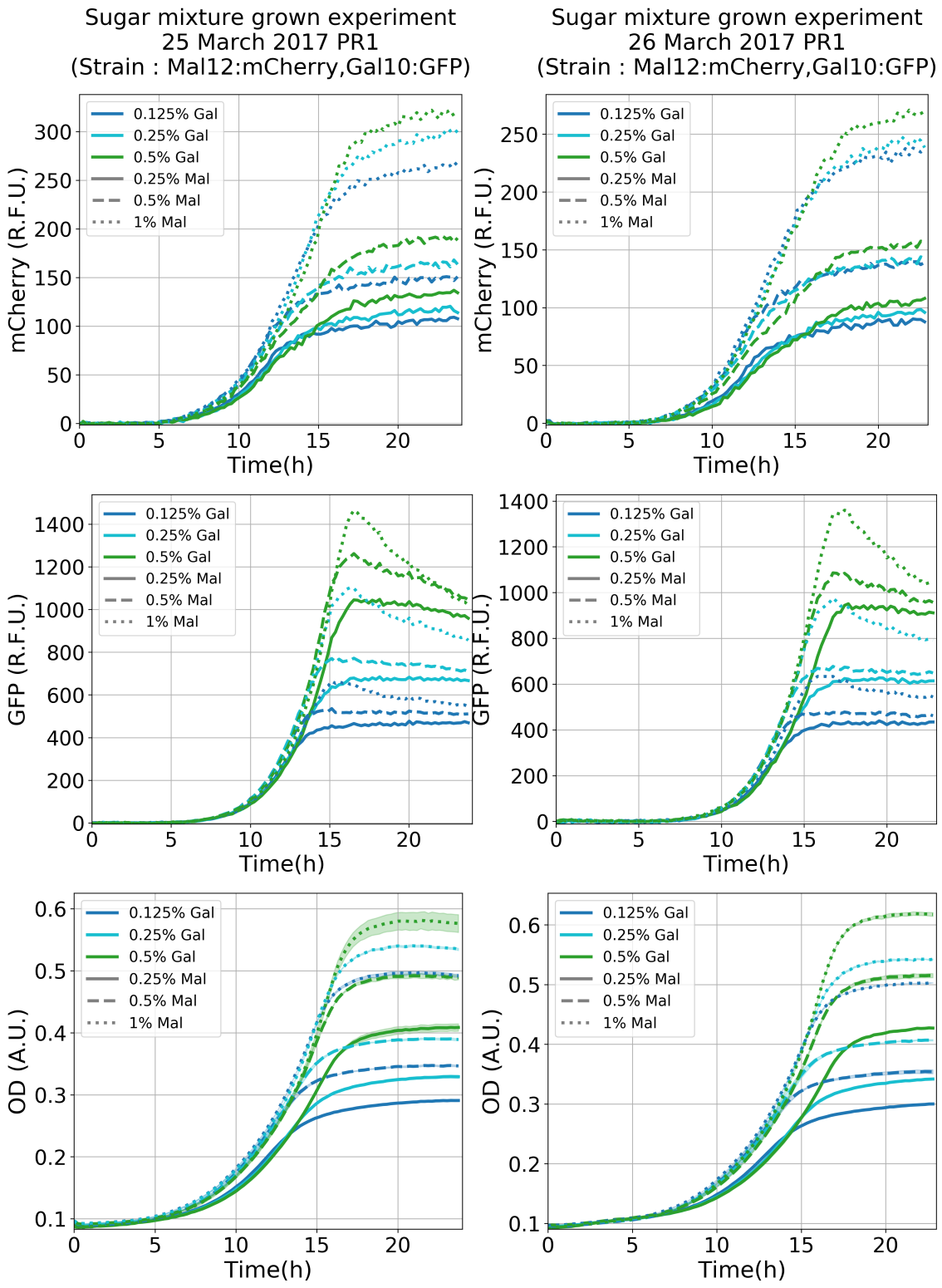


Figure 4.5: Fluorescent and absorbance data from yeast cultures grown in a binary-sugar mixture (Maltose: 0.25% - 1%, Galactose: 0.125% - 0.5%). Graphs on the left correspond to the experiment performed on March 25th, 2017, whereas those on the right are labels as the experiment performed on March 26th, 2017. The yeast strain used in this case is the double-tagged strain 710(Mal12:mCherry, Gal10:GFP). **Top**: Fluorescence levels of Mal12 enzyme tagged with mCherry. **Middle**: Fluorescence levels of Gal10 enzyme tagged with GFP. **Bottom**: Absorbance levels that indicate the amount of cellular mass.

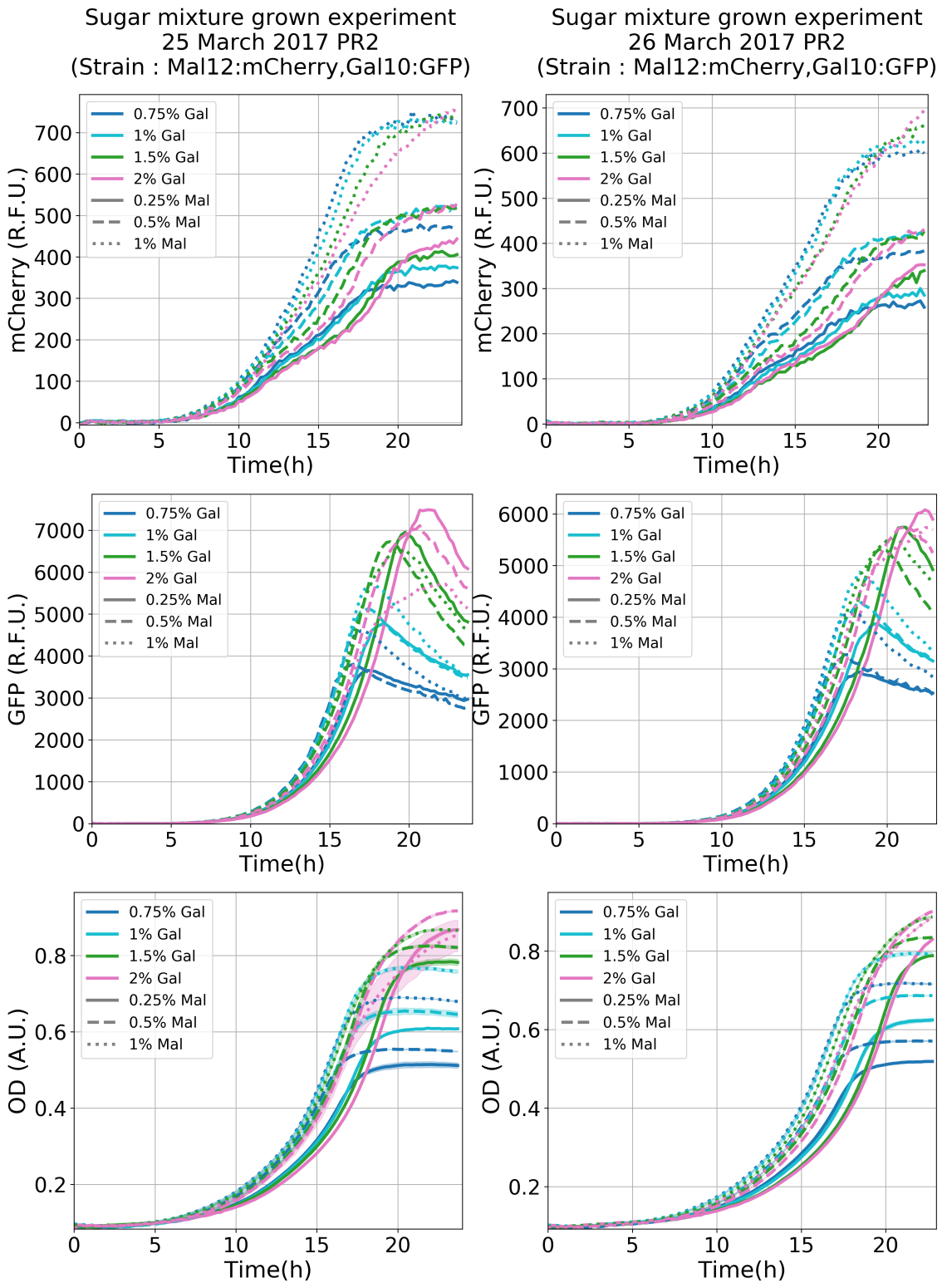


Figure 4.6: Fluorescent and absorbance data from yeast cultures grown in a sugar mixture (Maltose: 0.25% - 1%, Galactose: 0.75% - 2%). Graphs on the left correspond the experiment performed on March 25th, 2017, whereas those on the right are labels as the experiment performed on March 26th, 2017. The yeast strain used in this case is the double-tagged strain 710(Mal12:mCherry, Gal10:GFP). **Top**: Fluorescence levels of Mal12 enzyme tagged with mCherry. **Middle**: Fluorescence levels of Gal10 enzyme tagged with GFP. **Bottom**: Absorbance levels that indicate the amount of cellular mass.

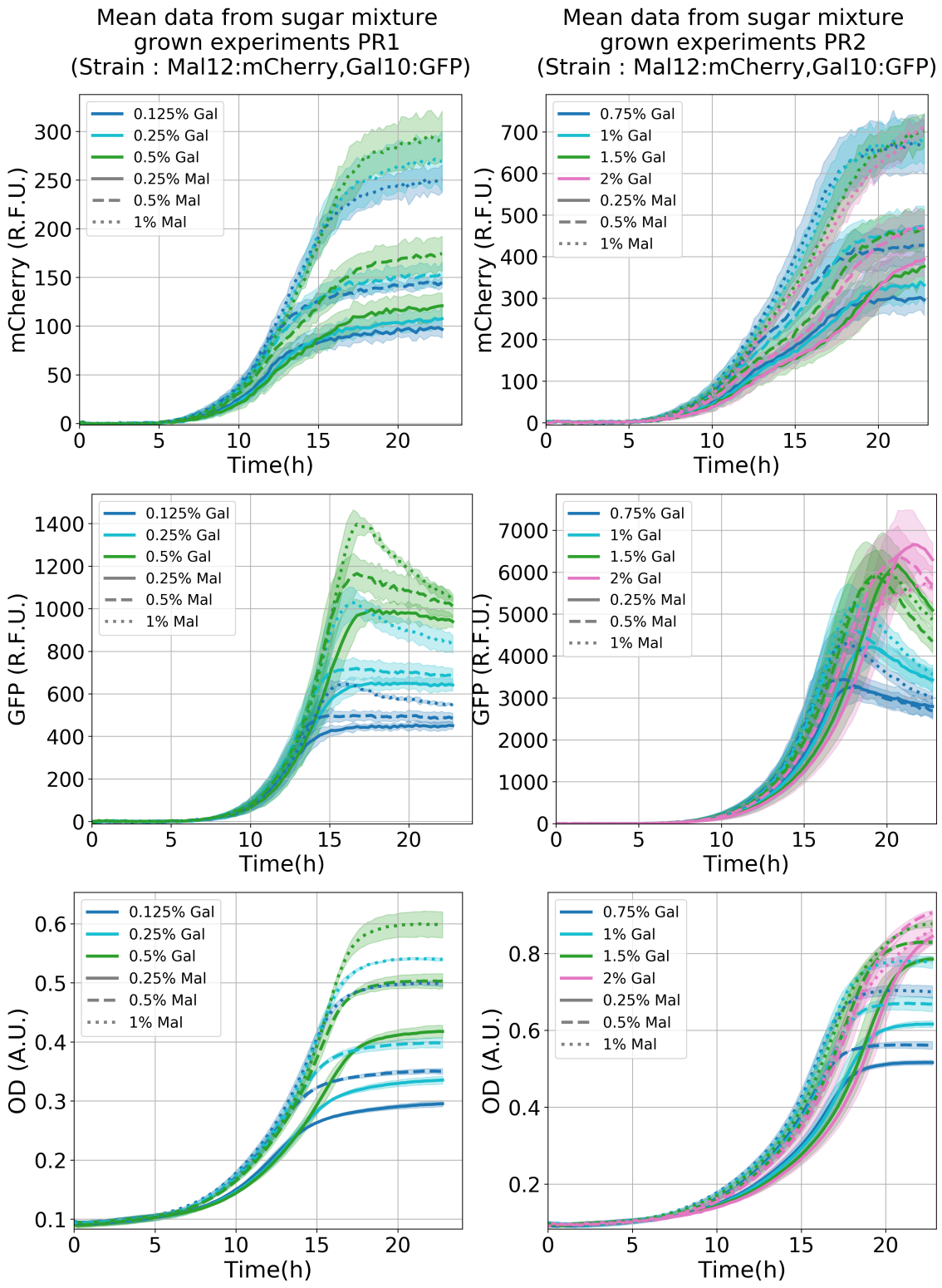


Figure 4.7: Mean values of fluorescent and absorbance data from yeast cultures grown in binary-sugar mixture experimental conditions: 1) Maltose: 0.25% - 1%, Galactose: 0.125% - 0.5%, and 2) Maltose: 0.25% - 1%, Galactose: 0.125% - 0.5%. Here the graphs on the left correspond to the data on the first experimental conditions, whereas the graphs on the right depict the data from the second experimental condition. Each curve is the mean of the two experiments shown in figures 4.5 and 4.6, with an error region corresponding to one standard deviation. **Top**: Fluorescence levels of Mal12 enzyme tagged with mCherry. **Middle**: Fluorescence levels of Gal10 enzyme tagged with GFP. **Bottom**: Absorbance levels that indicate the amount of cellular mass.

The dissimilarities found between GFP curves in both PR1 and PR2 are even more drastic. The data are still organised in groups in data set PR2 (although the 2% galactose - 1% maltose curve challenges this pattern), however, unlike both mCherry and GFP in PR1 and mCherry in PR2, in the GFP curves there is no clear organising pattern within the groups as can be observed in figure 4.6. It is possible that the dynamical pattern observed in PR1 (particularly in GFP) is lost in PR2 due to undefined metabolic dynamics revealed at high galactose concentrations.

The more complex and noisier patterns presented by the curves from experiment PR2, namely, the exponential phase dynamics and the lack of observable grouping patterns in the GFP data, have motivated our decision to work only with binary-sugar mixture experiment PR1.

We perform our experimental fit analyses with three individual experiments: For both single sugar conditions we used the experimental data corresponding to the single-tagged strains 680 (Mal12:mCherry) and 708 (Gal10:GFP), for maltose and galactose respectively, and the PR1 experiment from March 25th, 2017 on the double-tagged strain 710 (Mal12:mCherry, Gal10:GFP).

4.2 Parameterisation

In this section I will describe the parameterisation process used to fit the models to the experimental data for each experimental condition: single sugar, both maltose and galactose, as well as the mixed experiments. The objective of the parameterisation process is to converge into a set of parameters with which the model can satisfactorily reproduce the experimental time series of the fluorescence data (protein concentration) and the absorbance data (cellular mass).

Table 4.2: Parameter values obtained or otherwise derived from literature. Conversion factors were taken from Ho et al. (2018); Haddad and Lindegren (1953); White (1952); Day et al. (2004); Kim et al. (2015); Lowry et al. (1951); Reifenberger et al. (1997).

Parameter	Value range	Units	Notes
A_i	0.001-0.002	gprot/gdw	Derived from the un-induced abundance of Gal2p (transporter) multiplied by 1000 fold increase when induced (Narang et al., 1997; Boccazzi et al., 2006).
V_{Si}	101.07	gr/(gprot hr)	Glucose uptake as a function of transporter levels (Kim et al., 2015; Ho et al., 2018).
V_{Pi}	0.006	/hr	Induced abundance of Gal2p (transporter) (Boccazzi et al., 2006).
V_{Gi}	3.1-20.75	/hr	Assumed to be 0.01 V_{Si} (Narang et al., 1997).
K_{inhi}	0.18	gr/gdw	Calculated from maltose transporter activity with and without the presence of galactose (Jiang et al., 2000).
K_{Si}	0.4235-1.06	gr/gdw	Galactose uptake Km (Horák, 2013).
I_i	-	-	Speculated parameter.
K_{Pi}	0.048	gr/gdw	Near saturation intracellular sugar concentration. K_{Pi} is assumed to be in the same order of magnitude (Narang et al., 1997; Teusink et al., 1998).
Dcy_{Pi}	0.046 - .00433 0.00006*	/hr	Derived from the range of yeast protein half-lives (Horak and Wolf, 1997; Gancedo et al., 1982; Snapp, 2009). *Narang et al. (1997) assume this parameter to be 0.01 V_{Pi}.
Y_i	0.49	gdw/gr	Glucose biomass yield (Van Hoek et al., 1998).
V_B	0.42	/hr	Maximum specific growth (Van Hoek et al., 1998).
K_G	0.07	gr/gdw	Precursor concentration (Costenoble et al., 2007).
M	0.2-0.27	gprot/gdw	Grams of protein per grams of dry yeast (Siwiak and Zielenkiewicz, 2010).
Dcy_B	-	/hr	Death rate and/or dilution rate in a continuous culture scenario. Speculated value.

Parameter values in table 4.2 change their units depending on the model being used (Hill-RFO, Hill-RSO, Logistic-RFO and Logistic-RSO). The table shows the parameter units for the Hill-RFO model. With RSO models, the I_i parameter changes from unitless to gdw/gr units. With a Logistic model, the units of parameters K_G, M and Dcy_B change to $gdw^2/gr\,L$, $gprot\,L/gdw^2$ and $L/gdw\,hr$, respectively. The parameterisation process starts by taking parameter values from the literature, which serve as initial values and indicators of the order of magnitude that we should expect each parameter of the models to be in. The equations of the models we have developed, and described in Chapter 3, represent entire metabolic pathways and dynamics that take place at the population level. They are not intended to represent specific enzymatic reactions. Since most of the parameter values derived from the literature are rates of specific reactions, there are reservations regarding

how comparable they can be to the rate of the metabolic pathway as a whole. Table 4.2 lists the parameter values as well as the reference from where they were taken. The table also succinctly describes the assumptions made or the reasoning behind such parameter choice.

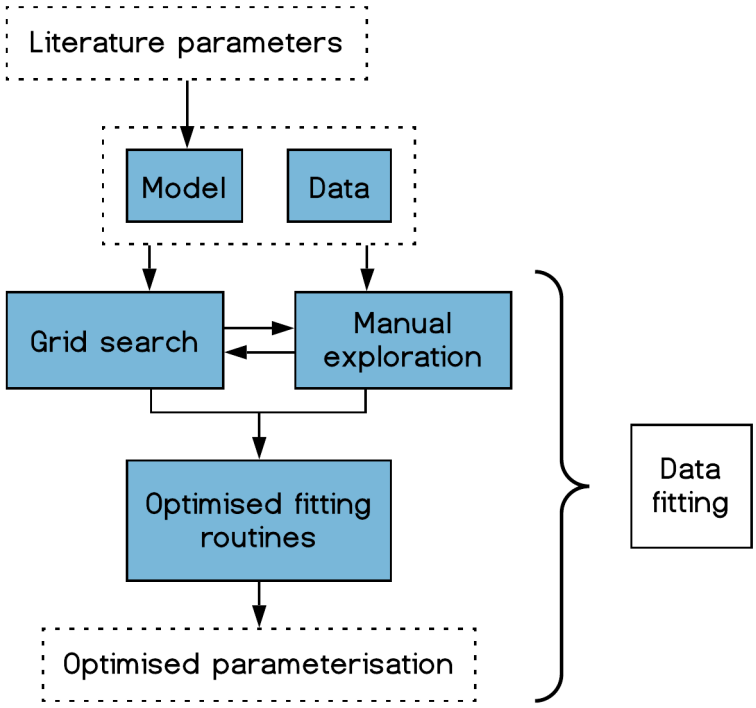


Figure 4.8: Data fitting flow diagram.

Figure 4.8 is a general diagram of the overall process followed in order to obtain an optimal parameterisation. In the following subsections I will detail the steps that I took to perform the parameterisation of the models described in Chapter 3.

4.2.1 Manual exploration and grid search

Initial attempts to optimise the parameterisation of the models using the method described in subsequent sections showed that the starting parameterisation needed to be pre-adjusted to the data in order to improve performance. Hence, once the initial parameter values were identified (table 4.2), I manually adjusted them to the data through visual evaluation (a process already described in section 3.2.3). As

I mentioned there, the goal of this parameter adjustment was to get the models' output closer to the sigmoidal curve shape that is characteristic of the experimental results, as well as to explore the models' capability to replicate the binary-mixture experiments' grouping pattern (figs. 4.3 to 4.6). Adjusting the parameters before performing the fitting routine described in the following section yielded better results, i.e. the model's output was closer to the experimental data, than doing the fitting routine without performing the adjustment.

Additionally, a grid search, also referred to as parameter sweep, was performed to further explore the parameter space. In order to perform the grid search, I selected a number of parameters of interest and created an array per parameter with manually selected values, usually differentiated by an order of magnitude when fitting to single sugar data. For some parameters, such as Y_i and V_B, an increase of an order of magnitude was considered unrealistic given their biological meaning and therefore a more moderate step size was used instead. These initial parameter values are shown in table 4.3. With the results of the grid search from the single sugar conditions, I narrowed the range of values for the grid search on binary-sugar mixture data.

Table 4.3: Selected parameters and manually specified values used during the grid search under single sugar and binary-sugar mixture conditions. Several grid searches were performed on each model for each condition. All of the parameters listed here were searched for, albeit not all in the same grid search.

Single Sugar		
Parameters		Values tested
A_i	$\in$	$\{1e^{-6}, 1e^{-5}, 1e^{-4}, 1e^{-3}, 1e^{-2}, 1e^{-1}\}$
V_{Si}	$\in$	$\{1e^{-1}, 1e^{0}, 1e^{1}, 1e^{2}, 1e^{3}, 1e^{4}\}$
V_{Pi}	$\in$	$\{1e^{-6}, 1e^{-5}, 1e^{-4}, 1e^{-3}, 1e^{-2}, 1e^{-1}\}$
K_{inhi}	$\in$	$\{1e^{-4}, 1e^{-3}, 1e^{-2}, 1e^{-1}, 1e^{0}, 1e^{1}\}$
K_{Si}	$\in$	$\{1e^{-4}, 1e^{-3}, 1e^{-2}, 1e^{-1}, 1e^{0}, 1e^{1}\}$
K_{Pi}	$\in$	$\{1e^{-4}, 1e^{-3}, 1e^{-2}, 1e^{-1}, 1e^{0}, 1e^{1}\}$
Dcy_B	$\in$	$\{0, 1e^{-3}, 1e^{-2}, 1e^{-1}, 1e^{0}, 1e^{1}\}$
V_B	$\in$	$\{1e^{-3}, 1e^{-2}, 1e^{-1}, 2e^{-1}, 1e^{0}\}$
Dcy_{Pi}	$\in$	$\{0, 1e^{-4}, 1e^{-3}, 6e^{-2}, 1e^{-1}\}$
Y_i	$\in$	$\{1e^{-3}, 1e^{-2}, 1e^{-1}, 4.9e^{-1}, 1e^{0}\}$
M	$\in$	$\{1e^{-3}, 1e^{-2}, 1e^{-1}, 3e^{-1}, 1e^{0}\}$

Sugar mixture		
Parameters		Values tested
V_{Si}	$\in$	$\{600, 800, 1000\}$
V_{Pi}	$\in$	$\{0.003, 0.005, 0.01\}$
V_{Gi}	$\in$	$\{7.5, 10, 12.5\}$
K_{inhi}	$\in$	$\{0.5, 5\}$
K_{Pi}	$\in$	$\{0.05, 0.15, 0.3\}$
I_i	$\in$	$\{0.01, 0.5, 0.15\}$
Dcy_{Pi}	$\in$	$\{0.0006, 0.06\}$
Y_i	$\in$	$\{0.1, 0.5\}$

As performance metric we used root mean square error (RMSE) in order to measure and evaluate the accuracy of the fit of each possible permutation within the

parameter space in relation to the experimental data.

Once the RMSE score had been computed for each permutation, the top scoring parameter set was stored and used as the input for the optimised fitting routine.

The same grid search procedure was used in order to find initial conditions of all state variables except the extracellular sugar concentration N_i, since those are explicitly given by the experimental data. Table 4.4 shows the values used for these grid searches. Once again, the RMSE score was computed in order to determine the best initial conditions.

Table 4.4: Initial conditions tested during grid search.

Single sugar		
State variable		Values tested
B	$\in$	$\{1e^{-3}, 3.25e^{-3}, 5.5e^{-3}, 7.75e^{-3}, 1e^{-2}, 3.25e^{-2}, 5.5e^{-2}, 7.75e^{-2}, 1e^{-1}, 3e^{-1}, 3.25e^{-1}, 5.5e^{-1}, 6.5e^{-1}, 7.75e^{-1}, 1\}$
G	$\in$	$\{1e^{-3}, 3.25e^{-3}, 5.5e^{-3}, 7.75e^{-3}, 1e^{-2}, 3.25e^{-2}, 5.5e^{-2}, 7.75e^{-2}, 1e^{-1}, 3e^{-1}, 3.25e^{-1}, 5.5e^{-1}, 6.5e^{-1}, 7.75e^{-1}, 1\}$
S_i	$\in$	$\{1e^{-6}, 3.25e^{-6}, 5.5e^{-6}, 7.75e^{-6}, 1e^{-5}, 3.25e^{-5}, 5.5e^{-5}, 7.75e^{-5}, 1e^{-4}, 3.25e^{-4}, 5.5e^{-4}, 7.75e^{-4}, 1e^{-3}, 5.5e^{-3}, 1e^{-2}\}$
P_i	$\in$	$\{1e^{-6}, 3.25e^{-6}, 5.5e^{-6}, 7.75e^{-6}, 1e^{-5}, 3.25e^{-5}, 5.5e^{-5}, 7.75e^{-5}, 1e^{-4}, 3.25e^{-4}, 5.5e^{-4}, 7.75e^{-4}, 1e^{-3}, 5.5e^{-3}, 1e^{-2}\}$

Sugar mixture		
State variable		Values tested
B	$\in$	$\{8e^{-2}, 1.5e^{-1}, 2.2e^{-1}, 2.9e^{-1}, 3.6e^{-1}, 4.3e^{-1}, 5e^{-1}\}$
G	$\in$	$\{8e^{-2}, 1.5e^{-1}, 2.2e^{-1}, 2.9e^{-1}, 3.6e^{-1}, 4.3e^{-1}, 5e^{-1}\}$
$\{S_1, S_2\}$	$\in$	$\{1e^{-7}, 5.5e^{-7}, 1e^{-6}, 5.5e^{-6}, 1e^{-5}, 5.5e^{-5}, 1e^{-4}, 5.5e^{-4}\}$
$\{P_1, P_2\}$	$\in$	$\{1e^{-7}, 5.5e^{-7}, 1e^{-6}, 5.5e^{-6}, 1e^{-5}, 5.5e^{-5}, 1e^{-4}, 5.5e^{-4}\}$

Additionally, we also explored using Latin-hypercubes sampling to obtain initial conditions and parameters. We took $100,000$ samples from uniform distributions within manually established ranges for each state variable and parameter. These $100,000$ samples were insufficient to provide comparable results in terms of RMSE scores, to the initial conditions and parameters yielded by the grid search technique, which were acceptable for our purposes. It is possible that the ranges determined by the uniform distribution of each model parameter were too broad. Hence, hindering the sampling process. It is also a possibility that more samples were needed to yield comparable RMSE scores to the grid search technique. However, larger sample sizes

exceeded the grid search's computational time. Therefore, we decided to continue using the grid search technique.

4.2.2 Fitting routine

Once a satisfactory set of initial parameter values was obtained, the model was parameterised through a fitting routine. To do this we employ the minimising algorithm L-BFGS-B available through the *scipy.optimize.minimize* function from the scipy library for Python 3.6. L-BFGS-B is a variant of the popular BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm with limited memory and the capacity to work within manually established bounds for each parameter. This last property is particularly useful since it prevents the algorithm from optimising values below zero, which are not biologically relevant. It is important to note that out of the eight state variables we used P_1, P_2 and B to compare them with the fluorescent (Mal12/mCherry and Gal10/GFP) and optical density data, respectively. The data sets are time series of approximately 25 hours and 100 points (103 in single sugar experiments and 101 in the binary-sugar mixture experiments).

The input for this algorithm are a function and the parameters. This input function, in our case must be a performance metric or objective function that outputs the goodness of fit of the model's output to the experimental data. Each iteration the parameter set is modified in the direction that reduces the output of the objective function, hence reducing the distance between the model's output and the experimental data.

The performance metric used int this process it the weighted RMSE. I have modified the RMSE formula to include a weighting factor w_i. The weighted RMSE formula reads

$$\text{Weighted RMSE} = \sqrt{\frac{1}{\sum_{i=1}^{m} w_i} \sum_{i=1}^{m} w_i \, (p_i - d_i)^2}, \tag{4.1}$$

where the vector weight w multiplies the errors by a different factor. This equation uses the definition of the weighted arithmetic mean or weighted average formula

$$\text{Weighted mean} = \frac{\sum_{i=1}^{m} w_i \, x_i}{\sum_{i=1}^{m} w_i}. \tag{4.2}$$

The reason we have decided to used a weighted version of the RMSE is to be able to separately weight the phases of the curve: lag, exponential and stationary. Once the squared errors are computed, they are multiplied by the weight w_i, the value of which depends on the phase of the curve the error belongs to. If the error belongs to the region of the curve that has been deemed to be more relevant, the contribution of that error to the overall score is increased. In order to do this, the weighting vector is constructed by multiplying the weights corresponding to the selected curve phase by a weighting factor. For instance, a weighting factor of 2 means that the errors within the selected region of the curve are worth twice as much as the errors of other regions. This would allow us to fit the entirety of the curve with differential values assigned to data points depending on the curve phase, and even fits on curve phases separately (e.g. an exponential only fit).

In order to correctly separate the three different phases I identified the inflexion points of each curve, which are the maximum and minimum values of the second derivative of each curve. In order to do this, I have use the software *platereader* developed by the Swain Lab to calculate the second derivative of all data-sets (mCherry, GFP and OD) and subsequently identify their maxima and minima. Manual refinements had to be made to correctly separate the curve phases.

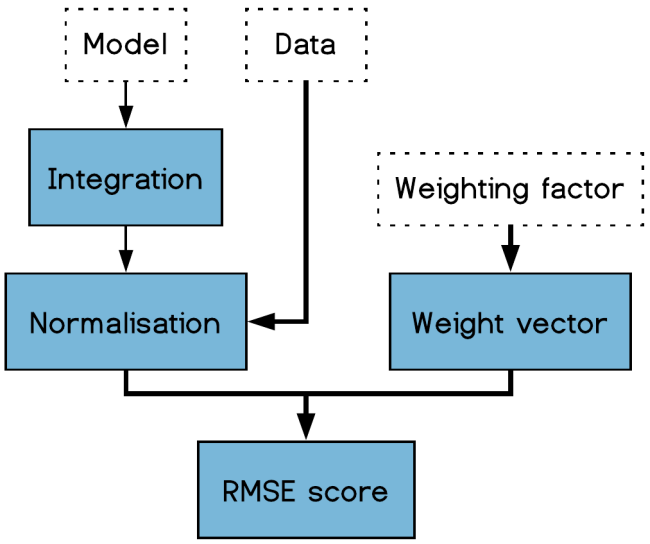


Figure 4.9: Flow diagram of the steps followed by the optimising routine at each iteration.

Figure 4.9 shows the steps followed by the input function at each iteration of the optimisation routine. The first step of the objective function is the integration of the model over the initial sugar concentrations of the experiment. In the RMSE context, this effectively gives us the predictions p_i for each initial extracellular sugar concentration. Subsequently, both the data and the model's output are normalised. We used a min-max normalisation, the formula of which reads

$$\text{normalised } x_i = \frac{x_i - \min}{\max - \min}. \tag{4.3}$$

Once both values have been normalised, the squared error is calculated, and subsequently multiplied by the weight vector. The weighed squared errors are then summed up and divided by the total amount of data points within the experimental data set. The total weighted RMSE score for any given experimental data set is given by the formula

$$\text{Total weighted RMSE} = \sqrt{\sum_{k=1}^{c} \frac{\sum_{j=1}^{n} \frac{\sum_{i=1}^{m} w_{kji} \left(p_{kji} - d_{kji}\right)^2}{n}}{c}}, \tag{4.4}$$

where data point d_{kji} corresponds to the point i, from the sugar concentration j and data type k; m is the number of points within a single time series, n is the number of all initial sugar concentrations that give rise to the different curves per experiment (seven for single sugar experiments and nine for sugar mixture experiments), and c is the number of data types in the experiment: two for single sugar experiments (mCherry or GFP, and OD), and three for sugar mixture experiments (mCherry, GFP, and OD). For instance, taking the maltose single sugar data, data point $d_{1,3,5}$ is the fifth point from the 1% maltose curve (the 3rd concentration in ascending order) from the fluorescent data (the first data type, the second being optical density). Together, data point d_{kji}, curve prediction p_{kji} and weight w_{kji} are used to compute the corresponding weighted error.

The optimisation method iterates through this process until the difference in the RMSE scores between steps falls below tolerance value $2.220446\,e^{-9}$ (default value), according to the formula

$$\frac{f_k - f_{k+1}}{max(|f_k|, |f_{k+1}|, 1)} \leq 2.220446\,e^{-9}, \tag{4.5}$$

where f_k is the RMSE score at iteration k. I followed this process to parameterise all of the models (described in detail in the previous chapter) for all experiments. In the following section I present the results of our parameterisations for each model.

4.3 Fitting results

In this section I present the fitting routine results obtained. Each of the four models was fitted to the experimental data introduced in a previous section. In total, each model was fitted to four different experiments: 1) Maltose single sugar data, 2) Galactose single sugar data, 3) Sugar mixture data (experimental data set labelled PR1), and 4) the three experimental data sets at once. All experimental data sets

were fitted with several weighting factors. For the fitting results presented here, the weighting factor affects the points belonging to the stationary phase of the curves. When the weighting factor is < 1, the fitting errors in the stationary phase are less significant than those in the other two regions of the curves. On the other hand, when the weighting factor is > 1, the errors in the stationary phase contribute to a higher extent to the overall RMSE score, which in principle signifies that they drive the fitting routine. Finally, a weighting factor $= 1$, is equal to a non-weighted RMSE computation, meaning that no curve phases is more important that any other.

Although different weighting factors were used in the optimisation routine, here we present and rank the resulting parameterisations based on their RMSE scores with a weighting factor $= 1$. Each phase and data type is scored separately for each model and experiment. The graphs presented correspond to the parameterisation with the smallest overall RMSE score for each experimental data set. Additionally, I take the best scoring PR1 parameterisation to reproduce all three experimental data sets by changing the initial values of the state variables through a grid search.

The initial values of the state variables (with the exception of the extracellular sugar concentrations N_i which are explicitly given by the experimental setup) were estimated manually, informed by value ranges used by Narang et al. (1997), whose models were discussed in the previous chapter. Additionally, grid search was used to inform the selection of initial values. Similarly, the initial parameter values used for each model were obtained by manually adjusting parameters found in the literature along with grid search. Table 4.2 shows the values found or otherwise derived from the literature for each parameter in our models.

It is worth mentioning that some grid searches were performed with models slightly different than the final four presented in the previous chapter. The grid searches were not performed after every modification to the models due to: 1) time constraints, and 2) through preliminary analyses, some modification were deemed too small to merit a new grid search being performed.

Unless stated otherwise, the only bounds implemented during the fitting are lower bounds set to 0 for every parameter, in order to prevent negative values, which are biologically meaningless, and for models with the RFO_i induction terms, parameters I_i are also upper bounded to avoid values above 1. The reasoning behind this decision is that in this case, I_i functions as a proportional factor that regulates the contribution of the sugar S_i to the activation of the metabolism cognate to the competing alternative. As such, we assume that sugar S_i cannot have more influence over the metabolic activation, than the competing sugar alternative the metabolism of which is being activated. For instance, an I_i value = 1 would mean that sugar S_i is equally effective at activating the competing sugar's metabolism as the competing sugar itself. On the other hand, if $I_i = 0$ then sugar S_i has no influence over the induction or activation of the competing sugar alternative's metabolism. This is not a problem for the RSO_i induction terms, since the contribution of sugar S_i to the activation of the competing sugar's metabolism is dependent on the concentration of the competing sugar.

Table 4.5: Overall RMSE scores from the four different models per experimental data set and fitting routine. Cells shaded in blue represent the best RMSE score for that experimental data set out of all models. Cells shaded in green are the best RMSE score for that experimental data set within that model. GS stands for grid search. The results under this column correspond to the PR1 parameterisation and grid search for new initial conditions.

Hill function biomass model RSO_i induction term	**All data**	**Single sugar**	**PR1**	**GS**
PR1	0.097897		0.057441	**0.056279**
Maltose	0.084844	**0.047408**		**0.068498**
Galactose	0.124461	**0.057331**		**0.064633**
All data	0.103719			**0.0633421**

Hill function biomass model RFO_i induction term	**All data**	**Single sugar**	**PR1**	**GS**
PR1	0.059324		**0.040205**	**0.04035**
Maltose	**0.05277**	**0.051341**		0.056897
Galactose	0.106383	0.069543		**0.055133**
All data	0.07664			**0.051332**

Logistic equation biomass and quadratic decay model RSO_i induction term	**All data**	**Single sugar**	**PR1**	**GS**
PR1	0.068124		**0.052153**	**0.052664**
Maltose	**0.078905**	**0.051796**		0.091442
Galactose	**0.065549**	**0.04798**		0.072602
All data	**0.071095**			0.073951

Logistic equation biomass and quadratic decay model RFO_i induction term with no upper bounds	**All data**	**Single sugar**	**PR1**	**GS**
PR1	0.081046		0.070712	**0.044997**
Maltose	0.07388	**0.045677**		**0.062399**
Galactose	0.108054	0.050985		**0.048301**
All data	0.088886			**0.052445**

Logistic equation biomass and quadratic decay model RFO_i induction term with upper bounds	**All data**	**Single sugar**	**PR1**	**GS**
PR1	**0.088543**		**0.070663**	-
Maltose	**0.066459**	**0.047365**		-
Galactose	**0.109978**	**0.040624**		-
All data	**0.090096**			-

I present only the best fit for each experimental data set, according to table 4.5. This is: 1) the Hill-RFO model fitted to experimental data set PR1, 2) the three experimental data sets at once, 3) the logistic-RFO model fitted to both maltose,

and 4) galactose single sugar experimental data sets. All other fitting results can be found in appendix A. Firstly, in each subsection I provide the initial values of the state variables used in the fitting routine. Subsequently, in each case, I present the RMSE score by curve phase and data type (fluorescence and absorbance), as well as by experiment. Finally, I present the graphs which depict the experimental data along with the parameterised model's output.

4.3.1 Fits to the PR1 data set

The best PR1 fit, measured by overall RMSE score, comes from the Hill-RFO model. In this case, the fitting algorithm has no upper bounds for any parameter, except the I_i parameters which had an upper bound of 1. Additionally, the lower bounds were set to 0, to avoid negative-valued parameters. Additionally, table 4.6 shows the initial values used during the fitting routine.

Table 4.6: Initial values of the state variables used during the PR1 fitting routine with the Hill-RFO model.

	B	**G**	$\boldsymbol{S_1}$	$\boldsymbol{S_2}$	$\boldsymbol{P_1}$	$\boldsymbol{P_2}$
PR1	0.1	0.3	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$

Table 4.7 shows the RMSE scores divided by section of the curve and by weighting factor used for the fit. Table 4.8 shows the RMSE scores of the lowest scoring fit (overall score) by section of the curve and data type.

Table 4.7: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 2 for all three experimental data sets.

Weight factor	2.0	0.5	10.0	5.0	1.0	100.0	0.1	I.P.	0.01
Overall RMSE	0.0402047	0.0410185	0.0434217	0.0437596	0.0450072	0.0478087	0.0602773	0.0904905	0.120234
Lag RMSE	0.014203	0.0201152	0.0123978	0.0117584	0.016948	0.0111846	0.0392858	0.0244699	0.0735935
Exponential RMSE	0.045735	0.0498759	0.0522993	0.0538437	0.0582028	0.0458202	0.0852931	0.122413	0.190627
Stationary RMSE	0.0505553	0.0464255	0.0526061	0.0520318	0.0490118	0.068062	0.0456265	0.0947722	0.040173

Table 4.8: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1). This parameterisation corresponds to the lowest overall RMSE score (w.f. 2).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.021991	0.052576	0.048537	**0.043219**	
GFP	0.008275	0.046597	0.071151	**0.049336**	
OD	0.007288	0.036600	0.015785	**0.023394**	
RMSE per phase	**0.014203**	**0.045735**	**0.050555**	**0.040205**	**Total RMSE score for the PR1 fit**

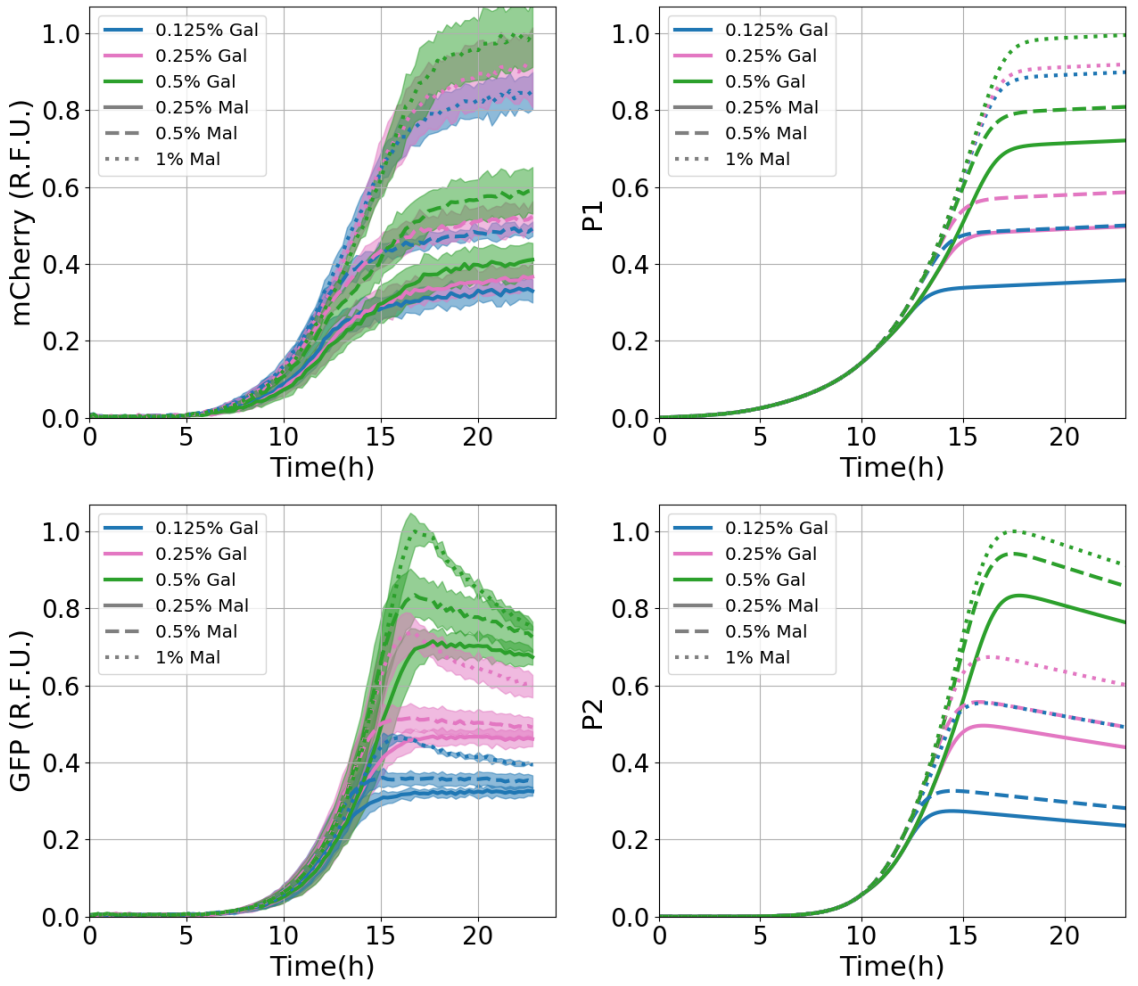


Figure 4.10: Results from the fit to sugar mixture experimental data set (PR1). This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of Mal12p. **Middle**: These graphs correspond to the fit to GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

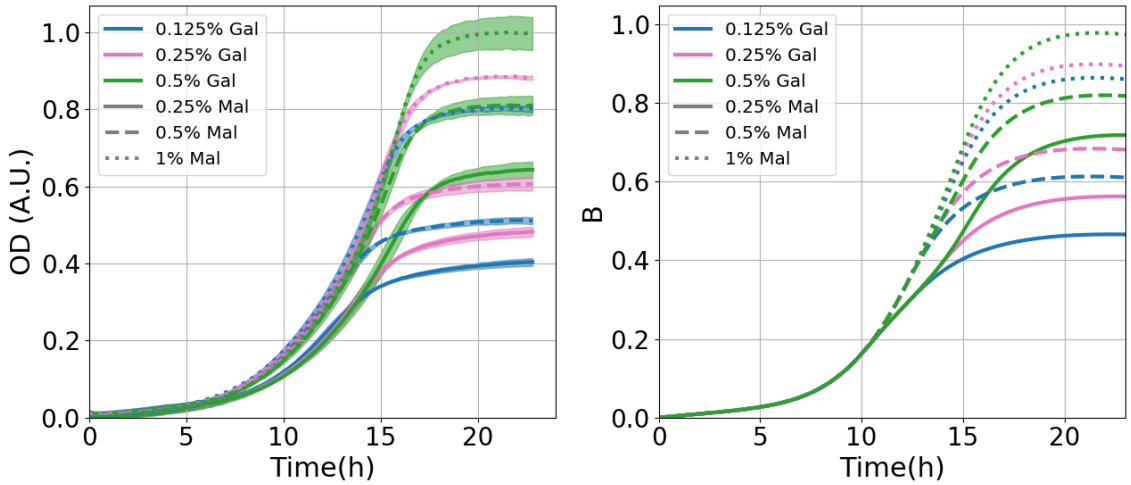


Figure 4.10: (cont.)

4.3.2 Fits to all three data sets

The best scoring parameterisation of all data sets corresponds to the best PR1 parameterisation of the Hill-RFO model. As mentioned earlier, this parameterisation is used to reproduce all three experimental data sets by changing the initial values of the state variables through a grid search. Table 4.9 shows the resulting initial condition.

Table 4.9: Initial values of state variables resulting from the grid search performed with the lowest overall RMSE PR1 parameterisation.

	B	**G**	S_1	S_2	P_1	P_2
PR1	0.08	0.22	$1e^{-3}$	$1e^{-3}$	$1e^{-3}$	$1e^{-5}$
Maltose	0.775	0.325	$1e^{-6}$	0	$1e^{-6}$	0
Galactose	0.775	0.325	0	$5.5e^{-3}$	0	$1e^{-4}$

Table 4.10 shows the RMSE scores divided by experiment and by section of the curve. Tables 4.11, 4.12, and 4.13 show the RMSE score divided by data type and section of the curve of PR1, maltose single sugar, and galactose single sugar experiments respectively.

Table 4.10: RMSE scores divided by curve phase and experimental data set (single sugar and sugar mixture).

	Lag phase	Exponential phase	Stationary phase	RMSE per experiment	
PR1	0.014897	0.046096	0.050374	**0.040350**	
Maltose	0.027098	0.081423	0.048456	**0.056897**	
Galactose	0.006112	0.071271	0.063260	**0.055133**	
RMSE per phase	**0.018199**	**0.067907**	**0.054428**	**0.051332**	**Total RMSE score**

Table 4.11: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.023813	0.052006	0.048079	**0.043140**	
GFP	0.007729	0.047803	0.071013	**0.049624**	
OD	0.006243	0.037212	0.016071	**0.023678**	
RMSE per phase	**0.014897**	**0.046096**	**0.050374**	**0.040350**	**Total RMSE score for the PR1 fit**

Table 4.12: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.038085	0.104629	0.047422	**0.069872**	
OD	0.004264	0.048085	0.049468	**0.039906**	
RMSE per phase	**0.027098**	**0.081423**	**0.048456**	**0.056897**	**Total RMSE score for the maltose fit**

Table 4.13: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
GFP	0.004940	0.069190	0.085131	**0.063401**	
OD	0.007092	0.073293	0.027503	**0.045382**	
RMSE per phase	**0.006112**	**0.071271**	**0.063260**	**0.055133**	**Total RMSE score for the galactose fit**

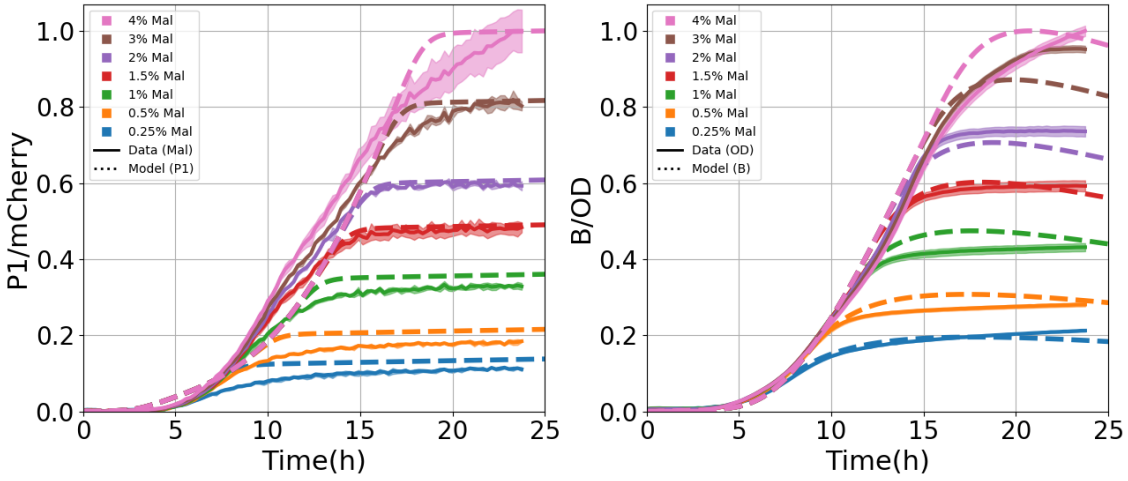


Figure 4.11: Fit to the single sugar experimental data sets with the lowest scoring PR1 parameterisation (overall RMSE) after performing a grid search for new initial state variables values (table 4.9). **Top**: The maltose data set. **Bottom**: galactose data set. Graphs to the left correspond to the fit to the fluorescence data (mCherry in the case of maltose, GFP in the case of galactose), whilst the graphs to the right present the results of the fit to the absorbance data (O.D.).

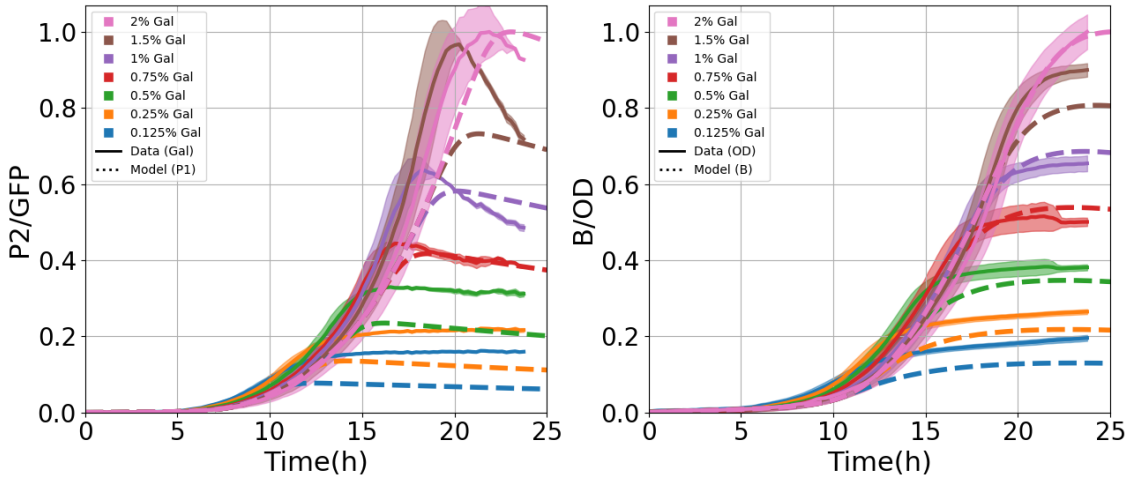


Figure 4.11: (cont.)

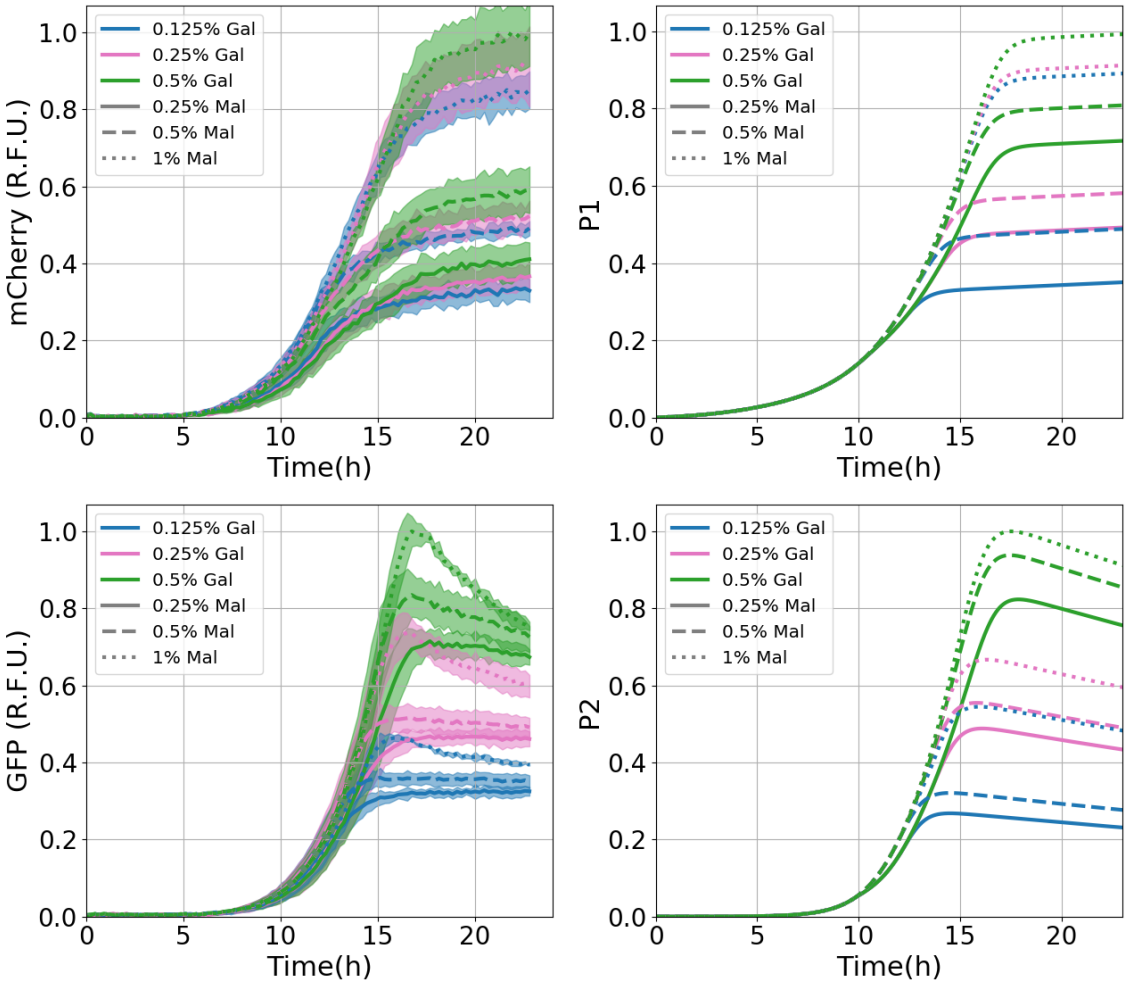


Figure 4.12: Fit to the binary-sugar mixture experimental data sets with the lowest scoring PR1 parameterisation (overall RMSE) after performing a grid search for new initial state variables values (table 4.9). **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

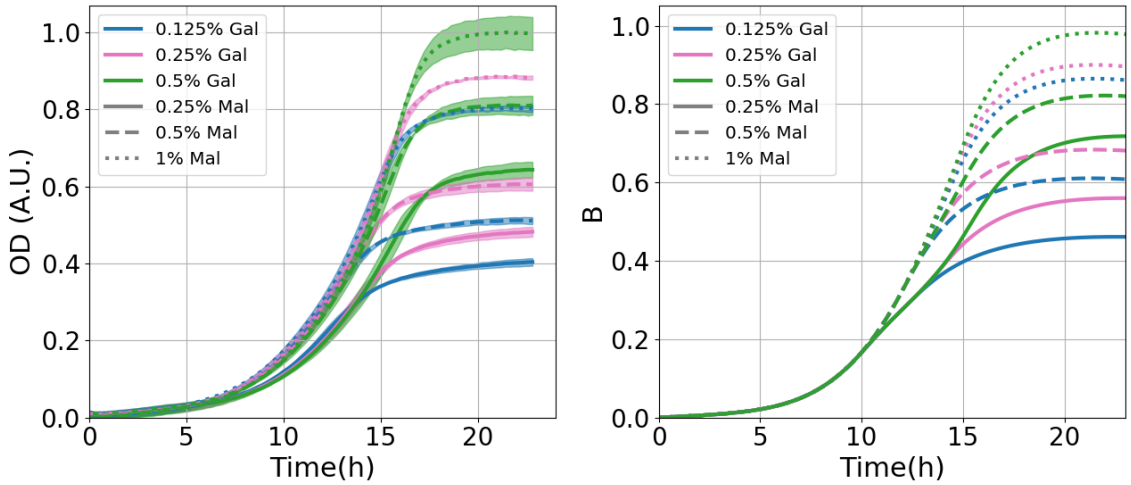


Figure 4.12: (cont.)

4.3.3 Fits to single sugar data sets

The best fits to both maltose and galactose data sets, measured by overall RMSE score, correspond to the logistic-RFO model. The initial conditions used for these fitting routines are specified in table 4.14.

Table 4.14: Initial values of the state variables used during the fitting routine to maltose and galactose, respectively.

	B	**G**	S_1	S_2	P_1	P_2
Maltose	0.1	0.1	$1e^{-4}$	0	$3e^{-3}$	0
Galactose	0.05	0.1	0	$1e^{-4}$	0	$2.5e^{-3}$

Fits for the maltose data set

For the maltose fitting routine, the fitting algorithm has lower bounds of 0 for all parameters, and no upper bounds except for parameters I_i, which has an upper bound of 1. Table 4.15 shows the RMSE scores divided by section of the curve and by weighting factor used for the fit. Table 4.16 shows the RMSE scores of the lowest scoring fit (overall score) by section of the curve and data type.

Table 4.15: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 5 for the maltose experimental data set.

Weight factor	5.0	1.0	10.0	2.0	0.5	0.1	I.P.
Overall RMSE	0.0456774	0.0470314	0.0475757	0.0480517	0.0483821	0.0571065	0.138003
Lag RMSE	0.0341704	0.0399212	0.0320045	0.033028	0.0418254	0.051641	0.034953
Exponential RMSe	0.0588438	0.0606472	0.0580573	0.066224	0.0629521	0.0752772	0.208984
Stationary RMSE	0.0403618	0.0369335	0.0489429	0.0380845	0.0361959	0.038079	0.110625

Table 4.16: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 5).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.040451	0.059209	0.038075	**0.046875**	
OD	0.026438	0.058476	0.042526	**0.044448**	
RMSE per phase	**0.034170**	**0.058844**	**0.040362**	**0.045677**	**Total RMSE score for the maltose fit**

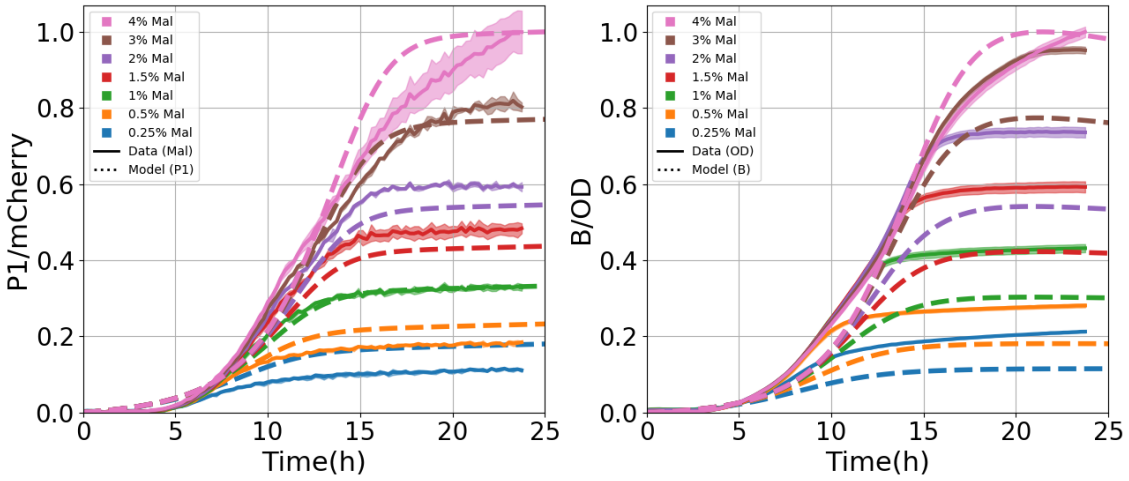


Figure 4.13: Fit to the to maltose experimental data set with no upper bounds. This corresponds to the parameterisation with the lowest overall RMSE score. The left graph corresponds to the fit to the mCherry fluorescence data, whilst the graph to the right presents the results of the fit to the absorbance data (O.D.).

Fits for the galactose data set

In the case of the galactose fit, the best fit is a result of parameterising the logistic-RFO model with lower bounds of 0 for all parameters, and upper bounds equal to ten times the initial parameter value, except for parameters I_i, which have an upper bound of 1. The reason for the implementation of upper bounds is that without them, for some weight factors, some parameters were optimised to more than 300 times the initial parameter value. This was not the case for the maltose fit,

which did not present this problem. Additionally, for the galactose fitting routine, the parameters corresponding to the absent sugar were explicitly ignored by the optimisation algorithm. When this metabolic branch was not ignored (as was the case for the maltose fit) the parameters were changed rather drastically. This stands in contrast to the maltose fit, in which the parameters corresponding to the absent sugar were hardly affected. Table 4.17 shows the RMSE scores divided by section of the curve and by weighting factor used for the fit. Table 4.18 shows the RMSE scores of the lowest scoring fit (overall score) by section of the curve and data type.

Table 4.17: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 1 for the galactose experimental data set.

Weight factor	1.0	5.0	2.0	0.5	10.0	0.1	I.P.
Overall RMSE	0.0406239	0.0432867	0.0432917	0.043908	0.0465057	0.0847988	0.188308
Lag RMSE	0.0171684	0.0141906	0.015175	0.0236251	0.0160985	0.0541885	0.0916783
Exponential RMSE	0.0524045	0.048932	0.0493312	0.0538257	0.0500648	0.127329	0.281704
Stationary RMSE	0.0437028	0.0550046	0.0543935	0.0482535	0.061014	0.0492296	0.136447

Table 4.18: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
GFP	0.016932	0.054940	0.057682	**0.047019**	
OD	0.017402	0.049740	0.022197	**0.033013**	
RMSE per phase	**0.017168**	**0.052404**	**0.043703**	**0.040624**	**Total RMSE score for the galactose fit**

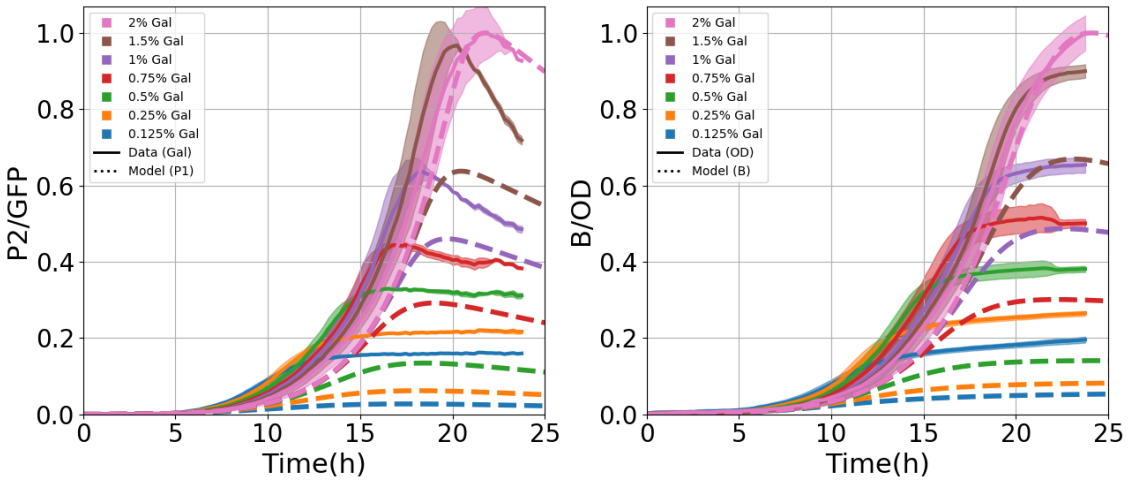


Figure 4.14: Fit to the to galactose experimental data set with no upper bounds. This corresponds to the parameterisation with the lowest overall RMSE score. The left graph corresponds to the fit to the GFP fluorescence data, whilst the graph to the right presents the results of the fit to the absorbance data (O.D.).

4.3.4 Conclusions

Table 4.5 shows the RMSE scores of each model divided by experimental data set and fitting routine. The Hill function biomass model with RFO_i induction term scored better for the PR1 data set as well as all three data sets in parallel. Logistic equation biomass and quadratic decay model with RFO_i induction term scored best at fitting the single sugar data sets. The maltose data set was fitted better with no upper bounds, whilst the galactose data set fit scored better with the implementation of upper bounds.

In general, the single sugar data sets scored better when the model was fitted to the single sugar data set exclusively rather than with other data sets (labelled as all data) or with the grid search method. The sugar mixture data set PR1 scored better through either an exclusive fit or the grid search method. Noteworthy is the fact that the grid search method delivers better scores on all data sets (both individually and collectively) for all models compared with the all data sets fitting routine, which fits the models with the three data sets collectively. The logistic equation biomass and quadratic decay model with RSO_i induction term is an anomaly in this regard since all but the PR1 data sets are better fitted with the all data sets fitting routine. Moreover, what seems to affect the overall RMSE score of the all data sets fitting routine and push it above the score yielded by the grid search method, is the fit to the galactose data set. The galactose data set score is noticeably larger than the score for maltose and PR1 data sets (in all models but for the logistic equation biomass and quadratic decay model with RSO_i induction term). On the other hand the grid search method seems to reach RMSE scores per experimental data set which are closer to one another and generally lower.

Regarding the binary-sugar mixture data set (PR1), it can be observed that the model that achieve the lowest RMSE score, the Hill-RFO model, struggles to reproduce the grouping pattern observed in the experimental data (section 4.1.2), particularly in the case of mCherry data (figures 4.10 and 4.12). In fact, both RFO

models (Hill and logistic) show difficulties reproducing this pattern (appendices A.2 and A.4). On the other hand, and despite scoring a higher RMSE score, both RSO models (Hill and logistic) perform better at reproducing and correctly separating the curves in the same group pattern as the one observed in the experimental data (appendices A.1 and A.3).

Generally speaking, based on the RMSE scores it can be observed that, for all models, the parameterisations presented here have the biggest difficulty fitting to the exponential phase of the curves, since the scores are generally higher in contrast to the stationary or lag phases. The higher rate of change in this region could account for this pattern. Additionally, the region around the second inflexion point (the separation point between the exponential and stationary phase) can differ between the model's output and the experimental data. This is particularly noticeable in the maltose data set (fig. 4.13). Moreover, the separation point, which was chosen based on the experimental data, could be different in the model's output, meaning that part of the model's stationary phase is captured as part of the exponential phase. This miss match could be a contributing factor driving the exponential RMSE scores higher. It is interesting to note that fitting the models to GFP data does not present this pattern. Fits to GFP data have a higher exponential phase RMSE score than the stationary phase score. The stationary phase of the curve presents the second highest RMSE scores, whilst the lag phase shows the lowest RMSE scores. The only anomalies in this regard are found in the galactose data set. Particularly, in the stationary phase of the curve of the GFP data set, which unlike the other experimental data sets, presents the highest RMSE scores out of the three curve sections. This is explained by the decay of the signal, which our model does not capture.

As I discussed in section 4.1.2, we suspect that the strong decay observed after the peak in the Gal10/GFP curves results from GFP's dynamics (half-life) rather than the enzyme's behaviour. Figure 4.1 shows how the two tagging proteins, mCherry and GFP, present different behaviour, even when they tag the same enzyme (Gal10

or Mal12) under the same experimental conditions. mCherry does not present the same strong decay from the peak as GFP does. This decay is not particularly well captured by our model, reflected by the higher RMSE scores observed in the Gal10/GFP curves' stationary phase fit. In contrast, our models had much more success fitting to the Mal12/mCherry curves' stationary phase, as indicated by the lower RMSE scores.

In this chapter, we determined the best model and the best parameter set based on the RMSE scores: The lower the score for a given experimental data set, the better the fit. The parameterised model with the lower score is considered the model that fits that data set the best. Our fitting methodology can be limiting in that it only provides us with a set of parameters. In contrast, procedures such as maximum likelihood estimation (MLE) or approximate Bayesian computation (ABC) could yield more information such as posterior probabilities of the models been tested, as well as a broader selection of parameter sets for the selected model (Ben Abdessalem et al., 2018). Furthermore, an improvement in model selection could be made by implementing penalised selection model criteria such as Akaike information criterion (AIC) and Bayesian information criterion (BIC), among others. Such tests aim to find a balance between the goodness of fit and model complexity (measured in the number of free parameters), penalising overly complex models (Kuha, 2004). These criteria could be implemented in our model selection process, particularly at later stages, e.g. choosing between RFO_i and RSO_i or between a Hill and logistic B equation. Selection criteria such as AIC or BIC could be useful even at the early stages of model structure selection (mutual inhibition, feed-forward inhibition, pooled inhibition). However, as mentioned in Chapter 3, our model structure choice was heavily informed by the literature on yeast regulatory mechanisms. Hence it would be unlikely to change.

As already discussed in section 4.3, several fits were performed for each model and experimental data set, each using a different weighting factor. The fitting results presented in this chapter correspond to the best fit. By this, I mean the set of

parameters that scored the lowest RMSE for a given model and a given experimental data set. It is evident that the best fits we could obtain with this method still do not capture all of the dynamics observed in the experimental data. Our models are simple and symmetric by design. This simplicity and lack of specificity come at the cost, which in this case is that even the best fits (according to RMSE score) still fail to replicate details in the experimental data.

I have mentioned already that our models do not include timescale separation, delays, dilution, nor fluorophore dynamics. Considering these elements could improve the model's goodness. In addition and as mentioned in Chapter 3, the P_i equations of our models represent the metabolic activity levels associated with sugar i. This approach is similar to model (3.1) in which equations e_i are described as the "lumped" system of inducible enzymes associated with the internalisation and catabolism their cognate sugars (Narang et al., 1997). A more detailed modelling of the specific metabolic pathways of interest would necessitate equations for each enzyme and metabolite in the chain, significantly increasing the model's size. Also, finding parameters for such a model would be challenging (Kholodenko, 2006). A reasonable expectation would be that a more complex and detailed model would fit the experimental data better. Subsequent iterations of our models could find greater success in fitting to the data by including some of the elements we decided to forgo, and in so doing, we could achieve a better understanding of the experimental data.

In this chapter, I have described the experimental data used for this project as well as the process through which the models introduced in the previous chapter were parameterised to fit the experimental data. Each data set has been fitted both individually and collectively. In addition, two approaches have been described in order to obtain a single parameterisation that manages to fit to all data sets. The grid search method yielded better scores than the " All data " fitting routine. Furthermore, both models with the RFO_i induction term generally yielded better scores for almost all data sets than their RSO_i counterparts. More specifically, the best overall collective fit is that of the Hill function biomass model with RFO_i

induction term through the grid search method.

In the next chapter I will take these model with the resulting parameterisations and analyse them as dynamical systems. I will study them under two different conditions: equal and unequal-value alternatives. We expect that this analysis will be informative about the decision-making dynamics that these models present.

Chapter 5

Dynamical systems analysis

In this chapter, I present the dynamical systems analysis, which includes bifurcation analysis, performed on the models developed in Chapter 3 (eq. (3.24)) and described in detail in section 3.2.4. The aim of these analyses is to study the dynamics and bifurcation of these models. In relation to the work presented in Chapter 4, these analyses should be considered as a separate study focused on the stability and dynamics of the models, not on their capacity to simulate experimental data from batch cultures. They can be considered as different systems.

To perform these analyses, we make the assumption that the extracellular nutrient concentrations are constant, which is necessary in order to achieve a steady-state in the system. By implementing this assumption, the models go from having eight state variables down to six (state variables N_i are now constants). Using these models, we analyse two different settings: 1) Equal-value alternatives case, and 2) unequal-value alternatives case. The equal-value alternatives case depicts a scenario in which there are two sources of the same nutrient with a constant concentration. This exploration of the model's dynamics presents similarities to the analysis presented by Pais et al. (2013) on house hunting honeybee swarms, or the studies of Dussutour et al. (2019) and Vogel et al. (2018) with slime mould. On the other hand, the unequal-value alternatives case represents a situation in which there are two different nutrients

at different and constant. One such scenario would a continuous culture in which new media is fed to the culture constantly, in an effort to maintain the nutrient concentration constant.

For the equal-value case, we present the results of the two RFO models (Hill-RFO, and Logistic-RFO). In the unequal-value case, call models are analysed. I present a description of the general approach that was taken in order to perform the bifurcation analysis, as well as introduce the tools used to perform such an analysis. Subsequently, I describe how these tools were used in order to obtain the results. Additionally, I also analyse the model under saturating conditions, as well as a variant of the unequal-value alternatives case by taking the equal-value alternatives parameterisation and varying a single parameter. These conditions are variants of the unequal-value alternatives case, and their results can be found in appendix B.5. The analyses presented by Pais et al. (2013) and Zabzina et al. (2014) on the equal-value alternatives case is characterised by the presence of pitchfork bifurcations, revealing that in such decision the state of deadlock can be broken by increasing the value of certain parameters. In particular, our results show parallels to Zabzina et al. (2014), which show that the deadlock state can be restored if the parameter value is increased further.

5.1 Analysis of the binary-sugar mixture decision of a cell culture

As previously mentioned, our interest lies in using our models to explore the cellular decision-making behaviour that a cell or a cell population presents when it encounters two sources of nourishment in its vicinity. In order to perform the dynamical systems analysis, and more specifically, the bifurcation and stability analysis, we must consider certain assumptions that are necessary to obtain a steady-state, which is necessary for any dynamic systems analysis. We consider the extracellular sugar

concentrations N_i as constant. Without this adjustment, the extracellular sugar concentration would be depleted over time, implying that a steady-state would not be reached until the nutrients, in our case the extracellular sugar concentrations N_i, are completely exhausted, which makes the steady-state rather uninteresting. In contrast, by making the extracellular sugar concentrations N_i constant, the system is allowed to reach a non-zero steady-state.

From a biological standpoint, one of the assumptions made is that the value of a carbon source to an organism is a function of the energetic contribution or growth advantage it grants to the organism, which itself can be seen as an aggregate of various factors such as the grams of biomass produced per gram of sugar consumed, the rate at which the sugar in question is internalised and metabolised, as well as its extracellular concentration (Perkins and Swain, 2009; Wang et al., 2015; Koirala et al., 2016). Other possible factors that can affect the cell's preferences are related to the role that a specific sugar plays in anabolic pathways. For instance, Shimada et al. (2013) argue that D-ribose has a higher value to *E. coli* than its energetic contribution would merit due to its role in nucleotide synthesis. One fundamental assumption in our analyses is that the model's parameters determine the sugar alternatives' value. The analyses performed can be divided into two different categories: 1) Equal-value alternatives, and 2) Unequal-value alternatives. The latter case is where two different sugars are available within the media. This would correspond to analysing our models with the parameters corresponding to the binary-sugar mixture experimental data. On the other hand, the equal-value alternatives case corresponds to the theoretical situation in which the system is presented with two sugars with the same parameters. This situation does not necessarily reflect single sugar experimental conditions, an environment in which one branch of the system remains inactive since there is only one sugar available. In the equal-value alternatives case, there are two extracellular sugars pools available to the system, which has two active branches with the same parameters. By branch I refer to the two sets of equations present in the models (N_i, S_i and P_i) which are structurally identical

to one another (they are only differentiated by their parameters).

5.1.1 Reparameterisation centred around the mean and the difference

We follow the approach presented by Pais et al. (2013) regarding the bifurcation analysis of a two-alternative decision problem. Pais et al. (2013) present a model, which analyses the value-based decisions of a honeybee swarm scouting and eventually choosing a new location in which to establish a new nest, uses two main parameters, namely the cross inhibition, and the perceived value or quality of the location in consideration. The two equations that define their model are structurally similar to the P_i equations in our models (explained in detail in Chapter 3). Similarly to ours, the two equations are symmetrical to one another, only distinguished by their parameterisation. Following their analysis and taking advantage of the structural symmetry of the model, most of the parameters in our models can be replaced by parameters that represent the mean value and the difference in their value between the two branches. The relationship between the mean and difference and the original parameters on each branch of the model are described by

$$\begin{aligned} p_1 &= \bar{p} + \Delta p \\ p_2 &= \bar{p} - \Delta p, \end{aligned} \tag{5.1}$$

where p_i is the a parameter corresponding to the ith branch, $\bar{p}$ is the mean value of the two parameters, and Δp corresponds to half the difference between the two parameters. Making these substitutions maintains the original parameter values, and allows us to perform the bifurcation analysis by varying the value of either the mean or the difference between parameter, rather than the absolute values. Parameters V_B, K_G, M and Dcy_B are unique parameters, and as such they were analysed as they are.

This reparameterisation is made in order to study how the system's dynamics change in response to changes to the mean of parameters and their difference. This is particularly relevant in the case of certain parameters of interest, which can be interpreted as indicators of the value of a sugar alternative to the yeast culture (K_{Pi}, V_{Pi}, Y_i), and the strength of inhibition (K_{inhi}). In this regard, analysing the mean and difference rather than the individual absolute values, serves to analyse the system's behaviour as the mean value of its alternatives changes, or as they become more distinct, which allows us to analyse the models in the context of value-sensitive decision-making, describe in detail in section 2.1.7.

5.1.2 Software used

In order to perform the dynamical systems analysis of the models we used MatCont, a MATLAB package for numerical bifurcation analysis of ODE systems (Dhooge et al., 2003). We use MATLAB version R2018a for these analyses (MATLAB, 2018). I tested all the different solvers available through MatCont and selected the solver ode15s, as it seems to handle our models the best in terms of efficiency and consistency. The ode15s is a solver that is particularly suited to solve stiff equations (MathWorks, 2020). Stiffness in the computational context refers to a system of coupled differential equations for which certain numerical methods are inefficient or unstable (Liu et al., 2019). Rathinam et al. (2003) mention two main reasons behind the stiffness of a system: 1) multiple timescales, and 2) the presence of variables of both small and large quantities. In our system, the state variables range over several orders of magnitude, which would cause stiffness. Hence, the necessity to employ numerical methods that can handle such systems.

Additionally, I used the *Dynamica* package for *Mathematica*, which similarly to MatCont is a package aimed to provide tools for the analysis of dynamical systems. Going into more detail, MatCont is a graphical software package for MATLAB that allows the user to perform continuation and numerical studies on parameterised

dynamical systems, specifically, bifurcation analysis, and it facilitates the computation of curves of equilibria, branch points, Hopf points, fold bifurcations and more. (Govaerts et al., 2019). MatCont performs its curve continuations by implementing a prediction-correction algorithm, which, as the name implies, predicts where the next point of a curve is, and then corrects it. The specific algorithm used is the Moore-Penrose continuation (Dhooge et al., 2003). Through the implementation of several test functions, MatCont can detect bifurcation points such as Hopf and fold bifurcations. MatCont also allows the user to monitor the eigenvalues of the system, which can be used to understand the stability of the solution (Strogatz, 1994). Similarly, Dynamica is a package build for Mathemathica for the analysis of dynamical systems, and it allows the user to compute vector fields, trajectories, flows, phase portraits, and bifurcation diagrams (Beer, 2016). An in-depth discussion of how either one of these software packages work goes beyond the scope of this work. For a more detail description see Dhooge et al. (2003); Govaerts et al. (2019) and Beer (2016).

I used Dynamica to obtain the trajectories and phase portraits in fig. 5.1, and especially in the initial stages of the bifurcation analysis for its ease of use to explore the parameter space. It also plays an important role during the single-parameter difference analysis (appendix B.5). Whilst MatCont requires the user to manually input the values for the state variables in order to integrate a particular solution, Dynamica does not. It does a sweep through the parameter space finding all the possible solutions to the systems. This is a particularly useful feature, especially when the solutions are not connected to one another, as is the case with the bifurcation analysis of our system under the single parameter difference case.

5.1.3 Colour coding in the bifurcation diagrams

The bifurcation diagrams in the following sections are all presented as either three- or two-dimensional diagrams that show the evolution of the system's dynamics as a

function of a parameter's value within the P_1, P_2 space. The choice of coordinates comes about because P_1 and P_2 represent the system's metabolic activity required to consume either sugar and therefore, their dynamics are a good representation of the system's commitment to either alternative. The colour of the solutions in the diagrams represents the stability of that solution: Blue represents a stable solution, in which the real part of all eigenvalues is negative. The resulting solution can either be a node or a spiral (depending on whether the eigenvalues are real or complex, respectively) (Strogatz, 1994). A stable solution means that all local trajectories lead to it. Green represents an unstable solution in which the real part of one of the eigenvalues is positive; this is an unstable solution. More specifically, within the P_1, P_2 plane, this solution behaves like a saddle point. Finally, red represents a solution in which the real part of two eigenvalues is positive. Within the P_1, P_2 plane, this solution is unstable. It can be either an unstable node or an unstable spiral, which depends on the eigenvalues being real or complex. This type of solution implies that within the P_1, P_2 plane, all local trajectories move away from it. A graphic representation of these solutions can be observed in the phase portraits in figure 5.1.

5.2 Equal-value alternatives

In this section I present the results obtained from applying dynamical systems analysis to the models in an equal-value alternatives decision problem. The equal-value alternatives case refers to the hypothetical situation where the system is presented with two sugar alternatives with the same value, in other words, the same sugar. This is simulated by parameterising the models with the same parameter values in both branches. This means that the parameters in the two branches of the models, as well as parameters V_{G1}, Y_1, V_{G2}, Y_2, have the same value. As mentioned in the previous section, this setting does not reflect the single sugar experiments used in the previous chapter. In that case, one of the branches would not present any relevant activity since there would only be one extracellular sugar pool N_i. The approach

put forward here is a theoretical exploration of the model's dynamics, similar to the bifurcation analysis presented by Pais et al. (2013) on house hunting honeybee swarms, or the studies of Dussutour et al. (2019) and Vogel et al. (2018) with slime mould.

The results displayed correspond to the two RFO models. The parameterisation used for these analyses is the initial parameterisation obtained through the manual estimation of parameters, as specified in Chapter 4. Similarly, the initial conditions are those specified in the previous chapter, obtained through grid search. To implement the equal-value alternatives setting, the parameters of the galactose branch were chosen and the two branches of the model were parameterised with them. It is noteworthy that analysing the difference parameters leads to no bifurcations. The bifurcations obtained emerge by changing the value of mean parameters, as well as unique parameters such as V_B. Additionally, these resulting bifurcations, regardless of the parameter that elicited them, present a consistent pattern in their shape and overall dynamics. Our results show a restoration of symmetry similar to that described by Vogel et al. (2018), where bifurcations along a unique stable solution present themselves in pairs: The initial branching point creates a pitchfork bifurcation which changes the stability of the unique solution from stable to unstable, and from which two new solutions emerge. The second bifurcation occurs at the second branching point where the two solutions which emerge at the first branching point collapse. After this point, only the initial unique solution remains with its stability restored. It is worth mentioning that not all the bifurcation diagrams are shown here, those missing can be found in appendix B.1.

5.2.1 Phase portraits of the equal-value alternatives case.

In all of the results presented, the dynamics of the system were analysed in the P_1, P_2 space. To reiterate what has been stated in previous chapters, they represent the metabolic activity of the pathway related to the consumption of their cognate

sugar. Biologically, this can be the concentration of a specific and essential enzyme in the metabolism of a particular sugar. The assumption being made here in regards to decision-making is that the commitment, or decision of the system, to either one of the sugar alternatives is expressed or represented by the relative levels of P_1 and P_2. Whereas the bifurcation portraits that constitute the bulk of the results presented in this chapter describe the dynamical changes undergone by the system as the value of one parameter is changed, the phase portraits presented in figure 5.1 present the state of the system with a particular parameterisation. For all intents and purposes, these phase portraits can be understood to be longitudinal slices taken from a bifurcation diagram at a particular parameter value. They depict all the solutions available to the system, as well as their stability in a two-dimensional P_1, P_2 plane. Additionally, the phase portraits also show several trajectories in that particular state of the system to illustrate the solutions' dynamics and stability. The portraits in figure 5.1 are qualitative in nature and are presented in order to illustrate the different dynamic states that the models produce under different parameterisations.

Portraits 5.1a and 5.1b depict a single stable solution, to which all trajectories are attracted to. From the decision-making perspective, this unique stable solution represents a decision-deadlock state in which the system is not committed to any single alternative. However, these two deadlock states can not be interpreted to be qualitatively equal. Portrait 5.1a shows how all trajectories in the space are driven to a near-zero deadlock state. This comes about as a result of the system, and in particular the induction terms of equations P_i, not being active. P_i has two sources: 1) the induction term which is sensitive to S_i, and 2) the basal activity represented by the term $A_i * V_{Pi}$, which is independent of the value of any state variable. This state of deadlock represents an inactive or un-induced cell culture, the metabolism of which is inactive. In this state, neither of the sugar alternatives are being properly utilised, therefore, the system is uncommitted. On the other hand, the deadlock state depicted in portrait 5.1b, due to its high P_i value can be interpreted as a

state in which both metabolic branches are active, hence, both sugar alternatives are being consumed simultaneously. The state depicted here can be interpreted as a commitment to both alternatives through simultaneous consumption.

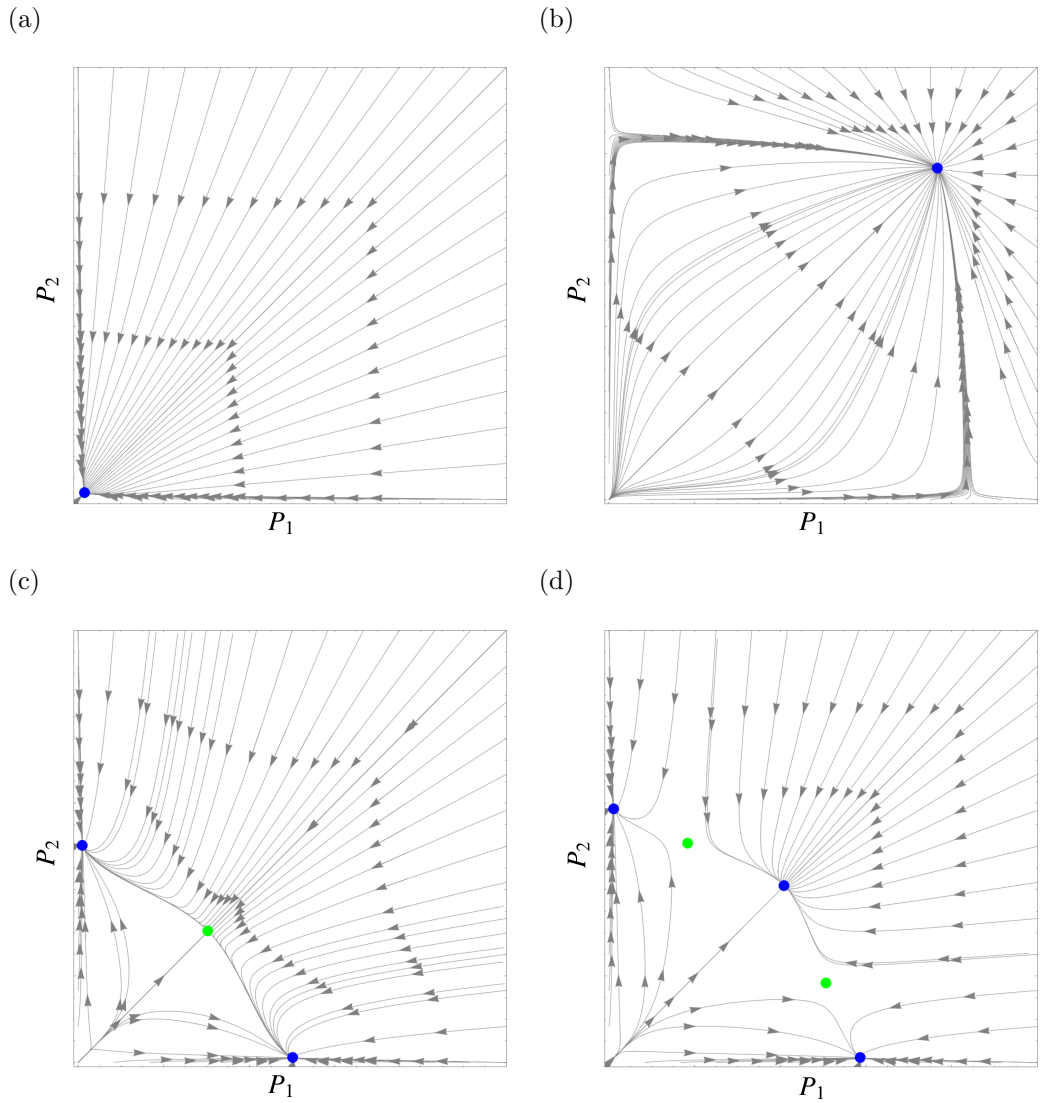


Figure 5.1: Phase portraits in the P_1, P_2 space, showing the different dynamics that the RFO models can present in the equal-value alternatives case. Blue dots represent stable attractors, green dots are saddle attractors, and red dots are unstable solutions.

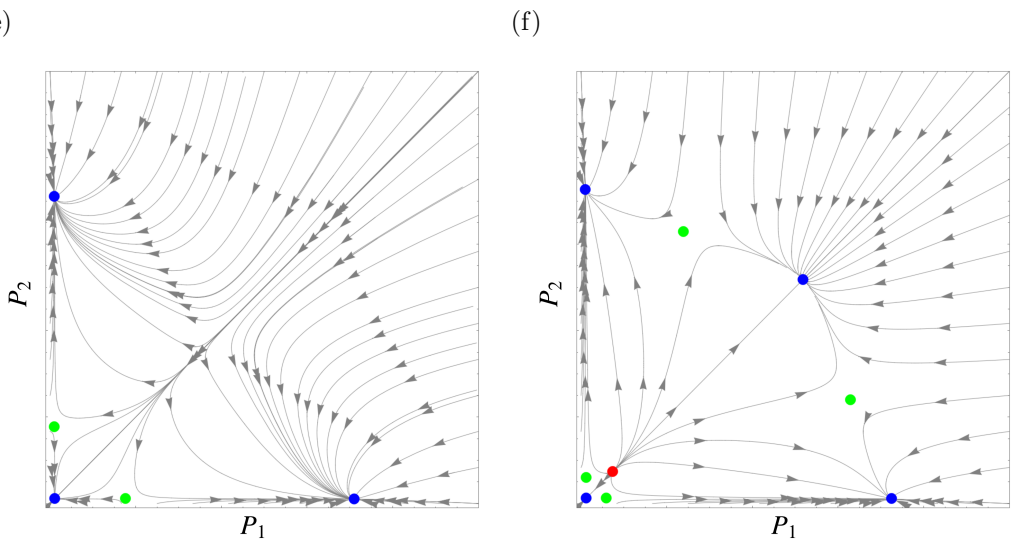


Figure 5.1: cont.

Finally, portrait 5.1f shows that, in some cases, all these solutions are available to the system. It is important to mention that not all bifurcation portraits exhibit all these behaviours. In particular the dynamics in phase portrait 5.1f are only seen in bifurcation diagrams such as those of figure 5.9. Portrait 5.1c shows a system in which a deadlock solution is unstable. The saddle node displayed in green pushes approaching trajectories to either one of the two stable solutions. These two stable solutions represent a commitment of the system to the consumption of one of the alternatives over the other. Portraits 5.1d and 5.1e show the coexistence of these two stable solutions, as well as a state of deadlock.

5.2.2 Hill function biomass model: RFO induction term

Here I present the results of the dynamical systems analysis on the Hill-RFO model. Table B.1 shows the values of the parameters in terms of mean and difference. The parameterisation used corresponds to the initial PR1 parameterisation of this model obtained through manual pre-visualisation. Table B.2 shows the initial values of the state variables used, which were taken from the experimental data fit initial conditions. The second row are the state variable values of the steady-state. These are obtained by following the process described in appendix B.3.

Parameters $\bar{A}$, $\bar{Dcy_P}$, $\bar{K_S}$, $\bar{I}$, ΔA, ΔDcy_P, ΔV_S, ΔV_P, ΔV_G, ΔK_S, ΔI, ΔK_P, ΔY, ΔN, K_G, M and Dcy_B have one stable solution across the entirety of the analysed interval. In contrast, the system presents bifurcations as a result of varying any other parameters.

Two parameters of particular interest are $\bar{K_P}$ and $\bar{K_{inh}}$. $\bar{K_P}$ (the mean of parameters K_{P1} and K_{P2}) represents the affinity that the induction of P_i has for its sugar substrates. P_i has an inversely proportional relation with K_Pi. Therefore, the lower this value is, the more sensitive the synthesis of P_i is to changes in sugar concentration and the faster it reaches saturation levels. In contrast, the larger this value is, the more sugar concentration is needed in order to induce the synthesis of P_i. On the other hand, $\bar{K_{inh}}$ (the mean of parameters K_{inh1} and K_{inh2}) is a parameter that regulates the strength of the inhibitory effect of G onto the synthesis of P_i, and it has a proportional relationship with the value of P_i. K_{inhi} affects G, which in turn affects K_Pi. Hence, the higher K_{inhi} is, the smaller the repressive effect of G over K_{Pi} is.

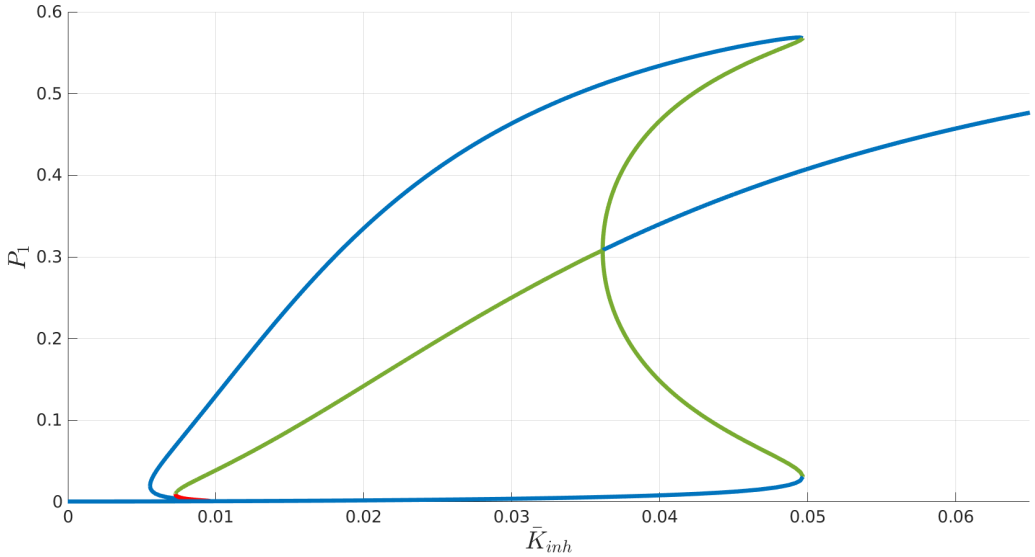


Figure 5.2: **Top**: Two-dimensional bifurcation diagram showing the value of P_1 as $\bar{K_{inh}}$ is changed. **Bottom**: Same bifurcation diagram showing only the deadlock solution to illustrate the stability changes.

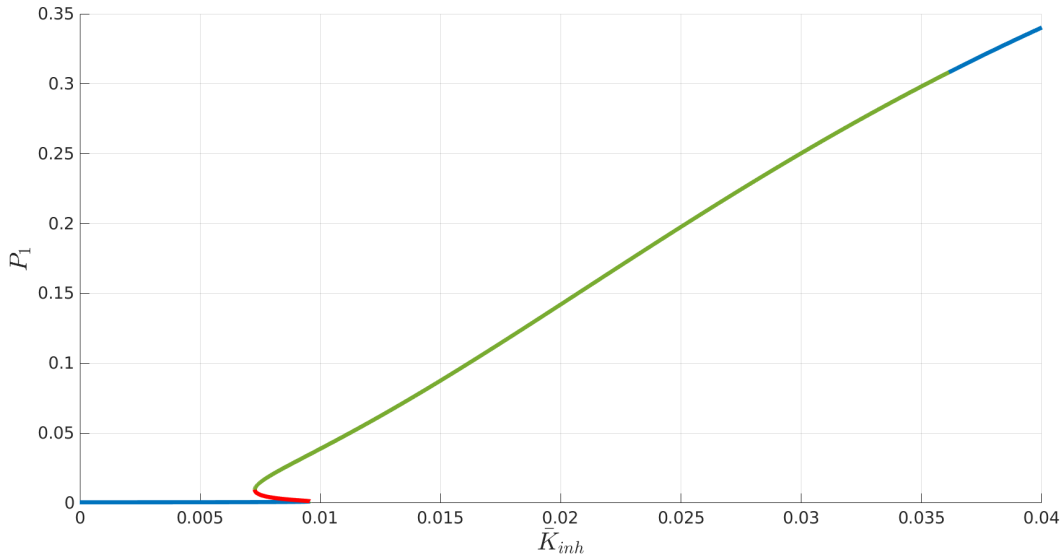


Figure 5.2: cont.

Figure 5.2 shows a two-dimensional bifurcation diagram (Top) that illustrates how the value of P_1 changes as $\bar{K}_{inh}$ is varied. Since is is an equal-value alternative case, the P_2 diagram is equivalent to this one. It can be observed that there are two different types of stable deadlock. In this context, deadlock means that both P_1 and P_2 have the same value, therefore, the system shows no preference between the sugar alternatives. It should also be noted that state variables S_1 and S_2 also have the same value. At low values of $\bar{K}_{inh}$ the deadlock state is characterised by having a near-zero P_i value. At this point $\bar{K}_{inh}$ value is approximating zero, which in turn strengthens the inhibitory effect of G, thereby reducing the synthesis rate of P_i whilst only the basal production of P_i (through the term $A_i * VPi$) continues unfettered. As the value of $\bar{K}_{inh}$ increases, a branch point is encountered, upon which the solution becomes unstable, first as a small region in which two eigenvalues turn from negative to complex conjugates with a positive real component; this is the red part of the solution. Subsequently, one of the complex eigenvalues with positive real part turns negative, which corresponds to the solution's green region. Within the P_1, P_2 space, this has the dynamic properties of a saddle. As the $\bar{K}_{inh}$ continues to increase the average value of the deadlock saddle solution increases until a second branch point is encountered. At this point the deadlock solution becomes completely stable. This can all be observed in the bottom diagram of figure 5.2, where the deadlock solution

has been isolated so that the bifurcations and stability changes can be observed better. The average P_i value of this stable deadlock solution is much larger than the stable deadlock region at lower $\bar{K}_{inh}$ values. As $\bar{K}_{inh}$ increases, the inhibition effect is reduced, allowing state variables P_1 and P_2 to increase. Biologically, this can be interpreted as reducing the affinity with which a repressor inhibits the expression of a given gene; the lower the affinity the lower the inhibition of protein synthesis through genetic expression. This stable deadlock solution reflects the state of the system in which both sugar alternatives are being exploited to the same extent. At the first branching point, not only does the stability of the deadlock solution change, but two other solutions emerge which are symmetrical to one another. Each solution depicts the commitment of the system to either one of the sugar alternatives, whilst the other is relatively ignored. These two solutions start as saddle solutions (the real part of one of the eigenvalues is positive) before becoming stable solutions as they encounter a limit point. This change in stability is indicated by the change from green to blue colouring. As $\bar{K}_{inh}$ increases, depending on which branch one follows, the value of either P_1 or P_2 increases above the P_i value of the saddle deadlock solution. In this region, there are two stable solutions available to the system. The deadlock solution behaves as a saddle which makes it push the trajectories toward either one of the two stable solutions, as can be seen in phase portrait 5.1c.

Whilst the analyses presented here are deterministic, in a naturalistic scenario, one would expect the system to show stochasticity. As Pais et al. (2013) mention in their work, when analysing equal-value problems with noise, if the system has multistability in the form of a pitchfork bifurcation, the system will find itself in or choose, either one of the stable solutions at random. The two-dimensional diagram in figure 5.2 shows how in one of the solutions the value P_1 increases as $\bar{K}_{inh}$ increases, whereas in the other stable solution its value remains both close to zero and relatively flat. The P_2 bifurcation diagram is identical to the P_1 diagram shown here. If one were to add noise terms to our model (similar to model (C.1) in appendix C) and ran the model with a $\bar{K}_{inh}$ value that allowed for one saddle deadlock solution

and two stable solutions (for instance $\bar{K}_{inh} = 0.03$), the P_i equations' output will equal either one of the two stable solutions in which P_1 is higher than P_2, or vice versa, but never the saddle solution. The deadlock state is only available if the model is run as deterministic, due to its stability.

Eventually, both of these stable solutions find a second limit point which changes their stability again back to that of a saddle, as one of the eigenvalues has its real component turn positive. These saddle solutions then curve back to collapse at the second branch point, leaving only the deadlock stable solution. This behaviour contrasts with that reported by Zabzina et al. (2014). Their system also shows multiple stable solutions as the system size parameter is increased. The difference is that their system does not present the second limit points in each of the branches that our system presents, therefore, the non-deadlock stable solutions never collapse. This means that as the system size parameter is increased, their system always presents three different stable solutions: a deadlock solution and two other in which the system is committed to either one of the two alternatives. It is important to note that this second bifurcation occurs at a lower $\bar{K}_{inh}$ value than the limit points at which the second and third solutions change their stability and move back. In a similar fashion, the initial bifurcation occurs at a higher $\bar{K}_{inh}$ value than that at which the two emergent solutions change their stability from saddle to stable. This results in two regions in which there are three stable solutions that the system can potentially find itself in. The two saddle solutions in between them will push any trajectories to either one of the stable solutions, as can be seen in the vector fields of phase portraits 5.1d. Additionally, it can be inferred from these vector fields, that the closer the stable and saddle solutions are to one another, the more likely it can be for a given trajectory to be pushed away from that solution and into a different stable attractor.

The system behaves differently as a function of changing the values of parameter $\bar{K}_P$, as can be seen in the bifurcation diagrams of figure 5.3, although the same pattern observed in the analysis of parameter $\bar{K}_{inh}$ presents itself again: stable

unique solutions flanking a multistable region in the centre, the start and end of which are determined by the emergence of a subcritical-like bifurcation at each side. The chief difference between the dynamics presented by varying parameter $\bar{K}_P$ in contrast with $\bar{K}_{inh}$ is the relation that the average P_i value has with the changing value of the parameter being analysed; in response to changes in $\bar{K}_{inh}$ the average P_i value increases as $\bar{K}_{inh}$ increases, whereas this relationship is inverted in response to $\bar{K}_P$, as it can be observed that the average P_i value decreases as $\bar{K}_P$ increases. As $\bar{K}_P$ starts to increase from zero the system finds itself in a single stable solution representing a deadlock state in which both alternatives are being used. Similarly, a multistable region in which two solutions, each of which represents the commitment to one of the alternatives over the other, emerges from a branch point as the initial deadlock solution becomes unstable.

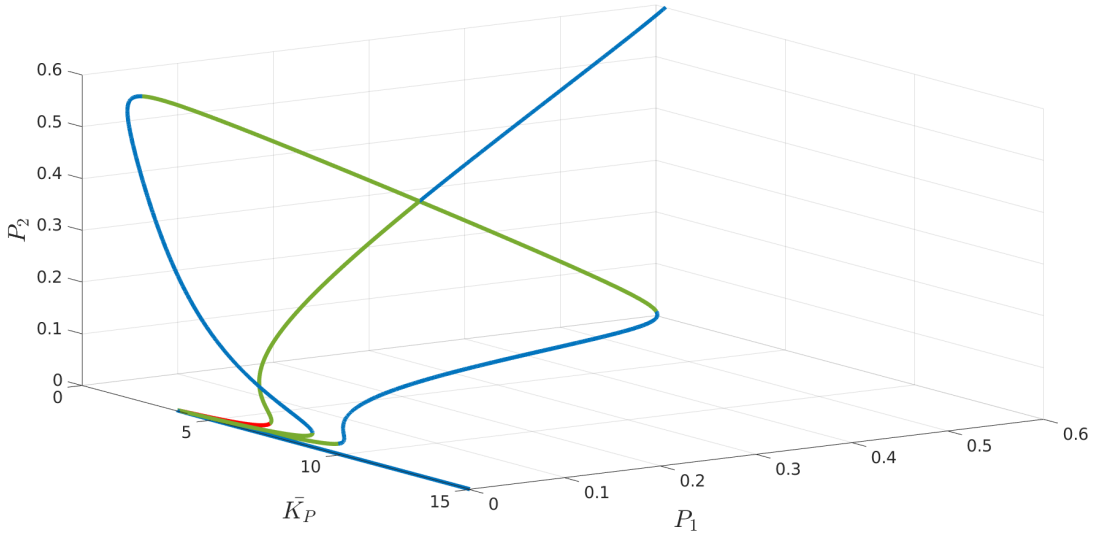


Figure 5.3: **Top**: Three-dimensional bifurcation diagram resulting from varying the $\bar{K}_P$ parameter in the P_1, P_2 space. **Bottom**: Two-dimensional bifurcation diagram showing the same solution, focusing on the value of P_1 as $\bar{K}_P$ is changed.

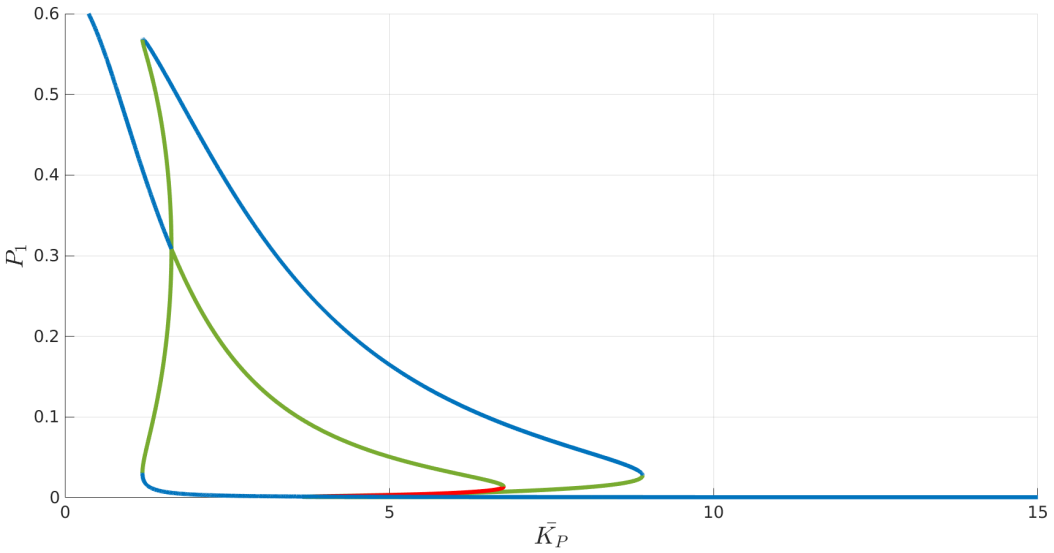


Figure 5.3: (cont.)

As $\bar{K}_P$ continues to increase, the average P_i value decreases, until a second bifurcation is found, upon which the multistable region comes to an end as the multiple solutions collapse, and the deadlock solution regains its stability. In this final deadlock state, the P_i values are dominated by the basal rate of production ($A_i\, V_P i$), since the output of the induction term is close to zero. Parameter $\bar{K}_P$ is inversely proportional to the affinity of the substrate, hence the higher the $\bar{K}_P$ value is, the more concentration of S_i is needed for P_i to be produced.

From these two sets of diagrams (figures 5.2 and 5.3), one can also observe that the solution which represents the deadlock state tends to display a "hook" or S-shaped curve as the system approaches one of the bifurcations, namely that which emerges from or resolves into the basal-activity deadlock state. This dynamic phenomenon is called hysteresis which can be defined as a jump between solutions as a parameter is changed (Seydel, 1988). Hysteresis loops are also responsible for lagging effects, or "switch" effects, as they are sometimes called in the context of genetic expression (Ozbudak et al., 2004b; Laurent and Kellershohn, 1999). It can be seen in both figures 5.2 and 5.3 that as the value of a parameter is increased the expression levels of P_i remain relatively static before a point of drastic change, which is called the hysteretic point, where the system jumps from that solution onto another one with

a significant and immediate change in P_i values. The solutions that emerge from the branching points show a similarly shaped curve near both bifurcations, which is what results in the regions where three stable solutions are available to the system. Phase portrait 5.1e depicts how the system's dynamics look in such regions.

A clear exception to this pattern is the system's dynamics in response to varying values of ΔK_{inh}. The diagrams in figure 5.4 show the dynamics yielded by varying parameter ΔK_{inh}. What can be observed is that as the parameter approaches the values at which either K_{inh1} or K_{inh2} approach zero the system presents, for a small range of values, hysteretic behaviour. The single solution presented by this diagram shows that the middle section is both stable and flat, which means that the system is quite resistant to changes in ΔK_{inh}, since for most of the interval explored the values of P_1 and P_2 are extremely close to one another, although they are not identical.

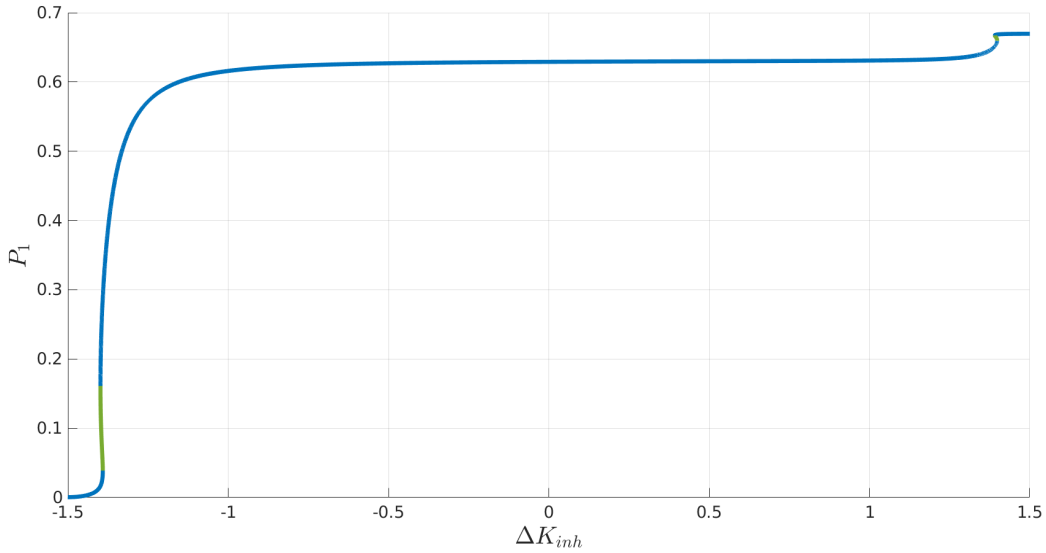


Figure 5.4: **Top**: Two-dimensional diagram resulting from varying the ΔK_{inh} parameter in the P_1, P_2 space. **Bottom**: Two-dimensional diagram showing the same solution, focusing on the value of P_1 as ΔK_{inh} is changed.

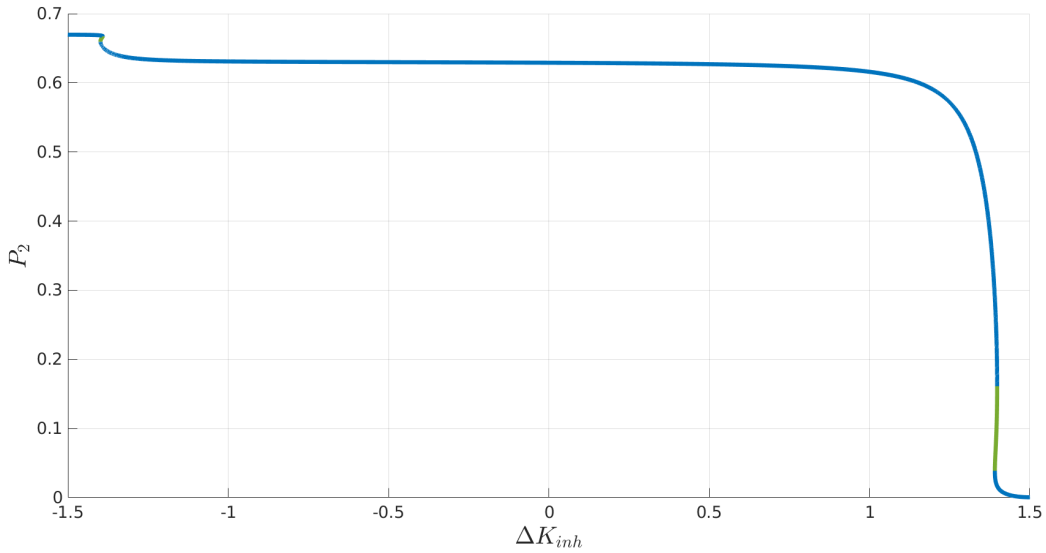


Figure 5.4: (cont.)

Eventually, these values start to diverge noticeably as either P_1 or P_2 drops in an almost straight line, depending on the sign of ΔK_{inh}. This drop occurs at $K_{inhi} \approx 0$. For instance, if ΔK_{inh} has a positive value as it increases toward the boundary value at which $K_{inh2} \approx 0$, then it is the value of P_2 which collapses, whilst P_1 increases slightly. The situation is mirrored if $K_{inh1} \approx 0$ as ΔK_{inh} is decreased. At these extremes, the solution changes from a stable one to a saddle. These regions are also bistable, since there is an overlap of the stable parts of the solution.

The pattern mentioned previously, namely a stable solution with a multistable region in the middle, is present in almost all of the analyses made. All the parameters that exhibited this pattern can be classified based on the relation that the average P_i value has to the value of the parameter: proportional as is the case of parameter $\bar{K}_{inh}$, or inversely proportional such as $\bar{K}_P$. Of the parameters whose results are not shown, parameters $\bar{Y}$ and $\bar{V}_G$ present an inversely proportional relationship with the average P_i value, such as $\bar{K}_P$, whereas parameters $\bar{V}_P$, $\bar{V}_S$ show a proportional relationship with the average P_i value, like $\bar{K}_{inh}$. Additionally, parameters $\bar{V}_P$, $\bar{V}_S$ have a more pronounced hysteresis loop, not dissimilar to that depicted in figure 5.9, where the multistable region encompasses four stable solutions, each one of the stable solutions that represent commitment to only one of the alternatives, and two

different stable deadlock solutions, one of high P_i value and one of low P_i value.

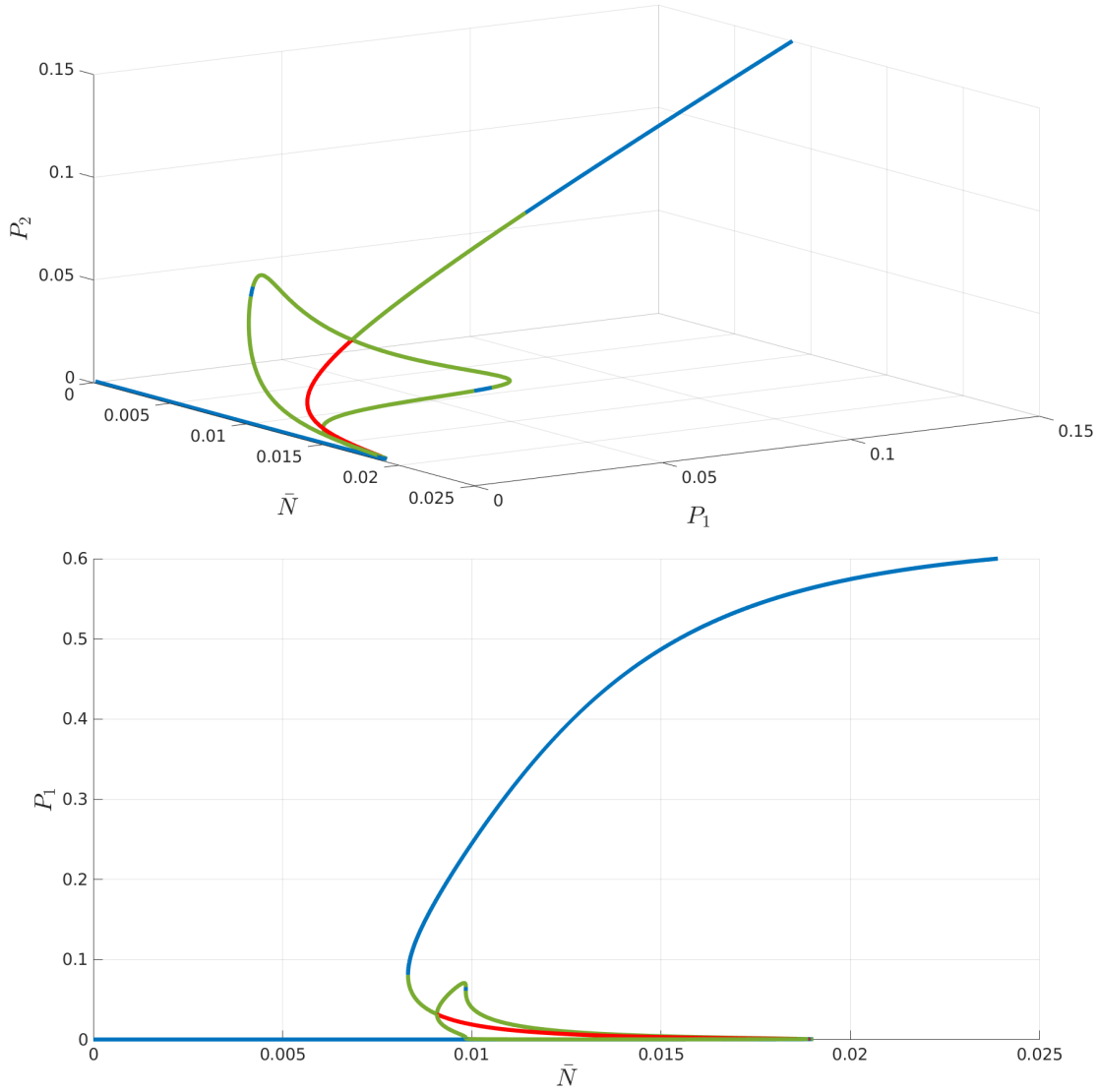


Figure 5.5: **Top**: Three-dimensional bifurcation diagram resulting from varying the $\bar{N}$ parameter in the P_1, P_2 space. **Bottom**: Two-dimensional bifurcation diagram showing the same solution, focusing on the value of P_1 as $\bar{N}$ is changed.

Another parameter of particular interest is $\bar{N}$. $\bar{N}$ is the mean extracellular sugar concentration. In all other analyses this quantity has been kept constant. However, by varying it we can analyse the dynamics of the system as the concentration of sugar in the medium changes. The diagrams in figure 5.5 reveal that, similarly to the dynamics of the system in relation to changes to parameters $\bar{V_P}$, $\bar{V_S}$, the deadlock solution in these diagrams also presents hysteresis. In this case, the multistable region flanked by bifurcations is present. However, in this case, the stable sections of the solutions that emerge at the branching points are relatively small, compared to

the length of these solutions. Also, the deadlock solution does not regain its stability at the second bifurcation. In this point, the solution turns from unstable to a saddle (in the P_1, P_2 space), and it is only after that this solution eventually regains its stability. It is also important to note that this behaviour occurs as parameter $\bar{N} \approx 0$, which means that the extracellular sugar pools are close to being depleted.

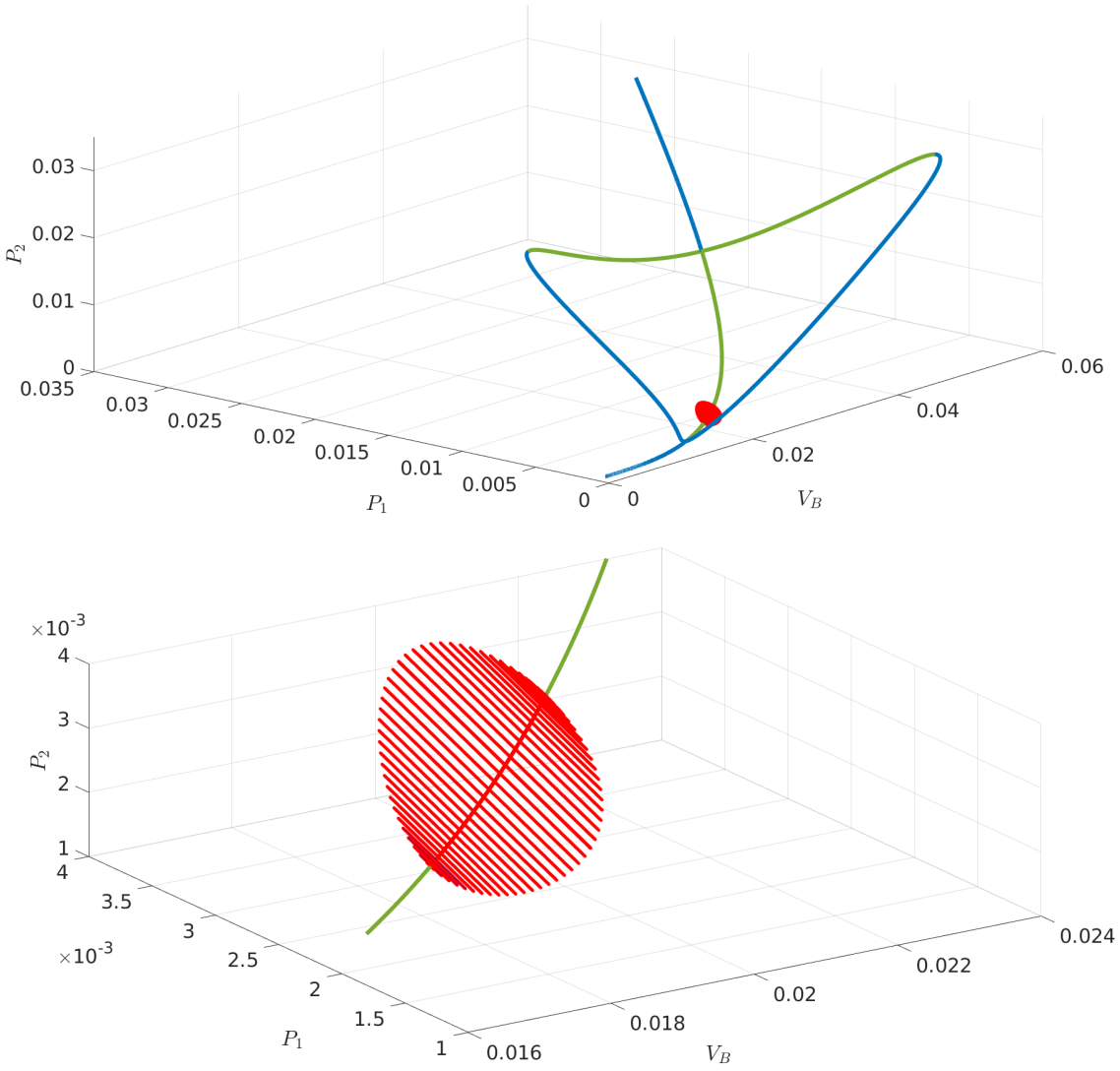


Figure 5.6: **Top**: Three-dimensional bifurcation diagram showing the dynamics of the system in the P_1, P_2 space as V_B is changed. What differentiates this bifurcation from others is the small unstable region in the deadlock solution that presents some periodic solutions. **Bottom**: Zoomed-in plot of the periodic solutions, in which $P_1 = P_2$.

Lastly, as parameter V_B is change, similar dynamics are observed. However, the bifurcation diagrams in figure 5.6 reveal that these solutions lack the characteristic hysteresis loop, and the multistable region. The deadlock solution in this case has

the shape of an exponential curve. Another remarkable feature is that the deadlock solution's unstable region has at each extreme a Hopf point, from which oscillatory behaviour emerges. This region is relatively small, and the oscillations occur at the middle of an unstable region, with two stable solutions available to the system. Moreover, the oscillations themselves are unstable, since two of the eigenvalues have a positive real component. It is noteworthy that this behaviour emerges at relative by small values of V_B. The initial parameter value is 0.9, whilst this dynamics emerges as the value is dropped to $0.04 - 0.01$. Additionally, in these periodic solutions $P_1 = P_2$.

5.2.3 Logistic equation biomass model: RFO induction term

In this subsection I present the results of the dynamical systems analysis on the logistic-RFO model. Table B.3 shows the values of the parameters in terms of mean and difference. This parameterisation corresponds to the initial PR1 parameterisation yielded by manual pre visualisation. Table B.4 shows the initial values of the state variables used, which were taken from the experimental data fit initial conditions. The second row are the state variable values of the steady-state. These are obtained by following the process described in appendix B.3.

Just as was done with the Hill-RFO model in the previous section, all the parameters were analysed. Parameters with only one stable deadlock solution are: $\bar{A}$, $\bar{Dcy_P}$, $\bar{I}$, $\bar{Y}$, ΔA, ΔDcy_P, ΔV_S, ΔV_P, ΔV_G, ΔK_S, ΔI, ΔK_P, ΔY, ΔN, M and Dcy_B. Compared to the Hill-RFO model, with the exception of $\bar{Y}$, the parameters that do exhibit more complex dynamics and bifurcations are the same, with the addition of parameter K_G, which in the case of the logistic-RFO model, does present bifurcations.

The results obtained are very similar between the two models. From a qualitative standpoint the same parameters show the same type of bifurcations with roughly the

same shape. Additionally, the changes in stability in each solution occur in similar locations. The bifurcation pattern observed in the previous model also presents itself here. The bifurcation diagrams can be separated into three different regions: 1) As the given parameter is increased from zero, the system presents a single stable solution which describes a state of deadlock. In this deadlock state, state variables P_1 and P_2 have the exact same value (similarly S_1 and S_2 also have the same value). 2) A multistable region. The deadlock state reaches a branch point. At this point, the stability of the deadlock solution changes from stable to unstable (either one or two eigenvalues change from a negative real number to a complex number with a positive real component). In addition, from this branch point two new symmetric solutions emerge. In each one of these solutions, one of the sugar alternatives is "chosen" over the other. This commitment is reflected by the fact that in each one of these solutions, the value of either P_1 or P_2 is larger than the other. These solutions start as unstable solutions since the real part of one of the eigenvalues is positive (this behaves as a saddle in the P_1, P_2 space). Subsequently, these solutions become stable for the majority of their length before changing their stability once again back to that of a saddle. Finally, 3) a second branch point is encountered where the deadlock solution becomes stable again, and the two non-deadlock solutions collapse. Other features of these bifurcations that are shared by the two models are the hysteresis loop and the multistable region that accompanies it, as well as the fact that, depending on the parameter being varied, the average P_i value decreases or increases proportionally to the change of the parameter. Parameters V_B, K_G, $\bar{V_S}$, $\bar{V_P}$, $\bar{K}_{inh}$ and $\bar{N}$ have a proportional relation with P_i, whereas parameters $\bar{V_G}$, $\bar{K_S}$ and $\bar{K_P}$ have an inversely proportional relation with P_i. There are two sources from which the value of P_i changes: the induction term which is sensitive to S_i values as well as to inhibition by G (this term is also sensitive to changes in almost all parameters), and the term $A_i * V_{Pi}$ which is the basal level of activity and is only sensitive to changes to the parameters that are explicitly in it. As the P_i value decreases in response to a changing parameter value, the contribution of the induction term is diminished, until eventually the P_i value is mostly dominated by the basal activity. This allows us

to roughly classify the deadlock states into two categories: basal activity deadlock, and induced deadlock. Basal activity deadlock can be interpreted as a system that has not committed to either one of the alternatives, whereas the induced deadlock can be seen as the system exploiting both alternatives simultaneously.

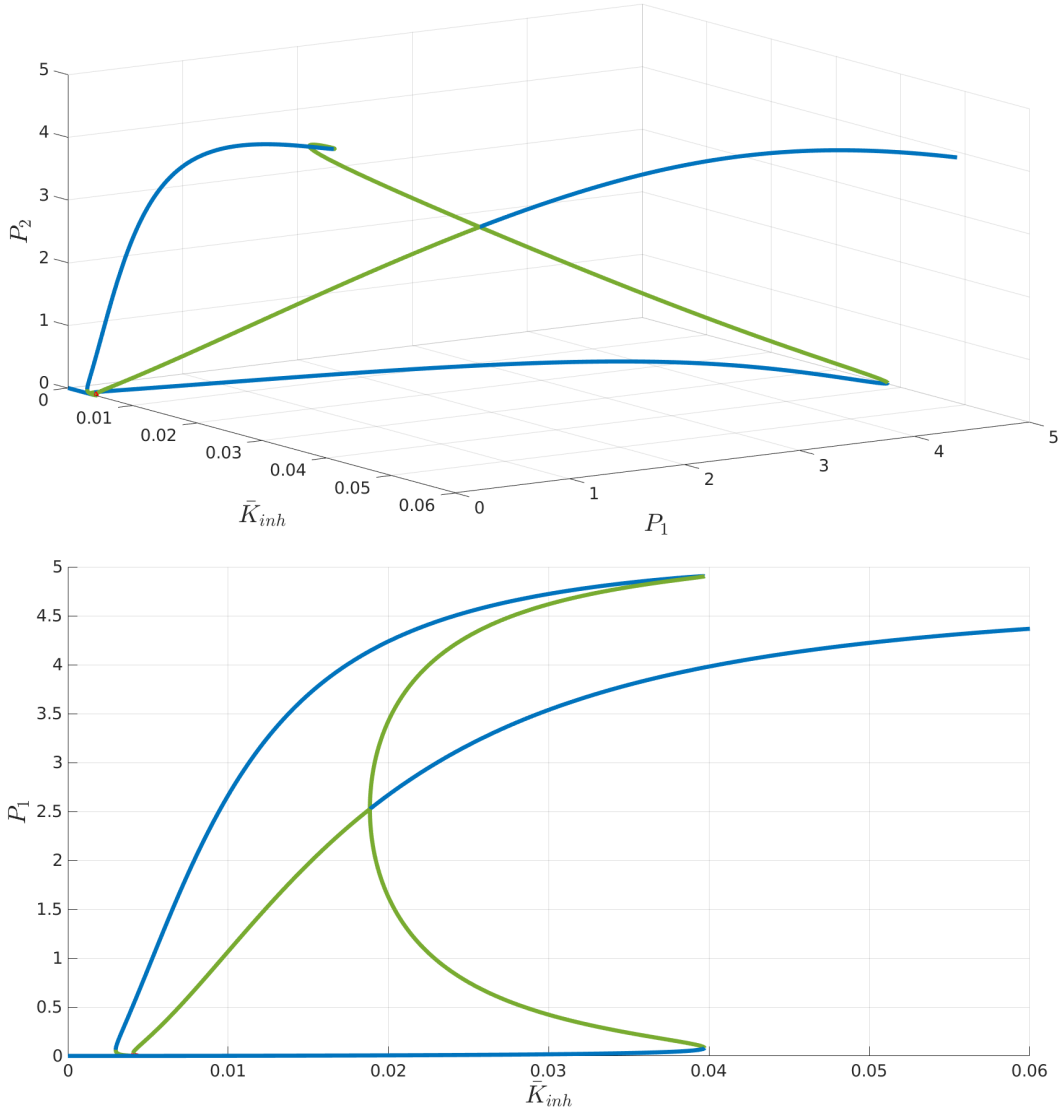


Figure 5.7: **Top**: Three-dimensional bifurcation diagram resulting from varying the $\bar{K}_{inh}$ parameter in the P_1, P_2 space. **Bottom**: Two-dimensional bifurcation diagram showing the same solution, focusing on the value of P_1 as $\bar{K}_{inh}$ is changed.

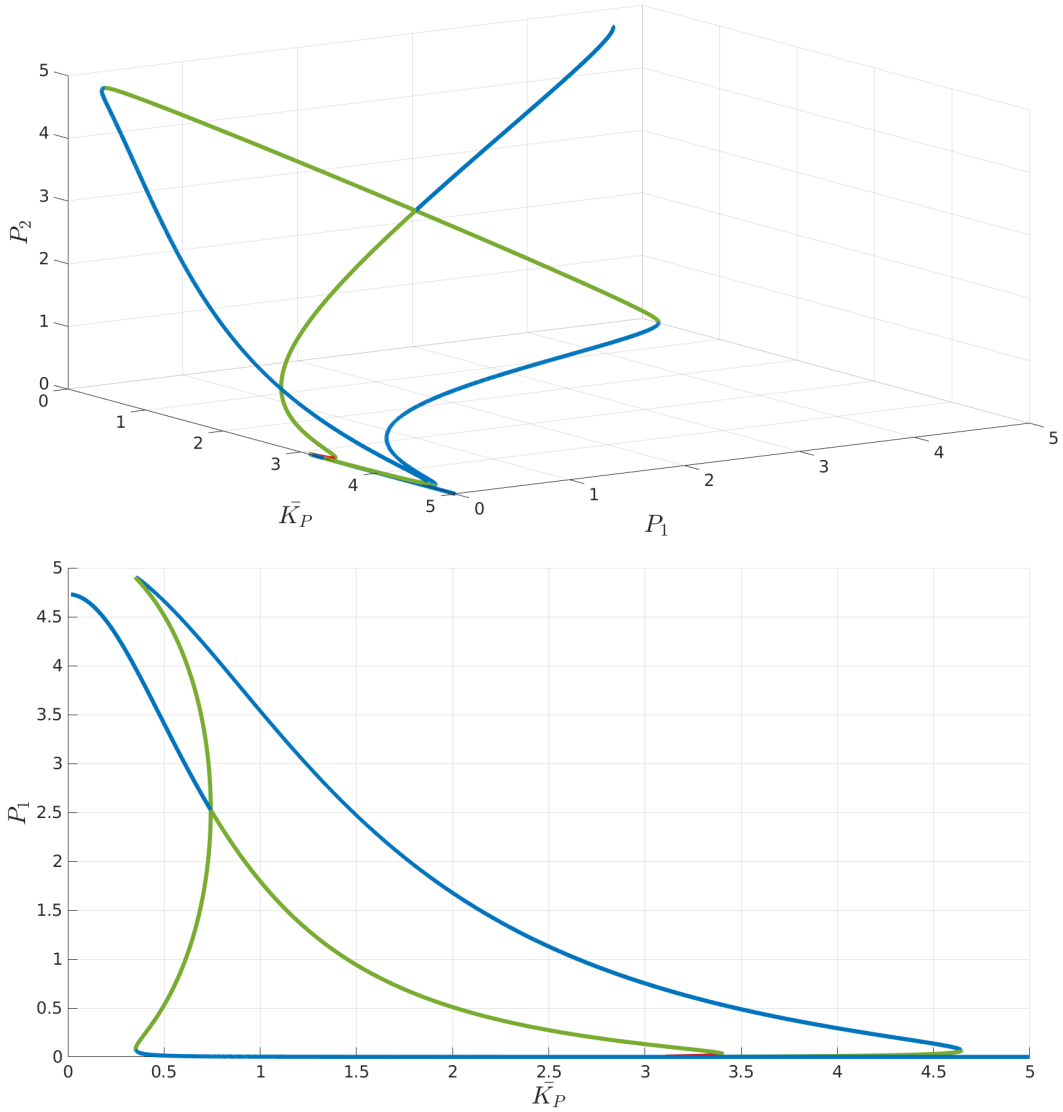


Figure 5.8: **Top**: Three-dimensional bifurcation diagram resulting from varying the $\bar{K}_P$ parameter in the P_1, P_2 space. **Bottom**: Two-dimensional bifurcation diagram showing the same solution, focusing on the value of P_1 as $\bar{K}_P$ is changed.

As mentioned before, all the parameters that present bifurcations are a different variation of this general pattern. Some parameters such as $\bar{V}_S$, $\bar{V}_P$ show a very pronounced S-shaped curve, as can be observed in figure 5.9. In contrast, parameters such as $\bar{K}_P$ and $\bar{K}_{inh}$ show a much less pronounced hysteresis loop, which takes place as the average P_i value diminishes. An important observation to make when looking at the two-dimensional diagram in this figure is that there are parameter values in which the system presents four different stable solutions it can take. Phase portrait 5.1f shows how a slice of the system's dynamics looks. It is immediately apparent that the basal-activity deadlock state is almost impossible to be reached by any

trajectory due to the fact that it is closely surrounded by two saddle solutions and an unstable solution. Any trajectory that is not already extremely close to this solution has a much higher chance to be pushed to one of the other three stable solutions.

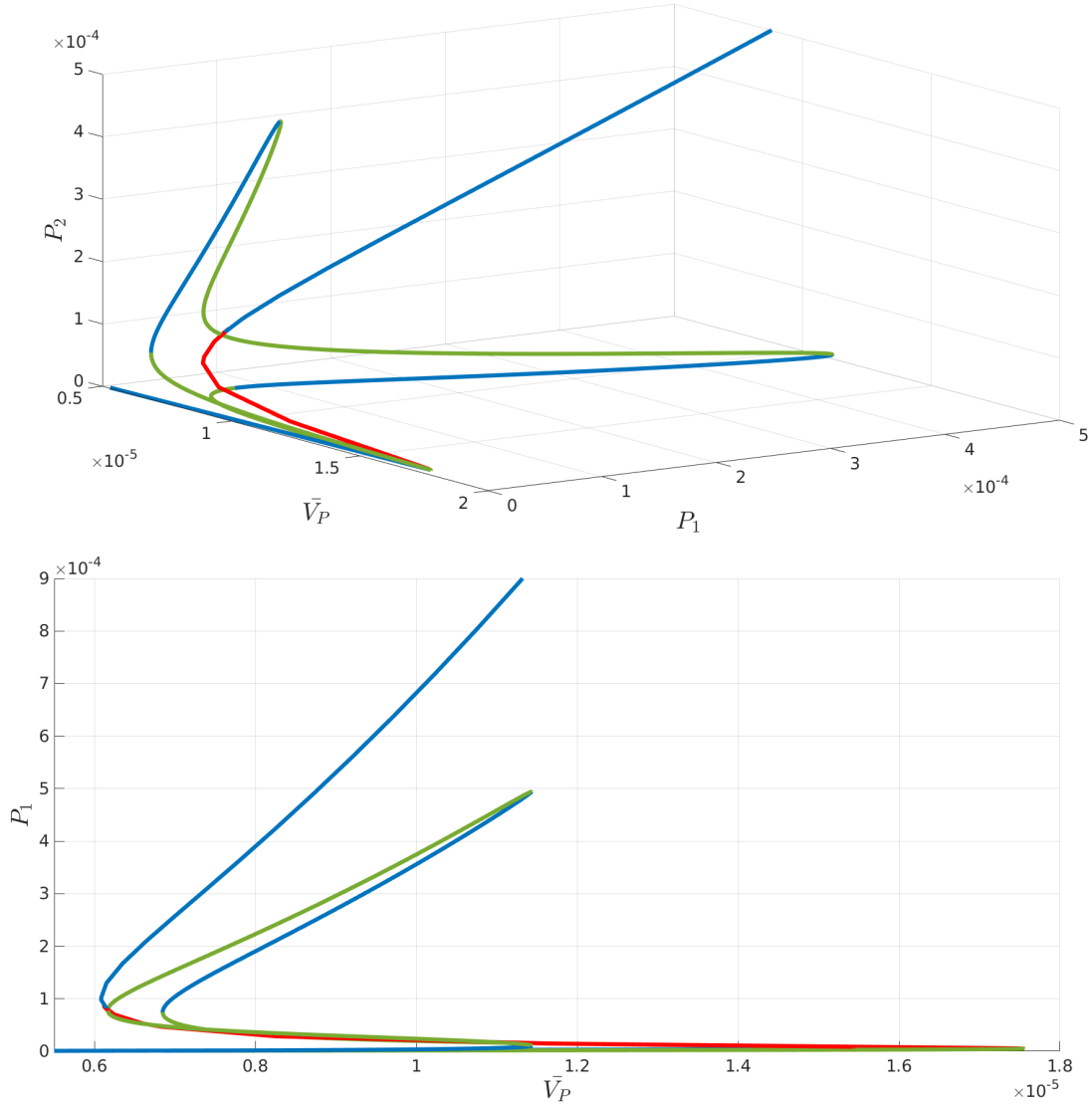


Figure 5.9: **Top**: Three-dimensional bifurcation diagram resulting from varying the $\bar{V}_P$ parameter in the P_1, P_2 space. **Bottom**: Two-dimensional bifurcation diagram showing the same solution, focusing on the value of P_1 as $\bar{V}_P$ is changed.

By varying parameter ΔK_{inh} the system presents a single solution with two hysteresis loops at the extremes, as can be seen in figure 5.10. This result in particular is very similar to that of the Hill-RFO model (figure 5.4). Another notable difference from this pattern is that of parameter V_B. Figure 5.11 shows how in this case, the system does not present any pitchfork bifurcations, but rather a single solution

which presents a couple of Hopf bifurcations within which oscillatory behaviour is observed. As is the case with the results from the Hill-RFO model, this behaviour takes place as V_B approaches zero, and the P_i value during these oscillations is relatively small compared to their steady-state value.

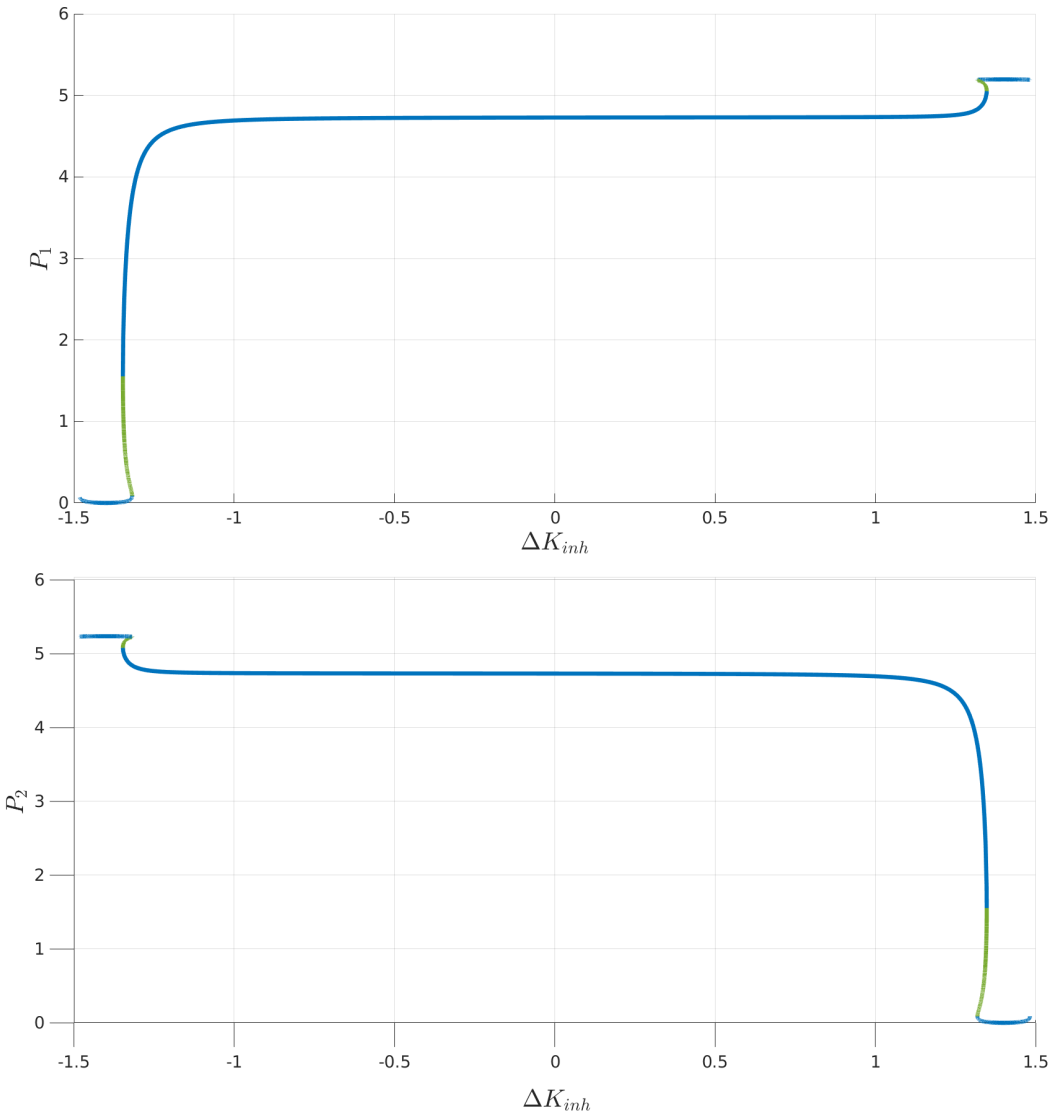


Figure 5.10: **Top**: Two-dimensional bifurcation diagram resulting from varying the ΔK_{inh} parameter in the P_1, P_2 space. **Bottom**: Two-dimensional bifurcation diagram showing the same solution, focusing on the value of P_1 as ΔK_{inh} is changed.

The periodic solutions are unstable since two eigenvalues have a positive real component. This unstable region is bookended by saddle regions, in which only one eigenvalue has a positive real part (the region at lower V_B is very small). In turn, these saddle regions are flanked by stable regions of the solution. It is unclear what this result could represent in terms of the biological system's behaviour. It is worth

noting that the change in stability and oscillations emerge at a low V_B (particularly considering that the literature value for V_B is 0.42). It is possible that a parameterisation that includes V_B values at which this behaviour is observed is not biologically plausible.

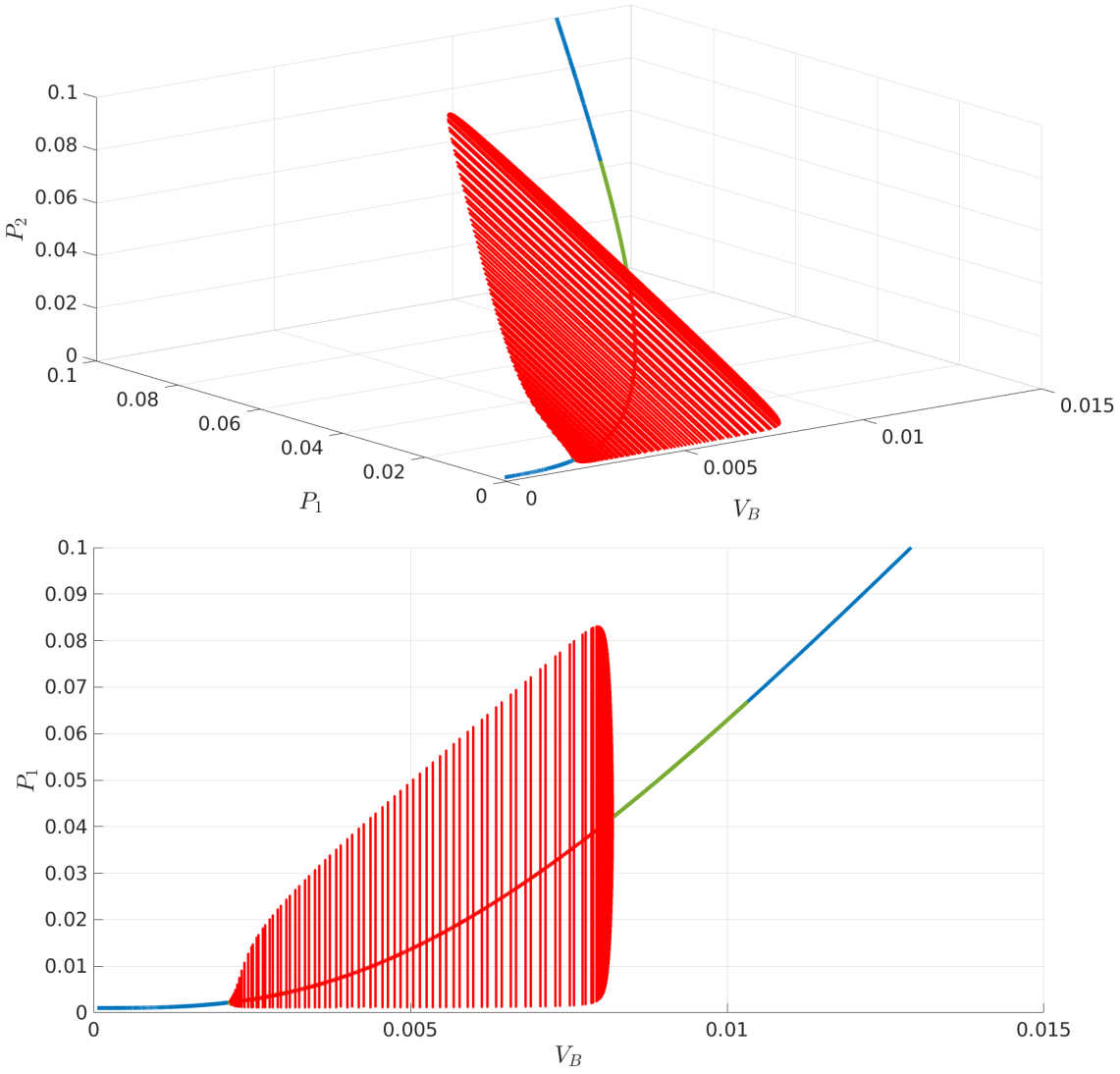


Figure 5.11: **Top**: Three-dimensional bifurcation diagram resulting from varying the V_B parameter in the P_1, P_2 space. **Bottom**: Two-dimensional bifurcation diagram showing the same solution, focusing on the value of P_1 as V_B is changed.

5.2.4 Final remarks on the equal-value alternatives analysis

In both of these models, the bifurcations were never found at parameter values close to the initial values (tables B.1 and B.3). Instead, most bifurcations occurred at values which differed in at least one order of magnitude from the initial values. In

the case of V_B and $\bar{K_P}$, these dynamics present themselves as the parameter value approaches zero. The dynamics observed in this section may occur at values that are improbable to be found in nature.

The dynamical systems analysis presented in this chapter was done under the assumption of the extracellular nutrient concentrations (N_i) remaining constant. This is why the, $\bar{N}$ bifurcation diagram (fig. 5.5 is a particular case because it is the only one in which this assumption is broken. In the case of parameter V_B, it is also worth mentioning that the oscillations shown by both models as a result of decreasing V_B values are only present within an unstable region of a deadlock solution at V_B values close to 0. The eigenvalues of these oscillations show that these are unstable since two of them remain positive.

Another particular case worth highlighting is that of parameter $\bar{Y}$ in the Hill-RFO model. Figure 5.12 shows that as the parameter $\bar{Y}$ is increased from zero, the first bifurcation takes place at ≈ 4. Parameter Y_i is the biomass yield of sugar i, also called yield coefficient. Biomass yield is defined as the grams of biomass produced per gram of substrate consumed (Hong, 1989). It is also generally accepted that the maximum biomass yield for glucose is approximately 0.5 grams of cell mass per gram of glucose consumed (Häggström et al., 2014; Van Hoek et al., 1998). Hence, it is unlikely for any substrate to surpass this, much less to have a value above 1. This means that our system's realistic behaviour in response to $\bar{Y}$ changes takes place before even the first bifurcation emerges, which is a stable deadlock solution. Similarly, the oscillatory behaviour in figure 5.11 could be taking place at a biologically unlikely parameter value range.

Lastly, it is essential to point out that changing the model's parameterisation affects the system's dynamics and not all features, such as bifurcations and stability changes, will remain present if the entire parameterisation is changed. Our work focused on the dynamical systems analysis of a single parameterisation per model. Further work is required to address questions such as how the model's dynamics

change and what features are preserved, changed, gained or lost, with different parameterisations.

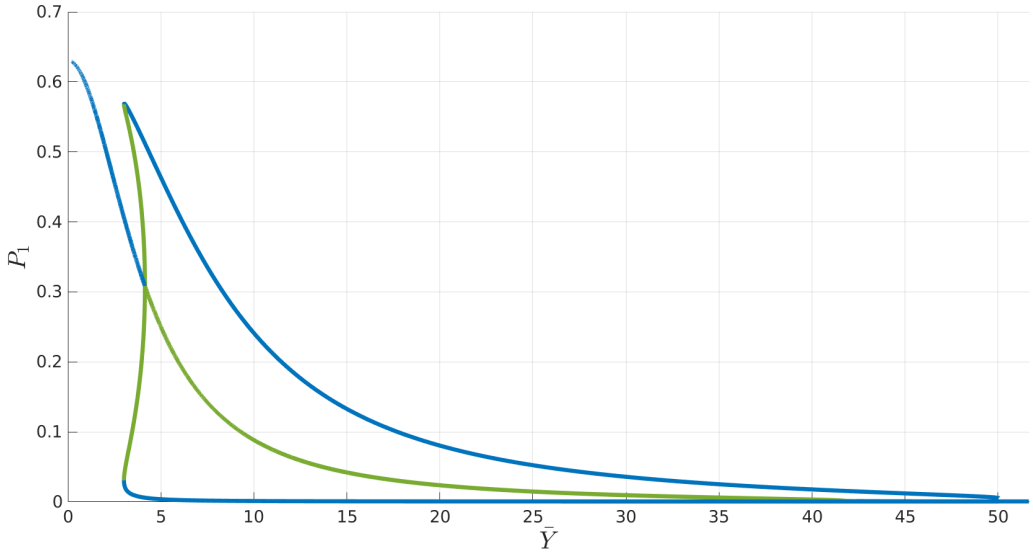


Figure 5.12: Two-dimensional bifurcation diagram showing the value of P_1 as $\bar{Y}$ is changed.

5.3 Unequal-value alternatives

The analysis of the unequal-value alternatives case consisted in analysing the models parameterised for a binary-sugar mixture environment. In this case, all four models were analysed for all different parameters. The parameterisation tested in each case was the parameterisation obtained through the experimental data fitting routine to data set PR1, in which the cellular culture was exposed to different concentrations of maltose and galactose simultaneously. Likewise, the initial values of the state variables used here are those specified in the previous chapter.

For the analysis, we followed the same procedure described in appendix B.3. In general, the analysis produced no bifurcations under any of the conditions tested for any parameter for any model. Almost all of the analyses showed a single stable solution with no critical points of note. All models, both in their steady-state and as the values of different parameters are varied, showed a preference for the alternative 1, a preference seen in the relatively large values of both S_1 and P_1, and the low

values of S_2 and P_2. A few parameters elicited a brief change in the stability of the solution as the value of the parameter being analysed approached a value of zero. Additionally, the models presented hysteretic behaviour in response to increasing values of parameter $\bar{Dcy}_P$. These are the results presented here.

5.3.1 Hill function biomass model: RSO induction term

Table B.5 shows the values of the parameters in terms of mean and difference. The original parameterisation was taken from the optimal parameterisation found by fitting to the experimental data in the previous chapter. The values used correspond to the best PR1 parameterisation for this model. Table B.6 shows the initial values of the state variables used, which were taken from the experimental data fit initial conditions, as well as the value of the state variables once the steady-state was reached.

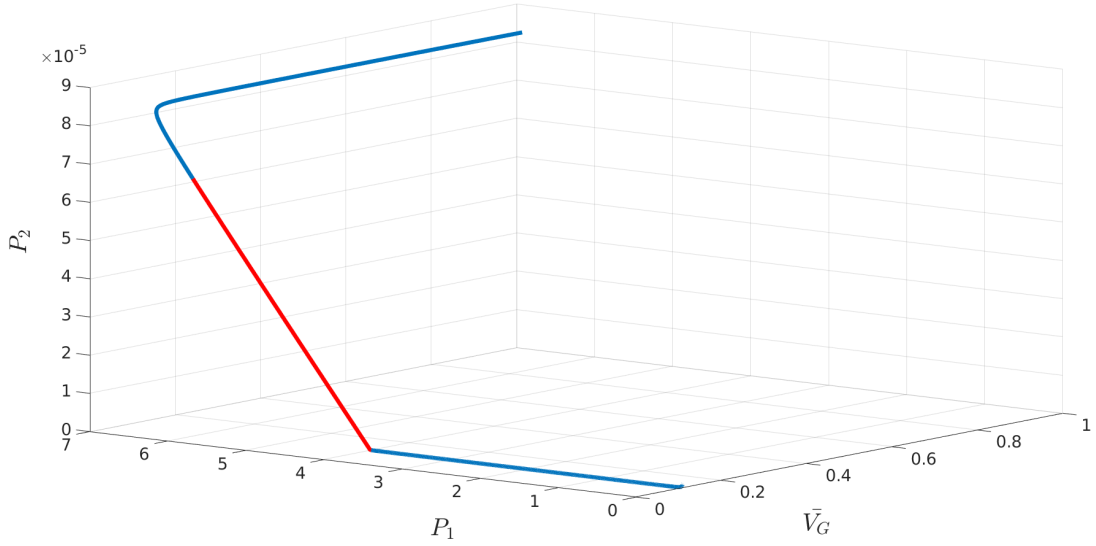


Figure 5.13: Three-dimensional diagram resulting from varying the $\bar{V}_G$ parameter in the P_1, P_2 space.

Two eigenvalues become complex conjugates with a positive real part, which in the P_1, P_2 space would signify an unstable focus or spiral. However, it must be stressed that the value region in which we encounter this behaviour corresponds to that in which, according to the equations (5.1), V_{G1} is close to 0. A similar solution to that

observed in figure 5.13 was obtained by varying the $\bar{Y}$ parameter. The solution is not only similar in the shape of the curve, but also the fact that this change in stability emerges as Y_1 is close to 0.

As the value of $\bar{V_G}$ decreases, the solution which starts with a high P_1 value relative to that of P_2, eventually reaches 0 for both P_1 and P_2. The unstable region covers most of the part of the solution in which both state variables decrease simultaneously. The unstable region regains stability once $P_2 = 0$ and, from there, the now stable solution continues to decrease the value of P_1 until this too is equal to 0. This drastic change in both state variables takes place over a relatively small range of $\bar{V_G}$ values, which is just at the threshold $V_1 \approx 0$.

The two points that encompass the unstable region are labelled as Hopf points by MatCont. However they do not lead to any oscillatory behaviour. For both parameters, $\bar{V_G}$ and $\bar{Y}$, as the value continues to decrease, the state variable G becomes negative, which is a direct consequence of V_{G1} and Y_1, respectively, becoming so small (according to equations (5.1)) that G decays at a higher rate that it is generated.

5.3.2 Hill function biomass model: RFO induction term

Table B.7 shows the values of the parameters in terms of mean and difference. The original parameterisation was taken from the optimal parameterisation found by fitting to the experimental data in the previous chapter. The values used correspond to the best PR1 parameterisation for this model. Table B.8 shows the initial values of the state variables used, which were taken from the experimental data fit initial conditions, as well as the value of the state variables once the steady-state was reached.

In this case, a full analysis of the model with this parameterisation was done. However, the steady-state yields has negative values for two state variables, S_2 and P_2 (table B.8). This is a biological impossibility, which renders any analysis derived

from it irrelevant, from the biological standpoint.

5.3.3 Logistic biomass equation model: RSO induction term

Table B.9 shows the values of the parameters in terms of mean and difference. The original parameterisation was taken from the optimal parameterisation found by fitting to the experimental data in the previous chapter. The values used correspond to the best PR1 parameterisation for this model. Table B.10 shows the initial values of the state variables used, which were taken from the experimental data fit initial conditions, as well as the value of the state variables once the steady-state was reached.

As shown in figure 5.14, with the parameterisation used for this analysis with this model, as parameter $\bar{V_G}$ is decreased, the system shows similar qualitative dynamics to previous models in response to changes to parameter $\bar{V_G}$. Additionally, the logistic-RSO model produces the same solution observed as parameter $\bar{V_G}$ is decreased in response to decrements in ΔV_G and ΔY. As either $\bar{V_G}$ or ΔV_G are varied, the unstable region of the solution presents itself where the value of V_{G1} approaches 0, whilst V_{G2} decreases or increases, respectively. Likewise, the unstable region of the solution obtained when parameter ΔY is decreased into the negative domain corresponds to the parameter Y_1 decreasing toward 0 and Y_2 increasing.

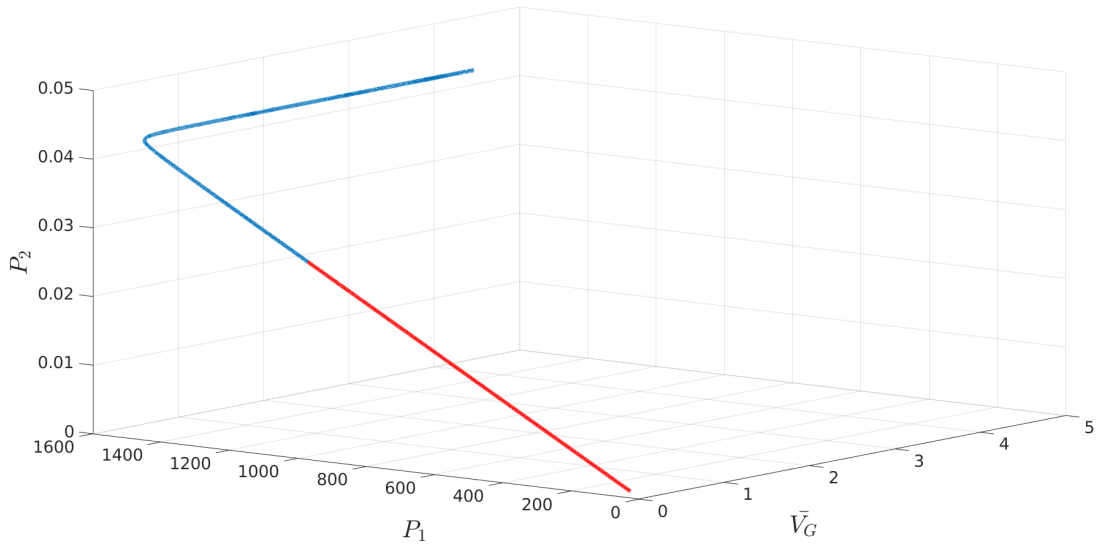


Figure 5.14: Three-dimensional bifurcation diagram resulting from varying the $\bar{V}_G$ parameter in the P_1, P_2 space.

5.3.4 Logistic biomass equation model: RFO induction term

Table B.11 shows the values of the parameters in terms of mean and difference. The original parameterisation was taken from the optimal parameterisation found by fitting to the experimental data in the previous chapter. The values used correspond to the best PR1 parameterisation for this model. More specifically, the parameterisation with upper bounds set to ten times the initial value was selected. Table B.12 shows the initial values of the state variables used, which were taken from the experimental data fit initial conditions, as well as the value of the state variables once the steady-state was reached.

The logistic-RFO model shows that as parameters ΔV_G, ΔY and $\bar{Y}$ decrease, there is a single solution which turns from stable to unstable in the P_1, P_2 space, as two eigenvalues turn from negative real numbers to complex conjugates with positive real components. Three-dimensional bifurcation diagrams for ΔV_G and $\bar{Y}$ are shown in figure 5.15. In this unstable state, the values of both P_1 and P_2 go to 0. The emergence of this stability change occurs as Y_1 and V_{G1} approach 0 according to equations (5.1). In the case of ΔV_G, ΔY, the decrement in these parameters not only

implies that V_{G1} and Y_1 decrease and approximate 0, but also that their counterparts increase. This is not the case when the mean $\bar{Y}$ decreases; in this case, both original parameters Y_1 and Y_2 decrease. However, even here, it is Y_1 that approaches 0.

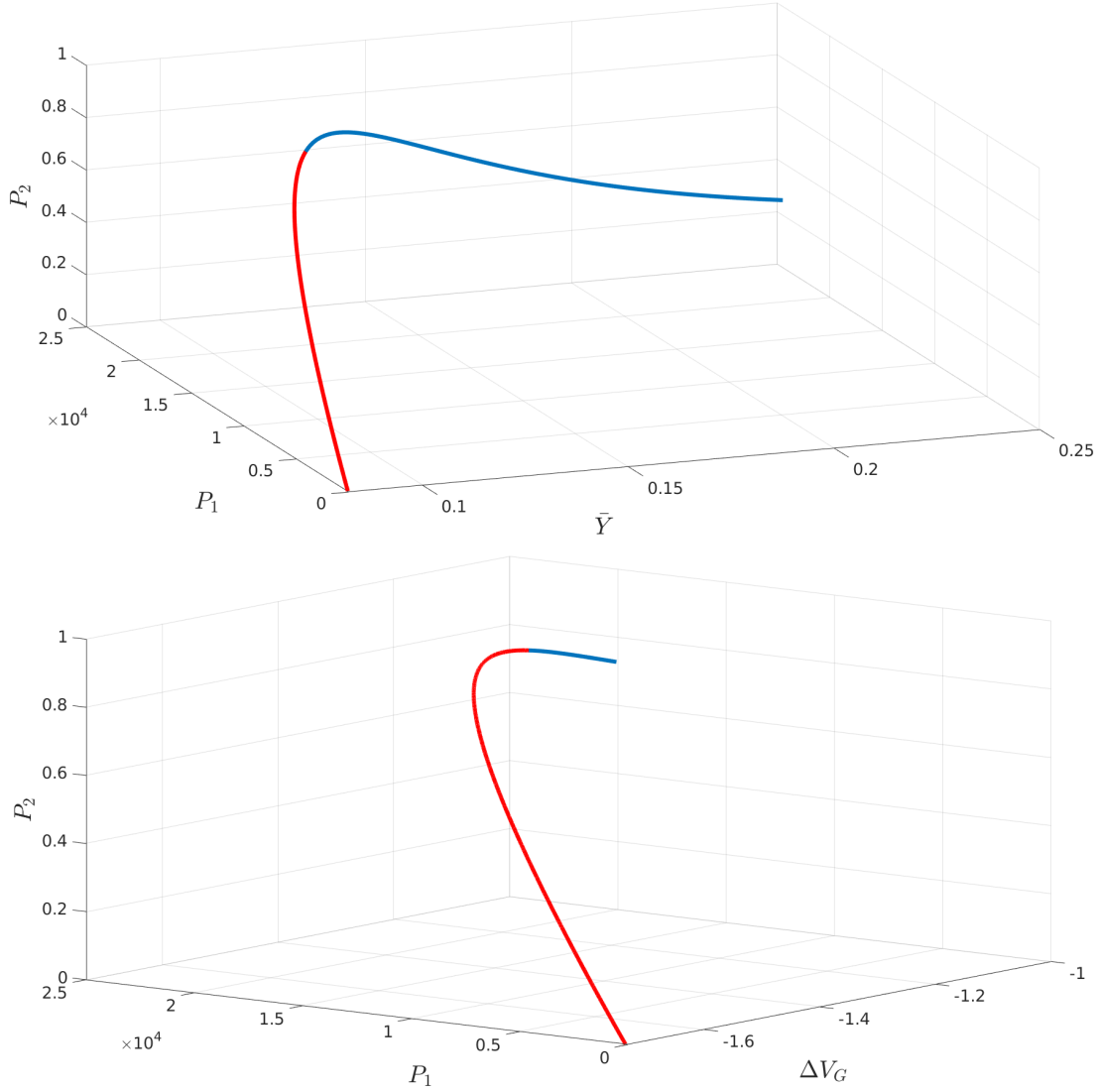


Figure 5.15: **Top**: Three-dimensional bifurcation diagram resulting from varying the $\bar{Y}$ parameter in the P_1, P_2 space. **Bottom**: Three-dimensional bifurcation diagram resulting from varying the ΔV_G parameter in the P_1, P_2 space.

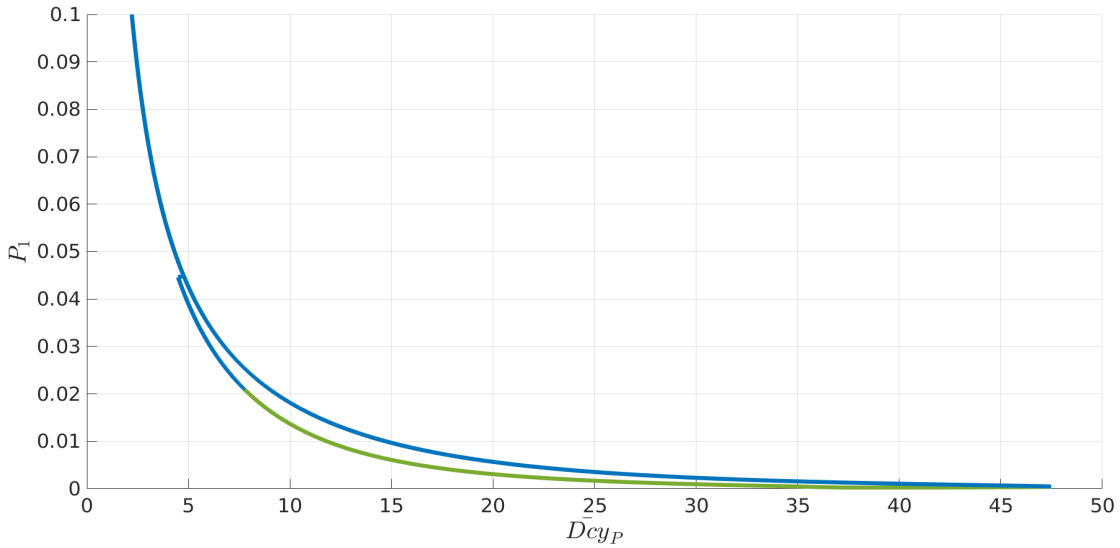
5.3.5 The effect of varying the decay parameter.

It can be noted that, under unequal-value alternatives conditions, all four models display a similar qualitative behaviour in response to changes of their parameters (without taking the Hill-RFO model into account due to its negatively-valued steady-state). In every model, the response elicited by the variation of most parameters is

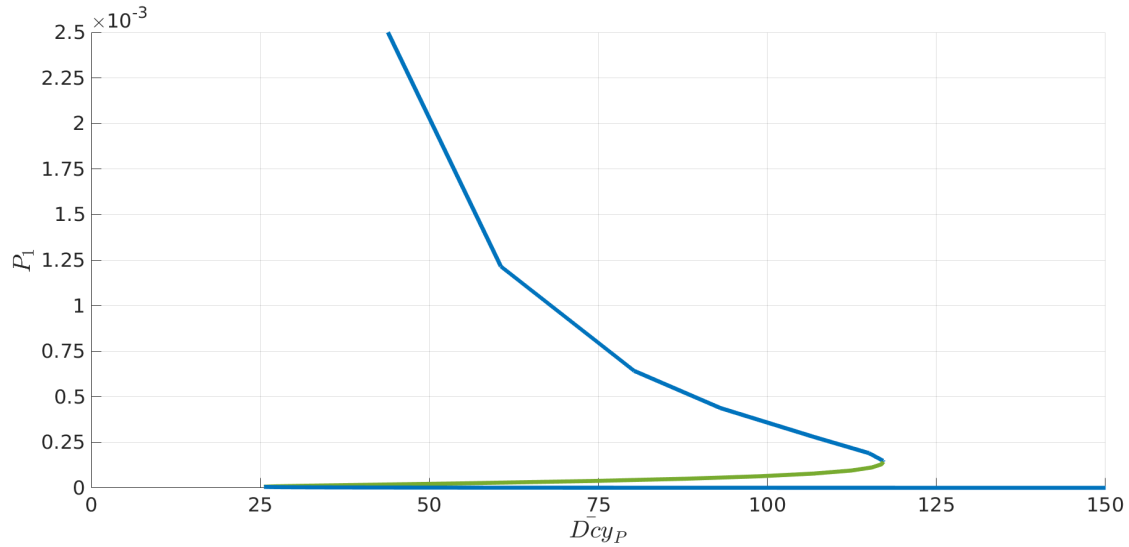
a single stable solution with a significant and noticeable bias for P_1. This is further discussed in subsection 5.5. The stability of this single solution changes from stable to unstable in very few cases, which are stated in each model's subsection. However, this pattern is broken in the case of parameter $D\bar{c}y_P$. All models display a characteristic hysteretic behaviour in response to increasing values of this parameter, as can be observed in figure 5.16. Particularly, diagrams 5.16b and 5.16c, corresponding to the logistic-RSO and logistic-RFO models respectively, display this phenomenon. As the $D\bar{c}y_P$ parameter is increased, the value of the stable solution decreases. A limit point changes the stability of the solution, from stable into a saddle, at which point the solution arcs back as the $D\bar{c}y_P$ diminishes. A second limit point is found, which restores the solution back to stable, at which point the solution continues to decrease the value of P_1, whilst $D\bar{c}y_P$ continues to increase. The logistic-RFO has a second instance of a hysteretic loop, which results in a region of $D\bar{c}y_P$ values in which the system can be in three different stable solutions, as can be noted in diagram 5.16c.

As $D\bar{c}y_P$ diminishes, and before reaching the first limit point, P_1 decreases whilst P_2 increases. Eventually, both P_1 and P_2 decrease their value in conjunction. Although all three models show this behaviour, only the logistic-RFO model breaks the P_1 bias, as in some sections of the hysteretic loops P_2 are larger than P_1. Finally, it is important to highlight the fact that all of these hysteretic behaviours come about at $D\bar{c}y_P$ values that are at least two and up to four orders of magnitude higher than the initial $D\bar{c}y_P$ value, the smallest difference being found in the logistic-RFO model.

(a) Hill-RSO



(b) Logistic-RSO



(c) Logistic-RFO

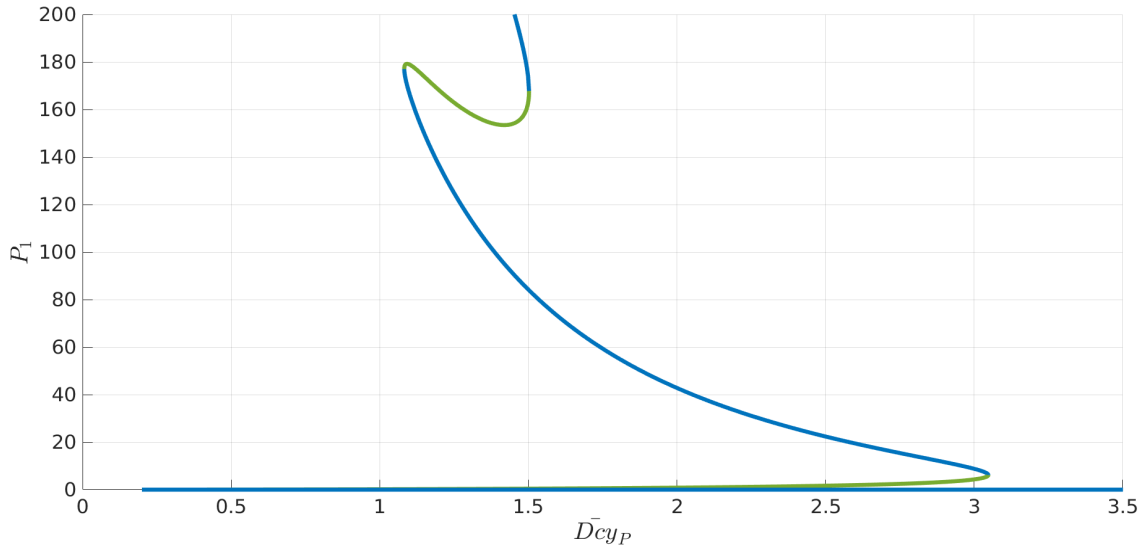


Figure 5.16: Two-dimensional bifurcation diagrams resulting from varying parameter $\bar{Dcy}_P$ in all models, excluding Hill-RFO.

5.3.6 Final remarks on the unequal-value alternatives analysis

Aside from the $D\bar{c}y_P$ analyses, for every model, except the Hill-RFO, a single bifurcation diagram is presented. In them, a single solution is observed as its' stability changes from stable to unstable at extreme parameter values. In all cases this occurs when one of the two parameters in the pair was approaching 0, e.g. in fig 5.14, the bifurcation appears as V_{G1} approaches 0. All bifurcation were found in response to changes of either $\bar{V_G}$, $\bar{Y}$, or ΔV_G. These parameters relate to the decay of S_i and its effects on G, Y_i being the biomass yield and V_{Gi} the rate of conversion. It is unlikely that such extreme values are found in a natural system. Moreover, the Hill-RFO model yielded negative S_2 and P_2 values at it's steady-state, which could suggest that the parameterisation used is not ideal for this type of analysis. This observation is particularly pertinent since the parameter set used is the result of a parameterisation effort to simulate a batch culture (discussed in Chapter 4), whereas the assumptions made to perform these dynamical systems analysis replicate the conditions of a continuous culture.

In general, for almost all the parameters tested, except those mentioned above, in every model analysed under unequal-value alternatives conditions, I did not find any bifurcations. When analysed, almost all parameters presented a single stable solution. A particularity of this unique solution is that P_1 is greater than P_2, generally speaking. Similarly, S_1 is also greater than S_2. This can be loosely interpreted as the system's preference for one sugar over the second sugar.

Due to the parameterisation used for all models, the alternative 1 is maltose. Aside from the parameterisation itself which, as mentioned was derived from the experimental data fitting results, described in Chapter 4, a potential reason for this observed bias is the choice of N_1 and N_2 values, which correspond to the 0.5% maltose and 0.25 % galactose condition in the experimental data. This difference in concentration could certainly tilt the system toward solutions that favour the consumption

of the alternative with the highest concentration.

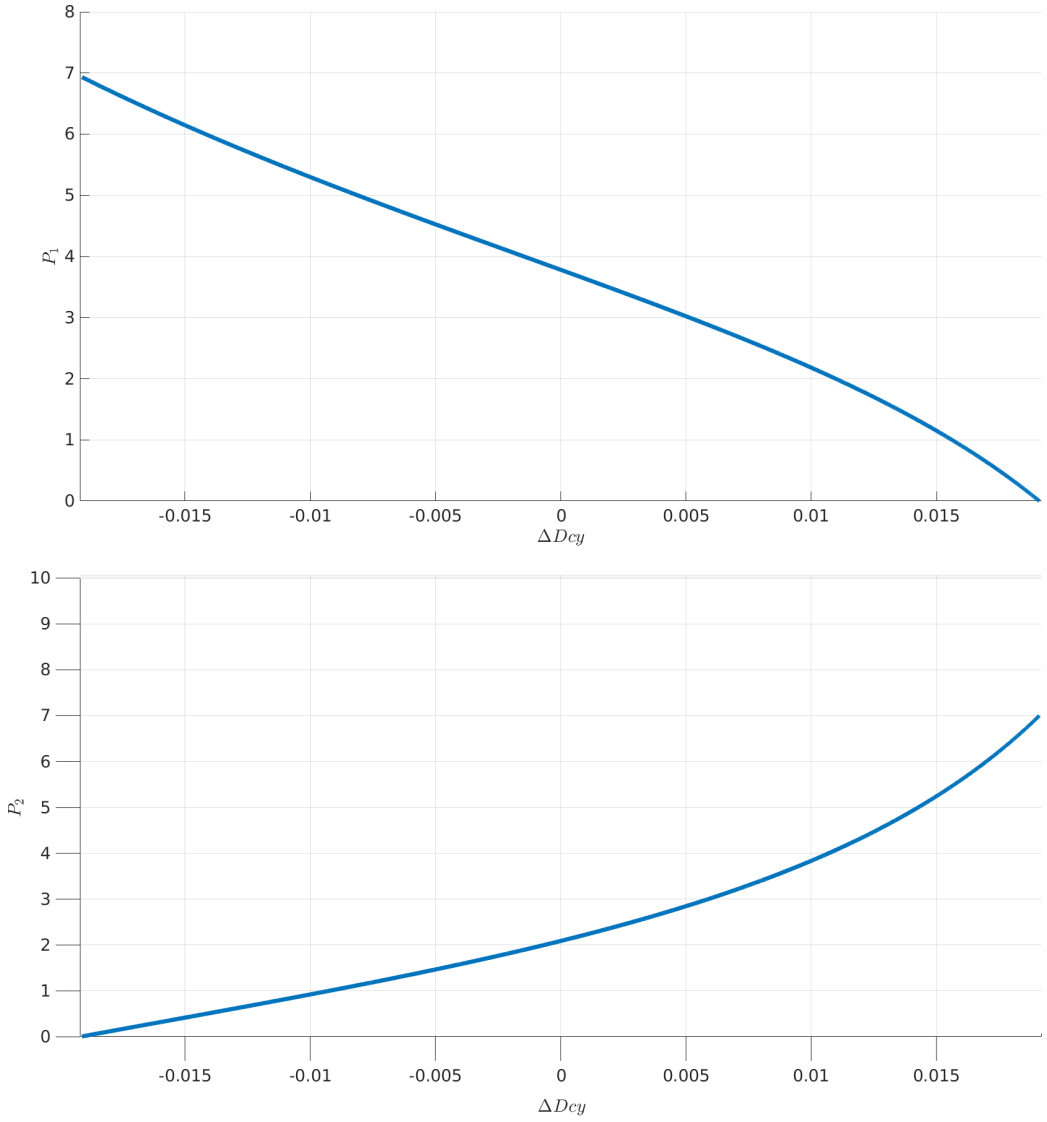


Figure 5.17: **Top**: Two-dimensional diagram showing how the value of P_1 is changed as a result of varying ΔDcy_P. **Bottom**: Two-dimensional diagram showing how the value of P_2 is changed as a result of varying ΔDcy_P. This is a result from model Hill-RSO. As ΔDcy_P changes from one extreme to the other, it can be observed that the preference of the system changes from one sugar to the other.

As discussed in Chapter 3, the models developed and analysed follow a global inhibition architecture. The effect of this inhibition is only present in the induction terms of state variables P_1 and P_2, by decreasing the activating effect of S_1 and S_2 proportional to the value of state variable G. However, with the parameterisations tested in the unequal-value alternatives case, our models did not present any change in stability nor drastic change in preference between P_1 and P_2 as a response of varying the strength of this inhibition. Similarly, the results of the analysis show

no change in response to varying values of cross-interaction between the two alternatives (mediated by parameters $\bar{I}$ and ΔI). We did, however, find changes in preference in response to varying values of ΔDcy_P, $\bar{Dcy}_P$ and ΔV_P. Respectively, these parameters relate to the rate of P_i decay, and the maximum rate of P_i synthesis. The diagrams in figure 5.17 clearly show this change of preference in model Hill-RSO under unequal-value alternative conditions. Although this diagram shows no change in stability, this result does suggest that decay dynamics can have a drastic effect on the system's alternative preference.

5.4 Saturated model

Given the conditions being simulated by the dynamical systems analysis, one can reduce the models by making assumptions of saturation wherever this assumption is appropriate. As explained in subsection 5.1, in order to perform the dynamical systems analysis the system needs to be able to reach a non-trivial steady-state, which is impossible under batch growth conditions due to the fact that the substrate concentration is always in decline. The only steady-state possible in such growth conditions is zero. For these reasons, the extracellular sugar concentrations N_i were considered to be constant. This setting resembles that of continuous growth conditions, in which the substrate present in the media is constantly being replenished as it is being consumed. Given this setting, one can assume that since N_i does not change, neither does S_i, nor G. The amount of biomass is equally considered to be kept as constant. Similarly, the Hill functions that populate these equations can be approximated to 1. The implementation of these assumptions results in the reduction of dimensionality in the model, from six equations to just two: P_1 and P_2. It becomes apparent that by doing this reduction, the structural differences between the Hill and the logistic models disappears. Therefore, the four models are reduced to two reduced models, differentiated by their induction terms. We analysed both these reduced version RFO and RSO models. The equations that define these mod-

els can be found in appendix B.4. Similarly, table B.13 shows the initial conditions used in their analysis.

Both of these models are analysed under unequal-value alternatives. In order to do this, I use the same parameterisations used for the analysis of the logistic-RSO (table B.9) and logistic-RFO (table B.11) models in the unequal-value alternatives case, for the saturated RSO and RFO model respectively. Also, I follow the same procedure as that described in appendix B.3. The initial conditions and the steady-state conditions that correspond to these parameterisations are found in table B.13. All parameters were analysed, and in all of them the system presented a single stable solution biased toward P_1. Similarly to the non-saturated models, for most of the parameters tested the system retains its bias toward P_1. The exceptions to this behaviour are parameter ΔDcy_P RSO saturated model and parameters $\bar{V_P}$, ΔV_P, $\bar{Dcy_P}$, and ΔDcy_P for the RFO saturated model.

Using an alternative parameterisation, specified in table B.14, the saturated RFO model showed multistability. In response to varying values of ΔDcy_P the system shows a hysteresis loop, similar to the effect discussed in subsection 5.3.5. In this case, it can be observed in figure 5.18 that the hysteresis loop occurs just as ΔDcy_P increases, going from the negative domain into the positive. The system presents bistability at the point (and in the vicinity of) where Dcy_{P1} and Dcy_{P2} are equal. It is important to note that, as can be seen in diagram 5.18, within the hysteretic region the stable solution coming from the left, and the saddle solution in between the two stable curves, are overlapping each other. This means that the saddle region of the curve would most likely push any trajectory down to the bottom stable solution. Finally, as the parameter ΔDcy_P is increased, the stable solution coming from the right initially presents a preference for P_1, as is the case with other parameters. However, right after the hysteresis loop, the second stable region of the solution reverts this preference, as it can be noticed that P_2 values increase whilst P_1 values decrease all the way down to zero. The initial conditions and the steady-state conditions used can be found in table B.15.

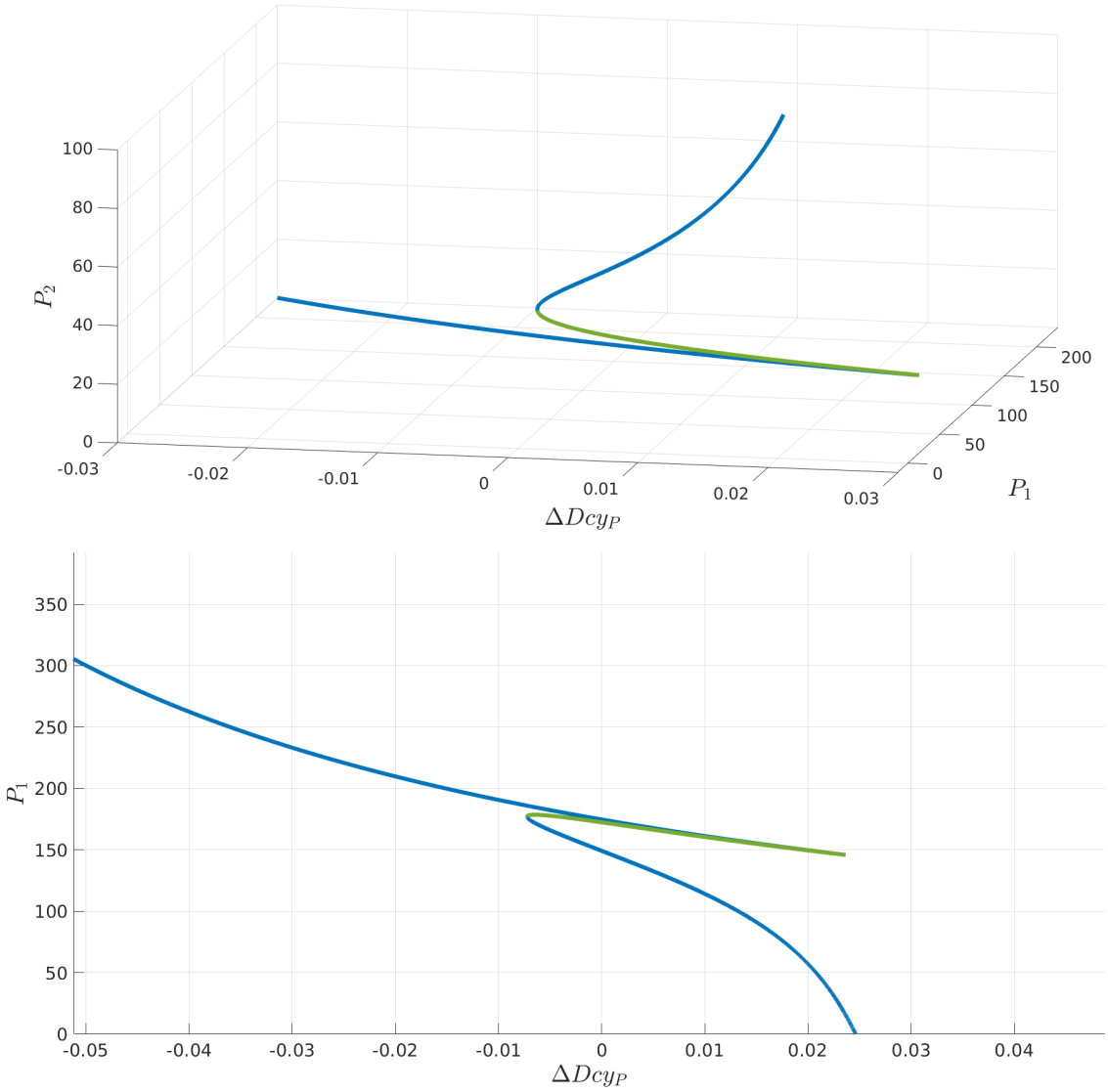


Figure 5.18: Bifurcation diagrams corresponding to parameter ΔDcy_P. **Top**: Three-dimensional view which shows the dynamics in the P_1, P_2 space. **Bottom** : Two-dimensional view focused on P_1. As ΔDcy_P changes from one extreme to the other, a hysteresis loop is encountered which also results in the preference of the system changing from one sugar to the other.

Comparing both the non-saturated models with the saturated models we find them qualitatively similar, although this simplification does alter the dynamics observed in the non-saturated models. The changes in stability found for some parameters in the non-saturated models are lost in the saturated models, as well as the hysteretic behaviour observed as parameter $D\bar{c}y_P$ is increased. Some form of hysteresis loop elicited by changes to the decay rates of P_1 and P_2 is recovered to some extent in the saturated models, albeit with a different parameterisation. In this case, it is present in the ΔDcy_P diagram.

5.5 Conclusions

In this chapter I have presented and discussed the results of the dynamical systems analysis performed on the models presented in Chapter 3. Different settings were investigated: the equal-value alternatives case, and unequal-value alternatives. The analysis of the saturated models or the analysis with a single-parameter difference can be considered variants of the unequal-value alternatives. For the unequal-value case we have used the optimal parameterisation obtained by fitting to the binary-sugar mixture experimental data set (PR1), discussed in Chapter 4.

As mentioned in section 5.3.6, the analysis under unequal-value alternatives conditions presented a single stable solution with no bifurcations in almost every parameter, with the notable exception of $\bar{V_G}$, $\bar{Y}$, $D\bar{cy}_P$, or ΔV_G, the results of which are discussed in their respective sections, and in general in section 5.3.6. Moreover, the change in preference without change in stability observed in response to varying parameter ΔDcy_P, $D\bar{cy}_P$ and ΔV_P, along with the lack of changes observed with parameters $K_i\bar{n}h$, ΔK_{inh}, $\bar{I}$ and ΔI, suggest that rate of P_i synthesis as well as its decay dynamics should be further explored as factors that affect the systems preferences, rather than inhibition or cross-induction, which do not seem to be an essential factor within our system. These observations are mirrored by the results of the saturated models' analysis.

Future iteration of these models could implement a form of inhibition to the decay dynamics. In yeast, the decay rate of membrane transporters is actively affected by inhibition pathways triggered by specific sugars (discussed in detail in Chapter 2). Enacting such modifications would allow the study of a different mechanism of inhibition (increasing membrane transporters' decay rate) that potentially, could affect the system's preference to a higher extent than the one currently explored in our models (reduced affinity for substrate), and reveal richer dynamics. Moreover, decay dynamics might not be the only mechanism worth expanding. Dilution in particular could potentially give rise to preference switching.

This research is inspired by the results of Pais et al. (2013). In their bifurcation analysis for the unequal-value case, their system presents multistability: as one of their parameters is increased, a single stable solution which starts close to a state of deadlock moves toward one of the alternatives, whilst from a critical point separated from this solution, a second solution emerges which veers toward the second, less valuable, alternative. In an attempt to exhibit such behaviour with these models, I analysed them under single-parameter difference conditions, the results of which are presented in appendix B.5. Under such conditions, a similar behaviour is presented by the Hill-RFO model (figures B.9 and B.10). The parameterisation used does not reflect any experimental data set. As such, it is unclear what experimental conditions or set of sugars could be differentiated from one another by just one parameter, making the relevance of this analytic approach (single-parameter difference) questionable.

When analysing the equal-value alternatives case, the models show an array of complex behaviours and bifurcations that lend themselves to an interpretation from a decision-making framework. As mentioned by Aidelberg et al. (2014), a cell that finds itself in a binary-sugar solution has four different behaviours it can adopt: 1) Consume both sugars simultaneously, 2) Consume one sugar exclusively, 3) Consume the second sugar exclusively, or 4) not consume any sugar. Other authors such as Beisel and Afroz (2015) and Koirala et al. (2016) have highlighted and explored these phenotypes as well, particularly in relation to the difference between a single cell versus the behaviour of the entire population. Most of the bifurcations found under an equal-value alternatives setting show that these options are available to the system. This is particularly evident in the phase portraits in figure 5.1. Although the solutions available to the system are dependent on the parameterisation of the model, they reflect the biological reality of a cell, or a cellular population.

Further efforts could be directed at not only different inhibition mechanism, but also different parameterisations. I decided to analyse these models using the same parameterisations achieved in Chapter 4. However, as mentioned before, Chapter 4 and

Chapter 5 should be considered separate studies that work with the same models. Analysing the models with other parameterisation would be beneficial, especially if the models were parameterised to simulate a continuous culture experiment, which is the experimental conditions we are replicating by the unequal-value alternative case by assuming that the extracellular nutrients are constant. The interest in analysing other parameterisations is reinforced by the results obtained by the unequal-value alternative case. Some state variables accumulated in vast amounts, whilst others remain at values close to zero, and in one case depicted in section 5.3.2, turned negative. It could be argued that this indicates the unsuitability of the parameterisation used for this analysis; Hence, the necessity to study other parameter sets.

Moreover, all our analyses were made by reparameterising the model to study the mean and the difference between parameters pairs rather than individual absolute values (section 5.1.1). As the mean or the difference of each pair of parameters changes, the value of both parameters is affected (e.g. as $\bar{K_P}$ is increased both K_{P1} and K_{P2} are increased in parallel). As mentioned before, this allows us to analyse our models in the context of value-sensitive decision-making. However, analysing the model's behaviour in response to changes in the absolute value of parameters could be worthwhile, as it could reveal different dynamics than the ones observed here.

Finally, although the equal-value alternatives case has no experimental corollary in the context of a yeast culture in a binary-mixture were the nutrients are homogenised in the liquid media, a decision problem between two equally valuable options can be experimentally realisable in other biological systems (Pais et al., 2013; Vogel et al., 2018; Dussutour et al., 2019). Our models are general enough so that they could be repurposed to study a different system in which the equal-value alternatives case represents a possible natural setting.

Chapter 6

Conclusions

Here I will present a summary of the results obtained by chapter in the thesis, as well as a brief description of future directions to possibly explore further.

6.1 Work presented

In this thesis, I present different mathematical models that explore the decision-making problem faced by a population of yeast cells that has two sugar sources available in its environment. This problem is divided into three related tasks: 1) model development, 2) fitting said model to experimental data, and 3) analysing the resulting parameterised model as a dynamical system. For this purpose, a literature review was conducted to properly understand the biology of yeast sugar preferences, as well as its inhibitory metabolism. This is presented in Chapter 2. Several models were developed and tested following the process presented and discussed in Chapter 3. From this process, a model with four variants was selected. In Chapter 4 these models were subsequently parameterised to fit to experimental data of yeast cells grown in maltose, galactose, and binary-sugar mixtures of maltose/galactose. In Chapter 5, the models are analysed as dynamical systems under different conditions. The results can be summarised as follows.

6.1.1 Model development

In Chapter 3, I describe the approach taken to the development of the model. In order to do this several models were built and tested, by following the mutual, feed-forward and pooled inhibition model architectures as they are presented by Bogacz et al. (2006). In accordance with the metabolic structure of yeast's carbon catabolic regulation network (as discussed in Chapter 2), the pooled inhibition architecture was chosen as a basis for the model. Our work is also motivated by the modelling approach presented by Pais et al. (2013), which focuses on concepts such as value-sensitivity and cross-inhibition, as well as their significance in decision-making dynamics. Both of these features are intrinsic properties of the cellular metabolism, their consumption dynamics, and in particular the carbon catabolite repression network in yeast (Horák, 2013).

One of our interests is to explore the effect of inhibition and alternative value in the decision-making process. Therefore, special interest was put into the induction terms and inhibition mechanism of the model, which takes cues from other modelling efforts (Narang et al., 1997; Chu et al., 2016). The resulting model, as well as its four variants, describe a system of eight ordinary differential equations. The equations describe the sugar concentration inside and outside of the cell, the lumped metabolic activity cognate to each sugar's consumption, the biomass, and the global inhibition signal. Unlike other ODE models similar to this, both sugars contribute to the activity of each metabolic pathway (with a lessened inducing effect over the competitor's metabolism). The resulting model makes use of two different equations for biomass, as well as two different induction terms (first- and second-order) in the equations that describe the metabolic activity, hence, four variants. These models are symmetric and non-specific such that the differences between the two alternatives are expressed through parameter values rather than the equations' structure. We conclude that in order for the model to appropriately reproduce binary-sugar mixture data, it requires a common resource term dependent on biomass as well

as a positive inducing effect of both sugar alternatives on the activation of both metabolic pathways.

6.1.2 Experimental data fit

In Chapter 4, the resulting models are parameterised in order to fit various experimental data sets, comprising cultures growing on single sugars (maltose and galactose), as well as in a binary-sugar mixture (maltose/galactose). In order to do this, I take parameter values from the literature as a starting point, and after some manual tuning to these values, the model is parameterised through a fitting routine that employs a weighted version of the RMSE formula as an objective function. A wide array of weighting factors was used on all model variants.

Out of the four variant models tested, the two variants that implement a first-order induction term, the logistic-RFO and the Hill-RFO models, present overall lower RMSE scores, indicating smaller errors between data and model, under all experimental conditions (single sugar and binary-sugar mixture). For both single sugar data sets, a closer fit (low RMSE score) is achieved by the logistic-RFO model, whereas the Hill-RFO model achieves the best fit to the binary-sugar mixture data set, as well as the best fit to all data sets with a single parameterisation. As mentioned in Chapter 4, the curves that represent metabolic activity in the binary-sugar mixture data display a particular group formation. With three different concentrations of both maltose and galactose, there are nine combinations. The metabolic activity shows that these nine curves arrange themselves in three groups of three curves each, the position of the groups is proportional to the concentration of their corresponding sugar, whilst the order of the curves within the groups is proportional to the concentration of their competing sugar. With this in mind, a visual examination shows that the model with the lowest RMSE score struggles to reproduce this grouping pattern. Actually, both the logistic-RFO and the Hill-RFO model variants present issues in this regard, whereas both second-order induction term model vari-

ants, logistic-RSO and Hill-RSO, are able to qualitatively reproduce this pattern, despite having a larger RMSE score. The results from both the logistic-RSO and the Hill-RSO models can be found in appendix A. Although, qualitatively these models seem able to reproduce the patterns present in the data, the model does not appear to be capable of refined numerical prediction. Additionally, it can be observed that galactose metabolism exhibits decay dynamics that the model is unable to capture. From these results it can be concluded that the logistic-RFO model fits the single sugar data better (lower RMSE score), and, whilst the Hill-RFO model yields a lower RMSE score for the binary-sugar mixture data, both of the second-order induction models yield a better group separation. The model is successful in reproducing the qualitative behaviour of all three experimental data sets with a single parameter set. The lowest scoring model is the Hill-RFO. Model Hill-RSO, although scoring slightly higher than its first-order counterpart, fits the single sugar data sets with similar goodness, and maintains the group separation in the binary-sugar mixture data.

The best fits still do not capture all the details observed in the experimental data. As mentioned in section 4.3.4, the fits' goodness could be improved by including in the model elements that were left out: separation of timescales, fluorophore dynamics, terms for dilution (similar to Narang et al. (1997)). Moreover, the equations that define our system are not detailed representations of the specific pathways involved in the internalisation and metabolism of the sugar alternatives. This lack of modelling detail, whilst it simplifies the model and its analysis, hinders the fit's performance. It is reasonable to expect that a more detailed model, although potentially harder to parameterise and analyse, would better replicate the experimental data.

6.1.3 Dynamical systems analysis

In Chapter 5, I perform dynamical systems analysis on the models. For this, two decision problems were recreated: 1) Equal-value alternatives, and 2) Unequal-value

alternatives. In both scenarios, the system's dynamics are examined in a situation in which sugar sources are available to the cell population, and the system finds itself in one of four different states: 1) the population consumes both sugars simultaneously, 2) and 3) it consumes either one of the sugar alternatives exclusively, and 4) it consumes neither. It is important to highlight that in order to perform dynamical systems analysis, it is assumed that the extracellular sugar concentration remains constant. With this assumption we guarantee that the systems will reach a non-zero steady-state, which is necessary to perform the dynamical systems analysis. These analyses should be considered separate from the experimental data fits performed in Chapter 4.

For this analysis, the models are reparameterised in such a way that the parameters are described in terms of their mean value, and the difference between them, such as is done in Pais et al. (2013). We are particularly interested in the dynamical changes elicited by parameters within the induction term, such as K_{Pi} that describes the affinity for the sugars; V_{Pi} which is the maximum P_i synthesis rate; I_i which controls the induction strength of the competing sugar, and K_{inhi} which describes the strength of the inhibition signal. Other parameters that we are interested in are those that modulate each sugar's contribution to the global inhibition signal, such as V_{Gi} and Y_i.

In the equal-value case, the Hill-RFO and the logistic RFO models are parameterised in such a way that the two alternatives represent the same sugar. We used parameters that describe galactose. The experimental setup fitted in Chapter 4 is not represented by this scenario. However, this case is similar to experimental setups used in the study of slime mould by Vogel et al. (2018), where the alternatives are spatially separated from one another. We find that parameters of interest do elicit bifurcations that change the system's preferences. Changes to parameters that modulate the inhibition strength (K_{inhi}) and parameters that are closely related to the value of the alternatives (K_{Pi}, V_{Pi} and Y_i) bring about these changes and allow for multistability. Similar to the symmetry-restoring bifurcations found by Zabzina

et al. (2014) in their analysis of biomass allocation in slime mould, as the mean of the parameters of interest is varied, two bifurcations appear: one breaking the symmetry of the initial deadlock state, and subsequently, a second bifurcation which restores the deadlock state. In these bifurcation diagrams, we find the four different consumption states described above. It is notable that, for all parameters analysed, the bifurcation region presents itself in a value range that does not coincide with the parameter values used to fit the experimental data in Chapter 4, and in some cases, it differs by several orders of magnitude. The analysis of both inhibition strength and affinity, K_{inhi} and K_{Pi} respectively, reveals that the bifurcation region occurs at high K_{Pi} values, which are inversely proportional to affinity. The lower the affinity is, the higher the concentration of sugar is needed to induce a metabolic response. Similarly, the bifurcation region is present at low K_{inhi} values. K_{inhi} values are inversely proportional to the inhibition strength. Together, these results suggest that under the equal-value conditions, the multistability prompted by the bifurcation region is present when the overall output of the induction term is lower than the parameter values used in Chapter 4, be it by increasing the strength of inhibition, or by decreasing the sugar's inducing effects. We observe a similar outcome when the maximum rate is decreased significantly.

In the unequal-value alternatives case, the system is parameterised with the parameters that reproduce the binary-sugar mixture data. Each one of the four models was analysed twice; once in its regular form, and a second time in a saturated form, which reduced the models' dimensionality from eight, down to two. In this case, we used the same parameterisation obtained in Chapter 4 from the binary-sugar mixture experimental data fit. We find that changes to inhibition and the parameters associated with the alternative's value in the manner implemented by our models elicited no bifurcations. In fact, from these analyses we see that in response to most parameters, the systems are mostly consistent in their output, which is a single stable solution with a bias for maltose. In the equal-value case, the single stable solution is described as a deadlock state in which $P_1 = P_2$. In this case,

however, the single solution shows a preference for the maltose, which is expressed by $P_1 \gg P_2$ and $S_1 \gg S_2$. However, the P_1/P_2 ratio does change in response to parameter variation, in both the saturated and unsaturated forms of the models. This bias is relatively robust, and there are few cases in which it is actually reversed, meaning that $P_2 \gg P_1$. For instance, In the logistic-RFO model (both saturated and not), changes to the mean and difference between V_{P1} and V_{P2} changes the bias from maltose, to galactose as V_{P1} decreases in value towards 0. Furthermore, in the unsaturated analysis, as parameter $\bar{Dcy}_P$, which is the mean value of the P_i decay rates, is varied, multistability is found as a consequence of a hysteresis loop. On the other hand, whilst no such behaviour was found in the saturated version, an alternative parameterisation shows a similar hysteresis loop as the difference between the two parameters is varied (ΔDcy_P). In both cases, this behaviour is present in parameter value ranges which are distinct from those obtained in Chapter 4. In the case of the $\bar{Dcy}_P$ parameter, the multistability is present at parameter values that are orders of magnitude larger than the initial $\bar{Dcy}_P$ value.

The presence of such behaviour in the analysis of the P_i decay rate could indicate that, by changing the decay rate of certain proteins, the system could exhibit multistability. This is supported by the literature that indicates that, in yeast, sugars inhibit each other through different mechanisms, including degradation of membrane transporters (Brondijk et al., 2001; Horak and Wolf, 1997; Horak et al., 2002; Jiang et al., 2000; Medintz et al., 1996). In Chapter 3, it was mentioned that, despite this, in the pursuit of simplicity the effects of inhibition in the models developed only affect the induction of P_i by diminishing the inducing effects of the sugars. In our models, inhibition does not target P_i decay rate. It is possible that, by restructuring the models so that the effects of inhibition affect the P_i decay rates, multistability would arise as a direct result of inhibitory mechanisms. Finally, with the parameters used in the unequal-value case, we do not see multistability emerging as a result of varying the values of parameters of interest such as inhibition strength, or sugar affinity and maximum rate of synthesis which can be seen as indicators of the

alternative's value.

In this chapter, I used the parameterisation obtained y the experimental data fit in Chapter 4. However, it is possible that a different set of parameters which describe a continuous-culture experiment would yield different results, particularly in the unequal-value case. Similarly, although the equal-value alternatives case does not correspond to a biological system in the context of a culture in a liquid media with homogenised sugar concentrations, this decision problem can be found in other biological systems, and its subsequent analysis would entail a different parameterisation.

Work by Pais et al. (2013) has explored and interpreted the multistability brought about by bifurcations, as a system making an all-or-nothing decision. In the context of house-hunting honeybees there is a concept called *decision thresholds*. If the number of bees committed to either one of the alternatives crosses this threshold, the decision is taken and the entire colony moves into this location to build their new nest. However, the interpretation of our results is rather different and follows more in line with that of Zabzina et al. (2014). Our system and its analysis do not have decision thresholds, nor do they describe an all-or-nothing decision; instead, the solutions to the system found in bifurcation analysis describe the state variables' values and the consumption state of the system. For instance, a solution in which $P_1 \gg P_2$ can be interpreted as preference or bias for maltose (S_1 will similarly be much larger than S_2). Loosely speaking, one could say that the system is "choosing" option 1 over option 2, even though both alternatives are being exploited at different rates. This allows for gradual preference changes as the P_1 and P_2 values change in response to varying certain parameters. Whereas other systems such as house hunting honeybees or foraging ants, the decisions are all-or-nothing, in the case of the yeast culture system we have modelled and studied, the decision is reflected in the ratio in which the sugars available to the culture are consumed. To conclude this point, the system we have studied here does not make decisions by "choosing" an alternative over the other unambiguously, but rather it "chooses" a phenotype,

which determines the ratio at which both alternatives are being consumed.

6.2 Future work

In this section, I will briefly touch upon two potential paths through which the current work presented in this thesis could be expanded, and its subject further explored.

6.2.1 Single-cell modelling

The experimental data sets used to parameterised the models correspond to studies made in yeast cultures at the populational level. Recent studies have highlighted the importance of carrying out single-cell investigations. The work of Beisel and Afroz (2015); Koirala et al. (2016); Solopova et al. (2013) and Ozbudak et al. (2004a) show that the behaviour and consumption dynamics observed in populations do not necessarily mirror the behaviour of individual cells. Koirala et al. (2016) results show that, when cultured in a binary-sugar mixture, individual cells can arrange themselves in four different subpopulations: one that consumes neither, one that consumes both sugars, and two subpopulations each committed to one of the two sugars available. This individual behaviour might be lost in a population-level analysis. Solopova et al. (2013) point to the fact that the concept of diauxic shift can be reinterpreted as the individual behaviour is examined. The traditional interpretation of diauxic shift is that a population of cells in a binary-sugar mixture will first consume their preferred sugar exclusively until depletion, before consuming non-preferred sugars. They do this by changing their molecular machinery, a process that creates a lag in growth whilst they change their metabolism. However, Solopova et al. (2013) suggest that individual cells vary in their response to the depletion of their preferred sugar. Their results indicate that, only a subpopulation of cells continues to divide on the second and less-preferred sugar, whilst the other remains inactive. Whilst the

overall population dynamics are not in question, the single-cell experimental data could prove useful in delivering more depth to the analysis of microbial consumption dynamics at the level of the individual, rather than the population. Our models are in principle, capable of describing an individual cell's behaviour, although they would need the implementation of a reproductive mechanism such as mitosis or in the case of yeast, budding. Analysing our models with single-cell data could allow us to validate this and contrast the results with population-level behaviour. Our models could be used for this purpose with some modifications, such as implementing cellular division or budding. A similar approach to that of Chu and Barnes (2016) could be particularly useful for this purpose.

6.2.2 Stochastic simulations

The models and their subsequent analysis presented in this thesis, have been entirely deterministic, which leaves the analysis of stochastic behaviour unexplored. Stochasticity is an intrinsic part of the cell's environment and their internal processes and further studies should address how the dynamics change when noise is added to the system (Perkins and Swain, 2009). Some stochastic simulations have been made and included in appendix C figure C.1. These analyses add a degree of noise to state variables B and P_i. However, the strength of these noise is arbitrary, but further studies could address other mechanisms for the study of stochasticity in the system such as by using the Gillespie algorithm, or with the master equation. Such approaches would provide a path to study the effects of the internal noise of the system.

Bibliography

J. H. Abel, B. Drawert, A. Hellander, and L. R. Petzold. GillesPy: A Python Package for Stochastic Model Building and Simulation. *IEEE Life Sciences Letters*, 2(3): 35–38, 2017. ISSN 2332-7685. doi: 10.1109/lls.2017.2652448.

M. Ackermann. A functional perspective on phenotypic heterogeneity in microorganisms. *Nature Publishing Group*, 13:497 – 508, 2015. doi: 10.1038/nrmicro3491.

D. G. Adams. Heterocyst formation in cyanobacteria. *Current Opinion in Microbiology*, 3(6):618–624, 2000. ISSN 13695274. doi: 10.1016/S1369-5274(00)00150-8.

D. Ahuatzi, A. Riera, R. Peláez, P. Herrero, and F. Moreno. Hxk2 Regulates the Phosphorylation State of Mig1 and Therefore Its Nucleocytoplasmic Distribution. *The Journal of Biological Chemistry*, 282(7):4485–4493, 2006. doi: 10.1074/jbc.M606854200.

G. Aidelberg, B. D. Towbin, D. Rothschild, E. Dekel, A. Bren, and U. Alon. Hierarchy of non-glucose sugars in Escherichia coli. 2014. doi: 10.1186/s12918-014-0133-z.

U. Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–61, 2007a. ISSN 1471-0056. doi: 10.1038/nrg2102.

U. Alon. *An Introduction to Systems Biology*. Taylor and Francis Group, 2007b.

U. Alon. Simplicity in biology. *Nature*, 446, 2007c.

R. P. Andrade, P. Kötter, K. D. Entian, and M. Casal. Multiple transcripts regulate glucose-triggered mRNA decay of the lactate transporter JEN1 from Saccha-

romyces cerevisiae. *Biochemical and Biophysical Research Communications*, 332 (1):254–262, 2005. ISSN 0006291X. doi: 10.1016/j.bbrc.2005.04.119.

B. W. Andrews and P. A. Iglesias. An information-theoretic characterization of the optimal gradient sensing response of cells. *PLoS Computational Biology*, 3(8): 1489–1497, 2007. ISSN 1553734X. doi: 10.1371/journal.pcbi.0030153.

Aristotle. Book 2, 2006. URL `http://classics.mit.edu/Aristotle/heavens.2.ii.html`. Accessed: 11-09-2019.

H. R. Arkes and P. Ayton. The sunk cost and concorde effects: Are humans less rational than lower animals? *Psychological Bulletin*, 125(5):591–600, 1999. ISSN 00332909. doi: 10.1037/0033-2909.125.5.591.

G. Balázsi, A. van Oudenaarden, and J. J. Collins. Cellular Decision Making and Biological Noise: From Microbes to Mammals. *Cell*, 144:910–925, 2011. doi: 10.1016/j.cell.2011.01.030.

L. Bandeira and P. Swain. Selected fluorescence and absorbance data sets from maltose and galactose grown modified yeast strains. *figshare*, 2020. doi: https://doi.org/10.6084/m9.figshare.12962444.v2.

R. D. Beer. *Dynamica (Version 1.0.8)*. 2016. URL `https://rdbeer.pages.iu.edu/`. Accessed: 19-01-2017.

C. L. Beisel and T. Afroz. Rethinking the Hierarchy of Sugar Utilization in Bacteria. *Journal of Bacteriology*, 198(3):374–6, nov 2015. ISSN 1098-5530. doi: 10.1128/JB.00890-15.

M. M. Belinchon and J. M. Gancedo. Xylose and some non-sugar carbon sources cause catabolite repression in Saccharomyces cerevisiae. *Archives of Microbiology*, 180(4):293–297, oct 2003. ISSN 0302-8933. doi: 10.1007/s00203-003-0593-9.

A. Ben Abdessalem, N. Dervilis, D. Wagg, and K. Worden. Model selection and parameter estimation in structural dynamics using approximate Bayesian com-

putation. *Mechanical Systems and Signal Processing*, 99:306–325, 2018. doi: 10.1016/j.ymssp.2017.06.017.

L. Bendrioua, M. Smedh, J. Almquist, M. Cvijovic, M. Jirstrand, M. Goksör, C. B. Adiels, and S. Hohmann. Yeast AMP-activated Protein Kinase Monitors Glucose Concentration Changes and Absolute Glucose Levels. *The Journal of Biological Chemistry*, 289(18):12863–12875, 2014. doi: 10.1074/jbc.M114.547976.

P. Boccazzi, Z. Zhang, K. Kurosawa, N. Szita, S. Bhattacharya, K. F. Jensen, and A. J. Sinskey. Differential gene expression profiles and real-time measurements of growth parameters in Saccharomyces cerevisiae grown in microliter-scale bioreactors equipped with internal stirring. *Biotechnology Progress*, 22(3):710–717, 2006. ISSN 87567938. doi: 10.1021/bp0504288.

R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4):700–765, 2006. ISSN 0033-295X. doi: 10.1037/0033-295X.113.4.700.

E. Bosdriesz, D. Molenaar, B. Teusink, and F. J. Bruggeman. How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. *FEBS Journal*, 282(10):2029–2044, 2015. ISSN 17424658. doi: 10.1111/febs.13258.

T. Bose, A. Reina, and J. A. R. Marshall. Collective decision-making. *Current Opinion in Behavioral Sciences*, 16:30–34, 2017. doi: 10.1016/j.cobeha.2017.03.004.

T. Bose, F. Bottom, A. Reina, and J. A. R. Marshall. Frequency-Sensitivity and Magnitude-Sensitivity in Decision-Making: Predictions of a Theoretical Model-Based Study. *Computational Brain & Behavior*, 2019a. ISSN 2522-0861. doi: 10.1007/s42113-019-00031-4.

T. Bose, A. Reina, and J. A. R. Marshall. Inhibition and excitation shape activity

selection : effect of oscillations in a decision-making circuit. *Neural Computation*, 31(5):870–896, 2019b.

S. Boulineau, F. Tostevin, D. J. Kiviet, P. Rein Ten Wolde, P. Nghe, and S. J. Tans. Single-Cell Dynamics Reveals Sustained Growth during Diauxic Shifts. *PLoS ONE*, 8(4), 2013. doi: 10.1371/journal.pone.0061686.

C. G. Bowsher and P. S. Swain. Environmental sensing, information transfer, and cellular decision-making. *Current Opinion in Biotechnology*, 28:149–155, 2014. ISSN 18790429. doi: 10.1016/j.copbio.2014.04.010. URL `http://dx.doi.org/10.1016/j.copbio.2014.04.010`.

C. Briat, C. Zechner, and M. Khammash. Design of a Synthetic Integral Feedback Circuit: Dynamic Analysis and DNA Implementation. *ACS Synthetic Biology*, 5: 1108–1116, 2016. doi: 10.1021/acssynbio.6b00014.

J. R. Broach. Nutritional Control of Growth and Development in Yeast. *Yeastbook*, 192:73–105, 2012. doi: 10.1534/genetics.111.135731.

T. H. C. Brondijk, W. N. Konings, and B. Poolman. Regulation of maltose transport in Saccharomyces cerevisiae. *Archives of Microbiology*, 176:96–105, 2001. doi: 10.1007/s002030100300.

R. Brückner and F. Titgemeyer. Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiology Letters*, 209(2):141–148, apr 2002. ISSN 03781097. doi: 10.1111/j.1574-6968.2002.tb11123.x.

F. J. Bruggeman, R. Planqué, D. Molenaar, and B. Teusink. Searching for principles of microbial physiology. *FEMS Microbiology Reviews*, 44(6):821–844, 2020. ISSN 15746976. doi: 10.1093/femsre/fuaa034.

M. Carlson. Glucose Repression in Yeast. *Current Opinion in Microbiology*, 2(2): 202–207, 1999. ISSN 13695274. doi: 10.1016/S1369-5274(99)80035-6.

G. P. Cereghino and I. E. Scheffler. Genetic analysis of glucose regulation in saccharomyces cerevisiae: control of transcription versus mRNA turnover. *The EMBO Journal*, 15(2):363–374, 1996. doi: 10.1002/j.1460-2075.1996.tb00366.x.

P. Chambers, A. Issaka, and S. P. Palecek. Saccharomyces cerevisiae JEN1 Promoter Activity Is Inversely Related to Concentration of Repressing Sugar. *Applied and Environmental Microbiology*, 70(1):8–17, 2004. doi: 10.1128/AEM.70.1.8-17.2004.

E. Chislenko. A Solution for Buridan ' s Ass *. *Ethics*, 126(2):283–310, 2016.

D. Chu and D. J. Barnes. The lag-phase during diauxic growth is a trade-off between fast adaptation and high growth rate. *Scientific Reports*, 6, 2016. doi: 10.1038/srep25191.

D. Chu, D. J. Barnes, T. M. Fuchs, M. Utz, and C. Pötzsche. The lag-phase during diauxic growth is a trade-off between fast adaptation and high growth rate. *Scientific Reports*, 6(1):25191, jul 2016. ISSN 2045-2322. doi: 10.1038/srep25191.

V. Chubukov, L. Gerosa, K. Kochanowski, and U. Sauer. Coordination of microbial metabolism. *Nature Reviews Microbiology*, 12(5):327–340, mar 2014. ISSN 1740-1526. doi: 10.1038/nrmicro3238.

W. S. Cooper and R. H. Kaplan. Adaptive â€œcoin-flipping": a decision-theoretic examination of natural selection for random individual variation. *Journal of Theoretical Biology*, 94(1):135–151, 1982. ISSN 10958541. doi: 10.1016/0022-5193(82)90336-8.

R. Costenoble, D. Müller, T. Barl, W. M. Van Gulik, W. A. Van Winden, M. Reuss, and J. J. Heijnen. 13C-Labeled metabolic flux analysis of a fed-batch culture of elutriated Saccharomyces cerevisiae. *FEMS Yeast Research*, 7(4):511–526, 2007. ISSN 15671356. doi: 10.1111/j.1567-1364.2006.00199.x.

I. D. Couzin. Collective cognition in animal groups. *Trends in Cognitive Sciences*, 13(1):36–43, jan 2009. ISSN 13646613. doi: 10.1016/j.tics.2008.10.002.

I. D. Couzin and J. Krause. Self-Organization and Collective Behavior in Vertebrates. *Advances in the Study of Behavior*, 32:1–75, 2003. ISSN 00653454. doi: 10.1016/S0065-3454(03)01001-5.

I. D. Couzin, C. C. Ioannou, G. Demirel, T. Gross, J. Colin, A. Hartnett, L. Conradt, S. A. Levin, and N. E. Leonard. Uninformed Individuals Promote Democratic Consensus in Animal Groups. *Science*, 334(6062):1578–1580, 2011.

A. J. Cowieson. The geometric framework: An in vivo approach. *Journal of Applied Poultry Research*, 23(2):295–300, 2014. ISSN 15370437. doi: 10.3382/japr.2014-00939.

T. D'amore, I. Russell, and G. G. Stewart. Sugar utilization by yeast during fermentation. *Journal of IndustriaI Microbiology*, 4:315–324, 1989.

P. Daran-Lapujade, M. L. A. Jansen, J. M. Daran, W. Van Gulik, J. H. De Winde, and J. T. Pronk. Role of Transcriptional Regulation in Controlling Fluxes in Central Carbon Metabolism of Saccharomyces cerevisiae: A chemostat culture study. *Journal of Biological Chemistry*, 279(10):9125–9138, 2004a. ISSN 00219258. doi: 10.1074/jbc.M309578200.

P. Daran-Lapujade, M. L. A. Jansen, J. M. Daran, W. Van Gulik, J. H. De Winde, and J. T. Pronk. Role of Transcriptional Regulation in Controlling Fluxes in Central Carbon Metabolism of Saccharomyces cerevisiae: A chemostat culture study. *Journal of Biological Chemistry*, 279(10):9125–9138, 2004b. ISSN 00219258. doi: 10.1074/jbc.M309578200.

A. Day, C. Schneider, and B. L. Schneider. Yeast cell synchronization. *Methods in molecular biology (Clifton, N.J.)*, 241:55 – 76, 2004.

R. De La Cruz, P. Guerrero, F. Spill, and T. Alarcón. The effects of intrinsic noise on the behaviour of bistable cell regulatory systems under quasi-steady state conditions. *Journal of Chemical Physics*, 143(7), 2015. ISSN 00219606. doi: 10.1063/1.4928575.

C. DeJuan and R. Lagunas. Inactivation of the galactose transport system in Saccharomyces cerevisiae. *FEBS Letters*, 207(2):258–261, 1986. ISSN 00145793. doi: 10.1016/0014-5793(86)81500-9.

J. Desponds, M. Vergassola, and A. M. Walczak. Hunchback Promoters Can Readout Morphogenetic Positional Information in Less Than a Minute. *bioRxiv*, pages 1–37, 2019. doi: 10.1101/676684.

A. Dhooge, W. Govaerts, and Y. A. Kuznetsov. MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs. *ACM Transactions on Mathematical Software*, 29(2):141–164, 2003. ISSN 00983500. doi: 10.1145/779359.779362.

M. R. Domingo-Sananes, O. Kapuy, T. Hunt, and B. Novak. Switches and latches: a biochemical tug-of-war between the kinases and phosphatases that control mitosis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1584): 3584–3594, 2011. ISSN 0962-8436. doi: 10.1098/rstb.2011.0087.

B. Drawert, A. Hellander, B. Bales, D. Banerjee, G. Bellesia, B. J. Daigle, G. Douglas, M. Gu, A. Gupta, S. Hellander, C. Horuk, D. Nath, A. Takkar, S. Wu, P. Lötstedt, C. Krintz, and L. R. Petzold. Stochastic Simulation Service: Bridging the Gap between the Computational Expert and the Biologist. *PLOS Computational Biology*, 12(12):e1005220, dec 2016. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1005220.

A. Dussutour, Q. Ma, and D. Sumpter. Phenotypic variability predicts decision accuracy in unicellular organisms. *Proceedings of the Royal Society B: Biological Sciences*, 286(1896), 2019. ISSN 14712954. doi: 10.1098/rspb.2018.2825.

H. El-Samad and M. Khammash. Modelling and analysis of gene regulatory network using feedback control theory. *International Journal of Systems Science*, 41(1): 17–33, 2010. ISSN 0020-7721. doi: 10.1080/00207720903144545.

S. Elmore. Apoptosis: A Review of Programmed Cell Death. *Toxicologic Pathology*, 35(4):495–516, 2007. ISSN 01926233. doi: 10.1080/01926230701320337.

P. Eraso and J. M. Gancedo. Catabolite repression in yeasts is not associated with low levels of cAMP. *European Journal of Biochemistry*, 141(1):195–198, 1984. ISSN 14321033. doi: 10.1111/j.1432-1033.1984.tb08174.x.

B. Ermentrout. *Simulating, Analyzing, and Animating Dynamical Systems: A Guide to XPPAUT for Researchers and Students*. SIAM, 2002. ISBN 978-0-89871-506-4.

R. Escalante-Chong, Y. Savir, S. M. Carroll, J. B. Ingraham, J. Wang, C. J. Marx, and M. Springer. Galactose metabolic genes in yeast respond to a ratio of galactose and glucose. *Proceedings of the National Academy of Sciences of the United States of America*, 112(5):1636–41, 2015. ISSN 1091-6490. doi: 10.1073/pnas.1418058112.

M. Farré, L. Kantiani, and D. Barceló. Microfluidic Devices: Biosensors. In *Chemical Analysis of Food: Techniques and Applications*, page 192. 2012. ISBN 9780123848628. doi: 10.1016/B978-0-12-384862-8.00007-8.

H. J. Federoff, T. R. Eccleshall, and J. Marmur. Carbon catabolite repression of maltase synthesis in Saccharomyces carlsbergensis. *Journal of Bacteriology*, 156 (1):301–307, 1983. ISSN 00219193.

J. E. Ferrell. Self-perpetuating states in signal transduction: Positive feedback, double-negative feedback and bistability. *Current Opinion in Cell Biology*, 14(2): 140–148, 2002. ISSN 09550674. doi: 10.1016/S0955-0674(02)00314-9.

J. M. Gancedo. Yeast Carbon Catabolite Repression. *Microbiology and Molecular Biology Reviews*, 62(2):334–361, 1998.

J. M. Gancedo. The Early Steps of Glucose Signalling in Yeast. *FEMS Microbiology Review*, 32:673–704, 2008. doi: 10.1111/j.1574-6976.2008.00117.x.

J. M. Gancedo, S. López, and F. Ballesteros. Calculation of half-lives of proteins in vivo. Heterogeneity in the rate of degradation of yeast proteins. *Molecular and Cellular Biochemistry*, 43(2):89–95, 1982. ISSN 03008177. doi: 10.1007/BF00423096.

F. J. Geske, R. Lieberman, R. Strange, and L. E. Gerschenson. Early stages of p53-induced apoptosis are reversible. *Cell Death and Differentiation*, 8(2):182–191, 2001. ISSN 13509047. doi: 10.1038/sj.cdd.4400786.

D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4): 403–434, 1976. ISSN 10902716. doi: 10.1016/0021-9991(76)90041-3.

B. Görke and J. Stülke. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nature Reviews Microbiology*, 6(8):613–24, 2008. ISSN 1740-1534. doi: 10.1038/nrmicro1932.

W. Govaerts, Y. A. Kuznetsov, H. G. E. Meijer, V. D. Witte, A. Dhooge, W. Mestrom, N. Neirynck, A. M. Riet, and B. Sautois. MATCONT: Continuation toolbox for ODEs in Matlab. 2019.

S. A. Haddad and C. C. Lindegren. A method for determining the weight of an individual yeast cell. *Applied microbiology*, 1(3):153–156, 1953. ISSN 00036919. doi: 10.1128/AEM.1.3.153-156.1953.

C. Häggström, U. Rova, T. Brandberg, and D. B. Hodge. Integration of Ethanol Fermentation with Second Generation Biofuels Technologies. *Biorefineries: Integrated Biochemical Processes for Liquid Biofuels*, pages 161–187, 2014. doi: 10.1016/B978-0-444-59498-3.00008-7.

T. D. Hanks and C. Summerfield. Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron*, 93:15–31, 2017. doi: 10.1016/j.neuron.2016.12.003.

Y. Hart, Y. E. Antebi, a. E. Mayo, N. Friedman, and U. Alon. Design principles of cell circuits with paradoxical components. *Proceedings of the National Academy of Sciences of the United States of America*, 109(21):8346–8351, 2012. ISSN 0027-8424. doi: 10.1073/pnas.1117475109.

Y. Hart, A. E. Mayo, O. Shoval, and U. Alon. Comparing Apples and Oranges:

Fold-Change Detection of Multiple Simultaneous Inputs. *PLoS ONE*, 8(3), 2013. ISSN 19326203. doi: 10.1371/journal.pone.0057455.

R. Hastie and T. Kameda. The robust beauty of majority rules in group decisions. *Psychological Review*, 112(2):494–508, 2005. ISSN 0033295X. doi: 10.1037/0033-295X.112.2.494.

J. Henke-Von Der Malsburg and C. Fichtel. Are generalists more innovative than specialists? A comparison of innovative abilities in two wild sympatric mouse lemur species. *Royal Society Open Science*, 5(8), 2018. ISSN 20545703. doi: 10.1098/rsos.180480.

P. Herrero, R. Fernfindez, and F. Moreno. Differential sensitivities to glucose and galactose repression of gluconeogenic and respiratory enzymes from Saccharomyces cerevisiae. *Arch Microbiol*, 143:216–219, 1985.

B. Ho, A. Baryshnikova, and G. W. Brown. Unification of Protein Abundance Datasets Yields a Quantitative Saccharomyces cerevisiae Proteome. *Cell Systems*, 6(2):192–205.e3, 2018. ISSN 24054720. doi: 10.1016/j.cels.2017.12.004.

B. M. Hogema, J. C. Arents, R. Bader, and P. W. Postma. Autoregulation of lactose uptake through the LacY permease by enzyme IIA(Glc) of the PTS in Escherichia coli K-12. *Molecular Microbiology*, 31(6):1825–1833, 1999. ISSN 0950382X. doi: 10.1046/j.1365-2958.1999.01319.x.

S. Hogg. Microbial Growth. In *Essential Microbiology*, chapter 5, pages 91–105. John Wiley & Sons, Ltd, 2005.

J. Hong. Yield Coefficients for Cell Mass and Product Formation. *Biotechnology and Bioengineering*, 33(1):506–507, 1989. ISSN 00223999. doi: 10.1016/0022-3999(83)90113-7.

K.-K. Hong, W. Vongsangnak, G. N. Vemuri, and J. Nielsen. Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated

systems level analysis. *Proceedings of the National Academy of Sciences*, 108(29): 12179–12184, jul 2011. ISSN 0027-8424. doi: 10.1073/PNAS.1103219108.

J. Horák. Regulations of sugar transporters: insights from yeast. *Current Genetics*, 59(1-2):1–31, may 2013. ISSN 0172-8083. doi: 10.1007/s00294-013-0388-8.

J. Horak and D. H. Wolf. Catabolite Inactivation of the Galactose Transporter in the Yeast Saccharomyces cerevisiae: Ubiquitination, Endocytosis, and Degradation in the Vacuole. *Journal of Bacteriology*, 179(5):1541–1549, 1997.

J. Horak, J. Regelmann, and D. H. Wolf. Two Distinct Proteolytic Systems Responsible for Glucose-induced Degradation of Fructose-1,6-bisphosphatase and the Gal2p Transporter in the Yeast Saccharomyces cerevisiae Share the Same Protein Components of the Glucose Signaling Pathway*. *The Journal of Biological Chemistry*, 277(10):8248–8254, 2002. doi: 10.1074/jbc.M107255200.

A. I. Houston, A. D. Higginson, and J. M. Mcnamara. Optimal foraging for multiple nutrients in an unpredictable environment. *Ecology Letters*, 14:1101–1107, 2011. doi: 10.1111/j.1461-0248.2011.01678.x.

Z. Hu, Y. Yue, H. Jiang, B. Zhang, P. W. Sherwood, and C. A. Michels. Analysis of the Mechanism by Which Glucose Inhibits Maltose Induction of MAL Gene Expression in Saccharomyces. *Genetics*, 154(1):121–132, 2000.

J. M. Hutchinson and G. Gigerenzer. Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural Processes*, 69 (2):97–124, 2005. ISSN 03766357. doi: 10.1016/j.beproc.2005.02.019.

N. Jamshidi and B. Ø. Palsson. Mass Action Stoichiometric Simulation Models: Incorporating Kinetics and Regulation into Stoichiometric Models. *Biophysical Journal*, 98:175–185, 2010. doi: 10.1016/j.bpj.2009.09.064.

H. Jiang, I. Medintz, and C. A. Michels. Two Glucose Sensing/Signaling Pathways Stimulate Glucose-induced Inactivation of Maltose Permease in Saccharomyces. *Molecular Biology of the Cell*, 8:1293–1304, 1997.

H. Jiang, I. Medintz, B. Zhang, and C. A. Michels. Metabolic Signals Trigger Glucose-Induced Inactivation of Maltose Permease in Saccharomyces. *Journal of Bacteriology*, 182(3):647–654, 2000.

M. Johnston, J. S. Flick, and T. Pextont. Multiple Mechanisms Provide Rapid and Stringent Glucose Repression of GAL Gene Expression in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, 14(6):3834–3841, 1994.

A. Kaniak, Z. Xue, D. Macool, J. H. Kim, and M. Johnston. Regulatory Network Connecting Two Glucose Signal Transduction Pathways in Saccharomyces cerevisiae. *Eukaryotic Cell*, 3(1):221–231, 2004. ISSN 15359778. doi: 10.1128/EC.3.1.221-231.2004.

A. B. Kao, N. Miller, C. Torney, A. Hartnett, and I. D. Couzin. Collective Learning and Optimal Consensus Decisions in Social Animal Groups. *PLoS Comput Biol*, 10(8), 2014. doi: 10.1371/journal.pcbi.1003762.

F. Kargi. Re-interpretation of the logistic equation for batch microbial growth in relation to Monod kinetics. *Letters in Applied Microbiology*, 48(4):398–401, apr 2009. ISSN 02668254. doi: 10.1111/j.1472-765X.2008.02537.x.

O. Kayikci and J. Nielsen. Glucose repression in Saccharomyces cerevisiaë. *FEMS Yeast Research*, 15, 2015. doi: 10.1093/femsyr/fov068.

A. Khmelinskii, P. J. Keller, A. Bartosik, M. Meurer, J. D. Barry, B. R. Mardin, A. Kaufmann, S. Trautmann, M. Wachsmuth, G. Pereira, W. Huber, E. Schiebel, and M. Knop. Tandem fluorescent protein timers for in vivo analysis of protein dynamics. *Nature Biotechnology*, 30(7):708–714, 2012. ISSN 10870156. doi: 10.1038/nbt.2281.

B. N. Kholodenko. Cell Signalling Dynamics in Time and Space. *Nature Reviews Molecular Cellular Biology*, 7(3):165–176, 2006.

D. Kim, J. Y. Song, and J. S. Hahn. Improvement of glucose uptake rate and production of target chemicals by overexpressing hexose transporters and transcriptional

activator Gcr1 in Saccharomyces cerevisiae. *Applied and Environmental Microbiology*, 81(24):8392–8401, 2015. ISSN 10985336. doi: 10.1128/AEM.02056-15.

C. J. L. Klein. Glucose and maltose metabolism in MIG1-disrupted and MAL-constitutive strains of Saccharomyces cerevisiae. *Food Technolgy and Biotechnology*, 35(4):287–292, 1997.

S. Koirala, X. Wang, and C. V. Rao. Reciprocal Regulation of L-Arabinose and D-Xylose Metabolism in Escherichia coli. *Journal of Bacteriology*, 198(3):386–93, 2016. doi: 10.1128/JB.00709-15.

A. Kuchina, L. Espinar, J. Garcia-Ojalvo, G. Rol, and M. Sü El. Reversible and Noisy Progression towards a Commitment Point Enables Adaptable and Reliable Cellular Decision-Making. *PLoS Comput Biol*, 7(11):e1002273, 2011. doi: 10.1371/journal.pcbi.1002273.

J. Kuha. AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods and Research*, 33(2):188–229, 2004. doi: 10.1177/0049124103262065.

E. Kussell and S. Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309(10):2075–2078, 2005. ISSN 19457197. doi: 10.1210/jcem-10-10-1361.

S. Lane, H. Xu, E. J. Oh, H. Kim, A. Lesmana, D. Jeong, G. Zhang, C. S. Tsai, Y. S. Jin, and S. R. Kim. Glucose repression can be alleviated by reducing glucose phosphorylation rate in Saccharomyces cerevisiae. *Scientific Reports*, 8, 2018. ISSN 20452322. doi: 10.1038/s41598-018-20804-4.

M. Laurent and N. Kellershohn. Multistability: A major means of differentiation and evolution in biological systems. *Trends in Biochemical Sciences*, 24(11):418–422, 1999. ISSN 09680004. doi: 10.1016/S0968-0004(99)01473-5.

K. Lemaire, S. Van De Velde, P. V. Dijck, and J. M. Thevelein. Glucose and Sucrose Act as Agonist and Mannose as Antagonist Ligands of the G Protein-

Coupled Receptor Gpr1 in the Yeast Saccharomyces cerevisiae. *Molecular Cell*, 16:293–299, 2004.

J. W. Lengeler. *Encyclopedia of Genetics*, volume 67. 2001. doi: 10.1016/S0300-9084(85)80227-3.

C. A. Lichten, R. White, I. B. Clark, and P. S. Swain. Unmixing of fluorescence spectra to resolve quantitative time-series measurements of gene expression in plate readers. *BMC Biotechnology*, 14, 2014. ISSN 14726750. doi: 10.1186/1472-6750-14-11.

M. Lindauer. Communication in Swarm-Bees Searching for a New Home. *Nature*, 179(4550):63–66, 1957.

M. Y. Liu, L. Zhang, and C. F. Zhang. Study on Banded Implicit Runge-Kutta Methods for Solving Stiff Differential Equations. *Mathematical Problems in Engineering*, 2019, 2019. ISSN 15635147. doi: 10.1155/2019/4850872.

T. Lodi, C. Donnini, and I. Ferrero. Catabolite repression by galactose in overexpressed GAL4 strains of Saccharomyces cerevisiae. *Journal of General Microbiology*, 137:1039–1044, 1991.

T. Long, K. C. Tu, Y. Wang, P. Mehta, N. P. Ong, B. L. Bassler, and N. S. Wingreen. Quantifying the integration of quorum-sensing signals with single-cell resolution. *PLoS Biology*, 7(3):0640–0649, 2009. ISSN 15449173. doi: 10.1371/journal.pbio.1000068.

O. H. . Lowry, N. J. . Rosebrough, A. L. . Farr, and R. J. Randall. Protein Measurement with the Folin Phenol Reagent. *The Journal of Biological Chemistry*, 193(1):265–275, 1951.

P. J. Macdonald, Y. Chen, and J. D. Mueller. Chromophore maturation and fluorescence fluctuation spectroscopy. *Analytical Biochemistry*, 421(1):291–298, 2013. doi: 10.1016/j.ab.2011.10.040.Chromophore.

R. Mahadevan, J. S. Edwards, and F. J. Doyle. Dynamic Flux Balance Analysis of Diauxic Growth in Escherichia coli. *Biophysical Journal*, 83:1331–1340, 2002.

J. A. Marshall, A. Favreau-Peigné, L. Fromhage, J. M. McNamara, L. F. Meah, and A. I. Houston. Cross inhibition improves activity selection when switching incurs time costs. *Current Zoology*, 61(2):242–250, 2015. ISSN 16745507.

J. A. Marshall, R. H. Kurvers, J. Krause, and M. Wolf. Quorums enable optimal pooling of independent judgements in biological systems. *eLife*, 8(i):1–14, 2019a. ISSN 2050084X. doi: 10.7554/eLife.40368.

J. A. Marshall, A. Reina, and T. Bose. Multiscale modelling tool: Mathematical modelling of collective behaviour without the maths. *PLoS ONE*, 14(9):1–16, 2019b. ISSN 19326203. doi: 10.1371/journal.pone.0222906.

J. A. R. Marshall. Comment on optimal policy for multi-alternative decisions. 2019.

J. A. R. Marshall, R. Bogacz, A. Dornhaus, R. Planqué, T. Kovacs, and N. R. Franks. On optimal decision-making in brains and social insect colonies. *Journal of The Royal Society Interface*, 6:1065 – 1074, 2009. doi: 10.1098/rsif.2008.0511.

N. Masuda, T. A. O'Shea-Wheller, C. Doran, and N. R. Franks. Computational model of collective nest selection by ants with heterogeneous acceptance thresholds. *Royal Society Open Science*, 2(6), 2015. ISSN 20545703. doi: 10.1098/rsos.140533.

MathWorks. ode15s, 2020. URL `https://uk.mathworks.com/help/matlab/ref/ode15s.html`. Accessed: 21-05-2020.

MATLAB. *version 9.4.0 (R2018a)*. The MathWorks Inc., Natick, Massachusetts, 2018.

J. M. McNamara, P. C. Trimmer, A. Eriksson, J. A. Marshall, and A. I. Houston. Environmental variability can select for optimism or pessimism. *Ecology Letters*, 14(1):58–62, 2011. ISSN 14610248. doi: 10.1111/j.1461-0248.2010.01556.x.

I. Medintz, H. Jiang, E. K. Han, W. Cui, and C. A. Michels. Characterization of the glucose-induced inactivation of maltose permease in Saccharomyces cerevisiae. *Journal of Bacteriology*, 178(8):2245–2254, 1996. ISSN 00219193. doi: 10.1128/jb.178.8.2245-2254.1996.

M. M. C. Meijer, J. Boonstra, A. J. Verkleij, and C. T. Verrips. Glucose Repression in Saccharomyces cerevisiae Is Related to the Glucose Concentration Rather Than the Glucose Flux*. *The Journal of Biological Chemistry*, 273(37):24102–24107, 1998.

R. Milo and R. Phillips. Cell biology by the numbers. In S. Scholl, K. Laurentiev, and N. Wolfe, editors, *Choice Reviews Online*, chapter 4, pages 286 – 288. Garland Science, Taylor & Francis Group, LLC, 2015. doi: 10.5860/choice.196525.

S. Mitri, S. Wischmann, D. Floreano, and L. Keller. Using robots to understand social behaviour. *Biological Reviews*, 88(1):31–39, 2013. ISSN 14647931. doi: 10.1111/j.1469-185X.2012.00236.x.

D. Molenaar, R. Van Berlo, D. De Ridder, and B. Teusink. Shifts in growth strategies reflect tradeoffs in cellular economics. *Molecular Systems Biology*, 5(323):1–10, 2009. ISSN 17444292. doi: 10.1038/msb.2009.82. URL `http://dx.doi.org/10.1038/msb.2009.82`.

H. D. Møller, K. S. Andersen, and B. Regenberg. A model for generating several adaptive phenotypes from a single genetic event. *Communicative and Integrative Biology*, 6(3), 2013. ISSN 19420889. doi: 10.4161/cib.23933.

J. Monod. The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 1949.

J. D. Murray. *Mathematical Biology : An Introduction*. Springer, proquest e edition, 2002.

I. Mytilinaios, M. Salih, H. K. Schofield, and R. J. Lambert. Growth curve prediction from optical density data. *International Journal of Food Microbiology*, 154(3):169–176, 2012. ISSN 01681605. doi: 10.1016/j.ijfoodmicro.2011.12.035.

A. Narang. The Dynamical Analogy Between Microbial Growth on Mixtures of Substrates and Population Growth of Competing Species. *Biotechnology and Bioengineering*, (3):3–8, 1998.

A. Narang and S. S. Pilyugin. Bacterial gene regulation in diauxic and non-diauxic growth. *Journal of Theoretical Biology*, 244:326–348, 2007. doi: 10.1016/j.jtbi.2006.08.007.

A. Narang, A. Konopka, and D. Ramkrishna. The Dynamics of Microbial Growth on Mixtures of Substrates in Batch Reactors. *Journal of Theoretical Biology*, 184: 301–317, 1997.

S. Neelamegham and G. Liu. Systems glycobiology: Biochemical reaction networks regulating glycan structure and function. *Glycobiology*, 21(12):1541–1553, 2011. ISSN 09596658. doi: 10.1093/glycob/cwr036.

N. R. Nené, J. Garca-Ojalvo, and A. Zaikin. Speed-dependent cellular decision making in nonequilibrium genetic circuits. *PLoS ONE*, 7(3):1–7, 2012. ISSN 19326203. doi: 10.1371/journal.pone.0032779.

J. Nevado, M. A. Navarro, and C. F. Heredia. Galactose inhibition of the constitutive transport of hexoses in Saccharomyces cerevisiae. *Yeast*, 9(2):111–119, feb 1993. ISSN 0749-503X. doi: 10.1002/yea.320090202.

A. M. New, B. Cerulus, S. K. Govers, G. Perez-Samper, B. Zhu, S. Boogmans, J. B. Xavier, and K. J. Verstrepen. Different Levels of Catabolite Repression Optimize Growth in Stable and Variable Environments. *PLoS Biology*, 12(1):e1001764, jan 2014. ISSN 1545-7885. doi: 10.1371/journal.pbio.1001764.

T. D. Nguyen-Huu, C. Gupta, B. Ma, W. Ott, K. Josić, and M. R. Bennett. Timing and Variability of Galactose Metabolic Gene Activation Depend on the Rate of Environmental Change. *PLoS Computational Biology*, 11(7):1–16, 2015. ISSN 15537358. doi: 10.1371/journal.pcbi.1004399.

S. C. Nicolis and J. L. Deneubourg. Emerging Patterns and Food Recruitment in Ants: an Analytical Study. *J. theor. Biol*, 198:575 – 592, 1999.

S. Nikolov, X. Lai, U. Liebal, O. Wolkenhauer, and J. Vera. Integration of sensitivity and bifurcation analysis to detect critical processes in a model combining signalling and cell population dynamics. *International Journal of Systems Science*, 41(1):81–105, 2010. ISSN 0020-7721. doi: 10.1080/00207720903147746.

A. Nilsson and J. Nielsen. Metabolic Trade-offs in Yeast are Caused by F1F0-ATP synthase. *Scientific Reports*, 2016. doi: 10.1038/srep22264.

J. T. Noel, S. S. Pilyugin, and A. Narang. The diffusive influx and carrier efflux have a strong effect on the bistability of the lac operon in Escherichia coli. *Journal of Theoretical Biology*, 256(1):14–28, 2009. ISSN 00225193. doi: 10.1016/j.jtbi.2008.09.003.

S. Novak, V. Zechner-krpan, and V. Mari. Regulation of Maltose Transport and Metabolism in Saccharomyces cerevisiae. *Food Technolgy and Biotechnology*, 42 (3):213–218, 2004.

A. Novick and M. Weiner. Enzyme Induction as an All-or-None Phenomenon. *Proceedings of the National Academy of Sciences of the Unites States of America*, 43 (7):553–566, 1957.

H. Olofsson, J. Ripa, N. Jonzén, H. Olofsson, J. Ripa, and N. Jonzen. Bet-hedging as an evolutionary game : the trade-off between egg size and number. *Proceedings of the Royal Society B*, 276(1669):2963–2969, 2009.

M. G. Ormerod. Introduction. In *Flow Cytometry - A Basic Introduction*, chapter 1. De Novo, 2008. URL `https://flowbook.denovosoftware.com/`. Accessed: 20-04-2020.

E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. van Oudenaarden. Multistability in the lactose utilisation network of Escherichia coli. *Nature*, 427 (6976):737–740, 2004a.

E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. Van Oudenaarden. Multistability in the lactose utilization network of Escherichia coli. *Nature*, 427 (6976):737–740, 2004b. ISSN 0028-0836. doi: 10.1038/nature02298.

S. Ozcan and M. Johnston. Function and regulation of yeast hexose transporters. *Microbiology and molecular biology reviews : MMBR*, 63(3):554–69, sep 1999. ISSN 1092-2172.

D. Pais, P. M. Hogan, T. Schlegel, N. R. Franks, N. E. Leonard, and J. A. R. Marshall. A Mechanism for Value-Sensitive Decision-Making. *PLoS ONE*, 8(9), 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0073216.

B. Ø. Palsson. *Systems Biology Simulation of Dynamic Network States*. Cambridge University Press, 2011.

M. Papamichos-Chronakis, T. Gligoris, and D. Tzamarias. The Snf1 kinase controls glucose repression in yeast by modulating interactions between the Mig1 repressor and the Cyc8-Tup1 co-repressor. *EMBO reports*, 5:368–372, 2004. doi: 10.1038/sj.embor.7400120.

B. L. Partridge. The structure and function of fish schools. *Scientific American*, 246(6):114 – 123, 1982. doi: 10.1201/9780203755990.

T. J. Perkins and P. S. Swain. Strategies for cellular decision-making. *Molecular systems biology*, 5(326):326, 2009. ISSN 1744-4292. doi: 10.1038/msb.2009.83.

T. Petit, J. A. Diderich, A. L. Kruckeberg, C. Gancedo, and K. Van Dam. Hexokinase regulates kinetics of glucose transport and expression of genes encoding hexose transporters in Saccharomyces cerevisiae. *Journal of bacteriology*, 182(23): 6815–8, dec 2000. ISSN 0021-9193.

T. Pfeiffer and A. Morley. An evolutionary perspective on the Crabtree effect. *Frontiers in Molecular Biosciences*, 1, 2014. ISSN 2296-889X. doi: 10.3389/fmolb.2014.00017.

T. Philippi and J. Seger. Hedging one's evolutionary bets, revisited. *Trends in Ecology and Evolution*, 4(2):41–44, 1989. ISSN 01695347. doi: 10.1016/0169-5347(89)90138-9.

A. Pirrone, T. Stafford, J. A. R. Marshall, D. Standage, and J. I. Gold. When natural selection should optimize speed-accuracy trade-offs. *Frontiers in Neuroscience*, 08, 2014. doi: 10.3389/fnins.2014.00073.

A. Pirrone, H. Azab, B. Y. Hayden, T. Stafford, and J. A. Marshall. Evidence for the speed-value trade-off: Human and monkey decision making is magnitude sensitive. *Decision*, 5(2):129–142, 2018a. ISSN 23259973. doi: 10.1037/dec0000075.

A. Pirrone, W. Wen, and S. Li. Single-trial dynamics explain magnitude sensitive decision making. *BMC Neuroscience*, 19(1):54, dec 2018b. ISSN 1471-2202. doi: 10.1186/s12868-018-0457-5.

J. R. Pomerening, E. D. Sontag, and J. E. Ferrell. Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nature cell biology*, 5(4):346–51, 2003. ISSN 1465-7392. doi: 10.1038/ncb954.

S. C. Pratt, E. B. Mallon, D. J. T. Sumpter, and N. R. Franks. Quorum Sensing , Recruitment , and Collective Decision-Making during Colony Emigration by the Ant Leptothorax albipennis. *Behavioral Ecology and Sociobiology*, 52(2):117–127, 2002.

R. Ratcliff. A theory of memory retrieval. *Psychological Review*, 85(2):59–108, 1978. ISSN 0033295X. doi: 10.1037/0033-295X.85.2.59.

R. Ratcliff and P. L. Smith. A comparison of sequential sampling models for 2-choice response. *Psychological Review*, 306(5703):1895, 2004.

R. Ratcliff and G. McKoon. Drift Diffusion Decision Model: Theory and data. *Neural Computation*, 20(4):873–922, 2008. ISSN 15378276. doi: 10.1016/j.biotechadv.2011.08.021.Secreted.

R. Ratcliff and J. N. Rouder. A Diffusion Model Account of Masking in Two-Choice Letter Identification. *Journal of Experimental Psychology: Human Pereeption and Performance*, 26(1):127–140, 2000. doi: 10.10371/0096-1523.26.1.127.

M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. *Journal of Chemical Physics*, 119:12784–12794, 2003. doi: 10.1063/1.1627296.

A. Reece, B. Xia, Z. Jiang, B. Noren, and R. McBride. Microfluidic Techniques for High Throughput Single Cell Analysis Graphical Abstract HHS Public Access. *Curr Opin Biotechnol*, 40:90–96, 2016. doi: 10.1016/j.copbio.2016.02.015.

E. Reifenberger, E. Boles, and M. Ciriacy. Kinetic characterization of individual hexose transporters of Saccharomyces cerevisiae and their relation to the triggering mechanisms of glucose repression. *European Journal of Biochemistry*, 245(2): 324–333, 1997. ISSN 00142956. doi: 10.1111/j.1432-1033.1997.00324.x.

A. Reina, J. A. R. Marshall, V. Trianni, and T. Bose. Model of the best-of-N nest-site selection process in honeybees. *PHYSICAL REVIEW E*, 95(5), 2017. doi: 10.1103/PhysRevE.95.052411.

A. Reina, T. Bose, V. Trianni, and J. Marshall. Psychophysical Laws and the Superorganism. *Scientific Reports*, 8(1), 2018. ISSN 20452322. doi: 10.1038/s41598-018-22616-y.

N. Rescher. Choice Without Preference. *Kant-Studien*, 51((1-4)):142–175, 1960. doi: 10.2307/j.ctt284tx3.4.

A. Rhee, R. Cheong, and A. Levchenko. The application of information theory to biochemical signaling systems. *Physical Biology*, 9(4), 2012. ISSN 14783967. doi: 10.1088/1478-3975/9/4/045011.

J. Ripa, H. Olofsson, and N. Jonzén. What is bet-hedging, really? *Proceedings of the Royal Society B: Biological Sciences*, 277(1685):1153–1154, 2009. ISSN 14712970. doi: 10.1098/rspb.2009.2023.

F. Rolland, J. Winderickx, and J. M. Thevelein. Glucose-sensing and -signalling mechanisms in yeast. *FEMS Yeast Research*, 2(2):183–201, 2002. ISSN 15671356. doi: 10.1016/S1567-1356(02)00046-6.

M. Salathé, J. Van Cleve, and M. W. Feldman. Evolution of stochastic switching rates in asymmetric fitness landscapes. *Genetics*, 182(4):1159–1164, 2009. ISSN 00166731. doi: 10.1534/genetics.109.103333.

K. R. Sanft, S. Wu, M. Roh, J. Fu, R. K. Lim, and L. R. Petzold. StochKit2: Software for discrete stochastic simulation of biochemical systems with events. *Bioinformatics*, 27(17):2457–2458, 2011. ISSN 13674803. doi: 10.1093/bioinformatics/btr401.

G. M. Santangelo. Glucose Signaling in Saccharomyces cerevisiae. *Microbiology and Molecular Biology Reviews*, 70(1):253–282, 2006. doi: 10.1128/MMBR.70.1.253-282.2006.

M. Santillán and M. C. Mackey. Influence of catabolite repression and inducer exclusion on the bistable behavior of the lac operon. *Biophysical Journal*, 86(3):1282–92, 2004. ISSN 0006-3495. doi: 10.1016/S0006-3495(04)74202-2.

T. Sasaki and S. C. Pratt. The Psychology of Superorganisms: Collective Decision Making by Insect Societies. *Annu. Rev. Entomol*, 63:259–75, 2018. doi: 10.1146/annurev-ento-020117.

Y. Schaerli, A. Munteanu, M. Gili, J. Cotterell, J. Sharpe, and M. Isalan. A unified design space of synthetic stripe-forming networks. *Nature Communications*, 5, sep 2014. ISSN 20411723. doi: 10.1038/ncomms5905.

H.-J. Schüller. Transcriptional control of nonfermentative metabolism in the yeast Saccharomyces cerevisiae. *Current Genetics*, 43(3):139–160, 2003. ISSN 0172-8083. doi: 10.1007/s00294-003-0381-8.

M. Schweizer and J. R. Dickinson. *The metabolism and molecular physiology of Saccharomyces cervisiae*. Taylor and Francis, 1 edition, 1999.

T. D. Seeley and P. Kirk Visscher. Quorum sensing during nest-site selection by honeybee swarms. *Behavioral Ecology and Sociobiology*, 56(6):594–601, 2004. ISSN 03405443. doi: 10.1007/s00265-004-0814-5.

T. D. Seeley, P. K. Visscher, T. Schlegel, P. M. Hogan, N. R. Franks, and J. A. R. Marshall. Stop Signals Provide Cross Inhibition in Collective Decision-Making by Honeybee Swarms. *Science*, 335(January):108–111, 2012.

J. Seger and H. J. Brockmann. What is bet-hedging? *Oxford Surveys in Evolutionary Biology*, 4:182 – 211, 1987.

R. Seydel. *From Equilibrium to Chaos. Practical Bifurcation and Stability Analysis.* Elsevier Science Publishing Co., Inc., 1988.

V. Shahrezaei and S. Marguerat. Connecting growth with gene expression: of noise and numbers A global regulation of gene expression by cellular growth. *Current Opinion in Microbiology*, 25:127–135, 2015. doi: 10.1016/j.mib.2015.05.012.

T. Shimada, A. Kori, and Ishihama. Involvement of the ribose operon repressor RbsR in regulation of purine nucleotide synthesis in Escherichia coli. *FEMS Microbiology Letters*, 344(2):159–165, jul 2013. ISSN 03781097. doi: 10.1111/1574-6968.12172.

O. Shoval, L. Goentoro, Y. Hart, A. Mayo, E. Sontag, and U. Alon. Fold-change detection and scalar symmetry of sensory input fields. *Proceedings of the National Academy of Sciences of the United States of America*, 107(36):15995–6000, 2010. ISSN 1091-6490. doi: 10.1073/pnas.1002352107.

L. N. Sierkstra, H. H. Sillje, J. M. Verbakel, and C. T. Verrips. The glucose-6-phosphate-isomerase reaction is essential for normal glucose repression in Saccharomyces cerevisiae. *European Journal of Biochemistry*, 214(1):121–127, 1993. ISSN 14321033. doi: 10.1111/j.1432-1033.1993.tb17903.x.

S. J. Simpson and D. Raubenheimer. A multi-level analysis of feeding behaviour: The geometry of nutritional decisions. *Philosophical Transactions of the Royal*

Society B: Biological Sciences, 342(1302):381–402, 1993. ISSN 09628436. doi: 10.1098/rstb.1993.0166.

S. J. Simpson, D. G. Le Couteur, D. E. James, J. George, J. E. Gunton, S. M. Solon-Biet, and D. Raubenheimer. The Geometric Framework for Nutrition as a tool in precision medicine. *Nutrition and Healthy Aging*, 4(3):217–226, 2017. ISSN 24519502. doi: 10.3233/NHA-170027.

K. Simpson-Lavy and M. Kupiec. Carbon Catabolite Repression in Yeast is Not Limited to Glucose. *Scientific Reports*, 9(1):6491, dec 2019. ISSN 2045-2322. doi: 10.1038/s41598-019-43032-w.

K. Simpson-Lavy, T. Xu, M. Johnston, and M. Kupiec. The Std1 Activator of the Snf1/AMPK Kinase Controls Glucose Response in Yeast by a Regulated Protein Aggregation. *Molecular Cell*, 68(6):1120–1133, 2017. ISSN 10974164. doi: 10.1016/j.molcel.2017.11.016.

M. Siwiak and P. Zielenkiewicz. A comprehensive, quantitative, and genome-wide model of translation. *PLoS Computational Biology*, 6(7):4, 2010. ISSN 1553734X. doi: 10.1371/journal.pcbi.1000865.

M. Slatkin. Hedging one'e evolutionary bets. *Nature*, 250(5469):704, 1974. ISSN 00280836. doi: 10.1038/250704a0.

E. L. Snapp. Fluorescent proteins: a cell biologist's user guide. *Trends in Cell Biology*, 19(11):649–655, 2009. ISSN 09628924. doi: 10.1016/j.tcb.2009.08.002.

A. Solopova, J. van Gestel, F. J. Weissing, H. Bachmann, B. Teusink, J. Kok, O. P. Kuipers, and b. R. Todd Klaenhammer. Bet-hedging during bacterial diauxic shift. *Proceedings of the National Academy of Sciences of the Unites States of America*, 111(20):7427–7432, 2013. doi: 10.1073/pnas.1320063111.

B. Sonnleitner, G. Locher, and A. Fiechter. Biomass determination. *Journal of Biotechnology*, 25(1-2):5–22, 1992. ISSN 01681656. doi: 10.1016/0168-1656(92)90107-K.

K. Steverson, H. K. Chung, J. Zimmermann, K. Louie, and P. Glimcher. Sensitivity of reaction time to the magnitude of rewards reveals the cost-structure of time. *Scientific reports*, 9(1):20053, 2019. ISSN 20452322. doi: 10.1038/s41598-019-56392-0.

S. H. Strogatz. *Nonlinear dynamics and Chaos*. Perseus Books, 1994.

J. Stülke and W. Hillen. Carbon catabolite repression in bacteria. *Current Opinion in Microbiology*, 2:195–201, 1999.

L. P. Sugrue, G. S. Corrado, and W. T. Newsome. Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nature Reviews Neuroscience*, 6(5):363–375, 2005. ISSN 1471003X. doi: 10.1038/nrn1666.

D. J. Sumpter, J. Krause, R. James, I. D. Couzin, and A. J. Ward. Consensus Decision Making by Fish. *Current Biology*, 18(22):1773–1777, 2008. ISSN 09609822. doi: 10.1016/j.cub.2008.09.064.

P. S. Swain, K. Stevenson, A. Leary, L. F. Montano-Gutierrez, I. B. Clark, J. Vogel, and T. Pilizota. Inferring time derivatives including cell growth rates using Gaussian processes. *Nature Communications*, 7(May):1–8, 2016. ISSN 20411723. doi: 10.1038/ncomms13766.

S. Tajima, J. Drugowitsch, and A. Pouget. Optimal policy for value-based decision-making. *Nature Communications*, 7(August):12400, 2016. ISSN 2041-1723. doi: 10.1038/ncomms12400.

S. Tajima, J. Drugowitsch, N. Patel, and A. Pouget. Optimal policy for multi-alternative decisions. *Nature Neuroscience*, 22(9):1503–1511, 2019. ISSN 15461726. doi: 10.1038/s41593-019-0453-9.

M. S. Talamali, J. A. Marshall, T. Bose, and A. Reina. Improving collective decision accuracy via time-varying cross-inhibition. *Proceedings - IEEE International Conference on Robotics and Automation*, 2019-May:9652–9659, 2019. ISSN 10504729. doi: 10.1109/ICRA.2019.8794284.

M. S. Talamali, T. Bose, M. Haire, X. Xu, J. A. Marshall, and A. Reina. Sophisticated collective foraging with minimalist agents: a swarm robotics test. *Swarm Intelligence*, 14(1):25–56, 2020. ISSN 19353820. doi: 10.1007/s11721-019-00176-9.

H. Tamaki. Glucose-Stimulated cAMP-Protein Kinase A Pathway in Yeast Saccharomyces cerevisiae. *Journal of Bioscience and Bioengineering*, 104(4):245–250, 2007. doi: 10.1263/jbb.104.245.

H. L. Tang, K. L. Yuen, H. M. Tang, and M. C. Fung. Reversibility of apoptosis in cancer cells. *British Journal of Cancer*, 100(1):118–122, 2009. ISSN 00070920. doi: 10.1038/sj.bjc.6604802.

B. Teusink, J. A. Diderich, H. V. Westerhoff, K. V. Dam, and M. C. Walsh. Intracellular Glucose Concentration in Derepressed Yeast Cells Consuming Glucose Is High Enough To Reduce the Glucose Transport Rate by 50%. *Journal of Bacteriology*, 180(3):556–562, 1998.

D. J. Timson. Galactose Metabolism in Saccharomyces cerevisiae. *Dynamic Biochemistry, Process Biotechnology and Molecular Biology*, pages 63–73, 2007. ISSN 0013-9424.

G. Tkačik and W. Bialek. Information Processing in Living Systems. *Annual Review of Condensed Matter Physics*, 7:89–117, 2016. ISSN 19475462. doi: 10.1146/annurev-conmatphys-031214-014803.

G. Tkačik and A. M. Walczak. Information transmission in genetic regulatory networks: A review. *Journal of Physics Condensed Matter*, 23(15), 2011. ISSN 1361648X. doi: 10.1088/0953-8984/23/15/153102.

B. D. Towbin, Y. Korem, A. Bren, S. Doron, R. Sorek, and U. Alon. Optimality and sub-optimality in a bacterial growth law. *Nature Communications*, 8, 2017. ISSN 20411723. doi: 10.1038/ncomms14123.

A. Tversky and D. Kahneman. Judgment under Uncertainty : Heuristics and Biases.

Science, 185(4157):1124–1131, 1974. ISSN 0036-8075. doi: 10.1126/science.185.4157.1124.

J. J. Tyson, K. C. Chen, and B. Novak. Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, 15(2):221–231, 2003. ISSN 09550674. doi: 10.1016/S0955-0674(03)00017-6.

M. Usher and J. L. McClelland. The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108(3):550–592, 2001. ISSN 0033295X. doi: 10.1037/0033-295X.108.3.550.

P. Van Hoek, J. P. Van Dijken, and J. T. Pronk. Effect of specific growth rate on fermentative capacity of baker's yeast. *Applied and Environmental Microbiology*, 64(11):4226–4233, 1998. ISSN 00992240. doi: 10.1128/AEM.64.11.4226-4233.1998.

W. A. Van Voorhies, J. Bauer, L. Brambilla, P. Duboc, and J. Francois. Robust Metabolic Responses to Varied Carbon Sources in Natural and Laboratory Strains of Saccharomyces cerevisiae. *PLoS ONE*, 7(1):e30053, jan 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0030053.

M. Vanoni, P. Sollitti, M. Goldenthal, and J. Marmur. Structure and Regulation of the Multigene Family Controlling Maltose Fermentation in Budding Yeast. *Progress in Nucleic Acid Research and Molecular Biology*, 37(C):281–322, 1989. ISSN 00796603. doi: 10.1016/S0079-6603(08)60701-1.

A. Varma and B. O. Palsson. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *Applied and Environmental Microbiology*, 60(10):3724–3731, 1994. ISSN 00992240.

A. Vazquez. How Fast Can We Grow? In *Overflow Metabolism*, page 42. 2018. ISBN 9780128122082. doi: 10.1016/b978-0-12-812208-2.00004-4.

J. W. Veening, O. A. Igoshin, R. T. Eijlander, R. Nijland, L. W. Hamoen, and O. P. Kuipers. Transient heterogeneity in extracellular protease production by Bacillus subtilis. *Molecular Systems Biology*, 4(184):1–15, 2008. ISSN 17444292. doi: 10.1038/msb.2008.18.

M. Vega, A. Riera, A. Fernández-Cid, P. Herrero, and F. Moreno. Hexokinase 2 Is an Intracellular Glucose Sensor of Yeast Cells That Maintains the Structure and Activity of Mig1 Protein Repressor Complex. *The Journal of Biological Chemistry*, 291(14):7267–7285, 2016. doi: 10.1074/jbc.M115.711408.

O. S. Venturelli, I. Zuleta, R. M. Murray, and H. El-Samad. Population Diversification in a Yeast Metabolic Program Promotes Anticipation of Environmental Shifts. *PLoS Biol*, 13(1):e1002042, 2015. doi: 10.1371/journal.pbio.1002042.

D. Vickers. Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, 13(1):37–58, 1970. ISSN 13665847. doi: 10.1080/00140137008931117.

P. Villa Martín, M. A. Muñoz, and S. Pigolotti. Bet-hedging strategies in expanding populations. *PLoS computational biology*, 15(4):e1006529, 2019. ISSN 15537358. doi: 10.1371/journal.pcbi.1006529.

D. Vogel, A. Dussutour, and J.-L. Deneubourg. Symmetry breaking and inter-clonal behavioural variability in a slime mould. *Biology Letters*, 14(12):20180504, dec 2018. ISSN 1744-9561. doi: 10.1098/rsbl.2018.0504.

A. Wald and J. Wolfowitz. Optimum Character of the Sequential Probability Ratio Test. *The Annals of Mathematical Statistics*, 19(3):326–339, 1948. ISSN 0003-4851. doi: 10.1214/aoms/1177730197.

J. Wang, E. Atolia, B. Hua, Y. Savir, R. Escalante-Chong, and M. Springer. Natural Variation in Preparation for Nutrient Depletion Reveals a Cost-Benefit Tradeoff. *PLoS Biology*, 13(1):1–31, 2015. ISSN 15457885. doi: 10.1371/journal.pbio.1002041.

X.-J. Wang. Probabilistic Decision Making by Slow Reverberation in Cortical Circuits. *Neuron*, 36(1):955–968, 2002. ISSN 14337339. doi: 10.1007/s00520-017-3845-y.

Y. Wang, M. Pierce, L. Schneper, C. G. Güldal, X. Zhang, S. Tavazoie, and J. R. Broach. Ras and Gpa2 Mediate One Branch of a Redundant Glucose Signaling Pathway in Yeast. *PLoS Biology*, 2(5):e128, may 2004. ISSN 1545-7885. doi: 10.1371/journal.pbio.0020128.

J. N. Weiss. The Hill equation revisited: uses and misuses. *The FASEB Journal*, 11:835–841, 1997.

A. Y. Weiße, D. A. Oyarzún, V. Danos, and P. S. Swain. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences of the United States of America*, 112(9):E1038–47, mar 2015. ISSN 1091-6490. doi: 10.1073/pnas.1416533112.

E. W. Weisstein. Logistic Equation, 2019. URL `http://mathworld.wolfram.com/LogisticEquation.html`. Accessed: 15-10-2019.

J. White. Variation in Water Content of Yeast Cells Caused by Varying Temperatures of Growth and by Other Cultural Conditions. *J Inst Brewing*, 68:47–50, 1952. doi: 10.1002/j.2050-0416.1952.tb02660.x.

G. Yagil and E. Yagil. On the Relation between Effector Concentration and the Rate of Induced Enzyme Synthesis. *Biophysical Journal*, 11(1):11–27, 1971. ISSN 00063495. doi: 10.1016/S0006-3495(71)86192-1.

T.-M. Yi, Y. Huang, M. I. Simon, and J. Doyle. Robust Perfect Adaptation in Bacterial Chemotaxis through Integral Feedback Control Published by : National Academy of Sciences Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences of the United States of America*, 97(9):4649–4653, 2016.

N. Yildirim, M. Santillán, D. Horike, and M. C. Mackey. Dynamics and bistability in a reduced model of the lac operon. *Chaos*, 14(2), 2004. doi: 10.1063/1.1689451.

C. You, H. Okano, S. Hui, Z. Zhang, M. Kim, C. W. Gunderson, Y.-P. Wang, P. Lenz, D. Yan, and T. Hwa. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500, 2013. doi: 10.1038/nature12446.

N. Zabzina, A. Dussutour, R. P. Mann, D. J. T. Sumpter, and S. C. Nicolis. Symmetry Restoring Bifurcation in Collective Decision-Making. *PLoS Comput Biol*, 10(12):e1003960, 2014. doi: 10.1371/journal.pcbi.1003960.

S. Zaman, S. I. Lippman, L. Schneper, N. Slonim, and J. R. Broach. Glucose regulates transcription in yeast through a network of signaling pathways. *Molecular Systems Biology*, 5, 2009. doi: 10.1038/msb.2009.2.

Appendix A

Supplementary information for Chapter 4

Here I present the results from the best fits to each data set by each one of the four models. I only omit the results that are presented in Chapter 4. The results are divided first by model and then by data set. In each model section I present a table with the initial conditions used for the fits to the data sets, as well as a table were I present the parameter sets that scored the lowest overall RMSE score in each data set fit (all three data sets, maltose, galactose and PR1). In each subsection of each model I present two tables; the first table displays the results of the different fits. The fits differ only in the weighting factor used, and they are ranked based on their overall RMSE score. Secondly, I include a table with the RMSE scores per data type (fluorescence and optical density) and section of the curve, corresponding to the lowest scoring (overall RMSE) fit. In the subsections "fits to all three data sets" and "Fits to all three data sets with PR1 parameters", in addition to the tables already mentioned, I present a table with the RMSE scores per data set (PR1, maltose and galactose) and section of the curve. Additionally, I present each a different table with the RMSE scores per data type (fluorescence and optical density) and section of the curve per data set (PR1, maltose and galactose). Finally, the subsection "Fits

to all three data sets with PR1 parameters" of each model, includes a table with the initial conditions yielded by applying a grid search with the lowest scoring PR1 parameters.

A.1 Hill function biomass model: RSO induction term

Table A.1: Initial values of the state variables used.

	B	G	S_1	S_2	P_1	P_2
PR1	0.1	0.3	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$
Maltose	0.1	0.2	$1e^{-4}$	0	$5e^{-3}$	0
Galactose	0.1	0.1	0	$1e^{-3}$	0	$2e^{-3}$

Table A.2: Different parameterisations resulting from fitting the model to different experimental data sets. These parameterisations correspond to those with the lowest overall RMSE score of their corresponding experimental data set. On each row corresponding to the parameter values, the lightest green indicates the lowest value, whereas the deepest green indicates the highest value of that parameter out of all the different parameterisations. Likewise, a red-blue colormap is used to differentiate between the highest and lowest normalised standard deviation, respectively.

Parameter	Normalised std.	Mean	All data	Maltose	Galactose	PR1	Initial Parameters
A_1	1.03e+00	8.64e-03	1.48e-02	6.33e-05	2.02e-02	1.23e-04	8.02e-03
V_{S1}	3.34e-04	6.00e+02	6.00e+02	6.00e+02	6.00e+02	6.00e+02	6.00e+02
V_{P1}	1.68e+00	8.32e-02	9.29e-03	3.32e-01	3.66e-02	3.30e-02	5.50e-03
V_{G1}	7.44e-02	9.58e-01	9.15e-01	1.08e+00	9.30e-01	9.29e-01	9.30e-01
K_{inh1}	4.95e-03	3.14e+00	3.15e+00	3.12e+00	3.15e+00	3.15e+00	3.15e+00
K_{S1}	1.54e+00	1.95e-01	6.06e-02	7.32e-01	6.00e-02	6.12e-02	6.00e-02
I_1	2.03e-04	1.30e+00	1.30e+00	1.30e+00	1.30e+00	1.30e+00	1.30e+00
K_{P1}	2.05e+00	6.70e-01	2.70e-02	3.13e+00	4.00e-02	1.15e-01	4.00e-02
Dcy_{P1}	2.22e+00	1.21e-04	1.00e-06	1.00e-06	1.01e-06	1.00e-06	6.00e-04
Y_1	3.77e-01	1.24e-01	7.73e-02	6.87e-02	1.60e-01	1.54e-01	1.60e-01
A_2	2.12e+00	3.77e-02	9.80e-05	1.80e-01	1.18e-06	1.00e-06	8.02e-03
V_{S2}	4.33e-05	5.00e+02	5.00e+02	5.00e+02	5.00e+02	5.00e+02	5.00e+02
V_{P2}	6.17e-01	9.64e-03	1.11e-02	1.15e-02	3.66e-03	1.79e-02	4.00e-03
V_{G2}	5.63e-01	1.53e+00	1.13e+00	1.15e+00	3.07e+00	1.15e+00	1.15e+00
K_{inh2}	2.60e-02	2.10e+00	2.08e+00	2.08e+00	2.20e+00	2.08e+00	2.08e+00
K_{S2}	1.46e-01	1.08e-01	1.10e-01	9.50e-02	1.34e-01	1.08e-01	9.50e-02
I_2	9.90e-05	2.42e+00	2.42e+00	2.42e+00	2.42e+00	2.42e+00	2.42e+00
K_{P2}	1.54e+00	1.58e-02	5.44e-04	1.00e-02	2.01e-06	5.84e-02	1.00e-02
Dcy_{P2}	2.18e+00	2.12e+00	1.98e-02	1.21e-01	1.04e+01	1.91e-02	6.00e-02
Y_2	1.04e+00	3.02e-01	6.96e-02	2.00e-01	8.55e-01	1.83e-01	2.00e-01
V_B	1.15e+00	1.47e+00	8.32e-01	2.96e-01	4.45e+00	8.56e-01	9.00e-01
K_G	9.41e-01	3.49e+00	1.87e+00	2.52e+00	9.35e+00	1.87e+00	1.85e+00
M	2.11e+00	3.60e+00	2.15e-01	8.12e-02	1.72e+01	2.13e-01	3.00e-01
Dcy_B	9.85e-01	1.90e-02	1.00e-06	4.32e-02	1.19e-04	2.63e-02	2.52e-02

Fits to all three data sets

Table A.3: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 1 for all three experimental data sets.

Weight factor	1.0	2.0	5.0	10.0	0.5	100.0	0.1	I.P.	0.01
Overall RMSE	0.104091	0.104123	0.107268	0.109684	0.110372	0.11147	0.134404	0.170686	0.248289
Lag RMSE	0.0555242	0.0464497	0.0570164	0.0528497	0.0636339	0.0489081	0.0871747	0.0913923	0.234673
Exponential RMSE	0.143232	0.144048	0.145867	0.147616	0.163536	0.148208	0.209096	0.232259	0.354643
Stationary RMSE	0.0943755	0.0980675	0.0999557	0.107275	0.0758468	0.11366	0.0536015	0.158442	0.0640245

Table A.4: RMSE scores divided by curve phase and experimental data set (single sugar and sugar mixture). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per experiment**	
PR1	0.057467	0.135772	0.088307	**0.099221**	
Maltose	0.050474	0.127136	0.062392	**0.086802**	
Galactose	0.058299	0.164160	0.122594	**0.122986**	
RMSE per phase	**0.055524**	**0.143232**	**0.094376**	**0.104091**	**Total RMSE score**

Table A.5: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.072323	0.152331	0.041313	**0.100237**	
GFP	0.059821	0.144133	0.120683	**0.113896**	
OD	0.033138	0.106408	0.084401	**0.080714**	
RMSE per phase	**0.057467**	**0.135772**	**0.088307**	**0.099221**	**Total RMSE score for the PR1 fit**

Table A.6: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.058549	0.140384	0.043381	**0.091319**	
OD	0.040832	0.112336	0.076835	**0.082037**	
RMSE per phase	**0.050474**	**0.127136**	**0.062392**	**0.086802**	**Total RMSE score for the maltose fit**

Table A.7: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
GFP	0.063330	0.163339	0.161337	**0.137501**	
OD	0.052790	0.164978	0.063474	**0.106511**	
RMSE per phase	**0.058299**	**0.164160**	**0.122594**	**0.122986**	**Total RMSE score for the galactose fit**

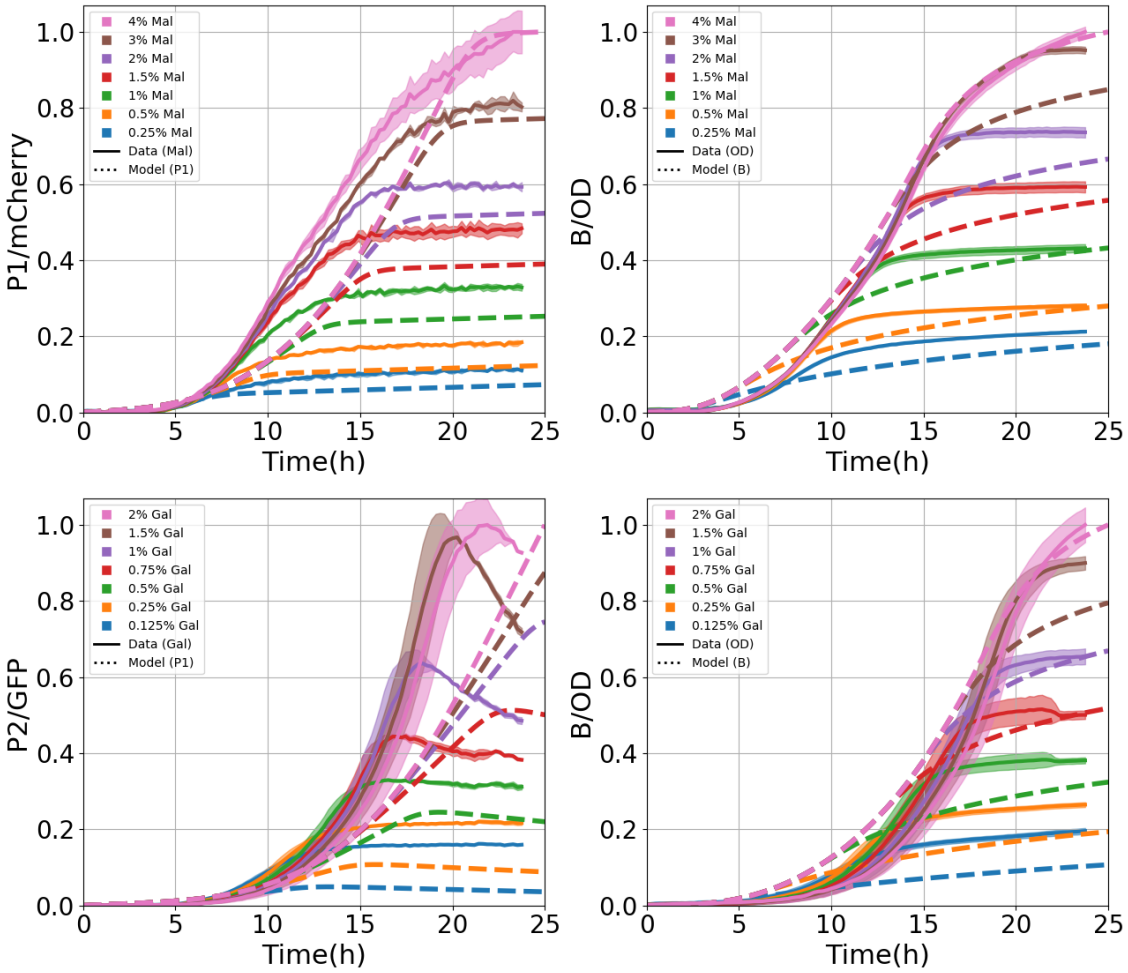


Figure A.1: Fit to the single sugar experimental data sets. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: The maltose data set. **Bottom**: galactose data set. Graphs to the left correspond to the fit to the fluorescence data (mCherry in the case of maltose, GFP in the case of galactose), whilst the graphs to the right present the results of the fit to the absorbance data (O.D.).

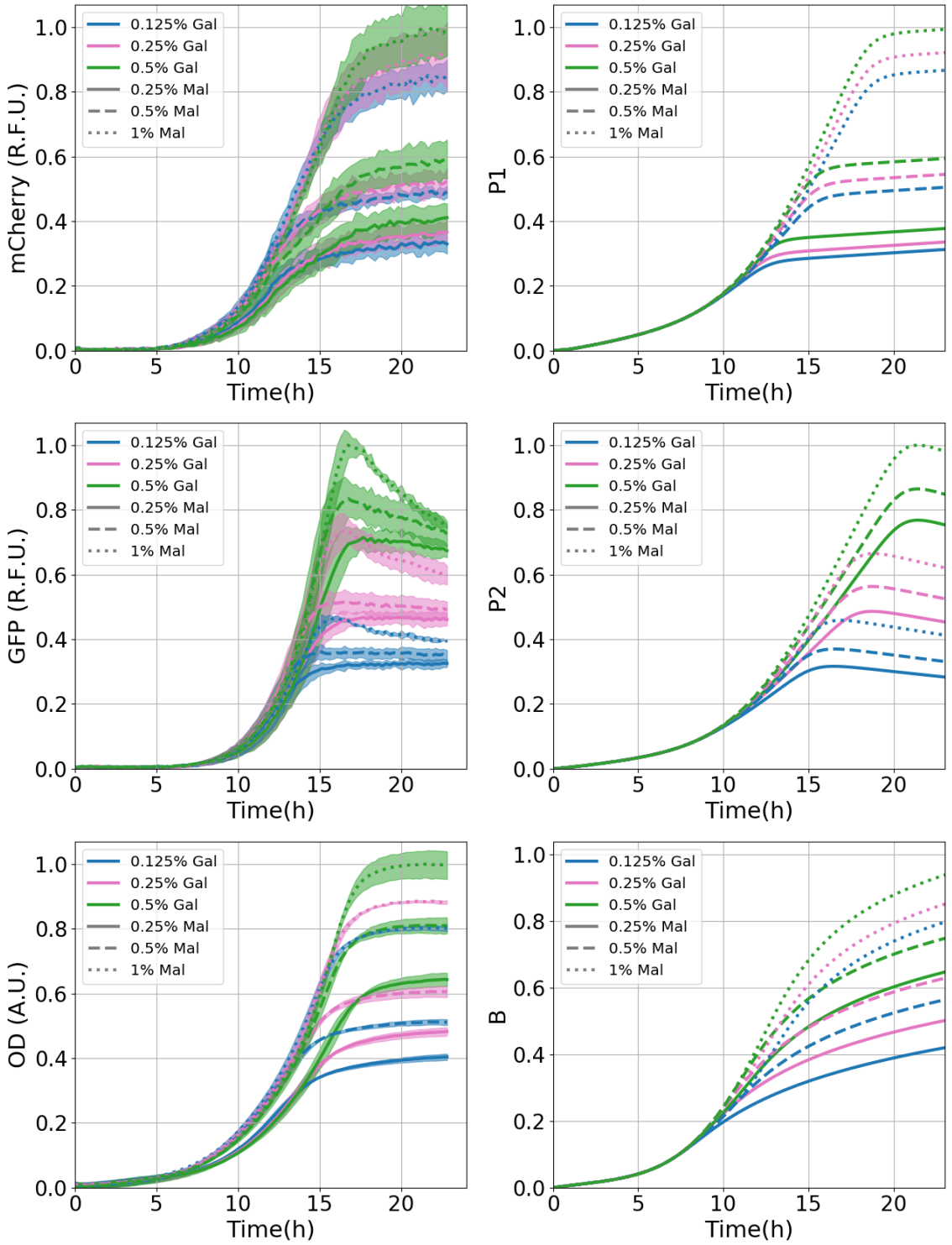


Figure A.2: Fit to the binary-sugar mixture experimental data set PR1. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

Fits to the maltose data set

Table A.8: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 1 for all the maltose experimental data set.

Weight factor	1.0	10.0	0.5	5.0	0.01	2.0	0.1	100.0	I.P.
Overall RMSE	0.0474082	0.0519579	0.0756722	0.0763427	0.0785248	0.0896218	0.0942474	0.0977484	0.147893
Lag RMSE	0.0213322	0.0172972	0.0603167	0.0524435	0.0674931	0.0786919	0.0914635	0.0683482	0.0834327
Exponential RMSe	0.0687776	0.0631326	0.109674	0.0913211	0.114571	0.111905	0.130558	0.104573	0.235122
Stationary RMSE	0.0394613	0.0617572	0.0388902	0.0799671	0.0285774	0.0733556	0.0351678	0.114269	0.0580856

Table A.9: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.025416	0.067587	0.039177	**0.047430**	
OD	0.016253	0.069948	0.039743	**0.047386**	
RMSE per phase	**0.021332**	**0.068778**	**0.039461**	**0.047408**	**Total RMSE score for the maltose fit**

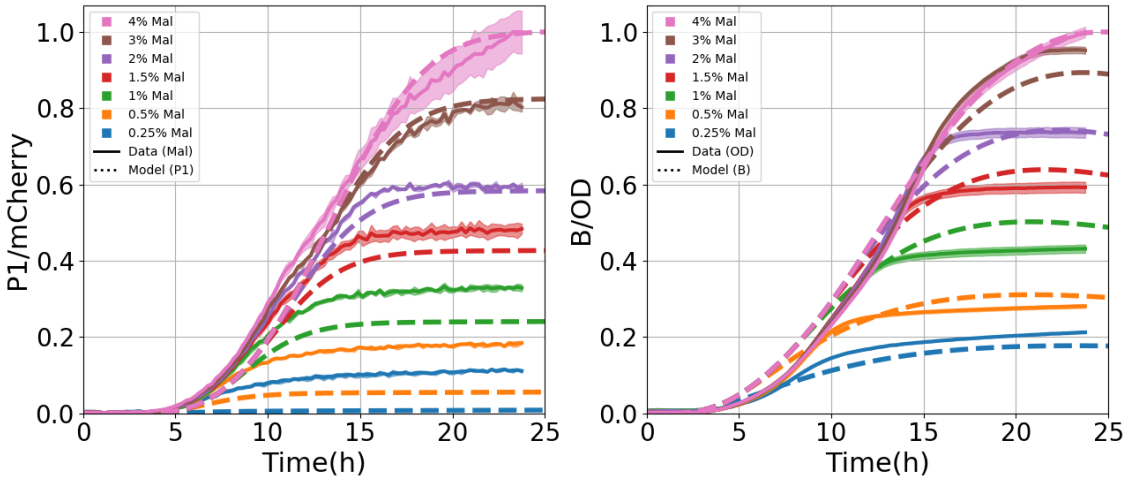


Figure A.3: Fit to the maltose experimental data set. This corresponds to the parameterisation with the lowest overall RMSE score. The left graph to the correspond to the fit to the mCherry fluorescence data, whilst the graph to the right presents the results of the fit to the absorbance data (O.D.).

Fits to the galactose data set

Table A.10: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 10 for the galactose experimental data set.

Weight factor	10.0	2.0	1.0	5.0	0.5	100.0	0.1	0.01	I.P.
Overall RMSE	0.0573314	0.0628947	0.0747704	0.0996862	0.106156	0.128901	0.16012	0.222813	0.235594
Lag RMSE	0.02382	0.0399613	0.0211453	0.0224111	0.0564365	0.0361024	0.11183	0.218037	0.125315
Exponential RMSE	0.0658191	0.0760077	0.0970452	0.11229	0.146283	0.135429	0.243411	0.312395	0.296449
Stationary RMSE	0.0704352	0.0670311	0.083108	0.129231	0.0960378	0.173788	0.0718323	0.061699	0.250855

Table A.11: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 10).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
GFP	0.032834	0.074313	0.078624	**0.065274**	
OD	0.007532	0.056053	0.061160	**0.048094**	
RMSE per phase	**0.023820**	**0.065819**	**0.070435**	**0.057331**	**Total RMSE score for the galactose fit**

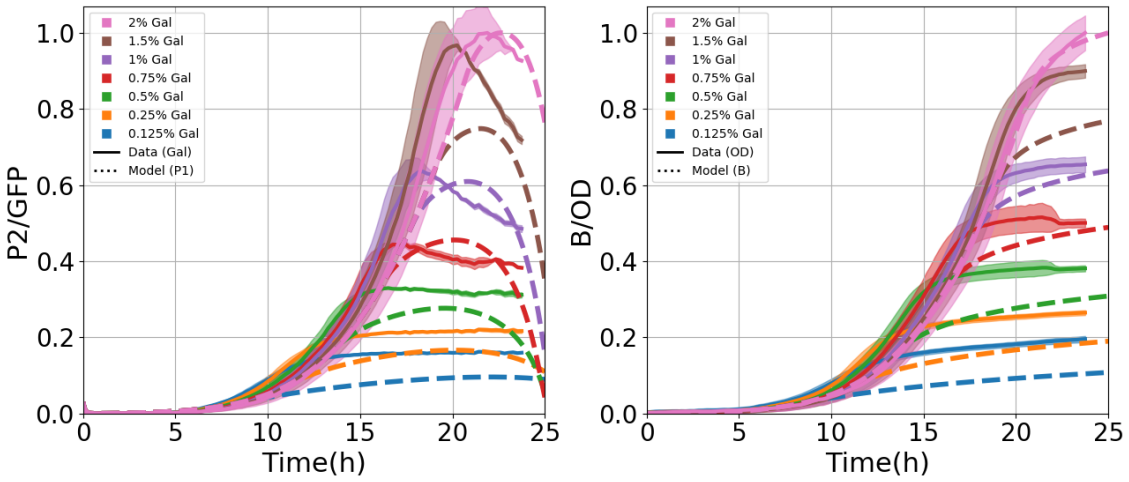


Figure A.4: Fit to the galactose experimental data set. This corresponds to the parameterisation with the lowest overall RMSE score. The left graph to the correspond to the fit to the GFP fluorescence data, whilst the graph to the right presents the results of the fit to the absorbance data (O.D.).

Fits to the PR1 data set

Table A.12: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 1 for the PR1 experimental data set.

Weight factor	1.0	2.0	0.1	5.0	10.0	0.5	100.0	I.P.	0.01
Overall RMSE	0.0574409	0.063546	0.0685307	0.0710509	0.0758844	0.0790331	0.0792223	0.100118	0.284653
Lag RMSE	0.011423	0.0173582	0.0439879	0.0231797	0.0208628	0.0402735	0.0197308	0.0489165	0.256925
Exponential RMSE	0.0797962	0.0861364	0.100539	0.0890029	0.0898743	0.113511	0.0904023	0.136631	0.419075
Stationary RMSE	0.0583136	0.0662835	0.0452363	0.0817672	0.0936092	0.0650541	0.101324	0.0949208	0.0380404

Table A.13: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.010859	0.106722	0.068294	**0.073420**	
GFP	0.006505	0.059918	0.073036	**0.054671**	
OD	0.015206	0.064207	0.014254	**0.038974**	
RMSE per phase	**0.011423**	**0.079796**	**0.058314**	**0.057441**	**Total RMSE score for the PR1 fit**

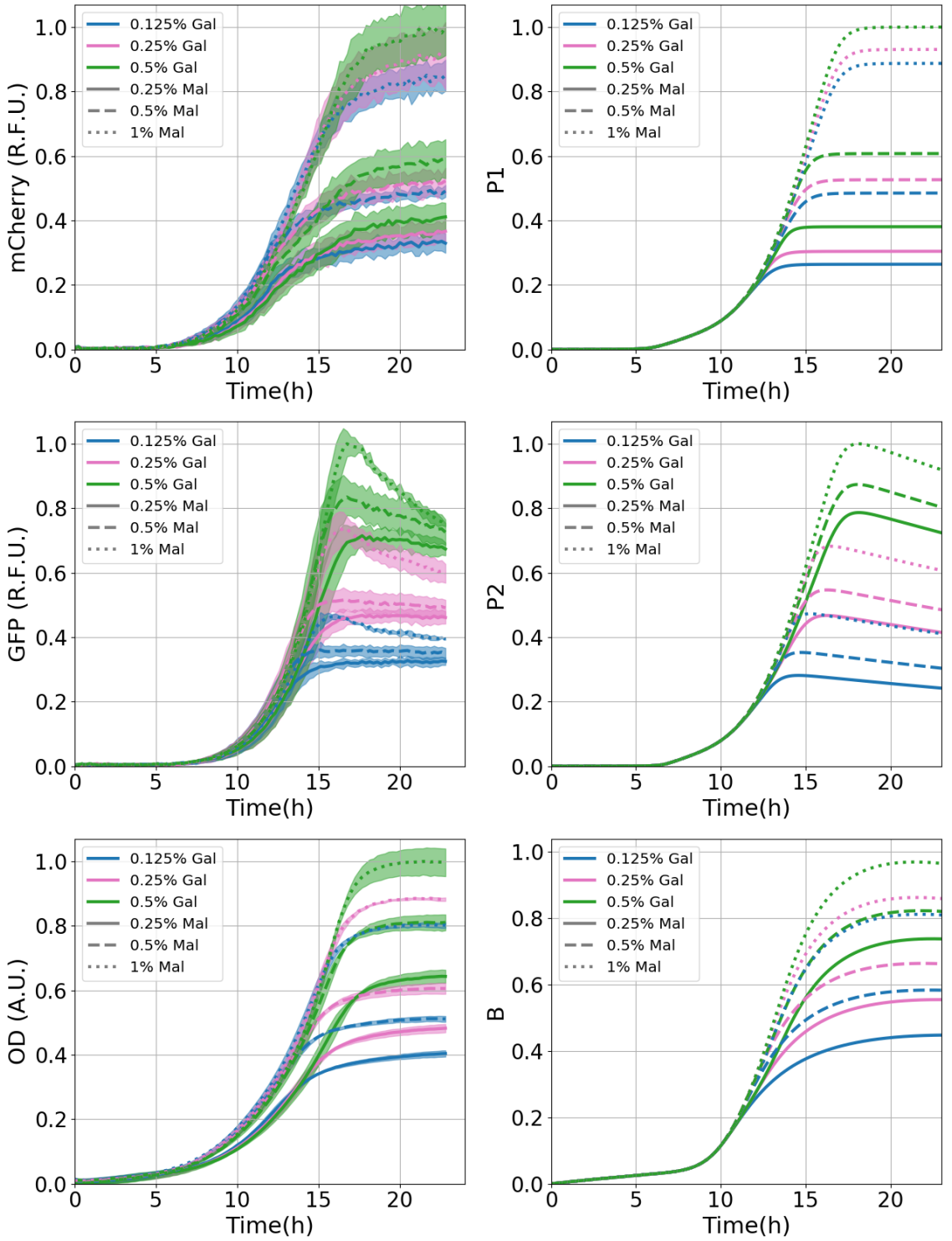


Figure A.5: Fit to binary-sugar mixture experimental data set PR1. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

Fits to all three data sets with PR1 parameters

Table A.14: Initial values of state variables resulting from the grid search performed with the lowest overall RMSE PR1 parameterisation.

	B	G	S_1	S_2	P_1	P_2
PR1	0.08	0.5	$1e^{-6}$	$1e^{-6}$	$1e^{-6}$	$5.5e^{-6}$
Maltose	0.55	0.325	$7.75e^{-4}$	0	$5.5e^{-6}$	0
Galactose	0.325	0.325	0	$7.75e^{-4}$	0	$5.5e^{-6}$

Table A.15: RMSE scores divided by curve phase for all three experimental data sets.

Overall RMSE	0.0633421
Lag RMSE	0.0172272
Exponential RMSE	0.0885707
Stationary RMSE	0.0624110

Table A.16: RMSE scores divided by curve phase and experimental data set (single sugar and sugar mixture).

	Lag phase	Exponential phase	Stationary phase	RMSE per experiment	
PR1	0.025838	0.073945	0.058022	**0.056279**	
Maltose	0.009002	0.102244	0.059507	**0.068498**	
Galactose	0.011904	0.087251	0.069121	**0.064633**	
RMSE per phase	**0.017227**	**0.088571**	**0.062411**	**0.063342**	**Total RMSE score**

Table A.17: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.015178	0.098711	0.069259	**0.070169**	
GFP	0.011111	0.053398	0.072158	**0.052222**	
OD	0.040608	0.061714	0.009808	**0.043026**	
RMSE per phase	**0.025838**	**0.073945**	**0.058022**	**0.056279**	**Total RMSE score for the PR1 fit**

Table A.18: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.010847	0.105891	0.061434	**0.070957**	
OD	0.006664	0.098462	0.057516	**0.065947**	
RMSE per phase	**0.009002**	**0.102244**	**0.059507**	**0.068498**	**Total RMSE score for the maltose fit**

Table A.19: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
GFP	0.006547	0.075090	0.094461	**0.069772**	
OD	0.015509	0.097913	0.025152	**0.059048**	
RMSE per phase	**0.011904**	**0.087251**	**0.069121**	**0.064633**	**Total RMSE score for the galactose fit**

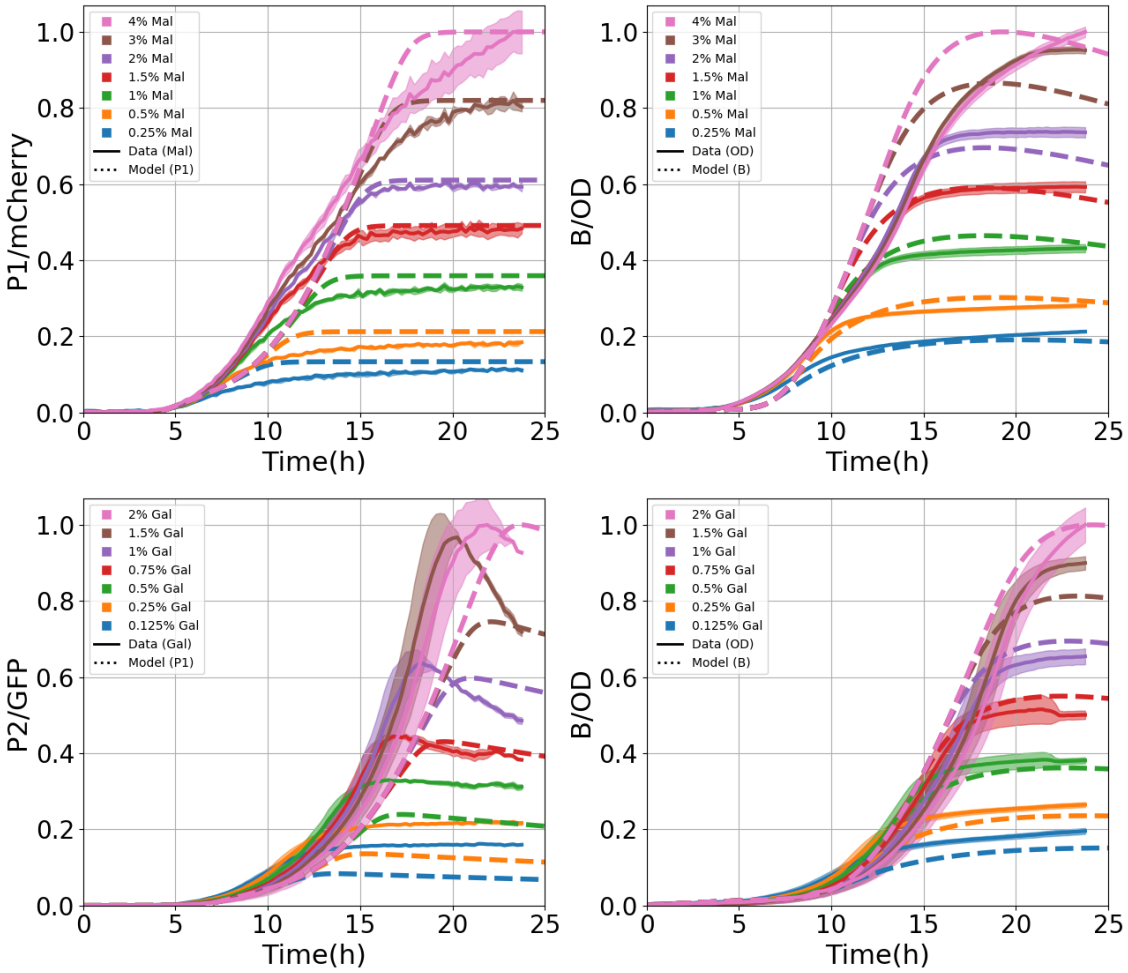


Figure A.6: Fit to the single sugar experimental data sets with the PR1 parameterisation with the lowest overall RMSE score after performing a grid search for new initial state variables values (table A.14). **Top**: The maltose data set. **Bottom**: galactose data set. Graphs to the left correspond to the fit to the fluorescence data (mCherry in the case of maltose, GFP in the case of galactose), whilst the graphs to the right present the results of the fit to the absorbance data (O.D.).

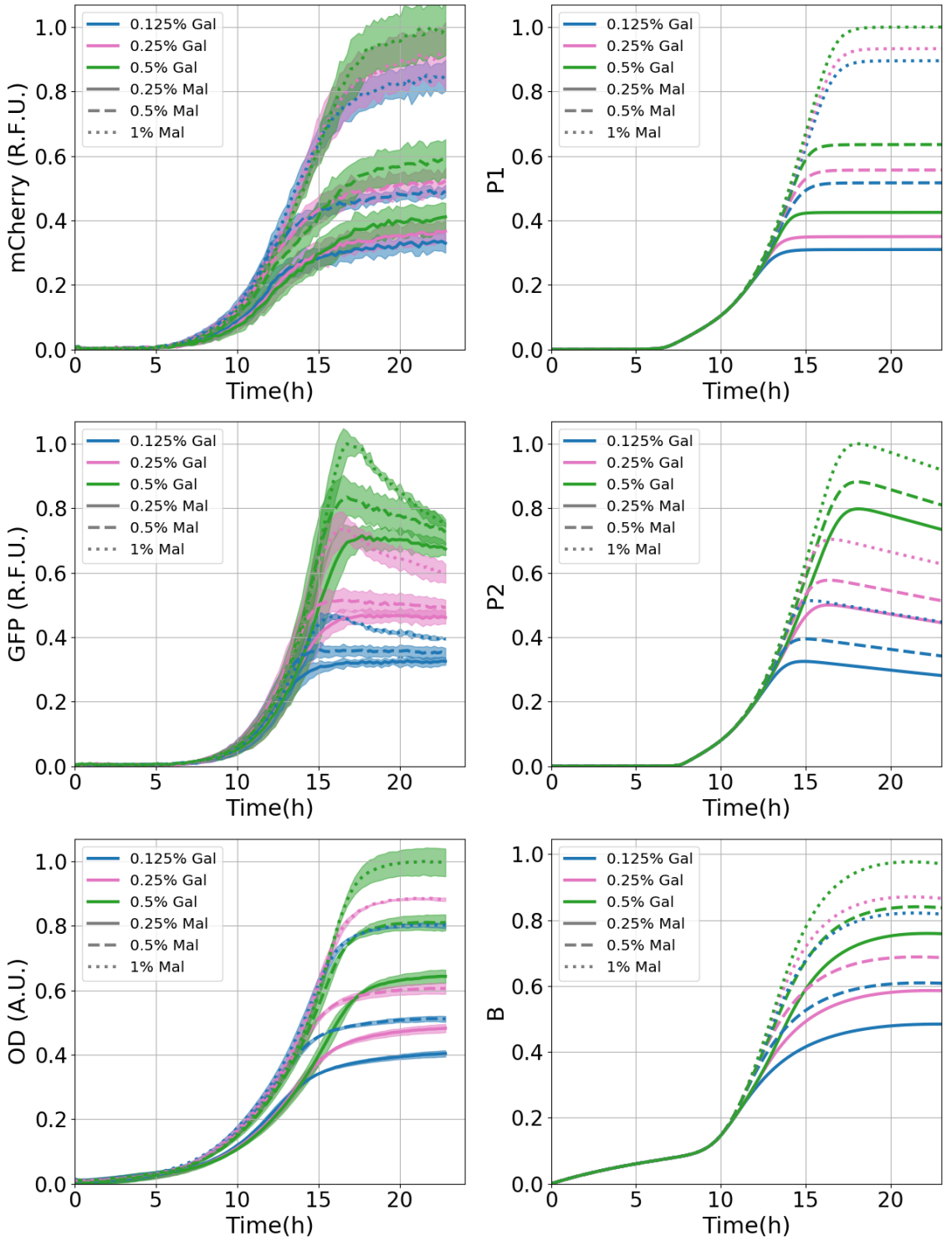


Figure A.7: Fit to the binary-sugar mixture experimental data set PR1 with the PR1 parameterisation with the lowest overall RMSE score after performing a grid search for new initial state variables values (table A.14). **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

A.2 Hill function biomass model: RFO induction term

Table A.20: Initial values of the state variables used.

	B	G	S_1	S_2	P_1	P_2
PR1	0.1	0.3	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$
Maltose	0.1	0.1	$1e^{-4}$	0	$1.5e^{-3}$	0
Galactose	0.1	0.1	0	$1e^{-4}$	0	$1.5e^{-3}$

Table A.21: Different parameterisations resulting from fitting the model to different experimental data sets. These parameterisations correspond to those with the lowest overall RMSE score of their corresponding experimental data set. On each row corresponding to the parameter values, the lightest green indicates the lowest value, whereas the deepest green indicates the highest value of that parameter out of all the different parameterisations. Likewise, a red-blue colormap is used to differentiate between the highest and lowest normalised standard deviation, respectively.

Parameter	Normalised std.	Mean	All data	Maltose	Galactose	PR1	Initial Parameters
A_1	8.48e-01	1.01e-02	2.03e-02	7.47e-03	3.62e-05	1.24e-02	8.02e-03
V_{S1}	9.42e-06	6.00e+02	6.00e+02	6.00e+02	6.00e+02	6.00e+02	6.00e+02
V_{P1}	1.38e+00	1.18e-02	5.18e-03	3.58e-02	3.62e-05	6.15e-03	6.50e-03
V_{G1}	3.11e-02	9.18e-01	8.93e-01	8.97e-01	9.30e-01	9.53e-01	9.30e-01
K_{inh1}	1.17e-01	6.88e-01	7.77e-01	5.83e-01	6.80e-01	7.13e-01	6.80e-01
K_{S1}	3.58e-01	9.19e-02	1.09e-01	1.30e-01	6.00e-02	6.91e-02	6.00e-02
I_1	6.92e-01	3.67e-01	6.79e-01	1.60e-01	1.60e-01	4.70e-01	1.60e-01
K_{P1}	9.14e-01	2.75e-01	1.39e-01	6.49e-01	2.00e-01	1.14e-01	2.00e-01
Dcy_{P1}	1.94e+00	1.53e-03	1.00e-06	1.26e-04	5.99e-03	1.00e-06	6.00e-04
Y_1	7.33e-01	3.01e-01	6.14e-01	1.31e-01	1.60e-01	3.00e-01	1.60e-01
A_2	1.24e+00	3.15e-03	4.38e-03	8.13e-03	9.63e-05	1.00e-06	8.02e-03
V_{S2}	1.07e-01	5.28e+02	5.00e+02	5.00e+02	6.13e+02	5.00e+02	5.00e+02
V_{P2}	9.03e-01	5.49e-03	1.15e-02	1.15e-04	3.11e-03	7.23e-03	4.00e-03
V_{G2}	1.25e+00	3.02e+00	1.12e+00	1.15e+00	8.66e+00	1.15e+00	1.15e+00
K_{inh2}	9.68e-02	8.01e-01	9.14e-01	7.50e-01	7.93e-01	7.49e-01	7.50e-01
K_{S2}	1.09e+00	3.67e-01	3.19e-01	9.50e-02	9.50e-01	1.06e-01	9.50e-02
I_2	1.37e+00	3.26e-01	9.83e-01	4.00e-02	4.00e-02	2.43e-01	4.00e-02
K_{P2}	1.43e+00	3.72e-01	1.15e+00	8.00e-02	1.33e-04	2.53e-01	8.00e-02
Dcy_{P2}	1.70e+00	1.70e-01	1.00e-06	6.03e-02	6.00e-01	1.85e-02	6.00e-02
Y_2	3.88e-01	1.58e-01	1.01e-01	2.00e-01	1.09e-01	2.20e-01	2.00e-01
V_B	1.01e+00	1.27e+00	6.53e-01	3.63e-01	3.16e+00	9.08e-01	9.00e-01
K_G	1.28e+00	5.28e+00	1.95e+00	1.95e+00	1.54e+01	1.81e+00	1.85e+00
M	1.47e+00	9.38e-01	2.32e-01	2.32e-01	3.00e+00	2.87e-01	3.00e-01
Dcy_B	3.84e-01	2.49e-02	1.37e-02	3.71e-02	2.52e-02	2.38e-02	2.52e-02

Fits to all three data sets

Table A.22: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 0.5 for all three experimental data sets.

Weight factor	0.5	10.0	1.0	2.0	5.0	0.1	100.0	I.P.	0.01
Overall RMSE	0.0824604	0.0848835	0.0885277	0.0928857	0.0976105	0.0989113	0.135942	0.155876	0.287217
Lag RMSE	0.0559184	0.0372071	0.0622602	0.0558478	0.0550506	0.0755177	0.0672058	0.0941286	0.24625
Exponential RMSE	0.113987	0.0994875	0.118751	0.120908	0.127754	0.147645	0.162968	0.19334	0.429766
Stationary RMSE	0.0654159	0.101654	0.0743863	0.0902524	0.0960818	0.0429933	0.156094	0.163253	0.0462863

Table A.23: RMSE scores divided by curve phase and experimental data set (single sugar and sugar mixture). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	Exponential phase	Stationary phase	RMSE per experiment	
PR1	0.023059	0.077660	0.064785	**0.059889**	
Maltose	0.040313	0.099202	0.035992	**0.065221**	
Galactose	0.084993	0.152010	0.085704	**0.112066**	
RMSE per phase	**0.055918**	**0.113987**	**0.065416**	**0.082460**	**Total RMSE score**

Table A.24: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.027131	0.064641	0.031381	**0.044345**	
GFP	0.027954	0.096470	0.095392	**0.079974**	
OD	0.008811	0.067886	0.050069	**0.048966**	
RMSE per phase	**0.023059**	**0.077660**	**0.064785**	**0.059889**	**Total RMSE score for the PR1 fit**

Table A.25: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.049503	0.096209	0.039322	**0.066465**	
OD	0.028279	0.102106	0.032320	**0.063953**	
RMSE per phase	**0.040313**	**0.099202**	**0.035992**	**0.065221**	**Total RMSE score for the maltose fit**

Table A.26: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
GFP	0.114152	0.181384	0.109105	**0.138846**	
OD	0.037641	0.115386	0.052789	**0.076414**	
RMSE per phase	**0.084993**	**0.152010**	**0.085704**	**0.112066**	**Total RMSE score for the galactose fit**

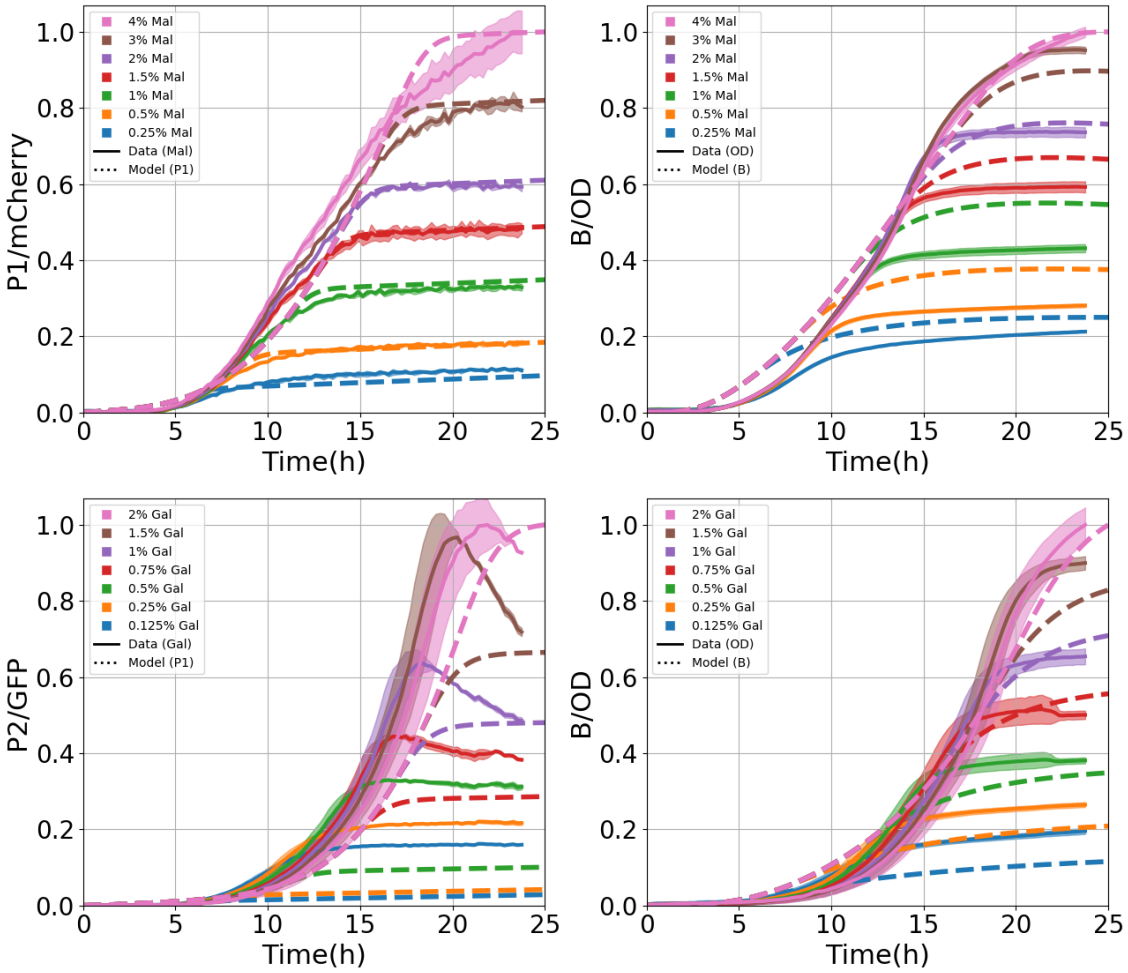


Figure A.8: Fit to the single sugar experimental data sets. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: The maltose data set. **Bottom**: galactose data set. Graphs to the left correspond to the fit to the fluorescence data (mCherry in the case of maltose, GFP in the case of galactose), whilst the graphs to the right present the results of the fit to the absorbance data (O.D.).

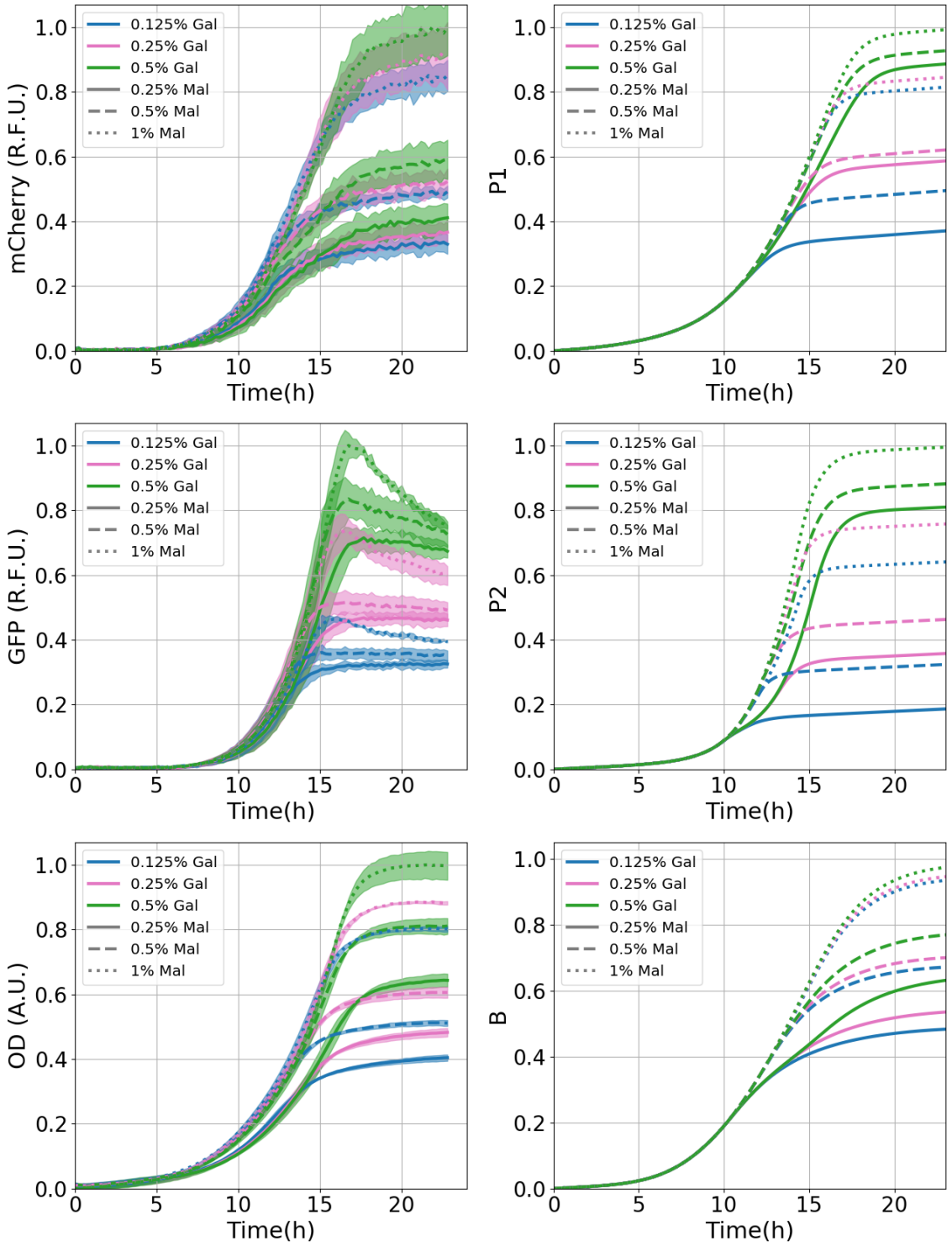


Figure A.9: Fit to the binary-sugar mixture experimental data set PR1. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

Fits to the maltose data set

Table A.27: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 1 for the maltose experimental data set.

Weight factor	1.0	2.0	0.5	0.1	5.0	10.0	I.P.
Overall RMSE	0.0513405	0.0515411	0.0518283	0.0539392	0.0605028	0.0629165	0.090673
Lag RMSE	0.0458228	0.0443715	0.0445128	0.0390225	0.0411822	0.0431632	0.0418205
Exponential RMSe	0.0662702	0.0646254	0.0696765	0.0783662	0.0750381	0.0723548	0.145038
Stationary RMSE	0.037631	0.0427105	0.0349616	0.0326237	0.0604571	0.0691171	0.0433554

Table A.28: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.064566	0.074782	0.038295	**0.061176**	
OD	0.005543	0.056490	0.036955	**0.039105**	
RMSE per phase	**0.045823**	**0.066270**	**0.037631**	**0.051341**	**Total RMSE score for the maltose fit**

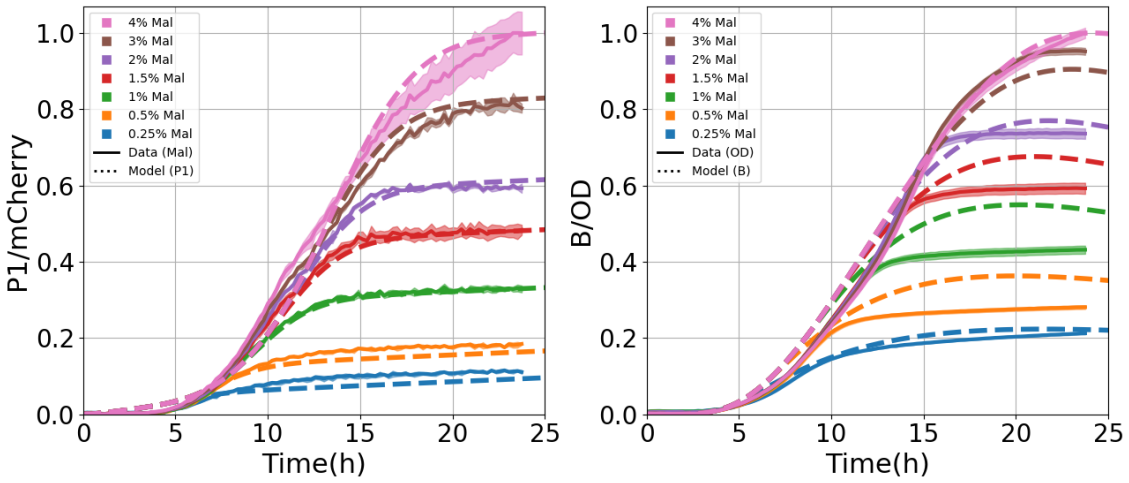


Figure A.10: Fit to the maltose experimental data set. This corresponds to the parameterisation with the lowest overall RMSE score. The left graph to the correspond to the fit to the mCherry fluorescence data, whilst the graph to the right presents the results of the fit to the absorbance data (O.D.).

Fits to the galactose data set

Table A.29: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 5 for the galactose experimental data set.

Weight factor	5.0	10.0	1.0	2.0	0.5	0.1	I.P.
Overall RMSE	0.0695431	0.0810662	0.103317	0.110676	0.118912	0.149314	0.237659
Lag RMSE	0.0181446	0.027247	0.0617838	0.0623169	0.0789625	0.127269	0.155669
Exponential RMSE	0.0860465	0.0952593	0.137743	0.134135	0.163249	0.219433	0.275898
Stationary RMSE	0.0823134	0.099491	0.0960887	0.121949	0.0976452	0.0503557	0.262856

Table A.30: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 5).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
GFP	0.013162	0.091231	0.099043	**0.078115**	
OD	0.022027	0.080528	0.061168	**0.059754**	
RMSE per phase	**0.018145**	**0.086047**	**0.082313**	**0.069543**	**Total RMSE score for the galactose fit**

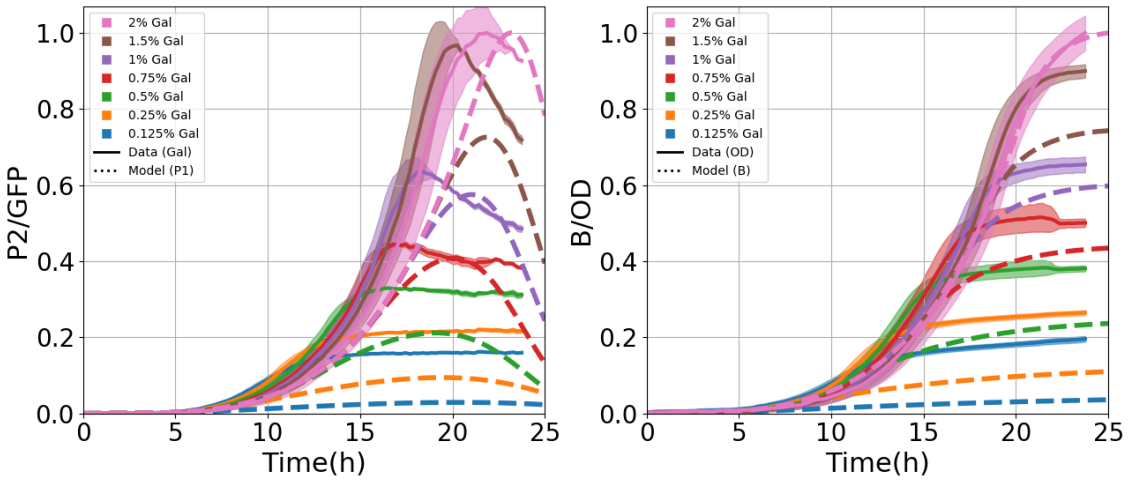


Figure A.11: Fit to the galactose experimental data set. This corresponds to the parameterisation with the lowest overall RMSE score. The left graph to the correspond to the fit to the GFP fluorescence data, whilst the graph to the right presents the results of the fit to the absorbance data (O.D.).

Fits to the PR1 data set

Results in section 4.3.1 of the main text.

Fits to all three data sets with PR1 parameters

Results in section 4.3.2 of the main text.

A.3 Logistic equation biomass model: RSO induction term

Table A.31: Initial values of the state variables used.

	B	G	S_1	S_2	P_1	P_2
PR1	0.3	0.3	$1e^{-7}$	$1e^{-7}$	$5.5e^{-6}$	$5e^{-6}$
Maltose	0.3	0.3	$1e^{-7}$	0	$5.5e^{-4}$	0
Galactose	0.3	0.3	0	$1e^{-7}$	0	$5.5e^{-6}$

Table A.32: Different parameterisations resulting from fitting the model to different experimental data sets. These parameterisations correspond to those with the lowest overall RMSE score of their corresponding experimental data set. On each row corresponding to the parameter values, the lightest green indicates the lowest value, whereas the deepest green indicates the highest value of that parameter out of all the different parameterisations. Likewise, a red-blue colormap is used to differentiate between the highest and lowest normalised standard deviation, respectively.

Parameter	Normalised std.	Mean	All data	Maltose	Galactose	PR1	Initial Parameters
A_1	1.07e+00	4.03e-02	1.00e-06	8.18e-02	6.37e-03	7.32e-02	8.02e-03
V_{S1}	1.13e-05	6.00e+02	6.00e+02	6.00e+02	6.00e+02	6.00e+02	6.00e+02
V_{P1}	1.57e+00	1.66e-02	6.42e-03	5.58e-02	3.19e-03	1.22e-03	5.50e-03
V_{G1}	2.15e-03	3.99e+00	3.98e+00	3.98e+00	3.98e+00	4.00e+00	3.98e+00
K_{inh1}	2.17e-04	3.50e+00	3.50e+00	3.50e+00	3.50e+00	3.50e+00	3.50e+00
K_{S1}	5.26e-02	6.09e-02	6.54e-02	6.05e-02	6.00e-02	5.78e-02	6.00e-02
I_1	6.30e-04	1.72e+00	1.71e+00	1.72e+00	1.72e+00	1.72e+00	1.72e+00
K_{P1}	4.46e-01	4.79e-02	7.36e-02	4.68e-02	5.00e-02	2.14e-02	5.00e-02
Dcy_{P1}	1.99e+00	1.61e-04	1.00e-06	1.00e-06	6.39e-04	1.00e-06	6.00e-04
Y_1	2.92e-01	1.34e-01	1.47e-01	7.57e-02	1.60e-01	1.52e-01	1.60e-01
A_2	2.00e+00	1.88e-03	1.00e-06	7.52e-03	1.00e-06	8.90e-06	8.02e-03
V_{S2}	3.81e-07	5.00e+02	5.00e+02	5.00e+02	5.00e+02	5.00e+02	5.00e+02
V_{P2}	6.79e-01	3.78e-03	6.51e-03	2.96e-03	5.03e-03	6.09e-04	2.00e-03
V_{G2}	2.23e-01	3.38e+00	3.00e+00	3.00e+00	3.00e+00	4.51e+00	4.50e+00
K_{inh2}	3.73e-01	4.23e-01	5.01e-01	5.00e-01	5.04e-01	1.86e-01	2.00e-01
K_{S2}	3.63e-02	9.70e-02	9.42e-02	9.50e-02	1.02e-01	9.66e-02	9.50e-02
I_2	3.21e-02	5.53e+00	5.45e+00	5.45e+00	5.45e+00	5.80e+00	5.80e+00
K_{P2}	1.78e-01	5.16e-02	5.48e-02	5.00e-02	3.99e-02	6.17e-02	5.00e-02
Dcy_{P2}	7.88e-01	2.76e-02	2.03e-02	6.00e-02	1.32e-02	1.70e-02	6.00e-02
Y_2	1.90e-01	1.04e-01	9.63e-02	1.10e-01	1.27e-01	8.09e-02	1.10e-01
V_B	2.30e-01	4.04e-01	3.86e-01	3.20e-01	3.75e-01	5.36e-01	4.00e-01
K_G	1.76e-04	1.00e+01	1.00e+01	1.00e+01	1.00e+01	1.00e+01	1.00e+01
M	1.21e-01	3.17e-01	2.88e-01	3.22e-01	2.90e-01	3.70e-01	3.00e-01
Dcy_B	1.66e+00	2.41e-04	8.34e-04	1.00e-06	1.00e-06	1.28e-04	1.00e-04

Fits to all three data sets

Table A.33: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 0.5 for all three experimental data sets.

Weight factor	0.5	1.0	2.0	0.1	100.0	5.0	10.0	0.01	I.P.
Overall RMSE	0.0700731	0.0711117	0.0774554	0.0779319	0.0807847	0.0808747	0.0896994	0.101845	0.174922
Lag RMSE	0.0230377	0.0243795	0.0226798	0.0234285	0.0246538	0.0231814	0.0229863	0.0205467	0.064793
Exponential RMSE	0.100332	0.0995365	0.1049	0.117285	0.09711	0.101458	0.104677	0.165633	0.266981
Stationary RMSE	0.0642922	0.0683282	0.0804955	0.0625743	0.0976748	0.0937602	0.112482	0.0571047	0.127736

Table A.34: RMSE scores divided by curve phase and experimental data set (single sugar and sugar mixture). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	Exponential phase	Stationary phase	RMSE per experiment	
PR1	0.016241	0.116814	0.061921	**0.076906**	
Maltose	0.031759	0.112664	0.064012	**0.077027**	
Galactose	0.017883	0.062134	0.066848	**0.053694**	
RMSE per phase	**0.023038**	**0.100332**	**0.064292**	**0.070073**	**Total RMSE score**

Table A.35: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.014191	0.105792	0.069524	**0.073545**	
GFP	0.017085	0.129994	0.067308	**0.085089**	
OD	0.017264	0.113343	0.046248	**0.071376**	
RMSE per phase	**0.016241**	**0.116814**	**0.061921**	**0.076906**	**Total RMSE score for the PR1 fit**

Table A.36: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.039467	0.135369	0.067403	**0.090232**	
OD	0.021440	0.084034	0.060431	**0.061028**	
RMSE per phase	**0.031759**	**0.112664**	**0.064012**	**0.077027**	**Total RMSE score for the maltose fit**

Table A.37: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
GFP	0.004835	0.070446	0.082793	**0.062825**	
OD	0.024824	0.052521	0.045636	**0.042651**	
RMSE per phase	**0.017883**	**0.062134**	**0.066848**	**0.053694**	**Total RMSE score for the galactose fit**

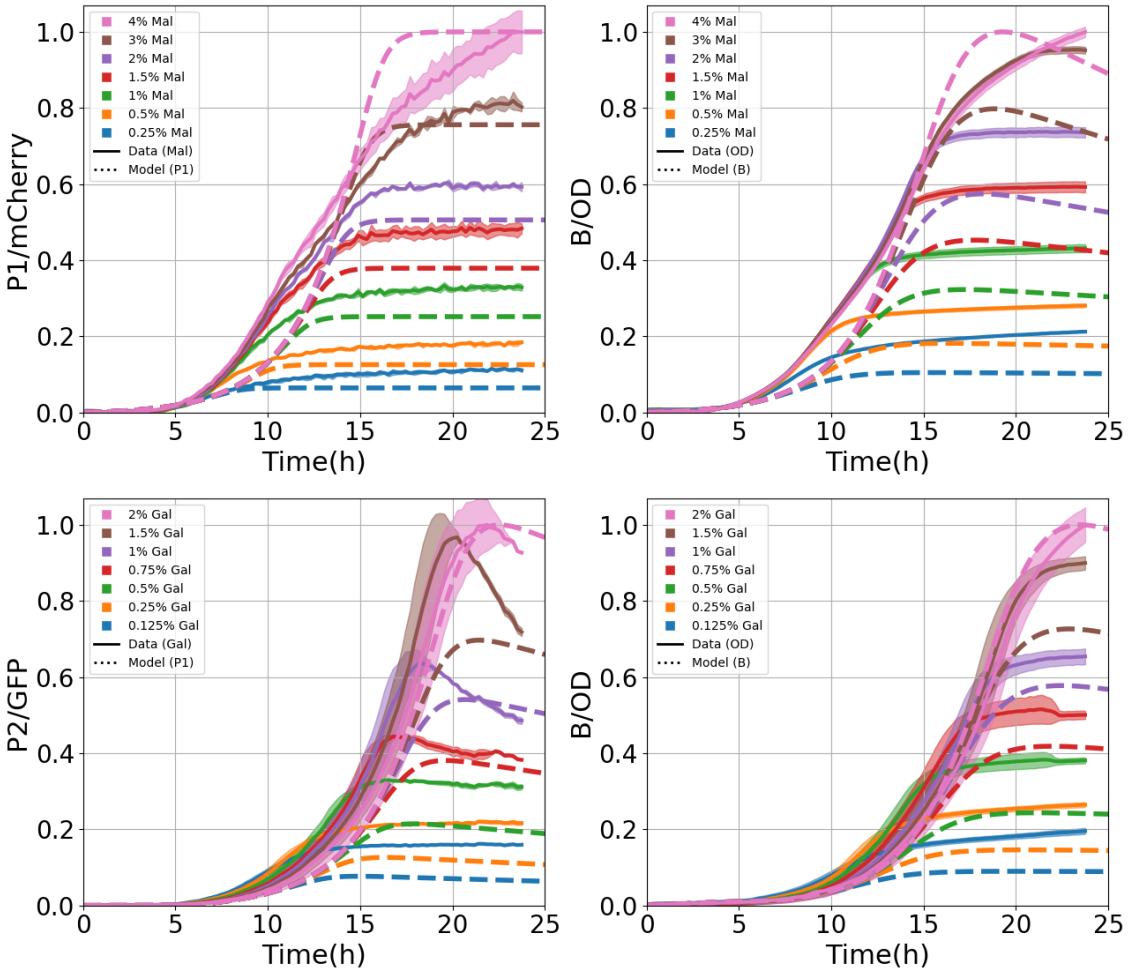


Figure A.12: Fit to the single sugar experimental data sets. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: The maltose data set. **Bottom**: galactose data set. Graphs to the left correspond to the fit to the fluorescence data (mCherry in the case of maltose, GFP in the case of galactose), whilst the graphs to the right present the results of the fit to the absorbance data (O.D.).

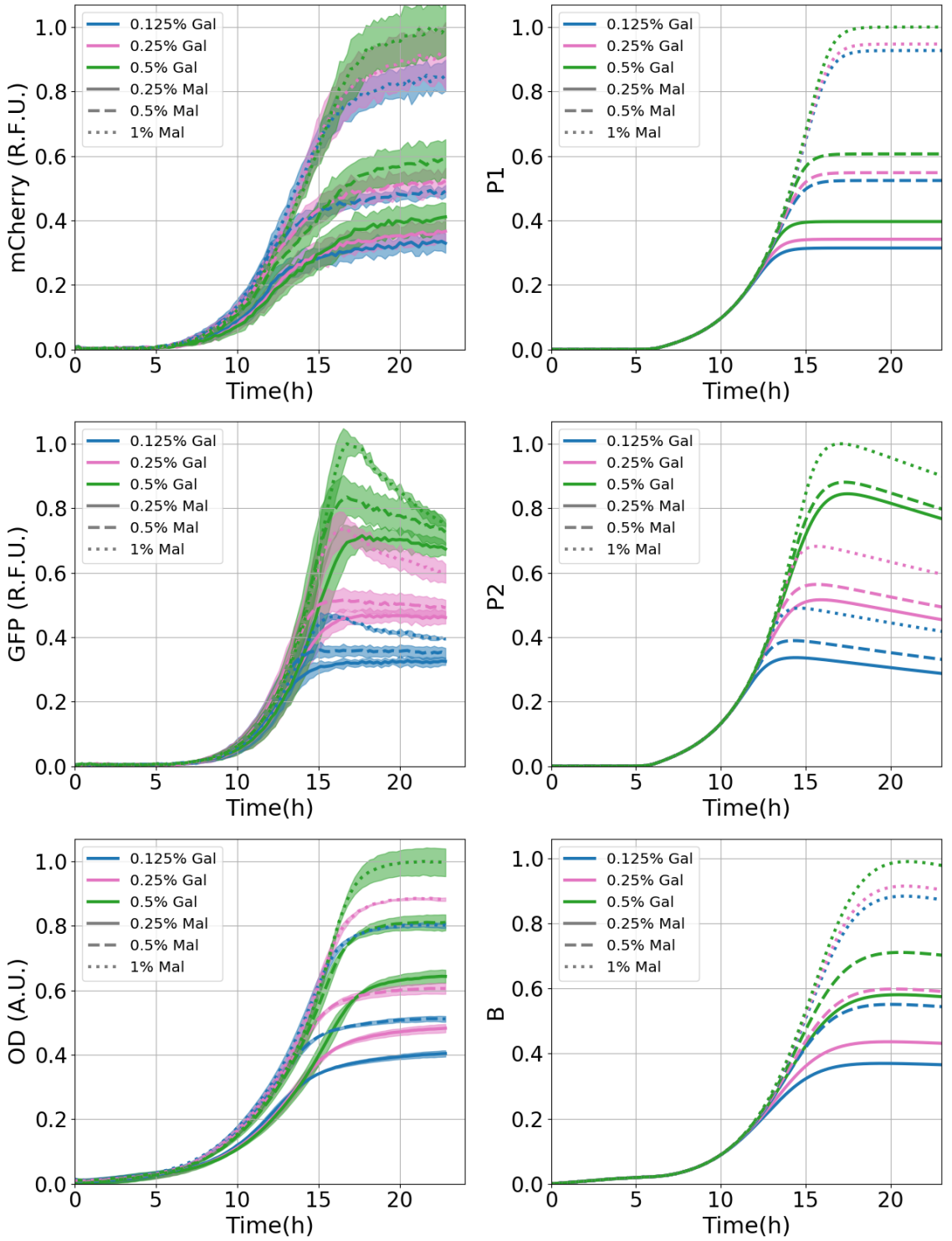


Figure A.13: Fit to the binary-sugar mixture experimental data set PR1. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

Fits to the maltose data set

Table A.38: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 0.5 for the maltose experimental data set.

Weight factor	0.5	2.0	5.0	10.0	0.01	0.1	1.0	100.0	I.P.
Overall RMSE	0.0517957	0.0518281	0.0519836	0.0520779	0.0530163	0.0558257	0.0608989	0.0662226	0.0766771
Lag RMSE	0.043115	0.042923	0.042909	0.0425437	0.0456995	0.0448849	0.0364798	0.0455448	0.0310459
Exponential RMSe	0.0665231	0.0668665	0.0672306	0.0676755	0.0672777	0.0737522	0.0863293	0.081199	0.110037
Stationary RMSE	0.0420018	0.0417725	0.0417822	0.0417898	0.0426315	0.0435372	0.0483994	0.0669976	0.0675734

Table A.39: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.052800	0.065888	0.036870	**0.053193**	
OD	0.030495	0.067152	0.046572	**0.050360**	
RMSE per phase	**0.043115**	**0.066523**	**0.042002**	**0.051796**	**Total RMSE score for the maltose fit**

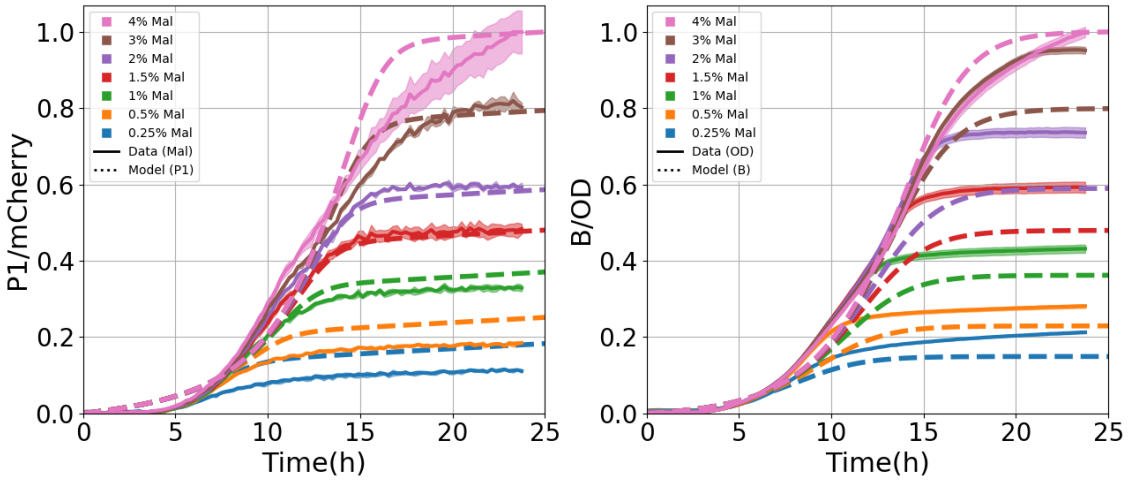


Figure A.14: Fit to the maltose experimental data set. This corresponds to the parameterisation with the lowest overall RMSE score. The left graph to the correspond to the fit to the mCherry fluorescence data, whilst the graph to the right presents the results of the fit to the absorbance data (O.D.).

Fits to the galactose data set

Table A.40: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 0.5 for the galactose experimental data set.

Weight factor	0.5	1.0	2.0	5.0	0.1	100.0	10.0	0.01	I.P.
Overall RMSE	0.0479797	0.0481152	0.0483163	0.0487317	0.0488931	0.0563953	0.0638978	0.0725844	0.155268
Lag RMSE	0.0148418	0.0145023	0.0144722	0.0164732	0.0166365	0.0160247	0.0169648	0.0235385	0.0461846
Exponential RMSE	0.0498682	0.0497496	0.0498035	0.0496689	0.0543969	0.0498804	0.0512135	0.106848	0.228695
Stationary RMSE	0.0647999	0.0652679	0.0656778	0.0662266	0.0627361	0.0824406	0.0966342	0.061926	0.133754

Table A.41: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
GFP	0.005494	0.051674	0.079010	**0.054598**	
OD	0.020258	0.047995	0.046428	**0.040288**	
RMSE per phase	**0.014842**	**0.049868**	**0.064800**	**0.047980**	**Total RMSE score for the galactose fit**

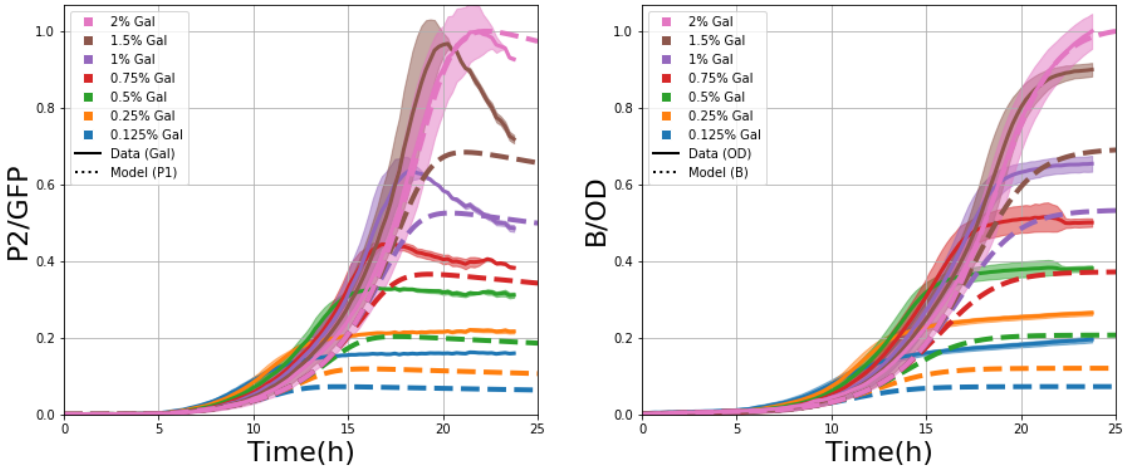


Figure A.15: Fit to the galactose experimental data sets. This corresponds to the parameterisation with the lowest overall RMSE score. The left graph to the correspond to the fit to the GFP fluorescence data, whilst the graph to the right presents the results of the fit to the absorbance data (O.D.).

Fits to the PR1 data set

This fit used a different set of initial conditions.

Table A.42: Initial values of the state variables used for the fit to sugar mixture experimental data set.

	B	G	S_1	S_2	P_1	P_2
PR1	0.1	0.3	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$

Table A.43: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 1 for the PR1 experimental data set.

Weight factor	1.0	0.5	2.0	5.0	0.1	10.0	I.P.
Overall RMSE	0.0521534	0.0542894	0.0550929	0.0578383	0.0626128	0.0924664	0.146682
Lag RMSE	0.0138483	0.0120577	0.0112981	0.0118762	0.0118577	0.0134769	0.037723
Exponential RMSE	0.0719791	0.0750445	0.0738379	0.0789274	0.0921208	0.0919895	0.213429
Stationary RMSE	0.0527937	0.0553619	0.0593801	0.0605412	0.0559841	0.130409	0.132559

Table A.44: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.015839	0.107109	0.059013	**0.071194**	
GFP	0.013273	0.043842	0.066223	**0.046490**	
OD	0.012177	0.046353	0.022214	**0.030498**	
RMSE per phase	**0.013848**	**0.071979**	**0.052794**	**0.052153**	**Total RMSE score for the PR1 fit**

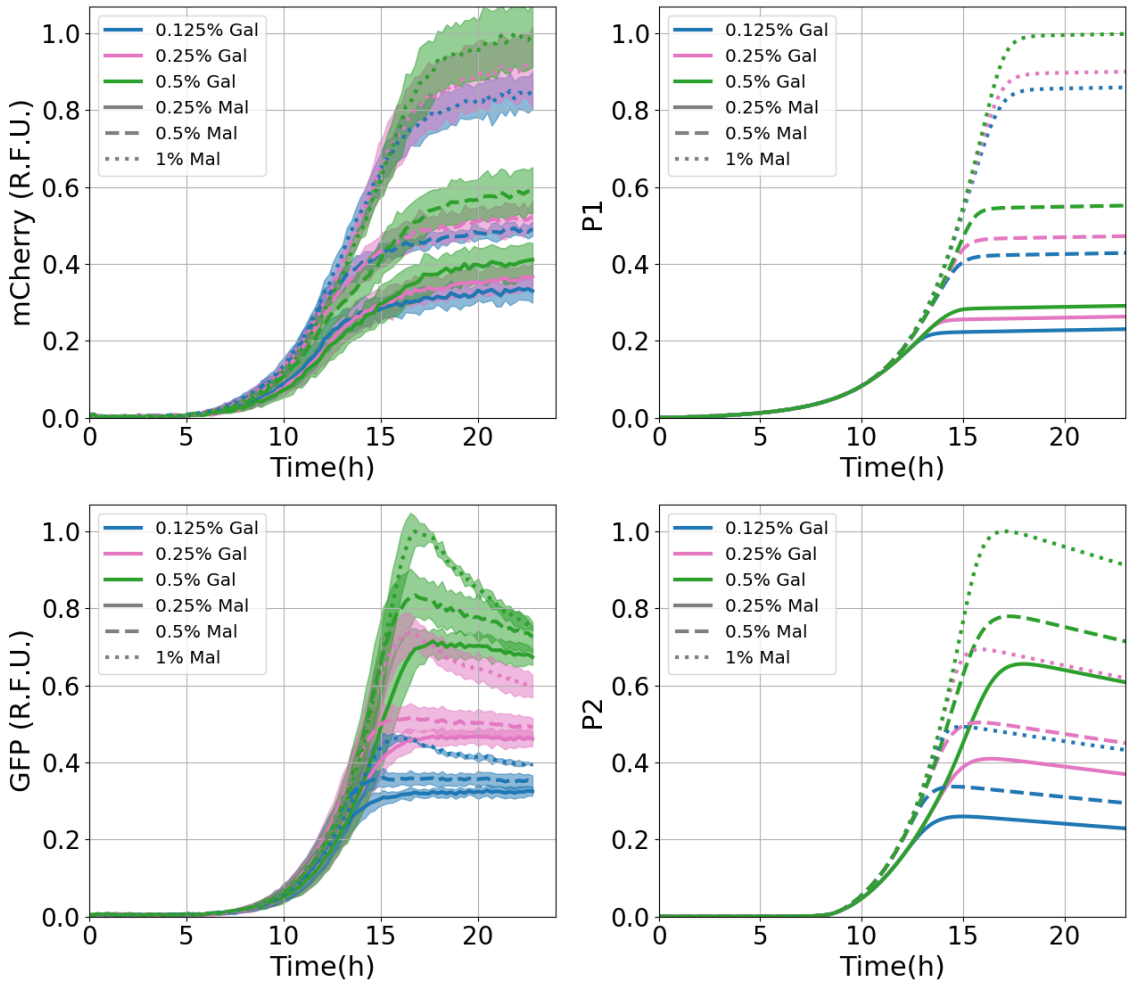


Figure A.16: Fit to the binary-sugar mixture experimental data set PR1. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

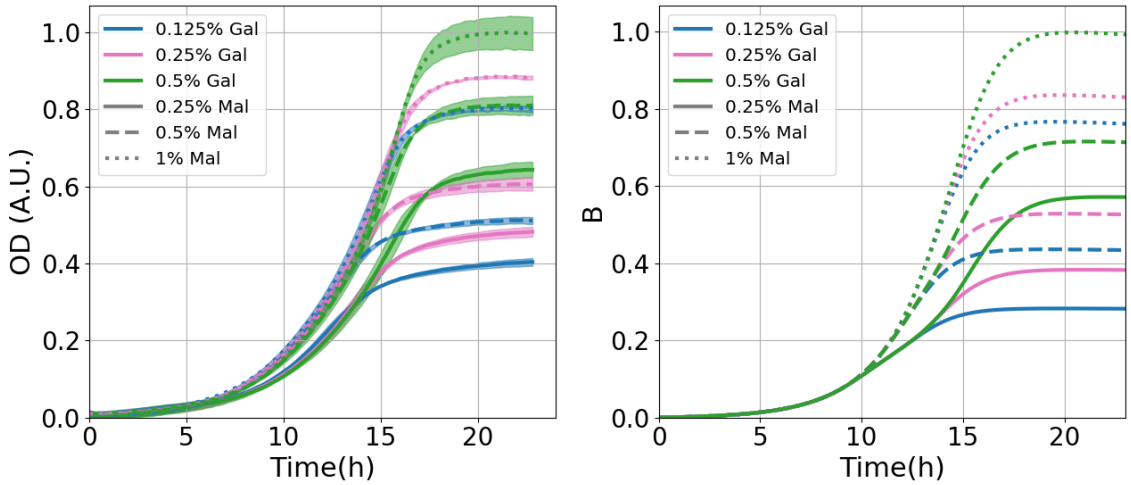


Figure A.16: (cont.)

Fits to all three data sets with PR1 parameters

Table A.45: Initial values of state variables resulting from the grid search performed with the lowest overall RMSE PR1 parameterisation.

	B	G	S_1	S_2	P_1	P_2
PR1	0.15	0.08	$1e^{-6}$	$1e^{-6}$	$1e^{-6}$	$1e^{-5}$
Maltose	1.0	0.325	$1e^{-6}$	0	$1e^{-6}$	0
Galactose	1.0	0.775	0	$5.5e^{-4}$	0	$7.75e^{-5}$

Table A.46: RMSE scores divided by curve phase.

Overall RMSE	0.0739507
Lag RMSE	0.0350063
Exponential RMSE	0.1055864
Stationary RMSE	0.0634996

Table A.47: RMSE scores divided by curve phase and experimental data set (single sugar and sugar mixture).

	Lag phase	Exponential phase	Stationary phase	RMSE per experiment	
PR1	0.014150	0.072923	0.052939	**0.052664**	
Maltose	0.025482	0.142677	0.063864	**0.091442**	
Galactose	0.053167	0.088153	0.072218	**0.072602**	
RMSE per phase	**0.035006**	**0.105586**	**0.063500**	**0.073951**	**Total RMSE score**

Table A.48: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.016552	0.107505	0.059104	**0.071471**	
GFP	0.012756	0.046925	0.066139	**0.047396**	
OD	0.012807	0.046843	0.023241	**0.031083**	
RMSE per phase	**0.014150**	**0.072923**	**0.052939**	**0.052664**	**Total RMSE score for the PR1 fit**

Table A.49: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.023087	0.158615	0.054423	**0.097730**	
OD	0.027671	0.124719	0.072079	**0.084688**	
RMSE per phase	**0.025482**	**0.142677**	**0.063864**	**0.091442**	**Total RMSE score for the maltose fit**

Table A.50: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
GFP	0.011590	0.076873	0.084857	**0.066444**	
OD	0.074291	0.098145	0.056834	**0.078277**	
RMSE per phase	**0.053167**	**0.088153**	**0.072218**	**0.072602**	**Total RMSE score for the galactose fit**

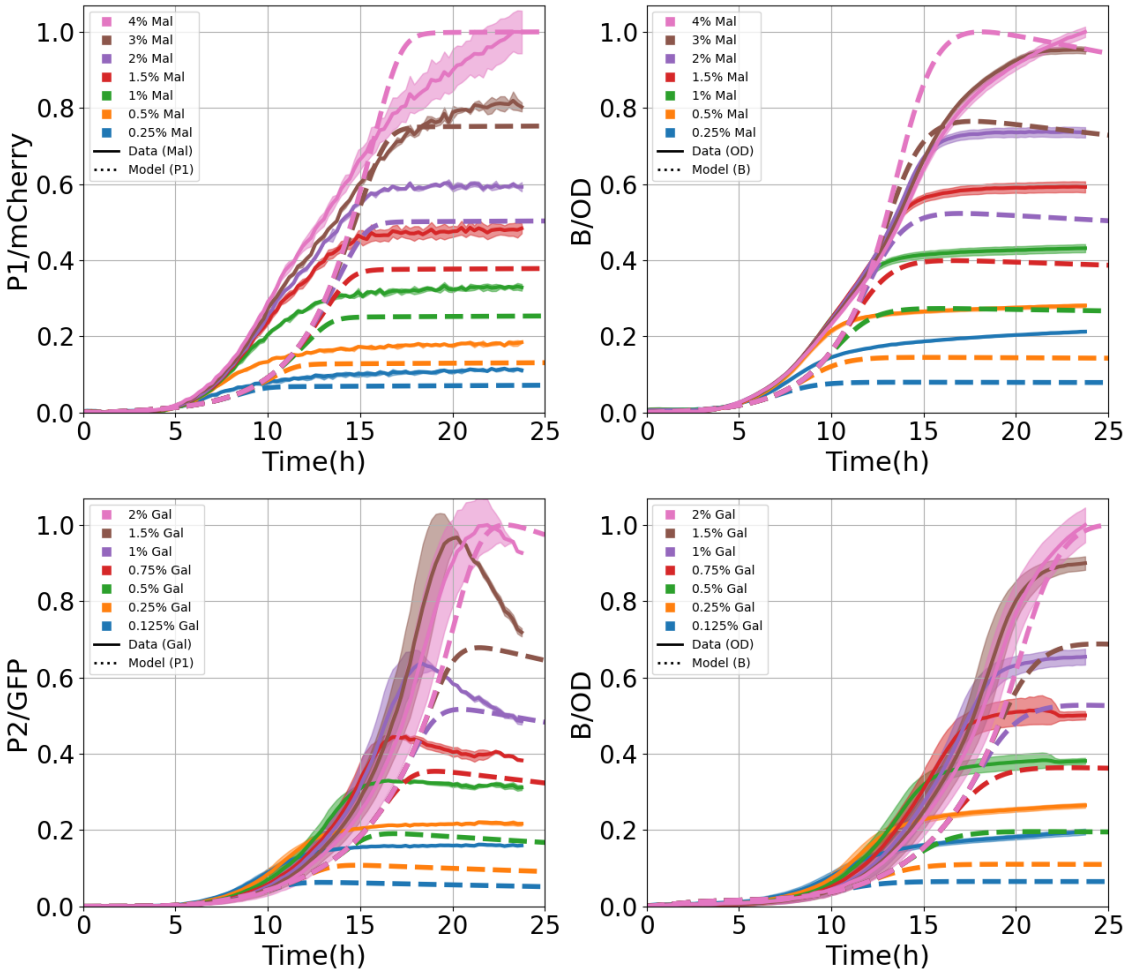


Figure A.17: Fit to the single sugar experimental data sets with the PR1 parameterisation with the lowest overall RMSE score after performing a grid search for new initial state variables values (table A.45). **Top**: The maltose data set. **Bottom**: galactose data set. Graphs to the left correspond to the fit to the fluorescence data (mCherry in the case of maltose, GFP in the case of galactose), whilst the graphs to the right present the results of the fit to the absorbance data (O.D.).

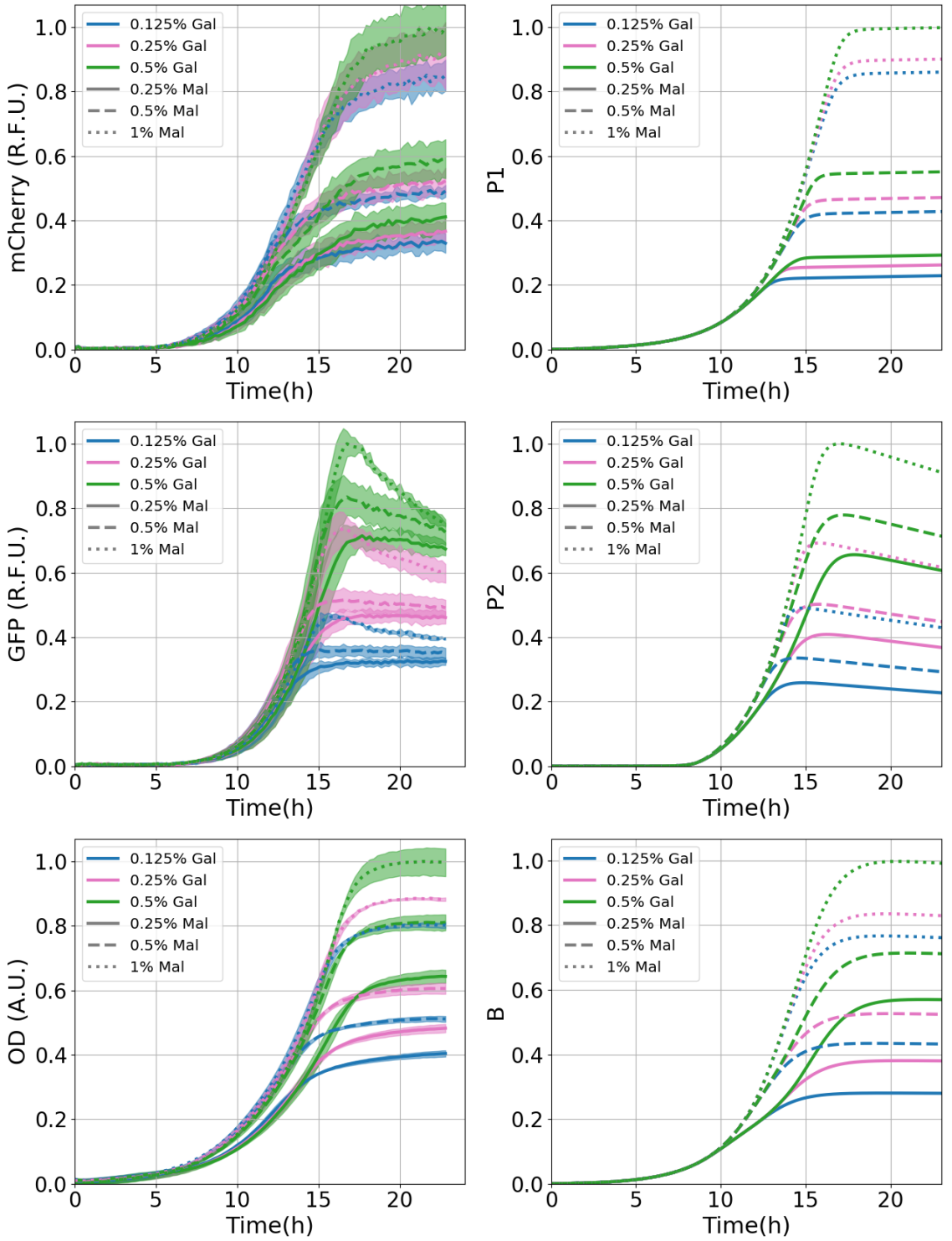


Figure A.18: Fit to the binary-sugar mixture experimental data set PR1 with the PR1 parameterisation with the lowest overall RMSE score after performing a grid search for new initial state variables values (table A.45). **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

A.4 Logistic equation biomass model: RFO induction term

With this model I present two sets of results: 1) The fitting algorithm has lower bounds of 0 for all parameters and upper bounds of 1 to I_i parameters, 2) and all parameters (except I_i have an upper bound equal to ten times the initial parameter value. The reason is that in the first set of fitting routines, for some weights factors, some parameters were optimised to more than 300 times the initial parameter value. Additionally, for the second set of results, when fitting to single sugar data sets, the parameters corresponding to the absent sugar were explicitly ignored by the optimisation algorithm. When this metabolic branch was not ignored (as it was the case for the first set of results presented), the parameters were changed drastically. This stands in contrast to the other models, in which the parameters corresponding to the absent sugar were hardly affected.

Table A.51: Initial values of the state variables used.

	B	**G**	$\boldsymbol{S_1}$	$\boldsymbol{S_2}$	$\boldsymbol{P_1}$	$\boldsymbol{P_2}$
PR1	0.1	0.1	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$
Maltose	0.1	0.1	$1e^{-4}$	0	$3e^{-3}$	0
Galactose	0.05	0.1	0	$1e^{-4}$	0	$2.5e^{-3}$

Table A.52: Different parameterisations resulting from fitting the model to different experimental data sets. These parameterisations correspond to those with the lowest overall RMSE score of their corresponding experimental data set. On each row corresponding to the parameter values, the lightest green indicates the lowest value, whereas the deepest green indicates the highest value of that parameter out of all the different parameterisations. Likewise, a red-blue colormap is used to differentiate between the highest and lowest normalised standard deviation, respectively. The parameterisations presented here are all resulting from fittings routines which are lower bounded to 0 with no upper bounds except for parameter I_i, which has a upper bound equal to 1.

Parameter	Normalised std.	Mean	All data	Maltose	Galactose	PR1	Initial Parameters
A_1	1.78e+00	3.01e-01	1.00e-06	8.56e-02	1.10e+00	1.78e-02	8.02e-03
V_{S1}	1.06e-05	6.00e+02	6.00e+02	6.00e+02	6.00e+02	6.00e+02	6.00e+02
V_{P1}	1.63e+00	3.17e-01	1.71e-01	4.61e-03	1.08e+00	7.46e-03	9.00e-03
V_{G1}	1.77e-01	8.69e-01	9.34e-01	6.40e-01	9.30e-01	9.72e-01	9.30e-01
K_{inh1}	6.26e-01	6.15e-01	8.03e-01	3.74e-02	8.00e-01	8.18e-01	8.00e-01
K_{S1}	6.49e-01	2.84e+00	1.91e+00	5.60e+00	1.90e+00	1.95e+00	1.90e+00
I_1	2.49e-01	1.96e-01	2.02e-01	1.60e-01	1.60e-01	2.64e-01	1.60e-01
K_{P1}	4.78e-01	1.00e-01	1.18e-01	4.96e-02	7.50e-02	1.58e-01	7.50e-02
Dcy_{P1}	2.00e+00	7.72e-01	1.00e-06	1.00e-06	3.09e+00	1.00e-06	6.00e-04
Y_1	1.87e-01	1.31e-01	1.44e-01	1.25e-01	1.00e-01	1.56e-01	1.00e-01
A_2	1.80e+00	2.17e-02	1.00e-06	4.18e-03	8.04e-02	2.38e-03	8.02e-03
V_{S2}	1.44e-06	7.00e+02	7.00e+02	7.00e+02	7.00e+02	7.00e+02	7.00e+02
V_{P2}	1.00e+00	6.56e-03	1.57e-02	2.40e-03	1.30e-03	6.82e-03	7.50e-03
V_{G2}	2.48e-01	7.16e-01	8.12e-01	8.00e-01	4.50e-01	8.02e-01	8.00e-01
K_{inh2}	5.09e-01	9.33e-01	6.91e-01	7.00e-01	1.65e+00	6.95e-01	7.00e-01
K_{S2}	5.83e-01	1.38e+00	1.00e+00	9.50e-01	2.59e+00	9.80e-01	9.50e-01
I_2	1.37e+00	8.39e-02	2.54e-01	1.50e-02	1.50e-02	5.16e-02	1.50e-02
K_{P2}	1.13e+00	1.35e-01	3.77e-02	2.00e-02	1.29e-01	3.53e-01	2.00e-02
Dcy_{P2}	1.38e+00	5.52e-02	1.00e-06	1.68e-01	2.80e-02	2.46e-02	6.00e-02
Y_2	4.59e-01	3.37e-01	2.78e-01	2.40e-01	5.68e-01	2.63e-01	2.40e-01
V_B	3.06e-01	3.84e-01	2.82e-01	3.70e-01	3.33e-01	5.52e-01	4.00e-01
K_G	7.35e-01	4.45e+00	2.03e+00	4.88e+00	8.93e+00	1.99e+00	2.00e+00
M	9.31e-01	1.34e+00	3.38e-01	1.94e+00	2.80e+00	2.68e-01	3.00e-01
Dcy_B	5.31e-01	1.71e-03	2.87e-03	1.31e-03	7.49e-04	1.92e-03	2.50e-03

Table A.53: Different parameterisations resulting from fitting the model to different experimental data sets. These parameterisations correspond to those with the lowest overall RMSE score of their corresponding experimental data set. On each row corresponding to the parameter values, the lightest green indicates the lowest value, whereas the deepest green indicates the highest value of that parameter out of all the different parameterisations. Likewise, a red-blue colormap is used to differentiate between the highest and lowest normalised standard deviation, respectively. The parameterisations presented here are all resulting from fittings routines which are lower bounded to 0 and upper bounded to ten times the initial value.

Parameter	Normalised std.	Mean	All data	Maltose	Galactose	PR1	Initial Parameters
A_1	1.80e+00	1.50e-02	3.96e-05	6.27e-02	8.02e-03	4.11e-03	8.02e-03
V_{S1}	9.90e-05	6.00e+02	6.00e+02	6.00e+02	6.00e+02	6.00e+02	6.00e+02
V_{P1}	9.05e-01	2.17e-02	5.27e-03	8.55e-03	9.00e-03	3.83e-02	9.00e-03
V_{G1}	2.49e-01	8.54e-01	1.02e+00	4.81e-01	9.30e-01	9.04e-01	9.30e-01
K_{inh1}	5.10e-01	6.75e-01	8.59e-01	6.15e-02	8.00e-01	8.56e-01	8.00e-01
K_{S1}	7.66e-01	2.95e+00	2.03e+00	6.99e+00	1.90e+00	1.93e+00	1.90e+00
I_1	8.71e-01	3.50e-01	8.63e-01	1.60e-01	1.60e-01	3.98e-01	1.60e-01
K_{P1}	7.14e-01	9.50e-02	1.02e-02	1.22e-01	7.50e-02	1.93e-01	7.50e-02
Dcy_{P1}	2.22e+00	1.21e-04	1.00e-06	1.00e-06	6.00e-04	1.00e-06	6.00e-04
Y_1	1.19e+00	2.61e-01	8.17e-01	9.78e-02	1.00e-01	1.65e-01	1.00e-01
A_2	1.78e+00	9.97e-03	1.00e-06	8.02e-03	4.11e-02	3.26e-04	8.02e-03
V_{S2}	2.93e-06	7.00e+02	7.00e+02	7.00e+02	7.00e+02	7.00e+02	7.00e+02
V_{P2}	1.67e+00	2.14e-02	8.96e-04	7.50e-03	2.03e-03	8.47e-02	7.50e-03
V_{G2}	1.51e-01	8.34e-01	1.04e+00	8.00e-01	6.95e-01	8.28e-01	8.00e-01
K_{inh2}	4.11e-01	7.77e-01	5.06e-01	7.00e-01	1.33e+00	6.52e-01	7.00e-01
K_{S2}	7.36e-01	1.79e+00	1.86e+00	9.50e-01	4.06e+00	1.15e+00	9.50e-01
I_2	1.99e+00	2.19e-01	1.00e+00	1.50e-02	1.50e-02	6.68e-02	1.50e-02
K_{P2}	1.36e+00	1.50e-01	7.61e-03	2.00e-02	7.21e-02	5.00e-01	2.00e-02
Dcy_{P2}	3.62e-01	3.89e-02	4.13e-02	6.00e-02	4.11e-02	2.74e-02	6.00e-02
Y_2	7.25e-01	5.49e-01	7.59e-01	2.40e-01	1.15e+00	3.29e-01	2.40e-01
V_B	4.69e-01	4.62e-01	3.13e-01	3.53e-01	3.81e-01	8.43e-01	4.00e-01
K_G	6.65e-01	3.93e+00	2.19e+00	6.13e+00	7.38e+00	1.97e+00	2.00e+00
M	1.09e+00	7.33e-01	1.24e+00	1.90e+00	1.90e-01	3.19e-02	3.00e-01
Dcy_B	1.21e+00	1.15e-03	3.52e-03	4.40e-04	1.21e-03	1.00e-06	2.50e-03

Fits to all three data sets (no upper bounds)

Table A.54: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 1 for all three experimental data sets.

Weight factor	1.0	2.0	0.5	5.0	10.0	0.1	100.0	I.P.	0.01
Overall RMSE	0.0852884	0.0870618	0.0874275	0.0875585	0.0981708	0.101973	0.110603	0.144765	0.163771
Lag RMSE	0.0495467	0.0484387	0.0491397	0.0469312	0.0470977	0.0557761	0.0432063	0.0621095	0.146924
Exponential RMSE	0.121136	0.123927	0.126303	0.12343	0.122437	0.15617	0.123805	0.215953	0.239346
Stationary RMSE	0.0685087	0.0709585	0.0675532	0.0745783	0.108183	0.0607889	0.13966	0.111255	0.0398685

Table A.55: RMSE scores divided by curve phase and experimental data set (single sugar and sugar mixture). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	Exponential phase	Stationary phase	RMSE per experiment	
PR1	0.046325	0.115046	0.074548	**0.083545**	
Maltose	0.042104	0.099994	0.051518	**0.069344**	
Galactose	0.058702	0.144179	0.076608	**0.100171**	
RMSE per phase	**0.049547**	**0.121136**	**0.068509**	**0.085288**	**Total RMSE score**

Table A.56: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.057969	0.148525	0.051466	**0.096728**	
GFP	0.054157	0.124138	0.109543	**0.100570**	
OD	0.012023	0.047297	0.044989	**0.038321**	
RMSE per phase	**0.046325**	**0.115046**	**0.074548**	**0.083545**	**Total RMSE score for the PR1 fit**

Table A.57: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.050497	0.113174	0.050094	**0.077175**	
OD	0.031552	0.084789	0.052903	**0.060507**	
RMSE per phase	**0.042104**	**0.099994**	**0.051518**	**0.069344**	**Total RMSE score for the maltose fit**

Table A.58: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 1).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
GFP	0.055632	0.127003	0.096582	**0.097558**	
OD	0.061620	0.159517	0.049086	**0.102716**	
RMSE per phase	**0.058702**	**0.144179**	**0.076608**	**0.100171**	**Total RMSE score for the galactose fit**

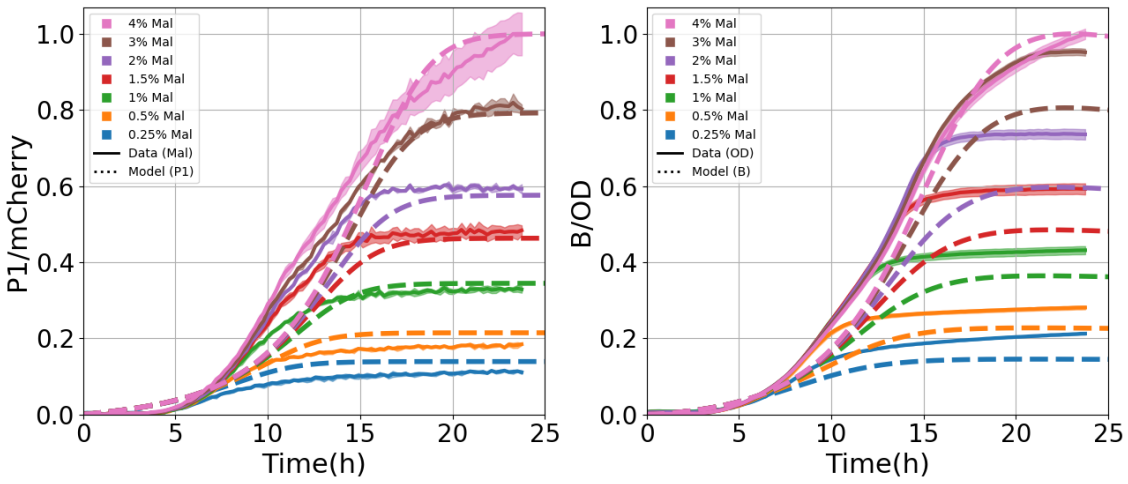


Figure A.19: Fit to the single sugar experimental data sets with no upper bounds. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: The maltose data set. **Bottom**: galactose data set. Graphs to the left correspond to the fit to the fluorescence data (mCherry in the case of maltose, GFP in the case of galactose), whilst the graphs to the right present the results of the fit to the absorbance data (O.D.).

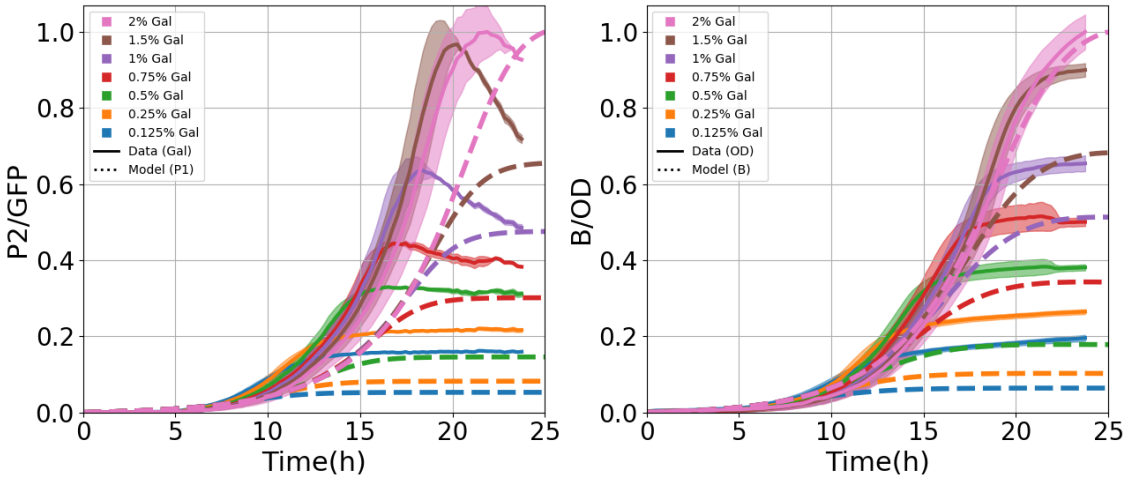


Figure A.19: (cont.)

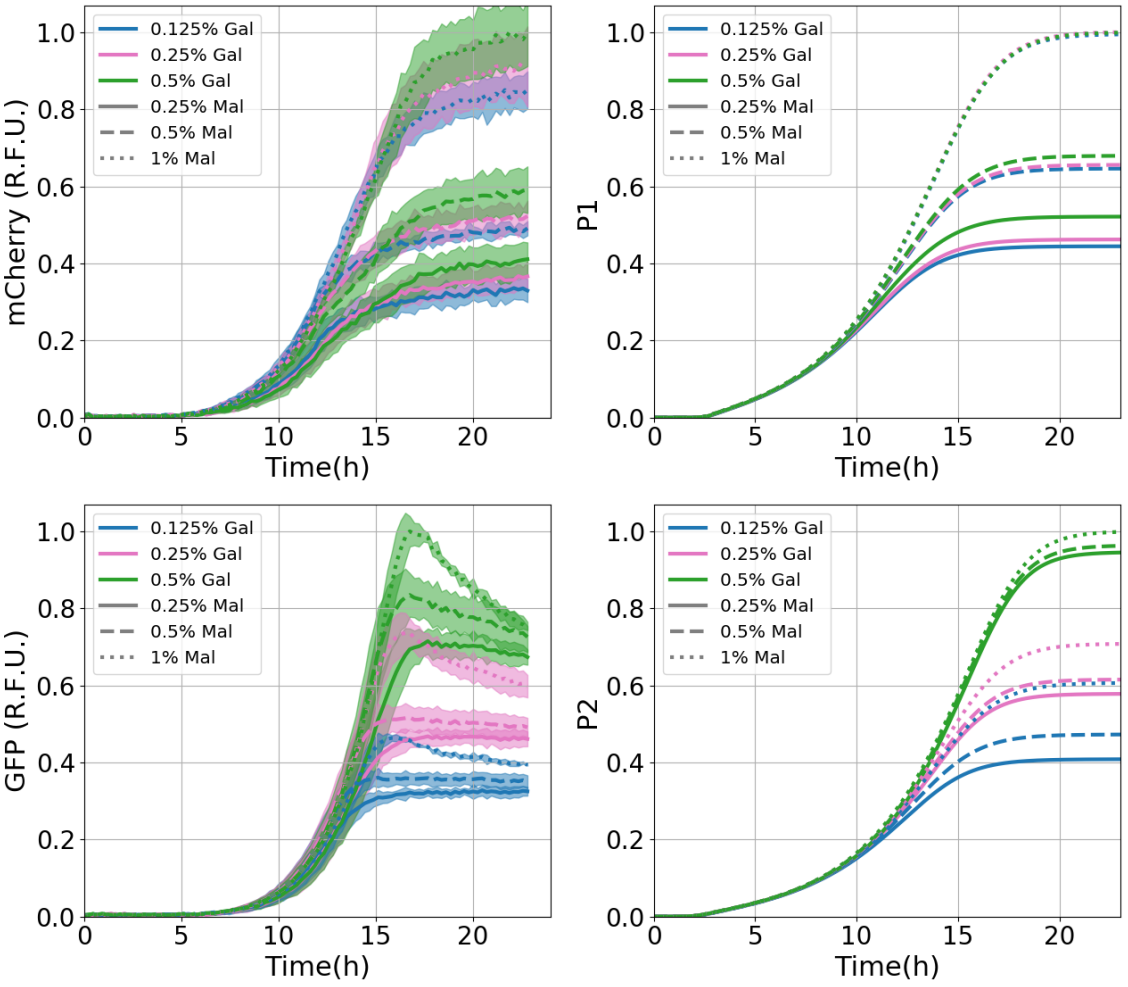


Figure A.20: Fit to the binary-sugar mixture experimental data set PR1 with no upper bounds. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

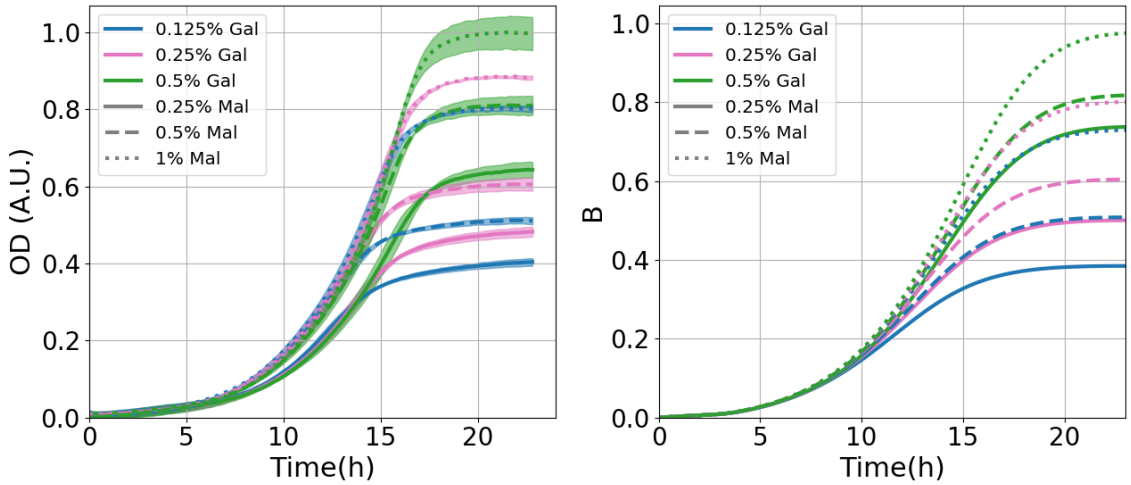


Figure A.20: (cont.)

Fits to all three data sets (with upper bounds)

Table A.59: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 2 for all three experimental data sets.

Weight factor	2.0	1.0	0.5	5.0	10.0	0.1	100.0	I.P.	0.01
Overall RMSE	0.0656933	0.0830956	0.0854282	0.0951636	0.0988267	0.105845	0.108844	0.144765	0.16949
Lag RMSE	0.0334793	0.0445871	0.0479708	0.0428837	0.0433451	0.057687	0.0457757	0.0621095	0.091811
Exponential RMSE	0.0946766	0.119079	0.123844	0.126856	0.122703	0.163006	0.117856	0.215953	0.27357
Stationary RMSE	0.0535004	0.0674304	0.0652329	0.096108	0.1112	0.0609151	0.139842	0.111255	0.0539529

Table A.60: RMSE scores divided by curve phase and experimental data set (single sugar and sugar mixture). This parameterisation corresponds to the lowest overall RMSE score (w.f. 2).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per experiment**	
PR1	0.030208	0.085525	0.048399	**0.059356**	
Maltose	0.032359	0.090614	0.054547	**0.063858**	
Galactose	0.037456	0.106609	0.057176	**0.073115**	
RMSE per phase	**0.033479**	**0.094677**	**0.053500**	**0.065693**	**Total RMSE score**

Table A.61: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1). This parameterisation corresponds to the lowest overall RMSE score (w.f. 2).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.021212	0.091252	0.045950	**0.060245**	
GFP	0.046229	0.110615	0.066741	**0.079220**	
OD	0.012270	0.037160	0.021485	**0.025775**	
RMSE per phase	**0.030208**	**0.085525**	**0.048399**	**0.059356**	**Total RMSE score for the PR1 fit**

Table A.62: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 2).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
mCherry	0.037664	0.075433	0.056387	**0.058561**	
OD	0.025994	0.103595	0.052642	**0.068748**	
RMSE per phase	**0.032359**	**0.090614**	**0.054547**	**0.063858**	**Total RMSE score for the maltose fit**

Table A.63: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 2).

	Lag phase	Exponential phase	Stationary phase	RMSE per data type	
GFP	0.010987	0.054111	0.069390	**0.051198**	
OD	0.051818	0.140723	0.041511	**0.089836**	
RMSE per phase	**0.037456**	**0.106609**	**0.057176**	**0.073115**	**Total RMSE score for the galactose fit**

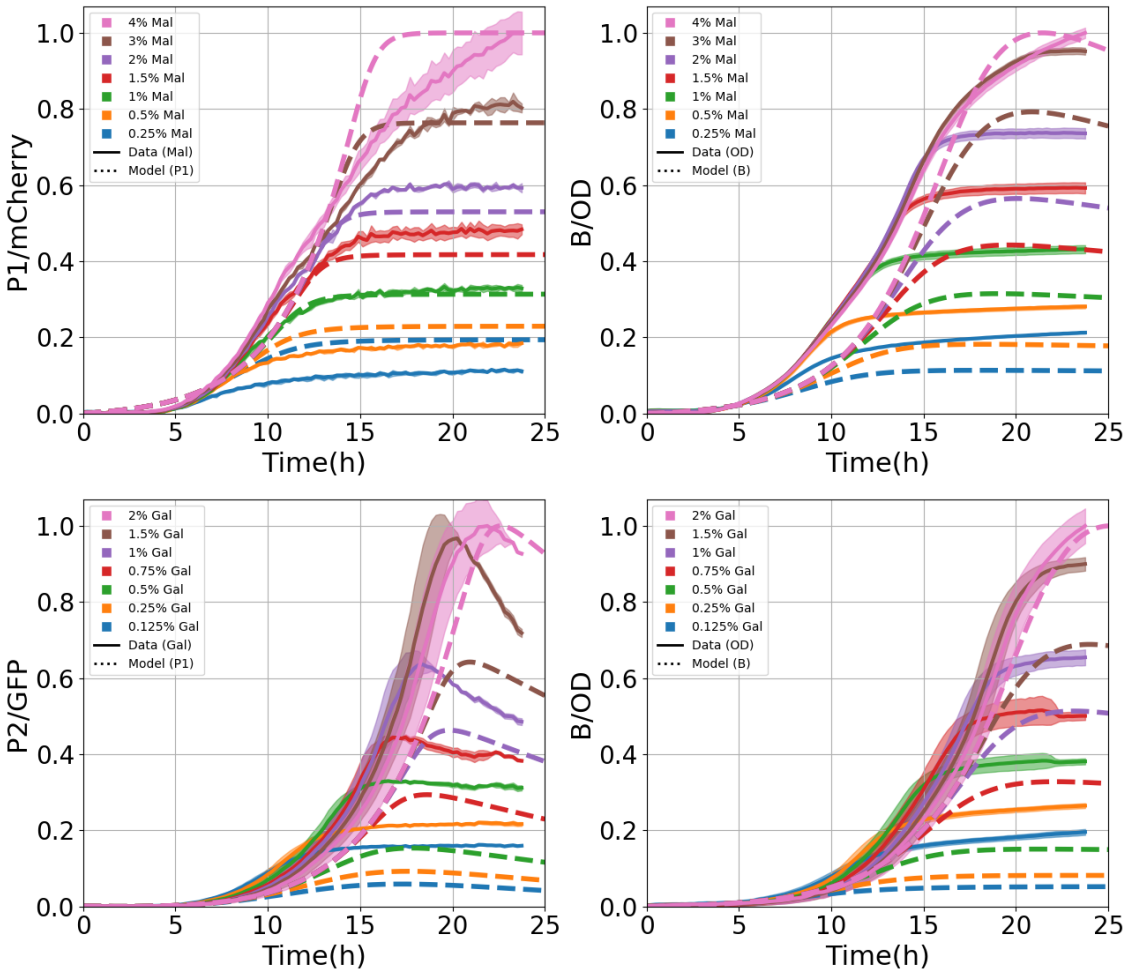


Figure A.21: Fit to the single sugar experimental data sets with upper bounds. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: The maltose data set. **Bottom**: galactose data set. Graphs to the left correspond to the fit to the fluorescence data (mCherry in the case of maltose, GFP in the case of galactose), whilst the graphs to the right present the results of the fit to the absorbance data (O.D.).

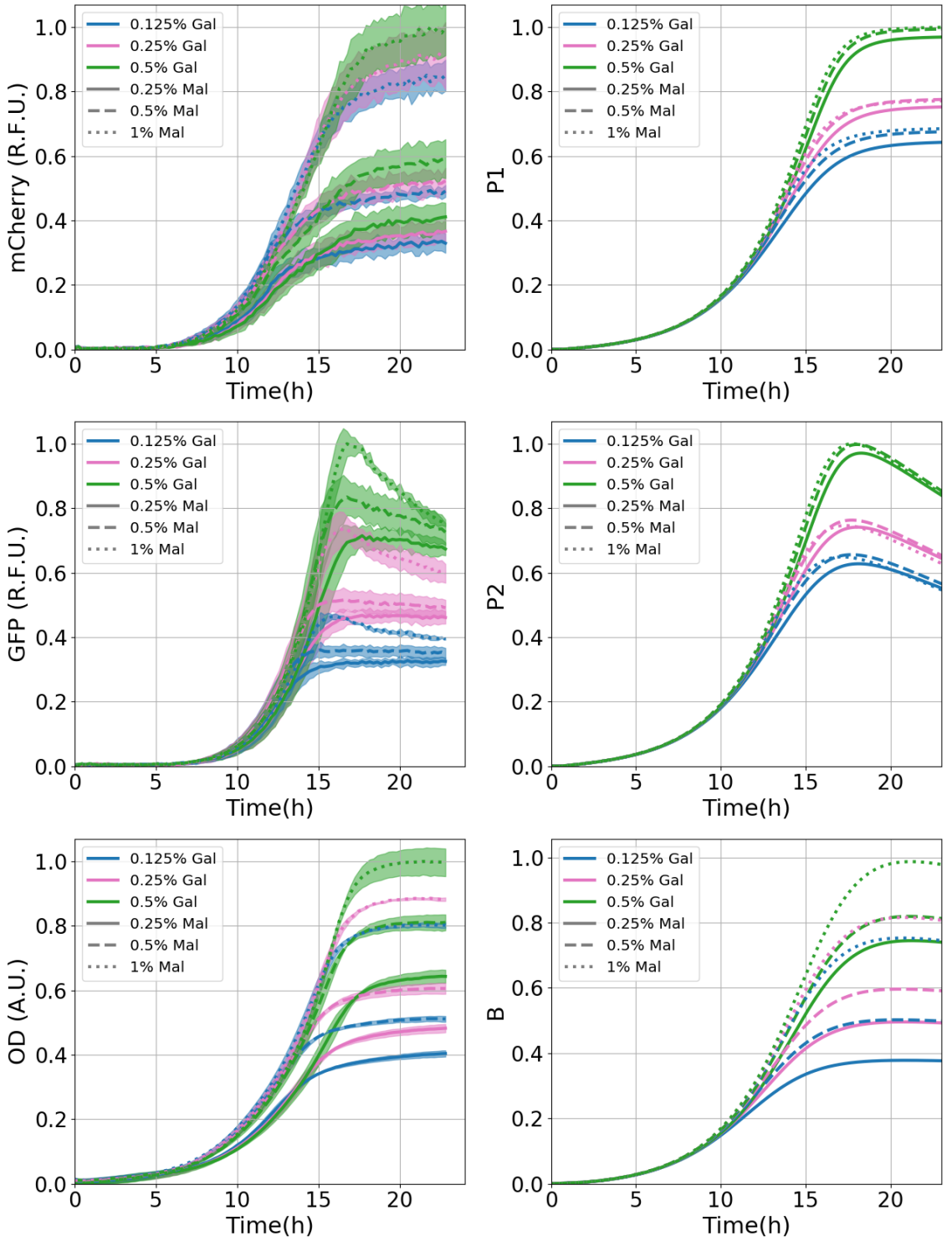


Figure A.22: Fit to the binary-sugar mixture experimental data set PR1 with upper bounds. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

Fits to the maltose data set (no upper bounds)

Results in section 4.3.3 of the main text.

Fits to the maltose data set (with upper bounds)

Table A.64: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 5 for the maltose experimental data set.

Weight factor	5.0	10.0	1.0	2.0	0.5	0.1	I.P.
Overall RMSE	0.0473648	0.0480467	0.0533344	0.0583014	0.078906	0.0810948	0.138003
Lag RMSE	0.0346402	0.033034	0.0486252	0.0525427	0.0471886	0.0466653	0.034953
Exponential RMSe	0.062299	0.0586447	0.0682538	0.0723451	0.120404	0.125454	0.208984
Stationary RMSE	0.0406098	0.0489389	0.0388675	0.046932	0.0442107	0.042577	0.110625

Table A.65: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 5).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.040470	0.062286	0.038680	**0.048351**	
OD	0.027606	0.062312	0.042452	**0.046357**	
RMSE per phase	**0.034640**	**0.062299**	**0.040610**	**0.047365**	**Total RMSE score for the maltose fit**

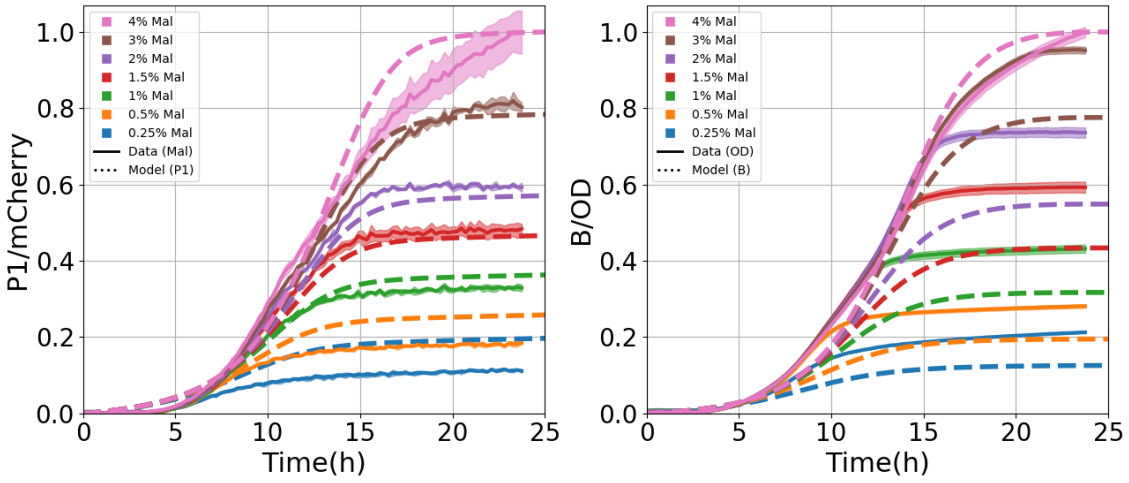


Figure A.23: Fit to the maltose experimental data set with upper bounds. This corresponds to the parameterisation with the lowest overall RMSE score. The left graph to the correspond to the fit to the mCherry fluorescence data, whilst the graph to the right presents the results of the fit to the absorbance data (O.D.).

Fits to the galactose data set (no upper bounds)

Table A.66: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 2 for the galactose experimental data set.

Weight factor	2.0	0.5	10.0	1.0	5.0	0.1	I.P.
Overall RMSE	0.050985	0.0510049	0.0624103	0.0676757	0.0759149	0.144894	0.18801
Lag RMSE	0.0130053	0.0100429	0.0298011	0.0393954	0.0475509	0.12114	0.0915962
Exponential RMSE	0.0585971	0.0657933	0.0852746	0.0827782	0.092109	0.210811	0.281049
Stationary RMSE	0.064774	0.0580936	0.0593742	0.0730463	0.0808952	0.0621815	0.136618

Table A.67: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose). This parameterisation corresponds to the lowest overall RMSE score (w.f. 2).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
GFP	0.005219	0.063293	0.081499	**0.059652**	
OD	0.017636	0.053491	0.041825	**0.040504**	
RMSE per phase	**0.013005**	**0.058597**	**0.064774**	**0.050985**	**Total RMSE score for the galactose fit**

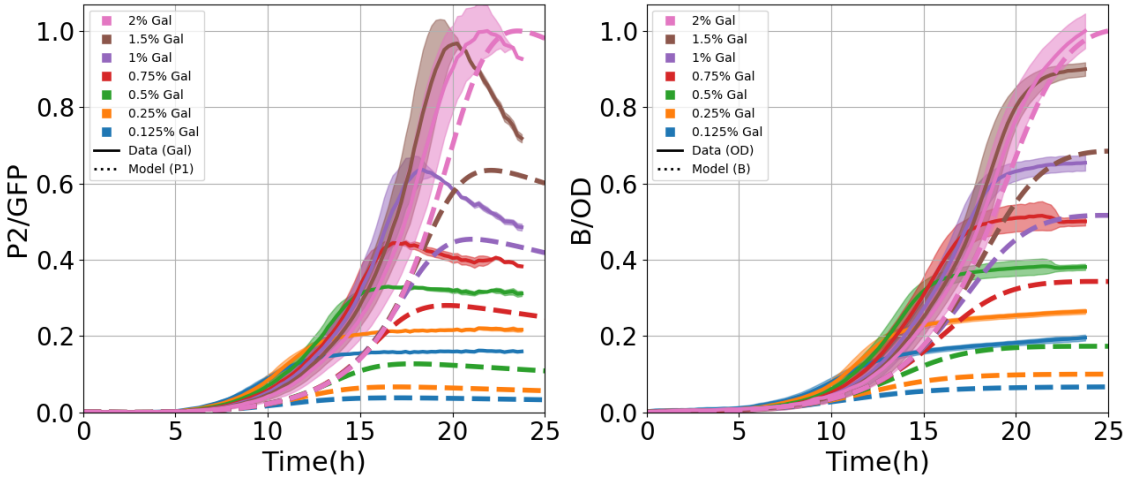


Figure A.24: Fit to the galactose experimental data set with no bounds. This corresponds to the parameterisation with the lowest overall RMSE score. The left graph to the correspond to the fit to the GFP fluorescence data, whilst the graph to the right presents the results of the fit to the absorbance data (O.D.).

Fits to the galactose data set (with upper bounds)

Results in section 4.3.3 of the main text.

Fits to the PR1 data set (no upper bounds)

Table A.68: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 0.5 for the PR1 experimental data set.

Weight factor	0.5	1.0	10.0	2.0	0.1	5.0	I.P.
Overall RMSE	0.0707121	0.072789	0.0736088	0.0793648	0.0919703	0.0943464	0.11541
Lag RMSE	0.0300851	0.0206366	0.0239096	0.0241216	0.0418072	0.0236257	0.0607717
Exponential RMSE	0.109127	0.111356	0.108792	0.124801	0.14648	0.147441	0.173592
Stationary RMSE	0.0467634	0.0553964	0.0620278	0.0523383	0.0465972	0.0663834	0.0783027

Table A.69: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.5).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.041650	0.090394	0.045144	**0.063097**	
GFP	0.007429	0.137492	0.062350	**0.087267**	
OD	0.030421	0.093011	0.025198	**0.058342**	
RMSE per phase	**0.030085**	**0.109127**	**0.046763**	**0.070712**	**Total RMSE score for the PR1 fit**

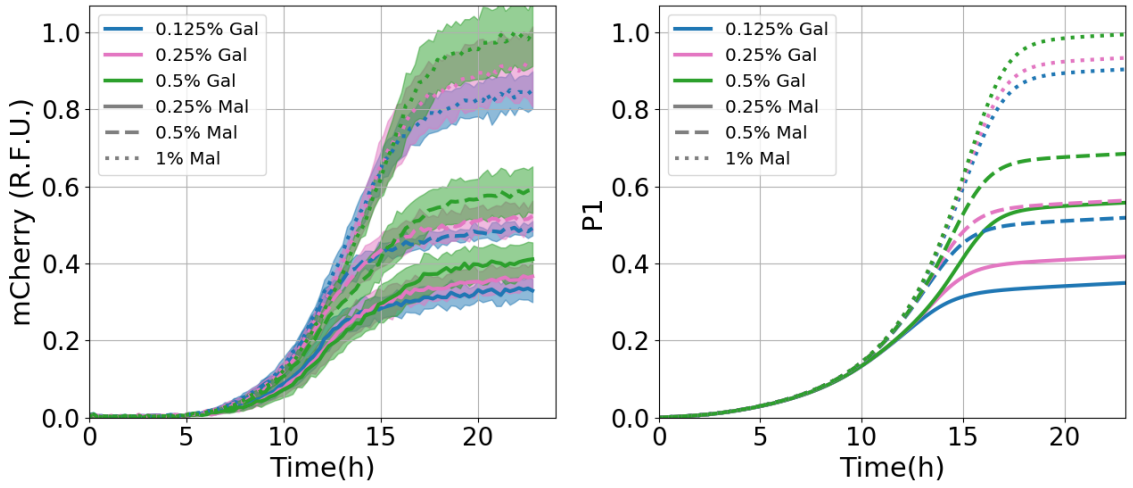


Figure A.25: Fit to binary-sugar mixture experimental data set PR1 with no upper bounds. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

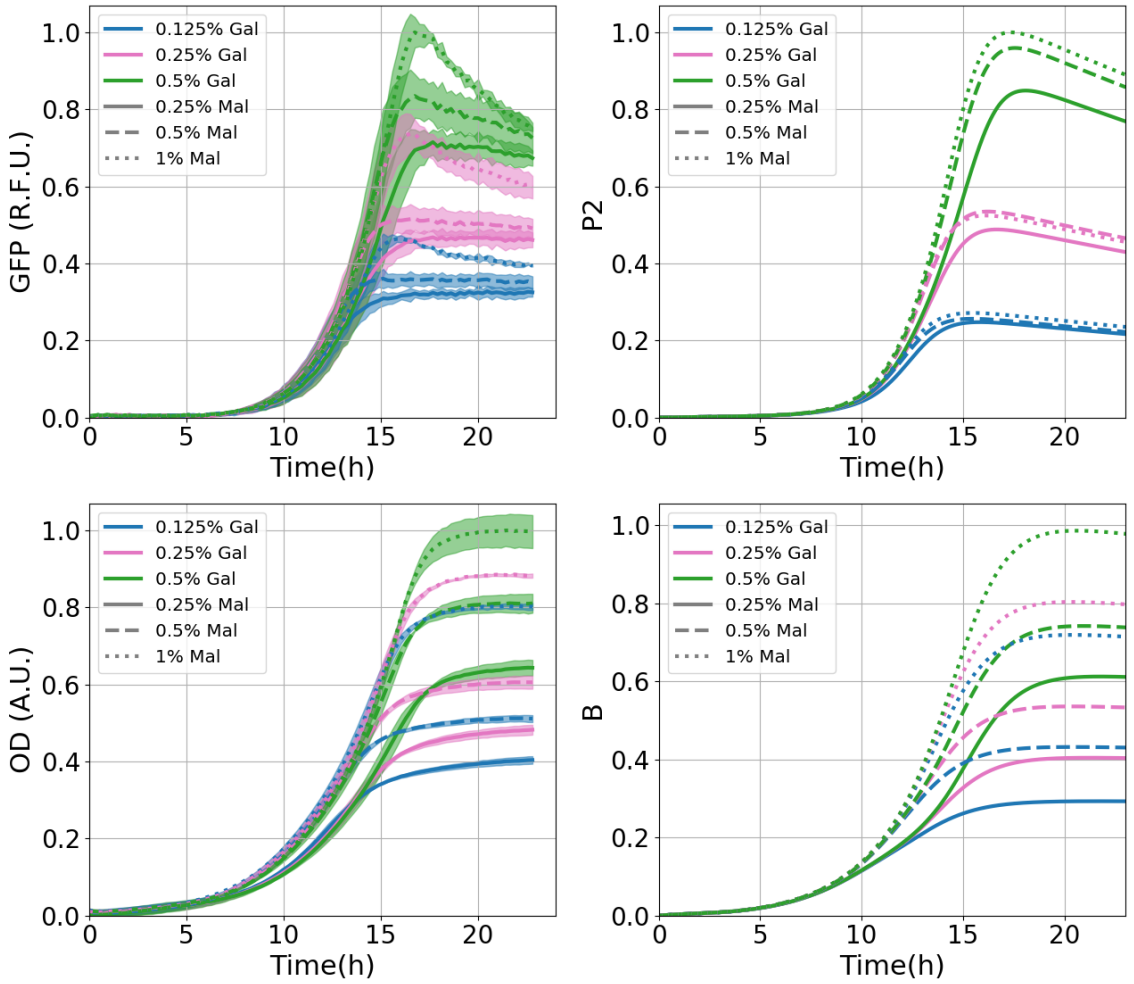


Figure A.25: (cont.)

Fits to the PR1 data set (with upper bounds)

Table A.70: RMSE scores divided by curve phase. The columns are sorted based on the overall RMSE score. The lowest scoring parameterisation corresponds to a weight factor of 0.1 for the PR1 experimental data set.

Weight factor	0.1	1.0	0.5	10.0	2.0	5.0	I.P.
Overall RMSE	0.0706632	0.0745766	0.075937	0.0871117	0.0882881	0.0887117	0.11541
Lag RMSE	0.0446738	0.0221325	0.0232154	0.0335581	0.042087	0.0369106	0.0607717
Exponential RMSE	0.107124	0.115066	0.117568	0.121824	0.127927	0.12786	0.173592
Stationary RMSE	0.0388413	0.0543603	0.0542034	0.0824509	0.0724402	0.0768033	0.0783027

Table A.71: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1). This parameterisation corresponds to the lowest overall RMSE score (w.f. 0.1).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.058842	0.100880	0.029550	**0.069552**	
GFP	0.025154	0.122642	0.055837	**0.079145**	
OD	0.043499	0.095960	0.023129	**0.062277**	
RMSE per phase	**0.044674**	**0.107124**	**0.038841**	**0.070663**	**Total RMSE score for the PR1 fit**

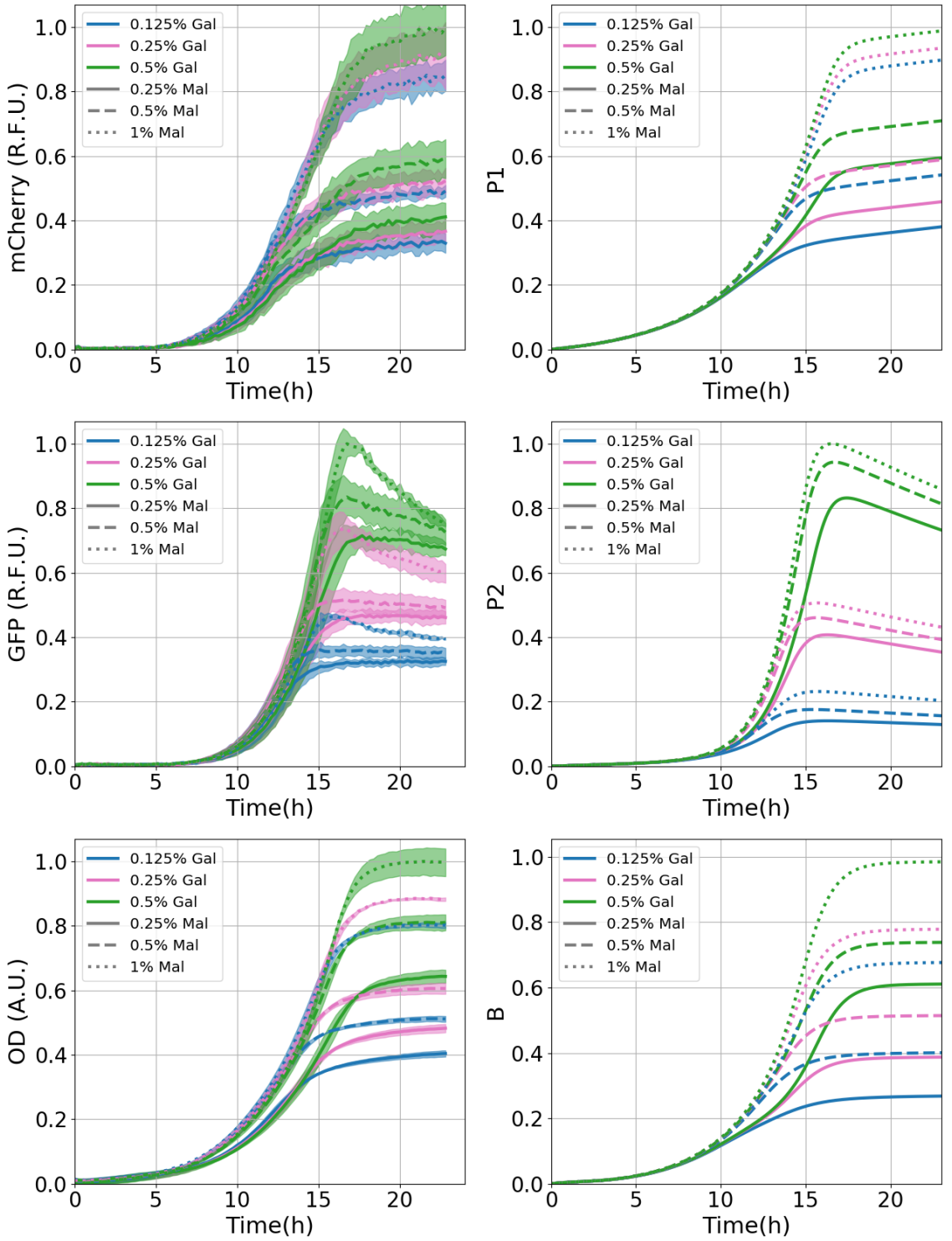


Figure A.26: Fit to binary-sugar mixture experimental data set PR1 with upper bounds. This corresponds to the parameterisation with the lowest overall RMSE score. **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

Fits to all three data sets with PR1 parameters (with upper bounds)

This results correspond to the lowest overall RMSE PR1 parameterisation with the lower and upper bounds.

Table A.72: Initial values of state variables resulting from the grid search performed with the lowest overall RMSE PR1 parameterisation.

	B	**G**	S_1	S_2	P_1	P_2
PR1	0.15	0.08	$1e^{-6}$	$1e^{-6}$	$1e^{-6}$	$5.5e^{-6}$
Maltose	1.0	0.55	$1e^{-2}$	0	$1e^{-5}$	0
Galactose	0.55	0.1	0	$1e^{-6}$	0	$1e^{-4}$

Table A.73: RMSE scores divided by curve phase.

Overall RMSE	0.05244491
Lag RMSE	0.0292053
Exponential RMSE	0.0734784
Stationary RMSE	0.0447143

Table A.74: RMSE scores divided by curve phase and experimental data set (single sugar and sugar mixture).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per experiment**	
PR1	0.031965	0.061611	0.035447	**0.044997**	
Maltose	0.030500	0.093019	0.045807	**0.062399**	
Galactose	0.024634	0.061228	0.051413	**0.048301**	
RMSE per phase	**0.029205**	**0.073478**	**0.044714**	**0.052445**	**Total RMSE score**

Table A.75: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the sugar mixture experimental data set (PR1).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.049485	0.068934	0.024907	**0.051059**	
GFP	0.022634	0.070588	0.055080	**0.053319**	
OD	0.010207	0.040661	0.010741	**0.024986**	
RMSE per phase	**0.031965**	**0.061611**	**0.035447**	**0.044997**	**Total RMSE score for the PR1 fit**

Table A.76: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (maltose).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
mCherry	0.038634	0.097107	0.036106	**0.063839**	
OD	0.019182	0.088742	0.053787	**0.060926**	
RMSE per phase	**0.030500**	**0.093019**	**0.045807**	**0.062399**	**Total RMSE score for the maltose fit**

Table A.77: RMSE scores divided by curve phase and data type (fluorescence and absorbance) for the single sugar experimental data set (galactose).

	Lag phase	**Exponential phase**	**Stationary phase**	**RMSE per data type**	
GFP	0.023005	0.068413	0.062065	**0.054959**	
OD	0.026161	0.053080	0.037875	**0.040564**	
RMSE per phase	**0.024634**	**0.061228**	**0.051413**	**0.048301**	**Total RMSE score for the galactose fit**

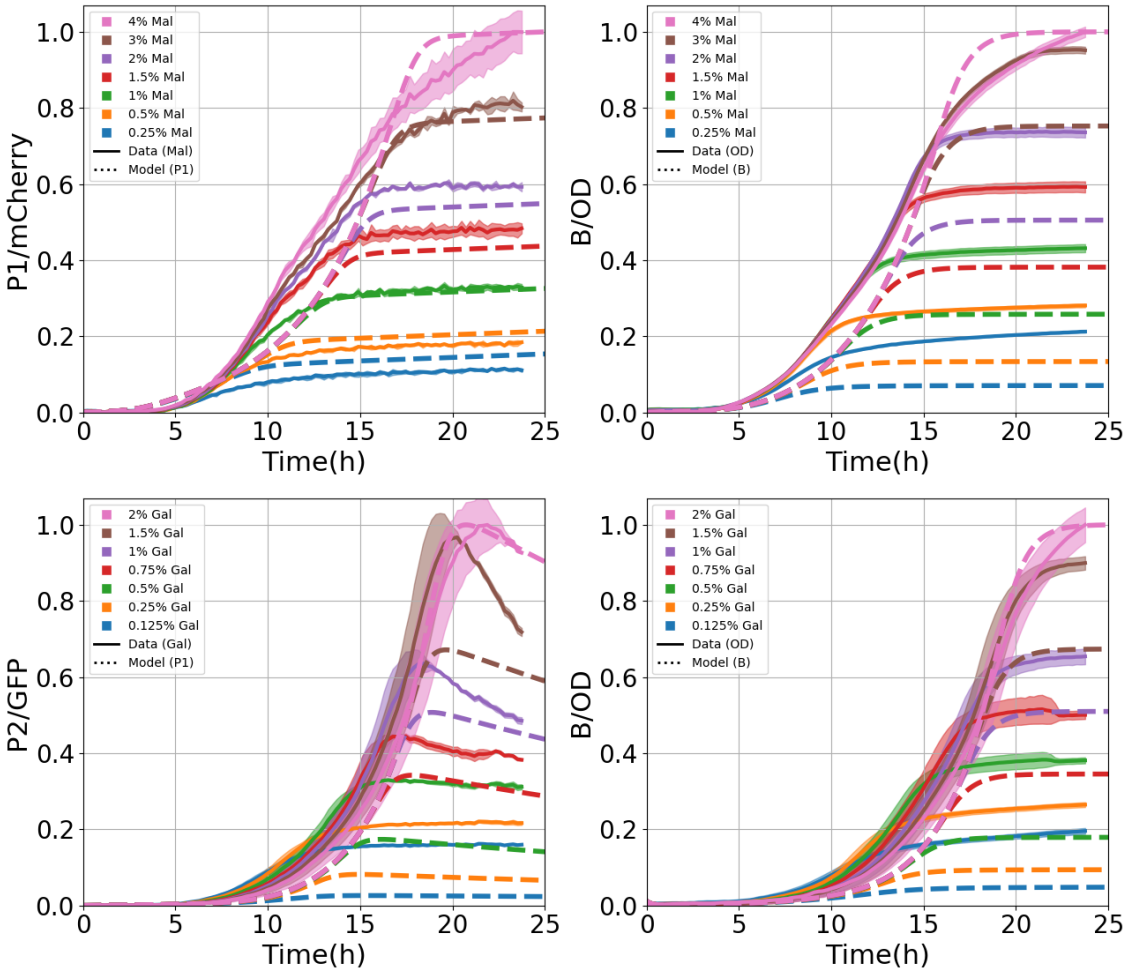


Figure A.27: Fit to the single sugar experimental data sets with the PR1 parameterisation (with upper bounds) with the lowest overall RMSE score after performing a grid search for new initial state variables values (table A.72). **Top**: The maltose data set. **Bottom**: galactose data set. Graphs to the left correspond to the fit to the fluorescence data (mCherry in the case of maltose, GFP in the case of galactose), whilst the graphs to the right present the results of the fit to the absorbance data (O.D.).

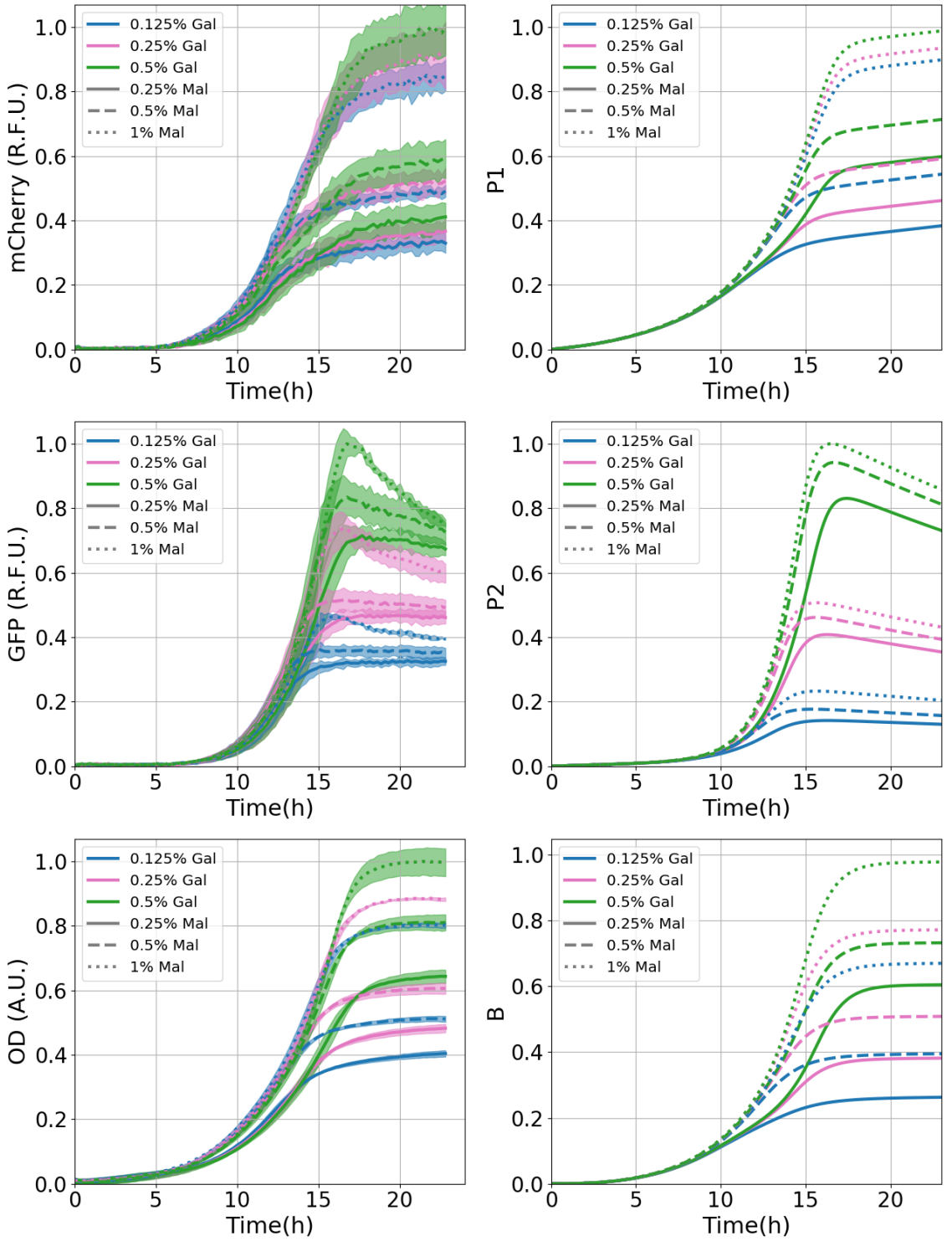


Figure A.28: Fit to the binary-sugar mixture experimental data set PR1 with the PR1 parameterisation (with upper bounds) with the lowest overall RMSE score after performing a grid search for new initial state variables values (table A.72). **Top**: These graphs correspond to the fit to the mCherry fluorescence data, which corresponds to the concentration of the Mal12p. **Middle**: These graphs correspond to the fit to the GFP fluorescence data, which corresponds to the concentration of the Gal10p. **Bottom**: These graphs present the results of the fit to the absorbance data (O.D.). Graphs to the left present the experimental data whilst graphs to the right present the model results.

Appendix B

Supplementary information for Chapter 5

B.1 Supplementary bifurcation diagrams under equal-value alternatives conditions

B.1.1 Hill-RFO model

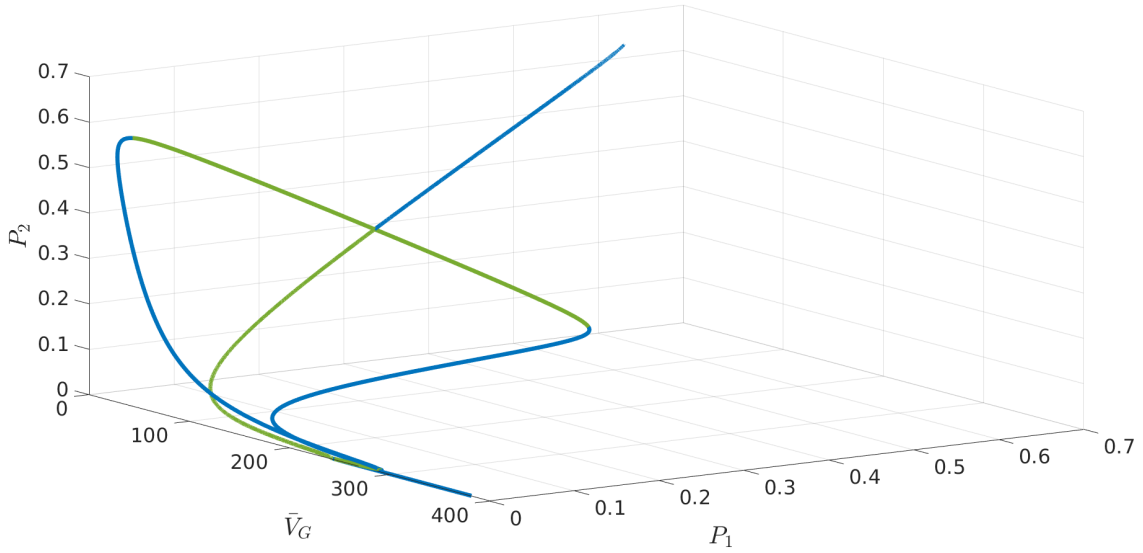


Figure B.1: Three-dimensional bifurcation diagram resulting from varying the $\bar{V}_G$ parameter in the P_1, P_2 space.

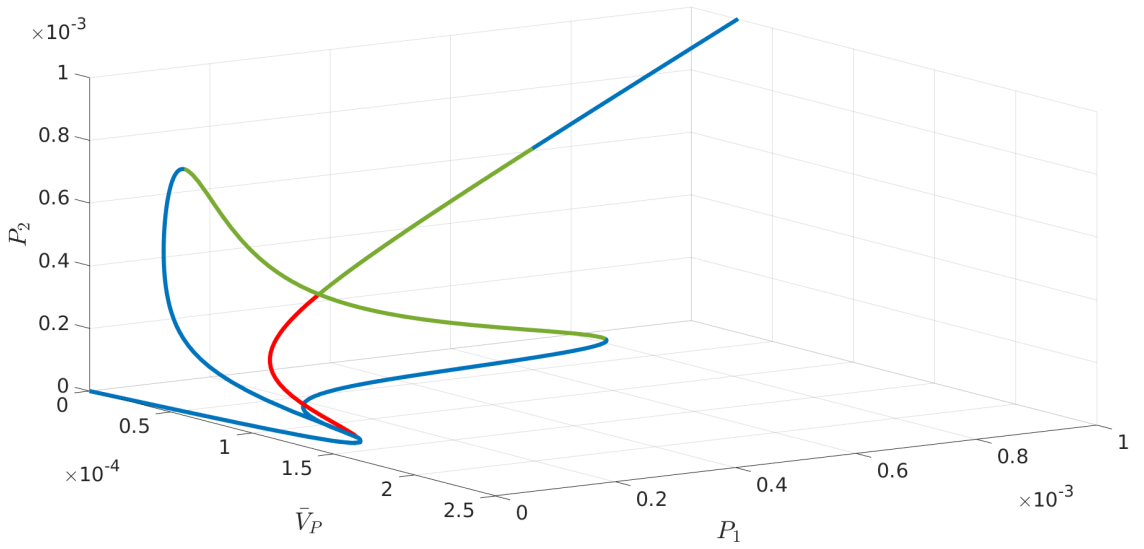


Figure B.2: Three-dimensional bifurcation diagram resulting from varying the $\bar{V}_P$ parameter in the P_1, P_2 space.

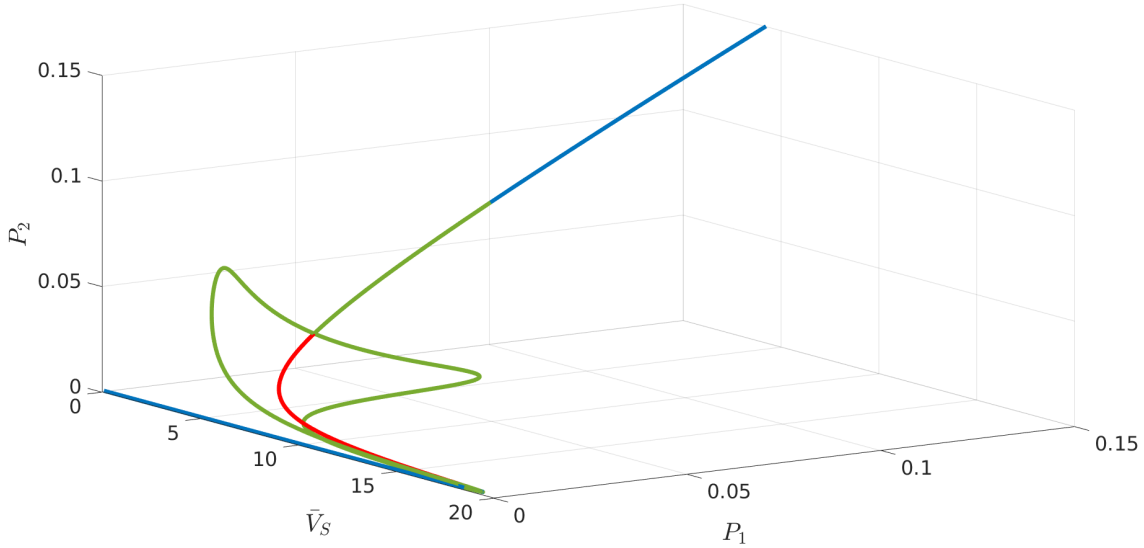


Figure B.3: Three-dimensional bifurcation diagram resulting from varying the $\bar{V}_S$ parameter in the P_1, P_2 space.

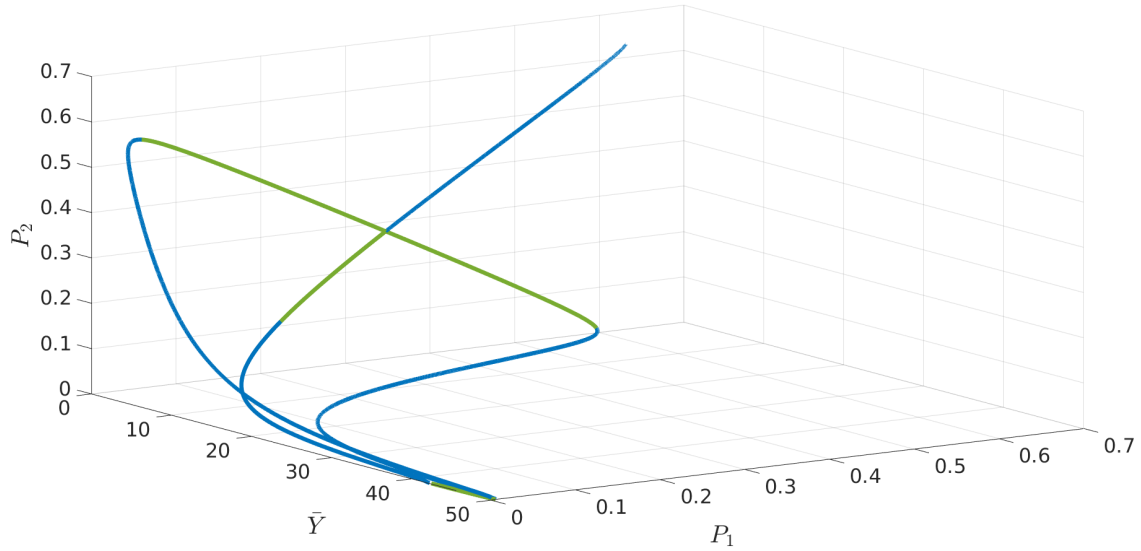


Figure B.4: Three-dimensional bifurcation diagram resulting from varying the $\bar{Y}$ parameter in the P_1, P_2 space.

B.1.2 Logistic-RFO model

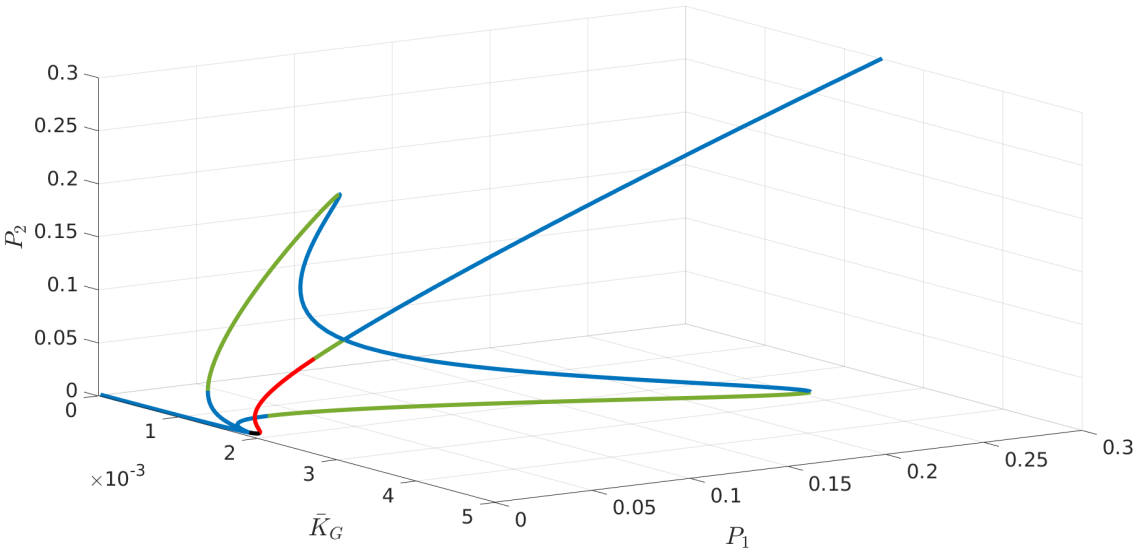


Figure B.5: Three-dimensional bifurcation diagram resulting from varying the $\bar{K}_G$ parameter in the P_1, P_2 space. In this diagram, the small black section of the solution that is near the first bifurcation point is an unstable region in which three eigenvalues have a positive real component.

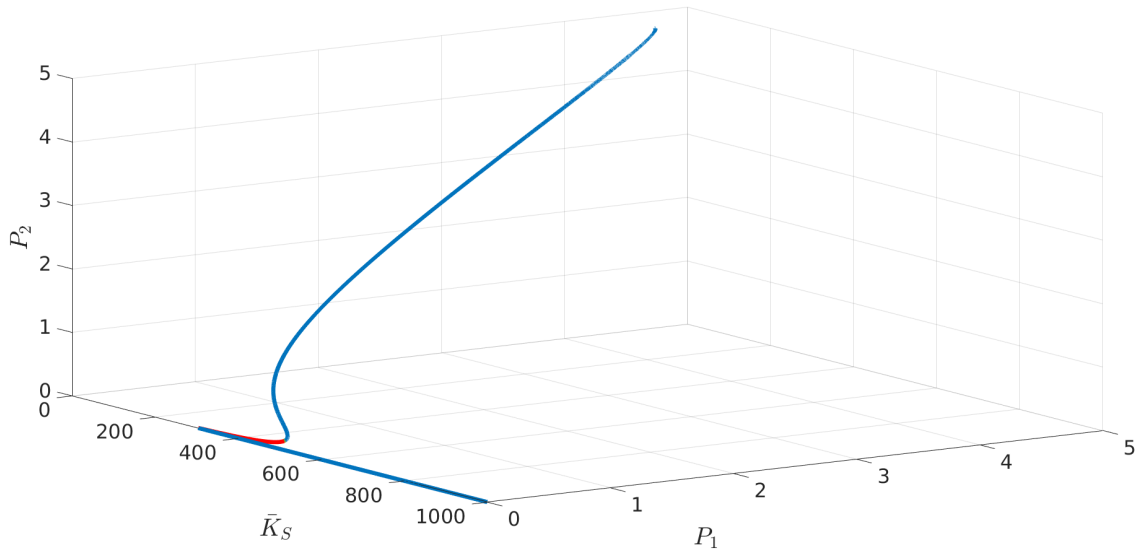


Figure B.6: Three-dimensional bifurcation diagram resulting from varying the $\bar{K}_S$ parameter in the P_1, P_2 space.

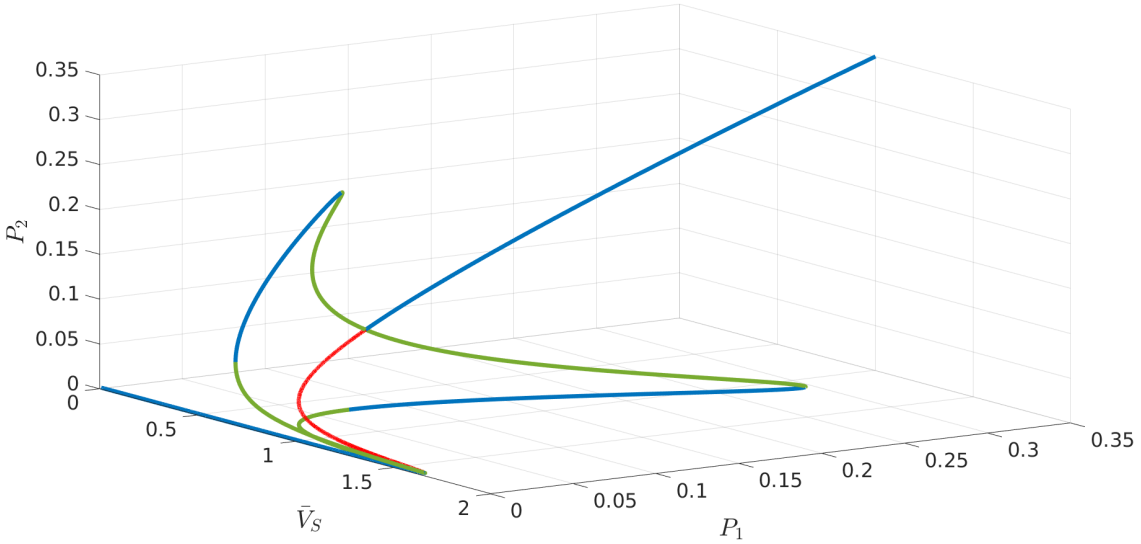


Figure B.7: Three-dimensional bifurcation diagram resulting from varying the $\bar{V}_S$ parameter in the P_1, P_2 space.

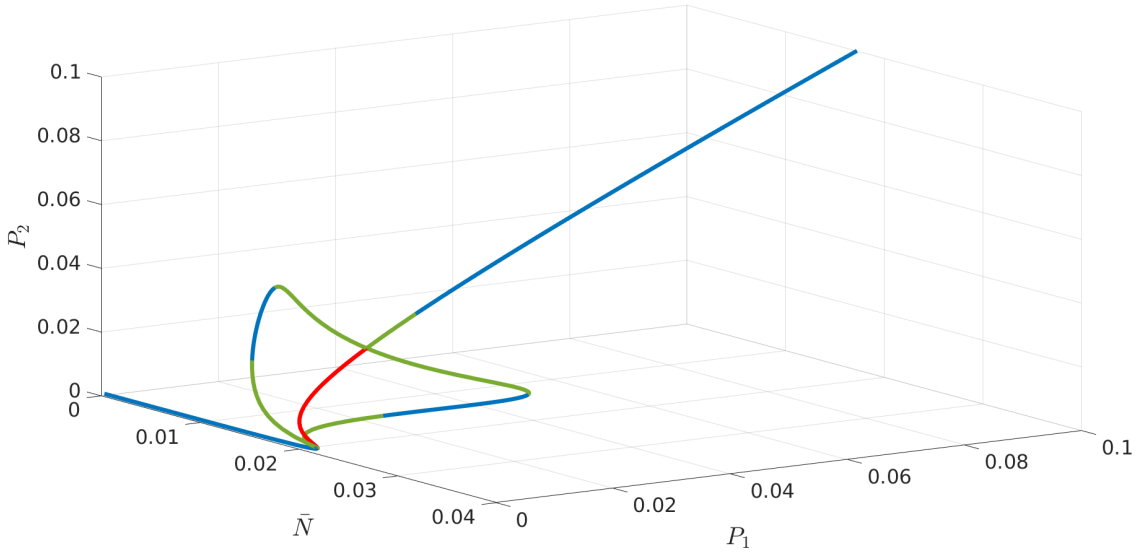


Figure B.8: Three-dimensional bifurcation diagram resulting from varying the $\bar{N}$ parameter in the P_1, P_2 space.

B.2 Parameters and initial conditions used in the dynamical systems analysis

B.2.1 Equal-value alternatives

Hill-RFO model

Table B.1: Mean and difference per parameter used for the analysis in MatConT for the Hill-RFO model under equal-value alternatives conditions.

Parameter	Value	Parameter	Value	Parameter	Value
$\bar{A}$	8.02e-3	ΔA	0.00e+0	V_B	9.00e-1
$\bar{V}_S$	5.00e+2	ΔV_S	0.00e+0	K_G	1.85e+0
$\bar{V}_P$	4.00e-3	ΔV_P	0.00e+0	M	3.00e-1
$\bar{V}_G$	1.15e+0	ΔV_G	0.00e+0	Dcy_B	2.52e-2
$\bar{K}_{inh}$	7.50e-1	ΔK_{inh}	0.00e+0		
$\bar{K}_S$	9.50e-2	ΔK_S	0.00e+0		
$\bar{I}$	4.00e-2	ΔI	0.00e+0		
$\bar{K}_P$	8.00e-2	ΔK_P	0.00e+0		
$\bar{Dcy}_P$	6.00e-2	ΔDcy_P	0.00e+0		
$\bar{Y}$	2.00e-1	ΔY	0.00e+0		
$\bar{N}$	1.47e+1	ΔN	0.00e+0		

Table B.2: Values of state variables used for the analysis of the Hill-RFO model under equal-value alternatives conditions. The first row shows the initial values of the variables, which were taken from the grid search process in order to perform the fit to the experimental data. The second row shows the values of the state variables in the steady-state.

	B	**G**	S_1	S_2	P_1	P_2
Initial conditions	0.1	0.3	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$
steady-state	35.70825	142.28717	278.34088	278.34088	0.62908	0.62908

Logistic-RFO model

Table B.3: Mean and difference per parameter used for the analysis in MatConT for the logistic-RFO model under equal-value alternatives conditions.

Parameter	**Value**	**Parameter**	**Value**	**Parameter**	**Value**
$\bar{A}$	8.02e-3	ΔA	0.00e+0	V_B	4.00e-1
$\bar{V}_S$	7.00e+2	ΔV_S	0.00e+0	K_G	2.00e+0
$\bar{V}_P$	7.50e-3	ΔV_P	0.00e+0	M	3.00e-1
$\bar{V}_G$	8.00e-1	ΔV_G	0.00e+0	Dcy_B	2.50e-3
$\bar{K}_{inh}$	7.00e-1	ΔK_{inh}	0.00e+0		
$\bar{K}_S$	9.50e-1	ΔK_S	0.00e+0		
$\bar{I}$	1.50e-2	ΔI	0.00e+0		
$\bar{K}_P$	2.00e-2	ΔK_P	0.00e+0		
$\bar{Dcy}_P$	6.00e-2	ΔDcy_P	0.00e+0		
$\bar{Y}$	2.40e-1	ΔY	0.00e+0		
$\bar{N}$	1.47e+1	ΔN	0.00e+0		

Table B.4: Values of state variables used for the analysis of the logistic-RFO model under equal-value alternatives conditions. The first row shows the initial values of the variables, which were taken from the grid search process in order to perform the fit to the experimental data. The second row shows the values of the state variables in the steady-state.

	B	**G**	S_1	S_2	P_1	P_2
Initial conditions	0.1	0.1	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$
steady-state	157.6747	5423.6589	5567.5225	5567.5225	4.7281985	4.7281985

B.2.2 Unequal-value alternatives

Hill-RSO model

Table B.5: Mean and difference per parameter used for the analysis in MatConT for the Hill-RSO model under unequal-value alternative conditions.

Parameter	Value	Parameter	Value	Parameter	Value
$\bar{A}$	6.1908e-5	ΔA	1.2182e-4	V_B	8.5639e-1
$\bar{V}_S$	5.5000e+2	ΔV_S	1.0000e+2	K_G	1.8664e+0
$\bar{V}_P$	2.5431e-2	ΔV_P	1.5063e-2	M	2.1275e-1
$\bar{V}_G$	1.0376e+0	ΔV_G	-2.1788e-1	Dcy_B	2.6292e-2
$\bar{K}_{inh}$	2.6151e+0	ΔK_{inh}	1.0697e+0		
$\bar{K}_S$	8.4670e-2	ΔK_S	-4.6961e-2		
$\bar{I}$	1.8581e+0	ΔI	-1.1250e+0		
$\bar{K}_P$	8.6509e-2	ΔK_P	5.6145e-2		
$\bar{Dcy}_P$	9.5517e-3	ΔDcy_P	-1.9101e-2		
$\bar{Y}$	1.6872e-1	ΔY	-2.8588e-2		
$\bar{N}$	1.1050e+1	ΔN	7.3000e+0		

Table B.6: Values of state variables used for the analysis of the Hill-RSO model under unequal-value alternative conditions. The first row shows the initial values of the variables, which were taken from the grid search process in order to perform the fit to the experimental data. The second row shows the values of the state variables in the steady-state.

	B	G	S_1	S_2	P_1	P_2
Initial conditions	0.1	0.3	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$
steady-state	32.57154	696.38876	4158.4934	0.038559146	6.9295177	8.2242453e-05

Hill-RFO model

Table B.7: Mean and difference per parameter used for the analysis in MatConT for the Hill-RFO model under unequal-value alternative conditions.

Parameter	Value	Parameter	Value	Parameter	Value
$\bar{A}$	6.1917e-3	ΔA	1.2381e-2	V_B	9.0752e-1
$\bar{V}_S$	5.4999e+2	ΔV_S	9.9992e+1	K_G	1.8099e+0
$\bar{V}_P$	6.6902e-3	ΔV_P	-1.0803e-3	M	2.8695e-1
$\bar{V}_G$	1.0514e+0	ΔV_G	-1.9686e-1	Dcy_B	2.3757e-2
$\bar{K}_{inh}$	7.3118e-1	ΔK_{inh}	-3.5450e-2		
$\bar{K}_S$	8.7456e-2	ΔK_S	-3.6700e-2		
$\bar{I}$	3.5629e-1	ΔI	2.2703e-1		
$\bar{K}_P$	1.8322e-1	ΔK_P	-1.3914e-1		
$\bar{Dcy}_P$	9.2341e-3	ΔDcy_P	-1.8466e-2		
$\bar{Y}$	2.6010e-1	ΔY	8.0364e-2		
$\bar{N}$	1.1050e+1	ΔN	7.3000e+0		

Table B.8: Values of state variables used for the analysis of the Hill-RFO model under unequal-value alternative conditions. The first row shows the initial values of the variables, which were taken from the grid search process in order to perform the fit to the experimental data. The second row shows the values of the state variables in the steady-state.

	B	**G**	$\boldsymbol{S_1}$	$\boldsymbol{S_2}$	$\boldsymbol{P_1}$	$\boldsymbol{P_2}$
Initial conditions	0.1	0.3	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$
steady-state	38.2009	1739.0998	5516.7008	-1.5023	10.9757	-0.0035

Logistic-RSO model

Table B.9: Mean and difference per parameter used for the analysis in MatConT for the logistic-RSO model under unequal-value alternative conditions.

Parameter	**Value**	**Parameter**	**Value**	**Parameter**	**Value**
$\bar{A}$	3.6596e-2	ΔA	7.3174e-2	V_B	5.3641e-1
$\bar{V}_S$	5.5000e+2	ΔV_S	9.9996e+1	K_G	9.9963e+0
$\bar{V}_P$	9.1238e-4	ΔV_P	6.0746e-4	M	3.6981e-1
$\bar{V}_G$	4.2518e+0	ΔV_G	-5.0795e-1	Dcy_B	1.2799e-4
$\bar{K}_{inh}$	1.8428e+0	ΔK_{inh}	3.3137e+0		
$\bar{K}_S$	7.7228e-2	ΔK_S	-3.8841e-2		
$\bar{I}$	3.7582e+0	ΔI	-4.0833e+0		
$\bar{K}_P$	4.1544e-2	ΔK_P	-4.0356e-2		
$\bar{Dcy}_P$	8.5129e-3	ΔDcy_P	-1.7024e-2		
$\bar{Y}$	1.1633e-1	ΔY	7.0900e-2		
$\bar{N}$	1.1050e+1	ΔN	7.3000e+0		

Table B.10: Values of state variables used for the analysis of the logistic-RSO model under unequal-value alternative conditions. The first row shows the initial values of the variables, which were taken from the grid search process in order to perform the fit to the experimental data. The second row shows the values of the state variables in the steady-state.

	B	**G**	$\boldsymbol{S_1}$	$\boldsymbol{S_2}$	$\boldsymbol{P_1}$	$\boldsymbol{P_2}$
Initial conditions	0.3	0.3	$1e^{-7}$	$1e^{-7}$	$5.5e^{-6}$	$5e^{-6}$
steady-state	4189.2964	919463.07	812454.33	23.752262	1548.02	0.042800646

Logistic-RFO model

Table B.11: Mean and difference per parameter used for the analysis in MatConT for the logistic-RFO model under unequal-value alternative conditions.

Parameter	Value	Parameter	Value	Parameter	Value
$\bar{A}$	2.2173e-3	ΔA	3.7817e-3	V_B	8.4318e-1
$\bar{V}_S$	6.5000e+2	ΔV_S	-1.0001e+2	K_G	1.9716e+0
$\bar{V}_P$	6.1516e-2	ΔV_P	-4.6423e-2	M	3.1884e-2
$\bar{V}_G$	8.6622e-1	ΔV_G	7.5877e-2	Dcy_B	1e-06
$\bar{K}_{inh}$	7.5400e-1	ΔK_{inh}	2.0428e-1		
$\bar{K}_S$	1.5378e+0	ΔK_S	7.8130e-1		
$\bar{I}$	2.3219e-1	ΔI	3.3081e-1		
$\bar{K}_P$	3.4649e-1	ΔK_P	-3.0702e-1		
$\bar{Dcy}_P$	1.3686e-2	ΔDcy_P	-2.7371e-2		
$\bar{Y}$	2.4716e-1	ΔY	-1.6411e-1		
$\bar{N}$	1.1050e+1	ΔN	7.3000e+0		

Table B.12: Values of state variables used for the analysis of the logistic-RFO model under unequal-value alternative conditions. The first row shows the initial values of the variables, which were taken from the grid search process in order to perform the fit to the experimental data. The second row shows the values of the state variables in the steady-state.

	B	**G**	$\boldsymbol{S_1}$	$\boldsymbol{S_2}$	$\boldsymbol{P_1}$	$\boldsymbol{P_2}$
Initial conditions	0.1	0.1	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$
steady-state	732442.75	2828779.8	13879005	179.149	23352.507	0.29244092

B.3 Bifurcation analysis in MatCont

The process followed in order to perform a bifurcation analysis in MatCont is as follows: We start by defining the dynamic system in MatCont. In order to do this, one must initialise MatCont in the MATLAB command window, which will cause several windows will emerge. The main window, labelled *MatCont* has a drop-down menu labelled *Select*, among the options one must chose *System*, and furthermore select either *New* or *Edit/Load*, depending on whether or not the system has been previously defined. We define the model as defined in Chapter 3, leaving the N_i equations out since the extracellular sugar concentration is considered to be constant. Following this we find the steady-state of the system. This was done by integrating over time until the state variable values no longer change significantly. This is evaluated manually.

In order to do this, once the system has been loaded, one selects the option *Initial point* → *Point*, in the sub-menu *Type*, found in the *MatCont* window. The *Integrator* window allows us to specify different hyper-parameters such as the maximum step size, the number of steps, as well as integrating method to be used, among others. As mentioned before, we used method ode15s. The number of steps is not necessarily relevant, as long as we choose an amount large enough for the system to reach a steady-state. The *Starter* window allows us to specify the value of the parameters of the system, as well as the initial value of the state variables and the time. For the results presented here we have chosen the initial parameterisation obtained through manual visualisation before performing the parameter optimisation, as well as the initial conditions for the state variables, all specified in the previous chapter.

Once the steady-state is found, we select this set of state variable values by choosing the last point of the time series that has just been integrated. We find this in the *Initial point* option from the *Select* drop-down menu in the *MatCont* window. We use this to analyse how this solution changes as the value of one parameter of interest is either increased or decreased. In order to do this, and once we have selected the steady-state as the initial point, we select the option *Initial point* → *Equilibrum*, found in the sub-menu *Type*, in the *MatCont* window. Subsequently, we select the parameter that we wish to change in the *Starter* window. In order to start the integration, in the *MatCont* window, we select either *Forward* or *Backward* in the *Compute* menu, depending on the direction that we wish to change the selected parameter. The *Continuer* window allows us to specify different hyper-parameters such as the maximum and minimum step size and the number of steps.

Once we have integrated over the parameter value range of interest, if our curve contains any critical points such as a *Branching point* or a *Limit Point* for instance, we can select for them in the *Initial point* option from the *Select* drop-down menu in the *MatCont* window, where one can find the first and last points of the series one has just integrated, as well as any critical points that have been detected during the computation. The bifurcations diagrams presented here, were obtained by first

selecting the critical point found during the integration of our steady-state solution. Then we choose the option *Curve* → *Equilibrium*, found in the sub-menu *Type*, in the *MatCont* window. This allows us to explore the different solutions that emerge from this point. Similarly, if the point to explore is a *Limit Point*, we can explore it by choosing the option *Curve* → *Equilibrium*, found in the sub-menu *Type*, in the *MatCont* window. In some cases, the solution yielded a *Hopf Point*, which usually leads to oscillatory behaviour. In order to analyse this, one can select this point, which automatically changes the type of curve, from *Curve* → *Equilibrium* to *Limit Cycle*. The analysis and interpretation of the bifurcation diagrams obtained is greatly informed by Strogatz (1994) and Seydel (1988).

B.4 Saturated model supplementary material

The reduced RFO (eq. B.1) and reduced RSO (eq. B.2) models read

$$\begin{aligned}
\frac{dP_1}{dt} &= V_{P1}\,A_1 - Dcy_{P1}\,P_1 + \\
&\qquad \frac{V_{P1}\,(\Omega - P_1 - P_2)\,(\Lambda_1\,P_1 + \Lambda_2\,P_2\,I_1)^2}{(\Lambda_1\,P_1 + \Lambda_2\,P_2\,I_1)^2 + (K_{P1}\,(1 + \frac{\Upsilon_1\,\Lambda_1\,P_1 + \Upsilon_2\,\Lambda_2\,P_2}{Kinh_1}))^2} \\
\frac{dP_2}{dt} &= V_{P2}\,A_2 - Dcy_{P2}\,P_2 + \\
&\qquad \frac{V_{P2}\,(\Omega - P_1 - P_2)\,(\Lambda_2\,P_2 + \Lambda_1\,P_1\,I_2)^2}{(\Lambda_2\,P_2 + \Lambda_1\,P_1\,I_2)^2 + (K_{P2}\,(1 + \frac{\Upsilon_2\,\Lambda_2\,P_2 + \Upsilon_1\,\Lambda_1\,P_1}{Kinh_2}))^2}
\end{aligned} \tag{B.1}$$

$$\frac{dP_1}{dt} = V_{P1}\,A_1 - Dcy_{P1}\,P_1 + \frac{V_{P1}\,(\Omega - P_1 - P_2)\,(\Lambda_1\,P_1 + \Lambda_1\,P_1\,\Lambda_2\,P_2\,I_1)^2}{(\Lambda_1\,P_1 + \Lambda_1\,P_1\,\Lambda_2\,P_2\,I_1)^2 + (K_{P1}\,(1 + \frac{\Upsilon_1\,\Lambda_1\,P_1 + \Upsilon_2\,\Lambda_2\,P_2}{Kinh_1}))^2}$$

$$\frac{dP_2}{dt} = V_{P2}\,A_2 - Dcy_{P2}\,P_2 + \frac{V_{P2}\,(\Omega - P_1 - P_2)\,(\Lambda_2\,P_2 + \Lambda_2\,P_2\,\Lambda_1\,P_1\,I_2)^2}{(\Lambda_2\,P_2 + \Lambda_2\,P_2\,\Lambda_1\,P_1\,I_2)^2 + (K_{P2}\,(1 + \frac{\Upsilon_2\,\Lambda_2\,P_2 + \Upsilon_1\,\Lambda_1\,P_1}{Kinh_2}))^2}, \tag{B.2}$$

where

$$\Omega = \frac{M\,V_B}{Dcy_B}\ , \qquad \Lambda_i = \frac{V_{Si}}{(Y_i\,V_{Gi} + V_B)}\ , \qquad \Upsilon_i = \frac{V_{Gi}\,Y_i}{V_B}. \tag{B.3}$$

Table B.13: Values of state variables used in the analysis of the saturated models under unequal-value alternative conditions. These values correspond to the test with the logistic-RFO (B.11) and the logistic-RSO (B.9) parameterisations. The first row shows the initial values of the variables, which were taken from the grid search process in order to perform the fit to the experimental data. The second row shows the values of the state variables in the steady-state.

	RFO		**RSO**	
	$\boldsymbol{P_1}$	$\boldsymbol{P_2}$	$\boldsymbol{P_1}$	$\boldsymbol{P_2}$
Initial conditions	$1e^{-5}$	$1e^{-5}$	$5.5e^{-6}$	$5e^{-6}$
steady-state	26882.909	0.42258	1548.7232	0.042904

Table B.14: Mean and difference per parameter used for the analysis in MatConT for the saturated RFO model under unequal-value alternative conditions.

Parameter	Value	Parameter	Value	Parameter	Value
$\bar{A}$	2.2372e-4	ΔA	-4.4543e-4	V_B	4.1891e-1
$\bar{V}_S$	6.5000e+2	ΔV_S	-1.0000e+2	K_G	2.0013e+0
$\bar{V}_P$	2.9657e-2	ΔV_P	3.5750e-2	M	3.0898e-1
$\bar{V}_G$	8.6980e-1	ΔV_G	1.2583e-1	Dcy_B	5.8852e-4
$\bar{K}_{inh}$	7.4811e-1	ΔK_{inh}	1.0572e-1		
$\bar{K}_S$	1.4259e+0	ΔK_S	9.5087e-1		
$\bar{I}$	8.3883e-2	ΔI	1.6776e-1		
$\bar{K}_P$	1.1110e-1	ΔK_P	-7.3583e-2		
$\bar{Dcy}_P$	1.2307e-2	ΔDcy_P	-2.4612e-2		
$\bar{Y}$	1.9533e-1	ΔY	-1.3558e-1		
$\bar{N}$	1.4700e+1	ΔN	0.0000e+0		

Table B.15: Values of state variables used in the analysis of the saturated RFO model under unequal-value alternative conditions with the alternative parameterisation shown in table B.14. The first row shows the initial values of the variables, which were taken from the grid search process in order to perform the fit to the experimental data. The second row shows the values of the state variables in the steady-state.

	RFO	
	P_1	P_2
Initial conditions	$1e^{-5}$	$1e^{-5}$
steady-state	219.92639	2.1349e-4

B.5 Analysis under single-parameter difference conditions

In this section I show the results from the single-parameter difference analysis. For this we have taken the same setup used in the equal-value alternatives case, then, one parameter tuple from the branches of the model was selected, and given the value from the unequal-value alternatives case. This produces a model in which only one of the parameter tuples has different values. With this setup, I proceed to analyse the model following the procedure described in appendix B.3. Additionally, the *Mathematica* package *Dynamica* was used in order to refine the search for all the system's solutions. As it was explained in the section 5.1.2, once a system is declared and parameterised in *Dynamica*, the bifurcation analysis does an search

for solutions through out the entire parameter space that was specified by the user. This contrasts with the procedure used in MatCont, where the starting point needs to be specified by the user before the system can be analysed. The system was analysed in *Dynamica* in order to discover the solutions that are disconnected from the initial steady-state solution. these solutions emerge from the appearance of a *Limit point* completely separated from the initial solution. The coordinates of this point can then be used as input in MatCont in order to numerically integrate that solution.

I present this analysis as a different kind of exploration of the models dynamics. This particular approach to the parameterisation of the model does not represent any one of the experimental conditions discussed in the previous chapter. For this, I have used the Hill-RFO model. All parameters of the model were analysed under different parameterisations. All parameterisations tested have one single non-zero difference parameter: ΔK_{inh}, ΔI, ΔK_P, ΔN, ΔV_P and ΔY. In the case presented here, the non-zero difference parameter is ΔV_P, which is set to $2.50e^{-3}$. This makes $V_{P1} = 6.50e^{-3}$, and $V_{P1} = 4.0e^{-3}$. Tables B.16 and B.17 show the parameterisation and the initial conditions used for these analysis, respectively. With this parameterisation, we encountered bifurcations and multistable behaviour in parameters Dcy_B, $\bar{V}_S$, $\bar{K}_{inh}$, ΔK_{inh}, $\bar{K}_P$ and $\bar{N}$.

Table B.16: Mean and difference per parameter used for the analysis in MatConT for the analysis of the Hill-RFO model with an equal-value alternatives parameterisation with non-zero valued ΔV_P parameter.

Parameter	**Value**	**Parameter**	**Value**	**Parameter**	**Value**
$\bar{A}$	8.02e-3	ΔA	0.00e+0	V_B	9.00e-1
$\bar{V}_S$	5.00e+2	ΔV_S	0.00e+0	K_G	1.85e+0
$\bar{V}_P$	5.25e-3	ΔV_P	2.50e-3	M	3.00e-1
$\bar{V}_G$	1.15e+0	ΔV_G	0.00e+0	Dcy_B	2.52e-2
$\bar{K}_{inh}$	7.50e-1	ΔK_{inh}	0.00e+0		
$\bar{K}_S$	9.50e-2	ΔK_S	0.00e+0		
$\bar{I}$	4.00e-2	ΔI	0.00e+0		
$\bar{K}_P$	8.00e-2	ΔK_P	0.00e+0		
$\bar{Dcy}_P$	6.00e-2	ΔDcy_P	0.00e+0		
$\bar{Y}$	2.00e-1	ΔY	0.00e+0		
$\bar{N}$	1.47e+1	ΔN	0.00e+0		

Table B.17: Values of state variables used for the analysis of the Hill-RFO model with an equal-value alternatives parameterisation with non-zero valued ΔV_P parameter. The first row shows the initial values of the variables, which were taken from the grid search process in order to perform the fit to the experimental data. The second row shows the values of the state variables in the steady-state.

	B	**G**	$\boldsymbol{S_1}$	$\boldsymbol{S_2}$	$\boldsymbol{P_1}$	$\boldsymbol{P_2}$
Initial conditions	0.1	0.3	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$
steady-state	35.710519	180.12817	436.73731	268.03771	0.98706754	0.60579053

The diagrams in figures B.9 and B.10 showcase the bifurcations yielded by varying parameters $\bar{K}_{inh}$ and $\bar{K}_P$, respectively. Similarly to the results from the equal-value alternatives case, these bifurcation portraits show a pattern that is maintained by the results corresponding to other parameters, which are not shown. In the case of $\bar{K}_{inh}$, as the value of the parameter is increased from zero, the unique stable solution solution, which can be defined as a deadlock state, reaches a critical point at which its stability changes briefly to an unstable or a saddle solution (in the P_i, P_2 space). There is an observable but relatively small region where this solution presents an S-shaped curve, moving backward until another critical point is found upon which the solution turns stable again. This also forces it to changes its direction back to the original orientation. The diagrams in figure B.9 show how at this point the original stable solution continues to grow in the direction of P_1, whilst the P_2 values of this solution remain close to zero. This can be seen as a commitment to the first alternative whilst ignoring the second one.

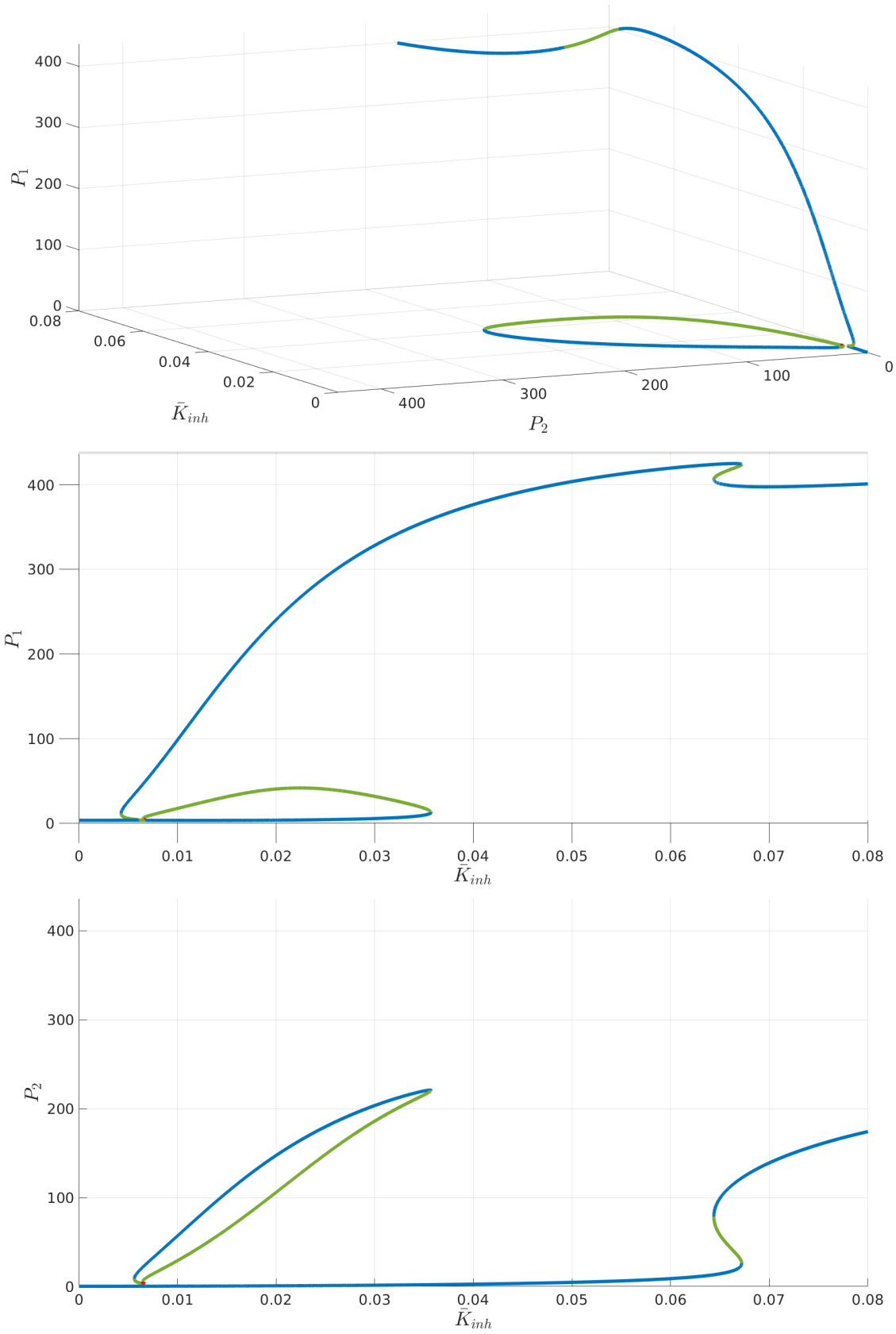


Figure B.9: **Top**: Three-dimensional diagram resulting from varying the $\bar{K}_{inh}$ parameter in the P_1, P_2 space. **Middle** and **Bottom**: Two-dimensional diagrams showing the same solution, focusing on the value of P_1 and P_2, respectively, as $\bar{K}_{inh}$ is changed.

Close to this unstable region, a second solution emerges. This solution has two main parts: 1) A stable part in which the values of P_2 increase, whilst the values of P_1 remain close to zero. Similar to the original solution, this stable region represents commitment to the second alternative. 2) A saddle solution which sits in between the two stable solutions. This secondary solution is circular in its overall shape. The stable and saddle parts of the curve emerge at a critical point from which the separate from one another before meeting again at a second critical point. In this interval of $\bar{K}_{inh}$ values, the system is multistable, as it has two available stable states: commitment to alternative one, and commitment to the second alternative. Since saddle attractor is closer to the latter, it is more likely that a randomly chosen set of state variables would converge toward the solution committed to alternative one. After this multistable region comes to an end, as $\bar{K}_{inh}$ continues to increase, the secondary solution is left behind, leaving the system only one solution available, which represents commitment to alternative one. Finally, the solution presents a second hysteresis loop, in the middle of which the stability changes briefly. For any larger values of $\bar{K}_{inh}$, there is only one stable solution in which the system maintains its commitment to alternative one.

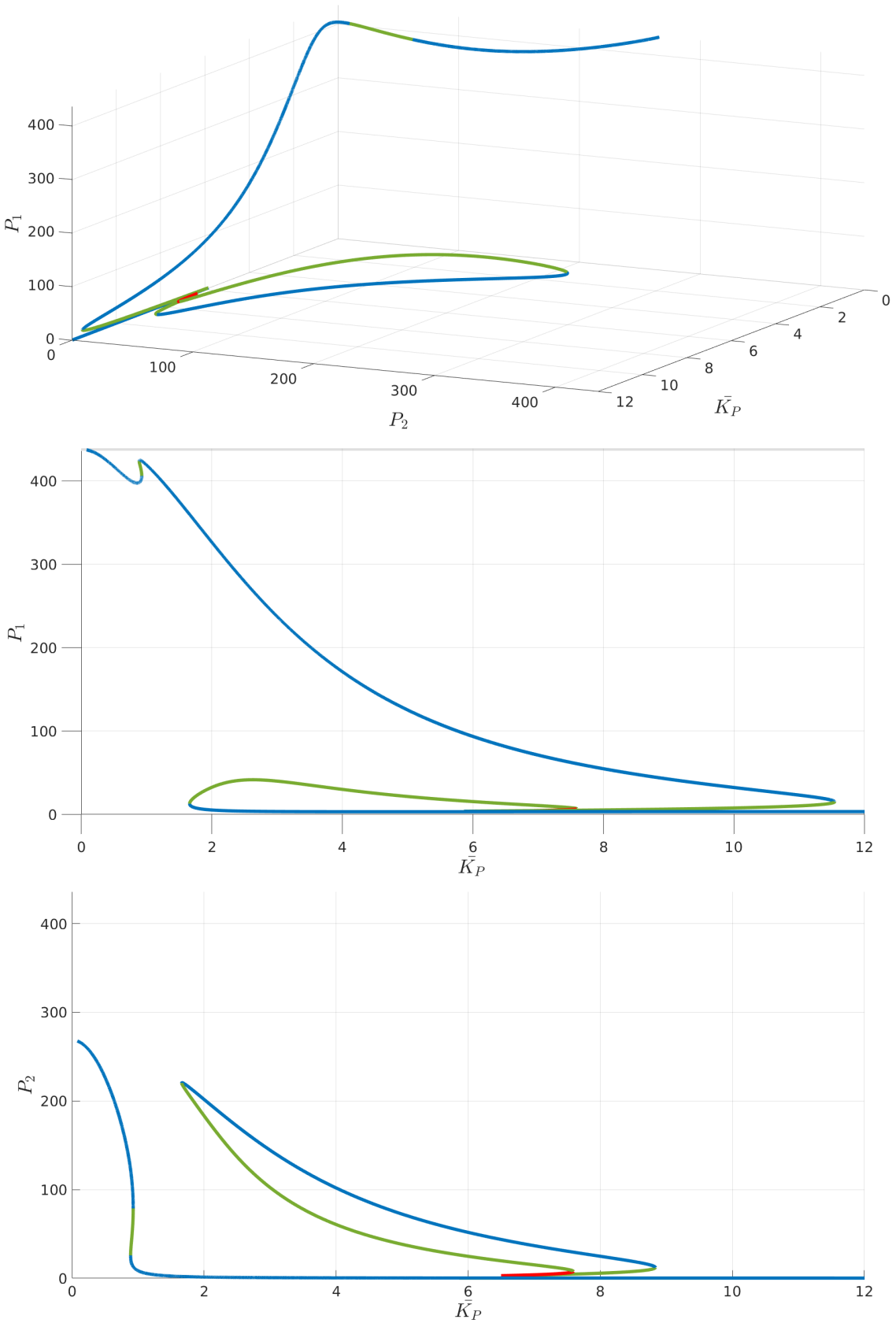


Figure B.10: **Top**: Three-dimensional diagram resulting from varying the $\bar{K}_P$ parameter in the P_1, P_2 space. **Middle** and **Bottom**: Two-dimensional diagrams showing the same solution, focusing on the value of P_1 and P_2, respectively, as $\bar{K}_P$ is changed.

The diagrams cognate to the analysis of parameter $\bar{K}_P$ show a similar pattern of bifurcations. And the relation between parameter value and P_i value is reversed from proportional, in the case of $\bar{K}_{inh}$ to inversely proportional.

It can be seen that with this parameterisation the system presents a bias toward P_1. This bias can be understood by observing that $V_{P1} > V_{P2}$. V_{Pi} is the maximum rate at which P_i is produced by the induction term. The larger the difference between these rates, the larger the difference in concentration between P_1 and P_2, since P_1 is being produced at a higher rate.

Appendix C

Stochastic simulations

In order to perform stochastic simulations a noise component was added to each equation of the Hill-RSO model. For the simulations shown in figure C.1 the noise parameter η_i at the end of each equation was arbitrarily selected.

$$
\begin{aligned}
dB &= (\, V_B \, \frac{G^2}{K_G^2 + G^2} - Dcy_B \, B \,)\, dt \; + \; \eta_B \, dW \\
dG &= (\, Y_1 \, V_{G1} \, S_1 + Y_2 \, V_{G2} \, S_2 - V_B \, G \, \frac{G^2}{K_G^2 + G^2} \,)\, dt \; + \; \eta_G \, dW \\
dN_i &= (\, -V_{Si} \, P_i \, \frac{N_i^2}{K_{Si}^2 + N_i^2} \,)\, dt \; + \; \eta_{N_i} \, dW \\
dS_i &= (\, V_{Si} \, P_i \, \frac{N_i^2}{K_{Si}^2 + N_i^2} - Y_i \, V_{Gi} \, S_i - V_B \, S_i \,)\, dt \; + \; \eta_{S_i} \, dW \\
dP_i &= (\, V_{Pi} \, A_i - Dcy_{Pi} \, P_i + V_{Pi} \, (M \, B - P_1 - P_2) \, RSO_i \,)\, dt \; + \; \eta_{P_i} \, dW.
\end{aligned}
\tag{C.1}
$$

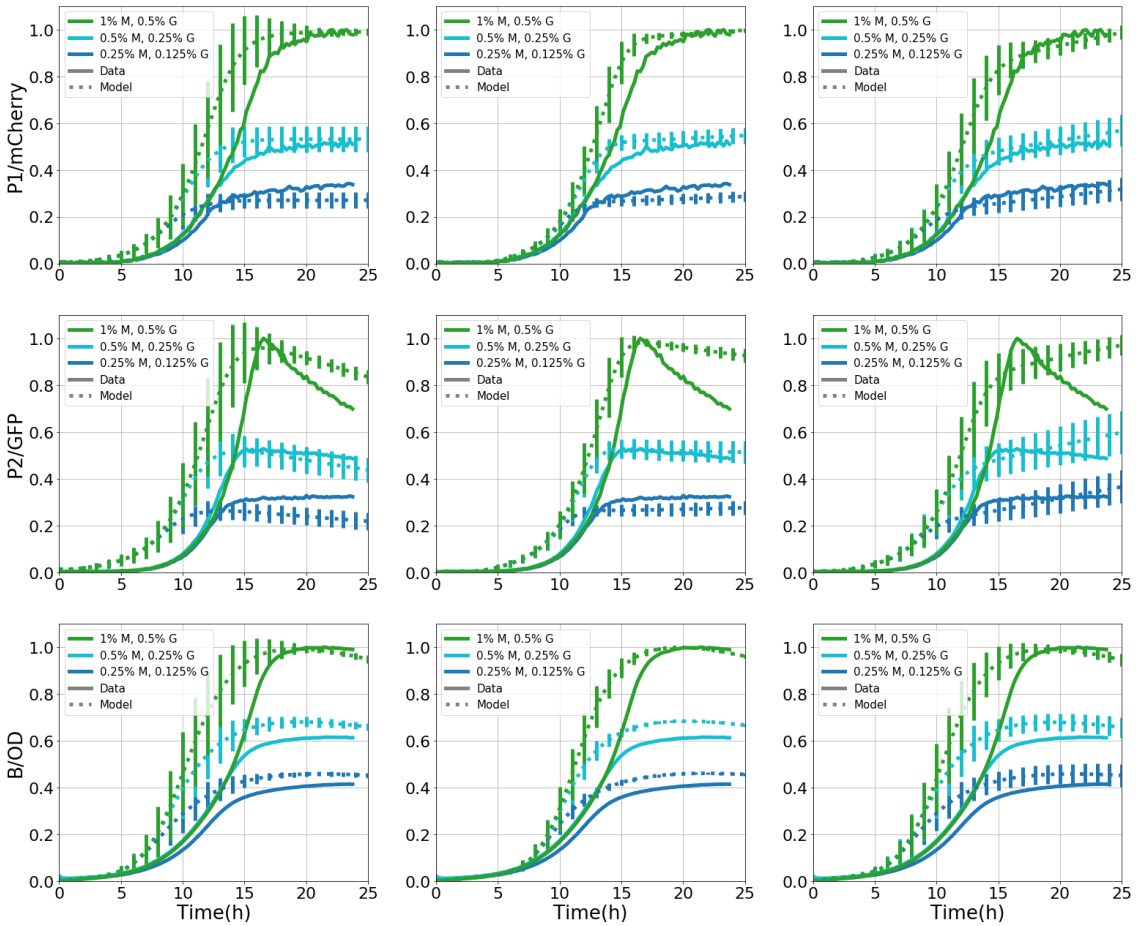


Figure C.1: Stochastic Hill-RSO model simulations. The parameterisation used here corresponds to the PR1 parameterisation shown in table A.2. Each row portrays a different data type (fluorescence and optical density), along with its corresponding model state variable. Each columns displays the results of using a different set of noise parameters: (Left) $B' \to 0.005$, $P_1' \to 0.00025$, $P_2' \to 0.00025$, (Middle) $B' \to 0.005$, $P_1' \to 0.0005$, $P_2' \to 0.0005$, (Right) $B' \to 0.01$, $G' \to 0.05$, $S_1' \to 0.05$, $S_2' \to 0.05$, $P_1' \to 0.0001$, $P_2' \to 0.0001$. All other noise parameters are $= 0$. Each parameterisation was simulated 1000 times. Each plot shows data curves from experimental data set PR1 as well as the mean of the 1000 stochastic simulations with error bars the length of a standard deviation. For visual clarity, only three out of the nine experimental conditions are being shown.