Characterising dynamical systems models of tumour-immune interactions



Ana Osojnik
St Anne's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2021

Za mami.

To Declan.

To Mel.

Acknowledgements

First of all, I would like to thank my academic supervisors, Prof Helen Byrne and Prof Eamonn Gaffney. Without their vision, expertise, enthusiastic guidance, endless support and encouragement this thesis would have never existed. Helen and Eamonn, you both have been a source of inspiration for me, academically and personally, and I am eternally grateful for your mentorship. Thank you for being so generous with your time, and for always giving me a motivation boost in our meetings.

I would like to thank my industrial supervisors, Dr James Yates and Dr Michael Davies, whose ideas, multidisciplinary expertise and guidance have been invaluable. I would also like to thank them for kindly hosting me on my visits to AstraZeneca.

I am hugely grateful to Dr Ben Lambert for his time, stimulating discussions and feedback. His invaluable expertise on parameter inference importantly influenced the work in Chapters 4 and 5.

I would also like to thank my examiners Prof Philip Maini, Prof Gary Mirams, Prof Sarah Waters, Prof David Gavaghan and Dr Enas Abu Shah for their time, comments and suggestions, and especially for interesting discussions about my work during the final viva and confirmation/transfer of DPhil status examinations, and to Dr Chris Farmer for helpful conversations about solving inverse problems.

I would like to thank the core teams of InFoMM EPSRC CDT and WCMB for their continuous efforts in running these two centres that have enriched my DPhil experience on so many levels.

A big thank you goes to all my friends, from InFoMM (special mention to Cohort 3), WCMB, and other wonderful people I am lucky to have met in Oxford and elsewhere, who have been supporting me. Thank you for sharing some of the best and also some difficult times with me.

To my partner, Declan, you are a magical human and my rock. I absolutely could not have done this without you. Thank you for everything, but especially for patiently and lovingly sticking by me at all times during these crazy few years, for always giving me solutions and for never stopping believing in me.

To Declan's family, especially to Debra and Maarten, thank you for all your kindness and support, and for always welcoming me into your family and your home with open arms.

And finally, to my family, I hope you know how grateful I am for your support in everything I have done from day one, even when that was far away from home, and thank you for always believing in me.

Abstract

Despite clinical successes, data from trials of cancer immunotherapies continue to highlight non-responders. A better understanding of tumour-immune interactions is needed to explain outcome variability. In this thesis, we investigate whether tumour-immune interaction models can explain variation in mouse tumour volume time series data, focusing on a well-known mathematical model by Kuznetsov et al. [88].

Exploiting a separation of timescales in the original model, and using matched asymptotics, we derive a new long-timescale approximation of the model, which differs from the quasi-steady-state approximation (QSSA) introduced by Kuznetsov et al. [88], but is in good agreement with numerical solutions of the original model. As well as exhibiting the three stages of immunoediting – elimination, equilibrium and escape, together with complex bifurcations and bistability, our reduced model highlights excitability, a feature not traditionally associated with tumour-immune models. We also identify different biophysical parameters that could be targeted with immunotherapy in order to control tumour size.

As the QSSA model is simpler, but still exhibits immunoediting, we use it for case studies investigating parameter identifiability for complex bifurcating and bistable models using *in silico* data, generated from the model with variable parameters, observables and noise levels. Using profile likelihoods, we show that identifiability varies across dynamically different datasets. Parameters related to the immune system are highly correlated, and can result in large parameter confidence intervals and bimodal likelihood surfaces.

We then explore a population approach, where the model is simultaneously calibrated to dichotomous data from two individuals, noting that variation in tumour outcomes can arise from bistability or bifurcations, representing distinct hypotheses for heterogeneous tumour responses. We demonstrate the hypotheses can be distinguished for most synthetic datasets considered, except for highly noisy tumour data, by comparing Bayesian population fits under each hypothesis.

This thesis provides a new perspective on how qualitative differences between tumour growth dynamics of responders and non-responders may be explained by bifurcating and/or bistable models. At the same time, it illustrates how integration of mathematical and statistical methods can facilitate development of simplified models, their analysis and calibration.

Contents

List of Abbreviations **iv**

1 Introduction **1**
1.1 Motivation . . . 1
1.2 Biological background . . . 4
1.2.1 Immune system and its activation . . . 4
1.2.2 Cancer, cancer-immunity cycle and immunoediting . . . 6
1.2.3 Mechanisms of immune evasion in cancer . . . 9
1.2.4 Immunotherapies . . . 12
1.3 Mathematical models of tumour-immune cell interactions . . . 13
1.3.1 Ordinary differential equations models . . . 13
1.3.1.1 Models including tumour and immune cells . . . 14
1.3.1.2 Models including multiple cell types and signalling molecules . . . 16
1.3.1.3 Parameter estimation . . . 18
1.3.1.4 Linking models to data . . . 23
1.3.2 Other modelling frameworks . . . 25
1.3.3 Summary and discussion . . . 27
1.4 Thesis outline . . . 28

2 An excitable model of tumour-immune interactions **31**
2.1 Introduction . . . 31
2.2 Mathematical model . . . 33
2.2.1 Kuznetsov et al.'s original five-compartment model . . . 34
2.2.2 Parameter assumptions and the dimensionless model . . . 37
2.2.3 Parameter values . . . 38
2.3 Matched asymptotic model approximation . . . 40
2.3.1 Asymptotic approximation of fast dynamics . . . 41
2.3.2 Asymptotic approximation of slow dynamics . . . 43
2.3.3 Comparison of numerical solutions . . . 44
2.4 Characterisation of model excitability . . . 45

2.4.1 Nullclines . . . 46
2.4.2 Separation of timescales and its implications . . . 48
2.5 Comparison with the QSSA model . . . 51
2.6 Discussion and conclusions . . . 53

3 Characterisation and parameter sensitivity of qualitative model behaviour 57
3.1 Introduction . . . 57
3.2 Characterisation of steady state solutions . . . 59
3.2.1 Asymptotic approximation of steady states . . . 60
3.2.2 Physicality of steady states . . . 61
3.2.3 Linear stability of steady states . . . 63
3.3 Bifurcation structure for default parameter values as the basal effector supply rate varies . . . 65
3.4 Changes in the bifurcation structure as default parameters vary . . . 72
3.4.1 Varying σ and δ . . . 72
3.4.2 Varying ρ . . . 75
3.4.3 Varying μ . . . 77
3.4.4 Varying η . . . 77
3.4.5 Summary . . . 78
3.5 Comparison with the QSSA model . . . 78
3.6 Discussion and conclusions . . . 80

4 Parameter identifiability of a model of tumour-immune interactions 85
4.1 Introduction . . . 85
4.2 Formulation of the inverse problem . . . 88
4.2.1 Forward model . . . 88
4.2.2 Observation models . . . 88
4.2.3 Likelihood . . . 90
4.3 Structural identifiability . . . 90
4.3.1 Definitions . . . 90
4.3.2 Methods and results . . . 91
4.4 Practical identifiability . . . 94
4.4.1 Methods . . . 96
4.4.1.1 Generation of synthetic data . . . 96
4.4.1.2 Simplified parameter estimation problem . . . 98
4.4.1.3 Profile likelihood approach . . . 99
4.4.2 Results . . . 103
4.4.2.1 Case study 4.1: elimination regime . . . 103

4.4.2.2 Case study 4.2: escape regime . . . 110
4.4.2.3 Case study 4.3: equilibrium regime . . . 118
4.5 Discussion and conclusions . . . 121

5 Distinguishing hypotheses for dichotomous tumour growth curves 129
5.1 Introduction . . . 129
5.2 Methods . . . 131
5.2.1 Generation of synthetic data . . . 132
5.2.2 Bayesian parameter estimation problem . . . 134
5.2.3 Parameter estimation . . . 136
5.2.3.1 Markov Chain Monte Carlo sampling . . . 137
5.2.3.2 Evaluation of sampling results . . . 139
5.2.4 Estimation of predictive accuracy . . . 142
5.2.5 Hypothesis testing . . . 145
5.3 Results . . . 145
5.3.1 Parameter inference . . . 146
5.3.1.1 Case studies for individual curves . . . 146
5.3.1.2 Case study 5.1: bistable regime, two curves . . . 151
5.3.1.3 Case study 5.2: bifurcating regime, two curves . . . 160
5.3.2 Comparison of fits . . . 167
5.4 Discussion and conclusions . . . 169

6 Discussion 176
6.1 Summary and discussion . . . 176
6.2 Future directions . . . 181
6.3 Final remarks . . . 185

Appendices 186
A Asymptotic approximation and conditions for physicality of the nonzero-tumour steady states of the asymptotically reduced model . 186
B Bifurcation structure of the asymptotically reduced model as parameter μ varies . . . 188
C Optimisation bounds . . . 189
D Profile likelihoods for case study 4.2b . . . 189
E Sensitivity-based identifiability . . . 189

Bibliography 192

List of Abbreviations

ABC	Approximate Bayesian Computation
ABM	agent-based model
ACT	adoptive cell therapy
APC	antigen-presenting cell
AZ	AstraZeneca
BCL	B cell lymphoma
CAR	chimeric antigen receptor
CI	confidence interval
CMA-ES	covariance matrix adaptation evolution strategy
CRI	credible interval
CTL	cytotoxic T cell
CTLA-4	cytotoxic T-lymphocyte-associated protein 4
CV	cross-validation
DC	dendritic cell
elim.	elimination
equi.	equilibrium
esc.	escape
FN	Fitzhugh-Nagumo
GM-CSF	granulocyte-macrophage colony-stimulating factor
HMC	Hamiltonian Monte Carlo
HPDI	highest posterior density interval
IFN-α	interferon-α
IFN-γ	interferon-γ
IL-2	interleukin-2
IO	input-output
IS	importance sampling
KDE	kernel density estimate
KLD	Kullback-Leibler divergence
LFO	leave-future-out
LOO	leave-one-out
MAP	maximum *a-posteriori*

MCMC	Markov Chain Monte Carlo
MDSC	myeloid-derived suppresor cell
MEM	mixed effects model(ling)
MH	Metropolis-Hastings
MHC	Major-Histocompatibility-Complex
MLE	maximum likelihood estimate, maximum likelihood estimation
NK	natural killer cell
NUTS	No-U-Turn Sampler
ODE	ordinary differential equations
PCA	principal component analysis
PD-1	programmed death protein 1
PD-L1	programmed death-ligand 1
PDE	partial differential equation
PDF	probability density function
PI	practical identifiability
PL	profile likelihood
PSIS	Pareto smoothed importance sampling
QSSA	quasi-steady-state-approximated, quasi-steady-state approximation
RHS	right-hand side
SDE	stochastic differential equation
SGI	structurally globally identifiable
SI	structural identifiability
SLI	structurally locally identifiable
SUI	structurally unidentifiable
SVD	singular value decomposition
TAA	tumour-associated antigen
TAM	tumour-associated macrophage
TCR	T cell receptor
TGF-β	transforming growth factor-β
Th	helper T cell
TME	tumour microenvironment
TNF-α	tumour necrosis factor-α
Treg	regulatory T cell
w.p.	with probability
WAIC	Watanabe-Akaike information criterion

Chapter 1

Introduction

1.1 Motivation

Cancer is a collection of diseases characterised by the development of abnormal cells that divide uncontrollably and spread to surrounding tissues [182]. It ranks as one of the leading causes of death worldwide, with millions of new cases diagnosed every year, and incidence and mortality rates rapidly growing [20]. Despite toxicity and limited specificity for cancer cells, chemotherapy and radiotherapy, together with surgery, remain the standard of care for cancer patients. The ability of cancer cells to break away from the primary site and metastasise makes cancer particularly challenging to control by these treatments alone, increasing the demand for alternative therapies.

Cancer development and progression is influenced by the immune system. Characterising this dependence has become an increasingly important research area. Numerous studies have demonstrated the dual role played by the immune system in cancer: it can protect organisms by recognising and eliminating malignant cells; it can promote tumour progression by exerting selective pressure on less *immunogenic* tumour cells, i.e. cells that evoke a weaker immune response [151]. These cells can be maintained in a state of immune-mediated equilibrium, but may, after further division and editing, eventually suppress anti-tumour activity and develop into clinically detectable tumours [165]. Dunn et al. [47] termed this process

immunoediting, proposing three possible outcomes of tumour-immune interactions, called '*the three Es of immunoediting*': tumour elimination, equilibrium, and escape.

Immunotherapies aim to harness the immune system to eradicate cancer cells, and target different mechanisms of tumour escape. Some immunotherapies have been approved for clinical use in the UK [51], although even the most successful of these have seen average response rates across tumour types of less than 50% [61, 100]. This suggests that a better understanding of tumour-immune interactions is needed to increase the pool of responding patients.

Our collaborators at AstraZeneca (AZ) have provided us with a dataset from a preclinical study, in which a type of immunotherapy was tested on mice (see Fig. 1.1). The data show tumour volume measurements divided into two groups: "non-responders" whose tumours continue to grow, and "responders" whose tumours shrink or are eradicated. Seeking to increase drug efficacy, AZ want to understand why this happens. Mathematical modelling has the potential to address this question; it can be used to interpret experimental observations and provide mechanistic insights into the complex nonlinear relations between a tumour and the immune system.

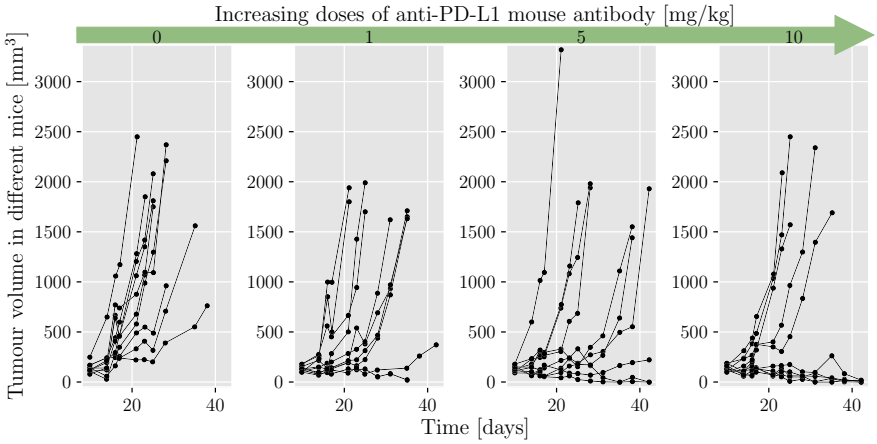


Figure 1.1: Plots show tumour volume data against time from a preclinical study with BALB/c mice. Every mouse is implanted with a tumour on day 0. Starting with day 6, they are administered twice weekly with different doses of an anti-PD-L1 mouse antibody for two weeks. Sizes of tumours are measured at different points in time, and the tumour evolution for each mouse is shown with a curve. Data provided by AstraZeneca.

A natural approach to modelling this dataset would be to split the mouse population into responders and non-responders, and fit different (phenomenological or data-driven) models to each group. In this thesis, we take a different approach, using a deterministic, time-dependent model that: (i) is minimally parametrised, in order to avoid overfitting and/or fixing parameters using other sources, (ii) is mechanistic, and (iii) exhibits the three E's of immunoediting (tumour elimination, equilibrium and escape) through bifurcations and/or bistability. A model with bifurcations can give rise to large qualitative changes in behaviour when parameters are varied across a threshold. If a model is bistable, then there exist parameter sets for which model solutions tend to different stable attractors (or a limit cycle) as initial conditions are varied. If the steady states correspond to different stages of immunoediting, then such changes in behaviour as model parameters or initial conditions are varied would represent transitions between tumour escape, control and eradication that are seen in the data.

The central question that arises from this modelling approach is whether models with the described characteristics can explain the dichotomous tumour growth data, and provide mechanistic insight into poor treatment responses. This is a challenging question, with many layers of complexity that stem from the rich model dynamics and the limited data available for model validation. The research in this thesis aims to increase understanding of these challenges. We address topics related to characterisation of the dynamics of bifurcating and/or bistable mathematical models of tumour-immune interactions and the feasibility of their calibration to dichotomous datasets. Our findings may stimulate the wider use of modelling in the development of new cancer immunotherapies, by showcasing, in particular, the ability of mechanistic models to generate new hypotheses about tumour-immune interactions, to explain existing datasets and to guide experimental design.

In the rest of this chapter we review the relevant literature. In Section 1.2 we provide an overview of the biology of cancer and the immune system. In Section 1.3

we review existing mathematical models of tumour-interactions, show how they have been used to inform immuno-oncology research, and introduce key concepts related to model calibration. The chapter ends in Section 1.4 with an outline of the research chapters that follow.

1.2 Biological background

1.2.1 Immune system and its activation

The immune system comprises organs, cells, and proteins that interact to defend the body against disease and infection. For smooth functioning, it must detect and neutralise pathogens and transformed cells, and also distinguish them from its own healthy cells (self).

Immune protection comprises several layers of defence. The first layer involves physical barriers, such as the skin and other epithelial surfaces that stop pathogens from entering the organism. The innate immune system is the second layer; it can respond quickly, via non-specific mechanisms, to pathogens, such as common microorganisms and bacteria, that breach the physical barriers. For example, macrophages are versatile phagocytotic cells involved in innate immunity. They reside in tissues, typically below epithelial cells, and watch for invaders by continually ingesting extra-cellular material, such as dying cells, debris or pathogens marked for destruction by other mechanisms. They can also release pro-inflammatory signals to attract other circulating immune cells, such as phagocytotic neutrophils and *natural killer cells* (NKs), to the site of infection. NKs are innate immune cells which target stressed or transformed cells that, unlike healthy cells, do not express *Major-Histocompatibility-Complex* (MHC) I molecules on their membranes. [157]

Adaptive immunity. Sometimes pathogens persist or are not recognised by the innate layer. In such cases, vertebrates can employ a third layer of defence called the adaptive immune system; it elicits a specific immune response against a

particular pathogen and is initiated, directed and assisted by innate immune cells. Its activation occurs mainly via *dendritic cells* (DCs). Like macrophages, DCs reside in infected tissues, where they engulf and process extra-cellular material. They can be activated by cytokines and chemokines expressed by other immune cells and dead pathogens, or by signals received via pattern recognition cellular receptors on DC membranes. Upon activation, DCs present pieces of the engulfed material on MHC I and II molecules, which are attached to their cell membranes. These presented molecules or *antigens* can induce a specific immune response; thus, DCs are classified as *antigen-presenting cells* (APCs). What distinguishes DCs from less effective APCs, such as macrophages, is their ability to travel to lymph nodes, where they can activate cells of the adaptive layer [115, 157].

The adaptive immune response is carried out by T cells and B cells. B cells are responsible for antibody-mediated immunity. Once activated, they produce proteins called antibodies. Since antibodies are cognate to antigens expressed by the detected pathogen via the MHC I molecules on their cell membranes, they can bind to them. This can neutralise the pathogen function or signal other cells of the innate immune system to destroy it. T cells are responsible for cell-mediated immunity, which plays a pivotal role in cancer. Therefore, henceforth, we will focus on the function of T cells (see [115, 157] for a more detailed review of B cell functions).

Every T cell expresses specific *T cell receptors* (TCRs) on its surface. These TCRs can bind to specific cognate antigens, on APCs for example. Variable specificities of TCRs across the T cell population enable T cells to collectively recognise and bind to millions of antigens. We distinguish two main subpopulations of T cells, *helper* and *cytotoxic T cells* (Ths and CTLs respectively). Ths help activate other cells of the immune system, while CTLs kill targeted non-self cells.

Naïve Ths are activated in the lymph nodes, when they recognise, via their TCRs, the cognate antigen presented on MHC II molecules of APCs. For successful

activation, co-stimulation is required upon binding through the TCR and the antigen-MHC complex. This is achieved via simultaneous binding of B7 proteins and CD28 molecules, expressed on the surface of activated APCs and Ths, respectively. Depending on the signalling in the local environment through cytokines and co-stimulatory receptors, naïve Ths differentiate into different subtypes (e.g. Th1, Th2, Th17, Tfh and others). Each subtype secretes different cytokines and growth factors, and enhances the recruitment and function of different immune cell types. Different Ths also promote their own proliferation and that of CTLs.

Similarly, CTL activation requires antigen presentation on DCs via MHC I. Details of CTL activation remain unclear, and may involve Ths. It is thought that efficient activation of long-lived CTLs requires Ths [157]. Once activated, CTLs differentiate into effector CTLs, which can kill target cells. They increase in number via rapid proliferation, and produce the cytokine *interleukin-2* (IL-2), which promotes differentiation of naïve CTLs into effector CTLs. They then leave the lymph nodes, enter the bloodstream, and travel to the site of infection. There they infiltrate the infected tissue, and scan cells until they find their target cell whose antigens, bound to MHC I molecules on the cell surface, match their specific TCR. Having found their match, effector CTLs kill the target cell via programmed death or apoptosis. After the immune attack dies down, some differentiated effector CTLs survive in the lymph nodes and at the tumour site. These become *memory T cells* that can rapidly respond if the invader is encountered again.

1.2.2 Cancer, cancer-immunity cycle and immunoediting

Cancer cells evolve from normal cells when they undergo gradual genetic and epigenetic (phenotypic) changes, which alter cell properties in ways that enhance their survival in tissues and spread [176]. According to Hanahan and Weinberg [65], the key characteristics of cancer cells may be summarised as ten hallmarks of cancer. Resisting cell death, sustaining proliferative signalling, evading growth suppressors,

activating invasion and metastasis, inducing angiogenesis, enabling replicative immortality, deregulating cellular energetics and avoiding immune destruction are identified as the core hallmarks or *"acquired functional capabilities that allow cancer to survive, proliferate, and disseminate"* [65]. Their acquisition is facilitated by enabling hallmarks; these are genomic instability and mutations, and tumour-promoting inflammation. In addition to tumour cells, many other cell types reside in malignant tissues. The hallmarks of cancer highlight that the active role played by the *tumour microenvironment* (TME) in tumour initiation and progression cannot be understood from the intrinsic properties of cancer cells alone. Below we outline the role of the immune system in cancer, as this is the focus in this thesis.

Tumour cells are coated with *tumour-specific antigens* (neoantigens), expressed exclusively on tumour cells as a result of mutated self-proteins, and *tumour-associated antigens* (TAAs), which may be expressed in larger quantities than on normal cells [63]. The presence of antigens on cancer cells induces anti-tumour responses via the immune mechanisms outlined in Section 1.2.1. However, the strength of these responses depends on antigen immunogenicity. For example,

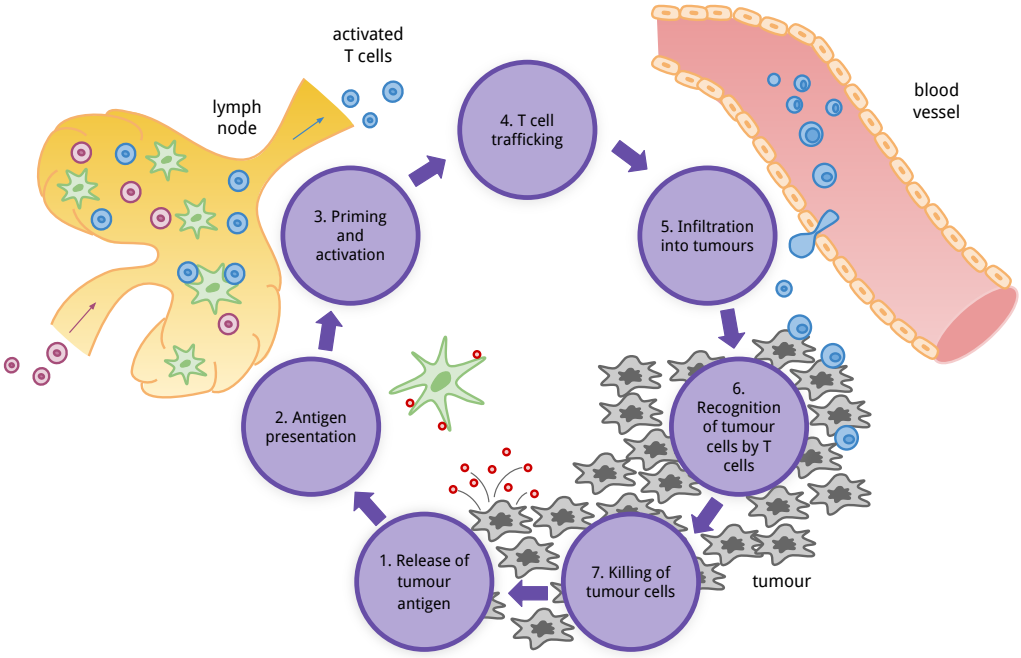


Figure 1.2: The cancer-immunity cycle. Figure based on Chen and Mellman [26].

neoantigens are thought to induce stronger immune responses than TAAs [162].

Chen and Mellman [26] proposed that maintenance of immunity to cancer can be encapsulated via a sequence of steps, termed the *cancer-immunity cycle*, which proceeds as follows (see Fig. 1.2). First, antigens and proinflammatory cytokines are released by dying tumour cells (1). DCs capture these antigens and present them to T cells, usually in the lymph nodes (2). These cells then become primed and activated (3) to mount an immune response against the tumour. Effector T cells subsequently traffic to (4) and infiltrate the tumour bed (5), until they find and bind to the target tumour cell with an antigen cognate to their TCR (6). Finally, the T cell kills the target tumour cell (7), and the cycle repeats.

In cancer patients, the cancer-immunity cycle is often disrupted, enabling tumour cells to survive or evade immune attack, and the disease to progress. The dynamic processes of tumour eradication by the immune system and tumour progression have been unified by Dunn et al. [47] under the concept of *immunoediting*, summarising the possible outcomes of an immune response to cancer as the "three E's of immunoediting":

(1) *Elimination* (immune surveillance). Transformed cells are destroyed by a competent immune system (optimal cancer-immunity cycle performance).

(2) *Equilibrium*. Some cancer cells escape immune attack due to abnormalities in cancer immunity. The surviving tumour cells are maintained in a state of immune-mediated dormancy, with some cells possibly undergoing further division and editing. During this period, new tumour cell variants with characteristics, genetic or epigenetic, that increase their resistance to immune destruction may be selected by the immune system.

(3) *Escape*. Immunologically sculpted tumours begin to grow progressively, become clinically detectable and establish an immunosuppressive TME.

Immunoediting provides a framework within which we can interpret different out-

comes of tumour-immune interactions. Of interest in this thesis, are mathematical models of tumour-immune interactions that capture these outcomes (we review such models in Section 1.3).

Thus far we have focused on the mechanisms by which the immune system is stimulated. It has been suggested that a balance between immunosupression and immunostimulation is needed to maintain tumour equilibrium [129], and that once the balance is shifted, escape or elimination will occur. In Section 1.2.3 we outline mechanisms by which the immune system may fail to be stimulated sufficiently or may be suppressed by tumour cells. This can shift the balance in favour of immunosuppression and enable tumour progression.

1.2.3 Mechanisms of immune evasion in cancer

Abnormalities in cancer immunity can occur at any stage in the cycle and may result in the tumour not being cleared by the immune system. For example, the tumour may not be recognised as foreign by T cells. This may occur if tumour cells lack mutations that produce immunogenic tumour antigens, or have defects in presenting tumour antigens on the cell surface, e.g. loss of expression of MHCs, or abnormalities in antigen processing [8]. Tumours can also evade immune attack by creating an immunosuppressive microenvironment [115]. This may occur through T cell exclusion from the tumour site (immunologically "cold" tumours), activation of co-inhibitory immune checkpoint pathways (immunologically "hot" tumours), or by recruiting cells with an immunosupressive phenotype [184]. In this section, we outline the role of the latter two mechanisms in cancer.

We also note that signalling via mutated cancer genes (i.e. intrinsic/oncogenic signalling) is thought to drive immunosuppression by modulating expression of certain cytokines, chemokines or membrane molecules which can, in turn, disrupt the recruitment and function of immune cells [15, 184]. However, how specific oncogenic

pathways interfere with immune signalling is not yet completely understood and remains an active area of research.

T cell exhaustion and negative co-stimulation. When tumour antigens are recognised and T cells are mobilised to destroy the tumour, their activity may become impaired when T cell stimulation is continual. This is termed *T cell exhaustion*, and is more likely to develop with prolonged antigen exposure or high antigen levels [61]. The key characteristics of exhausted T cells are overexpression of several inhibitory receptors (i.e. immune checkpoints), loss of effector functions (e.g. production of cytokines IFN-γ, IL-2 and TNF-α), and loss of proliferative capacity [15]. We will highlight the role of two checkpoints, *programmed death protein 1* (PD-1) and *cytotoxic T-lymphocyte-associated protein 4* (CTLA-4), which have been successfully exploited for development of immune checkpoint blockade therapies that inhibit signalling through these pathways.

Immunosuppression may occur during T cell activation in the lymph nodes. When a T cell is activated, expression of the CTLA-4 inhibitory receptor on its membrane is up-regulated. CTLA-4 binds with higher affinity to B7 proteins on activated APCs than the co-stimulatory CD28 receptor. Therefore, on repeated contact of the T cell and the APC, CTLA-4 outcompetes the binding of CD28. A co-inhibitory signal is transmitted, which dampens T cell activation [115, 122].

PD-1 is an inhibitory receptor which can be transiently expressed by activated T cells in response to sustained stimulation of their TCRs, cytokines and other signals [67]. Proinflammatory molecules released into the TME during immune attack (such as IFN-γ) drive expression of *programmed death-ligand 1* (PD-L1) on a diverse range of cells, including cancer cells which often overexpress it. When an effector T cell attaches with its TCR to a cognate antigen-MHC complex on a tumour cell, ligation of the T cell's PD-1 receptor with PD-L1 on the tumour cell also takes place. This, in turn, transmits a co-inhibitory signal, which triggers a

cascade of signalling events that inhibit the effector T cell cytotoxic activity by, for example, inducing apoptosis of effector T cells, facilitating their anergy and exhaustion, and inhibiting their proliferation. The PD-1/PD-L1 pathway is mainly responsible for dampening effector T cell activity in peripheral tissue, acting to minimise tissue damage in a healthy organism [68, 115].

Immunosuppressive immune cells. Different types of immune cells can contribute to an immunosuppressive TME. Their recruitment, phenotype and trafficking may be modulated by the tumour and other cells in the TME, and thus enable tumour progression and immune evasion.

Tumour cells may secrete a range of cytokines, including *transforming growth factor* β (TGF-β). TGF-β inhibits T cell activation, proliferation, differentiation and tumour cytotoxicity. It also induces differentiation of naïve Ths into *T regulatory cells* (Tregs) [115, 157]. Many tumours are infiltrated by Tregs, which produce immunoinhibitory cytokines such as IL-10, IL-35 and TGF-β [154].

Myeloid-derived suppressor cells (MDSCs) are a heterogeneous population of immature innate immune cells, such as DCs and macrophages, activated without maturing under pathological conditions. They can inhibit immune responses via numerous mechanisms [185]. Their expansion is promoted by tumour expression of *granulocyte-macrophage colony-stimulating factor* (GM-CSF) [15, 154]. Certain chemokines and chemokine receptors expressed by the tumour may also increase localisation of MDSCs and Tregs within the tumour [154].

Tumour-associated macrophages (TAMs) can also infiltrate the tumour bed in large numbers, and are correlated with poor prognosis [154]. Depending on the cytokine composition of the TME, their phenotype can switch between pro-inflammatory (M1-like) and anti-inflammatory (M2-like), with the latter typically being the dominant type in TAMs [124].

1.2.4 Immunotherapies

Immunotherapies aim to activate an immune response to the tumour by intervening at appropriate steps in the cancer-immunity cycle. We summarise several existing immunotherapies in Table 1.1. Many of these therapies, work by stimulating effector T cell functions. Immune checkpoint therapies have revolutionised the field of immunotherapy as they instead interrupt co-inhibitory signalling pathways, such as CTLA-4 and PD-1/PD-L1, which play an immunosuppressive role in the cancer-immunity cycle. Arguably the main issue is that, despite some success, only a fraction of patients respond to existing treatments [52, 100]. The significant challenges are to identify those tumours that will respond to a particular immunotherapy, to predict tumour responses, and to increase response rates among the large proportion of immunotherapy non-responders. It is thought the latter will require not only determination of optimal treatment scheduling and dosage,

Therapy	**Mechanism**	**Challenges**	**Ref.**
immune checkpoint blockade	counteract T cell exhaustion by using antibodies to block co-inhibitory checkpoint receptors (e.g. CTLA-4, PD-1/PD-L1) and their signalling	only a fraction of patients respond and show durable benefits	[61, 100, 121]
cancer vaccines	stimulate TAA-specific T cells, e.g. by increasing DC exposure to TAAs *ex vivo* (Sipuleucel-T)	identification of most suitable antigens	[64]
adoptive T cell transfer (ACT)	stimulate T cells by inducing their tumour specificity *ex vivo* via antigen-specific expansion or genetic engineering (e.g. CAR-T cells)	toxicity and cost	[6, 126]
oncolytic virus	selective replication of native or genetically modified viruses within tumour cells leads to their lysis and initiation of systemic anti-tumour immunity	clearance by the host before reaching tumour cells	[52]
cytokines	stimulate innate and adaptive immunity (e.g. IL-2, IFN-α)	toxicity at required large quantities, other negative effects (e.g. induction of regulatory checkpoints by IL-2)	[173]

Table 1.1: Types of available immunotherapies, their mechanisms of action and related challenges.

but also the design and identification of potent combination treatments in order to target multiple steps of cancer immunity [51, 80], e.g. stimulation of effector functions and also blockade of immunosuppression. As our understanding of cancer and immunosuppressive mechanisms increases, new immunotherapy drugs are being developed, with potential to expand the current repertoire of available therapies.

1.3 Mathematical models of tumour-immune cell interactions

Mathematical and computational modelling has been increasingly recognised as a tool to study complex systems, such as cancer. *In silico* models can be used to test and generate hypotheses, and can complement experimental work by interpreting associated data and guiding the design of new experiments. As such, modelling can provide mechanistic insights into different aspects of cancer and contribute to treatment improvement. Existing models typically focus on one or two hallmarks of cancer [9, 11]. Here we concentrate on models that account for tumour-immune cell interactions, and which could assist with the development of new immunotherapies.

Existing modelling frameworks vary from spatial to non-spatial models, deterministic to stochastic, continuous to discrete, single scale to multiscale [50, 105]. In this section, we review existing models describing tumour-immune interactions. In Section 1.3.1 we focus on models formulated using coupled ordinary differential equations (ODEs) as they are widely used, and are most relevant to the research presented in this thesis. We briefly review other approaches in Section 1.3.2.

1.3.1 Ordinary differential equations models

Well-mixed, time-dependent ODE models are more amenable to analysis and easier to validate than other frameworks. They can also generate mechanistic insight. They

neglect spatial effects, and describe population level dynamics deterministically, representing each species with a continuous, time-dependent variable.

1.3.1.1 Models including tumour and immune cells

The earliest models of tumour-immune interactions consist of two ordinary differential equations, describing how the density of tumour cells, $m(\tau)$, and immune cells, $e(\tau)$, changes over time, τ [50, 178]. The immune cell population typically represents CTLs or NKs, both of which can recognise and destroy cancer cells; we will also refer to them as "effector" cells. Tumour-immune interactions were initially modelled by analogy with predator-prey systems, with immune cells preying on cancer cells. This work inspired a number of extensions of the classical Lotka-Volterra models.

Many early models [35, 36, 55, 88, 158, 163] belong to the general family of two-equation models of tumour-immune interactions that was proposed by d'Onofrio [39]. They differ in the functional forms used to describe the tumour growth rate, immune predation of tumour cells, immune cell recruitment, proliferation and death rates. Numerical simulation and bifurcation analysis have been used to study the types of tumour-immune dynamics that different models exhibit.

Arguably, the most influential predator-prey-like model was proposed by Kuznetsov et al. [88]. They account for interactions between effector cells and tumour cells through the formation of an intermediate effector-tumour complex. The full model is reduced by applying the *quasi-steady-state approximation* (QSSA), a common technique for reducing models of chemical kinetics. Under this approximation, the number of complexes is assumed to be approximately constant, which they justify on the basis that formation and dissociation of complexes occurs on a shorter timescale ($\sim$ minutes) than cell proliferation and/or influx of effector cells ($\sim$ days). Under the QSSA, the following reduced model is obtained

$$\dot{e} = \sigma + \frac{\rho e m}{\eta + m} - \delta e - \mu e m, \tag{1.1a}$$

$$\dot{m} = \alpha m(1 - \beta m) - \kappa_2 e m, \tag{1.1b}$$

where the dimensionless variables e and m represent the numbers of free effector and tumour cells, respectively. Model parameters are estimated from experimental data for B cell lymphoma (BCL) in mice.

Kuznetsov et al.'s simplified model has been modified in many ways. Gałach [55] replaced the Michaelis-Menten effector influx with a bilinear function, thereby removing saturation in the immune response at large tumour size. Through convergence to steady states of different sizes (tumour-free, small-tumour, large-tumour) in different parameter regimes, Kuznetsov et al.'s model predicts tumour escape, elimination by effector cells, or tumour dormancy via damped oscillations. While Gałach's model also captures the three E's of immunoediting, it does not exhibit bistability between equilibrium and escape steady states, or the "sneaking through" phenomenon, where a small tumour lies outside the basin of attraction of the dormant steady state and escapes immune control; both phenomena are exhibited by the QSSA model (1.1).

Rihan et al. [137] introduced a time delay to the second and fourth terms of the right-hand side (RHS) of (1.1a), in order to account for the time between the recognition of non-self antigens by effector cells, the increase in effector numbers and their infiltration into the tumour. They found that a time delay less than a threshold value does not impact the stability of steady state solutions and therefore the model's immunoediting behaviour. By contrast, for a larger time delay nonzero steady states are destabilised via a Hopf bifurcation, and a stable limit cycle emerges. Khajanchi and Banerjee [78] added a time delay to the second term of the RHS of (1.1a) and showed that, as certain model parameters vary, the system can undergo period-doubling bifurcations which, in some regimes, lead to large-period oscillations or chaotic behaviour.

Roesch et al. [143] adapted Kuznetsov et al.'s model to investigate the impact of tumour-immune interactions on tumour responses to chemotherapy. They assumed

exponential tumour growth, and modified the tumour-immune interaction terms so that effector cells only interacted with tumour cells on its spherical surface (assumed), making $e \propto m^{0.75}$; the exponent of m is taken to be slightly larger than $\frac{2}{3}$ to account for the growth of B-cell-lymphomas in multiple nodes. In the treatment-free model the tumour either grows unboundedly or is eliminated at a tumour-free steady state, but the model cannot capture late tumour relapses (longer than two years). By adding linear loss terms in tumour and effector cell numbers to model chemotherapy effects, Roesch et al. used the model to investigate patient outcomes given chemotherapy (see Section 1.3.1.4 for further details).

1.3.1.2 Models including multiple cell types and signalling molecules

Other authors have developed mathematical models to describe tumour-immune interactions, including additional equations to represent other cell populations/ cytokines/drugs in the TME. Models of this type have been used to investigate different therapeutic targets, treatments and dosing regimens. Here we review these models, focusing on the insight they provide into cancer therapeutics.

In order to investigate ACT and IL-2 treatments, Kirschner and Panetta [79] included an ODE for the concentration of IL-2 cytokine, which stimulates effector cell proliferation. ACT is modelled as a constant influx of effector cells, and treatment with IL-2 is modelled as a constant source of IL-2. A bifurcation analysis of their model suggested that IL-2 therapy alone cannot eliminate the tumour, whereas ACT can, and the effect is enhanced when combining ACT with IL-2.

The role of immunosuppression in cancer progression has motivated the development of other models. Dritschel et al. [46] developed an ODE model to study the role of Ths and CTLs in anti-tumour immune responses. They added a Th compartment to Kuznetsov et al.'s model, and modelled immunosuppresion implicitly through a biphasic effector proliferation term. Their model exhibited immunoediting behaviour, and stable oscillations around the dormant steady state.

Model simulations suggested that Ths may be a more effective immunotherapeutic target than CTLs alone.

Robertson-Tessi et al. [138] developed a twelve-compartment model, which describes a multitude of cell types and cytokines in the TME, and includes multiple mechanisms of immunosuppression. By systematic numerical simulations, they showed that their model captures immunoediting behaviour, highlighted the importance of suppressive mechanisms at different stages of tumour development, and found Tregs to be significant at all stages. In further work, Robertson-Tessi et al. [139] showed that tumour responses to chemotherapy and immunotherapy depend on the balance between immunosuppressive and immunostimulatory effects.

Park et al. [123] recently proposed a model in which the immunosupressive dynamics of Robertson-Tessi et al.'s model are simplified, but Treg recruitment and function are retained. They investigated potential synergistic effects of chemotherapy and immunotherapeutic interventions. Chemotherapy depletes tumour and immune cells and can temporarily restore an immune response. The model by Park et al. [123] suggested that there exists a "Goldilocks window" of chemotherapy doses, in which maximal reduction in tumour size due to the drug and due to increased T-cell activation is achieved. Combining chemotherapy with DC vaccination was also found to increase the benefit further.

The recent success of immune checkpoint inhibitors has stimulated the development of several models that describe tumour responses to such therapies, alone and in combination [25, 82, 118, 130, 153]. The challenge in modelling checkpoint therapies is relating cellular dynamics with signalling through co-inhibitory receptors. Nikolopoulou et al. [118] developed a simple model for anti-PD-1 treatment, in which expression of PD-1 and PD-L1 molecules on tumour and immune cells is assumed to vary parametrically with tumour and immune cell numbers and anti-PD-1 drug concentration. Model simulations suggested that anti-PD-1 therapy alone cannot eradicate a tumour.

1.3.1.3 Parameter estimation

In many of the works mentioned above, parameter values are usually obtained from the literature, using data from different sources generated under different experimental conditions. In order to generate meaningful predictions, such models must be carefully calibrated and validated with datasets relevant to the context in which they will be used [177, 186]. In particular, model parameters need to be robustly estimated using such datasets [136].

Parameter estimation is an inverse problem that can be solved by minimising an objective function. As explained below, the inverse problem may be posed deterministically or statistically. First we introduce some notation following Stuart [164] and Lim [99].

Inverse problem notation. Let $\boldsymbol{\theta} \in \mathbb{R}^J$ be the vector of model parameters and $\mathbf{u} \in \mathbb{R}^{N_x K}$ the model solution (i.e. state variables over time) given as a map

$$\mathbf{u} = \boldsymbol{\mathcal{F}}(\boldsymbol{\theta}). \tag{1.2}$$

The task of computing $\boldsymbol{\mathcal{F}}(\boldsymbol{\theta})$ is called the *forward problem*, and its solution $\mathbf{u}$ is called the *forward model*. Let the forward model be $\mathbf{u} = [\mathbf{x}(\tau_1, \boldsymbol{\theta})^T, \ldots, \mathbf{x}(\tau_K, \boldsymbol{\theta})^T]^T$, where $\mathbf{x}(\tau, \boldsymbol{\theta})$ is the solution of a dynamical system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(\tau, \boldsymbol{\theta}), \boldsymbol{\theta}), \tag{1.3a}$$

$$\mathbf{x}(0, \boldsymbol{\theta}) = \mathbf{x}^0(\boldsymbol{\theta}), \tag{1.3b}$$

with time variable $\tau \in \mathbb{R}^+ \cup \{0\}$, time-dependent system state vector $\mathbf{x}(\tau, \boldsymbol{\theta}) \in \mathbb{R}^{N_x}$, and initial conditions $\mathbf{x}^0(\boldsymbol{\theta}) \in \mathbb{R}^{N_x}$. When the exact solution to the dynamical system is not accessible, a numerical integration scheme is used to obtain an approximate solution at discrete time points $(\tau_k)_{k=1}^K$, where K is the number of time points at which the numerical solution is computed. The system of equations and method of integration together determine the forward problem, and the forward model in this case is its numerical solution.

When $\boldsymbol{\theta}$ is unknown, or not directly observable, one must solve an inverse problem in order to recover it from observations of $\mathbf{u}$, which we denote as $\mathbf{s} = [\mathbf{y}_1, \ldots, \mathbf{y}_N]^T \in \mathbb{R}^{N_s}$. Here, for every $i \in \{1, \ldots, N\}$, $\mathbf{y}_i \in \mathbb{R}^{N_y}$ denotes observations taken at time τ_{k_i}, N_y is the number of observations made, N is the number of time points at which these observations are made and $N_s = N_y N$ is the total number of observations across all time points. We note that $\{\tau_{k_1}, \tau_{k_2}, \ldots, \tau_{k_N}\} \subseteq \{\tau_1, \tau_2, \ldots, \tau_K\}$. The definition of the inverse problem depends on our assumptions about how observations were generated. Most commonly, they are assumed to be of the form

$$\mathbf{s} = \boldsymbol{\mathcal{O}}(\mathbf{u}) + \boldsymbol{\nu} = \boldsymbol{\mathcal{O}}(\boldsymbol{\mathcal{F}}(\boldsymbol{\theta})) + \boldsymbol{\nu}, \tag{1.4}$$

where $\mathcal{O}(\mathbf{u})$ is the deterministic observation model, and $\boldsymbol{\nu} \in \mathbb{R}^{N_s}$ is additive random observation noise. However, uncertainty can be incorporated in other ways.

Deterministic framework. Parameter estimates are typically obtained by minimising an objective function that measures the agreement between observables, $\mathbf{s}$, and model predictions, $\mathcal{O}(\mathbf{u}) = \mathcal{O}(\mathcal{F}(\boldsymbol{\theta}))$, given a parameter set $\boldsymbol{\theta}$. A commonly used objective function is the weighted sum of squared residuals,

$$V(\boldsymbol{\theta}) = (\mathbf{s} - \boldsymbol{\mathcal{O}}(\mathbf{u}))^T \, \Xi^{-1} \, (\mathbf{s} - \boldsymbol{\mathcal{O}}(\mathbf{u})), \tag{1.5}$$

where Ξ is a matrix of symmetric, positive semidefinite weights, typically a diagonal matrix of known variances of independent observation errors [146].

Frequentist framework. Within a statistical framework, observations, $\mathbf{s}$, may be viewed as realisations of random variables. The joint *probability density function* (PDF) of observations or the likelihood, $p(\mathbf{s} \,|\, \boldsymbol{\theta})$, describes our statistical assumptions about the distribution of observations given specific values of $\boldsymbol{\theta}$, which are treated as constants, e.g. (1.4). From the likelihood, we can estimate the unknown parameters $\boldsymbol{\theta}$ via the *maximum likelihood estimator* (MLE)

$$\hat{\boldsymbol{\theta}}_{\text{MLE}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \, p(\mathbf{s} \,|\, \boldsymbol{\theta}). \tag{1.6}$$

Note that computing the maximum of the likelihood is equivalent to computing the maximum of the log-likelihood, $\mathcal{L}(\boldsymbol{\theta}) := \log p(\mathbf{s} \,|\, \boldsymbol{\theta})$. For Gaussian observations with mean $\boldsymbol{\mathcal{O}}(\boldsymbol{\mathcal{F}}(\boldsymbol{\theta}))$ and a known covariance matrix Ξ, it follows that

$$\mathcal{L}(\boldsymbol{\theta}) = -\frac{1}{2} V(\boldsymbol{\theta}) + \text{constant}. \tag{1.7}$$

In this case, $\hat{\boldsymbol{\theta}}_{\text{MLE}}$ is equivalent to the minimum of $V(\boldsymbol{\theta})$. We note that $\boldsymbol{\theta}$ is extended by the observation error covariance matrix, Ξ, if it is unknown.

Bayesian framework. Here unknown parameters, $\boldsymbol{\theta}$, are viewed as random variables. They are assigned a prior PDF, $p(\boldsymbol{\theta})$, that reflects beliefs/assumptions about their true values before observables are recorded. After observations are obtained, the prior is updated to a posterior PDF for $\boldsymbol{\theta}$, $p(\boldsymbol{\theta} \,|\, \mathbf{s})$, which is the joint distribution of $\boldsymbol{\theta}$ conditional on observations $\mathbf{s}$. From the posterior PDF, an estimate of $\boldsymbol{\theta}$ can be obtained, e.g. it may be estimated via the *maximum-a-posteriori* (MAP) *estimator* given by

$$\hat{\boldsymbol{\theta}}_{\text{MAP}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, p(\boldsymbol{\theta} \,|\, \mathbf{s}). \tag{1.8}$$

Sometimes the posterior mean or median computed using samples from $p(\boldsymbol{\theta} \,|\, \mathbf{s})$ are used. Using Bayes' rule, the posterior PDF may be written as

$$p(\boldsymbol{\theta} \,|\, \mathbf{s}) = \frac{p(\mathbf{s} \,|\, \boldsymbol{\theta})\, p(\boldsymbol{\theta})}{p(\mathbf{s})}, \tag{1.9}$$

whereby the posterior PDF is proportional to the prior PDF and the likelihood. Weak or non-informative priors do not have a large influence on the posterior, and the MAP estimate is, in such cases, close to the MLE [102]. From a mathematical perspective, specifying a prior can be viewed as equivalent to regularisation of a potentially poorly constrained MLE problem in order to make it well-posed and thus easier to solve [164]. The Bayesian approach thus provides a framework for carefully tracking the assumptions made, and makes it feasible to find a reasonable solution subject to its assumptions, when the system is poorly observed [99].

There are multiple sources of uncertainty in biological applications, including observation noise, discrepancies between the mathematical model and the biological system, and other sources of variability (e.g. inter-individual). Statistical frameworks allow for their systematic consideration. Assuming Gaussian observation errors, it is straightforward to define a MLE problem for estimating parameters of a deterministic ODE system from time-course data for a single individual. A more complex inverse problem can be formulated using the *mixed effects modelling* (MEM) framework. MEM is often used in pharmacometrics to calibrate models to data for a population of individuals, and account for the fact that different patients respond differently to the same drug [93]. A mathematical model (e.g. ODEs) is chosen to describe the dynamics of a biological process for each individual belonging to the population. Within this framework, we associate with each individual i a set of model parameters $\boldsymbol{\theta}^{(i)}$, which comprise "fixed" and "random" effects, i.e. $\theta_j^{(i)} = \theta_j^{(\text{pop})} + \lambda_j^{(i)}$. Here $\boldsymbol{\theta}^{(\text{pop})}$ are unknown parameters fixed for the entire population, while $\boldsymbol{\lambda}^{(i)}$ are the individual deviations from the population values. Typically, $\lambda_j^{(i)}$ are assumed distributed according to a zero-mean Gaussian distribution, with standard deviation ω_j, fixed over the population, i.e. $\lambda_j^{(i)} \sim \mathcal{N}(0, \omega_j^2)$. If $w_j = 0$, then the corresponding $\theta_j^{(i)}$ is a fixed population parameter; if ω_j are assumed unknown, then they describe inter-individual variability. The total number of parameters to estimate is lower than for separate fits of each individual's dataset. In MEMs, variability of data between individuals is explained by variability in the parameters with significant random effects.

The parameter estimation problems as described above are solved by minimising the objective function that emerges from the modelling and statistical assumptions. This may be challenging for multiple reasons [104]. The data are often noisy and/or scarce. For example, in the context of tumour-immune interactions, it is difficult to obtain time-resolved *in vivo* data describing properties of immune cells in the TME for a single subject, as is the case with AZ's data mentioned previously.

Parameter estimation may also be difficult due to uncertainty and/or complexity of the model and its behaviour, e.g. model discrepancy, over-/under-parametrisation with regards to data, nonlinearity, bifurcations [95, 142]. It also depends on any additional statistical constraints, such as the prior [164]. Each of these factors affects the shape of the objective function, and therefore whether its minimum is unique and straightforward to identify.

There are many methods for solving parameter estimation problems [136]. Optimisation methods seek the single set of parameter values for which the model best fits the data. The methods range from local to global, gradient-based to gradient-free, deterministic to stochastic. Inference-based approaches quantify uncertainty in parameter estimates, but may be computationally intensive, as the objective function must be explored in multi-dimensional parameter space. For problems formulated within a Bayesian framework, the posterior distribution of parameters, $p(\boldsymbol{\theta} \mid \mathbf{s})$, can be approximated with popular methods that are generally used for sampling multi-dimensional probability distributions, such as *Markov chain Monte Carlo* (MCMC) [5, 91] or *Approximate Bayesian Computation* (ABC) [167]; the latter is used when the PDF (up to a constant) cannot be evaluated easily.

By analysing identifiability of the inverse problem, potential difficulties with parameter estimation may be resolved before attempting to fit real data and/or before the data are collected [146]. Several methods can be used to determine *structural and practical identifiability* of a problem, i.e. whether its parameter values can be uniquely recovered, and with low uncertainty [28, 111, 134]. Structural identifiability considers perfect measurements (resulting from the deterministic, time-continuous observation model, $\mathcal{O}(\cdot)$) and is thus related only to the structure of the dynamical system and the types of observables, while practical identifiability also considers noise and the times at which measurements are made.

As Chapters 4 and 5 of this thesis involve identifiability analysis and parameter estimation, we review the relevant methods in more detail in these chapters.

1.3.1.4 Linking models to data

As more treatments and data become available, there has been a push for better integration of tumour-immune interactions models and data [127, 172]. This has led to the development of simple ODE models, which incorporate some mechanistic details and can be readily related to data [22, 34, 82, 85, 87, 88, 137, 143, 153, 179, 180]. Here we summarise studies that provide a more data-informed insight into the outcomes of tumour-immune interactions.

We first note that Kuznetsov et al. [88] parametrised their QSSA model using time-series data of tumour cell numbers in mice with BCL_1 lymphoma in the spleen. The logistic tumour growth parameters were estimated separately using data from non-chimeric mice which could not elicit an immune response. The constant effector supply rate was estimated without fitting to data, while other immune parameters were calibrated using data from chimeric mice with immune systems. These estimates have been used in many subsequent modelling studies (see, for example, [34, 46, 143, 150]).

In [88] only tumour measurements were used to estimate model parameters. By contrast, de Pillis et al. [34] used experimental data from human and murine cytotoxicity assays to estimate tumour cell kill parameters in their model, in which the effector population is decomposed into NKs and CTLs. Other model parameters were fixed at values found in the literature.

Roesch et al. [143] calibrated a modified version of Kuznetsov et al.'s model to progression-free survival data from a clinical trial of different chemotherapy dosing. They used the maximum entropy principle to fit five (of fourteen) model parameters. Some of these parameters were assumed to be therapy-, study- and/or patient-specific, the latter of which were assumed statistically distributed, with

means and variances fitted to the data. Paradoxically, clinical data show that increasing chemotherapy intensity can hinder its effectiveness, potentially due to its negative impact on the immune system that hence fails to control any remaining tumour cells. Roesch et al.'s model was able to explain these data by incorporating immune system effects and variability in patient-specific parameters.

Wilkie and Hahnfeldt [179] explored the dual role of the immune system in cancer progression by modifying a simple logistic model of tumour growth. They modelled immune system effects indirectly by imposing a dynamic carrying capacity, which depends on the balance between immunosuppressive and immunostimulatory signals in the TME (i.e. balance between tumour and immune cells). After fixing some parameters from the literature, and with the simplifying assumptions of immune homeostasis and negligible immune predation at detectable tumour sizes, they used MCMC to estimate tumour growth parameters and the homeostatic immune cell number from time-course data for subcutaneous fibrosarcoma grown in immunodeficient mice. Multiple parameter sets were found to fit the data equally well, suggesting that the model is non-identifiable. Nevertheless, they conducted a simulation study using the resulting parameter sets, and showed the model can predict a separation of tumour growth curves into escape and elimination as a result of variability in parameter values.

Such heterogeneity in tumour growth is seen in AZ's data (see Fig. 1.1) from a preclinical trial with an immune checkpoint inhibitor. As multiple trials suggest that the efficacy of checkpoint inhibitors may be increased with combination treatments, several recent modelling studies have investigated their synergistic effects.

Serre et al. [153] investigated the combination of immune checkpoint blockade and radiotherapy. They developed a semi-mechanistic model that describes immune response via interactions among tumour cells, their antigens and effector cells. Their model qualitatively reproduces tumour growth curves that show poor responses to radiotherapy or checkpoint blockade monotherapies, and improved responses

to combinations of these treatments due to their synergy. While the simulated survival curves were in good agreement with results from preclinical studies, the model lacks rigorous calibration of parameters and quantitative validation.

Kosinsky et al. [82] studied data from a preclinical study in which radiotherapy was combined with an anti-PD-L1 drug. Their model describes the time evolution of healthy tumour cells, radiotherapy-damaged tumour cells, naïve T cells and CTLs. Eight (of twenty-five) model parameters were fitted using the nonlinear MEM framework. Their results suggested that heterogeneity in tumour responses to treatment can be best explained through variability in the T cells' ability to infiltrate the tumour upon antigen exposure. From simulations of the calibrated model they concluded that administering anti-PD-(L)1 prior to, or concurrently with radiotherapy, has greater benefit than post radiotherapy.

Modelling work by Radunskaya et al. [130] was motivated by data from a mouse study, in which treatment with an anti-PD-L1 drug and an inhibitor of *Bruton's tyrosine kinase* (BTK), which slows down tumour growth, was shown to be more effective than each drug alone. Their mathematical model (with eleven ODEs) was adapted from Köse et al. [81] and describes interactions among dendritic, effector and tumour cells in the spleen, the blood and the tumour. The model attributes the dichotomy in tumour volume data to variability in the tumour growth rate and the maximum kill rate of tumour cells by immune cells. These two parameters were fitted to data, while other parameters were fixed at values estimated in [81].

1.3.2 Other modelling frameworks

Other approaches have been used to model tumour-immune interactions and account for their complex and multiscale nature.

Partial differential equations (PDEs) can capture how tumour and immune cells interact in spatial domains and can account for cell movement [1, 54, 89, 90, 97, 98, 107]. For example, Matzavinos et al. [107] developed a spatial extension of

Kuznetsov et al.'s original model, by assuming cells move via diffusion, and that immune cells also move chemotacticaly down the spatial concentration gradients of cytokines/chemokines produced by complexes during immune cell activation.

Stochastic differential equation (SDE) models have been used to consider the impact of inherent randomness in tumour-immune interactions on the time-evolution of interacting cell populations, effects which may be particularly influential at small cell numbers. They may account for intrinsic stochasticity [4, 16, 23], associated with variability of birth/death processes, cell genotype, phenotype and/or function, or extrinsic stochasticity [40, 94, 183], which may arise from variability in environmental factors such as oxygen and nutrient supply, the state of the host's immune system, temperature, radiation, or both [17, 101].

In contrast to continuum models, where cells/molecules are treated as densities, *agent-based models* (ABMs) consider individual cells as discrete entities (i.e. agents). The behaviour of each agent (i.e. changes in state and on/off-lattice location) and interactions with neighbouring agents are modelled by deterministic or stochastic rules. While being computationally demanding, ABMs can capture heterogeneity and randomness of cell interactions, particularly at small cell numbers, and spatial effects. The TME's multiscale nature can be incorporated into ABMs by their coupling to ODEs or PDEs that describe the evolution of TME components, such as oxygen, chemokines, cytokines; such models are termed "hybrid" ABMs. We refer the reader to Norton et al. [119] for a recent review of (hybrid) ABMs.

Kinetic (cellular) theory for active particles has been used to develop models that describe large systems of interacting entities and account for the fact that biological function is statistically distributed across cells of the same population, and differs among different populations [11]. Each cell is characterised by a particular internal state. From a stochastic description of cell interactions, integro-differential equations are derived that describe the time evolution of distributions of internal

cell states over each interacting population of cells. We refer the reader to Bellomo and Delitala [10] for a review of such models.

1.3.3 Summary and discussion

ODEs are widely used to examine the spatially-averaged dynamics of tumour cells and immune cells. To date, such models have been primarily used for *in silico* testing and to generate hypotheses that may be tested with new experiments. The hypotheses relate to tumour-immune interactions, different immunotherapeutic targets, and scheduling and dosing of cancer therapies. Despite various attempts to relate (semi-)mechanistic ODE models to experimental data, models that have been robustly validated against data remain scarce [127]. Progress has been hindered by the lack of appropriate immunological data and inter-individual variability.

Other modelling frameworks are better suited to investigating how spatial effects, processes that act on different space/time scales, stochasticity and/or heterogeneity (e.g. within a cell population) impact tumour-immune dynamics and outcomes. However, the resulting models are typically more complex than ODEs, and more difficult to analyse and validate. More detailed datasets are also needed to validate these models. For example, to quantitatively validate ABMs, spatio-temporal data are needed [119]. As such data are difficult to obtain, ABM simulations are typically qualitatively compared to histological images at single time points.

There is a need to bridge the gap between modelling and data in order to tackle major challenges in cancer immunotherapy. Multiscale mechanistic models must be devised with the view of constraining them with available data [186]. Modellers must also work with experimentalists and clinicians to design experiments which collect data that enable model validation [174]. There are also opportunities for modellers to better exploit large datasets being generated from single-cell RNA sequencing, mass spectrometry and mass cytometry [3, 105]. Such datasets provide a lot of information about the tumour and its microenvironment, but usually only

at a single time point. Close collaboration and advances in both modelling and experimentation will therefore be needed to move modelling in (immuno-)oncology from hypothesis testing/generation to prediction in preclinical and clinical settings [2].

In this thesis we will address challenges related to ODE models, focusing on the analysis of model behaviour and the estimation of model parameters from data.

1.4 Thesis outline

The research presented in this thesis is motivated by AZ's experimental data (see Fig. 1.1). We focus on simple mechanistic ODE models that can capture the dichotomous response seen in the tumour growth curves. More specifically, we are interested in models with bifurcating and/or bistable behaviour that exhibit the three stages of immunoediting (tumour elimination, equilibrium and escape) via stable steady states of different tumour sizes in different regions of parameter space. To answer questions about immune system characteristics that may lead to tumour escape, such a model requires a minimum of two compartments, one describing tumour cell growth, and the other describing immune cell dynamics. In this thesis, we focus on one of the earliest and most influential tumour-immune interactions models with these properties that was proposed by Kuznetsov et al. [88]. The remainder of this thesis comprises four research chapters in which we characterise different versions of this model from a dynamical systems perspective and from a parameter identification perspective.

Chapter 2 provides new insights into the original, five-compartment model by Kuznetsov et al. [88]. This model is of interest as its QSSA version exhibits the three Es of immunoediting, and it has been widely studied in the literature. Exploiting a separation of timescales in the model, we perform a formal matched asymptotic analysis in order to derive a new, two-dimensional, long-timescale

approximation of the original model that differs from the QSSA model. We analyse phase-plane dynamics of the reduced model to show it is excitable, a feature not previously reported for similar models of tumour-immune interactions.

In **Chapter 3** we perform a bifurcation analysis via analytical and numerical techniques in order to fully characterise the impact of excitability on model behaviour and its implications for tumour-immune interactions. We focus on characterising the model's steady states and limit cycles and show that the model is descriptive of immunoediting. We perform a systematic parameter sensitivity analysis to demonstrate how tumour growth curves may change as patient-specific parameters and/or immunotherapy parameters vary and to show that the resulting dynamics are consistent with dichotomous datasets.

The reduced model accurately describes Kuznetsov et al.'s full model in the limit considered in Chapters 2 and 3, but the model's complex dynamics render it difficult to validate against experimental data. In **Chapter 4** we investigate the feasibility of estimating the model parameters of the QSSA model from observables of tumour numbers (recall that the QSSA model is simpler to analyse but still exhibits immunoediting via bifurcations and bistability, and therefore allows for investigation of model identifiability in the presence of bifurcations). We first note the model's structural identifiability, which is a property of the model that describes whether parameters may be identified from continuous and noise-free observables. We then explore the model's practical identifiability within a frequentist framework, using noisy and time-spaced synthetic data, chosen to exhibit different qualitative behaviours, which correspond to different stages of immunoediting.

Since model parameters are poorly identifiable for some datasets, we investigate in **Chapter 5** a parameter fitting approach where multiple tumour growth curves are fitted simultaneously and certain unknown parameters are assumed fixed over a population of curves. This approach exploits data from other individuals to inform estimation of their parameters. While it is similar to mixed effects modelling [93],

we perform a full Bayesian inversion and do not prescribe a statistical model for the distribution of individual parameters. We propose and test two hypotheses, generated by the model's bifurcation structure, that can provide different biophysical explanations for the dichotomy in tumour growth curves. Hypothesis 1 (H1) asserts that the initial conditions are identical and attributes the dichotomous behaviour to differences in the model parameters. Hypothesis 2 (H2) asserts that the model parameters for both individuals are identical and lie in a region of parameter space that exhibits bistability, and that the dichotomous behaviour is due to differences in the initial conditions. In order to establish whether the QSSA model can be used to robustly attribute dichotomy in the synthetic data to particular sources of variability, we investigate the quality and quantity of data needed to distinguish between these two hypotheses.

This thesis ends in **Chapter 6**, where we summarise the main conclusions of the research and outline possible directions for future work.

Chapter 2

An excitable model of tumour-immune interactions

The research presented in this chapter has been published in the *Journal of Theoretical Biology* [120].

2.1 Introduction

A multitude of mathematical models have been developed over the years in order to increase our understanding of how the dynamic interplay between cancer and the immune system affects cancer growth. However, few of these models have been successfully parametrised. Simple mechanistic models that reproduce experimental data, but are not overparameterised, are of particular interest, as they are easier to analyse and calibrate, and therefore more appropriate to use as predictive tools. They may be direct mathematical translations of biological knowledge (see Section 1.3.1), or they may be simplifications of more complex models [72, 88]; mathematical techniques, such as asymptotic analysis, may be used to derive reduced models using biologically-justified assumptions.

As described in Chapter 1, Kuznetsov et al. [88] derived a two-equation, quasi-steady-state-approximated (QSSA), model from a five-equation model. The original model describes interactions between effector and tumour cells mediated by the formation of intermediate effector-tumour complexes. The model incorporates

immune killing of tumour cells, tumour-induced suppression and anergy of effector cells, and limits on immune recruitment and proliferation when the tumour grows large. Kuznetsov et al. [88] performed model reduction by assuming that the timescale on which tumour and effector cells form complexes is much shorter than the timescale of other processes. Under this QSSA, the model reduces to two ODEs that describe the dynamics of the effector and tumour cells; the complexes vary parametrically with these variables. The reduced model has been calibrated to murine data [156]. Kuznetsov et al.'s analysis of the QSSA model's bifurcations reveals how variation in the influx rate and death rate of effector cells affects system dynamics; the model exhibits tumour escape, equilibrium or elimination, or coexistence (bistability) of small-tumour or tumour-free steady states with escape. The occurrence of sneaking through and immunostimulatory effects in this model was found to depend on the rate at which tumour cells inactivate effector cells.

A large body of work stems from the QSSA model of Kuznetsov et al. [88] and provides insight into its behaviours and those of its numerous modifications [33, 34, 46, 55, 137, 143]. However, few models of tumour-immune interactions have considered the role of effector-tumour conjugates. Matzavinos et al. [107] analysed Kuznetsov et al.'s original model for a different non-dimensionalisation and different parameter values. As they varied the rate of conjugate formation, they identified parameter regimes with bistability between a large-amplitude limit cycle and a small-tumour steady state. Such behaviour is not evident in the QSSA model with simplified conjugate dynamics. While this study and others [1, 42, 77] highlight the important role of effector-tumour conjugate kinetics in Kuznetsov et al.'s original model, the findings are based on numerical results, with little analytical support. A comprehensive mathematical investigation of Kuznetsov et al.'s original model is required to understand the validity of the QSSA of conjugate dynamics that led to the derivation of their reduced model.

In this chapter we extend the studies of Kuznetsov et al. [88] and Matzavinos

et al. [107] to increase understanding of how tumour-immune cell interactions impact tumour growth dynamics. We explore the validity of the QSSA in the context of Kuznetsov et al.'s original model using asymptotic analysis. Excitability in tumour-immune dynamics emerges naturally from a new asymptotic reduction of the original model. To our knowledge, excitability has not traditionally been identified as a feature of existing models of tumour-immune interactions. However, it provides a natural mathematical framework for mechanistic understanding of model-based predictions of tumour-immune dynamics. This chapter serves as a basis and motivation for Chapter 3, where we investigate the impact of excitability on the long-term dynamics and their sensitivity to changes in parameter values.

The remainder of this chapter is organised as follows. In Section 2.2 we describe the original model of Kuznetsov et al. [88]. In Section 2.3 we exploit the separation of timescales, which arises due to fast conjugate dynamics, to derive a new asymptotic reduction of the original model. The reduced model we obtain differs from the QSSA model derived in [88], but agrees with numerical simulations of the original model where the asymptotic approximation is valid. Using a combination of asymptotic and phase-plane techniques, we show in Section 2.4 that our reduced model is excitable and we investigate the origins of its excitability. By comparing the phase-plane behaviour of our reduced model to that of Kuznetsov et al.'s QSSA model, we show in Section 2.5 that the latter is not excitable in the parameter regime of interest. We discuss our findings and outline directions for future work in Section 2.6.

2.2 Mathematical model

In this section, we summarise the original model from Kuznetsov et al. [88]. We then introduce the assumptions we make about parameter values, and use them to nondimensionalise the model. This will facilitate its simplification and analysis of

its long-term behaviour in later sections of this chapter and in Chapter 3. We end by stating the parameter values that will be used for model analysis.

2.2.1 Kuznetsov et al.'s original five-compartment model

Kuznetsov et al. [88] proposed a system of five, time-dependent, nonlinear ordinary differential equations (ODEs) to describe temporal interactions between effector and tumour cells in a well-mixed (spatially-uniform) region, which represents the interface of the tumour and its microenvironment (see Fig. 2.1). Here the generic term effector (immune) cells refers to all cell types that can destroy tumour cells when activated, noting that these cells are anticipated to be predominantly CTLs. In their model, tumour and effector cells may exist in a functional state, where they interact with other cells, or in a dysfunctional state, where they are unable to interact with other cells. Accordingly, the tumour cell population is split into normal tumour cells, $M(t)$, and lethally hit (marked for death) tumour cells, $M^*(t)$, while the effector cell population is divided into activated effector cells, $E(t)$ and inactivated effector cells, $E^*(t)$. The model also includes an intermediate state,

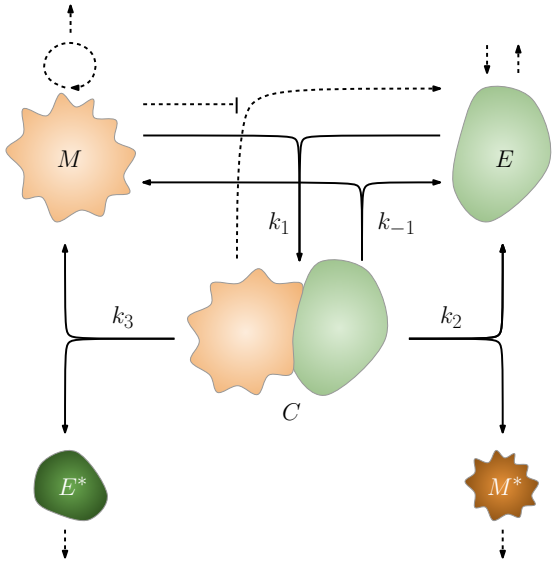


Figure 2.1: Schematic showing how effector (green) and tumour (orange) cells interact in the model of Kuznetsov et al. [88]. Kinetic interactions are denoted with solid arrows, while birth, death and migration processes are denoted with dashed ones. Arrows with a straight line in place of an arrow head represent inhibition of a process.

$C(t)$, associated with the formation of conjugates, when an activated effector cell and a normal tumour cell come in contact.

The evolution of the normal tumour cell population is governed by their proliferation, and programmed cell death due to interactions with effector cells. Proliferation is modelled via a logistic growth term $aM(1-bM)$, where the positive constant a (units: day^{-1}) is the basal growth rate, and the reciprocal of the positive parameter b (units: cells^{-1}) is the carrying capacity of the tumour site. Tumour cell programmed death, induced by effector cells, is modelled as a transition from the functional state, M, to the dysfunctional state, M^*, that is assumed to occur via effector-tumour conjugates, C; the inactivation of effector cells occurs in a similar manner. Conjugates are assumed to form at rate k_1EM, which is proportional to the functional effector and tumour cell numbers, with constant of proportionality k_1 (units: cell^{-1}day^{-1}). The resulting conjugate may dissociate into (see solid arrows in Fig. 2.1): (i) a tumour cell, M, and an activated effector cell, E, at rate k_{-1} (units: day^{-1}); (ii) a lethally hit tumour cell, M^*, and an activated effector cell, E, at rate k_2 (units: day^{-1}); (iii) an inactivated (suppressed) effector cell, E^*, and a tumour cell, M, at rate k_3C, proportional to conjugate numbers, with constant of proportionality k_3 (units: day^{-1}). Dysfunctional cells are assumed to arise only by transitioning, via conjugates, from their respective functional compartments, and their ultimate fate is death or emigration from the region. Programmed death of lethally hit tumour cells, M^*, is assumed to occur at rate d_3 (units: day^{-1}).

The evolution of the activated effector cells is assumed to be dominated by their migration from the lymph nodes to the tumour site, natural cell death, and inactivation by tumour cells, as described above. Migration to the tumour is modelled via a supply term with two parts: a continuous baseline supply of activated effector cells to the tumour at a constant rate s (units: cells day^{-1}), and a tumour-stimulated rate, $\frac{fC}{g+M}$, that accounts for an increase in the rate of effector recruitment, proliferation, and infiltration due to their interactions with tumour

cells within conjugates; f (units: cells day^{-1}) and g (units: cells) are positive constants, such that f/g is the maximum effector supply rate per complex, while g is the tumour cell concentration, at which the influx rate per complex is half-maximal. The factor $(g+M)^{-1}$ ensures that the tumour-stimulated rate of influx saturates as the number of tumour cells increases, and accounts for limitations in the rate of stimulation of the immune system, or in the rate of transport to and through the tumour. Activated effector cells are assumed to die at rate d_1 (units: day^{-1}). As for $M^*(t)$, dysfunctional effector cells are assumed to die/migrate out at rate d_2 (units: day^{-1}).

By combining the processes outlined above, we recover the following ODEs for the different cell species that were proposed by Kuznetsov et al. [88]:

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \underbrace{aM(1-bM)}_{\text{net proliferation}} - \underbrace{k_1EM}_{\substack{\text{conjugate}\\ \text{formation}}} + \underbrace{(k_{-1}+k_3)C}_{\substack{\text{conjugate}\\ \text{dissociation}\\ \text{w/o damage to}\\ \text{tumour cell}}}, \tag{2.1a}$$

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \underbrace{s}_{\substack{\text{constant}\\ \text{supply}}} + \underbrace{\frac{fC}{g+M}}_{\substack{\text{tumour-}\\ \text{stimulated}\\ \text{supply}}} - \underbrace{d_1E}_{\text{death}} - \underbrace{k_1EM}_{\substack{\text{conjugate}\\ \text{formation}}} + \underbrace{(k_{-1}+k_2)C}_{\substack{\text{conjugate}\\ \text{dissociation}\\ \text{w/o damage to}\\ \text{effector cell}}}, \tag{2.1b}$$

$$\frac{\mathrm{d}C}{\mathrm{d}t} = \underbrace{k_1EM}_{\substack{\text{conjugate}\\ \text{formation}}} - \underbrace{(k_{-1}+k_2+k_3)C}_{\text{conjugate dissociation}}, \tag{2.1c}$$

$$\frac{\mathrm{d}M^*}{\mathrm{d}t} = \underbrace{k_2C}_{\substack{\text{tumour cell}\\ \text{marking}\\ \text{for death}}} - \underbrace{d_3M^*}_{\substack{\text{programmed}\\ \text{death}}}, \tag{2.1d}$$

$$\frac{\mathrm{d}E^*}{\mathrm{d}t} = \underbrace{k_3C}_{\substack{\text{effector cell}\\ \text{inactivation}}} - \underbrace{d_2E^*}_{\substack{\text{death or}\\ \text{migration}}}. \tag{2.1e}$$

We close equations (2.1a–e) by prescribing the following initial conditions:

$$E(0)=E_0, \quad M(0)=M_0, \quad C(0)=0, \quad E^*(0)=0, \quad \text{and} \quad M^*(0)=0, \tag{2.1f}$$

where the positive constants E_0 and M_0 denote the initial concentrations of free tumour cells and effector cells, respectively. No conjugates, C, and dysfunctional

cells, M^* and E^*, are present initially. We note from equations (2.1) that the dynamics of M^* and E^* are slave to those of M, E and C. Since the system behaviour can be determined from equations (2.1a–c), we henceforth focus on these.

2.2.2 Parameter assumptions and the dimensionless model

We nondimensionalise equations (2.1a–c) and (2.1f) by rescaling model variables as follows:

$$e = \frac{E}{\hat{E}}, \quad m = \frac{M}{\hat{M}}, \quad c = \frac{C}{\hat{C}}, \quad \tau = k_2 t, \tag{2.2a}$$

where $\hat{E}$ and $\hat{M}$ are positive constants (see Table 2.1). We fix $\hat{C} = \frac{k_1 \hat{E} \hat{M}}{K}$ with $K = k_{-1} + k_2 + k_3$, in order to balance terms in equation (2.1c) for C. We introduce the following dimensionless parameter groupings

$$\sigma = \frac{s}{k_2 \hat{E}}, \quad \rho = \frac{f k_1}{k_2 K}, \quad \eta = \frac{g}{\hat{M}}, \quad \delta = \frac{d_1}{k_2}, \quad \mu = \frac{k_3 k_1 \hat{M}}{k_2 K}, \quad \kappa_1 = \frac{k_1 \hat{M}}{K}, \tag{2.2b}$$

$$\alpha = \frac{a}{k_2}, \quad \beta = b\hat{M}, \quad \kappa_2 = \frac{k_1 \hat{E}}{K} = \kappa_1 \frac{\hat{E}}{\hat{M}}. \tag{2.2c}$$

Additionally, we identify a small dimensionless parameter

$$0 < \varepsilon = \frac{k_2}{k_{-1} + k_2 + k_3} = \frac{k_2}{K} \ll 1. \tag{2.2d}$$

The choice of ε is based on the observation that while effector cells constantly scan and interact with tumour cells in order to locate their target cell, only a small proportion of these interactions result in the marking of tumour cells for death, or inactivation of effector cells. For this to happen, the tumour antigen must be cognate to the TCR, but this is not the case at every effector-tumour interaction due to heterogeneity in repertoires of tumour antigens [19, 26, 131] and TCRs at the tumour site [44]. The average kinetics of tumour cell programmed death and effector cell inactivation are therefore slow in comparison to conjugate formation and dissociation without damage. In terms of parameter values, we thus assume that the (per cell) rates k_{-1} and $k_1 \hat{E}$, are large in comparison to k_2 and

k_3, i.e. $k_1\hat{E} = \mathcal{O}(k_{-1})$, $k_2 \ll k_{-1}$, and $k_3 = \mathcal{O}(k_2) \ll k_{-1}$. Given our interest in the long-term behaviour, we scale the model with one of the longer timescales, choosing k_2^{-1} as in [88]. Following their reasoning for applying the QSSA, we assume that kinetics of processes such as cell death, proliferation and migration to the tumour microenvironment are much slower than kinetics of conjugate formation and dissociation without damage, and we assume they are approximately comparable to those of tumour cell programmed death and effector cell inactivation; we therefore have $\frac{s}{\hat{E}}$, $\frac{fk_1}{K}$, d_1, $a = \mathcal{O}(k_2) \ll k_{-1}$. In summary, there is a separation of timescales, with several events associated with the complexes occurring rapidly. We will see in Section 2.2.3 that our parameter assumptions are consistent with the parameter values from [88] (see Table 2.1). We note that these assumptions represent a distinguished limit, and estimates of parameters k_1 and k_{-1} are needed to justify its physical relevance.

Under the above assumptions, model equations (2.1) transform to give:

$$\frac{\mathrm{d}e}{\mathrm{d}\tau} = \sigma + \frac{\rho c}{\eta + m} - \delta e - \mu c - \frac{\kappa_1}{\varepsilon}(em - c), \tag{2.3a}$$

$$\frac{\mathrm{d}m}{\mathrm{d}\tau} = \alpha m(1 - \beta m) - \kappa_2 c - \frac{\kappa_2}{\varepsilon}(em - c), \tag{2.3b}$$

$$\frac{\mathrm{d}c}{\mathrm{d}\tau} = \frac{1}{\varepsilon}(em - c), \tag{2.3c}$$

with initial conditions

$$e(0) = \frac{E_0}{\hat{E}} =: e_0, \quad m(0) = \frac{M_0}{\hat{M}} =: m_0, \quad \text{and} \quad c(0) = 0. \tag{2.3d}$$

Henceforth, we use dimensionless variables and parameters.

2.2.3 Parameter values

We use the dimensional parameter values given in Table 2.1, which were estimated by Kuznetsov et al. [88]. Growth curves of splenic B cell lymphoma in mice [156] were used to estimate several of these parameters. The logistic growth parameters, a and b, were estimated from tumour growth data for a non-chimeric mouse, where

the tumour mass was assumed to grow in the absence of an immune response. Assuming that the system is in steady state in tumour absence, with the steady state number of effector cells in the region being $\tilde{E} = 3.2 \times 10^5$, and that the lifetime of CTLs is approximately 30 days, i.e. $d_1 \approx 1/30$ days^{-1}, the constant influx rate s was estimated to be $s \approx \tilde{E} d_1$. Other parameters were estimated using data for a chimeric mouse with an immune system. Values for scalings of variables, $\hat{E}$ and $\hat{M}$, were chosen from experiments to be the same order-of-magnitude as the numbers of tumour cells, this is 10^6 cells.

The corresponding dimensionless parameter values are computed in Table 2.2, by assuming that the ratio $\frac{K}{k_1} = 10^6$ cells, or equivalently $\frac{K}{k_1 \hat{E}} \approx 1$. This agrees with the assumptions stated in Section 2.2.2, and implies that $\kappa_1 = \kappa_2 = 1$.

Parameter	**Description**	**Value**
a	net tumour growth rate	0.18 day^{-1}
b	inverse carrying capacity	2.0×10^{-9} cell^{-1}
d_1	per cell death rate of activated effector cells	0.0412 day^{-1}
f	'maximum" tumour-stimulated effector supply rate per complex	1.245×10^5 cells day^{-1}
g	Michaelis-Menten-type constant	2.019×10^7 cells
k_2	rate of effector inactivation	0.1101 day^{-1}
k_3	rate of tumour kill	3.422×10^{-4} day^{-1}
s	constant effector supply rate to the tumour site	1.3×10^4 cells day^{-1}
$\hat{E}, \hat{M}, \hat{C}, \frac{K}{k_1}$	typical tumour cell numbers	10^6 cells

Table 2.1: Summary of the dimensional parameters that appear in the dimensional model (2.1), as estimated in [88].

Parameter	**Description**	**Value**
α	net tumour growth rate	1.6348
β	inverse carrying capacity	0.002
δ	per cell death rate of activated effector cells	0.374
ρ	"maximum" tumour-stimulated effector supply rate per complex	1.131
η	Michaelis-Menten-type constant	20.19
μ	rate of effector inactivation	0.003
σ	constant effector supply rate to the tumour site	0.118
κ_1, κ_2	ratios of variable scales	1

Table 2.2: Summary of the dimensionless parameters that appear in the nondimensional model (2.3). Estimates of these values are computed from dimensional parameters in Table 2.1 using the parameter groupings (2.2).

2.3 Matched asymptotic model approximation

The dependence of equations (2.3) on the small parameter $0 < \varepsilon \ll 1$ means that the system operates on at least two timescales. Exploiting this, we apply the method of matched asymptotic expansions to equations (2.3) in order to derive approximate equations that are simpler to analyse.

We first rewrite the model as a slow-fast system of ordinary differential equations, with a clear separation of timescales. In particular, we seek to eliminate terms in (2.3a) and (2.3b) that scale with $\frac{1}{\varepsilon}$. We define

$$v(\tau) := c(\tau), \qquad \text{and} \qquad \boldsymbol{w}(\tau) := \begin{bmatrix} w_1(\tau) \\ w_2(\tau) \end{bmatrix} = \begin{bmatrix} e(\tau) + \kappa_1 c(\tau) \\ m(\tau) + \kappa_2 c(\tau) \end{bmatrix}, \tag{2.4a}$$

where w_1 and w_2 are, in descriptive terms, the total number of effector and tumour cells respectively, and v is the number of conjugates. With significant algebra, we arrive at a slow-fast system of ODEs for $\boldsymbol{u}(\tau) := [v(\tau), \boldsymbol{w}(\tau)^T]^T$ of the desired form

$$\varepsilon \frac{\mathrm{d}v}{\mathrm{d}\tau} = f(v, \boldsymbol{w}), \tag{2.5a}$$

$$\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}\tau} = \boldsymbol{g}(v, \boldsymbol{w}), \tag{2.5b}$$

with initial conditions

$$v(0) = 0, \qquad \boldsymbol{w}(0) = [w_{1,0}, w_{2,0}]^T = [e_0, m_0]^T, \tag{2.5c}$$

and

$$f(v, \boldsymbol{w}) := em - c = (w_1 - \kappa_1 v)(w_2 - \kappa_2 v) - v, \tag{2.5d}$$

$$\boldsymbol{g}(v, \boldsymbol{w}) := \begin{bmatrix} \sigma + \frac{\rho v}{\eta + w_2 - \kappa_2 v} - \delta(w_1 - \kappa_1 v) - \mu v \\ \alpha(w_2 - \kappa_2 v)(1 - \beta(w_2 - \kappa_2 v)) - \kappa_2 v \end{bmatrix}. \tag{2.5e}$$

The singular nature of system (2.5) means that we cannot construct a uniformly valid solution by taking the limit $\varepsilon \to 0$. We must consider separately approximations on the short and the long timescales, and match them to obtain a uniformly valid approximation of the full model solution $\boldsymbol{u}(\tau)$. We denote the inner and outer solutions by $\boldsymbol{u}_{\mathrm{I}}(\tau_f)$ and $\boldsymbol{u}_{\mathrm{O}}(\tau)$ respectively, where τ_f denotes the rescaled

time variable on the short timescale, proportional to the slow time variable τ via $\tau_f := \tau/\varepsilon$. The composite matched solution $\boldsymbol{u}_{\mathrm{C}}(\tau)$ [70] is then given by

$$\boldsymbol{u}_{\mathrm{C}}(\tau) = \boldsymbol{u}_{\mathrm{I}}(\tau/\varepsilon) + \boldsymbol{u}_{\mathrm{O}}(\tau) - \boldsymbol{u}_{\mathrm{overlap}}, \tag{2.6a}$$

where, to satisfy the initial conditions (2.3d), we impose Prandtl's matching condition

$$\lim_{\tau_f \to \infty} \boldsymbol{u}_{\mathrm{I}}(\tau_f) = \lim_{\tau \to 0} \boldsymbol{u}_{\mathrm{O}}(\tau) = \boldsymbol{u}_{\mathrm{overlap}}. \tag{2.6b}$$

We determine the inner and outer solutions in Sections 2.3.1 and 2.3.2.

2.3.1 Asymptotic approximation of fast dynamics

To analyse behaviour on the short timescale, we rescale time in equations (2.5) via $\tau_f = \tau/\varepsilon$, and obtain the system

$$\frac{\mathrm{d}v}{\mathrm{d}\tau_f} = f(v, \boldsymbol{w}), \tag{2.7a}$$

$$\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}\tau_f} = \varepsilon \boldsymbol{g}(v, \boldsymbol{w}). \tag{2.7b}$$

Let us suppose that

$$v_{\mathrm{I}}(\tau_f; \varepsilon) = \sum_{i=0}^{\infty} \varepsilon^i v_{\mathrm{I}}^{(i)}(\tau_f), \quad \text{and} \quad \boldsymbol{w}_{\mathrm{I}}(\tau_f; \varepsilon) = \sum_{i=0}^{\infty} \varepsilon^i \boldsymbol{w}_{\mathrm{I}}^{(i)}(\tau_f), \tag{2.8}$$

are asymptotic expansions in ε, where subscript I denotes the inner (short-timescale) model solution. By substituting from (2.8) in (2.7) and equating terms of $\mathcal{O}(\varepsilon^i)$, we obtain a sequence of differential equations. At leading order in ε we obtain

$$\frac{\mathrm{d}v_{\mathrm{I}}}{\mathrm{d}\tau_f} = f\left(v_{\mathrm{I}}, \boldsymbol{w}_{\mathrm{I}}\right), \tag{2.9a}$$

$$\frac{\mathrm{d}\boldsymbol{w}_{\mathrm{I}}}{\mathrm{d}\tau_f} = 0, \tag{2.9b}$$

with initial conditions

$$v_{\mathrm{I}}(0) = 0, \qquad \boldsymbol{w}_{\mathrm{I}}(0) = \boldsymbol{w}_0 = [w_{1,0}, w_{2,0}]^T, \tag{2.9c}$$

where $\boldsymbol{u}_{\mathrm{I}}(\tau_f;\varepsilon) \sim \boldsymbol{u}_{\mathrm{I}}^{(0)}(\tau_f)$. For simplicity henceforth $\boldsymbol{u}_{\mathrm{I}}$ will always represent the leading order approximation $\boldsymbol{u}_{\mathrm{I}}^{(0)}$. Equations (2.9) imply that

$$\boldsymbol{w}_{\mathrm{I}}(\tau_f) := \boldsymbol{w}_{\mathrm{I}}(0) = \boldsymbol{w}_0 = [e_0, m_0]^T = \text{constant} \tag{2.10}$$

while the dynamics of the conjugates, v_{I}, are governed by

$$\frac{\mathrm{d}v_{\mathrm{I}}}{\mathrm{d}\tau_f} = Av_{\mathrm{I}}^2 - Bv_{\mathrm{I}} + C, \tag{2.11a}$$

$$A = \kappa_1\kappa_2, \;\; B = 1 + \kappa_2 w_{1,\mathrm{I}} + \kappa_1 w_{2,\mathrm{I}}, \;\; C = w_{1,\mathrm{I}}w_{2,\mathrm{I}}, \tag{2.11b}$$

where $w_{1,\mathrm{I}}$ and $w_{2,\mathrm{I}}$ are constants corresponding to initial conditions (2.9c). For positive initial conditions $\boldsymbol{w}_0 > \boldsymbol{0}$ and $v_0 = 0$, we deduce that v_{I} tends to a positive steady state $v_{\mathrm{I}}^* = v^\dagger(\boldsymbol{w}_0)$ satisfying $f(v_{\mathrm{I}}^*, \boldsymbol{w}_0) = 0$ as $\tau_f \to \infty$, where we define

$$v^\dagger(\boldsymbol{w}) = \frac{1}{2\kappa_1\kappa_2}\left(1 + \kappa_1 w_2 + \kappa_2 w_1 - \sqrt{(1 + \kappa_1 w_2 + \kappa_2 w_1)^2 - 4\kappa_1\kappa_2 w_1 w_2}\right). \tag{2.12}$$

We conclude that, on the short timescale at leading order in ε, the number of conjugates, v_{I}, rapidly relaxes from 0 to $v_{\mathrm{I}}^* = v^\dagger(\boldsymbol{w}_0)$, while the total numbers of tumour and effector cells undergo negligible changes and are so conserved at $\boldsymbol{w}_{\mathrm{I}}^* = \boldsymbol{w}_0$. The numbers of free effector and tumour cells, e and m, are depleted due to conjugate binding to their steady state values $e_{\mathrm{I}}^* = e^\dagger(\boldsymbol{w}_0)$ and $m_{\mathrm{I}}^* = m^\dagger(\boldsymbol{w}_0)$, respectively, where

$$e^\dagger(\boldsymbol{w}) = w_1 - \kappa_1 v^\dagger(\boldsymbol{w}) = \frac{1}{2\kappa_2}\left(\kappa_2 w_1 - 1 - \kappa_1 w_2 + \sqrt{4\kappa_2 w_1 + (\kappa_2 w_1 - 1 - \kappa_1 w_2)^2}\right), \tag{2.13a}$$

$$m^\dagger(\boldsymbol{w}) = w_2 - \kappa_2 v^\dagger(\boldsymbol{w}) = \frac{1}{2\kappa_1}\left(\kappa_1 w_2 - 1 - \kappa_2 w_1 + \sqrt{4\kappa_1 w_2 + (\kappa_1 w_2 - 1 - \kappa_2 w_1)^2}\right). \tag{2.13b}$$

Note that when the steady state $[e_{\mathrm{I}}^*, m_{\mathrm{I}}^*, c_{\mathrm{I}}^*]^T$ is reached, $\frac{\mathrm{d}c_{\mathrm{I}}}{\mathrm{d}\tau_f} = \frac{\mathrm{d}v_{\mathrm{I}}}{\mathrm{d}\tau_f} = f(v_{\mathrm{I}}^*, \boldsymbol{w}_{\mathrm{I}}^*) = 0$, or $v_{\mathrm{I}}^* = c_{\mathrm{I}}^* = e_{\mathrm{I}}^* m_{\mathrm{I}}^*$, and expressions (2.12) and (2.13) satisfy $v^\dagger(\boldsymbol{w}) = e^\dagger(\boldsymbol{w})m^\dagger(\boldsymbol{w})$.

2.3.2 Asymptotic approximation of slow dynamics

We now focus on the long timescale, for which the dynamics of v are slave to the dynamics of $\boldsymbol{w}$, i.e. v changes parametrically with respect to $\boldsymbol{w}$. On this timescale the total effector and tumour cell numbers, w_1 and w_2, vary significantly, since $\mathcal{O}(\varepsilon)$ terms in (2.7) become non-negligible. Formally, we consider the long-timescale system (2.5), and seek solutions for v_{O} and $\boldsymbol{w}_{\mathrm{O}}$ of the form

$$v_{\mathrm{O}}(\tau;\varepsilon) = \sum_{i=0}^{\infty} \varepsilon^i v_{\mathrm{O}}^{(i)}(\tau), \quad \text{and} \quad \boldsymbol{w}_{\mathrm{O}}(\tau;\varepsilon) = \sum_{i=0}^{\infty} \varepsilon^i \boldsymbol{w}_{\mathrm{O}}^{(i)}(\tau). \tag{2.14}$$

At leading order in ε, we obtain the following equations that approximate dynamics on the slow manifold,

$$\frac{\mathrm{d}\boldsymbol{w}_{\mathrm{O}}}{\mathrm{d}\tau} = \boldsymbol{g}(v_{\mathrm{O}}, \boldsymbol{w}_{\mathrm{O}}), \tag{2.15a}$$

$$f(v_{\mathrm{O}}, \boldsymbol{w}_{\mathrm{O}}) = 0. \tag{2.15b}$$

In equation (2.15b), $v_{\mathrm{O}} = v^{\dagger}(\boldsymbol{w}_{\mathrm{O}})$ is the positive solution to the quadratic (2.5d) as defined in (2.12), except that $\boldsymbol{w}_{\mathrm{O}}$ now evolves on the long timescale. Initial conditions for the outer (slow manifold) solution are given by $\lim_{\tau\to 0} = [v^{\dagger}(\boldsymbol{w}_0), \boldsymbol{w}_0^T]^T$, equivalent to the inner solution steady state $\boldsymbol{u}_{\mathrm{I}}^* = [v_{\mathrm{I}}^*, \boldsymbol{w}_{\mathrm{I}}^{*T}]^T$ given in Section 2.3.1, by which the matching condition (2.6b) is automatically satisfied. Thus, on the long timescale, the model reduces to

$$\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}\tau} = \boldsymbol{g}(v^{\dagger}(\boldsymbol{w}), \boldsymbol{w}), \tag{2.16}$$

or equivalently

$$\frac{\mathrm{d}w_1}{\mathrm{d}\tau} = \sigma + \frac{\rho e^{\dagger}(\boldsymbol{w}) m^{\dagger}(\boldsymbol{w})}{\eta + m^{\dagger}(\boldsymbol{w})} - \delta e^{\dagger}(\boldsymbol{w}) - \mu e^{\dagger}(\boldsymbol{w}) m^{\dagger}(\boldsymbol{w}), \tag{2.17a}$$

$$\frac{\mathrm{d}w_2}{\mathrm{d}\tau} = \alpha m^{\dagger}(\boldsymbol{w})(1 - \beta m^{\dagger}(\boldsymbol{w})) - \kappa_2 e^{\dagger}(\boldsymbol{w}) m^{\dagger}(\boldsymbol{w}), \tag{2.17b}$$

where we drop the subscript O notation for simplicity. Since equations (2.17) are highly nonlinear, we simplify the steady state and linear stability analysis in later

sections, by rewriting them in terms of the dimensionless variables e and m:

$$\begin{bmatrix} \dot{e} \\ \dot{m} \end{bmatrix} = \begin{bmatrix} 1+\kappa_1 m & \kappa_1 e \\ \kappa_2 m & 1+\kappa_2 e \end{bmatrix}^{-1} \underbrace{\begin{bmatrix} \sigma + \frac{\rho e m}{\eta + m} - \delta e - \mu e m \\ \alpha m(1-\beta m) - \kappa_2 e m \end{bmatrix}}_{=: \boldsymbol{q}(e,m)}. \tag{2.18}$$

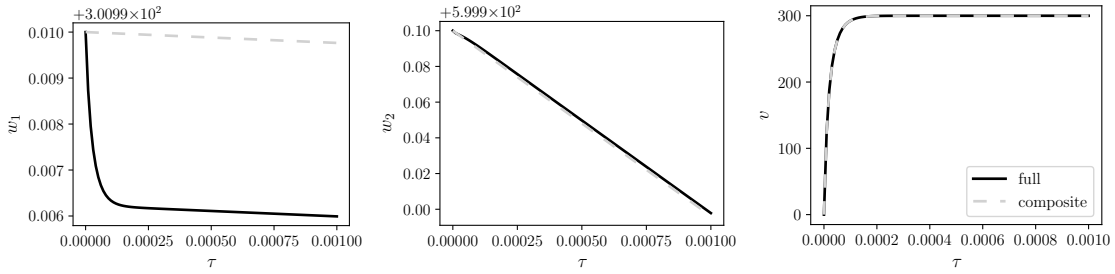
This contrasts with Kuznetsov et al.'s QSSA model which corresponds to equations (2.3) with $c = em$.

2.3.3 Comparison of numerical solutions

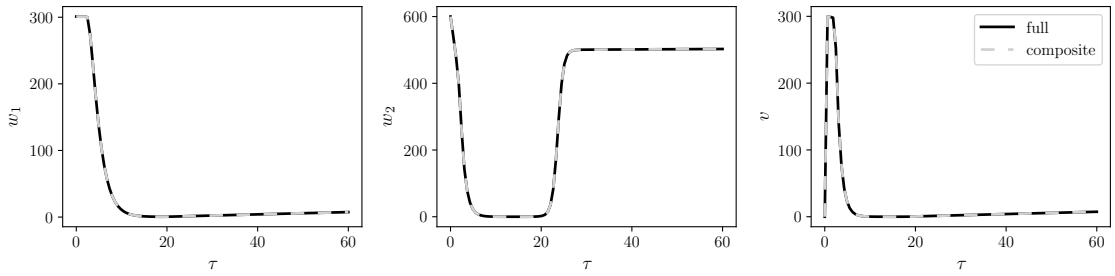
To confirm the accuracy of the asymptotic reduction, we compare the solution of the full system (2.5), equivalent to (2.3), with the composite solution (2.6). We solve the full system by integrating equations (2.5) using the Python function `scipy.integrate.odeint`, a wrapper for the `lsoda` solver from Fortran's library `odepack`. This ODE solver automatically switches between stiff and nonstiff methods, noting the system is stiff when the dynamics are fast. To construct the composite solution, we use the same solver to generate numerical solutions for the fast and slow systems, whereby integration of the governing equations for both systems is performed with respect to the slow timescale τ. We then combine numerical solutions using the matching condition (2.6).

The results presented in Fig. 2.2 show that the full and composite solutions are in good agreement for small and large times, i.e. during the fast and slow dynamics. As anticipated, the relative error in the composite solution compared to the full solution is of $\mathcal{O}(\varepsilon)$. We conclude that the matched asymptotic approximation of the full solution is accurate within the expected error range, and note that the solutions relax rapidly onto the slow manifold dynamics. Based on these findings, we assert that the reduced model characterises the long-term dynamics of the full model, by which we motivate its analyses in Section 2.4 and Chapter 3.

Further information about how the full and asymptotic long-timescale solutions, i.e. solutions of equations (2.5) and (2.17) respectively, compare with solutions of



(a)



(b)

Figure 2.2: Series of plots showing good agreement between the composite and full solutions of equations (2.5) on (a) the short timescale ($0 \leq \tau \leq \varepsilon/10$), where the inner solution dominates, and on (b) the long timescale ($\varepsilon/10 \leq \tau \leq 60$), where the outer solution dominates. The left, middle and right panel plots respectively show the solutions for model variables $w_1(\tau)$, $w_2(\tau)$ and $v(\tau)$ over time. The solutions are computed numerically as described in Section 2.3.3, for initial conditions $w_1(0) = 301$, $w_2(0) = 600$, $v(0) = 0$, and default parameter values with $\varepsilon = 10^{-2}$.

the QSSA model from [88] are provided in Section 2.5, and also in Section 3.5 of Chapter 3.

2.4 Characterisation of model excitability

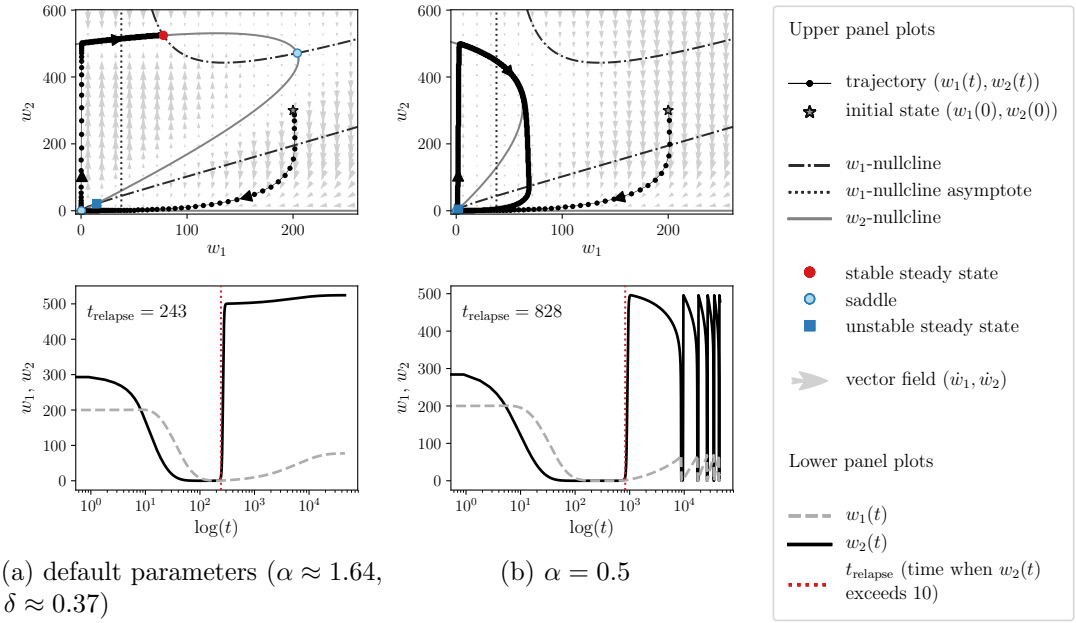
We now characterise the long-term behaviour of the composite solution by analysing the outer equations (2.17) using approaches similar to those used to study the *Fitzhugh-Nagumo* (FN) *equations*, which model excitability in neuronal dynamics [116]. A system is excitable if a small perturbation from a stable steady state results in a large excursion before the system relaxes back to the same or a different steady state. Alternatively, the system may relax onto a limit cycle with large excursions. In this section, we show that our reduced system (2.17) exhibits excitable dynamics.

2.4.1 Nullclines

On the w_1- and w_2-nullclines of the slow system (2.17), we have $\dot{w}_1 = 0$ and $\dot{w}_2 = 0$, respectively. We find that the w_1-nullclines are a subset of solutions to a quadratic in w_2, which has polynomials in w_1 as its coefficients. The w_2-nullclines are the line $w_2 = 0$ and a subset of solutions to another quadratic in w_2, with coefficients independent of w_1. For the sake of brevity, we do not present the complex, implicit, equations for these nullclines. For default parameter values in Fig. 2.3a, the nontrivial w_2-nullcline has a shape of a rotated and skewed U that, together with the $w_2 = 0$ line, resembles a cubic; this is intersected in different ways by the w_1-nullcline. The resulting nullcline geometry is similar to that of the FN model, where a straight line nullcline associated with one dependent variable intersects a cubic nullcline associated with the other dependent variable [116].

We investigate how robust the nullclines and behaviour of the system (2.17) are to small perturbations in parameter values. The U shape of the non-trivial w_2-nullcline is preserved under changes in tumour kinetic constants α (see Figs. 2.3a and 2.3b) or β. Under variation of parameters associated with the immune system and its response to the tumour, the shape of the w_1-nullcline ranges from approximately linear, as in the FN model, to curved (see Figs. 2.3a, 2.3c and 2.3d), in which it intersects the w_2-nullcline, hence the number and stability of the steady states also vary. The shape of the w_2-nullcline is therefore robust to parameter changes, whereas that of the w_1-nullcline is not.

All solution trajectories presented in Fig. 2.3 exhibit excitable behaviour, with large excursions in the phase plane and switching kinetics. A typical trajectory exhibits rapid vertical dynamics, during which the tumour grows (or shrinks) quickly, until it reaches the w_2-nullcline. The trajectory then moves approximately horizontally on this nullcline, with the number of effector cells either increasing or decreasing on a slow timescale. When the horizontal direction can no longer be sustained, the trajectory is pushed onto a different branch of the w_2-nullcline via



(a) default parameters ($\alpha \approx 1.64$, $\delta \approx 0.37$)

(b) $\alpha = 0.5$

(c) $\delta = 0.11$

(d) $\delta = 1.1$

Figure 2.3: Behaviour of the long-timescale model (2.17), shown via (w_1, w_2) phase portraits and dynamic plots of example solutions, is excitable, and remains excitable as individual parameters are varied. Plot (a) shows model behaviour for default parameter values (see Table 2.2). In plot (b), the tumour growth rate parameter α is varied, whereas the effector death rate parameter δ is varied in plots (c) and (d), with the other parameters fixed at their default values in all plots. (w_1, w_2) phase portraits in upper panels are equipped with nullclines, steady states, flow field $(\dot{w}_1, \dot{w}_2)$ and an example solution trajectory. The latter, starting at $w_1(0) = 200$, $w_2(0) = 300$, is obtained numerically using the same method as in Section 2.3.3. The time step between plotted trajectory points is uniform; points far apart indicate fast dynamics, whereas points close together show slow dynamics. In the lower panels the solution trajectory is plotted against time for $0 < t \leq 36331$ days. The red dotted vertical line denotes the time, t_{relapse} (units: days), which we define to be the earliest time, after the first shrinkage, at which the tumour cell density, w_2, exceeds the chosen threshold value of 10×10^6 cells; t_{relapse} increases when δ or α are decreased. We note that the trajectory simulation time period (36331 days) exceeds the lifespan of chimeric mice (approx. 1000 days [42]), for which default parameters were estimated. Trajectory simulations beyond 1000 days are not physically realistic for a chimeric mouse, but are generally informative for characterising phase plane dynamics. We speculate that the qualitative dynamics observed here may operate on shorter timescales in a different parameter regime.

fast vertical movement. The rapid switches between nullcline branches continue until a stable steady state or limit cycle is reached. We note that changes in the shape of the w_1-nullcline under single parameter variations do not greatly impact large phase-plane excursions of trajectories, although they do impact long-term behaviour.

2.4.2 Separation of timescales and its implications

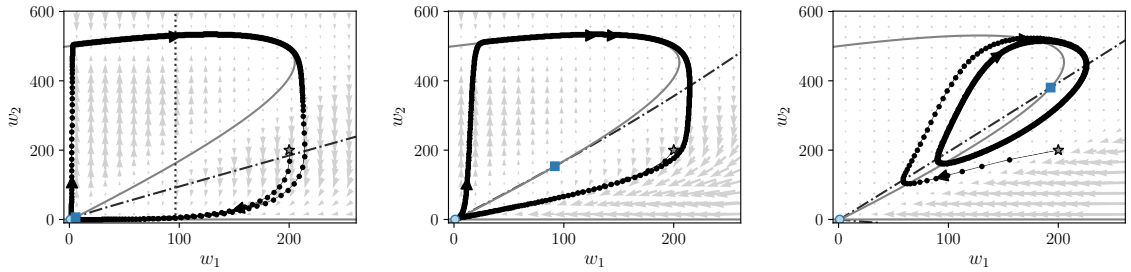
In Fig. 2.3, we observe that the transient fast kinetics arise due to the dominance of the vertical direction $(0, \dot{w}_2)$ in the flow field, except in the vicinity of the w_2-nullcline, where $\dot{w}_1$ is significant. This indicates a separation of timescales, whereby w_2 evolves on a shorter timescale than w_1. As model parameters are varied, the length of time trajectories spend on the w_2-nullcline also varies. As we decrease the pro-tumour/anti-immune parameters, such as δ, the rate of effector cell death/migration out of the tumour, or α, the tumour growth rate, the time to tumour relapse (i.e. once tumour cell numbers exceed a threshold value of 10×10^6 cells after the first shrinkage) increases. While this indicates that the difference in timescales changes as model parameters vary, the excitable dynamics seem to be robust to small changes in individual parameters.

Inspection of the default parameter values in Table 2.2 suggests the model parameters can be rescaled in terms of two positive small parameters, $\zeta = \mathcal{O}(10^{-1})$ and $\xi = \mathcal{O}(10^{-3})$, which we treat independently in order to preserve the richness of the model when their values are close to zero. We rescale the model parameters as follows

$$\sigma = \zeta\hat{\sigma}, \quad \delta = \zeta\hat{\delta}, \quad \eta = \frac{\hat{\eta}}{\zeta}, \quad \mu = \xi\hat{\mu}, \quad \beta = \xi\hat{\beta}, \tag{2.19}$$

with $\alpha = \mathcal{O}(1)$ and $\rho = \mathcal{O}(1)$. Under (2.19), equations (2.17) transform to give

$$\dot{w}_1 = \zeta\left(\hat{\sigma} + \frac{\rho e m}{\hat{\eta} + \zeta m} - \hat{\delta}e - \frac{\xi}{\zeta}\hat{\mu}em\right) =: \zeta g_1(w_1, w_2), \tag{2.20a}$$



(a) $\zeta = 0.1$ (b) $\zeta = 1$ (c) $\zeta = 10$

Figure 2.4: Dynamics of the long-timescale model (2.20), shown via (w_1, w_2) phase portraits, evolve from excitable to non-excitable as ζ increases from 0.1 to 10; ζ is a dimensionless parameter that scales parameters associated with effector cells (see equation (2.19)). Large excursions in the phase plane in plots (a) and (b) transform into regular oscillations in plot (c). The plots are generated using $\hat{\delta} = \hat{\sigma} = 3$, while other parameters are kept at their default values. Trajectories start at $w_1(0) = w_2(0) = 200$. See Fig. 2.3 for legend.

$$\dot{w}_2 = \alpha m - \kappa_2 e m - \xi \alpha \hat{\beta} m^2 =: g_2(w_1, w_2), \tag{2.20b}$$

where we drop the notation $e = e^\dagger(\boldsymbol{w})$ and $m = m^\dagger(\boldsymbol{w})$ from (2.17) for simplicity.

Equations (2.20) are singular, and, in the limiting case $0 < \xi < \zeta \ll 1$, we can apply asymptotic methods, as in Section 2.3, to further approximate the fast and slow dynamics. For brevity, we only sketch this analysis here. In the limit $\zeta, \xi \to 0$, the leading order fast manifold equations for $\boldsymbol{w}_\mathrm{F}$ are

$$\dot{w}_{1,\mathrm{F}} = 0, \tag{2.21a}$$

$$\dot{w}_{2,\mathrm{F}} = \alpha m^\dagger(\boldsymbol{w}_\mathrm{F}) - \kappa_2 e^\dagger(\boldsymbol{w}_\mathrm{F}) m^\dagger(\boldsymbol{w}_\mathrm{F}). \tag{2.21b}$$

From equations (2.21), we note that the total number of effector cells, $w_{1,\mathrm{F}}$, is approximately constant on the fast manifold, while the dynamics of the total number of tumour cells, $w_{2,\mathrm{F}}$, are governed by the difference between exponential growth and immune-induced cell death in equation (2.3b) for $\dot{m}$. We deduce that, on the fast timescale, the total number of tumour cells diverges to infinity if $w_{1,\mathrm{F}} < w_1^\times := \frac{\alpha}{\kappa_2}(1 + \frac{\alpha\kappa_1 w_2}{1+\alpha})$, and is attracted to the tumour-free steady state if $w_{1,\mathrm{F}} > w_1^\times$, i.e. enough effector cells are recruited relative to the tumour size.

Once a trajectory that follows the fast manifold reaches the neighbourhood of the w_2-nullcline, it starts moving on the slow manifold of the outer equations

(2.20). Rescaling time via $\tilde{\tau} = \zeta\tau$ in equations (2.20), yields the slow manifold equations for $\boldsymbol{w}_{\mathrm{S}}$ with respect to $\tilde{\tau}$:

$$\dot{w}_{1,\mathrm{S}} = g_1(w_{1,\mathrm{S}}, w_{2,\mathrm{S}}), \tag{2.22a}$$

$$g_2(w_{1,\mathrm{S}}, w_{2,\mathrm{S}}) = 0, \tag{2.22b}$$

where equation (2.22b) constrains the trajectory to the w_2-nullcline, i.e. the dynamics of $w_{2,\mathrm{S}}$ are slave to those of $w_{1,\mathrm{S}}$. In contrast to the asymptotic solutions derived in Section 2.3, the trajectories do not remain on the slow manifold, but jump between the fast and slow manifolds. When moving on the slow manifold, i.e. on a branch of the w_2-nullclines, the number of total effector cells changes slowly, which eventually causes a change in the attracting steady state of the fast manifold equations. When $w_{1,\mathrm{F}} \approx w_1^{\times}$, solution trajectories switch via fast dynamics to a different branch of the w_2-nullclines (see Fig. 2.3). We conclude that the separation of timescales and the logistic growth assumption allow tumour cells to exploit small changes in the effector cell numbers to rapidly expand to carrying capacity in this parameter regime.

In Fig. 2.4 we show how the system dynamics may change from excitable to non-excitable as the value of ζ increases. For smaller values of ζ, transitions between the fast and slow manifolds are sharp, and the trajectory moves on the above manifolds (see Fig. 2.4a). For the case shown, the trajectory is a limit cycle, with a large period and amplitude. For larger values of ζ, the excitable dynamics are less pronounced; they resemble classic oscillations, with smaller excursions and periods of oscillation, and they do not follow the nullclines closely (see Fig. 2.4c). In this case, the timescale of effector dynamics approaches those of the tumour, rather than the tumour dominating and the immune system slowly catching up, as in the excitable regime.

In summary, our results show that the model's excitable behaviour can be attributed to a combination of the robust, cubic-like, shape of the w_2-nullcline,

the separation of timescales, and the logistic growth assumption. Excitability is observed for $\zeta \lesssim \mathcal{O}(10^{-1})$, and a considerable inflation of ζ is needed to eliminate the excitable behaviour (see Figs. 2.3 and 2.4). We conclude that in a neighbourhood of the default parameter regime the system (2.17) is inherently prone to excitable behaviour, analogously to the FN model.

2.5 Comparison with the QSSA model

Here we discuss how the long-timescale approximation (2.17) (or equivalently (2.18)) of the original model (2.3) differs from Kuznetsov et al.'s QSSA model, given in its dimensionless form as

$$\frac{\mathrm{d}e}{\mathrm{d}\tau} = \sigma + \frac{\rho e m}{\eta + m} - \delta e - \mu e m, \tag{2.23a}$$

$$\frac{\mathrm{d}m}{\mathrm{d}\tau} = \alpha m(1 - \beta m) - \kappa_2 e m, \tag{2.23b}$$

with $c = em$, initial conditions $e(0) = e_0$ and $m(0) = m_0$, and dimensionless parameter groupings defined in (2.2).

We first note that the equation $c = em$ in the QSSA model does not generally satisfy the initial conditions of the full model (2.3) since, for some positive e_0 and m_0, $c(0) = e(0)m(0) = e_0 m_0$ is non-zero. As explained in Section 2.3, this is also the case for the solution of the long-timescale (outer) model equations (2.18) (or equivalently (2.17)). Close to $\tau = 0$, the full model has a boundary layer, in which the appropriate approximation is the solution of the inner equations (2.9) with arbitrary initial conditions. In order to compare solutions on the long-timescale, without loss of generality, we simulate the long-timescale, QSSA and full models using the steady state of the inner problem as initial conditions, so that discrepancies in the boundary layer do not occur. For initial conditions of the full model such that $[e(0), m(0), c(0)] = [e_0, m_0, 0]$, the steady state of the inner problem is $[e_{\mathrm{I}}^*, m_{\mathrm{I}}^*, c_{\mathrm{I}}^*] = [e_0 - \kappa_1 v^\dagger([e_0, m_0]), m_0 - \kappa_2 v^\dagger([e_0, m_0]), v^\dagger([e_0, m_0])]$ (see

Section 2.3.1 for details). Equivalently, in terms of variables w_1, w_2 and v, for initial conditions $[w_1(0), w_2(0), v(0)] = [e_0, m_0, 0]$, the corresponding steady state of the inner problem is $[w^*_{1,\mathrm{I}}, w^*_{2,\mathrm{I}}, v^*_{\mathrm{I}}] = [e_0, m_0, v^\dagger([e_0, m_0])]$.

For equivalent parameter values and initial conditions, solutions of the long-timescale and the QSSA models differ greatly (see Fig. 2.5b). Solutions of the QSSA model do not follow those of the full model (2.3) closely, while those of the reduced model do. We conclude that the QSSA model is not an accurate

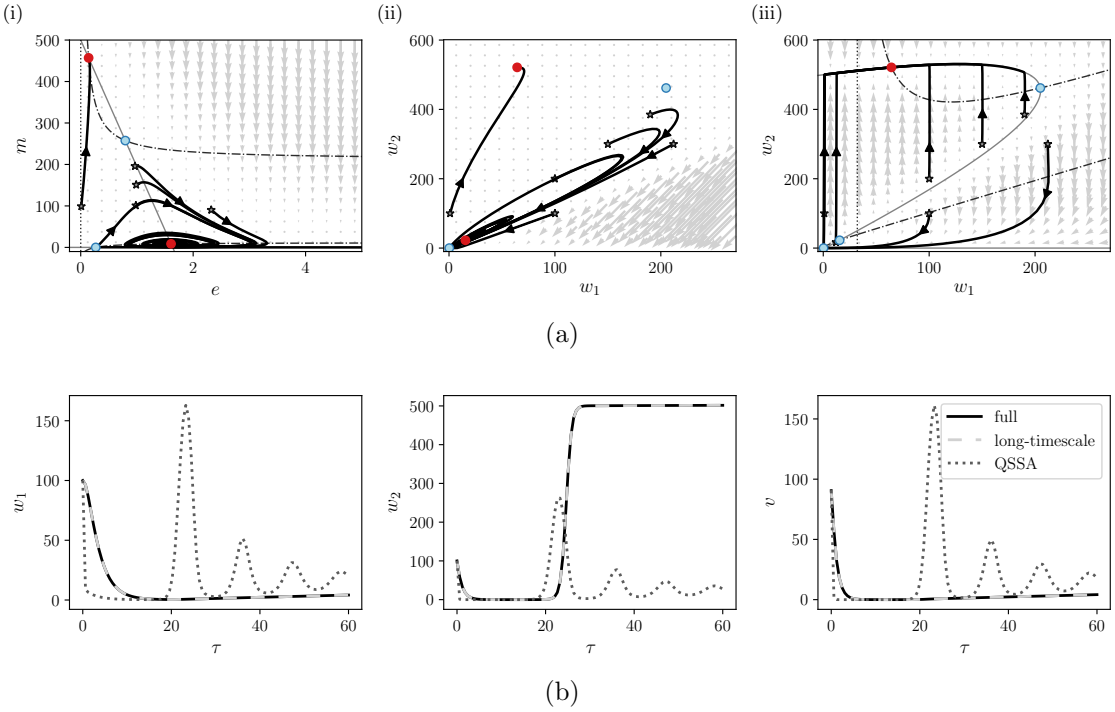


Figure 2.5: Series of plots (a) showing phase portraits of (i) the QSSA model (2.23) in (e, m) phase space, (ii) the same model in (w_1, w_2) phase space, and (iii) the asymptotic long-timescale model (2.17) in (w_1, w_2) phase space. The QSSA model (see (i) and (ii)) does not exhibit excitable dynamics that are observed in model (2.17) (see (iii) and also Figs. 2.3 and 3.4); trajectories do not follow nullclines, and no limit cycles are observed. See Fig. 3.4 for phase portrait legend. In all three plots, equivalent initial conditions, appropriately transformed and fixed at the steady state of inner model equations (2.9) as described in the main text, are used to compute solution trajectories via the methods described in Section 2.3.3. Trajectories in plot (a)(ii) are generated by numerically solving equations (2.23) first; the obtained solutions for e and m are then transformed via $w_1 = e + \kappa_1 em$ and $w_2 = m + \kappa_2 em$ to compute numerical solutions for total effector and tumour cell numbers, w_1 and w_2. The series of plots (b) show different solutions for model variables (i) $w_1(\tau)$, (ii) $w_2(\tau)$ and (iii) $v(\tau)$ over time $(0 \leq \tau \leq 60)$ for the one set of initial conditions from (a), given as $w_1(0) = w_2(0) = 100$, $v = v^\dagger(\boldsymbol{w}(0)) = 90.48$. Solid black, dashed light grey and dotted dark grey curves respectively correspond to solutions of the full model (2.3), the long-timescale model (2.17) and the QSSA model (2.23). The QSSA solution does not agree with the matching full and long-timescale solutions. For all plots in (a) and (b), parameter σ is fixed at $\sigma = 0.1$, and other parameters at their default values (see Table 2.2).

approximation of the full model in the considered parameter regime.

In Fig. 2.5a we present phase portraits for both models in terms of w_1 and w_2 (total effector and tumour cell numbers), and for the QSSA model in terms of e and m (free effector and tumour cells). While our long-timescale model exhibits excitable behaviour, the QSSA model does not, neither in the (e, m) phase portrait nor in the (w_1, w_2) one. We detect differences in the stability of the steady states of the two models: all solutions of our reduced model evolve to a large-tumour steady state, while some solutions of the QSSA model exhibit damped oscillations to a small-tumour steady state. Differences in steady state stability motivate our analysis of the reduced model's bifurcation structure in Chapter 3.

2.6 Discussion and conclusions

We investigated the five-equation mathematical model by Kuznetsov et al. [88] to increase understanding of the impact that nonlinear tumour-immune interactions have on tumour growth dynamics. We focused on how simplification of effector-tumour conjugate dynamics affects model behaviour. We performed a new asymptotic reduction of the model, and characterised its behaviour using phase-plane techniques. We identified parameter regimes for which the model is excitable, a feature of tumour-immune interactions that, to our knowledge, has not been previously reported for existing models of tumour-immune interactions.

In Kuznetsov et al.'s original model a separation of timescales arises because the rates of conjugate formation and dissociation with both cells in the conjugate remaining functional are assumed much faster than all other processes included in the model. We exploited this separation of timescales to derive a two-dimensional matched asymptotic approximation of the original model that is more accurate in this distinguished limit than the quasi-steady-state-approximated (QSSA) model proposed by Kuznetsov et al. [88]. On the short timescale, the total numbers of

tumour and effector cells do not vary at leading order, while conjugates relax to a steady state. This allows the full model to be simplified on the long timescale, where conjugates vary parametrically with respect to the total number of effector and tumour cells. Comparison of the asymptotically reduced model with the full model on both timescales and in multiple parameter regimes showed that it preserves excitable dynamics of the full model. This contrasts with the behaviour of the QSSA model as, for example, we found differences in long-term model outcomes; we characterise this further in Chapter 3. This result highlights that special care is needed when using quasi-steady-state assumptions to approximate mathematical models. We note that parameter estimates for the rates associated with conjugate dynamics are needed to justify our asymptotic reduction; this may motivate further investigation of alternative distinguished limits of Kuznetsov et al.'s original model.

On the long-timescale, and for parameters in a neighbourhood of the default parameter regime, the tumour-immune dynamics of our reduced model are excitable because of the structure of nullclines and the large ratio of timescales between effector and tumour cell dynamics. Excitability manifests through large phase-plane excursions, characterised by alternation of short periods of large and rapid changes in total tumour cell numbers, and long periods of slow variation in total effector cell numbers, during which tumour cell numbers remain close to zero or the carrying capacity. The switch from slow to fast dynamics is driven by exponential growth of the tumour when effector cell numbers are much smaller than tumour cell numbers. This behaviour illustrates how the underlying saturating growth (here modelled by logistic growth) may contribute to excitability. It also provides a possible explanation for why tumours relapse after periods of remission. Long-term tumour recurrence is exhibited by the model through long periods with low tumour cell numbers, followed by rapid (exponential) tumour regrowth. We note that the length of the remission period and, hence, the excitable dynamics, are sensitive to small parameter changes, particularly when cell numbers are small. This suggests

that stochastic effects may have a significant impact on the model's excitable dynamics, and may, at small cell numbers, lead to tumour eradication, rather than recurrence.

In our study we used parameter values that were estimated in [88] by fitting the QSSA model to murine data [156]. We stress that physically realistic parameter regimes for the full model must be identified by calibrating the model to murine and/or patient data. Further, for biological realism, parameter values must be tuned so that the system dynamics occur on timescales which are shorter than the lifespan of the host (for a chimeric mouse, this is approx. 1000 days [42]). In our work, longer simulation times are sometimes used. This is done in order to evaluate the behaviour of the reduced model, and provide a basis for future characterisation of model behaviour in alternative parameter regimes of biological relevance, or with stochastic effects, where oscillation periods and/or times to tumour extinction/recurrence may be shorter than those presented in this thesis. We note that the long timescales associated with the slow dynamics in the model may make excitable features of tumour-immune interactions difficult to identify, as they may exceed the lifespan of the animal or patient. The field of cardiac electrophysiology has enjoyed great success in calibrating models with excitable dynamics and using them to generate quantitative predictions [177]. Progress in this field has come gradually with improvements in experiment design (e.g. by using complex stimulus protocols) and parameter estimation (e.g. by fitting additional quantities derived from measurements). We hope that calibration of models in immuno-oncology will follow a similar path in future.

There are several ways in which the original model could be modified to better describe tumour-immune interactions. We could generalise the logistic tumour growth term to include competition with other cells in the tumour region for resources, growth factors and space. The assumption of a constant supply of activated cytotoxic T cells (CTLs), $s > 0$, is only valid when the region has been

previously exposed to tumour antigens, or if the organism has been treated with a cancer vaccine or adoptive cell therapy (e.g. [79]). We could omit the baseline effector influx term in a treatment-free model [75, 79, 84, 96] where activation of CTLs is predominantly tumour-dependent. Our excitable model does not capture immunosuppression at large tumour sizes, since the number of effector cells does not decrease as the tumour increases in size; this could also be an area for further exploration.

The simplifying assumptions of the original model by Kuznetsov et al. [88] allowed us to reduce it to a two-dimensional tumour-effector model that more faithfully preserves features of the original model in the distinguished limit of fast conjugate dynamics. While the model has several limitations and requires further experimental validation, it revealed how excitability may underpin tumour-immune interactions. We also identified the need to investigate further the impact of excitability on the predicted long-term dynamics; this will be the focus of Chapter 3.

Chapter 3

Characterisation and parameter sensitivity of qualitative model behaviour

The research presented in this chapter has been published in the *Journal of Theoretical Biology* [120].

3.1 Introduction

Parameters in ODE models of tumour-immune interactions may vary for many reasons. Parameters estimated from clinical data will have uncertainties associated with measurement errors or modelling discrepancy; they may also vary between patients, depending on the aggressiveness and immunogenicity of their tumours, and the strength of the immune responses. Administration of chemo- and/or immunotherapy may also alter more model parameters. It is therefore crucial to investigate the effects of parameter variations on model behaviour in order to understand the full range of behaviours a model may exhibit, and predictions it can make. Parameter sensitivity analyses can also be used to explore the potential benefits of (immuno)therapies, or combinations thereof, as the mechanisms targeted by the therapy may be represented by varying the corresponding model parameters. For example, Kirschner and Panetta [79] investigated the impact of administering the cytokine interleukin-2 (IL-2) and adoptive cell therapy (ACT), separately and

in combination, on long-term outcomes. In their model, parameters representing the (constant) supply rates of IL-2 and immune cells were varied to study the effects of IL-2 and ACT on the model's bifurcation structure. In this way, ACT as a monotherapy or combined with IL-2 was found to be more beneficial for the patient than IL-2 alone. Dritschel et al. [46] performed a comprehensive bifurcation analysis of their model to demonstrate that targeting Ths with ACT may be more effective than targeting CTLs. Similarly, Besse et al. [13] investigated the outcomes of cessation of treatment with a tyrosine kinase inhibitor, imatinib, in chronic myelogenous leukemia. They found that, depending on the immune system efficacy, treatment with imatinib can lead to cancer remission by inhibiting leukemic cell proliferation.

In Chapter 2 we derived a new approximation of the original model of Kuznetsov et al. [88], valid in the limit of fast effector-tumour conjugate formation and dissociation. We showed that excitable dynamics emerge from the reduced model. Studies of systems with excitable features, such as the Fitzhugh-Nagumo (FN) model, have shown that they tend to exhibit complex bifurcation structures [32, 83]. In this chapter we investigate how excitability influences the long-term outcomes of our reduced model, and its implications for our understanding of tumour-immune dynamics and therapy. Our aims are to characterise how the efficacy of the immune system affects tumour outcomes, and to demonstrate that phenomena such as immunoediting may be captured in the bifurcating behaviour of our excitable model. Our approach is similar to [13, 46, 79]. We perform a comprehensive parameter sensitivity analysis of the system's long-term behaviour, focusing on parameters related to the immune system that may vary between patients or with immunotherapy. We place these insights within the context of immuno-oncology, and the theoretical prospect of implementing tumour control.

The remainder of this chapter is organised as follows. In Section 3.2 we use asymptotic and linear stability analysis to identify and characterise the steady

states of the reduced model. We confirm and extend these results numerically in Sections 3.3 and 3.4, where we describe the system's global bifurcation structure for different values of model parameters relating to effector cell supply and decay, σ and δ, respectively. We also explore the sensitivity of the bifurcation structure to changes in μ, ρ and η, parameters associated with the immune response to the tumour, in order to gain insight into tumour responses to different immunotherapies. We compare our results to analogous ones from Kuznetsov et al.'s QSSA model in Section 3.5. We summarise our findings in Section 3.6, where we also explain how they can be used to understand tumour-immune interactions and to identify immunotherapeutic targets, represented by model parameters, that may successfully eliminate a tumour.

3.2 Characterisation of steady state solutions

Identifying the steady state solutions of the slow system is most tractable when analysing equations (2.18) for variables e and m. The steady state solutions solve $\boldsymbol{q}(e, m) = 0$ and are identical to those of the full system (2.3). We find that there is *a tumour-free steady state* $(e_1^*, 0)$, with

$$e_1^* = \frac{\sigma}{\delta}, \tag{3.1}$$

and *up to three nonzero-tumour steady states* $(e_{2,3,4}^*, m_{2,3,4}^*)$ (physically realistic when both e^* and m^* are real and nonnegative) given by

$$e_{2,3,4}^* = \frac{\alpha}{\kappa_2}(1 - \beta m) \qquad \text{and} \qquad m_{2,3,4}^* = m, \tag{3.2}$$

where m is a root of the cubic

$$p(m) := a_3 m^3 + a_2 m^2 + a_1 m + a_0 \tag{3.3a}$$

with coefficients

$$\begin{aligned} a_0 &= \eta\left(\tfrac{\sigma\kappa_2}{\alpha} - \delta\right), & a_1 &= \tfrac{\sigma\kappa_2}{\alpha} + \rho - \mu\eta - \delta + \delta\eta\beta, \\ a_2 &= -\mu + (\mu\eta + \delta - \rho)\beta, & a_3 &= \mu\beta. \end{aligned} \tag{3.3b}$$

3.2.1 Asymptotic approximation of steady states

Recalling that μ and β are small for the default parameter values (see Table 2.2), we introduce the artificial small parameter $\xi = \mathcal{O}(10^{-3})$ as before, and write $\mu = \xi\hat{\mu}$ and $\beta = \xi\hat{\beta}$. This allows us to approximate the roots m of the cubic (3.3a), and estimate the nonzero-tumour steady states $(e^*_{2,3,4}, m^*_{2,3,4})$. We detail this analysis in Appendix A, and summarise the key results below.

When $0 < \xi \ll 1$, there are at most three types of steady state solutions:

(i) the *tumour-free steady state* $(e^*_1, 0) = (\frac{\sigma}{\delta}, 0)$,

(ii) up to one *intermediate-sized-tumour steady state* (e^*_2, m^*_2), where

$$e^*_2 \sim \frac{\alpha}{\kappa_2} = \mathcal{O}(1) \quad \text{and} \quad m^*_2 \sim -\frac{a_0}{\hat{a}_1} = \mathcal{O}(1), \tag{3.4}$$

and

(iii) up to two *large-tumour steady states* $(e^*_{3,4}, m^*_{3,4})$, of the form

$$e^*_{3,4} \sim \frac{\alpha}{\kappa_2}(1 - \hat{\beta}\hat{m}_{3,4}) = \mathcal{O}(1) \quad \text{and} \quad m^*_{3,4} \sim \frac{1}{\xi}\hat{m}_{3,4} = \mathcal{O}(\xi^{-1}), \tag{3.5a}$$

$$\text{where} \qquad \hat{m}_{3,4} = \frac{-\hat{a}_2 \pm \sqrt{\hat{a}_2^2 - 4\hat{a}_3\hat{a}_1}}{2\hat{a}_3}. \tag{3.5b}$$

The coefficients in equations (3.4) and (3.5) are $\mathcal{O}(1)$, and related to the original model parameters via

$$\hat{a}_1 = \frac{\sigma\kappa_2}{\alpha} + \rho - \delta, \qquad \hat{a}_2 = -\hat{\mu} + \hat{\beta}(\delta - \rho), \qquad \hat{a}_3 = \hat{\mu}\hat{\beta}. \tag{3.6}$$

In regions of parameter space where the steady states are physically realistic and stable, we may identify them with the elimination, equilibrium/dormancy and escape phases associated with the three Es of immunoediting. In Section 3.2.2 we identify regions of parameter space in which the different steady state solutions exist and are physically realistic. In Sections 3.2.3 and 3.3 we examine their stability analytically and numerically, respectively.

3.2.2 Physicality of steady states

It is straightforward to note that the tumour-free steady state $(\frac{\sigma}{\delta}, 0)$ exists for all positive values of σ and δ. From (3.4) we deduce that the intermediate-sized-tumour steady state is nonnegative if

$$\delta^\diamond(\sigma, \alpha, \kappa_2) \leq \delta < \rho + \delta^\diamond(\sigma, \alpha, \kappa_2), \tag{3.7a}$$

$$\text{where} \qquad \delta^\diamond(\sigma, \alpha, \kappa_2) := \frac{\sigma\kappa_2}{\alpha}, \tag{3.7b}$$

or equivalently $\frac{a_0}{\eta} \leq 0 < \hat{a}_1$. We use (3.5) to derive similar conditions for the large-tumour steady states (see Appendix A for details). In short, provided $\hat{m}_{3,4} \leq \beta^{-1}$, there is a unique large-tumour steady state solution (e_3^*, m_3^*) when $\hat{a}_1 < 0$, two such solutions when $0 < \hat{a}_1 < \hat{a}_2^2/4\hat{a}_3 =: \hat{a}^\#$ and $\hat{a}_2 < 0$, and none otherwise.

In Fig. 3.1 we plot the boundaries of the physically realistic regions in the (σ, δ) plane, predicted by the asymptotic analysis. We compare them to a colour plot of the (σ, δ) plane, which shows how the number of physically realistic steady states

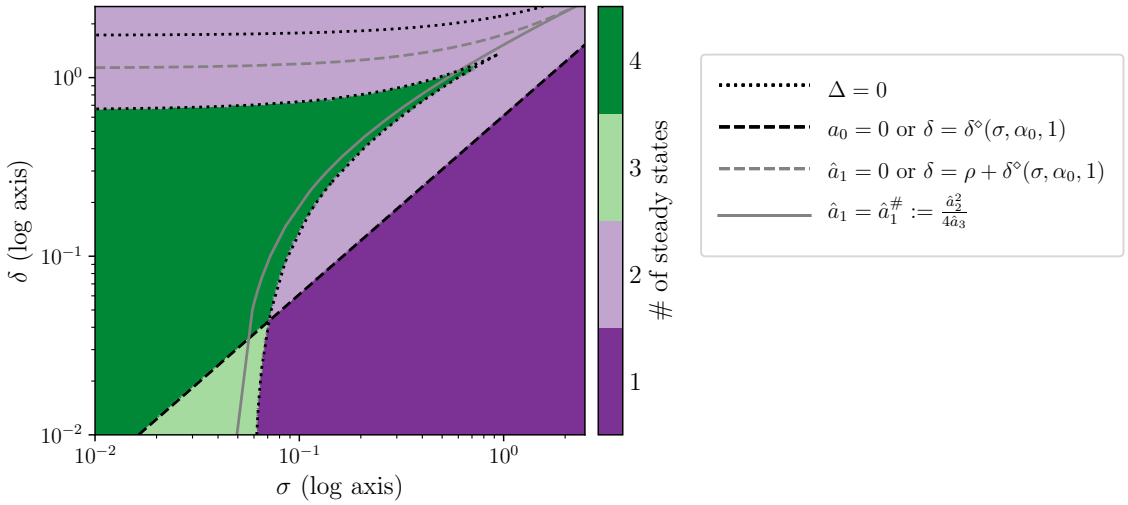


Figure 3.1: The number of physically realistic steady states in system (2.18) changes as model parameters vary. Accordingly, the (σ, δ) parameter space is decomposed into distinct regions which are colour-coded in line with the number of physically realistic steady states that exist. The plot is generated by numerically computing the number of physically realistic steady states of system (2.18) at different values of σ and δ that lie on a triangular mesh of the (σ, δ) plane, with other parameters fixed at their default values (see Table 2.2). We also plot analytically obtained region boundaries (see Section 3.2.2 and Appendix A). The black curves, $\Delta = 0$ and $a_0 = 0$, trace region borders well, while the grey curves, $\hat{a}_1 = 0$ and $\hat{a}_1 = \hat{a}_1^\#$, capture them only qualitatively.

varies in this parameter space. The colour plot is computed numerically by sweeping through points in the (σ, δ) parameter space using the function `DensityPlot` in Mathematica [181]. The line $a_0 = 0$ accurately captures the lower border of the region, in which the intermediate-sized-tumour steady state is physically realistic, while the curve $\hat{a}_1 = 0$ only qualitatively captures the upper border of this region. The border $\hat{a}_1 = \hat{a}_1^{\#}$, where the asymptotic large-tumour steady states (e_3^*, m_3^*) and (e_4^*, m_4^*) collide and become complex conjugates, shows similar inaccuracies (see Fig. 3.1). These discrepancies occur because the asymptotic approximations of the nonzero-tumour steady states, given in Section 3.2.1, lose accuracy in and close to parameter regimes where the roots of the cubic (3.3a) collide, e.g. as the parameters approach the curve $\hat{a}_1 = 0$.

To avoid these difficulties, we study the multiplicity of roots of the original cubic (3.3a). We identify parameter sets for which multiplicity of the cubic roots changes by considering the discriminant,

$$\Delta := a_2^2 a_1^2 - 4a_3 a_1^3 - 4a_2^3 a_0 - 27a_3^2 a_0^2 + 18a_3 a_2 a_1 a_0, \tag{3.8}$$

of the cubic (3.3a). Using Mathematica we can numerically compute curves in (σ, δ) parameter space on which $\Delta = 0$. These curves separate regions which have different numbers of physically realistic steady state solutions (see Fig. 3.1). At the lower boundary of the dark green region, roots m_3^* and m_4^* collide and become complex, while roots m_2^* and m_4^* collide at the upper boundary of the dark green region. On the black dotted curve in the light purple region, the complex roots m_2^* and m_4^* collide again and become real, but since they are real and negative (unrealistic), the steady state multiplicity is unchanged. Curves on which steady states coalesce indicate parameter regimes where saddle-node-type bifurcations occur, i.e. when there is a collision of two physically realistic steady states, one stable, and the other a saddle; we demonstrate this in Sections 3.3 and 3.4.

3.2.3 Linear stability of steady states

To demonstrate the system exhibits immunoediting behaviour via the three types of steady states identified in Section 3.2.1, our goal in this section is to identify parameter regions in which physically realistic steady states are stable.

We determine steady state stability via linear stability analysis, complemented with numerics. As in Section 3.2.1, we assume that parameters μ and β scale with the small parameter $0 < \xi \ll 1$. We introduce a second small parameter $0 < \nu \ll \xi$, and seek solutions of the form

$$e(\tau) = e^* + \nu e_1(\tau) + \mathcal{O}(\nu^2), \qquad m(\tau) = m^* + \nu m_1(\tau) + \mathcal{O}(\nu^2), \tag{3.9}$$

where (e^*, m^*) are steady states. By substituting (3.9) into equations (2.18), equating terms of $\mathcal{O}(\nu)$ to zero and neglecting terms of $\mathcal{O}(\xi)$ or higher orders of ν, we obtain a linearised system $\dot{\boldsymbol{x}} = \mathcal{J}(e^*, m^*)\boldsymbol{x}$ for $\boldsymbol{x} = (e_1, m_1)$, with Jacobian matrix

$$\mathcal{J}(e^*, m^*) = \underbrace{\begin{bmatrix} 1 + \kappa_1 m^* & \kappa_1 e^* \\ \kappa_2 m^* & 1 + \kappa_2 e^* \end{bmatrix}}_{=:\mathcal{B}(e^*, m^*)}^{-1} \underbrace{\begin{bmatrix} \frac{\rho m^*}{\eta + m^*} - \delta & \frac{\rho \eta e^*}{(\eta + m^*)^2} \\ -\kappa_2 m^* & \alpha - \kappa_2 e^* \end{bmatrix}}_{=:\mathcal{A}(e^*, m^*)} \tag{3.10}$$

where (e^*, m^*) is a leading order steady state approximation from Section 3.2.1. The eigenvalues λ_1 and λ_2 of $\mathcal{J}(e^*, m^*)$ predict exponentially decaying behaviour of the linearised system, and therefore linear stability of the steady state (e^*, m^*) whenever $\max_{i=1,2} \mathrm{Re}(\lambda_i) < 0$.

For the tumour-free steady state $(e_1^*, m_1^*) = (\frac{\sigma}{\delta}, 0)$

$$\lambda_1 = -\delta \qquad \text{and} \qquad \lambda_2 = \alpha - \frac{\sigma \kappa_2}{\delta}, \tag{3.11}$$

so this steady state is linearly stable when the parameters satisfy

$$\delta < \frac{\sigma \kappa_2}{\alpha} = \delta^\diamond(\sigma, \alpha, \kappa_2), \tag{3.12}$$

This means that the ratio of the basal effector influx rate, σ, and effector death rate, δ, must be larger than the net tumour growth rate, α, for the tumour-free steady state to be stable, i.e. for tumour eradication.

Considering (3.7) and (3.12), we observe that the region in which the intermediate-sized-tumour steady state (e_2^*, m_2^*) is physically realistic is separated from the region in which the tumour-free steady state is stable by the line $\delta = \delta^\diamond(\sigma, \alpha, \kappa_2)$. This implies that the tumour-free and the intermediate-sized-tumour steady states may undergo a transcritical bifurcation when $\delta = \delta^\diamond(\sigma, \alpha, \kappa_2)$, provided the intermediate-sized-tumour steady state is stable once it has emerged as physically realistic.

When characterising the local stability of the intermediate-sized-tumour steady state, we consider the trace and the determinant of the Jacobian evaluated at this steady state. We note that

$$\det(\mathcal{J}(e^*, m^*)) = \frac{\det(\mathcal{A}(e^*, m^*))}{\det(\mathcal{B}(e^*, m^*))}, \quad \text{and} \quad \operatorname{tr}(\mathcal{J}(e^*, m^*)) = \frac{\operatorname{tr}(\tilde{\mathcal{B}}(e^*, m^*)\mathcal{A}(e^*, m^*))}{\det(\mathcal{B}(e^*, m^*))}, \tag{3.13}$$

where $\tilde{\mathcal{B}}(e^*, m^*) = \det(\mathcal{B}(e^*, m^*))(\mathcal{B}(e^*, m^*))^{-1}$, and $\det(\mathcal{B}(e^*, m^*)) = 1 + \kappa_2 e^* + \kappa_1 m^* > 0$ for any nonnegative steady state. Therefore, the signs of $\det(\mathcal{J}(e^*, m^*))$ and $\operatorname{tr}(\mathcal{J}(e^*, m^*))$ are determined by the signs of $\det(\mathcal{A}(e^*, m^*))$ and $\operatorname{tr}(\tilde{\mathcal{B}}(e^*, m^*)$ $\mathcal{A}(e^*, m^*))$ respectively. When $(e^*, m^*) = (e_2^*, m_2^*)$ from (3.4), we have

$$\det(\mathcal{A}(e_2^*, m_2^*)) = \frac{\alpha}{\eta\rho}\left(\frac{\sigma\kappa_2}{\alpha} - \delta + \rho\right)^2 m_2^* \tag{3.14}$$

which is positive when m_2^* is physically realistic. Stability of (e_2^*, m_2^*) thus reduces to satisfying the trace condition

$$\operatorname{tr}(\tilde{\mathcal{B}}(e_2^*, m_2^*)\mathcal{A}(e_2^*, m_2^*)) = -\frac{\alpha\eta\left(\frac{\sigma\kappa_2}{\alpha} - \delta\right)}{\hat{a}_1}\left(\kappa_1 - \frac{1}{\eta\rho}\hat{a}_1^2\right) - (\alpha + 1)\frac{\sigma\kappa_2}{\alpha} < 0. \tag{3.15}$$

We conclude that in regions where $\delta > \delta^\diamond(\sigma, \alpha, \kappa_2)$ and inequality (3.15) are simultaneously satisfied, the intermediate-sized-tumour steady state is stable.

When $\delta = \delta^\diamond(\sigma, \alpha, \kappa_2)$,

$$\operatorname{tr}(\tilde{\mathcal{B}}(e_2^*, m_2^*)\mathcal{A}(e_2^*, m_2^*)) = -(\alpha + 1)\frac{\sigma\kappa_2}{\alpha} < 0, \tag{3.16}$$

so our analysis predicts that the tumour-free and intermediate-sized-tumour steady states undergo a transcritical bifurcation on this curve. Moreover, when $\operatorname{tr}(\tilde{\mathcal{B}}(e_2^*, m_2^*)$

$\mathcal{A}(e_2^*, m_2^*)) = 0$, the Jacobian eigenvalues are purely imaginary, and the intermediate-sized-tumour steady state is predicted to undergo a Hopf bifurcation. We numerically validate these predictions in Section 3.3, and note that further work is required to demonstrate existence of the Hopf bifurcation.

Since exact analytical expressions for the roots of the cubic (3.3a) are not practical to work with, we have characterised the stability of nonzero-tumour steady states using their asymptotic approximations (3.4) and (3.5), as shown above for the intermediate-sized-tumour steady state. Similar analysis for the large-tumour steady states is difficult due to the complexity of the asymptotic expressions (3.5) (see Appendix A). Therefore, in Section 3.3, we numerically identify parameter regimes in which these steady states are stable.

3.3 Bifurcation structure for default parameter values as the basal effector supply rate varies

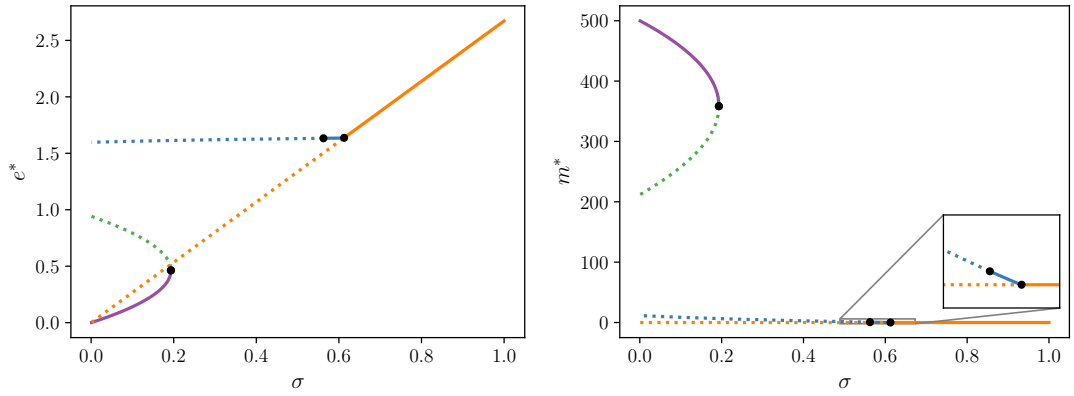
Our aims here are to use numerical methods to validate analytical results from Section 3.2, and to detect global bifurcations not evident from the linear stability analysis. We focus on σ, the nondimensional constant supply rate of effector cells, as the key bifurcation parameter. A larger value of σ corresponds to an immune system that is more effective at recruiting effector cells into the tumour microenvironment, and the recruited cells better infiltrate the tumour. In practice, σ may vary between patients since those properties of the immune system on which this parameter depends, such as the specificity of the repertoire of effector cells and the degree of their infiltration of the tumour, exhibit high inter-patient variability [53, 145]. The parameter σ may also vary in response to immunotherapies, such as vaccination and adoptive T cell therapy, both of which provide an external boost to the immune system [52]. These two therapies enhance the population of tumour antigen specific effector T cells, by increasing exposure to tumour antigens away

from the tumour microenvironment described by our system (e.g. in lymph nodes or *ex vivo*).

The system's bifurcation structure is depicted with bifurcation diagrams showing how (scaled) steady state tumour and effector cell numbers vary with σ (see Figure 3.2). We generate these diagrams numerically, by considering discrete values of σ that lie in the interval $[0, 1]$, while fixing other parameters at their default values in Table 2.2. For each value of σ, we calculate the steady states of system (2.18), and their corresponding eigenvalues and eigenvectors to characterise local stability. Moreover, by perturbing a steady state along the direction of an eigenvector associated with a positive/negative eigenvalue, and using the perturbed point as an initial state for forward/backward integration of equations (2.18), we also generate unstable/stable manifolds of the computed steady states. We plot the steady states and manifolds in (w_1, w_2) phase portraits at different values of σ (see Figs. 3.4 and 3.5), and observe that in some regions of parameter space the phase portraits exhibit limit cycles. By tracing where limit cycles emerge and where steady states change multiplicity or stability as σ varies, we identify five regions with qualitatively different model behaviour, separated by four bifurcation points σ^{SH}, σ^{Hopf}, σ^{SO} and σ^{T} (see schematic in Fig. 3.3). The behaviours in the five regions (A)–(E) are described below:

(A) $0 < \sigma < \sigma^{\mathrm{SH}}$: The system is monostable; for all positive initial conditions, trajectories evolve to the larger of the large-tumour steady states (*tumour escape*). There are also three unstable steady states; the tumour-free and the smaller of the large-tumour steady states are saddles, linked via a heteroclinic connection (see Figs. 3.4A and 3.5A).

(AB) $\sigma = \sigma^{\mathrm{SH}}$: The large-tumour steady states collide at a *saddle-node homoclinic bifurcation* (or infinite period bifurcation) [117]; the saddle and node annihilate each other, and the tumour-free steady state admits a large-amplitude, stable homoclinic orbit.



(a) number of free effector cells

(b) number of free tumour cells

Figure 3.2: The long-timescale system (2.18) (or equivalently (2.17)) exhibits the three Es of immunoediting through stable steady states of varying tumour cell densities. We can see this in the bifurcation diagrams, which are obtained numerically by sweeping through the σ parameter space, and computing the system steady states and their stability at discrete values of σ between 0 and 1. Other parameters are fixed at their default values (see Table 2.2). Plots (a) and (b) respectively show the steady state numbers of free effector and tumour cells. Purple and green curves denote large-tumour steady states, blue curves denote intermediate-sized-tumour steady states, and orange curves denote tumour-free steady states. Local stability of a steady state is distinguished by linestyle, whereby solid curves denote stable solutions, while dashed curves denote unstable ones. Black points mark bifurcations of steady state solutions, i.e. where solution branches change stability or collide.

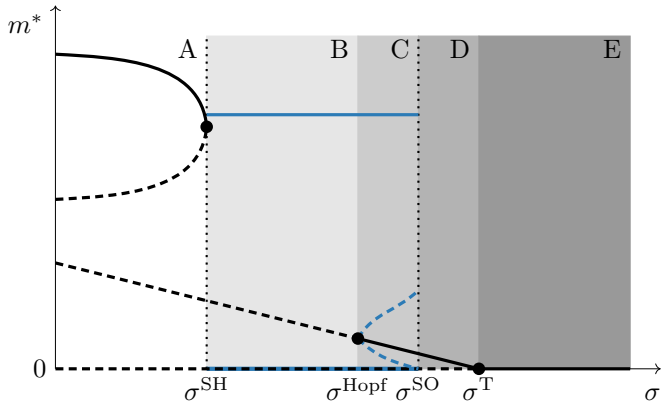
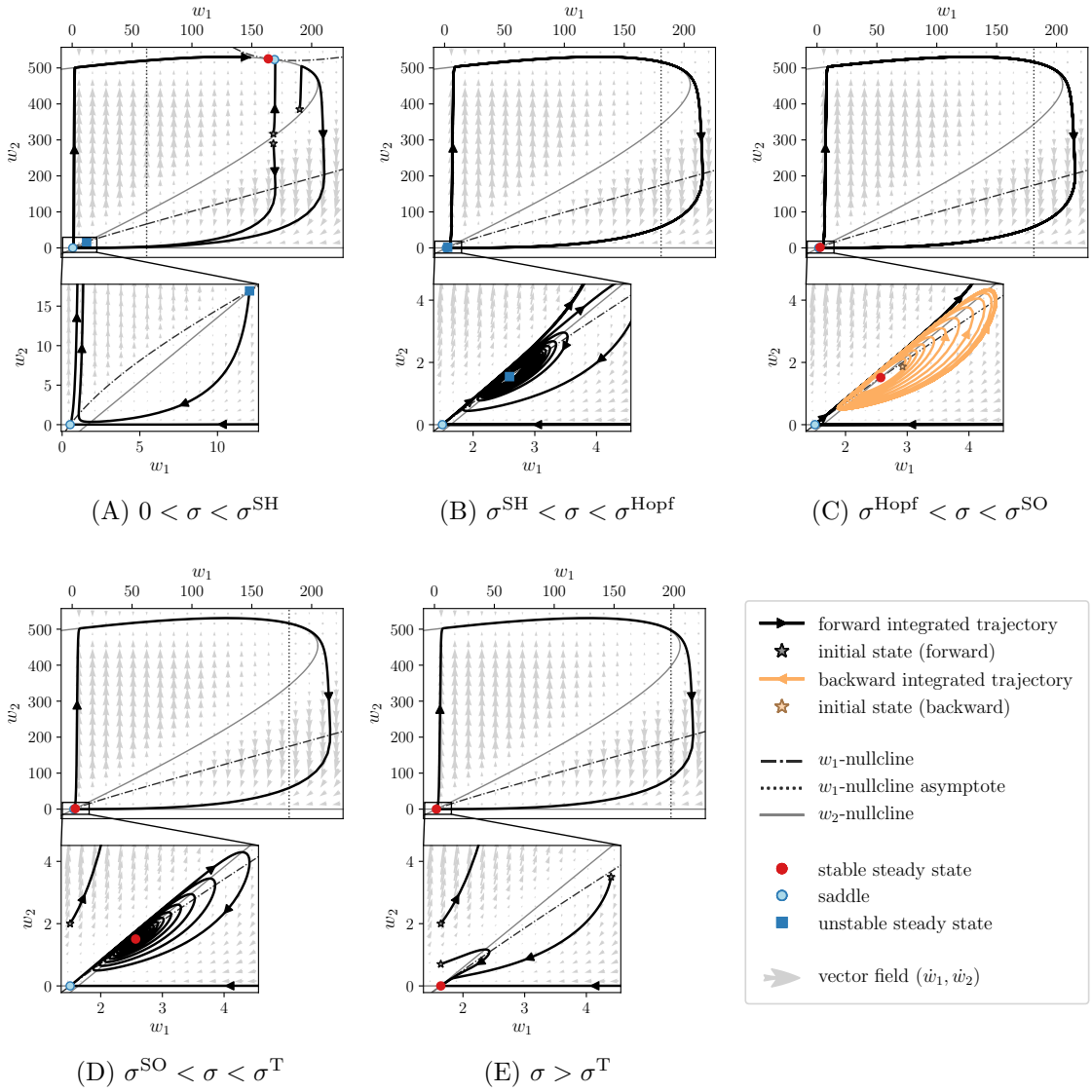


Figure 3.3: Schematic (not to scale) of the bifurcation diagram for scaled tumour cell numbers in Fig. 3.2b, indicating how periodic solutions arise in the long-timescale system (2.17) with respect to its steady states, as the parameter σ is varied. Black curves indicate tumour cell numbers of steady state solutions, whereas blue curves indicate maximum and minimum tumour cell numbers of periodic solutions. Local stability of solutions is distinguished by linestyle, whereby solid curves denote stable solutions, while dashed curves denote unstable ones. As σ is increased, the system undergoes the following bifurcations: SH = saddle-node homoclinic bifurcation, Hopf = subcritical Hopf bifurcation, SO = saddle-node of orbits, T = transcritical bifurcation. Bifurcations of steady state solutions are denoted with black dots, while bifurcations of periodic solutions are denoted with dotted vertical lines. These bifurcations split the parameter space into regions of qualitatively different model behaviours with respect to its steady states and orbits; the different regions are shaded with grey of varying intensities.



(A) $0 < \sigma < \sigma^{\mathrm{SH}}$ (B) $\sigma^{\mathrm{SH}} < \sigma < \sigma^{\mathrm{Hopf}}$ (C) $\sigma^{\mathrm{Hopf}} < \sigma < \sigma^{\mathrm{SO}}$

(D) $\sigma^{\mathrm{SO}} < \sigma < \sigma^{\mathrm{T}}$ (E) $\sigma > \sigma^{\mathrm{T}}$

Figure 3.4: As σ is varied in Regions (A)–(E) of the parameter space (see the schematic in Fig. 3.3), phase portrait plots (A)–(E) of the long-timescale model (2.17) in (w_1, w_2) plane reveal that the model exhibits rich bifurcating behaviours, which include emergence of orbits and large excursions in phase space. The values of σ for which the phase portraits are generated lie within intervals indicated in respective plot captions, while other parameters are fixed at their default values (see Table 2.2). The lower panel plots show model behaviour in the vicinity of the tumour-free steady state by magnifying this area of phase space. Solid black curves show forward integrated trajectories (stable), whereas solid orange curves denote trajectories integrated backward in time (unstable) in order to uncover the dynamics close to unstable orbits. For schematics of these phase portrait plots, see Fig. 3.5.

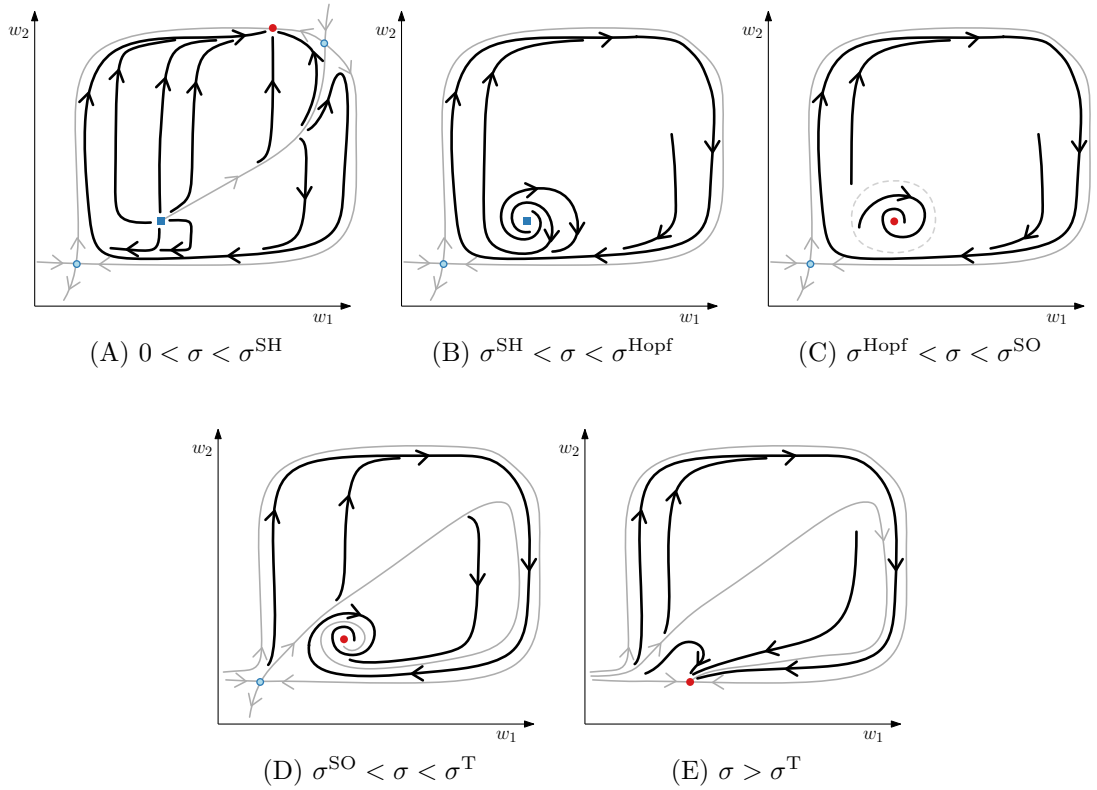


Figure 3.5: Schematic diagrams (not to scale) of phase portraits in Fig. 3.4 as σ varies. The vertical axes represent total tumour cell numbers, w_1, while the horizontal axes represent total effector cell numbers, w_2. Solid grey curves denote stable and unstable manifolds, solid black curves show example trajectories, and dashed grey circles represent unstable orbits. As in Fig. 3.4, stable, unstable and saddle nodes are respectively marked with red discs, blue squares and light blue discs with an outline of a darker shade of blue.

(B) $\sigma^{\mathrm{SH}} < \sigma < \sigma^{\mathrm{Hopf}}$: For all positive initial conditions trajectories evolve to the stable homoclinic orbit associated with the tumour-free steady state, via *oscillations with large amplitudes* (see Figs. 3.4B and 3.5B).

(BC) $\sigma = \sigma^{\mathrm{Hopf}}$: There is a *subcritical Hopf bifurcation*, at which the intermediate-sized-tumour steady state becomes stable and is surrounded by an unstable, small-amplitude limit cycle that emerges from it.

(C) $\sigma^{\mathrm{Hopf}} < \sigma < \sigma^{\mathrm{SO}}$: The system exhibits *bistability* between the intermediate-sized-tumour steady state and the homoclinic orbit; trajectories starting within the region of the phase space mapped out by the unstable, small-amplitude limit cycle, oscillate towards the intermediate-sized-tumour steady state (*tumour dormancy*); otherwise the system evolves towards the homo-

clinic orbit via *large-amplitude oscillations* (see Figs. 3.4C and 3.5C).

(CD) $\sigma = \sigma^{\mathrm{SO}}$: There is a *saddle-node of limit cycles* [117], at which the stable homoclinic and the unstable Hopf orbits collide, and then annihilate each other.

(D) $\sigma^{\mathrm{SO}} < \sigma < \sigma^{\mathrm{T}}$: The system is monostable; trajectories evolve towards the stable intermediate-sized-tumour steady state (*tumour dormancy*) via damped oscillations, possibly preceded by a large excursion (see Figs. 3.4D and 3.5D).

(DE) $\sigma = \sigma^{\mathrm{T}}$: There is a *transcritical bifurcation*, at which the unstable tumour-free and the stable intermediate-sized-tumour steady state exchange stability, and the latter steady state becomes physically unrealistic (negative).

(E) $\sigma > \sigma^{\mathrm{T}}$: The system is monostable; all trajectories evolve to the tumour-free steady state (*tumour elimination*). (See Figs. 3.4E and 3.5E)

The local bifurcations established numerically agree with the analytical predictions from Section 3.2. Further work is needed to prove the existence of the identified global bifurcations, and the local subcritical Hopf bifurcation (e.g. weakly nonlinear analysis); such analyses are beyond the scope of this thesis.

We have shown how the system dynamics change as the basal supply rate of effector cells, σ, and thereby the strength of the immune system, is increased. When σ is small (in Region (A)), effector cells at sufficient numbers infiltrate tumours and decrease their volume, but the response cannot be sustained; tumour escape eventually occurs for all tumours. In Region (B) effector cells infiltrate all tumours, and markedly decrease their size. As this happens, effector cell numbers also decline, enabling the immune-suppressed tumour to relapse and rapidly grow towards its carrying capacity until the immune system is stimulated again. In this way, the cycle repeats, with tumour and effector cells oscillating in an excitable manner as they traverse phase space in the vicinity of the large-amplitude homoclinic orbit, passing through regions with very small and large tumour mass. Increasing σ further (Region (C)) leads to the emergence of a small region of phase space

within the homoclinic loop, in which the immune system controls the tumour size – tumour equilibrium or dormancy. In Region (C) there is a fine balance between tumour and effector cell numbers, and all solutions starting within it oscillate towards the dormant steady state as the amplitude of oscillations decreases. In order to enter this dormant region, the tumour must be suppressed in a specific way, since otherwise, the system is attracted to the homoclinic loop of escape and elimination, and we say that the tumour is "sneaking through". As σ increases, the basin of attraction of the small tumour steady state increases so that in Region (D), the immune response is strong enough to drive all tumours to a dormant state, albeit after a large phase-plane excursion (see Fig. 3.4D). Ideally, we would seek to increase σ (e.g. via therapy) so that the system enters Region (E), where the model predicts all tumours will be eradicated by the immune system.

Taken together, our results suggest that, as model parameters vary, the system exhibits the three Es of immunoediting – elimination, equilibrium and escape. Our analysis reveals that a therapy which increases the basal effector supply rate (e.g. adoptive cell therapy or cancer vaccine) could push the system from tumour escape to elimination or equilibrium at the intermediate-sized-tumour steady state. Such changes in system behaviour occur via a series of local bifurcations (saddle-node, Hopf and transcritical).

The system also admits stable oscillatory solutions with large phase-plane excursions, arising from its intrinsic excitability as shown in Section 2.4. We have demonstrated numerically that these solutions emerge and disappear via global bifurcations of a homoclinic orbit, a feature that is common in excitable systems, but cannot be detected via local analysis. The excitable homoclinic orbit is seen as the main driver of excitable dynamics in all regions, as large excursions are observed not only in regions where the orbit exists, but also in other regions, where the associated (disconnected) steady state manifolds retain the approximate shape of the orbit.

3.4 Changes in the bifurcation structure as default parameters vary

We now consider how variation of other model parameters affects the bifurcation structure. We focus on the parameters μ, ρ and η, which together determine the extent of the immune response to the tumour. This is because they may be altered by immunotherapies that impact tumour-immune interactions directly, or reverse the effects of tumour-modulated immunosuppression; for example, immune checkpoint therapies that inhibit the PD-1/PD-L1 axis may, in a tumour-dependent manner, decrease T cell apoptosis, exhaustion and anergy, while also increase T cell proliferation [68]. At the same time, we investigate the efficacy of such treatments across populations of patients whose immune systems (in the absence of a tumour) can vary in strength; a patient might have a compromised immune system (corresponding to a small σ or a large effector death rate δ); they may have been vaccinated or undergone adoptive cell therapy (large σ).

3.4.1 Varying σ and δ

We first study the bifurcation structure when σ and δ co-vary in a neighbourhood of the default parameter regime. This provides a reference for benchmarking model outcomes associated with variation of parameters linked to immunotherapy. We numerically track the number of physically realistic steady states, and their (local) stability, in order to partition the (σ, δ) plane (see Figure 3.6); transitions from one region to another indicate qualitative changes in the system dynamics. Using analytical results from Section 3.2, we also compute curves on which the existence and stability of steady state solutions change. In the (σ, δ) plane, these curves are in good agreement with the numerically determined region boundaries. This confirms our analysis and increases our insight into the behaviour of our model. We identify seven distinct regions listed in Table 3.1. Regions (1)–(4) exhibit

behaviours equivalent to those in Regions (A)–(E), as related in Table 3.1, and described in both Section 3.3 and Fig. 3.3. Behaviours in Regions (5)–(7) are described as follows:

(5) The system is monostable; for all positive initial conditions trajectories evolve to the large-tumour steady state (*tumour escape*). The only other steady state is the tumour-free one, which is a saddle.

(6) The system exhibits bistability between the tumour-free steady state (*tumour elimination*) and the larger of the large-tumour steady states (*tumour escape*); depending on initial conditions trajectories evolve to one of these two stable steady states. The smaller of the large-tumour steady states is a saddle.

(7) The system exhibits bistability between the intermediate-sized-tumour steady state (*tumour equilibrium*) and the larger of the large-tumour steady states (*tumour escape*). The smaller of the large-tumour steady states and the tumour-free steady state are both saddles.

Region No.	**Denotation in (σ, δ) parameter space of Fig. 3.6 and the left panels of Figs. 3.7, 3.8 and B.1**	**Description in Sec. 3.3 and phase portrait in Figs. 3.4 and 3.5**	**Qualitative behaviour**
(1)	yellow with diagonal lines	(E)	elimination
(2)	yellow with cross-hatching	joint (C) and (D)	equilibrium (or oscillations – bistable)*
(3)	orange	(B)	oscillations
(4)	yellow below upper curve $\Delta = 0$	(A)	escape
(5)	yellow above upper curve $\Delta = 0$	**	escape
(6)	blue with diagonal lines	**	escape or elimination – bistable
(7)	blue with cross-hatching	**	escape or equilibrium – bistable

Table 3.1: Distinct regions of (σ, δ) parameter space, as denoted with different colours and hatching (see the second column) in Fig. 3.6 and in the left-panel plots of Figures 3.7, 3.8 and B.1, exhibit different qualitative behaviours that range from elimination to escape (see the fourth column). The third column indicates to which regions from the bifurcation schematic in Fig. 3.3, where only σ varies, some regions of the (σ, δ) parameter space in Fig. 3.6 correspond. *Region (2) exhibits two distinct qualitative behaviours, and should be split into two disjoint parameter subregions depending on whether the stable homoclinic orbit with large amplitude and period exists. In the parameter subregion where this orbit exists, trajectories may evolve to an oscillatory solution in the vicinity of the orbit. **For a description of model behaviour in Regions (5)–(7) see Section 3.4.1. Phase portraits for these regions are omitted for brevity.

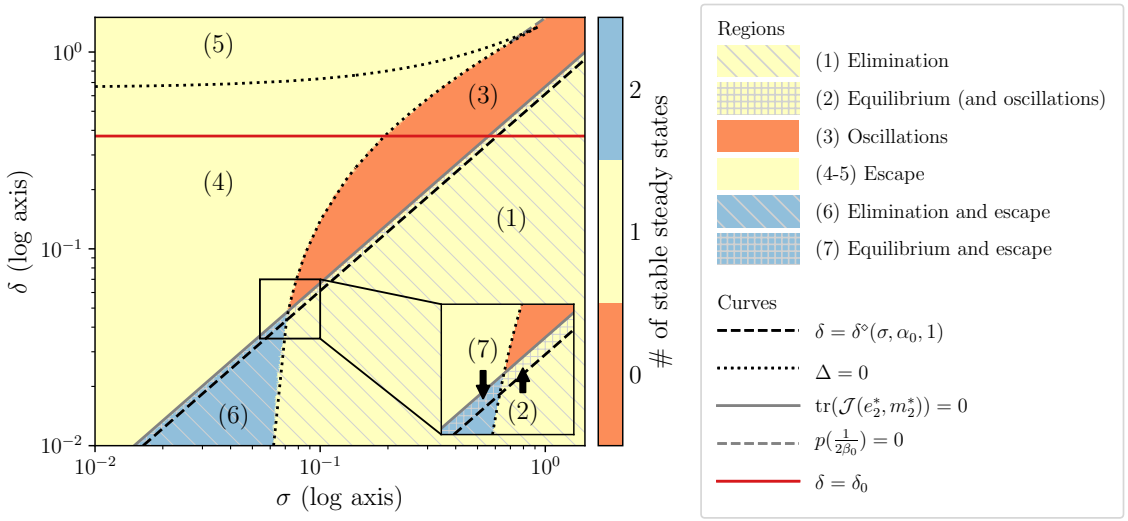


Figure 3.6: The long-timescale model (2.17) exhibits qualitatively different behaviours as parameters σ and δ are varied, and the model undergoes bifurcations. The plot shows how the (σ, δ) parameter space is split into distinct regions depending on the existence and stability of the model's steady states. Regions with different numbers of stable steady states are distinguished using a colour map. Parameters other than σ and δ are fixed at their default values (see Table 2.2). Boundaries of distinct regions are traced by different curves on which bifurcations may occur. On the dotted black curves $\Delta = 0$, and the multiplicity of steady states changes (see Section 3.2.2). Dashed black lines indicate where the local stability of the tumour-free steady state changes and the intermediate-sized-tumour steady state becomes physically unrealistic, i.e. when $\delta = \delta^\diamond(\sigma, \alpha, \kappa_2)$. Solid grey lines indicate where the local stability of the latter steady state changes, i.e. when $\mathrm{tr}(\mathcal{J}(e_2^*, m_2^*)) = 0$ (see Section 3.2.3). Dashed grey curves, for example by the transition between Regions (3) and (5) for large δ, indicate where a large-tumour steady state of $\mathcal{O}(\xi^{-1})$, with $0 < \xi \ll 1$ (as in Section 3.2.1), changes stability; in analysis not presented in this thesis due to limited space, we derive that this is given by the curve $p((2\beta)^{-1}) = 0$, where p is the cubic (3.3a). Hatching with diagonal lines and grey cross-hatching (between solid grey and dashed black lines, see the magnified area in the plot) respectively denote regions in which the tumour-free and intermediate-sized-tumour steady states are locally stable. The red horizontal line marks the default value of δ, $\delta_0 = 0.374$ (see Figs. 3.2 and 3.3 for bifurcation diagrams at this value of δ).

In Fig. 3.6 we see that varying δ typically has a similar effect to varying σ. The model predicts that therapies that increase σ and/or decrease δ can drive the system to regions of parameter space where tumour elimination or dormancy occur for all initial conditions (Region (1) and a subregion of Region (2) respectively).

We now investigate how varying the tumour-immune interactions parameters, μ, ρ and η, affects model behaviour in the (σ, δ) parameter space.

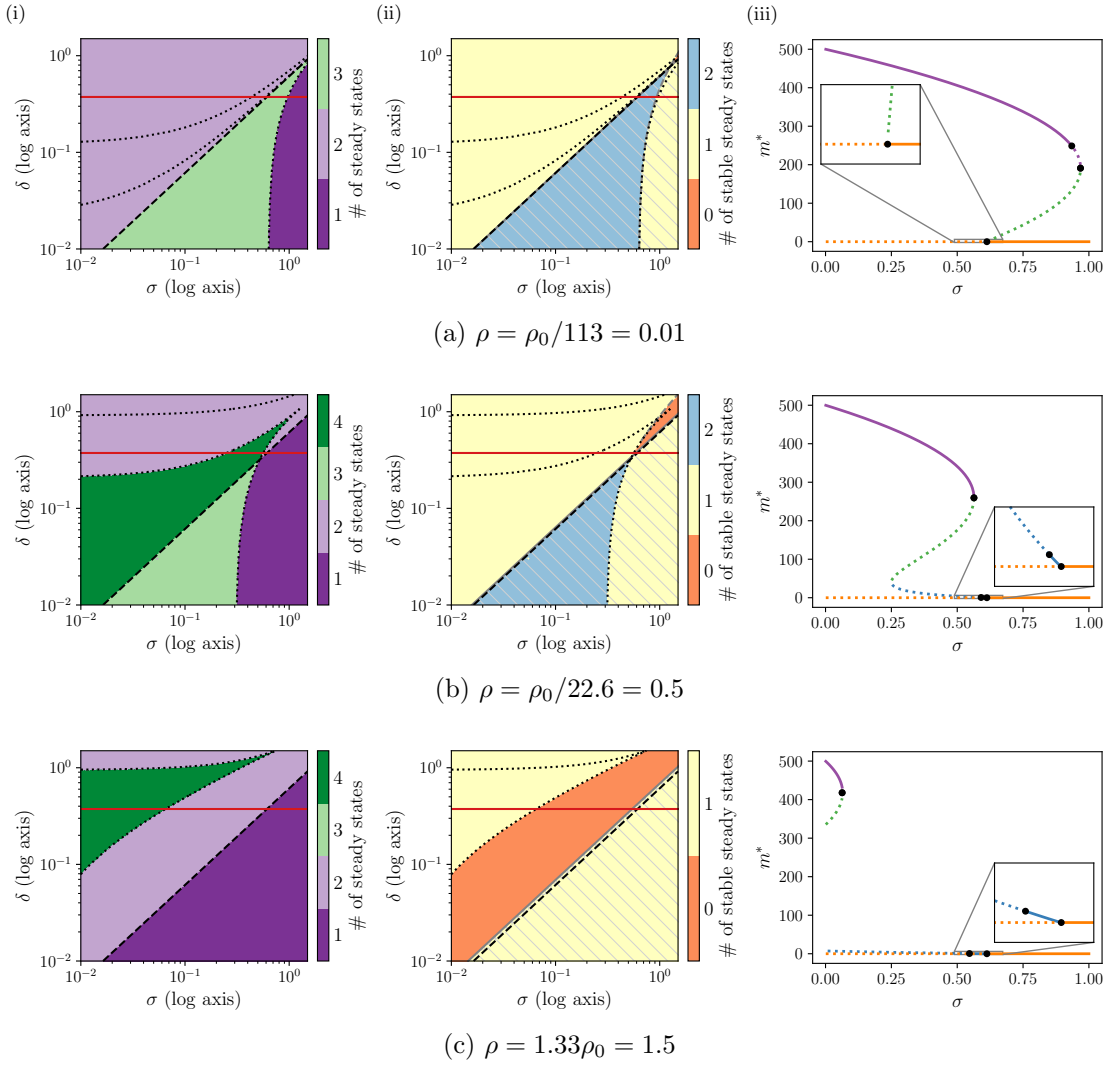
3.4.2 Varying ρ

As ρ, the tumour-stimulated supply rate of effector cells, decreases, we observe an increase in the region in which a locally stable large-tumour steady state exists (see panel (iii) of Fig. 3.7); the tumour cell numbers of this steady state decrease as σ increases, and it collides with the unstable large-tumour steady state (saddle point) at a larger value of σ. The distinction between the nonzero-tumour steady states, as described in Section 3.2.1, is less pronounced for smaller values of ρ, and eventually one nonzero-tumour steady state is lost (e.g. at $\rho = 0.01$, see panels (i) and (iii) of Fig. 3.7). This is consistent with our earlier asymptotic analysis (see Section 3.2.2), which predicts that the region in which the intermediate-sized-tumour steady state is physically realistic (see equation (3.7)), shrinks as ρ is reduced.

When ρ is small, the presence of tumour cells does not stimulate effector cells and their proliferation sufficiently to eliminate or control the tumour, e.g. due to low tumour immunogenicity or immunosuppression. We see that, within the plotted (σ, δ) parameter region, decreasing ρ reduces the sizes of Regions (2) and (7) where the intermediate-sized-tumour steady state is mono- or bistable, and dormancy may occur (see panels (ii) and (iii) of Fig. 3.7). Since the size of the region of (σ, δ) parameter space with a locally stable large-tumour steady state (Regions (4)–(6)) increases, Region (1), exhibiting tumour elimination, contracts, while Region (6), with bistability of escape and elimination, expands.

Conversely, by increasing ρ, effector cells increase in numbers at a greater rate in response to the tumour, and eradicate it more effectively. Therefore, as ρ increases, there is a reduction in the sizes of Regions (4)–(6) where tumour escape may occur. Region (1), with tumour elimination, expands, but only up to its maximum size below the line $\delta = \delta^{\diamond}(\sigma, \alpha, \kappa_2)$ (see Section 3.2.1). Above this line, Region (3), with stable, large-amplitude, oscillatory solutions, is considerably magnified. While Regions (2) and (7), in which tumour dormancy may occur, also appear, these are

small, and large-amplitude oscillations or escape are the dominant behaviours (see panel (ii) of Fig. 3.7c).

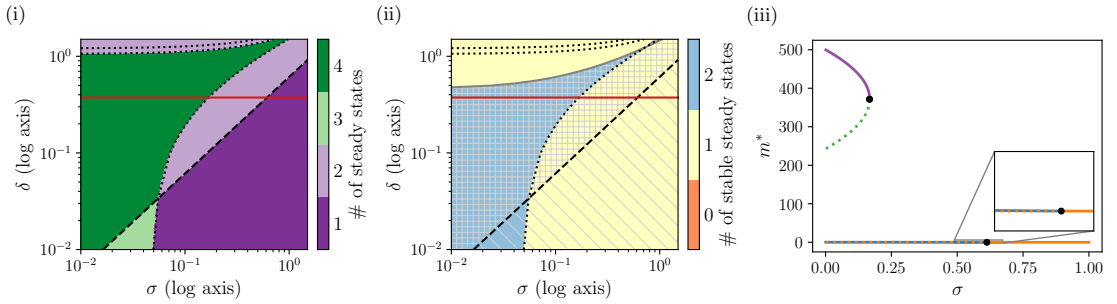


(a) $\rho = \rho_0/113 = 0.01$

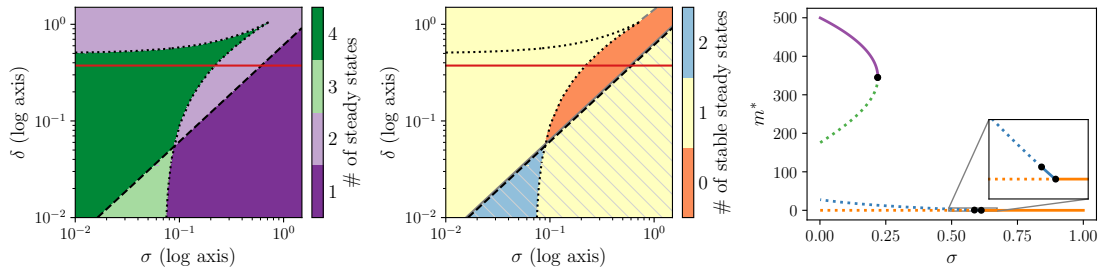
(b) $\rho = \rho_0/22.6 = 0.5$

(c) $\rho = 1.33\rho_0 = 1.5$

Figure 3.7: Series of plots (a)–(c) showing how the bifurcation structure of the long-timescale system (2.17) changes as ρ is varied from its default value $\rho_0 = 1.131$. As ρ is increased, Regions (4)–(6), in which the model may exhibit tumour escape, shrink due to which Region (1) with tumour elimination expands up to its maximum size below the black dashed line, $\delta = \delta^{\diamond}(\sigma, \alpha, \kappa_2)$. Above this line, while small regions in which tumour dormancy may occur (Regions (2) and (7)) emerge, the dominating region is Region (3) in which the model exhibits large-amplitude oscillations. For each value of ρ, we present three bifurcation diagrams: plots of (σ, δ) parameter space showing (i) regions with different numbers of physically realistic steady states (see Fig. 3.1 for legend and details), and (ii) regions with different numbers of stable steady states (see Fig. 3.6 for legend and details); (iii) bifurcation diagrams of tumour steady state numbers, m^*, as σ is varied and δ is fixed at its default value (see Fig. 3.2 for legend and details).



(a) $\eta = \eta_0/50 = 0.404$



(b) $\eta = 2\eta_0 = 40.38$

Figure 3.8: Series of plots (a) and (b) showing how the bifurcation structure of the long-timescale system (2.17) changes as η is varied from its default value $\eta_0 = 20.19$. As η is decreased, Regions (2) and (7), in which the dormant steady state of system (2.17) is locally stable, expand. For each value of η, we present three bifurcation diagrams: plots of (σ, δ) parameter space showing (i) regions with different numbers of physically realistic steady states (see Fig. 3.1 for legend and details), and (ii) regions with different numbers of stable steady states (see Fig. 3.6 for legend and details); (iii) bifurcation diagrams of tumour steady state numbers, m^*, as σ is varied and δ is fixed at its default value (see Fig. 3.2 for legend and details).

3.4.3 Varying μ

Increasing μ, the rate of effector cell inactivation, has a similar effect on model behaviour as decreasing ρ, but with greater sensitivity (see Appendix B). This is expected since the term $\frac{\rho e m}{\eta + m}$ in model equations (2.3) may saturate, whereas the related term $-\mu e m$ does not.

3.4.4 Varying η

Changes in parameter η, which determines the size of the tumour at which effector proliferation starts saturating, have a different effect on the model dynamics compared to changes in either ρ or μ. Varying η mainly affects the zero-trace/Hopf bifurcation curve (solid grey), where the intermediate-sized-tumour steady state changes

stability (see Fig. 3.8). Decreasing η therefore increases the size of Regions (2) and (7) that exhibit mono- and bistability of tumour dormancy, respectively, and reduces the size of Region (3) with stable oscillatory solutions. We conclude that a therapy which reduces η (i.e. increases the threshold at which limitations in immune response start to occur), may push the system to a region where tumour dormancy is possible; due to the prospect of region bistability this may hold only for a subset of the phase space.

3.4.5 Summary

We observe that in (σ, δ) parameter space the line $\delta = \delta^{\diamond}(\sigma, \alpha, \kappa_2)$, which is only affected by changes in the tumour growth rate, α, or the scaling parameter, κ_2, stratifies patients with parameters in (σ, δ) parameter space into those with strong immune responses (immunocompetent or vaccinated), whose tumours may be eliminated, and those with weaker responses (immunocompromised), whose tumours may at best be controlled. Assuming that the immune state of the patient, characterised by parameters σ and δ, is known, we note that changing parameters ρ and μ separately or simultaneously is beneficial for immunocompetent patients. For immunocompromised patients, manipulating η is more effective, as it is possible to achieve tumour dormancy. Ultimately, to achieve tumour elimination for immunocompromised patients, the basal function of the immune system must be improved, for example by targetting the parameters σ and/or δ.

3.5 Comparison with the QSSA model

In Chapter 2 we showed that Kuznetsov et al.'s QSSA model [88] does not exhibit excitable features. Here we discuss how the bifurcation structure of the long-timescale reduced model (2.17) (or equivalently (2.18)) differs from that of the QSSA model (2.23). While the bifurcation structure of the QSSA model has been

previously studied [88, 137], our aim here is to put it in the context of our analysis of the reduced model.

Steady states, their asymptotic approximations (3.4) and (3.5), and the conditions for local stability of the tumour-free steady state (3.12) are equivalent in model (2.18) and the QSSA model (2.23). We derive a condition, analogous to (3.15), for stability of the intermediate-sized-tumour steady state in the QSSA model. The Jacobian of the QSSA model, at leading order in ξ, is given by $\mathcal{J}(e^*, m^*) = \mathcal{A}(e^*, m^*)$, with $\mathcal{A}$ defined in (3.10). For $(e^*, m^*) = (e_2^*, m_2^*)$ as in (3.4), we obtain that

$$\det(\mathcal{A}(e_2^*, m_2^*)) > 0 \quad \text{as in (3.14), and} \quad \operatorname{tr}(\mathcal{A}(e_2^*, m_2^*)) = -\frac{\sigma\kappa_2}{\alpha} < 0, \qquad (3.17)$$

for physically realistic (e_2^*, m_2^*) (see Section 3.2.3). If the intermediate-sized-tumour steady state is physically realistic, then it is stable (see Fig. 3.9). This differs from model (2.18), in which the trace condition (3.15) is not trivially satisfied.

Parameter regions in which tumour elimination may occur correspond in both models (see yellow and blue regions with diagonal lines in Figs. 3.6 and 3.9). Due to differences in local stability of the intermediate-sized-tumour steady state, mono- and bistable regions with dormancy (yellow and blue, both with cross-hatching) are larger in the QSSA model than in model (2.18); tumour equilibrium is therefore more plausible in the QSSA model. Simulations of the QSSA model in Fig. 2.5 and in [88] show larger basins of attraction of the dormant steady state than in model (2.18). The homoclinic orbit that arises in model (2.18) (see Figs. 3.4B and 3.4C) does not exist in the QSSA model; this is shown by a straightforward application of the Dulac-Bendixson criterion in [88].

These findings have implications for cancer therapeutics. The QSSA model predicts that tumour elimination or dormancy in patients with different levels of immunocompetency may be easier to achieve, as more of the parameter space and the phase space exhibit the desired behaviour than in our reduced model, and

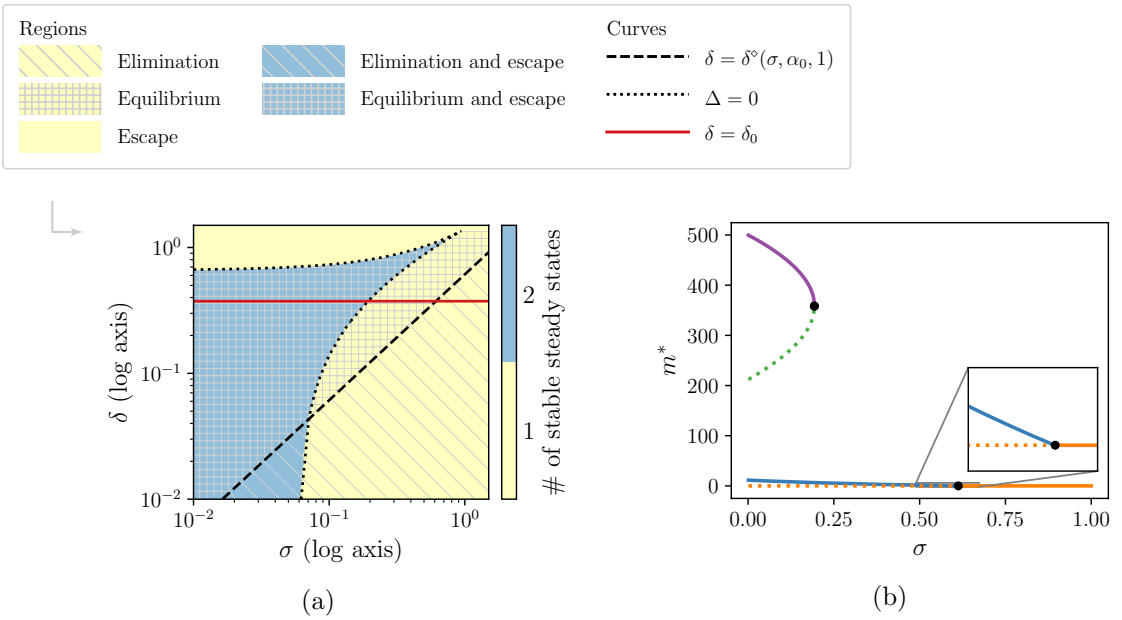


Figure 3.9: Bifurcation structure of the QSSA model (2.23) differs from that of the long-timescale model (2.18) as σ and δ are varied in the vicinity of the default parameter regime (see Table 2.2); the key difference is that the intermediate-sized-tumour steady state is stable when physically realistic. Plot (a) illustrates this by showing how the (σ, δ) parameter space is split into distinct regions depending on existence and stability of steady states in the QSSA model. Plot (b) is a bifurcation diagram of steady state tumour cell numbers in the QSSA model as σ is varied, and δ is fixed at its default value, $\delta_0 = 0.374$. Plots for the long-timescale model (2.17) that are comparable to (a) and (b) are Figs. 3.2 and 3.6 respectively; see these figures for details on how (a) and (b) are generated and legend.

make these regions more accessible. Moreover, the absence of the large-amplitude homoclinic orbit in the QSSA model suggests long periods of tumour remission are less dominant in the explored parameter space; however, there exist regimes showing temporary remission in the QSSA model [88].

3.6 Discussion and conclusions

In this chapter, we extended our analysis of the long-timescale approximation of Kuznetsov et al.'s original model [88] of tumour-immune interactions to better understand its predicted long-term dynamics. We systematically characterised the sensitivity of qualitative model behaviour to changes in parameters associated with the action of the immune system. We revealed that excitability can give rise to complex dynamics not reported in analyses of alternative simplifications of

Kuznetsov et al.'s model.

Like Kuznetsov et al.'s QSSA model, our reduced model exhibits the three Es of immunoediting – elimination, equilibrium, and escape [47]. We investigated how the number and nature of steady state and limit cycle solutions change as we vary σ, the basal rate at which effector cells are introduced to the tumour microenvironment, and δ, their death rate, in order to identify parameter regions in which each stage of immunoediting may arise via a (locally) stable steady state. For large σ and small δ (immunocompetent or vaccinated patient), tumour elimination is predicted. For small σ and large δ (immunocompromised patient), the tumours escape immune surveillance (and grow to carrying capacity). For intermediate values of σ and δ, the predicted tumour responses are diverse; bistability of the tumour elimination/equilibrium and escape solutions can occur; monostability of a homoclinic orbit with a large amplitude, or bistability of the latter with the tumour equilibrium solution can also arise. We conclude that inter-subject variability in immune function, particularly in the supply and death rates of effector cells, or in the initial numbers of tumour and effector cells, can result in large qualitative differences in how a particular tumour progresses. Experiments have indicated similar qualitative dichotomies in tumour progression curves between subjects that respond to cancer therapy, and those that do not. For example, we can observe this in preclinical studies testing the efficacy of immune checkpoint inhibitors of the PD-1/PD-L1 axis alone (such as AstraZeneca's dataset in Fig. 1.1), in combination with a Bruton's tyrosine kinase inhibitor [149] or with radiotherapy [44]. Variability in model parameter values and variability in initial cell numbers represent two different hypotheses via which our model may explain such heterogeneity in tumour responses. This idea provides the basis for the work in Chapter 5.

Our analysis has shown that the intrinsic excitability of the model derived in Chapter 2 generates a complex, non-intuitive, bifurcation structure that includes emergence of a homoclinic orbit. Model solutions traversing the vicinity of this

orbit undergo long-term oscillations that demonstrate long-term tumour recurrence. Short-term oscillations, similar to those observed in leukemias [29, 56, 108, 141], can be seen in the model's excitable parameter regime and away from it; here the system may evolve via damped oscillations to a nonzero-tumour steady state or to a limit cycle with smaller periods and amplitudes than for the large homoclinic orbit. Only damped oscillations about a steady state and no limit cycle solutions exist for Kuznetsov et al.'s QSSA model [88]. However, using a re-parameterised version of Kuznetsov et al.'s original model, Matzavinos et al. [107] showed that the model may exhibit a large-amplitude limit cycle, damped oscillations about a dormant steady state, or bistability of these two types of dynamics. These dynamics qualitatively resemble those identified in the present study of the reduced model, which shows clearly how transitions to oscillatory solutions of this type may arise in different parameter regimes. Long- and short-term oscillatory behaviours, similar to those of our reduced model, have been observed in the tumour-immune interactions model by Kirschner and Panetta [79] that accounts for the presence of the cytokine IL-2. We note further that in parameter regimes where the homoclinic orbit and equilibrium steady state solution are bistable, our reduced model exhibits sneaking through, whereby very small tumours may escape (in the vicinity of the orbit) while larger ones may be controlled, and, counter-intuitively, immunostimulation can be damaging. In Kuznetsov et al.'s QSSA model these phenomena were shown to depend on μ, the rate at which tumour cells inactivate effector cells, while in our model they are attributed to the model's excitable properties. Overall, we have demonstrated that our reduced model is consistent with a range of clinically observed phenomena, providing support for the hypothesis that excitable dynamics may underpin tumour-immune interactions.

By varying the values of parameters that regulate the effector cell dynamics, we used our model to investigate how tumour cell numbers may be controlled via immunotherapy in a heterogeneous population of patients. As previously suggested,

immunotherapies that improve the competence of a patient's immune system (i.e. increase baseline effector supply rate or decrease effector death rate), such as adoptive cell therapy or vaccination, could control or eliminate a tumour predicted to escape in the absence of treatment. Other immunotherapies, such as immune checkpoint therapies, may directly target the immune response to the tumour. Treatments that target ρ, the maximum rate at which effector cells are recruited by the tumour, and μ, the rate at which tumour cells inactivate effector cells, are mainly beneficial to immunocompetent patients that lie in a bistable region of tumour escape and elimination. For immunocompromised patients, the most promising therapy was found to be one that targets η, the tumour cell density at which increased effector cell proliferation and infiltration due to the tumour become limited. Reducing this parameter may drive the system to a monostable parameter region exhibiting tumour equilibrium, or to a bistable region with only a subset of smaller tumours under immune control. We note, however, that tumour control in the latter region would be difficult to achieve (e.g. via surgery, in which the effector and tumour cell numbers are manipulated) due to the small size of the basin of attraction of the dormant steady state. For these patients, the system may be moved to a monostable tumour equilibrium regime via a therapy that targets μ or ρ and η. The model also predicts that it may be more effective to give the majority of immunocompromised patients a combination therapy that initially increases the patient's immunocompetence; this way tumour elimination/control may be achieved or become more easily attainable with subsequent immunotherapies that directly modulate the immune response to the tumour.

In practice, immunotherapy will likely impact multiple parameters, and combination (immuno)therapies may be synergistic or antagonistic [90, 109, 144]. Moreover, extrinsic factors, such as variability in drug concentration due to blood vessel density [43], or effector cell trafficking and infiltration of the tumour, may cause stochastic fluctuations in model parameters [69]. Such random effects may be

dominant at small cell numbers. These observations do not diminish the insight our results provide, rather they suggest possible directions for future investigation. First, it would be interesting to explore how intrinsic and extrinsic stochasticity impacts the model's excitable and bifurcating dynamics. Previous studies have shown that noise in dynamical systems may lead to stochastic bifurcations [41, 94]. Noise may also significantly influence the probability of tumour recurrence or eradication [23]. Second, there is a practical question about the feasibility of tuning model parameters and (combination) therapies with sufficient accuracy to drive a patient and their tumour into a desired parameter region. We address the problem of model calibration in Chapter 4.

Through a systematic dynamical systems analysis of our simplified model of tumour-immune interactions, based on analytical and numerical results, we were able to gain a thorough mechanistic understanding of the qualitative behaviour of the original model. The model captures a variety of clinically observed phenomena through dynamics that are excitable and have a rich bifurcation structure. The model has thus demonstrated the complexity of tumour-immune interactions, and the heterogeneity in their final outcomes. Preferred strategies for controlling tumour size may differ between patients depending on the strength of their immune systems. This result supports personalised approaches to cancer therapy, in which mathematical models could play an important role in future if patient-specific parameters can be reliably estimated [80, 140]. We emphasize that understanding the intricate behaviours of tumour-immune interactions models is one of the pivotal steps towards overcoming difficulties associated with their validation, and, in the longer term, using validated models to generate patient-specific predictions.

Chapter 4

Parameter identifiability of a model of tumour-immune interactions

4.1 Introduction

Given the complexity of interactions of the tumour with its microenvironment, mathematical models of cancer incorporate multiple nonlinear terms in order to describe the biology and generate mechanistic insights. However, increased model complexity typically requires more and/or better quality data for validation of the model, which presents modellers with the challenge of trying to balance model complexity and limitations of experimental data. Natural questions that arise from this problem are as follows: What model complexity is sufficient to describe the existing data? What type and quality of data are sufficient to robustly validate a given model? Are the existing data good enough to calibrate a model of choice?

Motivated by the dichotomous dataset of tumour growth curves provided by our collaborators in AstraZeneca (see Fig. 1.1), we investigate in this chapter the feasibility of calibrating bifurcating and/or bistable models of tumour-immune interactions, such as those presented in previous chapters. These models incorporate dynamics of both tumour and immune cells since functioning of the immune system plays a central role in response or non-response to immunotherapies. The data, however, only provide information on tumour growth dynamics since immune cell

data are difficult to collect. This makes the modelled system only partially observed, and therefore non-trivial to fit to tumour data, particularly due to its complex, nonlinear behaviour. The question at hand is as follows: Is it possible to uniquely, or with low uncertainty, estimate model parameters from tumour growth data for a single subject? Addressing questions of model identification is a pivotal step towards establishing whether a mathematical model can explain partial data in a quantitative way.

We focus on the quasi-steady-state-approximated (QSSA) model by Kuznetsov et al. [88]. It comprises two coupled differential equations and describes interactions between tumour and immune cells. The model is descriptive of the three E's of immunoediting through steady state solutions that correspond to tumour elimination, equilibrium or escape. The existence and stability of these steady state solutions may change as model parameters vary; in certain parameter regimes, model solutions may approach different steady states as initial conditions vary across a threshold – bistability. We choose the QSSA model for analysis due to these desirable behaviours and because its equations and dynamics are of manageable complexity, unlike those of the excitable model presented in Chapter 2.

Our approach considers model calibration from the viewpoint of structural and practical identifiability of model parameters. ***Structural or a priori identifiability*** (SI) is an intrinsic property of a model and its parametrisation alone [146]. It considers whether unknown model parameters can be uniquely determined from perfect observables, e.g. continuous and noise-free tumour growth curves in our case. It is therefore independent of the quality and quantity of data. In addition to the model structure, practical identifiability also considers experimental conditions, i.e. measurement schedule and uncertainty. ***Practical identifiability*** (PI) concerns properties of the parameter estimation problem, the aim of which is to find parameters that minimise some measure of error between model forecasts and observations [102, 134, 146]. Numerical optimisation or inference algorithms are

typically used to search for optimal parameters on an error surface that is defined over a multi-dimensional parameter space. Practical identifiability of model parameters, i.e. uncertainty of their estimates, hence depends on the shape and curvature of this multi-dimensional surface which is influenced by the model's structural properties and the data. Structural identifiability is a prerequisite for practical identifiability [133, 146]. Both should, in principle, be investigated before collecting experimental data and performing parameter fitting in order to understand how to reparametrise the model, what types of data are needed and what experimental conditions ensure that the data fitting procedure is well-defined. While this is not guaranteed for real data even when the model is identifiable, good model identifiability increases the likelihood of producing unique parameter estimates with low uncertainty. Such analysis also equips one with a better understanding of the problem, enabling one to identify and address difficulties that arise during model calibration to data.

In this chapter, we consider the structural and practical identifiability of Kuznetsov et al.'s model in order to determine whether model parameters can be recovered from tumour observables generated from the model. In Section 4.2, we define the parameter estimation inverse problem. In Section 4.3, we formally define SI, summarise methods for testing SI and establish SI of Kuznetsov et al.'s model. We consider PI within the framework of maximum likelihoods estimation (MLE) in Section 4.4. Synthetic datasets that exhibit different qualitative behaviours, corresponding to the different stages of immunoediting, with different observables and noise levels, are used to infer model parameters and explore the parameter estimation problem using a profile likelihood approach. This allows us to draw conclusions about the model's PI given different types of data. In Section 4.5 we discuss our results and outline directions for future work.

4.2 Formulation of the inverse problem

In this section we define the MLE problem of calibrating the QSSA model of tumour-immune interactions proposed by Kuznetsov et al. [88]. We adopt terminology and notation introduced in Section 1.3.1.3.

4.2.1 Forward model

Let $\mathbf{x}(\tau, \boldsymbol{\theta}) = [e(\tau, \boldsymbol{\theta}), m(\tau, \boldsymbol{\theta})]^T$ be the system state vector denoting the number of effector cells, $e(\tau, \boldsymbol{\theta})$, and the number of tumour cells, $m(\tau, \boldsymbol{\theta})$, $\boldsymbol{\theta} = [e_0, m_0, \sigma, \rho, \eta, \mu, \delta, \alpha, \beta]^T$ be the unknown system parameters and $\mathbf{x}^0(\boldsymbol{\theta}) = [e_0, m_0]^T$ the system initial conditions. The dynamical system is given by

$$\dot{e} = \sigma + \frac{\rho e m}{\eta + m} - \mu e m - \delta e =: f_1(\mathbf{x}(\tau, \boldsymbol{\theta}), \boldsymbol{\theta}), \tag{4.1a}$$

$$\dot{m} = \alpha m(1 - \beta m) - e m =: f_2(\mathbf{x}(\tau, \boldsymbol{\theta}), \boldsymbol{\theta}), \tag{4.1b}$$

with initial conditions

$$e(0, \boldsymbol{\theta}) = e_0, \quad \text{and} \quad m(0, \boldsymbol{\theta}) = m_0. \tag{4.1c}$$

The forward model is the numerical solution of these equations. Understanding the forward problem is crucial for solving the inverse problem [99]. Equations (4.1) have been studied in [88, 137] and also in Chapters 2 and 3.

4.2.2 Observation models

We assume that observations, $\mathbf{s} = [\mathbf{y}(\tau_{k_1}, \boldsymbol{\theta}), \ldots, \mathbf{y}(\tau_{k_N}, \boldsymbol{\theta})]^T \in \mathbb{R}^{N_s}$, are generated according to (1.4), i.e. $\mathbf{s} = \boldsymbol{\mathcal{O}}(\boldsymbol{\mathcal{F}}(\boldsymbol{\theta})) + \boldsymbol{\nu}$. They are determined by the forward and observation models, and additive random measurement noise defined as $\boldsymbol{\nu} = [\mathbf{w}(\tau_{k_1}), \ldots, \mathbf{w}(\tau_{k_N})]^T \in \mathbb{R}^{N_s}$. We consider cases where

$$\boldsymbol{\mathcal{O}}(\mathbf{u}) = [\mathbf{h}(\mathbf{x}(\tau_{k_1}, \boldsymbol{\theta})), \ldots, \mathbf{h}(\mathbf{x}(\tau_{k_N}, \boldsymbol{\theta}))]^T, \tag{4.2}$$

with $\mathbf{h}(\cdot) : \mathbb{R}^{N_x} \mapsto \mathbb{R}^H$, so that

$$\mathbf{y}(\tau, \boldsymbol{\theta}) = \underbrace{\mathbf{h}(\mathbf{x}(\tau, \boldsymbol{\theta}))}_{=:\tilde{\mathbf{y}}(\tau,\boldsymbol{\theta})} + \mathbf{w}(\tau), \tag{4.3}$$

where $\tilde{\mathbf{y}}(\tau, \boldsymbol{\theta})$ denotes the noise-free observable (or equivalently model prediction of the observable) at time τ. We define two observation models, which differ in the number of observed components of system (4.1).

Tumour data only. As experimental data typically record only tumour volume over time for different subjects (e.g. AZ's data in Fig. 1.1), here only $m(\tau, \boldsymbol{\theta})$ is measured for each subject. We assume that observation errors are additive Gaussian with zero mean on the logarithmic scale, i.e. $w(\tau) \sim \mathcal{N}(0, \xi_m^2)$. Hence we define the observation model, $h_1(\cdot)$, as follows

$$h_1(\tau, \mathbf{x}(\tau, \boldsymbol{\theta})) = \log\left(m(\tau, \boldsymbol{\theta}) + \frac{1}{16}\right) \qquad \text{for } \tau \in T, \tag{4.4}$$

where T is the set of times at which tumour cell numbers are recorded. In general, we assume that observations are not made at $\tau = 0$, i.e. $0 \notin T$. Note that we have added $\frac{1}{16}$ to the measurement since there exist parameter regimes in which the solution of the system (4.1) satisfies $m(\tau, \boldsymbol{\theta}) \approx 0$ for some $\tau \geq 0$. Since not all components of the system state vector $\mathbf{x}(\tau, \boldsymbol{\theta})$ are observed in this case, it means that our system is partially observed.

Tumour and effector data. We also consider hypothetical cases in which both tumour and effector data are recorded, in order to investigate how model identifiability changes when data on effector cell numbers are available. Our observation operator, $\mathbf{h}_2(\cdot)$, is therefore defined as follows

$$\mathbf{h}_2(\tau, \mathbf{x}(\tau)) = \log\left(\mathbf{x}(\tau) + \frac{1}{16}\right) \qquad \text{for } \tau \in T, \tag{4.5}$$

where T is the set of times at which tumour and effector data are recorded. Observation errors are independent and identically distributed according to $\mathcal{N}(0, \xi_e^2)$ and $\mathcal{N}(0, \xi_m^2)$, respectively, for effector and tumour cell measurements.

4.2.3 Likelihood

In this chapter we consider the problem of estimating $\boldsymbol{\theta}$ from observations $\mathbf{s}$ using MLE, which requires specification of the likelihood. The distribution assumed for observation errors results in a multivariate Gaussian likelihood,

$$p(\mathbf{s} \mid \boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^k |\Xi|}} \exp\left(-\frac{1}{2} V(\boldsymbol{\theta})\right), \tag{4.6}$$

where $V(\boldsymbol{\theta}) := (\mathbf{s} - \boldsymbol{\mathcal{O}}(\boldsymbol{\mathcal{F}}(\boldsymbol{\theta})))\, \Xi^{-1} (\mathbf{s} - \boldsymbol{\mathcal{O}}(\boldsymbol{\mathcal{F}}(\boldsymbol{\theta})))$ as in (1.5), and the covariance matrix Ξ is diagonal with entries ξ_m^2 and ξ_e^2. For simplicity, we assume that the covariance matrix is known, and note that this is sufficient for our purposes of learning about the inverse problem generated by the forward and observation models, but may not be satisfactory when working with real experimental data. In our setup, the MLE problem is equivalent to

$$\hat{\boldsymbol{\theta}}_{\text{MLE}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, V(\boldsymbol{\theta}). \tag{4.7}$$

4.3 Structural identifiability

In this section, we define structural identifiability (SI), outline different methods for SI analysis, and present results of SI analysis for the inverse problem of Kuznetsov et al.'s QSSA model and observables of tumour numbers.

4.3.1 Definitions

The purpose of SI analysis is to determine whether the inverse problem of recovering parameters from ideal experimental data is (i) uniquely solvable, or if there are (ii) a finite, or (iii) an infinite number of parameter vectors that generate identical observables [28, 147]. SI thus depends on the forward model and types of observations recorded, and not the experimental data.

An ideal observable is defined by

$$\tilde{y}(\tau, \boldsymbol{\theta}) := h(\mathbf{x}(\tau, \boldsymbol{\theta})) \qquad \text{for all } \tau \geq 0, \tag{4.8}$$

where $\mathbf{x}(\tau, \boldsymbol{\theta})$ is the exact, time-continuous solution of equations which define the forward problem. For a continuous map $h(\cdot)$, $\tilde{y}(\tau, \boldsymbol{\theta})$ is thus noise-free and continuous in time τ. Here we consider one-dimensional maps $h(\cdot)$ but the definitions generalise to multidimensional maps $\mathbf{h}(\cdot)$.

A parameter vector $\bar{\boldsymbol{\theta}} := [\bar{\theta}_1, \ldots, \bar{\theta}_J]^T$ is called *indistinguishable* from parameter vector $\boldsymbol{\theta} := [\theta_1, \ldots, \theta_J]^T$ if it gives rise to the same observables, i.e.

$$\tilde{y}(\tau, \boldsymbol{\theta}) = \tilde{y}(\tau, \bar{\boldsymbol{\theta}}) \qquad \text{for all } \tau \geq 0. \tag{4.9}$$

We define three identifiability properties of parameter θ_j that characterise the number of solutions to (4.9). Parameter θ_j is *structurally locally identifiable* (SLI) if, for almost any $\bar{\boldsymbol{\theta}} \in \Omega \subset \mathbb{R}^J$, there exists a neighbourhood $V(\bar{\boldsymbol{\theta}})$ for which $\boldsymbol{\theta} \in V(\bar{\boldsymbol{\theta}})$ and (4.9) implies $\theta_j = \bar{\theta}_j$. It is *structurally globally identifiable* (SGI) if, for almost any $\bar{\boldsymbol{\theta}} \in \Omega$, such a neighbourhood is almost the entire parameter state space Ω (except for a set of measure zero). Parameter θ_j is *structurally unidentifiable* (SUI) if for almost any $\bar{\boldsymbol{\theta}} \in \Omega$ no such neighbourhood $V(\bar{\boldsymbol{\theta}})$ exists [103, 175].

The model is SLI/SGI if model parameters θ_j are SLI/SGI for all $j \in \{1, \ldots, J\}$. If θ_j is SUI for at least one $j \in \{1, \ldots, J\}$, then the model is SUI [171, 175].

4.3.2 Methods and results

A variety of methods for testing SI of models are available (see [28, 111] for reviews). While different methods are generally applicable to linear and nonlinear models, their performances differ due to the greater algebraic complexity of nonlinear models.

Taylor series approach. Observables in the neighbourhood of $\tau = \tau_0$ can be represented by their Taylor series expansion

$$\tilde{y}(\tau, \boldsymbol{\theta}) = \tilde{y}(\tau_0, \boldsymbol{\theta}) + \tau \dot{\tilde{y}}(\tau_0, \boldsymbol{\theta}) + \frac{\tau^2}{2!} \ddot{\tilde{y}}(\tau_0, \boldsymbol{\theta}) + \ldots \tag{4.10}$$

In the Taylor series approach, uniqueness of this representation guarantees SI of the system [28, 128]. Let $\bar{\boldsymbol{\theta}}$ be a vector of unknown parameters that is indistinguishable from $\boldsymbol{\theta}$. Formally, a sufficient condition for SGI is

$$a_k(\boldsymbol{\theta}) = a_k(\bar{\boldsymbol{\theta}}) \qquad \Rightarrow \qquad \boldsymbol{\theta} = \bar{\boldsymbol{\theta}} \qquad k \in \{1, 2, \ldots, k_{\max}\}, \tag{4.11}$$

where $a_k(\boldsymbol{\theta}) := \lim_{\tau \to \tau_0^+} \tilde{y}^{(k)}(\tau, \boldsymbol{\theta})$ are Taylor coefficients of the model output, and $k_{\max}$ is the smaller positive integer for which the system (4.11) gives a solution for parameters $\boldsymbol{\theta}$. While the Taylor series approach is simple to understand and implement, it is difficult to determine the number of derivatives (e.g. $k_{\max}$) beforehand. Moreover, nonlinearities in the model may result in complex algebraic expressions that may be computationally infeasible to solve symbolically.

Differential algebra approaches. The common aim of differential algebra approaches is to summarise the input-output behaviour of the system $\{\dot{\mathbf{x}} - \mathbf{f}(\mathbf{x}(\tau, \boldsymbol{\theta}), \boldsymbol{\theta}), \tilde{y} - h(\mathbf{x}(\tau, \boldsymbol{\theta}))\}$ with a differential polynomial, or a set of thereof, called the *input-output* (IO) *relation*. For a one-dimensional observable, the IO relation may be written in the form

$$\zeta_{N+1}(\tilde{y}) = \sum_{i=1}^{N} \zeta_i(\tilde{y}) b_i(\boldsymbol{\theta}), \tag{4.12}$$

where $(\zeta_i(\tilde{y}))_{i=1}^{N+1}$ are differential polynomials in terms of $\tilde{y}$ and its derivatives, $b_i(\boldsymbol{\theta})$ are rational functions in $\boldsymbol{\theta}$, and N is some positive integer. Equations of the IO relation have a special form that involves only observables $\tilde{y}$ and unknown parameters $\boldsymbol{\theta}$, having eliminated non-observed variables with differential elimination algorithms (see [18]); these procedures are analogous to Gauss-elimination in the case of linear systems of equations. We remark that the functions $b_i(\boldsymbol{\theta})$ fully embody the parameter dependence of the IO relation [12], similar to Taylor series coefficients $a_k(\boldsymbol{\theta})$. Introducing $\bar{\boldsymbol{\theta}}$, a parameter vector that is indistinguishable from $\boldsymbol{\theta}$, and since then $\tilde{y}^{(k)}(\tau, \boldsymbol{\theta}) = \tilde{y}^{(k)}(\tau, \bar{\boldsymbol{\theta}})$ for all $\tau \geq 0$ and $k \in \mathbb{N}$, one may deduce

an alternative version of the IO relation as follows:

$$\sum_{i=1}^{N}(b_i(\boldsymbol{\theta}) - b_i(\bar{\boldsymbol{\theta}}))\zeta_i(\tilde{y}) = \zeta_{N+1}(\tilde{y}) - \zeta_{N+1}(\tilde{y}) = 0. \tag{4.13}$$

Provided the polynomials $(\zeta_i(\tilde{y}))_{i=1}^{N}$ are linearly independent, one can extract a nonlinear system of algebraic equations involving $\boldsymbol{\theta}$ and $\bar{\boldsymbol{\theta}}$ given by

$$b_i(\boldsymbol{\theta}) = b_i(\bar{\boldsymbol{\theta}}) \qquad \text{for} \quad 1 \leq i \leq N. \tag{4.14}$$

Compared to systems obtained with the Taylor series approach, (4.14) is algebraically simpler to solve, and the number of its solutions determines SI.

Variants of this approach [7, 12, 38] mainly differ in the way linear independence of $(\zeta_i(\tilde{y}))_{i=1}^{N}$ and solutions to (4.14) are established. Several software packages are available, but most require expert mathematical knowledge, e.g. `DifferentialAlgebra` package in Maple™ [106]. Exceptions are the DAISY software [12] and the COMBOS online app [110], but they do not provide any intermediate results and can be computationally expensive.

Results for Kuznetsov et al.'s model. To determine the SI of Kuznetsov's model given tumour cell numbers as observables, we used a differential algebra-based approach by Bearup et al. [7], which is suited for polynomial or rational nonlinear systems such as (4.1). Symbolic computations are used to determine linear independence of monomials $(\zeta_i(\tilde{y}))_{i=1}^{N}$ and solutions to (4.14). This method can be easily implemented in Maple™, but it is not fully automated and requires some knowledge of differential algebra to perform simple checks of the methodology.

For brevity, the intermediate steps of our computations in Maple™ are not presented here. We simply report that, given observation model $h_1(\cdot)$, model (4.1) is SGI since all parameters $\boldsymbol{\theta}$ are found to be SGI. We verified this result for unknown kinetic parameters with the COMBOS software (results not shown for brevity). We remark that SI for $h_1(\cdot)$ automatically implies SI for observation model $\mathbf{h}_2(\cdot)$, which also includes effector cell numbers as observables.

4.4 Practical identifiability

We have shown that Kuznetsov et al.'s QSSA model is structurally identifiable. In this section, we study the model's data-based or *practical identifiability* (PI).

PI is closely related to parameter estimation, which aims to find parameter values for which the model best describes the observed data. In Chapter 1 we introduced different frameworks within which the parameter estimation problem can be defined. If the minimum of the objective function that measures the prediction error is unique, then the model is said to be *practically identifiable* [102, 146].

For problems posed within a frequentist framework, we say that a model is *practically identifiable* if its maximum likelihood estimate (MLE), $\hat{\boldsymbol{\theta}}_{\text{MLE}}$, is unique. In practice, it is important to determine uncertainties in parameter estimates and their possible non-identifiabilities, as these affect uncertainties in model predictions [132]. Uncertainty in a parameter estimate $\hat{\theta}_j$ is defined by its *confidence interval* (CI) $[\theta_j^-, \theta_j^+]$ and an associated *confidence level* $(1-\alpha)$, which represents the proportion of possible CIs $[\theta_j^-, \theta_j^+]$ that contain the true value θ_j^*. In other words, at confidence level $(1-\alpha)$, $(1-\alpha)\%$ of sampled datasets will give CIs that contain the true parameter value. Even though the MLE of a parameter may be unique, a large CI indicates non-identifiability of the parameter. To account for this, Raue et al. [134] redefined PI as follows: a parameter is *practically identifiable* if its CI is finite, and practically non-identifiable if its CI is infinite.

Computation of CIs is challenging, particularly in the case of non-identifiabilities. For example, in a two-dimensional parameter space, practical non-identifiabilities of parameters manifest via infinitely extended, relatively flat valleys on the likelihood surface (see Fig. 4.1), within which the value of the likelihood remains below a desired threshold value [134]. For a small number of parameters one may compute the likelihood surface to detect such valleys, whereas for larger models this is not feasible. CIs based on quadratic approximations of the likelihood around the MLE

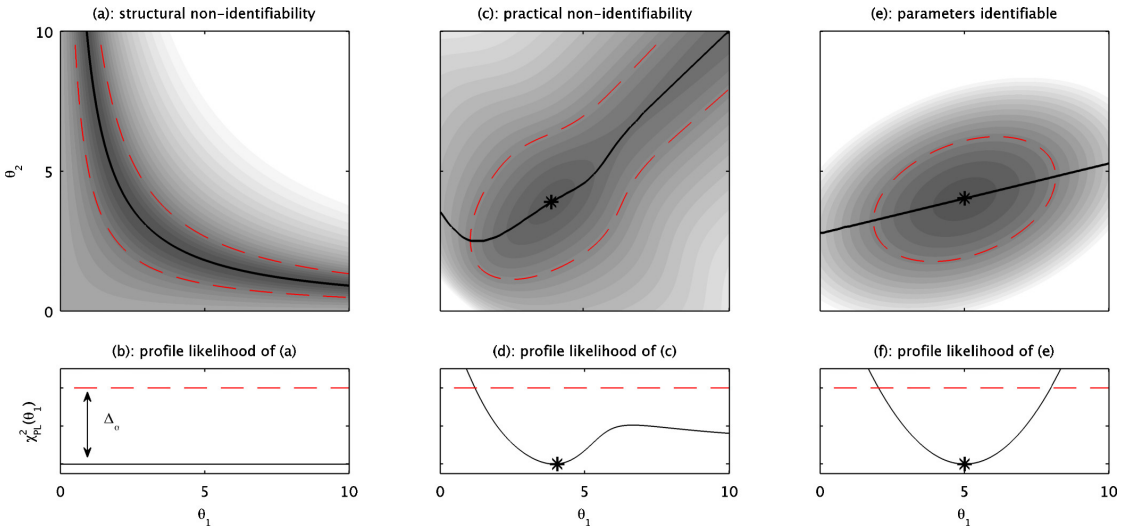


Figure 4.1: Comparison of typical likelihood landscapes in the cases of structural non-identifiability, practical non-identifiability and identifiability of model parameters. Upper panel density plots show the likelihood in (θ_1, θ_2) parameter space; darker shades correspond to lower values of the negative log-likelihood and the minimum is marked with an asterisk. Lower panel plots show the corresponding profile likelihood (solid black curves; denoted as $\chi^2_{\mathrm{PL}}(\theta_1)$ in this figure, but denoted as $\mathrm{PL}_1(\theta_1)$ later in this chapter) as θ_1 is varied, i.e. the minimum of the negative log-likelihood at different values of θ_1. The dashed red curves denote the likelihood threshold values that are used to assess likelihood-based confidence intervals with confidence level α (see Section 4.4.1.3). (This figure is republished with permission of American Institute of Physics, from [132]; permission conveyed through Copyright Clearance Center, Inc.)

are more computationally efficient but prone to large inaccuracies in the case of asymmetries and non-identifiabilities [71].

Raue et al. [134] proposed to use the method of *profile likelihoods* (PLs), a common tool in statistical inference [125], to detect flat regions and determine parameter identifiability. For different values of a parameter θ_j, the likelihood is minimised with respect to all other unknown parameters, $\theta_{i\neq j}$ (see Section 4.4.1.3). In this way, one traces the curvature of the likelihood along its flattest regions as θ_j varies in some interval. A U-shaped profile likelihood curve that exceeds a chosen threshold twice implies a finite CI for θ_j and its PI, otherwise θ_j is non-identifiable (see Fig. 4.1).

Here synthetic data generated from Kuznetsov et al.'s QSSA model are used as observables, and we adopt a frequentist MLE approach to fit the model to the data. Our aim is to determine how sensitive recovery of parameters and model dynamics

is to different types of output data, which may be viewed as PI analysis [134] or a form of inverse sensitivity analysis [31, 71, 92]. We exploit PLs to explore the resulting inverse problems, and identify conditions under which model parameters can be inferred from the data. We describe our methodology in Section 4.4.1, and present results in Section 4.4.2.

We note that PI of a parameter within a Bayesian framework may be defined in an analogous manner to the frequentist framework, whereby a large credible interval (analogous to a CI) of parameter's marginal posterior distribution implies its non-identifiability. Bayesian approaches to parameter identifiability have been used before [31, 71, 74, 135, 155, 168], using Markov Chain Monte Carlo methods. Building upon our knowledge of the MLE problem investigated below, we explore the Bayesian versions of the inverse problem related to Kuznetsov et al.'s QSSA model in Chapter 5.

4.4.1 Methods

In this section we describe the methodology we use to investigate PI of Kuznetsov et al.'s QSSA model. Our workflow is summarised in Fig. 4.2, and individual steps are detailed below.

4.4.1.1 Generation of synthetic data

We define a synthetic dataset for a set of parameter values (one subject) to be the tumour cell growth measurements, and, if available, also effector cell measurements, at discrete time points, and concatenated into a vector $\mathbf{s}$. For timepoints $T_N := \{n\Delta t \,|\, 1 \leq n \leq N-1, k \in \mathbb{N}, \Delta t = \frac{10}{N}\}$, observation model, $\mathbf{h}(\cdot)$, and a measurement error with standard deviation $\xi_m = \xi_e = \xi$, a synthetic dataset is generated as follows:

(1) Pick a parameter regime $\boldsymbol{\theta}^*$.

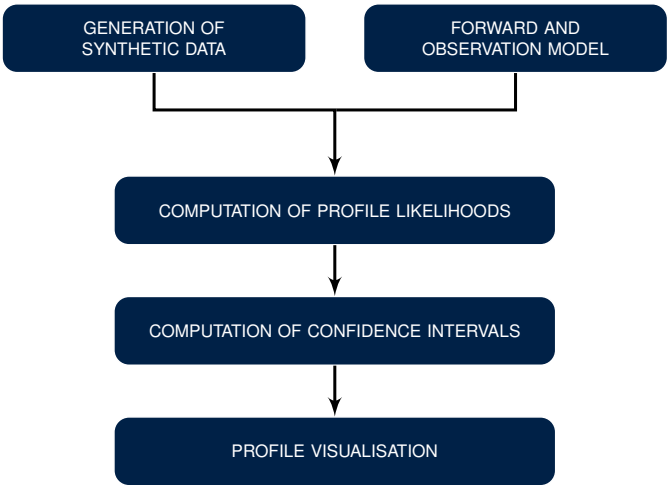


Figure 4.2: Schematic diagram summarising the workflow used to determine practical identifiability of a model given synthetic data.

(2) Compute $\boldsymbol{\mathcal{O}}(\boldsymbol{\mathcal{F}}(\boldsymbol{\theta}^*))$ by first numerically solving the forward problem for $\tau \in T_N$ to obtain $\boldsymbol{\mathcal{F}}(\boldsymbol{\theta}^*)$, and then using equation (4.2) to transform the solution to $\boldsymbol{\mathcal{O}}(\boldsymbol{\mathcal{F}}(\boldsymbol{\theta}^*))$.

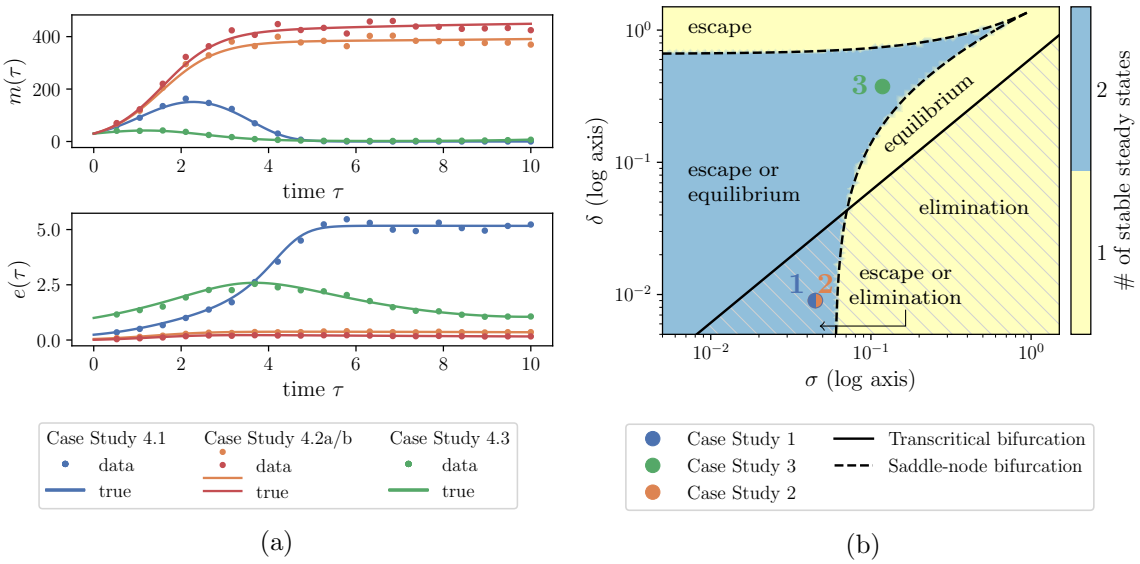
(3) Simulate $\boldsymbol{\nu}$, the vector of random variables $\mathcal{N}(0, \xi^2)$, and compute the vector of synthetic data $\mathbf{s} = \boldsymbol{\mathcal{O}}(\boldsymbol{\mathcal{F}}(\boldsymbol{\theta}^*)) + \boldsymbol{\nu}$.

We generate three different types of datasets, choosing parameter regimes for which the model exhibits different qualitative behaviours, each representative of one of the three stages of immunoediting – tumour elimination, equilibrium, or escape. The parameter regimes are stated in Table 4.1 and are indicated on the two-dimensional bifurcation diagram in Fig. 4.3b. Fig. 4.3a shows the corresponding forward model solutions from which synthetic datasets were generated, and exemplary datasets for $T = T_{20}$, $\xi = 0.03$ and $\mathbf{h}(\cdot) = \mathbf{h}_2(\cdot)$. For each case study, we generate datasets under noise levels $\xi \in \{0.03, 0.05, 0.1, 0.15, 0.2\}$, and observation models $\mathbf{h}_2(\cdot)$ and $h_1(\cdot)$.

This setup allows us to investigate how identifiability differs between datasets exhibiting uncontrolled tumour growth and tumour shrinkage, and as recorded observables and noise levels in the data vary.

Case study No.	Description of (σ, δ) regime	Qualitative behaviour	Parameter θ^*_{sub} e_0	m_0	σ	δ
4.1	elim. and esc. - bistable	elim.	0.24	30	0.045	0.009
4.2a	elim. and esc. - bistable	esc., close to separatrix	0.045	30	0.045	0.009
4.2b	elim. and esc. - bistable	esc., away from separatrix	0.01	30	0.045	0.009
4.3	equi. and esc. - bistable	equi.	1	30	0.118	0.374

Table 4.1: Parameter values used for generation of synthetic datasets.



(a) (b)

Figure 4.3: Plot (a) shows an example of three synthetic datasets that exhibit immunoediting behaviour for noise level $\xi = 0.03$ and observation model $\mathbf{h}_2(\cdot)$. These datasets are used for our case studies. Point markers denote the generated synthetic data, while solid curves denote forward model simulations from which the data were generated. Plot (b) shows the (σ, δ) bifurcation diagram of the Kuznetsov model denoting disjoint regions of the (σ, δ) parameter space with qualitatively different model behaviour, for other parameters fixed at default values (see Table 2.2). Values of σ and δ used to compute curves in plot (a) are denoted with coloured point markers, numbered according to which case study in Table 4.1 they correspond to.

4.4.1.2 Simplified parameter estimation problem

We formulated the inverse problem of estimating unknown model parameters in Section 4.2. We anticipate that features of our problem, such as high nonlinearity of the mathematical model (i.e. the forward problem), combined with partial observations (for observation model $h_1(\cdot)$) and a relatively large number of parameters for the amount of data, will render the likelihood landscapes complicated and intractable. For this reason, we first perform PI analysis on simpler problems and increase complexity where appropriate. We outline below the simplified problems

that we explore in Section 4.4.2.

We consider a subset of the model parameters, $\boldsymbol{\theta} = [e_0, m_0, \sigma, \rho, \eta, \mu, \delta, \alpha, \beta]^T$, and recall that our focus is on parameter identifiability for a bifurcating and/or bistable model. In line with this, we focus on parameters, denoted as $\boldsymbol{\theta}_{\text{sub}}$, for which the model's immunoediting behaviour via bifurcations and bistability is preserved as $\boldsymbol{\theta}_{\text{sub}}$ varies, with other parameters fixed at their default values (see Table 2.2). Fig. 4.3b shows that this is the case for variation of σ and δ [88, see also Chapter 3]; as these two parameters vary the model passes through mono- and bistable regions and admits steady states that correspond to the different stages of immunoediting. We therefore include these two parameters in $\boldsymbol{\theta}_{\text{sub}}$. We preserve bistability by also including the initial conditions, e_0 and m_0, in $\boldsymbol{\theta}_{\text{sub}}$. Thus the reduced set of unknown model parameters is $\boldsymbol{\theta}_{\text{sub}} = [e_0, m_0, \sigma, \delta]^T$. We seek to infer $\boldsymbol{\theta}_{\text{sub}}$ from synthetic datasets using MLE.

4.4.1.3 Profile likelihood approach

The general idea of the *profile likelihood* (PL) approach is to explore the parameter space for each parameter individually in the direction of the least increase of the objective function $V(\boldsymbol{\theta})$, which defines the inverse problem, starting from the global minimum. The profile of the objective function $V(\boldsymbol{\theta})$, computed for each parameter $\theta_i \in \boldsymbol{\theta}$, is formally defined as

$$\text{PL}_i(p) := \min_{\boldsymbol{\theta}_{-i}} V(\boldsymbol{\theta} \,|\, \theta_i = p)\,, \tag{4.15}$$

where $\boldsymbol{\theta}_{-i} = [\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_J]^T$ is the vector of nuisance model parameters $\theta_j \in \boldsymbol{\theta}$ for $j \neq i$, which does not include the parameter of interest, θ_i. To compute the full profile curve, parameters $\boldsymbol{\theta}_{-i}$ must be optimised for a series of values of θ_i within an interval of interest around the global minimum, hence ensuring to find values of $V(\boldsymbol{\theta})$ as small as possible within this interval.

We use Algorithm 1, which is similar to the direct step method proposed in Raue et al. [134], to compute the profiles of $V(\boldsymbol{\theta})$. Profile computation starts at the objective function minimum, $\boldsymbol{\theta}^0$ (equal to $\hat{\boldsymbol{\theta}}_{\text{MLE}}$), which is found using a minimisation routine. From this point, we vary θ_i in the positive/negative direction so that the increase in the objective function at the new value of θ_i remains below a chosen threshold. The new value of θ_i is then used to initialise minimisation at

Algorithm 1: Algorithm for computation of the profiles of the objective function $V(\boldsymbol{\theta})$ in the direction of a parameter $\theta_i \in \boldsymbol{\theta}$. The formulation below is similar to the direct step method proposed by Raue et al. [134].

Required parameters: parameter location i, initial guess $\boldsymbol{\theta}^{\text{ini}}$, maximum step size $\varepsilon^{\max}$, minimum step size $\varepsilon^{\min}$, lower interval bound $\beta_{\min}$, upper interval bound $\beta_{\max}$, q, Δ_α, maximum number of points per branch N_{points}

$\boldsymbol{\theta}^0 = \textsc{Optimise}(V(\boldsymbol{\theta}),\, \boldsymbol{\theta}^{\text{ini}})$ ▷ *Find the objective function minimum*
$\text{LPL}_i = \textsc{ComputeProfileBranch}(\boldsymbol{\theta}^0, \beta_{\min}, -1)$ ▷ *Compute left profile branch*
$\text{RPL}_i = \textsc{ComputeProfileBranch}(\boldsymbol{\theta}^0, \beta_{\max}, 1)$ ▷ *Compute right profile branch*

function $\textsc{ComputeProfileBranch}(\boldsymbol{\theta}^0,\, \beta,\, \gamma)$
Inputs: initial parameter set $\boldsymbol{\theta}^0$, interval bound β, direction of parameter sweep $\gamma \in \{1, -1\}$
Outputs: profile PL_i
$j = 0$ ▷ *Initialise iteration counter*
$\text{PL}_i(\theta_i^0) = V(\boldsymbol{\theta}^0)$ ▷ *Evaluate profile at the initial point*
while $j < N_{\text{points}}$ **and** $\gamma\theta_i^j < \gamma\beta$ **do**
$j = j + 1$ ▷ *Increase iteration counter*
$\varepsilon = \min\{\varepsilon^{\max}, |\beta - \theta_i^{j-1}|\}$ ▷ *Set step size to maximum*
while $|\text{PL}_i(\theta_i^{j-1}) - V(\boldsymbol{\theta}^{j-1} \,|\, \theta_i = \theta_i^{j-1} + \gamma\varepsilon)| > q \cdot \Delta_\alpha$ **and** $\varepsilon > \varepsilon^{\min}$ **do**
$\varepsilon = \max\{\frac{\varepsilon}{2}, \varepsilon^{\min}\}$ ▷ *Halve step size if change in $V(\boldsymbol{\theta})$ sufficiently small*
$\theta_i^j = \theta_i^{j-1} + \gamma\varepsilon$ ▷ *Perform step of size ε in θ_i-direction*
$\boldsymbol{\theta}_{-i}^j = \textsc{Optimise}(V(\boldsymbol{\theta} \,|\, \theta_i = \theta_i^j),\, \boldsymbol{\theta}_{-i}^{j-1})$ ▷ *Find the minimum at the new θ_i*
$\text{PL}_i(\theta_i^j) = V(\boldsymbol{\theta}^j)$ ▷ *Evaluate profile at the minimum found*
return PL_i

function $\textsc{Optimise}(F(\boldsymbol{\theta}),\, \boldsymbol{\theta}^{\text{ini}})$
Inputs: function to minimise $F(\boldsymbol{\theta})$, initial guess $\boldsymbol{\theta}^{\text{ini}}$
Outputs: $\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, F(\boldsymbol{\theta})$
▷ *Minimise $F(\boldsymbol{\theta})$ using a chosen optimisation algorithm with starting guess $\boldsymbol{\theta}^{\text{ini}}$*
return $\hat{\boldsymbol{\theta}}$

the next step. For computational reasons we do not allow the step to drop below a minimum step size, $\varepsilon^{\min}$, even when the criterion determining the maximum objective function increase is not satisfied. For parameters where the MLE is close to zero and the surrounding confidence region is small, we perform stepping on a logarithmic scale in order to zoom in on the confidence region.

Optimisation. We use a local, gradient-based, solver for all optimisations. This is adequate for finding minima when we move along the profile, since stepping is gradual and the solver starts close to the minimum. During initial optimisation, the true data-generating parameter values are used as starting guesses. While this is sufficient for determining identifiability given synthetic data generated from the model, it may be difficult to find the global minimum for real experimental data where starting guesses may be poor and the global minimum may not be unique.

A Gaussian likelihood, with known variance, yields a nonlinear least-squares objective function $V(\boldsymbol{\theta})$, and many solvers (e.g. [24, 114] and references in [113]) are adapted to these types of problems. In particular, we use the trust-region-reflective solver implemented in the `scipy.optimize.least_squares` function in Python. We choose relatively large optimisation bounds to cover a large region of parameter space (see Appendix C). As the value of the tumour initial condition, m_0, is at least an order of magnitude larger than values of e_0, σ and δ in the chosen parameter regimes (see Table 4.1), we choose a scaling factor of 10 for m_0, and 1 for other parameters. Objective function evaluation involves computing the forward model by solving equations (4.1) using `scipy.integrate.solve_ivp`, which implements a fourth-order Runge-Kutta method. We also supply the solver with a function that computes the derivative of residuals, $\mathbf{s} - \mathcal{O}(\mathcal{F}(\boldsymbol{\theta}))$, with respect to the parameters, i.e. $\frac{\mathrm{d}(\mathbf{s}-\mathcal{O}(\mathcal{F}(\boldsymbol{\theta})))}{\mathrm{d}\boldsymbol{\theta}} = -\frac{\mathrm{d}\mathcal{O}(\mathcal{F}(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}}$. Sensitivities of model predictions, $\frac{\mathrm{d}\mathbf{x}(\tau,\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}}$ for $\tau \in T_N$, are needed to compute the residual derivative. Following recommendations in [136],

we compute the sensitivities with sensitivity equations

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\mathrm{d}\mathbf{x}(\tau,\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}} = \frac{\partial \mathbf{f}(\mathbf{x}(\tau,\boldsymbol{\theta}),\boldsymbol{\theta})}{\partial \mathbf{x}}\frac{\mathrm{d}\mathbf{x}(\tau,\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}} + \frac{\partial \mathbf{f}(\mathbf{x}(\tau,\boldsymbol{\theta}),\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \tag{4.16}$$

which are stated using notation of Section 1.3.1.3. By extending forward model equations (4.1) by (4.16), we solve sensitivity equations simultaneously with the forward model for $\tau \in T_N$.

Confidence intervals. To determine uncertainty in inferred parameters, we use a threshold in the computed PL to determine the CI around the MLE. In particular, we compute likelihood-based CIs, based on the distributional properties of the Wilk's likelihood ratio statistic [125, 134] given by

$$W = 2(\max_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}) - \max_{\boldsymbol{\theta}_{-A}} \mathcal{L}(\boldsymbol{\theta}_{-A},\, \boldsymbol{\theta}_A = \boldsymbol{\theta}_0)). \tag{4.17}$$

Here subscripts A and $-A$ denote disjoint partitions of the set of unknown parameters $\{\theta_1, \theta_2, \ldots, \theta_J\}$; $\boldsymbol{\theta}_A$ are parameters of interest, fixed at $\boldsymbol{\theta}_0$, while $\boldsymbol{\theta}_{-A}$ are nuisance parameters that are optimised out. For reasonably regular likelihoods [125], $W \sim \chi^2_{J-R}$, as the sample size tends to ∞, where J is the dimension of the total parameter space of $\boldsymbol{\theta}$ and R is the dimension of the parameter space of $\boldsymbol{\theta}_{-A}$. It follows that a $(1-\alpha)\%$ confidence region, which holds simultaneously for all parameters, is given by $\{\boldsymbol{\theta} \,|\, V(\boldsymbol{\theta}) - V(\hat{\boldsymbol{\theta}}_{\mathrm{MLE}}) < \chi^2_{J,\,1-\alpha}\}$, while a $(1-\alpha)\%$ CI that holds pointwise for some parameter θ_i is given by $\{\theta_i \,|\, \mathrm{PL}_i(\theta_i) - V(\hat{\boldsymbol{\theta}}_{\mathrm{MLE}}) < \chi^2_{1,\,1-\alpha}\}$. Approximate bounds for these regions can be calculated by finding the points at which the PL intersects with the objective function threshold, $\chi^2_{J,1-\alpha}$ for a simultaneous confidence region or $\chi^2_{1,1-\alpha}$ for a pointwise CI. Following Simpson et al. [155], we determine these points by applying cubic interpolation to the PLs.

Profile visualisation. For a specific dataset, we use pair plots to present the PLs for unknown model parameters. The diagonal shows the PL as a single parameter varies. Off-diagonal plots show how nuisance parameters vary along the PLs in

order to distil any parameter dependencies. We also overlay PLs in (σ, δ) bifurcation diagrams, such as Fig. 4.3b, in order to relate the landscape of $V(\boldsymbol{\theta})$ to the model's bifurcation structure. Bifurcation diagrams are computed as in Chapter 3. For each case study, we plot, on the same axes, PLs for different observation models and different levels of observation noise. We also compare the errors in the MLEs and the widths of the likelihood-based confidence regions, in order to determine how noise and availability of effector data affects identifiability. We perform checks of the so-called predictive fit, where we consider how forward model solutions along the profile vary within the obtained CIs.

4.4.2 Results

We explore profiles for synthetic datasets, generated as explained in Section 4.4.1.1. For each dataset, we compute the profile of the objective function $V(\boldsymbol{\theta})$, separately for each parameter in $\boldsymbol{\theta}_{\text{sub}} = [e_0, m_0, \sigma, \delta]^T$. Our strategy is to first analyse profiles for datasets that include tumour and effector cells (observation model $\mathbf{h}_2(\cdot)$) and low noise levels ($\xi = 0.03$) in order to understand the problem when the system is fully observed. We then increase the standard deviation of noise in the data ($\xi = 0.05, 0.1, 0.15$ and 0.2) and/or assume only tumour cell measurements are available (observation model $h_1(\cdot)$).

4.4.2.1 Case study 4.1: elimination regime

Profiles for the fully observed system and low noise. Profiles for all parameters have a unique minimum, which lies close to the true parameter values used to generate the data (see Fig. 4.5). The profiles are U-shaped with small 95% CIs, and thus indicate good identifiability of parameters for this dataset.

Profiles for high noise levels and/or the partially observed system. As the standard deviation of observation noise increases, the profiles for all parameters

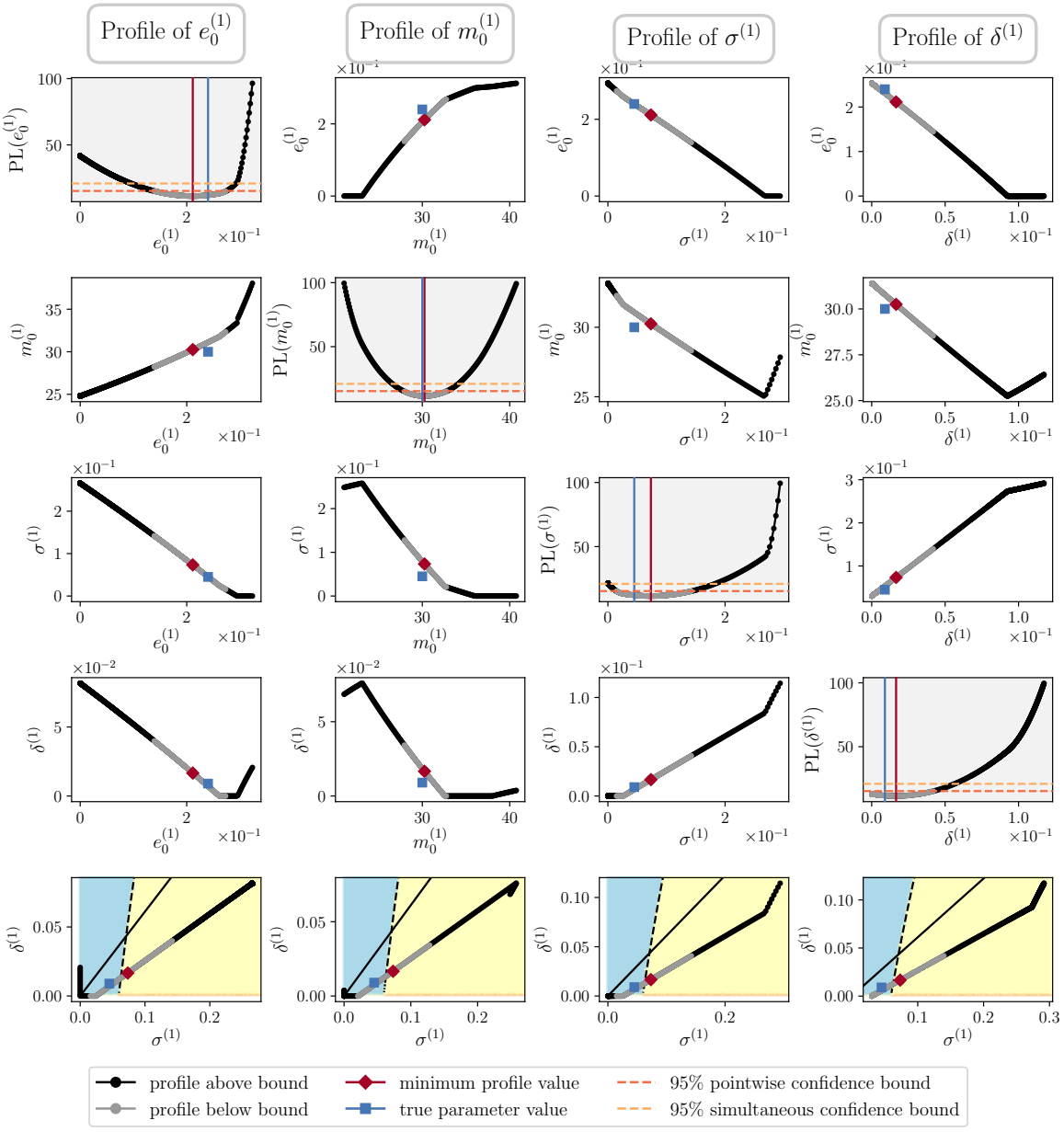


Figure 4.4: Series of plots showing profile likelihoods of $V(\boldsymbol{\theta})$ for parameters $\boldsymbol{\theta}_{\text{sub}} = [e_0, m_0, \sigma, \delta]^T$ given **elimination** data (**case study 4.1**), observation model $h_1(\cdot)$ and observation noise standard deviation $\xi = 0.03$. Diagonal plots with light grey background show the computed profile values as the parameter of interest varies, while off-diagonal plots in each column show how other (nuisance) parameters vary along the profile in the same column. Plots in the bottom row show the (σ, δ) bifurcation diagrams overlaid with profiles. Parts of profiles with values below the 95% pointwise confidence threshold $\chi^2_{1,0.95}$ are plotted in grey, whereas those that exceed it are plotted in black. For interpretation of colours in the bifurcation diagrams in the bottom row, see Fig. 4.3b. The shown profiles are U-shaped, L-shaped or bathtub-like for different parameters. They extend along a straight line in (σ, δ) space that is parallel to the transcritical bifurcation line on which the model's zero-tumour steady state becomes unstable. The variation of nuisance parameters along profiles indicates correlations among parameters. Confidence intervals lie in regions of parameter space in which the model exhibits a stable zero-tumour steady state. MLEs mostly correspond well with the true, data-generating parameters.

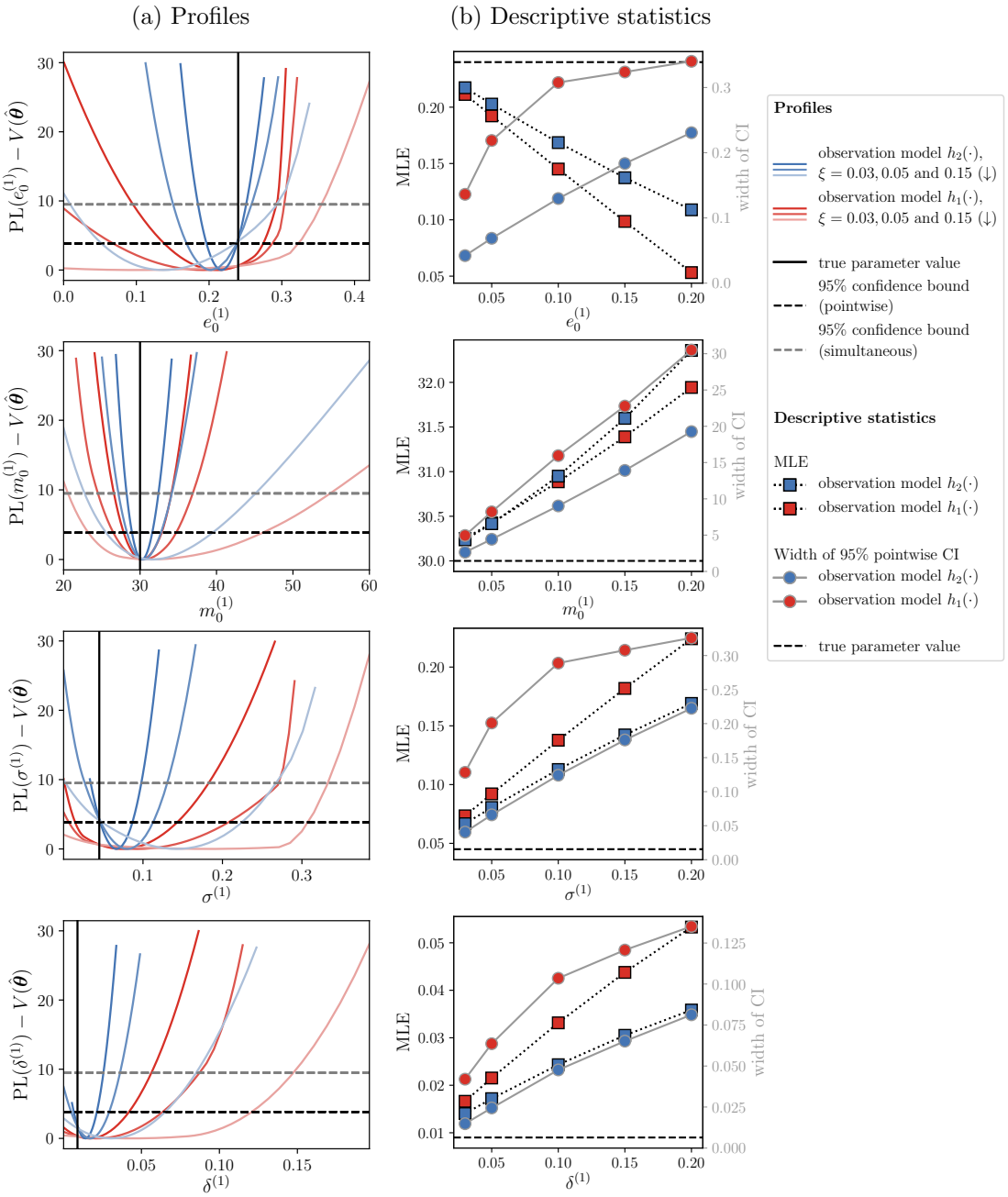


Figure 4.5: Series of plots showing the impact of observation noise standard deviation ($\xi \in [0.03, 0.2]$) and observation models with and without effector cell data ($\mathbf{h}_2(\cdot)$ and $h_1(\cdot)$) on profile likelihoods for data generated from the **elimination** regime (**case study 4.1**). Panel (a) shows profile values translated by $V(\hat{\boldsymbol{\theta}})$ as parameters of interest in $\boldsymbol{\theta}_{\text{sub}} = [e_0, m_0, \sigma, \delta]^T$ individually vary, while panel (b) shows how the corresponding maximum likelihood estimates (MLEs) and the size of confidence intervals (CIs) vary with increasing ξ, split by different observation models. Curves and points plotted in blue correspond to observation model $\mathbf{h}_2(\cdot)$, while those in red correspond to $h_1(\cdot)$. As noise level increases, profiles flatten and confidence intervals increase, but remain bounded for e_0, σ and δ.

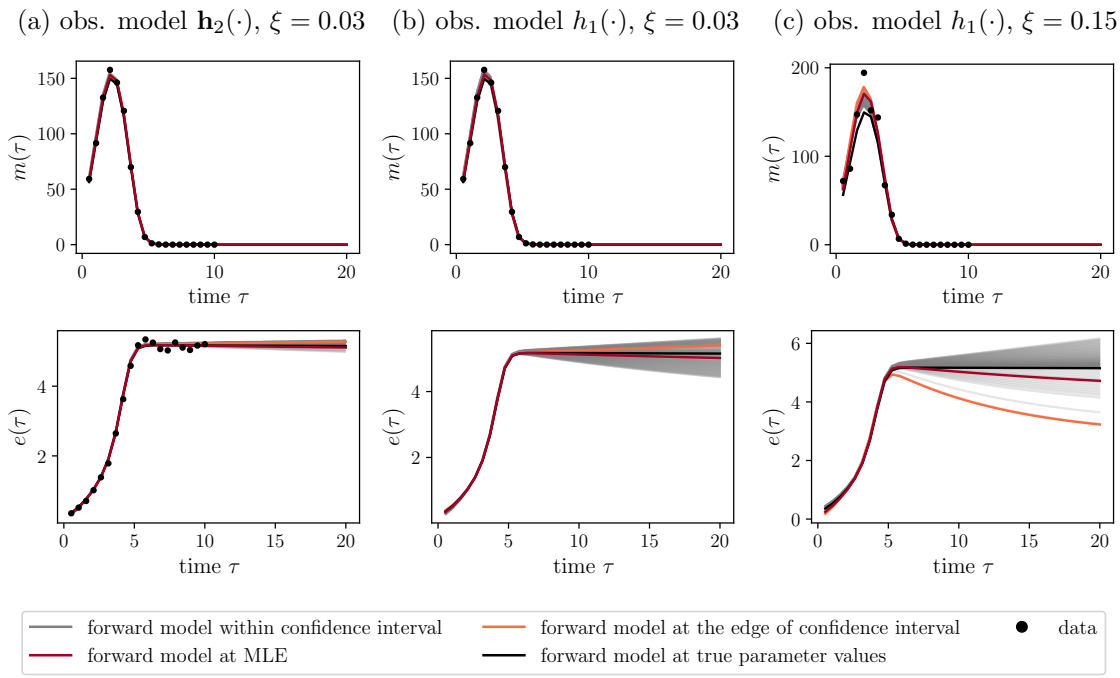


Figure 4.6: Series of plots showing the sensitivity of forward model solutions for parameters within the 95% pointwise confidence interval for σ given **elimination** data (**case study 4.1**). Dynamics corresponding to the elimination stage of immunoediting are successfully recovered for different quality/quantity of measurements.

undergo some level of flattening, and CIs increase (see Fig. 4.5). For the cases considered, the 95% CIs are finite, suggesting overall practical identifiability in this regime, at least up to the width of CIs.

For the fully observed system, the size of the CIs typically grows approximately linearly (see plot of CI width in Fig. 4.5b). However, for some parameters the CIs extend away from the true parameter value in a biased way, which can result in deformation of the regular U-shaped profiles to an asymmetric L-shape. For example, the upper bounds of CIs for the effector initial condition, e_0, all fall close to 0.3, while the lower ones extend increasingly to zero, biasing the MLE towards values that are smaller than the true value. Similar behaviour can be seen in profiles for parameters σ and δ, where the CIs increase more rapidly away from zero and, thus, the MLE for these parameters is larger than the true value. Such effects are less noticeable for the tumour initial condition, m_0.

Removing effector measurements from the dataset results in considerable flat-

tening of the profiles (see Fig. 4.5), and their shape becomes bathtub-like, i.e. flat in most of the CI. As a result, the MLE may be difficult to identify exactly and more prone to bias. Also, observation model $h_1(\cdot)$ has different effects on CIs at different noise levels. In comparison to the fully observed case, CIs for m_0 increase negligibly at small noise levels, but their size is much greater at larger noise levels (e.g. 64% increase in width of CI from 14 for $\mathbf{h}_2(\cdot)$ to 23 for $h_1(\cdot)$ when $\xi = 0.15$). For parameters σ, δ and e_0, availability of effector cell data has a greater impact on identifiability at small noise levels and a lesser effect at large noise levels. For $\xi \leq 0.05$, the size of their CIs increases by more than 100% for all three parameters compared to the fully observed system, but the growth of their CIs is slower for $\xi > 0.05$. This suggests that uncertainty in these three parameters is bounded, and we investigate the reasons for this next.

Variation of nuisance parameters and the impact on profiles. How the optimised parameters vary along a profile for a particular parameter provides information about how parameter relationships affect the curvature of the profiles. The trends below are characteristic for data generated from the elimination parameter regime across all noise levels and observation models.

Variation of nuisance parameters along the profiles indicates dependencies among the parameters (see off-diagonal scatter plots in Fig. 4.4). The most pronounced parameter dependency is between σ and δ, which are positively correlated. As a result, all profiles initially lie along a straight line with a positive slope in the (σ, δ) plane (see bottom panel of Fig. 4.4), which spans parameter regions that exhibit elimination (blue and yellow regions below the line $\delta = \delta^{\diamond}(\sigma, \alpha, \kappa_2)$). Initial conditions, e_0 and m_0, are both strongly negatively correlated with σ and δ along this line; for example, in the e_0 profile, as e_0 increases, σ and δ decrease.

Away from the minimum (red diamond) in the positive/negative direction, profile values increase along the (σ, δ) line. Once profiles reach the edge of the

positive quadrant in the (σ, δ) plane, they move along $\delta = 0$ until they cross the line $\delta = \delta^{\diamond}(\sigma, \alpha, \kappa_2)$. They then switch to the edge $\sigma = 0$, and enter a region of bistability between escape and equilibrium steady states (blue region above $\delta = \delta^{\diamond}(\sigma, \alpha, \kappa_2)$), where elimination behaviour is not possible. Due to parameter dependencies, other parameters adjust to compensate for the constraint on the parameter at the edge of the parameter space; parameter dependencies thus appear different on the edge.

Once other parameters cannot compensate sufficiently, such as in regions with different qualitative behaviour, profile values increase sharply. This may lead to CIs becoming sharply bounded (see profiles of e_0 and m_0 in Fig. 4.4). This bound lies just before (for $\xi \leq 0.05$) or approximately where the profile enters the bistable region of escape and equilibrium (for $\xi \geq 0.1$, not shown); this is expected as it is difficult to find a fit in a region of parameter space in which the model does not exhibit tumour elimination. Similarly, as e_0 is negatively correlated with σ and δ, we observe sharp increases in the profile value when the edge of phase space $e_0 = 0$ is reached (see profiles for σ or δ in Fig. 4.4).

Sources of bias in MLEs and boundedness of confidence regions. We have seen that bias in MLEs and boundedness of CIs of e_0, σ and δ are observed particularly when the quality and quantity of data are poor. We attribute this to the combination of informativeness of data, parameter dependencies, physically realistic parameter bounds and the forward model's bifurcation structure.

When data are very informative and parameters highly identifiable (e.g. for the fully observed system with $\xi = 0.03$) the confidence region is small. Crucially, it is considerably smaller than the intersection of the physically realistic parameter region and the region with the matching bifurcation structure. Henceforth we will call boundaries of such small, well-defined confidence regions "data-based" bounds.

Linear dependencies among the parameters mean that as we move along the

profile of a particular parameter to the edges of the confidence region, other parameters also vary. In this way the profiles eventually reach the edge of the physically realistic region for a particular parameter or a bifurcation curve, on the other side of which lies a parameter region in which qualitative model behaviour does not match the data. When either of these two events occur, the profile value is likely to increase at a faster rate, and exceed the confidence threshold, which leads to the confidence region becoming abruptly bounded. This explains the transformation of profiles from a U shape to an L or bathtub shape as the noise level increases, and boundedness of regions of uncertainty for poorly identifiable parameters. Moreover, since the resulting CIs are not centred around the true parameter value, the bias in the MLE tends to be in the direction of the centre of the CI, which may be particularly influential when there is little curvature in the profile within the confidence region due to poor data.

Forward model solutions. Forward model solutions for parameter values within the 95% pointwise CI of the profile of σ show that, for different elimination datasets, tumour dynamics are identified to a good degree of accuracy as the range of solutions around the data is small (see Fig. 4.6). Effector dynamics are identified well for $0 < \tau < 5$, but uncertainty increases for $\tau \geq 5$ when effector data are not considered and noise levels increase. By contrast, the tumour is eliminated and remains so beyond the time period of data, which suggests elimination tumour dynamics are easily identifiable.

Summary. Despite strong correlations between model parameters, elimination data define a unique U-shaped region within the physically realistic parameter space. This region extends along a straight line in parameter space, and, as quality and/or quantity of the data reduces, its shape elongates, flattens, while remaining bounded. Together with small uncertainties in predicted solutions, this suggests a degree of PI in this regime. However, for poor quality data the source of PI is

related to parameter dependencies, physical parameter constraints and the model's bifurcation structure, which may hinder identification of the MLE to better accuracy than the size of the confidence region.

4.4.2.2 Case study 4.2: escape regime

We now detail results for data exhibiting escape behaviour. Unless otherwise stated, the results refer to case study 4.2a. Since the results for case study 4.2b are similar, we only consider it briefly when describing the sensitivity of forward model solutions that are close to the separatrix.

Profiles for the fully observed system. Parameters e_0, m_0 and σ exhibit good identifiability as their profiles are U-shaped, have small errors and small CIs relative to the true parameter values (see Fig. 4.8a). By contrast, parameter δ has an L-shaped profile and a larger CI (of width 0.048, which compares with 0.018 for σ). The global minimum (MLE) for δ lies outside the computation range, at $\delta = \mathcal{O}(10^{-12})$, whereas the true value is $\delta = 0.009$. Taken together, these results indicate poor identifiability of δ, a trend which extends to other datasets from regime 4.2a.

Within CIs, the profiles span bistable parameter regions (see bottom panel of Fig. 4.7), where, for certain initial conditions, the model exhibits growth to a large-tumour steady state. While the size of the CI for parameter m_0 increases linearly with increasing observation noise, CI growth with increasing noise is minimal for other parameters and strictly bounded within the described bistable parameter regions (see Fig. 4.8), in a similar way as for elimination data.

In the sense of finiteness of CIs, we conclude that parameters are identifiable for data with $\xi \leq 0.2$ and observation model $\mathbf{h}_2(\cdot)$, but there is some evidence that parameter δ may be difficult to identify. We also note that profile shapes outside CIs may impact the ability to recover parameters in practice; we discuss this later.

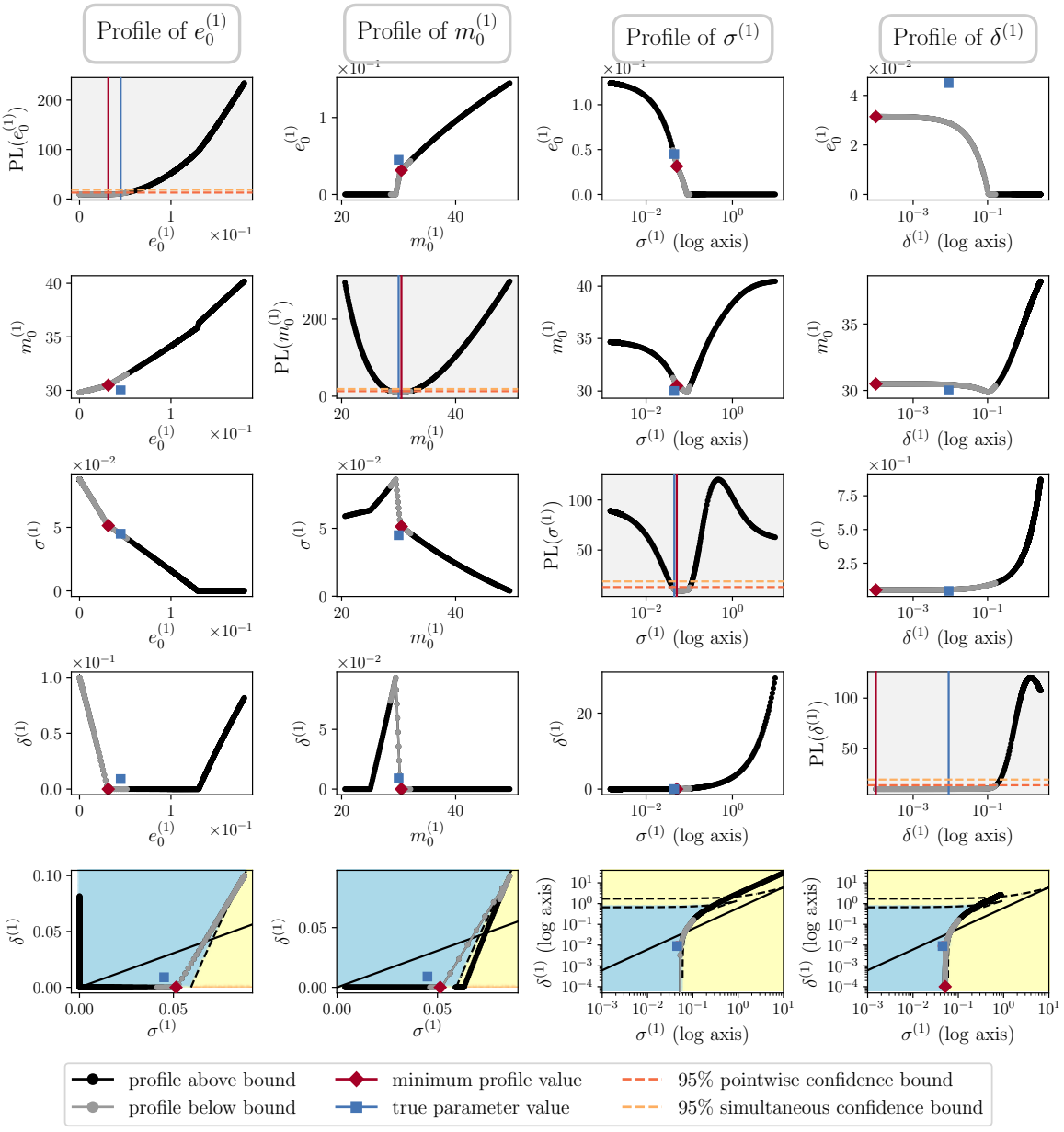


Figure 4.7: Series of plots showing profile likelihoods of $V(\boldsymbol{\theta})$ for parameters $\boldsymbol{\theta}_{\text{sub}} = [e_0, m_0, \sigma, \delta]^T$ given **escape** data (**case study 4.2a**), observation model $h_1(\cdot)$ and observation noise standard deviation $\xi = 0.03$. Diagonal plots with light grey background show the computed profile values as the parameter of interest varies, while off-diagonal plots in each column show how other (nuisance) parameters vary along the profile in the same column. Plots in the bottom row show (σ, δ) bifurcation diagrams overlaid with profiles. Parts of profiles with values below the 95% pointwise confidence threshold $\chi^2_{1,0.95}$ are plotted in grey, whereas those that exceed it are plotted in black. For interpretation of colours in the bifurcation diagrams in the bottom row, see Fig. 4.3b. The shown profiles for σ and δ are double-valleyed, while those for e_0 and m_0 appear U-shaped, and only explore the global minimum. The profiles trace a straight line in (σ, δ) space that is parallel to the bifurcation curve on which the model's large-tumour steady states undergo a saddle-node bifurcation. The confidence intervals span the bistable (blue) parameter regions. Expect for δ, MLEs correspond well to the true, data-generating parameter values.

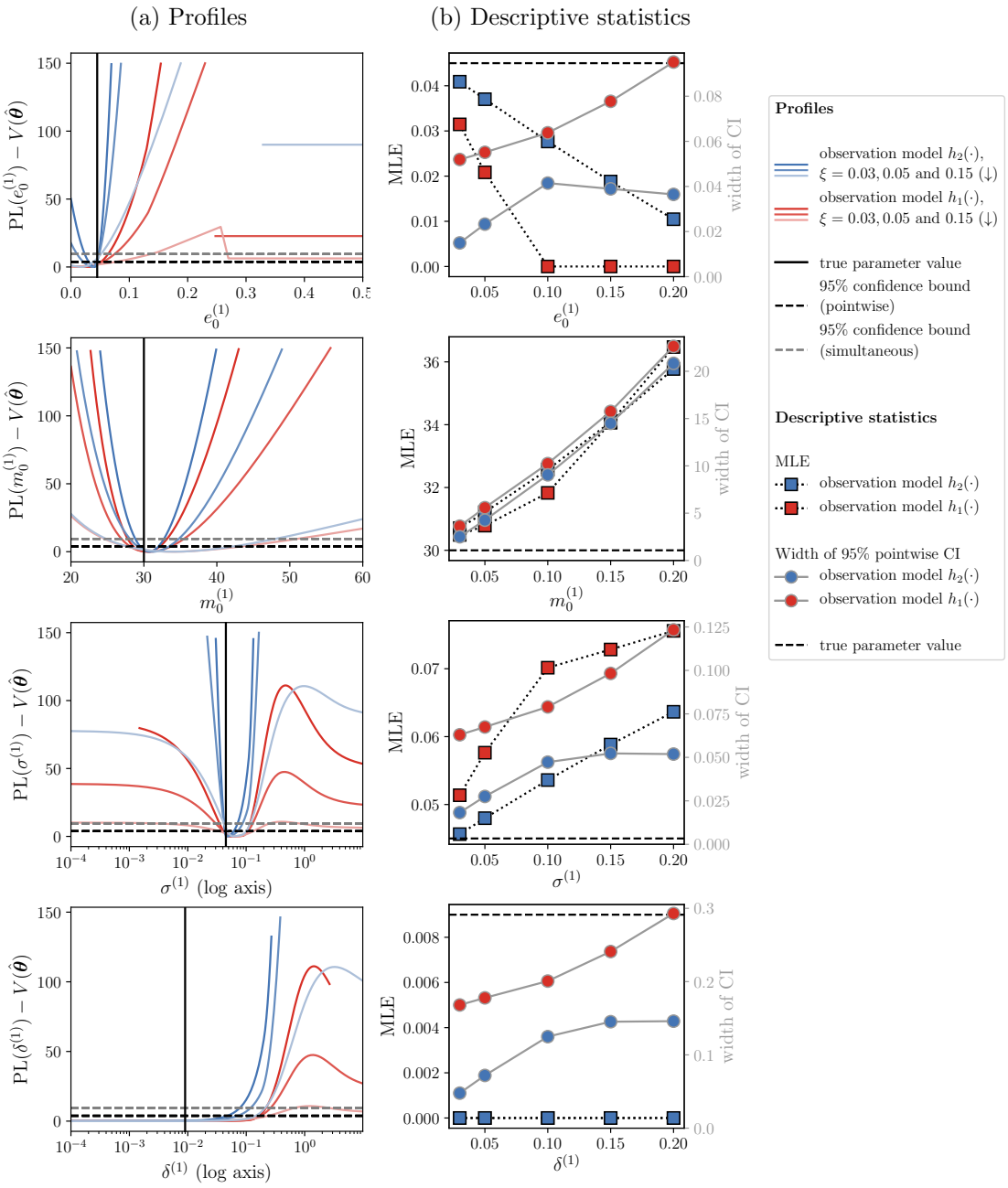


Figure 4.8: Series of plots showing the impact of observation noise standard deviation ($\xi \in [0.03, 0.2]$) and observation models with and without effector cell measurements ($\mathbf{h}_2(\cdot)$ and $h_1(\cdot)$) on profile likelihoods for data generated from the **escape** regime (**case study 4.2a**). Panel (a) shows profile values translated by $V(\hat{\boldsymbol{\theta}})$ as parameters of interest in $\boldsymbol{\theta}_{\text{sub}} = [e_0, m_0, \sigma, \delta]^T$ individually vary, while panel (b) shows how the corresponding maximum likelihood estimates (MLEs) and the size of confidence intervals (CIs) vary with increasing ξ, split by different observation models. Curves and points plotted in blue correspond to observation model $\mathbf{h}_2(\cdot)$, while those in red correspond to $h_1(\cdot)$. Profiles for σ and δ have a double-valleyed shape, and flatten as noise level increases. When effector data are unavailable, the difference in profile value between the local minima decreases, and, as noise levels increase, the simultaneous confidence region becomes disjoint. Eventually, as the two disjoint regions (valleys) merge, the confidence region becomes unbounded. Profiles for e_0 exhibit discontinuous jumps to the local minimum.

Variation of nuisance parameters and their impact on profiles. Along the σ profile, within the CI, σ and δ are mostly linearly dependent (except close to $\delta = 0$; see Fig. 4.7). This is similar to the elimination regime, except that the slope of the line on which the profile lies is different; the line is approximately parallel to the boundary of the bistable region on which the determinant of the cubic is zero ($\Delta = 0$) and a saddle-node bifurcation occurs. As σ decreases, the profile collapses onto $\delta = 0$, influencing the CI as in the elimination regime; at $\xi = 0.03$, the lower boundary of the CI corresponds to where δ hits the edge of the physically feasible parameter space, and moves only slightly as noise levels increase. The CI for σ is also sharply bounded above. At higher noise levels and/or for a partially observed system, this is again due to the negative correlation of σ and δ with e_0; as σ increases, e_0 decreases approximately linearly until $e_0 = 0$ at $\sigma \approx 0.1$, which is where the upper bound of the CI lies.

Bifurcation structure and the distant local minimum. As σ increases, the sharp increase in the profile corresponds to it leaving the CI and entering a region in which a good fit is difficult to find. This happens within the region of bistability of escape and equilibrium steady states (see bottom panel of Fig. 4.7) and beyond it, when the profile enters the monostable region (yellow) above the zero-discriminant curve. In the monostable region the profile attains a local maximum and then decreases, extending into another flat valley containing a very distant local minimum with $\sigma, \delta = \mathcal{O}(10^3)$; this minimum is difficult to identify due to the flatness of $V(\boldsymbol{\theta})$. We conclude that the double-valleyed profile shape reflects the existence of disjoint regions in the bifurcation diagram, in each of which there is a stable, large-tumour steady state solution (i.e. escape). Within these regions the profile is flat; it is steep when it crosses bifurcation curves that impact model behaviour in the region of the phase plane relevant to the data (i.e. close to the emerging small-tumour steady state where the trajectory starts, and to the large-tumour steady states

where the trajectory ends) and/or when the edge of the physically realistic region of parameter space is reached, as described above.

Due to the strong positive correlation between σ and δ, the profile for δ has a similar double-valleyed shape to that of σ. We cannot observe this shape of the likelihood surface in the same way from the profiles of e_0 and m_0. These mainly explore the likelihood surface in the neighbourhood of the MLE. However, for less informative data, as the profiles move away from the MLE, they may become attracted by the distant minimum, and may suddenly jump to it (e.g. profile of e_0 for $\xi = 0.15$ and $\mathbf{h}_2(\cdot)$, or for $h_1(\cdot)$; see Fig. 4.8). Thereafter the profiles are completely flat, because optimisation terminates when σ and δ reach upper optimisation bounds (see Appendix C). While this means that the profile does not reach the distant minimum exactly, the value of $V(\boldsymbol{\theta})$ after the jump is similar to that in the vicinity of the local minima in the profiles for σ and δ, deeming our computations sufficient to determine parameter identifiability from such profiles. We note also that this demonstrates how the presence of a local minimum may prevent the optimiser from finding the MLE for certain starting guesses, even for well-constrained confidence regions.

Profiles for the partially observed system. Removing availability of effector data strongly influences parameter identifiability in the escape regime. Profiles become considerably flatter inside and outside the CI compared to the fully observed case, but they retain their shape. For parameters σ, δ and e_0, profile values at the distant local minimum fall close to or below the confidence thresholds $\chi^2_{0.95,1}$ and $\chi^2_{0.95,4}$ for $\xi \geq 0.15$ (see Fig. 4.8), and for e_0 they are more easily attracted by the distant local minimum. If the profile falls below the threshold the second time, the confidence regions become disjoint/bimodal. If the peak between the two valleys also takes values below the threshold, the disjoint regions merge, and the CI may be very large and potentially unbounded (see red profiles of σ and δ in

Fig. 4.8), implying practical non-identifiability. Our computed profiles show that this happens for $\xi > 0.15$, if we consider the simultaneous threshold $\chi^2_{0.95,4}$. For escape regime 4.2b, which has similar double-valleyed profiles, this already occurs at lower noise levels (see Fig. D.1).

Forward model solutions and the impact of proximity to the separatrix. For a fully observed system and low noise, forward model solutions within the CI correspond to the data well (see Fig. 4.10a). The best-fitting solution (red) and the data-generating solution (black) agree quantitatively and qualitatively, tending to a large-tumour steady state; this also holds beyond the time window of the data ($\tau > 10$), where variability is small. As the level of observation noise increases and/or only tumour data are considered, the best-fitting solution does not agree qualitatively with the data-generating solution (see Figs. 4.10b and 4.10c). While the two match quantitatively for $\tau < 10$, they diverge at later times; the best-fitting solution and other solutions within the CI tend to the tumour-free steady state, rather than to the large-tumour one, with large uncertainty in the predicted dynamics for escape data from regime 4.2a.

We postulate that finding such closely fitting solutions with different qualitative behaviour is possible due to the model's bistable and bifurcating structure. For example, solutions may start with slightly different initial conditions, which lie on different sides of the phase plane separatrix, and will follow the separatrix for some time before departing from it and converging to different steady states (see Fig. 4.9). The same would apply to slightly different parameter values, which lie on different sides of a bifurcation curve. Both situations would explain well-fitting solutions that tend to different steady states, particularly when data have been generated using parameters/initial conditions that lie close to a bifurcation curve/separatrix.

To demonstrate how proximity to the separatrix affects model predictions, we consider data generated from parameter regime 4.2b, in which e_0 lies further away

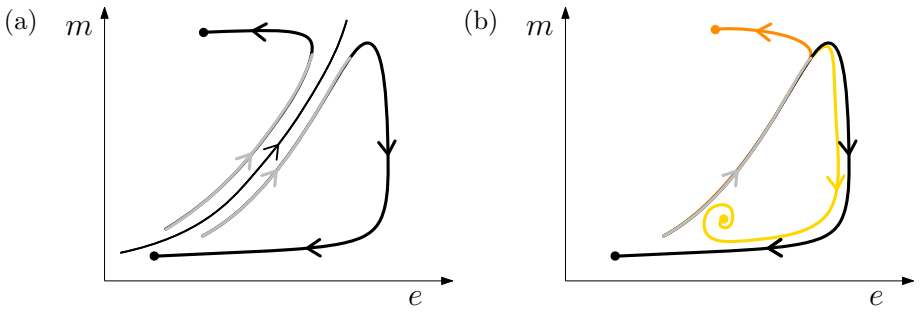


Figure 4.9: Schematics showing how a dataset may fit well to trajectories with different long-term behaviour as they converge to different steady states. Plot (a) shows a phase plane with a separatrix and two trajectories (bold curves) that are both initially similar when they follow the separatrix. Plot (b) shows trajectories for different parameter sets (black, orange, yellow) close to a bifurcation that are also all similar in the initial time period.

from the separatrix than in regime 4.2a. We find that parameter identifiability in the two regimes is very similar, with parameters in regime 4.2b being marginally less identifiable (see Fig. D.1). Despite this, forward model solutions within the confidence region of the σ profile in regime 4.2b indeed demonstrate considerably better qualitative and quantitative agreement with the data-generating solution than in regime 4.2a (see Figs. 4.10d–f). Uncertainty in tumour cell dynamics is small even when effector cell measurements are not considered and noise levels are high. Effector dynamics are identified to a good degree of accuracy for the fully observed system, whereas, without effector data, variability increases, but the solutions remain qualitatively close to the true dynamics.

Our results for case studies 4.2a and 4.2b highlight that proximity to a separatrix can significantly affect the uncertainty in predicted solutions for some datasets. Away from the separatrix, the CIs are less likely to span regions with long-term qualitative behaviour that does not match the data, and more likely to recover solutions which accurately reproduce the true escape-type dynamics.

Summary. For the fully observed system, CIs are finite for all parameters which means that, in theory, we can recover them within the width of their CIs, but bias in estimates can be large, particularly for parameter δ. An influential local minimum is detected in the profiles, affecting modality of the likelihood. Profiles

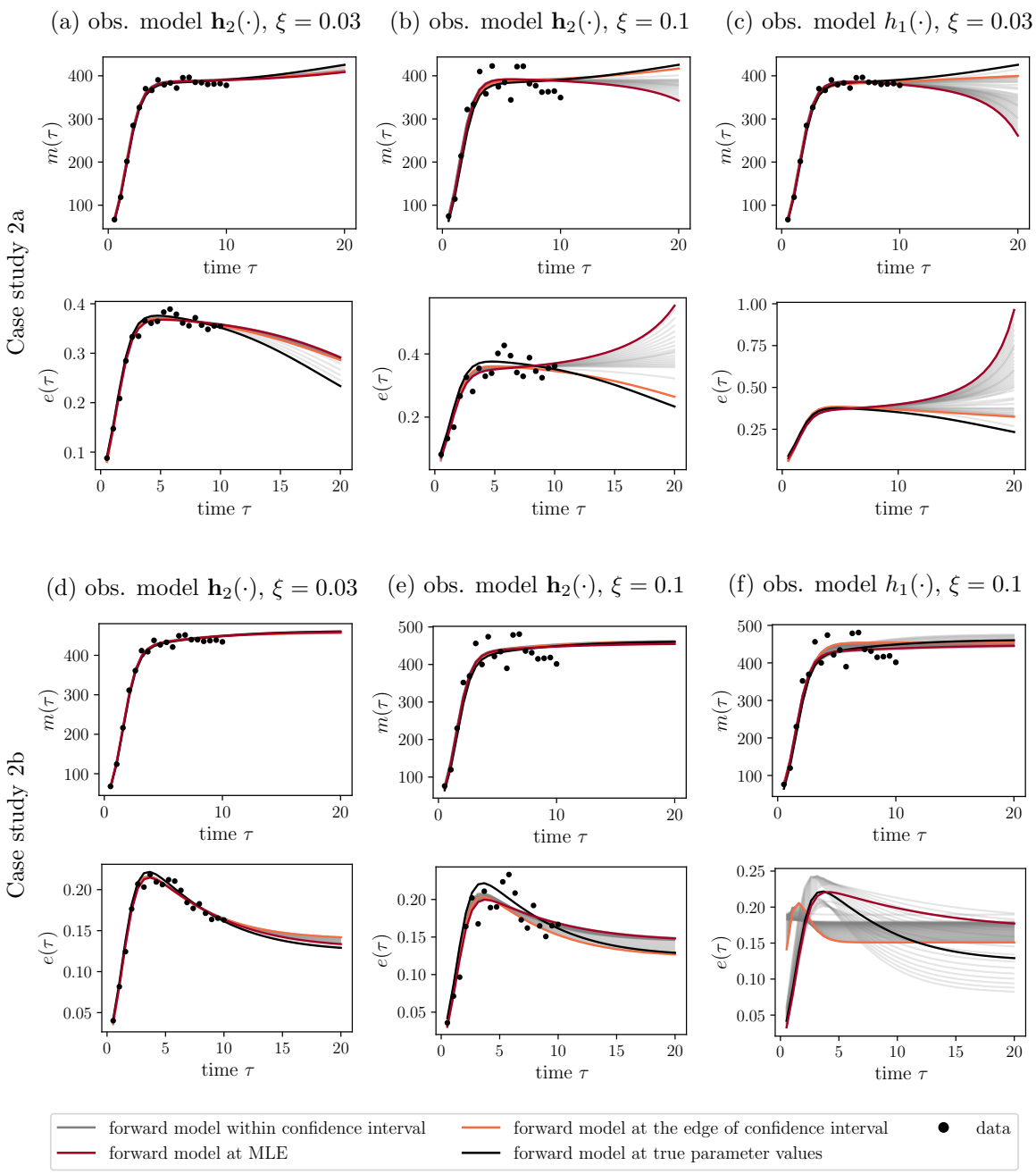


Figure 4.10: Series of plots showing the sensitivity of forward model solutions for parameters within the 95% pointwise confidence interval for σ given **escape** data. Upper panel plots correspond to **case study 4.2a**, while those in the bottom panel correspond to **case study 4.2b**. Recovery of dynamics corresponding to the escape stage of immunoediting is particularly sensitive to whether the data-generating regime lies in proximity of a phase space separatrix. They are successfully recovered for data generated from a regime that is away from the separatrix (4.2b), whereas the predicted forward model solutions show great variability and divergence to different steady states for data from a regime that lies close to the separatrix (4.2a).

exhibit high sensitivity to the availability of effector cell measurements, and immune system parameters may become unidentifiable at higher noise levels. The recovered forward model solutions can be highly uncertain for parameters in the vicinity of a phase plane separatrix. Together, this suggests poorer overall identifiability for escape data compared to elimination data.

4.4.2.3 Case study 4.3: equilibrium regime

Identifiability for the fully observed system and low noise levels. For $\mathbf{h}_2(\cdot)$ and $\xi = 0.03$, all profiles have unique minima close to the true parameter values (see Fig. 4.13). The CIs are comparable in size to those for the elimination and escape data. Uncertainty in the predicted dynamics within the CI of σ is negligible (see Fig. 4.12a), which implies good identifiability of parameters and dynamics in this regime.

Identifiability for high noise level and/or partially observed system. Profiles in the equilibrium regime are more resilient to decreased quality/quantity of data (see Fig. 4.13). There are only small differences compared to when effector data are not included. Increasing observation noise decreases the accuracy of recovery for all parameters. Sensitivity of parameter uncertainties to noise is consistently greater than to omission of effector data. Together, this demonstrates that tumour data alone provide a lot of information about the tumour-immune dynamics. Indeed there is a negligible change in the range of forward model solutions along the σ profile for the inverse problem with/without effector data (see Figs. 4.12b and 4.12c). We explain this by noting that tumour-immune interactions have a large influence on tumour growth dynamics in this regime; the effector cells cause the tumour size to undergo damped oscillations as it relaxes to an immune-mediated equilibrium. As the tumour is constantly responding to changes in effector cell numbers, measurements of tumour size oscillations provide sufficient information to recover parameters related to the immune system (and hence effector

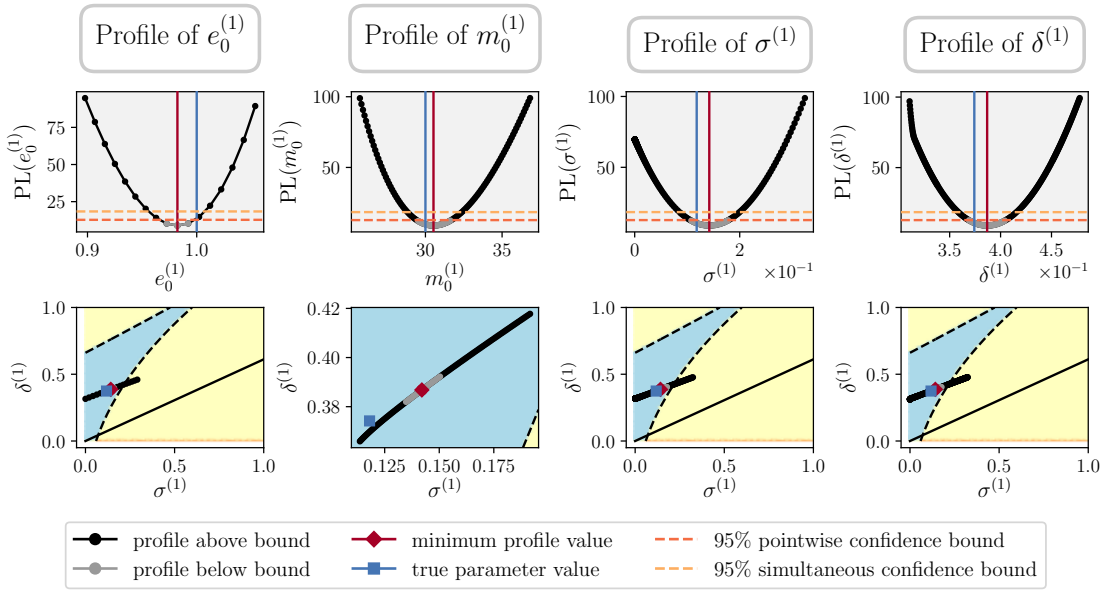


Figure 4.11: Series of plots showing profile likelihoods of $V(\boldsymbol{\theta})$ for parameters $\boldsymbol{\theta}_{\text{sub}} = [e_0, m_0, \sigma, \delta]^T$ given **equilibrium** data (**case study 4.3**), observation model $h_1(\cdot)$ and observation noise standard deviation $\xi = 0.03$. Top row plots show the computed profile values as the parameter of interest varies. Bottom row plots show (σ, δ) bifurcation diagrams overlaid with profiles. Parts of profiles that have values below the 95% pointwise confidence threshold $\chi^2_{1,0.95}$ are plotted in grey, and those that exceed it in black. For interpretation of colours in bifurcation diagrams in the bottom row, see Fig. 4.3b. The shown profiles are U-shaped and trace a straight line in (σ, δ) space, parallel to the transcritical bifurcation line on which the model's small-tumour steady state emerges. MLEs mostly correspond well to the true data-generating parameter values.

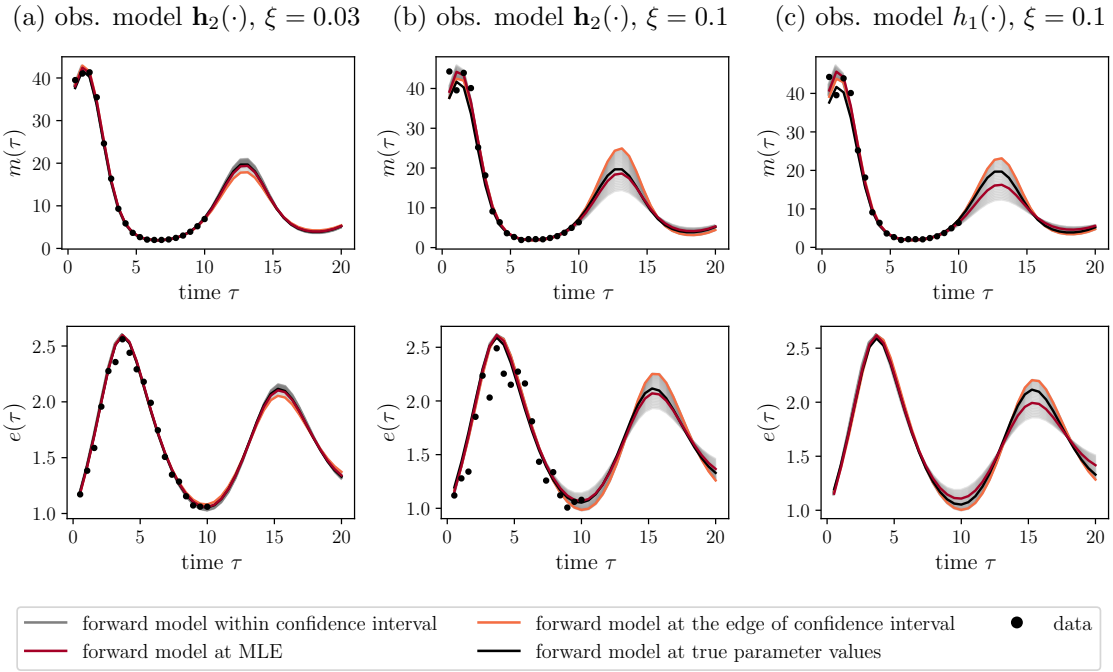


Figure 4.12: Series of plots showing the sensitivity of forward model solutions for parameters within the 95% pointwise confidence interval for σ given **equilibrium** data (**case study 4.3**). Dynamics corresponding to the equilibrium stage of immunoediting are successfully recovered for different quality/quantity of measurements.

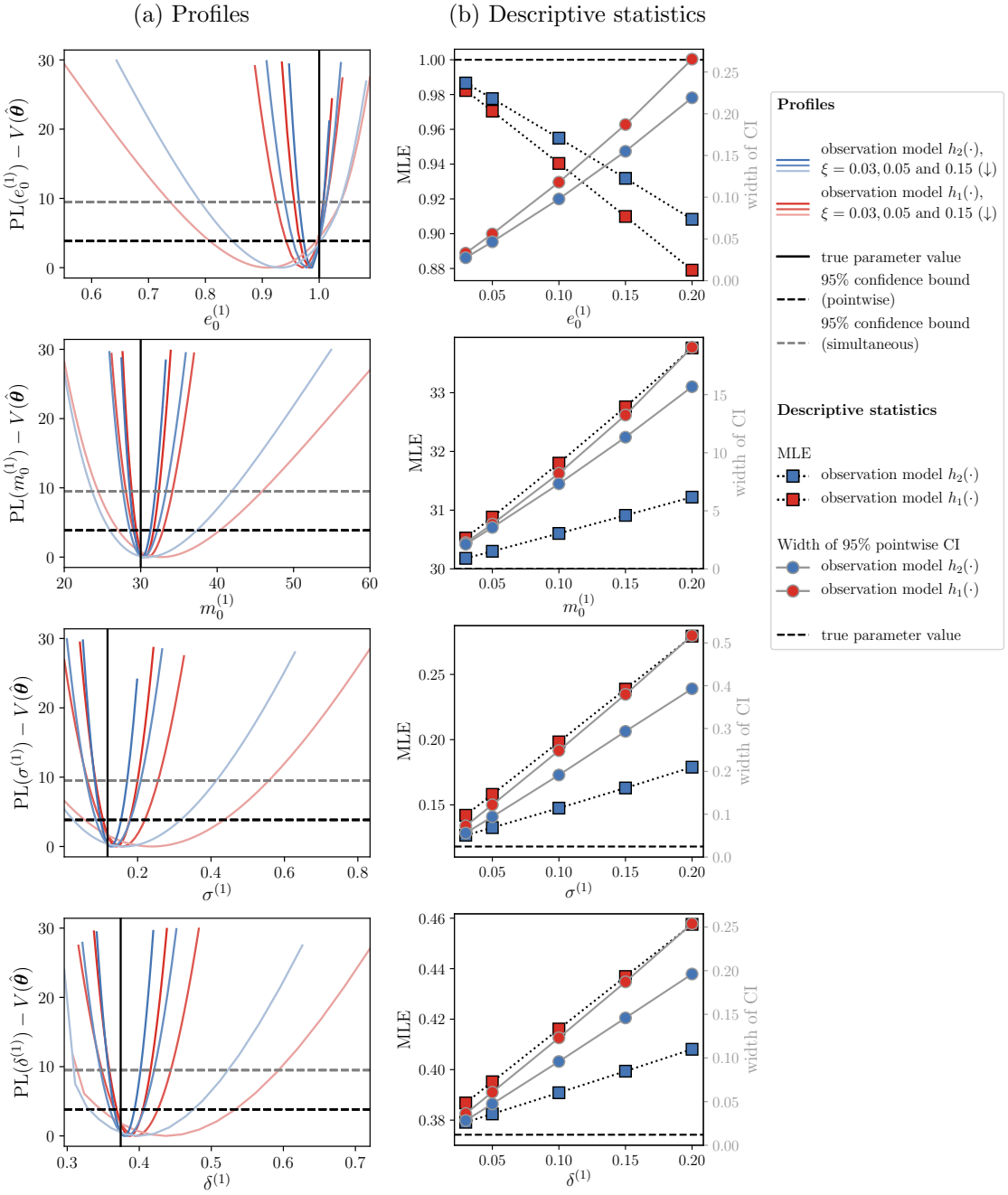


Figure 4.13: Series of plots showing the impact of observation noise standard deviation ($\xi \in [0.03, 0.2]$) and observation models with and without effector cell measurements ($\mathbf{h}_2(\cdot)$ and $h_1(\cdot)$) on profile likelihoods for data generated from the **equilibrium** regime (**case study 4.3**). Panel (a) shows profile values translated by $V(\hat{\boldsymbol{\theta}})$ as parameters of interest in $\boldsymbol{\theta}_{\text{sub}} = [e_0, m_0, \sigma, \delta]^T$ individually vary, while panel (b) shows how the corresponding maximum likelihood estimates (MLE) and the size of confidence intervals vary with increasing ξ, split by different observation models. Curves and points plotted in blue correspond to observation model $\mathbf{h}_2(\cdot)$, while those in red correspond to $h_1(\cdot)$. As noise levels increase, profiles flatten and confidence intervals increase. The effect of unavailability of effector data is small.

cell dynamics) to a similar level of accuracy as when effector cell data are also provided.

Variation of parameters and their relation to the bifurcation structure. Profiles for e_0, σ and δ trace straight lines in (σ, δ) space, as in previous case studies, with a positive correlation between σ and δ, and negative correlations of the two with e_0. As in the elimination regime, the traced line in (σ, δ) space is approximately parallel to the line $\delta = \delta^{\diamond}(\sigma, \alpha, \kappa_2)$, on which the small-tumour steady state undergoes a transcritical bifurcation with the zero-tumour steady state and becomes physically unrealistic and unstable; this suggests a way in which the bifurcation structure of the model impacts the MLE landscape. Within the CI, profiles span the bistable region of the (σ, δ) parameter space in which both the equilibrium and escape steady states are stable (blue above the line $\delta = \delta^{\diamond}(\sigma, \alpha, \kappa_2)$), and, for high noise, also the region in which only the equilibrium steady state is stable (yellow). Since the true parameter values lie further from the boundaries of the physically realistic parameter space than for regimes 4.1 and 4.2, and since the equilibrium regime exhibits good identifiability, the physical boundaries and the bifurcation structure have less of an effect on the size of the confidence region. Only at high noise levels, when the CIs extend to $\sigma = 0$ (see profiles of e_0 and δ in Fig. 4.11), do the CIs have sharp bounds; otherwise they are "data-based".

Summary. Model parameters and tumour-immune dynamics are robustly identifiable from equilibrium-type data. Tumour data alone contain sufficient information about the full system for parameter identifiability.

4.5 Discussion and conclusions

In this chapter we investigated identifiability of the quasi-steady-state-approximated (QSSA) model of tumour-immune interactions by Kuznetsov et al. [88]. Our aim was to identify experimental settings under which parameters, and system dynamics, of

this bistable and bifurcating model can be recovered. We characterised structural and practical aspects of model identifiability for tumour growth curves, which correspond to three stages of immunoediting (tumour elimination, equilibrium and escape) [47]. Our study was motivated by AZ's data from an immunotherapeutic preclinical study (see Fig. 1.1), as this and similar datasets [44, 88, 149] typically show tumour responses that are consistent with immunoediting.

Structural identifiability. Using differential algebra-based methods, we showed that the QSSA model is structurally identifiable for observables of tumour numbers, so all model parameters and initial conditions can be uniquely determined from tumour growth data that are noise-free and continuous in time (this automatically implies identifiability when both tumour and effector data are observed). While structural identifiability does not guarantee parameter identifiability from data of lesser quality and quantity, it is a prerequisite for practical identifiability.

Practical identifiability. Using profile likelihoods [125, 134], we investigated a series of inverse problems in which a subset of kinetic parameters and initial conditions from the QSSA model are estimated by applying maximum likelihood estimation (MLE) to synthetic data generated from the model. These data varied in their qualitative dynamics, quality and quantity. The accuracy, robustness and uncertainties of the parameter estimates and the predicted model dynamics varied markedly for different datasets.

We computed likelihood-based confidence regions in each dimension of parameter space, extended to include the initial conditions. For good quality, informative data capturing tumour and effector dynamics (e.g. parameter regimes exhibiting equilibrium or elimination), confidence regions are small and narrow U-shaped valleys, which are well-constrained by the data, findings consistent with the model's structural identifiability.

For poor quality and/or poorly informative data, the bottom of the valleys flattens, extending along correlations between specific parameters associated with the immune system (i.e. the supply and death rates of effector cells, σ and δ, and the initial number of effector cells, e_0). The valleys span regions of the bifurcation diagram with model behaviour that matches the data. They sometimes extend to their edges, and, rather than being bounded solely by the data, they appear sharply bounded by physical constraints and/or curves on which the behaviour of the matching model changes markedly. When any of the parameters are close to the boundary of the matching region, the confidence regions appear small when compared to those associated with parameters that are away from the boundary. MLEs also become biased towards the centre of the bounded confidence region, particularly when local identifiability within the flat valley is poor. While such confidence regions may be easy to identify, MLEs could be difficult to determine to better accuracy.

Our results thus suggest that confidence regions are heavily reliant on the combination of parameter relationships, the bifurcation structure of the model and physical constraints on model parameters. We conclude that, for complex models, the size of (finite) confidence intervals alone may not be the most useful predictor of how well the true parameters may be recovered. The geometry of the parameter space and sources of uncertainty vary between parameter regimes, affecting the accuracy of parameter estimates in different ways. For example, identifiability given data from the regime exhibiting tumour escape is poor, even though the confidence regions are small compared to those from other regimes. This is because the profiles are very flat within the confidence intervals, which appear to be constrained by physical constraints and the bifurcation structure, rather than the data. Extra care is therefore needed with interpreting inference results in such models.

Good identifiability for equilibrium and elimination data, poor identifiability for escape data. We compared and characterised identifiability for different datasets by the way in which profiles and forward model solutions change as quality and quantity of data vary. We found that data from regimes in which tumour-immune interactions have a large influence on model dynamics are most informative for parameter estimation and enable correct recovery of tumour-immune dynamics. The escape regime is the least identifiable, followed by the elimination regime, while the equilibrium regime is the most identifiable of the three.

In Kuznetsov et al.'s model, multiple regions of the bifurcation diagram are consistent with each stage of immunoediting, as seen in Chapter 3. The interplay of this complexity, parameter dependencies and poorly informative and/or noisy data exhibiting tumour escape resulted in double-valleyed profile likelihoods. The two valleys correspond to two disjoint regions of parameter space, one consistent with the true data-generating parameter values. Given only tumour data, differences between model fits associated with global and the local minima (i.e. valleys) decrease. With increasing noise in the data, the minima give rise to disjoint confidence regions (i.e. bimodal likelihood), which eventually merge and produce unbounded confidence regions that imply practical non-identifiability. Recovery of immune parameters in the escape regime thus shows greater sensitivity to the lack of effector data than in the equilibrium and elimination regimes. We speculate this may be because data exhibiting logistic growth to a large-tumour steady state are largely insensitive to effector dynamics. However, this may also be due to small values of effector parameters. In future work it would be interesting to investigate identifiability of escape dynamics generated from other parameter regimes.

Due to the model's bistable and bifurcating structure, the predicted forward model solutions are sensitive to escape data generated close to the phase plane separatrix. In such cases model predictions are highly uncertain; at long times they may exhibit a wide range of behaviours, some qualitatively different from

the data-generating solution. This is because it can be difficult to distinguish between model solutions that initially follow the separatrix but eventually converge to different steady states. We anticipate similar findings for regimes generated close to bifurcation curves, but postpone this for future work.

Context of identifiability in systems with bifurcations. Roesch and Stumpf [142] studied parameter inferrability for 1-D ODEs that correspond to the normal forms of co-dimension 1 bifurcations, and have one unknown parameter and one unknown initial condition. They found that regions of high curvature of the likelihood were closely related to the bifurcation structure, and that the level of noise and data density affected parameter inferrability quantitatively, but not qualitatively. Our observations for Kuznetsov et al.'s model are consistent with their findings, as we also find that the local and the global shape of likelihood landscapes are both important in systems with bifurcations. The work in this chapter extends their analysis to a system which undergoes multiple bifurcations (mainly co-dimension 1), has multiple unknown parameters and is partially observed. We showed that these properties add to the complexity of likelihood surfaces and our ability to characterise them.

Methodological challenges. We identified two main challenges for characterising parameter identifiability of Kuznetsov et al.'s model. The first was the dimensionality of its parameter space. We initially focused on four key parameters chosen so that the model exhibits tumour escape, equilibrium and elimination as they vary. We then used profile likelihoods to summarise the shape of the likelihood surface in one dimension by optimising over the other key parameters. In future work other parameters should also be considered.

The ease of profile visualisation and previous analyses of model behaviour enabled us to relate the profile shape to the model's bifurcation structure, highlighting our second challenge. Unlike the simple models considered in [142], for Kuznetsov

et al.'s QSSA model the parameter space cannot be split into unique regions, each yielding distinct qualitative behaviours. Local minima that emerge may result in disjoint confidence regions and a multimodal likelihood surface. This raises questions about the sufficiency of defining practical identifiability [134] by finiteness of confidence intervals. Multimodality also causes practical difficulties for methods that use the gradient of the log-likelihood to search parameter space, since, when starting in the basin of a local minimum, the gradient directs the parameter search towards the local minimum, making recovery of the global minimum difficult.

Connection to sensitivity-based identifiability and model "sloppiness". To determine the local parameter sensitivity of noise-free model outputs and local parameter dependencies we could have used sensitivity-based methods. Several methods use the sensitivity matrix to rank parameters according to how sensitive model outputs are to small changes in their values (see [111] for a review of sensitivity-based methods). Parameters ranked the lowest are considered unidentifiable and tend to be discarded. Our preliminary studies indicate that these methods would enable us to quantify poor sensitivity of escape curves to some of the immune parameters (see Appendix E). However, we note that while discarding/fixing parameters makes no difference locally, in a bifurcating system it does so globally as it can limit the range of model behaviour.

We also note that sensitivity varies between regions of parameter space, as does the spectrum of eigenvalues of the sensitivity matrix. For example, the largest eigenvalue spectrum is in the escape regime (see Appendix E). Gutenkunst et al. [62] described a large span of eigenvalues as model "sloppiness". It may result in large parameter uncertainties along eigenvectors associated with the smallest eigenvalues, but not necessarily affect the uncertainty of model predictions. Consistent with this, in profile likelihoods for Kuznetsov et al.'s model, we observed parameter dependencies, confidence intervals along those, and, in some cases, good

identifiability of model dynamics despite large parameter uncertainties, but the likelihood surfaces were U-shaped and had unique minima for *in silico* data. In future work we could more comprehensively characterise model "sloppiness" in different regions of parameter space, which is important as "sloppiness" may hinder the use of the model beyond the data used for calibration and in new contexts.

Methodological recommendations. We conclude that for bifurcating systems the ultimate challenge is methodological, how to relate the local and the global properties of the likelihood surface to characterise model identifiability. For a model deemed structurally identifiable, our recommendation is to use a combination of different methods where possible. Prior to profile likelihoods, sensitivity-based methods can be used to gain an intuition as to which parameters model outputs may be insensitive to or which are correlated. Profile likelihoods enable simultaneous exploration of the local and, to an extent, also the global structure for different datasets. However, global characterisation could be improved by using stochastic methods, such as the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) or Markov Chain Monte Carlo (MCMC), to identify local minima [136]. As suggested in [168], profiles can be computed around those minima and merged, which would provide a more integrated view of model identifiability.

Summary and final remarks. By performing a systematic parameter identifiability analysis using *in silico* data, we have shown that the rich bifurcation structure of Kuznetsov et al.'s QSSA model, parameter dependencies and constraints, and varying levels of informativeness, quality and quantity of data generated from different parameter regimes of the model, result in non-trivial log-likelihood surfaces, which may be multimodal. These findings suggest caution is needed with parameter estimation in complex systems.

Based on our analysis, we conclude that it would be difficult to robustly fit Kuznetsov et al.'s model to tumour growth data alone, especially when growth

is uncontrolled. In future, we could use simpler models to calibrate to such data, formulate the parameter estimation problem differently or improve experimental design and data collection. For example, we could calibrate the model using a population approach, where data from multiple individuals are fitted simultaneously, and some parameters are identical between individuals. We explore this in Chapter 5.

Chapter 5

Distinguishing hypotheses for dichotomous tumour growth curves

5.1 Introduction

In Chapter 4 we investigated the problem of fitting the QSSA model by Kuznetsov et al. [88] to data for a single individual. We now consider a parameter estimation problem where data from multiple individuals that exhibit dichotomous tumour growth are fitted simultaneously. We seek to establish whether alternative mechanistic hypotheses for the dichotomy in tumour responses to treatment, represented by variability in model parameters and/or initial conditions, can be distinguished based on their ability to explain the data.

Previous studies of the QSSA model [88, 137, Chapters 2 and 3] demonstrate that dichotomous tumour growth curves may be generated by the model in two ways: (i) by using the same initial conditions and parameter values that lie in different regions of parameter space, separated by bifurcations, or (ii) by using the same parameter values and considering initial conditions that lie in different regions of the phase plane, and that evolve to different steady states. In other words, dichotomous model solutions may be driven either by bifurcations or bistability.

Suppose we have data from two tumours which belong to different individuals and exhibit different qualitative behaviours; one tumour shrinks, while the other

grows uncontrollably. Motivated by the bifurcation structure of Kuznetsov et al.'s model, we propose two alternative hypotheses that may explain the dichotomy in these data (see Fig. 5.1):

Hypothesis 1 (H1): $\boldsymbol{\theta}^{(1)}_{3:} \neq \boldsymbol{\theta}^{(2)}_{3:}, \qquad \mathbf{x}^0(\boldsymbol{\theta}^{(1)}) = \mathbf{x}^0(\boldsymbol{\theta}^{(2)})$

The tumour growth curves have the same initial conditions, but are generated from parameter regimes which exhibit qualitatively different behaviours.

Hypothesis 2 (H2): $\boldsymbol{\theta}^{(1)}_{3:} = \boldsymbol{\theta}^{(2)}_{3:}, \qquad \mathbf{x}^0(\boldsymbol{\theta}^{(1)}) \neq \mathbf{x}^0(\boldsymbol{\theta}^{(2)})$

The tumour growth curves come from the same, bistable parameter regime, but were generated with different initial conditions.

The superscript i denotes individual i ($i = 1, 2$). The subscript denotes all model parameters except initial conditions (1st and 2nd entry of parameter vector $\boldsymbol{\theta}$).

For each hypothesis we formulate an inverse problem, where the unknown model parameters, $\boldsymbol{\theta}$, are split into two categories: population parameters, which are assumed fixed over the population, i.e. $\theta_j^{(\text{pop})} = \theta_j^{(i)}$ for all individuals i, and have a single copy $\theta_j^{(\text{pop})}$ in $\boldsymbol{\phi}$ (vector of all parameters to be inferred in the Bayesian scheme); and individual parameters, which vary between individuals and have a copy of $\theta_j^{(i)}$ in $\boldsymbol{\phi}$ for each individual i. Such inverse problems enable one to exploit the abundance of data and information from different individuals, without fixing specific parameter values, and, thus, imposing heavy constraints on the problem.

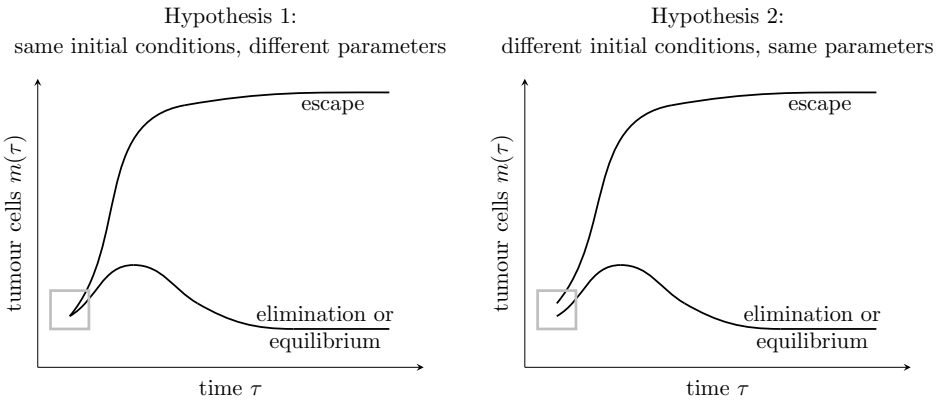


Figure 5.1: Schematic depicting two hypotheses via which the QSSA model by Kuznetsov et al. [88] may explain the dichotomy in tumour growth curves.

Recall that a popular approach to simultaneous model calibration using population data is *mixed effects modelling* (MEM), where individual and population model parameters are estimated [93]. Our approach and the inverse problems we formulate (as outlined above) are closely related to MEM. However, we make no assumptions about the distribution of individual parameters. We thus consider a less constrained problem than MEM, in order to explore whether fits according to different hypotheses are feasible, and avoid potential difficulties with non-convergence of MEMs that are difficult to explore within this framework. Our approach is fully Bayesian and enables us to systematically track and increase problem complexity, building upon maximum likelihood estimation (MLE) results for single curves from Chapter 4.

In this chapter, we aim to understand posterior landscapes for the population fitting inverse problems under the two proposed hypotheses, and to address the following fundamental question: Given dichotomous data, can we distinguish between the proposed hypotheses? We consider case studies for which dichotomous synthetic data are generated from the model according to one of the two hypotheses; model fits under both hypotheses are compared based on their predictive accuracy.

The chapter is organised as follows. We present the workflow for Bayesian inference and model comparison in Section 5.2. In Section 5.3, we present results of Bayesian inference for data from a single individual (as in Chapter 4) to inform inference for two curves. We then perform hypothesis comparison for synthetic data comprising two different curves, generated either from a bistable parameter regime, or from a bifurcating parameter regime. In Section 5.4 we discuss our findings and possible directions for future work.

5.2 Methods

We now introduce the methodology we use to compare hypotheses H1 and H2 for dichotomous synthetic datasets generated from Kuznetsov et al.'s QSSA model. We

consider two types of synthetic data, described in Section 5.2.1. The comparison workflow has four main components (see Fig. 5.2 and Sections 5.2.2–5):

(1) *Bayesian inverse problem formulation* under each hypothesis (H1 and H2, or true and alternative), and the choice of prior distribution,

(2) *Parameter inference* using Markov Chain Monte Carlo (MCMC) methods, and evaluation of sampling performance,

(3) Estimation of *out-of-sample predictive accuracy* for each posterior fit, and

(4) *Hypothesis testing* to determine whether the difference in predictive accuracy between fits according to different hypotheses is significant.

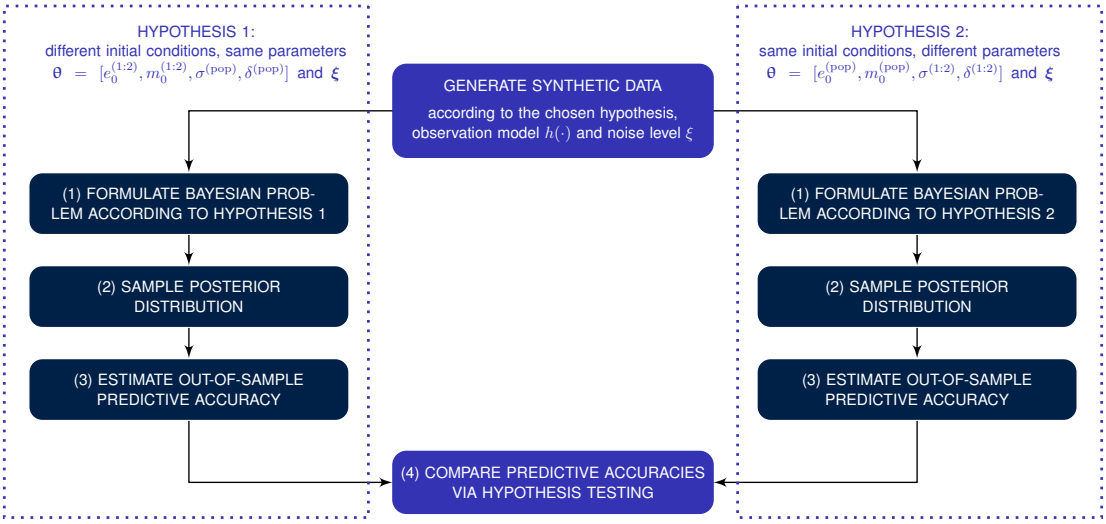


Figure 5.2: Diagram showing the workflow for comparing predictive accuracy of model fits to a synthetic dataset under Hypotheses 1 and 2.

5.2.1 Generation of synthetic data

In this chapter we study data from two individuals generated from the forward and observation models (see Section 5.2.2), using the procedure outlined in Chapter 4, and assuming there are no population parameters. To ensure a dichotomy in tumour responses, we choose parameter sets so that each individual's data exhibit different qualitative behaviour. Specifically, data of individuals 1 and 2 exhibit tumour elimination and escape, respectively.

Case study	Parameter regime	True hypothesis	Individual i	Individual parameters $\theta^{(i)}$ $\sigma^{(i)}$	$\delta^{(i)}$	$e_0^{(i)}$	$m_0^{(i)}$
5.1	bistable	2	1	0.045	0.009	0.24	30.0
			2	0.045	0.009	0.045	30.0
5.2	bifurcating	1	1	0.05	0.03	0.5	30.0
			2	0.05	0.6	0.5	30.0

Table 5.1: Parameter values used to generate data for case studies 5.1 and 5.2.

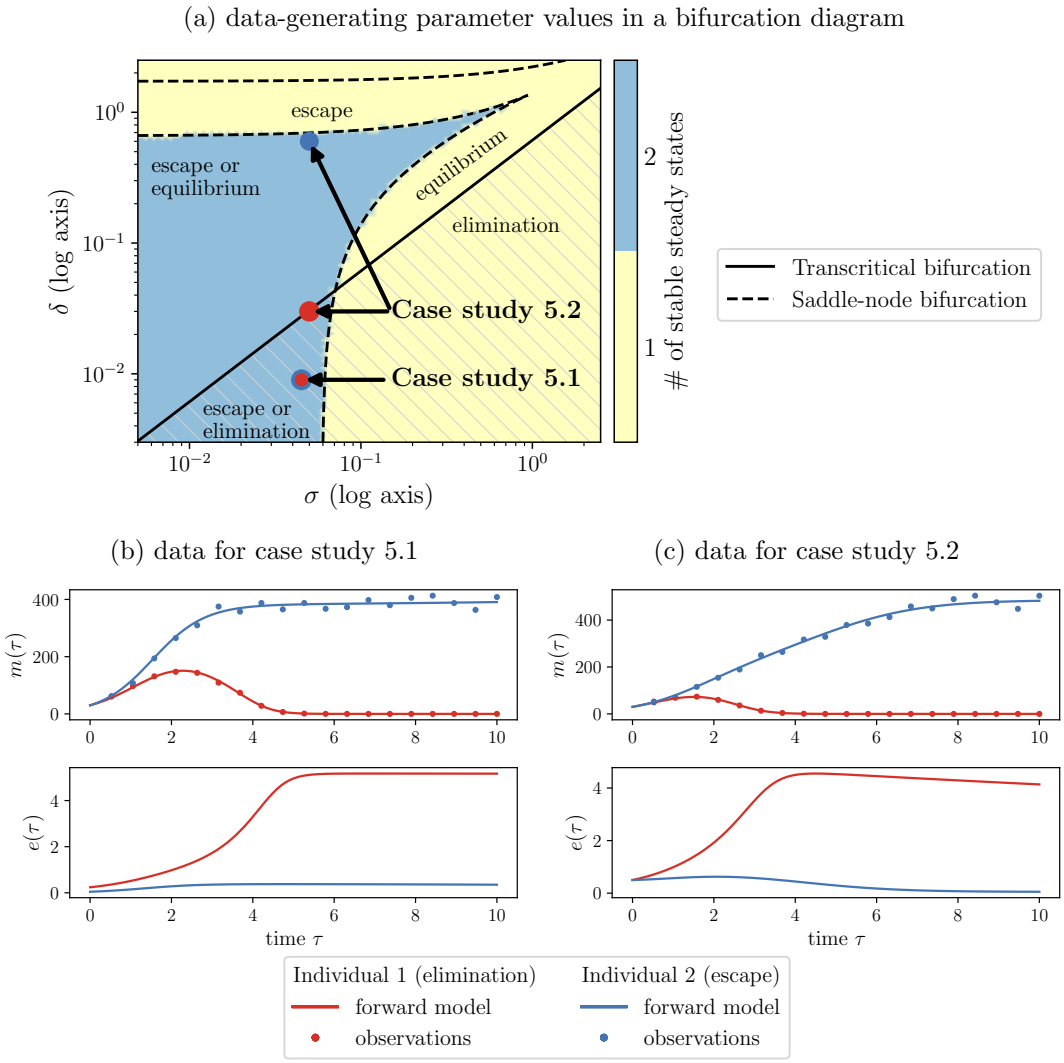


Figure 5.3: Plot (a) shows the (σ, δ) bifurcation diagram of the Kuznetsov model denoting disjoint regions of the (σ, δ) parameter space with qualitatively different model behaviour, with other parameters fixed at default values (see Table 2.2). Coloured point markers denote values of σ and δ used to generate datasets for case studies 5.1 and 5.2 (see Table 5.1). Blue and red point markers respectively denote parameter regimes that elimination-type data of individual 1 and escape-type data of individual 2 are generated from. Synthetic data, and data-generating forward model solutions, corresponding to these regimes, are shown in plots (b) and (c), respectively, for case studies 5.1 and 5.2.

We perform two case studies. For case study 5.1, data are generated according to H2 (considered true), with different values for e_0, the effector initial condition. For case study 5.2, data are generated according to H1 (considered true), and the two curves differ in either effector supply rate, σ, or effector death rate, δ. The parameter values used are stated in Table 5.1, and visualised in Fig. 5.3, together with exemplary datasets.

5.2.2 Bayesian parameter estimation problem

We define the inverse problems for fitting pooled datasets under each hypothesis by extending the MLE problem for parameters σ, δ, e_0 and m_0 from Chapter 4. We first adapt it to account for data from multiple individuals, and to allow for some population parameters, $\boldsymbol{\theta}^{(\text{pop})}$, which are shared between individuals. We then formulate the problem within a Bayesian framework in order to use efficient MCMC techniques to infer model parameters and their uncertainty.

Parameters. Let P be the number of individuals for which we have data. Let $\boldsymbol{\theta}^{(\text{pop})}$ be the vector of all population parameters, and $\boldsymbol{\theta}^{(i)}$ the vector of parameters unique to individual i. Let $\boldsymbol{\theta} = [\boldsymbol{\theta}^{(\text{pop})^T}, \boldsymbol{\theta}^{(1)^T}, \ldots, \boldsymbol{\theta}^{(P)^T}]^T$ be the stacked vector of all unknown model parameters. For example, when fitting under H1 we assume that σ and δ are population parameters, and initial conditions, e_0 and m_0, are individual parameters. For $P = 2$, this implies

$$\boldsymbol{\theta}^{(\text{pop})} = [\sigma^{(\text{pop})}, \delta^{(\text{pop})}]^T, \quad \boldsymbol{\theta}^{(1)} = [e_0^{(1)}, m_0^{(1)}]^T, \quad \boldsymbol{\theta}^{(2)} = [e_0^{(2)}, m_0^{(2)}]^T. \tag{5.1}$$

When fitting under H2, the parameter roles are reversed, and

$$\boldsymbol{\theta}^{(\text{pop})} = [e_0^{(\text{pop})}, m_0^{(\text{pop})}]^T, \quad \boldsymbol{\theta}^{(1)} = [\sigma^{(1)}, \delta^{(1)}]^T, \quad \boldsymbol{\theta}^{(2)} = [\sigma^{(2)}, \delta^{(2)}]^T. \tag{5.2}$$

Forward model. As in Section 4.2, we define the forward model for individual i as $\mathbf{u}^{(i)} = \mathcal{F}^{(i)}(\boldsymbol{\theta})$, where the forward model is the solution to Kuznetsov et al.'s ODE system (4.1) at times $\tau_1, \ldots, \tau_K$, i.e. $\mathbf{u}^{(i)} = [\mathbf{x}^{(i)}(\tau_1, \boldsymbol{\theta}), \ldots, \mathbf{x}^{(i)}(\tau_K, \boldsymbol{\theta})]^T$. Here

only parameters $\boldsymbol{\theta}^{(i)}$ and $\boldsymbol{\theta}^{(\text{pop})}$ are needed to compute the solution $\mathbf{u}^{(i)}$. For a pooled dataset, the forward model $\mathbf{u} = \mathcal{F}(\boldsymbol{\theta})$ includes a copy of the forward model for each individual, i.e. $\mathbf{u} = [\mathbf{u}^{(1)^T}, \ldots, \mathbf{u}^{(P)^T}]^T$, and $\mathcal{F}(\cdot)$ is defined accordingly.

Observation model. The vector of pooled observations is $\mathbf{s} = [\mathbf{s}^{(1)^T}, \ldots, \mathbf{s}^{(P)^T}]^T$, where $\mathbf{s}^{(i)}$ are observations for individual i as in Chapter 4. We assume that $\mathbf{s} = \mathcal{O}(\mathbf{u}) + \boldsymbol{\nu} = \mathcal{O}(\mathcal{F}(\boldsymbol{\theta})) + \boldsymbol{\nu}$, where

$$\mathcal{O}(\mathbf{u}) = [\underbrace{\mathbf{h}(\mathbf{x}^{(1)}(\tau_{k_1}, \boldsymbol{\theta})), \ldots, \mathbf{h}(\mathbf{x}^{(1)}(\tau_{k_N}, \boldsymbol{\theta}))}_{\mathbf{s}^{(1)}}, \ldots, \underbrace{\mathbf{h}(\mathbf{x}^{(P)}(\tau_{k_1}, \boldsymbol{\theta})), \ldots, \mathbf{h}(\mathbf{x}^{(P)}(\tau_{k_N}, \boldsymbol{\theta}))}_{\mathbf{s}^{(N)}}]^T. \tag{5.3}$$

Maps $\mathbf{h}(\cdot)$ are defined as in Chapter 4, and $\boldsymbol{\nu}$ is random measurement noise.

Likelihood. We again assume that $\boldsymbol{\nu}$ is Gaussian, with zero mean and standard deviations ξ_e or ξ_m, for effector and tumour measurements, i.e. $\boldsymbol{\nu} \sim \mathcal{N}(\mathbf{0}, \Xi)$, where the covariance matrix Ξ is diagonal with entries ξ_e^2 and ξ_m^2. The data

Parameter ϕ_i associated with	**Prior 1**	**Prior 2**	**Prior 3**
σ	Lognormal$(0.75, 1.5)$	Cauchy$(0, 0.5)$ on $(0, \infty)$	$\mathcal{N}(0, 1)$ on $(0, \infty)$
δ	Lognormal$(0.75, 1.5)$	Cauchy$(0, 0.5)$ on $(0, \infty)$	$\mathcal{N}(0, 1)$ on $(0, \infty)$
e_0	Lognormal$(0.75, 1.5)$	Cauchy$(0, 0.5)$ on $(0, \infty)$	Cauchy$(0, 1)$ on $(0, \infty)$
m_0	Lognormal$(0.75, 1.5)$	Lognormal$(0.75, 1.5)$	Lognormal$(0.75, 1.5)$
ξ_e, ξ_m	Lognormal$(-1, 1)$	Cauchy$(0, 1)$ on $(0, \infty)$	Cauchy$(0, 1)$ on $(0, \infty)$
Used for case studies	4.1, 4.2a & 4.3 (individual)	4.1, 4.2a & 4.3 (individual)	5.1 & 5.2 (two individuals)

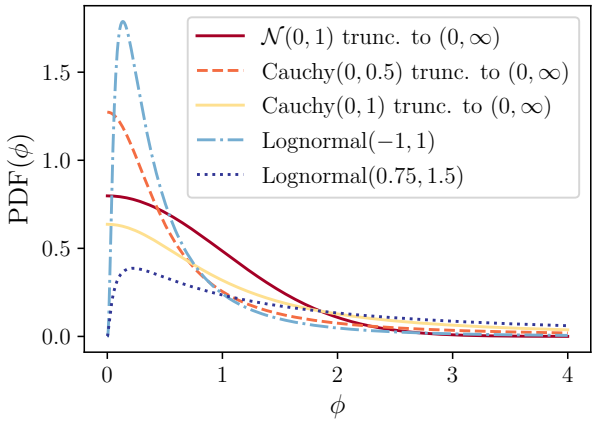


Table 5.2 & Figure 5.4: Summary of different prior distributions used in Chapter 5.

likelihood, $p(\mathbf{s} \mid \boldsymbol{\phi})$, is therefore Gaussian. We assume that ξ_e or ξ_m are fixed over the population, and are unknown, as this is normally the case when working with real datasets. We estimate ξ_e or ξ_m simultaneously with $\boldsymbol{\theta}$, making the vector of unknown parameters $\boldsymbol{\phi} = [\boldsymbol{\theta}^T, \boldsymbol{\xi}^T]^T$. Depending on whether effector dynamics are observed (observation model $\mathbf{h}_2(\cdot)$) or not (observation model $h_1(\cdot)$), we have that $\boldsymbol{\xi} = [\xi_e, \xi_m]^T$ or $\xi = \xi_m$, respectively.

Prior distribution. Within the Bayesian framework all unknown parameters $\boldsymbol{\phi}$ are considered to be random variables, and we must assign them a prior distribution, $p(\boldsymbol{\phi})$. We consider different priors (see Table 5.2) and their influence on inference results (see Section 5.3.1.1). Since all parameters are positive, we consider three alternative types of prior distributions: lognormal, Gaussian and Cauchy distributions, the latter two distributions being truncated to $(0, \infty)$.

5.2.3 Parameter estimation

We aim to infer model parameters and their uncertainty for the Bayesian inverse problems described in Section 5.2.2, given dichotomous synthetic datasets. In other words, we seek to compute the posterior distribution of unknown parameters, $\boldsymbol{\phi}$, given the observables, $\mathbf{s}$, i.e.

$$p(\boldsymbol{\phi} \mid \mathbf{s}) = \frac{p(\mathbf{s} \mid \boldsymbol{\phi})\, p(\boldsymbol{\phi})}{\int_{\boldsymbol{\phi}} p(\mathbf{s} \mid \boldsymbol{\phi})\, p(\boldsymbol{\phi})\, \mathrm{d}\boldsymbol{\phi}}. \tag{5.4}$$

Bayesian inversion requires integration over the parameter space to compute the normalising constant. For likelihoods whose analytic expressions are not accessible (e.g. due to nonlinearity of ODEs), and for high-dimensional parameter spaces, for which numerical integration is infeasible, the posterior must be approximated. Sampling from the posterior using *Markov chain Monte Carlo* (MCMC) methods is one of the most powerful techniques for doing this [164]. We describe below the MCMC algorithm we use, and also explain how we evaluate its performance.

5.2.3.1 Markov Chain Monte Carlo sampling

Metropolis-Hastings (MH) algorithm. MH methods are popular MCMC samplers that generate a Markov chain of parameters whose stationary distribution is equivalent to the target posterior distribution. MH methods rely crucially upon a proposal distribution $q(\boldsymbol{\phi} \mid \tilde{\boldsymbol{\phi}})$, considered to be a Markov transition kernel, which is used to propose a new candidate sample given the value of the current sample. The candidate sample is accepted with acceptance probability that is determined by comparing the values of the unnormalised posterior PDF, $\pi(\boldsymbol{\phi}) = p(\mathbf{s} \mid \boldsymbol{\phi})\, p(\boldsymbol{\phi})$ (likelihood $\times$ prior), at the current and the candidate sample. Attempts to move to a sample point with higher posterior probability than the current point are accepted, while attempts to move to points with lower posterior probability are sometimes rejected. The MH algorithm is outlined in Algorithm 2.

Exploration of the parameter space and convergence to $p(\boldsymbol{\phi} \mid \mathbf{s})$ using Algorithm 2 can be very slow. It greatly depends on how similar the proposal is to the posterior; if the two are very different, then few samples are accepted. The proposal has to be carefully tuned to the problem, and enable the generated Markov chain to efficiently move around the parameter space. This is particularly difficult to achieve for high-dimensional spaces and complex landscapes (e.g. with multiple modes) [14].

Hamiltonion Monte Carlo (HMC) algorithm. HMC is a variant of the MH algorithm that uses the sampled density function's differential structure to

Algorithm 2: Metropolis-Hastings algorithm

$\boldsymbol{\phi}^0 \sim \pi_0(\boldsymbol{\phi})$ ▷ *Initialise parameter values*

for $i = 1 : N_{\text{samples}}$ **do**

  $\boldsymbol{\phi}^* \sim q(\boldsymbol{\phi} | \boldsymbol{\phi}^{i-1})$ ▷ *Sample a candidate from the proposal*

  $\alpha(\boldsymbol{\phi}^*, \boldsymbol{\phi}^{i-1}) = \min\left\{1, \frac{\pi(\boldsymbol{\phi}^*) q(\boldsymbol{\phi}^{i-1} \mid \boldsymbol{\phi}^*)}{\pi(\boldsymbol{\phi}^{i-1}) q(\boldsymbol{\phi}^* \mid \boldsymbol{\phi}^{i-1})}\right\}$ ▷ *Compute acceptance probability*

  $u \sim U(0, 1)$ ▷ *Sample a uniform random variable*

  if $u \leq \alpha(\boldsymbol{\phi}^*, \boldsymbol{\phi}^{i-1})$ **then** $\boldsymbol{\phi}^i = \boldsymbol{\phi}^*$ ▷ *Accept the candidate w.p.* $\alpha(\boldsymbol{\phi}^*, \boldsymbol{\phi}^{i-1})$

  else $\boldsymbol{\phi}^i = \boldsymbol{\phi}^{i-1}$ ▷ *Reject the candidate w.p.* $1 - \alpha(\boldsymbol{\phi}^*, \boldsymbol{\phi}^{i-1})$

move to unexplored regions of parameter space. As such, it can efficiently navigate high-dimensional spaces. By analogy with a physical system, HMC uses for the proposal an approximate Hamiltonian dynamics simulation based on numerical integration that is then corrected by performing an MH acceptance step.

As in a physical system, a candidate point has a position, $\boldsymbol{\phi} \in \mathbb{R}^L$, and an (auxiliary) momentum variable $\mathbf{m} \in \mathbb{R}^L$, so that each dimension of $\mathbf{m}$ complements a dimension of the parameter vector $\boldsymbol{\phi}$. The algorithm draws from a joint density

$$p(\mathbf{m}, \boldsymbol{\phi}) = p(\mathbf{m} \,|\, \boldsymbol{\phi}) p(\boldsymbol{\phi}) = e^{-H(\mathbf{m}, \boldsymbol{\phi})}, \tag{5.5}$$

which is such that the target posterior density for $\boldsymbol{\phi}$ is recovered if we marginalise out the momentum $\mathbf{m}$. The joint PDF can be written in terms of the Hamiltonian function, $H(\mathbf{m}, \boldsymbol{\phi})$, independent of the parametrisation, and defined as follows

$$H(\mathbf{m}, \boldsymbol{\phi}) = -\log p(\mathbf{m}, \boldsymbol{\phi}) = -\log p(\mathbf{m} \,|\, \boldsymbol{\phi}) - \log p(\boldsymbol{\phi}) =: K(\mathbf{m} \,|\, \boldsymbol{\phi}) + U(\boldsymbol{\phi}). \tag{5.6}$$

The value of the Hamiltonian at a point $(\mathbf{m}, \boldsymbol{\phi})$ is called the energy at that point, and may be decomposed into its so-called kinetic energy, $K(\mathbf{m} \,|\, \boldsymbol{\phi}) := -\log p(\mathbf{m} \,|\, \boldsymbol{\phi})$, and potential energy, $U(\boldsymbol{\phi}) := -\log p(\boldsymbol{\phi})$ [14]. Since the Hamiltonian captures the geometric properties of the sampled landscape, we use it to generate paths that rapidly move along the posterior surface. At every MH step, a momentum $\mathbf{m}^*$ is sampled (e.g. via $\mathcal{N}(\mathbf{0}, M)$ with M as the mass matrix) and the Hamiltonian equations,

$$\frac{\partial \boldsymbol{\phi}}{\partial t} = \frac{\partial H}{\partial \mathbf{m}} = \frac{\partial K}{\partial \mathbf{m}}, \qquad \frac{\partial \mathbf{m}}{\partial t} = -\frac{\partial H}{\partial \boldsymbol{\phi}} = -\frac{\partial K}{\partial \boldsymbol{\phi}} - \frac{\partial U}{\partial \boldsymbol{\phi}}, \tag{5.7}$$

are numerically integrated to compute the desired trajectory. To control the numerical integration error in a way that preserves the value of the Hamiltonian (i.e. energy), a type of symplectic integrator called the *leapfrog integrator* is used, where particle momentum is updated differently to particle position [14]. The

end point of the generated path, $(\mathbf{m}^*, \boldsymbol{\phi}^*)$, is used as the candidate point in the sampler, and is accepted with probability

$$\alpha(\mathbf{m}, \boldsymbol{\phi}, \mathbf{m}^*, \boldsymbol{\phi}^*) = \min(1, \exp(H(\mathbf{m}, \boldsymbol{\phi}) - H(\mathbf{m}^*, \boldsymbol{\phi}*))). \tag{5.8}$$

The performance of HMC is highly sensitive to the step size and the chosen period of integration of the Hamiltonian equations. It requires careful, and often computationally expensive, tuning. The No-U-Turn Sampler (NUTS) [73], implemented as part of the Stan package [160], extends HMC by performing its efficient automatic tuning. Step size is tuned during the warmup stage and fixed during subsequent sampling. The number of sufficient warmup iterations for tuning depends on the complexity of the posterior and is chosen by the user.

We use the NUTS HMC in Stan [160] for our analysis. For each sampled posterior, we run 10 MCMC chains with settings `adapt_delta=0.99` and `maxtreedepth=15`; for details see [159]. During the warmup stage (tuning), we generate 3000 parameter samples per chain which are discarded. The subsequent 5000 samples per each stationary/converged chain comprise our inferred posterior.

5.2.3.2 Evaluation of sampling results

We characterise posterior surfaces and evaluate performance of posterior sampling in four ways: (1) by visualising posterior samples, computing their summary statistics and performing posterior predictive checks, (2) by performing diagnostic checks for convergence of MCMC chains to the stationary distribution, (3) by comparing marginal posterior distributions to the prior to quantify the information gain from the dataset, and (4) by comparing marginal posterior distributions to profile likelihoods to identify potential difficulties and/or bias in HMC sampling, and to establish the influence of the prior. The details are described below.

Posterior characterisation. We visualise HMC-generated marginal posterior distributions using histograms and plots of *kernel density estimates* (KDE) for each ϕ_i. Non-truncated KDEs are estimated using the `KDEUnivariate` function from the Statsmodels Python package [152], while truncated ones are estimated using the `SmoothKernelDistribution` function from Mathematica [181]. For each marginal posterior distribution, $p(\phi_i \,|\, \mathbf{s})$, we record the *highest posterior density interval* (HPDI); these are types of *credible intervals* (CRIs) (Bayesian analogue of confidence intervals). HPDI with confidence $100(1-\alpha)\%$ is the narrowest interval that covers $100(1-\alpha)\%$ of the marginal distribution. We calculate them using the `hdi` function from the ArviZ Python package [86]. We note that these HPDIs do not correct for multimodality of the posterior and are, therefore, less accurate in such cases. We also perform posterior predictive checks, where we look at distributions of model forecasts, $\mathcal{F}(\boldsymbol{\Phi}_{\text{sampled}})$, evaluated at the sampled parameters.

MCMC diagnostics. We check for convergence of the MCMC chains generated via HMC by (i) visually inspecting the chains for each ϕ_i, (ii) inspecting the distribution of the log-likelihood and log-posterior, (iii) checking whether the rank-normalised split-$\widehat{R}$ statistic [170] is satisfactory for each ϕ_i, and (iv) checking for any divergent MCMC transitions. The $\widehat{R}$ statistic compares the variance of all chains together with the variance of individual chains. If all chains are mixing well, then $\widehat{R}$ converges to 1; $\widehat{R} > 1.01$ indicates one or more chains have not converged to the target posterior. We use the `rhat` function from the ArviZ package [86] to compute $\widehat{R}$ for each parameter separately. Divergent HMC transitions occur if the numerically approximated Hamiltonian path departs from the true one. If present, they may indicate insufficient exploration of problematic regions in the posterior.

Kullback-Leibler divergence (KLD). We use the KLD statistic [125] to determine how different the prior and posterior distributions, $p(\boldsymbol{\Phi})$ and $p(\boldsymbol{\Phi} \,|\, \mathbf{s})$,

respectively, are. For parameter ϕ_i, the KLD is given by

$$\mathrm{D}_{\mathrm{KL}}(\phi_i) = \int_{\phi_i} \log\left(\frac{p(\phi_i \,|\, \mathbf{s})}{p(\phi_i)}\right) p(\phi_i \,|\, \mathbf{s})\, \mathrm{d}\phi_i = \mathbb{E}_{p(\phi_i \,|\, \mathbf{s})}\left[\log\left(\frac{p(\phi_i \,|\, \mathbf{s})}{p(\phi_i)}\right)\right] \tag{5.9a}$$

$$\approx \frac{1}{N_{\mathrm{samples}}} \sum_{j=1}^{N_{\mathrm{samples}}} (\log(\hat{p}(\phi_i^{(j)} \,|\, \mathbf{s})) - \log(p(\phi_i^{(j)}))). \tag{5.9b}$$

Equation (5.9a) is defined only when, for all ϕ_i, if $p(\phi_i) = 0$ then $p(\phi_i \,|\, \mathbf{s}) = 0$, which holds since $p(\phi_i \,|\, \mathbf{s}) \propto p(\phi_i) \int_{\boldsymbol{\Phi}_{-i}} p(\mathbf{s} \,|\, \boldsymbol{\Phi}) p(\boldsymbol{\Phi}_{-i}) \mathrm{d}\boldsymbol{\Phi}_{-i}$. (5.9b) gives the estimate of KLD, where $(\phi_i^{(j)})_{j=1}^{N_{\mathrm{samples}}}$ are samples from $p(\phi_i \,|\, \mathbf{s})$ computed via the HMC. The marginal probability densities $p(\phi_i^{(j)} \,|\, \mathbf{s})$ are evaluated using a KDE based on posterior samples, denoted as $\hat{p}(\phi_i^{(n)} \,|\, \mathbf{s})$. Kernel density estimation is performed in Mathematica (`SmoothKernelDistribution`) [181], for most cases using Gaussian kernels truncated to $(0, \infty)$ as all parameters are positive. For posteriors with distant modes and poor mixing, the kernel densities for parameters ϕ_i associated with e_0, σ and δ are, instead, estimated on a logarithmic scale using non-truncated Gaussian kernels. In this case we estimate the KDE of $\tilde{\phi}_i = \log(\phi_i)$ and then transform it to the linear scale by multiplying the obtained density by the Jacobian of the transformation, $\frac{\mathrm{d}\exp(\tilde{\phi}_i)}{\mathrm{d}\tilde{\phi}_i} = \exp(\tilde{\phi}_i) = \phi_i$.

Profile likelihoods. Using the methodology from Chapter 4, if we sample the posterior for a given dataset then we also compute profile likelihoods (PLs) for the corresponding MLE problem (i.e. without the prior and with a known observation standard deviation, $\boldsymbol{\xi}$). This enables us to compare the posterior and the likelihood surfaces and, together with other diagnostics, to uncover potential biases and difficulties with sampling, and the impact of the prior. Our aim is to minimise such effects to sample posteriors that closely resemble the likelihood. Our approach here is similar to studies by Raue et al. [135], Simpson et al. [155], and Vanlier et al. [168], where joint Bayesian and frequentist parameter inference, involving MCMC sampling and PLs, was adopted.

We note that when we fit the model to a particular dataset under the true hypothesis, we know the true parameter values at which we can initialise the PL computation. When fitting the model to a dataset under the alternative hypothesis, the true parameter values are unknown. In this case, we first run MCMC to compute the corresponding posterior and obtain the MAP parameter estimates. The mode of the posterior is used as an initial guess for finding the MLE, which is then used as an initial guess for PL computation.

5.2.4 Estimation of predictive accuracy

In this section we summarise the methodology we will use from Vehtari et al. [169].

For every calibrated model, we would like to quantify how well it performs on unseen data (i.e. to measure its predictive accuracy). Ideally one would compute the 'expected log pointwise predictive density' (elpd) for a new dataset $\tilde{\mathbf{s}}$ of N data points taken one at a time, given an old dataset $\mathbf{s}$, where elpd is defined as follows

$$\text{elpd} = \sum_{i=1}^{N} \int \log\, p(\tilde{s}_i \,|\, \mathbf{s})\, p_{\text{true}}(\tilde{s}_i) \mathrm{d}\tilde{s}_i.$$

Here $p_{\text{true}}(\tilde{s}_i)$ is the unknown distribution representing the true data-generating process for a new i-th point $\tilde{s}_i$. If we have limited data, we can approximate elpd using *cross-validation* (CV), where the model is re-fitted to different training subsets. Since this is computationally expensive for Bayesian models, we use either the *Watanabe-Akaike information criterion* (WAIC) or approximate CV, such as the *leave-one-out* (LOO) *estimate* of the out-of-sample predictive fit.

WAIC is defined as

$$\text{elpd}_{\text{WAIC}} = \sum_{i=1}^{N} \log\, p(s_i \,|\, \mathbf{s}) - \phi_{\text{WAIC}}, \quad \text{with} \quad \phi_{\text{WAIC}} = \sum_{i=1}^{N} \text{Var}_{\boldsymbol{\Phi} \,|\, \mathbf{s}}[\log\, p(s_i \,|\, \boldsymbol{\Phi})]. \tag{5.10}$$

Here $p(s_i \,|\, \mathbf{s}) = \int p(s_i \,|\, \boldsymbol{\Phi}) p(\boldsymbol{\Phi} \,|\, \mathbf{s})\, \mathrm{d}\boldsymbol{\Phi} \approx \frac{1}{N_{\text{samples}}} \sum_{j=1}^{N_{\text{samples}}} p(s_i \,|\, \boldsymbol{\Phi}^{(j)})$, $\boldsymbol{\Phi}^{(j)}$ is a posterior (MCMC) sample and N_{samples} is the sample size. $\text{Var}_{\boldsymbol{\Phi} \,|\, \mathbf{s}}[\cdot]$ is the variance

under the parameter posterior. ϕ_{WAIC} as defined is the estimated effective number of parameters, which represents the penalty for overfitting to training data (analogous to Akaike or deviance information criteria). Gelman et al. [58] and Vehtari et al. [169] empirically estimated that WAIC becomes unreliable when $\text{Var}_{\boldsymbol{\Phi} \mid \mathbf{s}}[\log p(s_i \mid \boldsymbol{\Phi})] > 0.4$ for some i.

The LOO estimate is defined as

$$\text{elpd}_{\text{LOO}} = \sum_{i=1}^{N} \log p(s_i \mid \mathbf{s}_{-i}), \tag{5.11}$$

where $\mathbf{s}_{-i} = [s_1, \ldots, s_{i-1}, s_{i+1}, \ldots]^T$ and $p(s_i \mid \mathbf{s}_{-i}) = \int p(s_i \mid \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathbf{s}_{-i}) \, \mathrm{d}\boldsymbol{\theta}$ are log pointwise LOO predictive densities. They may be computed by re-sampling the posterior via MCMC without the i-th point, which is called *LOO cross-validation* (LOO-CV). As this is computationally expensive, we approximate the densities using *Pareto smoothed importance sampling* (PSIS) [169] where possible, for which only the posterior distribution for the full dataset is needed. We note also that LOO and WAIC are asymptotically equivalent [169].

Pareto smoothed importance sampling (PSIS). Pointwise log predictive densities, $p(s_i \mid \mathbf{s}_{-i})$, are approximated using log predictive densities evaluated at

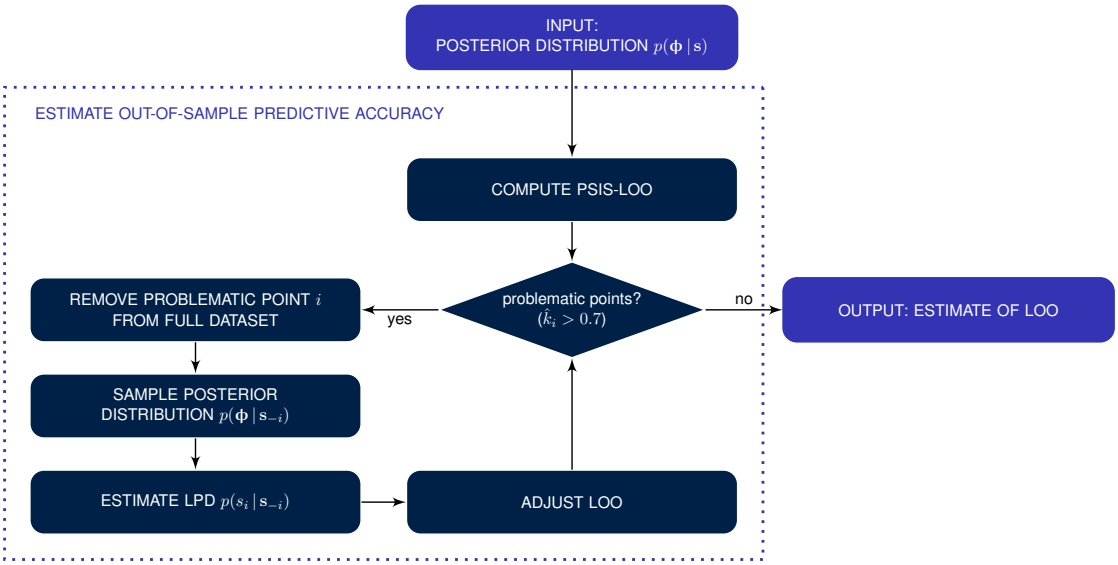


Figure 5.5: Diagram showing the process of estimating the leave-one-out (LOO) predictive accuracy for a Bayesian model fit using samples from its posterior distribution $p(\boldsymbol{\Phi} \mid \mathbf{s})$.

posterior samples given the full dataset, $p(s_i \mid \boldsymbol{\Phi}^{(s)})$, and re-weighting them via *importance sampling* (IS). In particular, the IS approximation is

$$p(s_i \mid \mathbf{s}_{-i}) \approx \frac{\sum_{j=1}^{N_{\text{samples}}} p(s_i \mid \boldsymbol{\Phi}^{(j)}) r_i^{(j)}}{\sum_{j=1}^{N_{\text{samples}}} r_i^{(j)}} = \frac{1}{\frac{1}{N_{\text{samples}}} \sum_{j=1}^{N_{\text{samples}}} r_i^{(j)}}, \tag{5.12}$$

for raw importance sampling weights defined as

$$r_i^{(j)} = \frac{1}{p(s_i \mid \boldsymbol{\Phi}^{(j)})} \propto \frac{p(\boldsymbol{\Phi}^{(j)} \mid \mathbf{s}_{-i})}{p(\boldsymbol{\Phi}^{(j)} \mid \mathbf{s})}. \tag{5.13}$$

Weights (5.13) follow from a well-known result by Gelfand et al. [57] provided the data points are conditionally independent; this is the case for the observation model defined in Section 5.2.2. The computation (5.12) may be unstable if the ratios $r_i^{(j)}$ have long right tails and, therefore, high or infinite variance. We can stabilize the weights using PSIS, where, separately for each data point s_i, a generalised Pareto distribution is fitted to the largest 20% of raw weights $r_i^{(j)}$, and the largest ones are replaced by the expected values of order statistics of the fitted Pareto distribution. The smoothed weights, denoted as $w_i^{(j)}$, are used in place of $r_i^{(j)}$ in (5.12). We refer the reader to [169] for more details on PSIS-LOO.

There are two main advantages of PSIS. First, it reduces the variance of importance weights by introducing a small bias. Second, it enables the user to diagnose when PSIS may be unreliable. This is considered the case when the estimated shape parameter, $\hat{k}_i$, of the fitted Pareto distribution for some data point i exceeds 0.7, an empirically determined threshold suggested in [169]. This often occurs when the model is not robust or there exist influential data points. For data points with $\hat{k}_i > 0.7$ the authors advise estimating $p(s_i \mid \mathbf{s}_{-i})$ by re-sampling the posterior $p(\boldsymbol{\Phi} \mid \mathbf{s}_{-i})$ with the problematic point i held out, which is the approach we adopt here (see Fig. 5.5). The method we use to estimate elpd_{LOO} is thus a combination of LOO-CV and PSIS-LOO, which is computationally much cheaper for a small number of problematic data points than LOO-CV. The PSIS-LOO method is implemented as function `loo` in the ArviZ package [86], and we adapt it to implement adjustment for problematic points.

5.2.5 Hypothesis testing

This section builds on ideas proposed by Vehtari et al. [169] and Lambert [91].

Once the elpd (LOO or WAIC) values have been computed, we perform a one-sided hypothesis test to determine which fitting hypothesis is statistically more likely to explain our dataset. The test compares the following hypotheses:

Null hypothesis (H_0): $\text{elpd}_{\text{true}} - \text{elpd}_{\text{alternative}} \leq 0$

We cannot distinguish which hypothesis is more likely, or data are more likely to have been generated from the alternative hypothesis.

Alternative hypothesis (H_1): $\text{elpd}_{\text{true}} - \text{elpd}_{\text{alternative}} > 0$

Data are more likely to have been generated from the true hypothesis.

Since $\text{elpd}_{\text{true}} - \text{elpd}_{\text{alternative}} = \sum_{i=1}^{N_{\text{points}}} (\text{elpd}_{\text{true}}^{(i)} - \text{elpd}_{\text{alternative}}^{(i)})$, our test statistic,

$$T = \frac{\text{elpd}_{\text{true}} - \text{elpd}_{\text{alternative}}}{\sqrt{N_{\text{points}}}\,\text{SE}[\text{elpd}_{\text{true}}^{(i)} - \text{elpd}_{\text{alternative}}^{(i)}]}\,, \tag{5.14}$$

is effectively distributed according to a standard normal distribution $\mathcal{N}(0,1)$ when $N_{\text{points}} > 30$, and otherwise according to the student's T-distribution, $t_{N_{\text{points}}-1}$, with $N_{\text{points}} - 1$ degrees of freedom. $\text{elpd}^{(i)}$ is the estimated pointwise LOO or WAIC value for the i-th data point, and SE is the sample standard error. For $T = x$, we compute the probability that T is more extreme than the value of our test statistic under the null hypothesis, $\mathbb{P}(T > x \,|\, H_0)$. This is the p-value: if it is smaller than a chosen small significance threshold, then we are unlikely to observe the value of T under the null hypothesis, and we can reject it. This means that the elpd values of the fits are statistically different, and therefore the fits and fitting hypotheses associated to them are distinguishable for our dataset.

5.3 Results

Our main results relate to hypothesis comparison for dichotomous pooled synthetic datasets (see Section 5.2.1). We rely on the reliability of inference using HMC

for each dataset and a set of modelling/fitting assumptions. In Section 5.3.1.1, we evaluate the performance of HMC on inference problems defined by individual curves. In Sections 5.3.1.2–3, we consider results of parameter estimation for our pooled datasets. In Section 5.3.2, we address the question of distinguishing the data-generating hypothesis from alternative ones for these datasets.

5.3.1 Parameter inference

5.3.1.1 Case studies for individual curves

Here we test HMC on datasets for a single individual from Chapter 4 (case studies 4.1, 4.2a and 4.3). We compare results for MLE from Chapter 4 and the equivalent Bayesian problem, in which $\boldsymbol{\xi}$, standard deviations of observation noise, are also unknown and estimated (see Section 5.2.2). We consider two different priors for the unknown parameters $\boldsymbol{\phi}$ (see priors 1 and 2 in Table 5.2 and Fig. 5.4).

Case studies 4.1 and 4.3: elimination and equilibrium. For brevity, we do not show results for these two case studies, and only summarise them below. For both case studies, we obtain MCMC chains that converge to a unimodal stationary distribution for all noise levels and priors; split-$\widehat{R}$ values are all acceptable (< 1.01) and there are no divergent iterations. However, depending on the prior used, we observe discrepancies between the marginal posterior distributions and the *profile likelihoods* (PLs) in the elimination regime. The lognormal prior (prior 1) prevents sampling of values close to zero, and introduces bias compared to the PL, particularly for parameters $\delta^{(1)}$, $e_0^{(1)}$ and $\sigma^{(1)}$, for which the PLs may have large probabilities for values close to zero. Cauchy priors truncated to $(0, \infty)$ (prior 2) assign greater probability to values close to zero (see Fig. 5.4). This remedies discrepancies in that region, and has a negligible influence on the likelihood. Small differences remain for observation model $h_1(\cdot)$, particularly at larger noise levels. In the equilibrium regime, we observe good agreement between the marginal posteriors

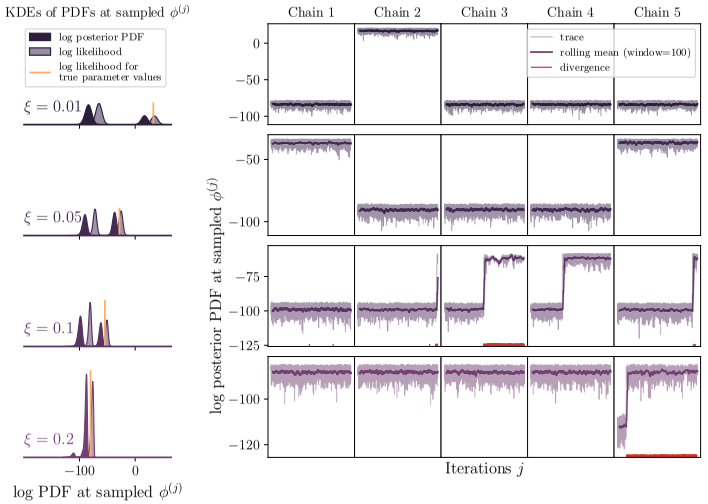
and PLs for all priors tested. This may be because the domain boundaries have a lesser impact (see Chapter 4) and because the problem has better identifiability.

Case study 4.2a: escape. For escape data, under prior 1 not all chains converge to the same mode or they jump between modes (see Figs. 5.6a–b). When chains converge to different modes or remain at a less probable mode for long time periods, they mix poorly, which is detected by large split-$\widehat{R}$ values (see Figs. 5.7a–b).

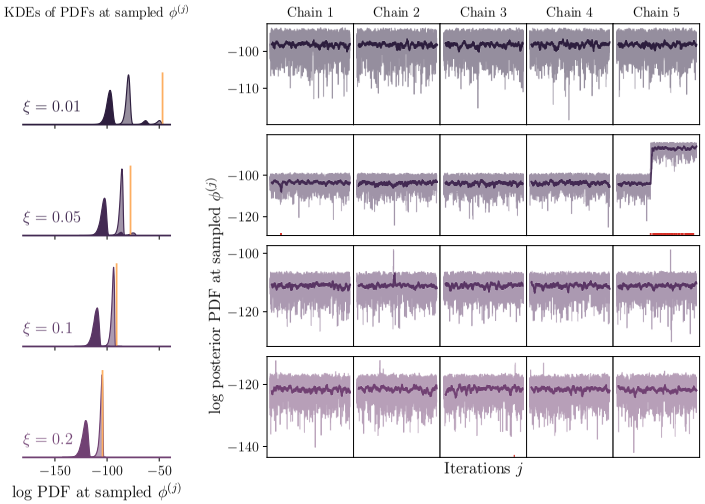
Poor mixing biases the posterior sampled, as seen by comparison with PLs. In Chapter 4 we showed that the negative log-likelihood surface generated by the escape data has a global minimum (close to the data-generating parameter values) and a distant local minimum (disjoint parameter region). Its double-valleyed shape should give rise to a bimodal posterior when the distant mode provides a good fit (i.e. when the difference of the weighted sum of squares, $V(\boldsymbol{\theta}) - V(\hat{\boldsymbol{\theta}})$, lies below the confidence bound $\chi^2_{4,0.95}$). This is the case for data without effector cell measurements and high noise levels ($\xi \geq 0.15$). In other cases $V(\boldsymbol{\theta})$ is not small enough and the distant mode should therefore only be visited very rarely by the MCMC. In our simulations, under prior 1, some chains visit only the distant mode, implying poor performance of the HMC under this setup and sampling bias.

We attribute the difficulties with posterior sampling to the bimodal likelihood surface, the prior and the design of the HMC algorithm. The probability of visiting the two modes based on the likelihood may be significantly influenced by the chosen prior, and can lead to improbable modes becoming more probable and being visited more frequently by MCMC chains (e.g. for a prior placing greater weight on larger values of σ and δ). Sampling from multimodal distributions is a common problem of many samplers and optimisation techniques [136, 187]. The design of the HMC algorithm makes transitioning between modes particularly difficult. Since it uses the gradient in the posterior to move around, it can tune to the mode it starts in and explore this mode very well, but it may struggle to transition between different

(a) prior 1, observation model $\mathbf{h}_2(\cdot)$



(b) prior 1, observation model $h_1(\cdot)$



(c) prior 2, observation model $h_1(\cdot)$

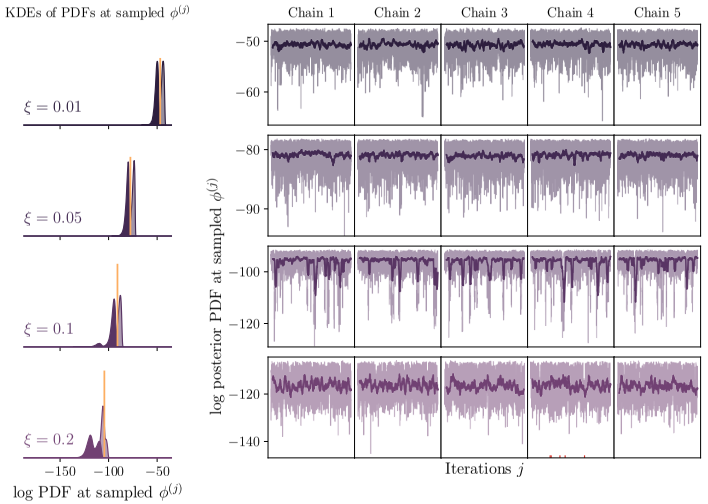


Figure 5.6: Series of plots showing (left) the distribution of log posterior PDF (dark shade) and log-likelihood (light shade) for all samples across 10 MCMC chains, and (right) log posterior PDF for MCMC chains 1-5 plotted against MCMC iteration number, when data generated from the **escape** regime (**case study 4.2a**) are used for inference. Plots (a-c) differ in the observation model and prior used. MCMC chains are more likely to get stuck at the distant mode under prior 1. Orange vertical lines denote the log-likelihood evaluated at the true parameter values. Red marks denote divergent iterations.

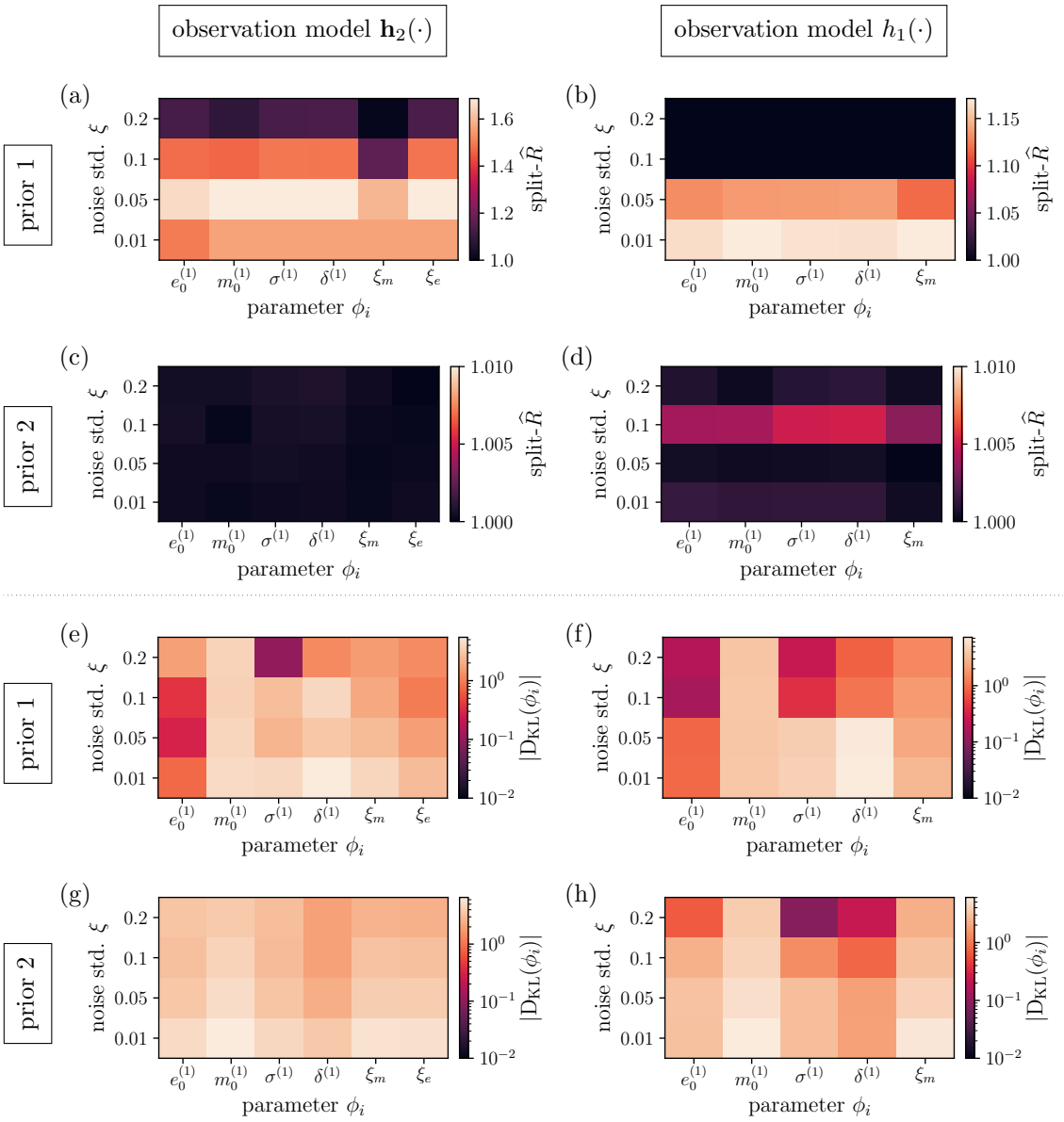


Figure 5.7: Series of heatmaps showing how values of (a-d) split-$\widehat{R}$ and (e-h) Kullback-Leibler divergence (KLD) vary for inferred posterior distributions given data from the **escape** regime (**case study 4.2a**). The heatmaps differ in the prior and observation model used. Each heatmap shows variation with different noise levels and for different parameters. Values of $\widehat{R} > 1.01$ indicate poor mixing of the posterior chains for prior 1. Small KLDs ($\leq \mathcal{O}(10^{-1})$) indicate similarity to the prior, and therefore poor identifiability.

modes, especially when they are isolated and the region between them has low posterior probability; this is called an energetic barrier [187]. This is what happens for observation model $h_1(\cdot)$ under prior 1 (see Fig. 5.6). Noting that energy is preserved in any HMC trajectory, when a chain first explores the distant mode, large kinetic energies are needed to propose samples from the region between the

modes of low posterior density or the other mode, which is therefore unlikely. If the chain switches to a different mode, divergent MCMC iterations may occur as the HMC is tuned to the geometry of the mode it initially starts in (see Figs. 5.6a–b). Because of this, and the long tail of prior 1, we observe that the distant and less probable mode is more frequently sampled than the more probable mode at the MLE for most noise levels (see Fig. 5.6b).

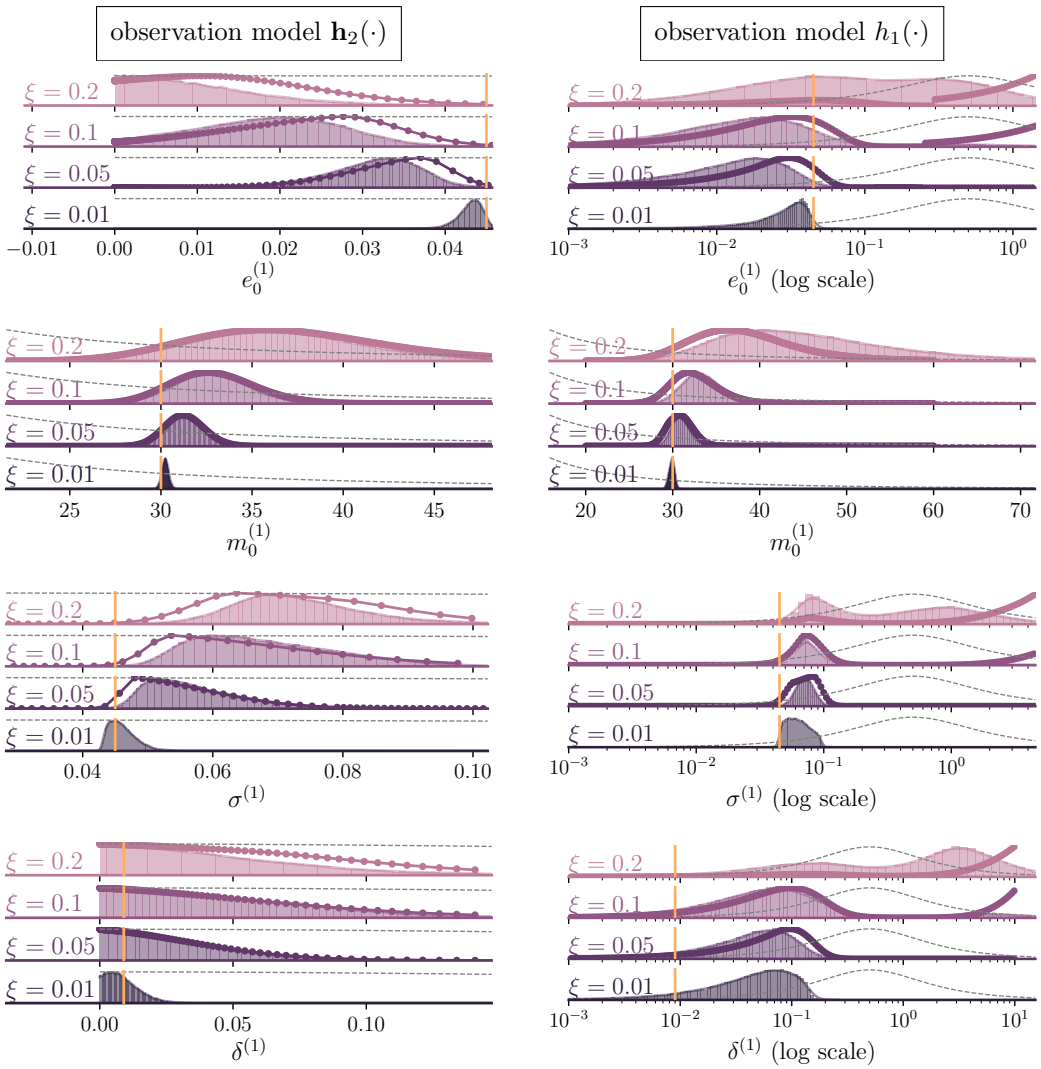


Figure 5.8: Series of plots showing marginal posterior distributions and profile likelihoods (PLs) given **escape** (**case study 4.2a**) datasets with different noise levels and **prior 2**. Marginal posteriors are unimodal, and become bimodal as noise increases for observation model $h_1(\cdot)$. There are small discrepancies between the posterior and PLs for unimodal distributions, while discrepancies increase for bimodal distributions due to computational difficulties. Note that PLs and priors are scaled so their maxima in the interval of interest correspond with the maximum of the marginal posterior density. Legend: light shaded histogram = marginal posterior, solid curve with points = PL, dashed curve = prior, orange vertical line = true parameter value.

By using a prior with a greater density around the MLE (prior 2), some of these difficulties may be overcome. Chains converge to the more probable mode and the distant mode is not sampled for all noise levels, given observation model $\mathbf{h}_2(\cdot)$ (results not shown), and, for $\xi < 0.1$, given observation model $h_1(\cdot)$ (see Fig. 5.6c). We identify no issues with HMC sampling, as split-$\widehat{R}$ values are satisfactory (see Figs. 5.7c–d) and there are no divergences. Large KLDs for all parameters (see Figs. 5.7g–h) indicate that the posterior is significantly different from the prior and that information content in these datasets is significant. For data without effector cells and with $\xi \geq 0.1$, chains move between modes and generate a bimodal distribution. Because the likelihood is very flat for high noise levels (i.e. poor identifiability, see Chapter 4), and since, together with the prior, it generates a bimodal distribution with well-connected modes (i.e. the energy barrier between the modes is sufficiently low that the HMC can breach it; see Fig. 5.8), the HMC performs better than in previous cases. The prior is influential as the marginal posterior exhibits features of the prior (low KLD in Fig. 5.7h) and differs from the PL.

These results suggest that the prior can markedly contribute to the estimation problem being well-constrained; it enables approximate solutions to be found, but they may have large uncertainties. The results also suggest that special care is needed when imposing a prior for non-trivial, but flat, likelihood surfaces, as interactions between the likelihood and the prior may lead to unexpected artefacts in the posterior.

5.3.1.2 Case study 5.1: bistable regime, two curves

Here we consider a dataset for two individuals generated from a bistable parameter regime of Kuznetsov et al.'s QSSA model. The parameters used to generate the data correspond to Hypothesis 2 (H2); they differ between individuals only in the

effector initial condition, e_0, and exhibit elimination and escape (see Fig. 5.3b). We describe and analyse inference results for fits under H1 and H2.

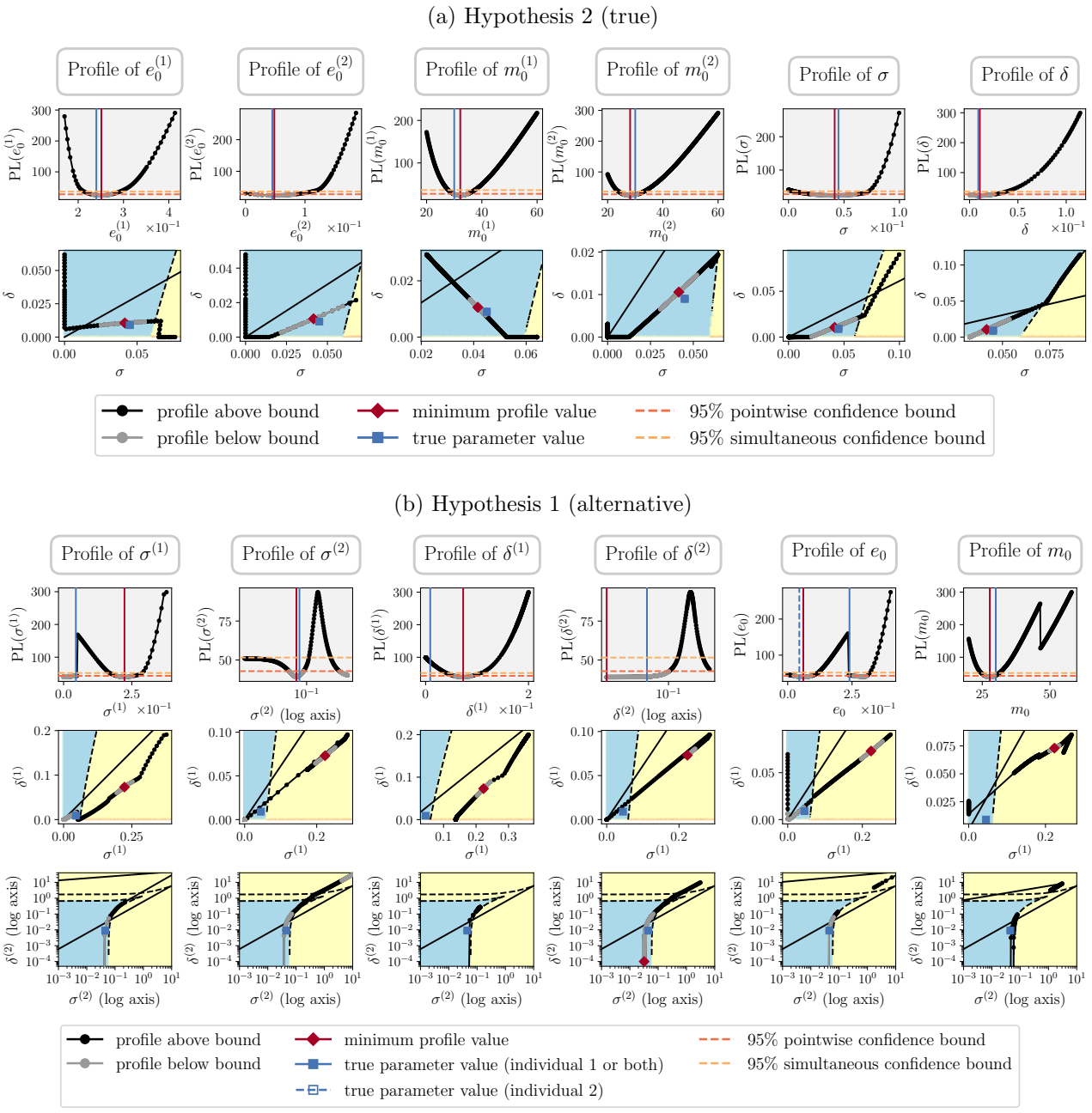


Figure 5.9: Series of plots showing profile likelihoods for a dataset generated from a **bistable** parameter regime under **case study 5.1**, with observation standard deviation $\xi = 0.05$ and for observation model $h_1(\cdot)$, fitted according to (a) H2 (true), and (b) H1 (alternative). Bottom panels show which bifurcation regions the profiles explore. Profiles for H1 fit exhibit multimodality, while those for H2 fit do not. The mode close to the origin is better adjusted for data of individual 2 (elimination), while the distant one is better adjusted for individual 1 (escape).

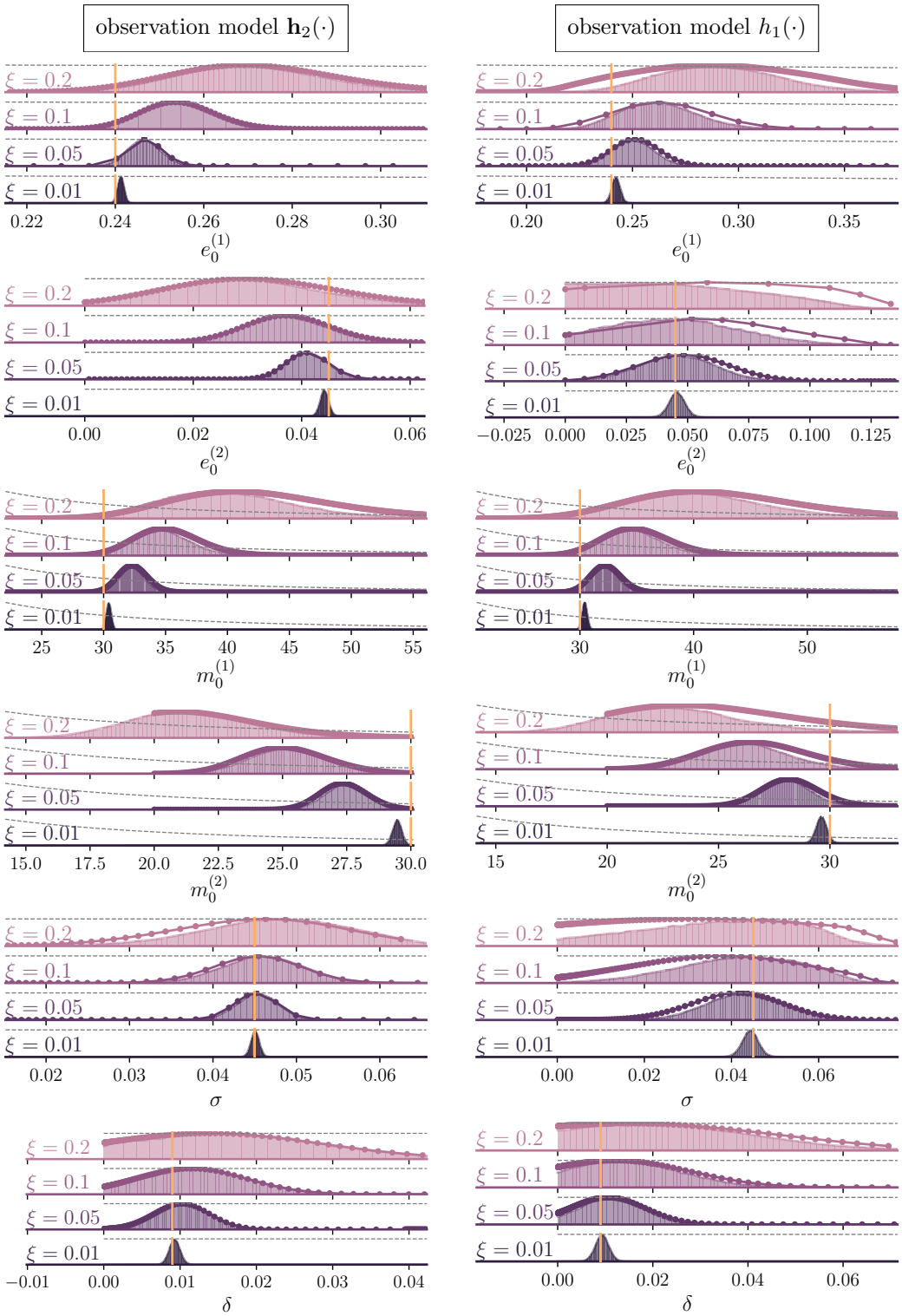


Figure 5.10: Series of plots showing marginal posterior distributions and profile likelihoods (PLs) given datasets generated from a **bistable** regime (**case study 5.1**) with different noise levels, and given **prior 3** and **H2** (true). PLs mostly match marginal posteriors, but there are slight discrepancies as effector data are removed and noise is increased. Note that PLs and priors are scaled so their maxima in the interval of interest correspond with the maximum of the marginal posterior density. Legend: light shaded histogram = marginal posterior, solid curve with points = PL, dashed curve = prior, orange vertical line = true parameter value (for the population parameters, i.e. those fixed over the population during fitting, the true data-generating values of the corresponding individual parameters of both individuals are marked).

Fit under Hypothesis 2 (true). MCMC chains for H2 converge well to a unimodal distribution for all noise levels and observation models (see Fig. 5.10). HMC performs well for this problem as diagnostics are satisfactory, with split-$\widehat{R} < 1.01$ (results not shown), and the marginal posterior distributions typically matching PLs for all parameters $\boldsymbol{\theta}$.

Unimodal marginal posteriors imply identifiability up to the widths of *credible intervals* (CRIs). Identifiability of population parameters $\sigma^{(\text{pop})}$ and $\delta^{(\text{pop})}$ is expected to be better than when the model is fitted to individual curves (case studies 4.1 and 4.2a), since information from both individuals is used for inference. Population parameters indeed have smaller CRIs due to greater curvature of the likelihood for the pooled dataset. This also affects the CRIs of individual initial conditions, $e_0^{(1)}$, $e_0^{(2)}$, $m_0^{(1)}$ and $m_0^{(1)}$, which are smaller (see Fig. 5.13).

In case studies 4.1 and 4.2a parameters σ and δ exhibited strong positive linear dependence. While $\sigma^{(\text{pop})}$ and $\delta^{(\text{pop})}$ depend linearly along PLs of different parameters, the dependencies lie along lines in $(\sigma^{(\text{pop})}, \delta^{(\text{pop})})$ space that have different slopes (see Fig. 5.9a), and are thus less pronounced for this problem. This, and the disappearance of the double-valleyed PL shape, present when fitting the escape curve alone, may contribute to increased identifiability.

The predicted model solutions inferred within the CRI match the data-generating solution within the time interval of the data (see Fig. 5.14a), but, despite additional data, a small proportion of predicted solutions for individual 2 diverge and tend to a different steady state. This is because the escape curve (individual 2) is in close proximity of the phase-plane separatrix (see Section 4.4.2.2). Data collection at later times may remedy this.

Fit under Hypothesis 1 (alternative). For H1, the ground truth is unknown, as we are fitting a model that is misspecified for the data. Despite this, we find a reasonable fit (see Fig. 5.14a). The inferred effector dynamics (e.g. for individual

1) differ from the data-generating dynamics. Therefore, while the fit is good for observation model $h_1(\cdot)$, the discrepancies in fitted dynamics are much larger for $\mathbf{h}_2(\cdot)$. H1 fits for $\mathbf{h}_2(\cdot)$ are thus expected to be easily distinguished from fits under the true hypothesis (see Section 5.3.2).

The problem of fitting the model under H1 retains many features of the problem of fitting the model to individual curves, but some are more pronounced. For observation model $\mathbf{h}_2(\cdot)$, PLs of $\sigma^{(1)}$ and $\delta^{(1)}$, associated with individual 1 (elimination), are U-shaped, while those of $\sigma^{(2)}$ and $\delta^{(2)}$, associated with individual 2 (escape), are double-valleyed, as seen in Chapter 4. Again, one valley lies at a global minimum and the other at a distant local minimum (results not shown). For observation model $h_1(\cdot)$, already at low noise levels, both valleys lie within the confidence region, but remain disjoint (see Fig. 5.9b). As noise level increases, the distant minimum may become more probable than the minimum close to the origin (results not shown). Therefore, while acceptable fits may be found under the alternative hypothesis, some parameters are less identifiable; this contrasts with the expectation that identifiability increases as there is extra information in the pooled dataset. On the other hand, population parameters, $e_0^{(\mathrm{pop})}$ and $m_0^{(\mathrm{pop})}$, are slightly better constrained than for individual fits (see Fig. 5.13), although the trends are not as clear-cut as for H1.

We note that each valley corresponds to fits that are better adjusted to one individual's data (see bottom two rows of Fig. 5.9b). In the valley close to the origin the estimated parameter values for individual 2 (escape) are closer to its data-generating parameter values than those of individual 1 (elimination), and vice versa for the distant valley. The individual whose data better fits the model depends on data quality and quantity.

The double-valleyed shape of the negative log-likelihood surface under H1 generates a bimodal posterior. There is a large sharp peak in the profiles of $\sigma^{(2)}$ and $\delta^{(2)}$ that separates the confidence regions, making it extremely difficult for

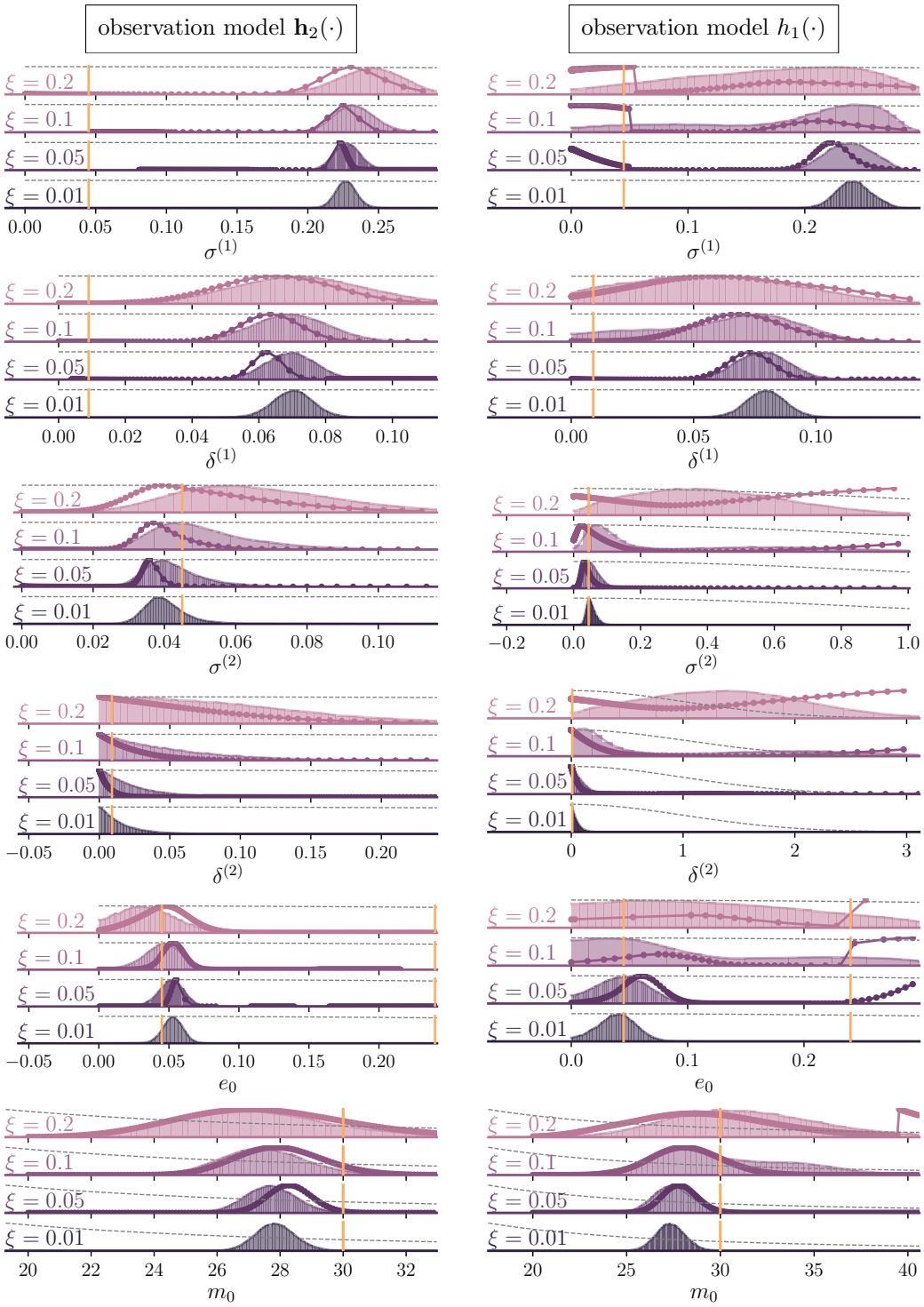


Figure 5.11: Series of plots showing marginal posterior distributions and profile likelihoods (PLs) given datasets generated from a **bistable** regime (**case study 5.1**) with different noise levels, and under **prior 3** and **H1** (alternative). For most noise levels, the posteriors are unimodal, whereas for observation model $h_1(\cdot)$ and $\xi \geq 0.1$, the distant mode is also sampled. Marginal posteriors and PLs exhibit discrepancies, particularly for bimodal cases. Note that PLs and priors are scaled so their maxima in the interval of interest correspond with the maximum of the marginal posterior density. Legend: light shaded histogram = marginal posterior, solid curve with points = PL, dashed curve = prior, orange vertical line = true parameter value (for the population parameters, i.e. those fixed over the population during fitting, the true data-generating values of the corresponding individual parameters of both individuals are marked).

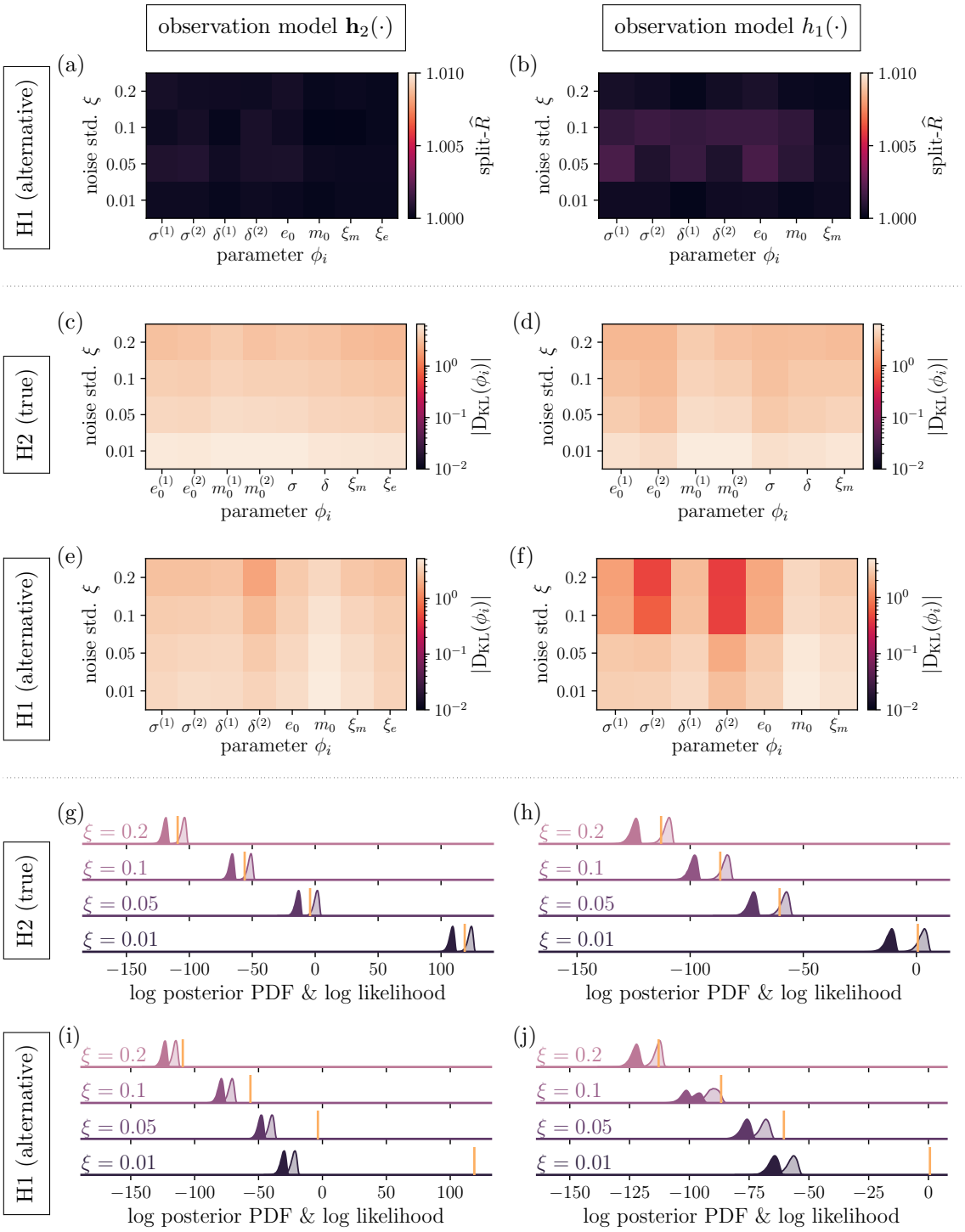


Figure 5.12: Series of plots showing how values of (a-b) split-$\widehat{R}$, (c-f) Kullback-Leibler divergence and (g-j) log-posterior and log-likelihood distributions, vary for inferred posteriors given data from the **bistable** regime (**case study 5.1**), fitted under prior 3. The plots differ in the observation model used and whether the model is fitted according to the true or alternative hypothesis. Each plot shows variation with different noise levels (and for different parameters in heatmaps). In plots (g-j) dark and light shaded KDEs correspond to log-posterior and log-likelihood PDFs, respectively, while orange lines denote the log-likelihood at true parameter values. Despite a level of bimodality in the posterior for the alternative hypothesis, the $\widehat{R}$ values are satisfactory. For the H1 fit under observation model $h_1(\cdot)$, KLDs of $\sigma^{(2)}$ and $\delta^{(2)}$ are low for $\xi \geq 0.1$, and the log-posterior distribution at $\xi = 0.1$ is slightly bimodal and biased.

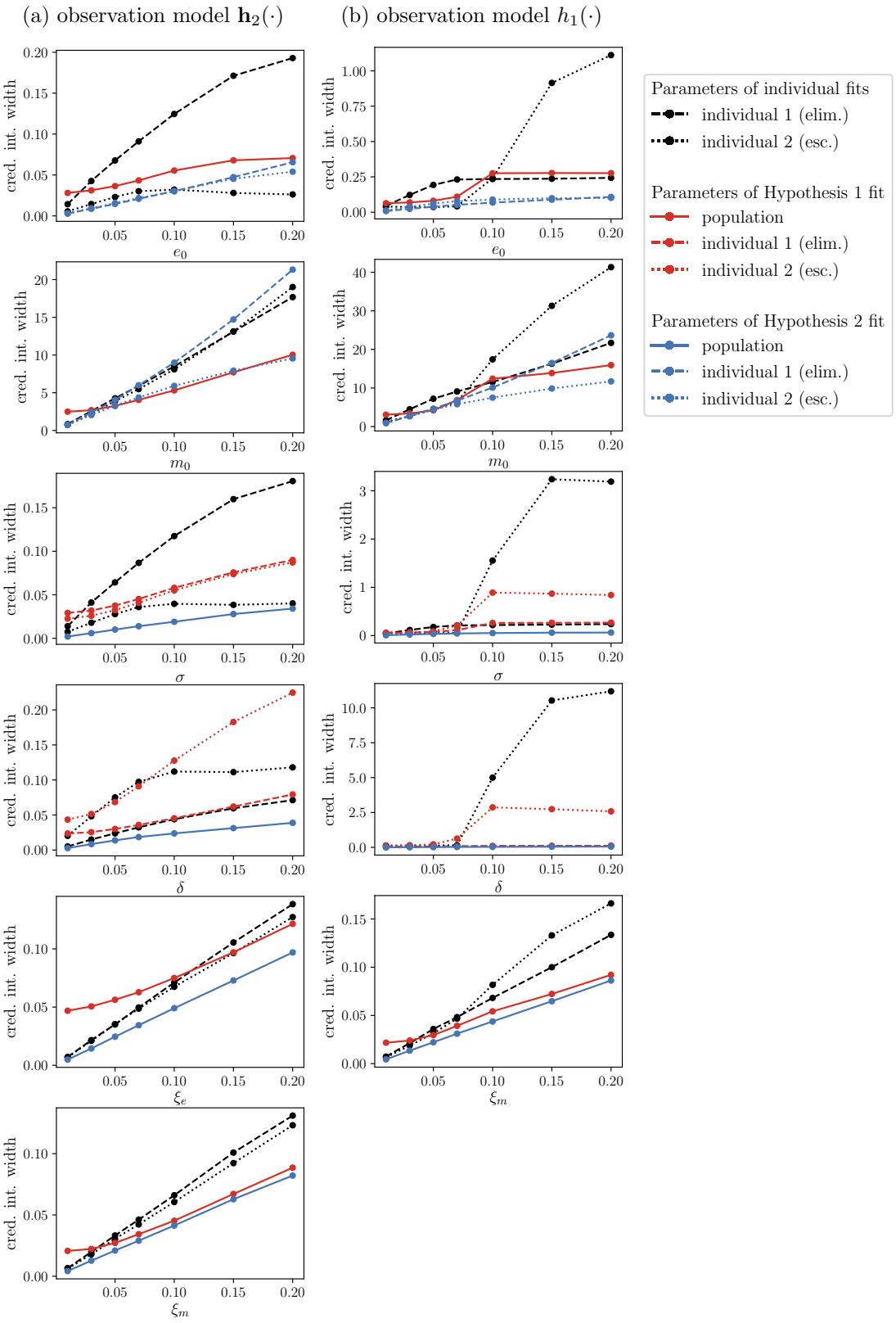


Figure 5.13: Series of plots showing the width of credible intervals for model fits to the pooled dataset for **case study 5.1** under H1 and H2 hypotheses, and for separate fits to each individual's data in this dataset. Credible intervals decrease for population fits under the true (H2) hypothesis in comparison to individual fits, whereas the decrease in uncertainty is not definitive for fits under the alternative hypothesis (H1). We note that large jumps in uncertanty correspond to multimodal distributions, for which the computed credible intervals are unreliable.

HMC to jump between confidence regions (isolated modes). Therefore, we simplify the inverse problem and choose a prior that assigns greater probabilities to smaller values of $\sigma^{(1)}$, $\sigma^{(2)}$, $\delta^{(1)}$, $\delta^{(2)}$ and e_0, so that we mainly recover the mode closer to the origin (see prior 3 in Table 5.2). This way we artificially constrain the Bayesian problem in order to find a reasonable solution via MCMC. The prior is chosen based on findings of Section 5.3.1.1, but further tuned to the problem manually.

We obtain MCMC chains with satisfactory split-$\widehat{R}$ diagnostics (see Fig. 5.12). We distinguish identifiable and poorly identifiable cases (e.g. for observation model $h_1(\cdot)$ and $\xi \geq 0.1$). For identifiable cases, only the preferred mode (closer to the origin) is sampled (Fig. 5.11). For less identifiable cases, values of KLD are smaller and indicate similarity of the posterior and the prior. Since the likelihood surface is much flatter for these cases, and since the prior has a sufficiently long tail, the HMC can transition between the modes and there is slight multimodality of the marginal posterior for some parameters (see Fig. 5.11). There is a small level of bias in sampling at $\xi = 0.1$, since the log-posterior PDF of the more frequently sampled mode is smaller than that of the less frequently sampled one (see Fig. 5.13j).

Since we plan to use the posterior samples for comparison of fits under H1 and H2, we are mainly interested in the differences in log-likelihood values at sampled points. Due to likelihood flatness for poorly identifiable cases, differences in log-likelihood values between the two modes are small. Any bias in sampling due to limitations of the HMC will have a minimal impact on the distribution of the log-likelihood at sampled points. We also note that when one mode is not sampled or sampled less frequently due to our chosen prior, we interpret this as the nature of the prior-constrained Bayesian inverse problem that we have posed, rather than sampling bias.

We note that some discrepancies between marginal posteriors and PLs can be observed (see Fig. 5.11). These are anticipated, particularly for multimodal likelihoods, for which the PLs may be difficult to compute to good accuracy (e.g.

there may be discontinuous jumps in PLs when the optimiser is suddenly attracted to a different mode). For poorly identifiable cases, the prior may also strongly influence the posterior, and cause discrepancies with the PL. We conclude that agreement of marginal posteriors and PLs is sufficient for our purposes, while acknowledging that further work is needed to resolve when differences arise.

5.3.1.3 Case study 5.2: bifurcating regime, two curves

We now consider a dataset for two individuals generated from two different (σ, δ) parameter regimes of Kuznetsov et al.'s QSSA model. The parameter values correspond to Hypothesis 1 (H1); they differ between individuals only in the effector cell death rate, δ, but give rise to elimination and escape (see Fig. 5.3c).

For both hypotheses, we recover unimodal posteriors (see Figs. 5.16 and 5.17) with good diagnostics (not shown). For H1, when effector data are not available, we notice a marked improvement in identifiability of all parameters compared to fitting individual curves separately (see Fig. 5.18). When effector data are available, this is mostly the case for population parameters, $e_0^{(\text{pop})}$ and $m_0^{(\text{pop})}$. For H2, parameters σ, δ, $e_0^{(1)}$ and $e_0^{(2)}$ have CRIs that are, at small noise levels, comparable to analogous parameters of H1 fit, but they increase more slowly as observation noise increases, and are, at high noise levels, much smaller in comparison (see Fig. 5.18). Slow growth of uncertainty can also be seen for $m_0^{(1)}$ and $m_0^{(2)}$ for H2. We attribute this to the shape of the parameter region in which the H2 fit is found, noting that no region of parameter space will perfectly correspond to the synthetic dataset.

For H1, the MLE/MAP (interchangeable as they match, see Fig. 5.16) of $\sigma^{(2)}$ and $\delta^{(2)}$, associated with individual 2, lie close to the true parameter values (see Fig. 5.15a), and the predicted model solutions within the CRI fit the data of individual 2 well, with small variability (see Fig. 5.14b). In $(\sigma^{(1)}, \delta^{(1)})$ space, associated with individual 1, PLs lie along a straight line, and span the bistable parameter regions and the monostable region with a stable zero-tumour steady

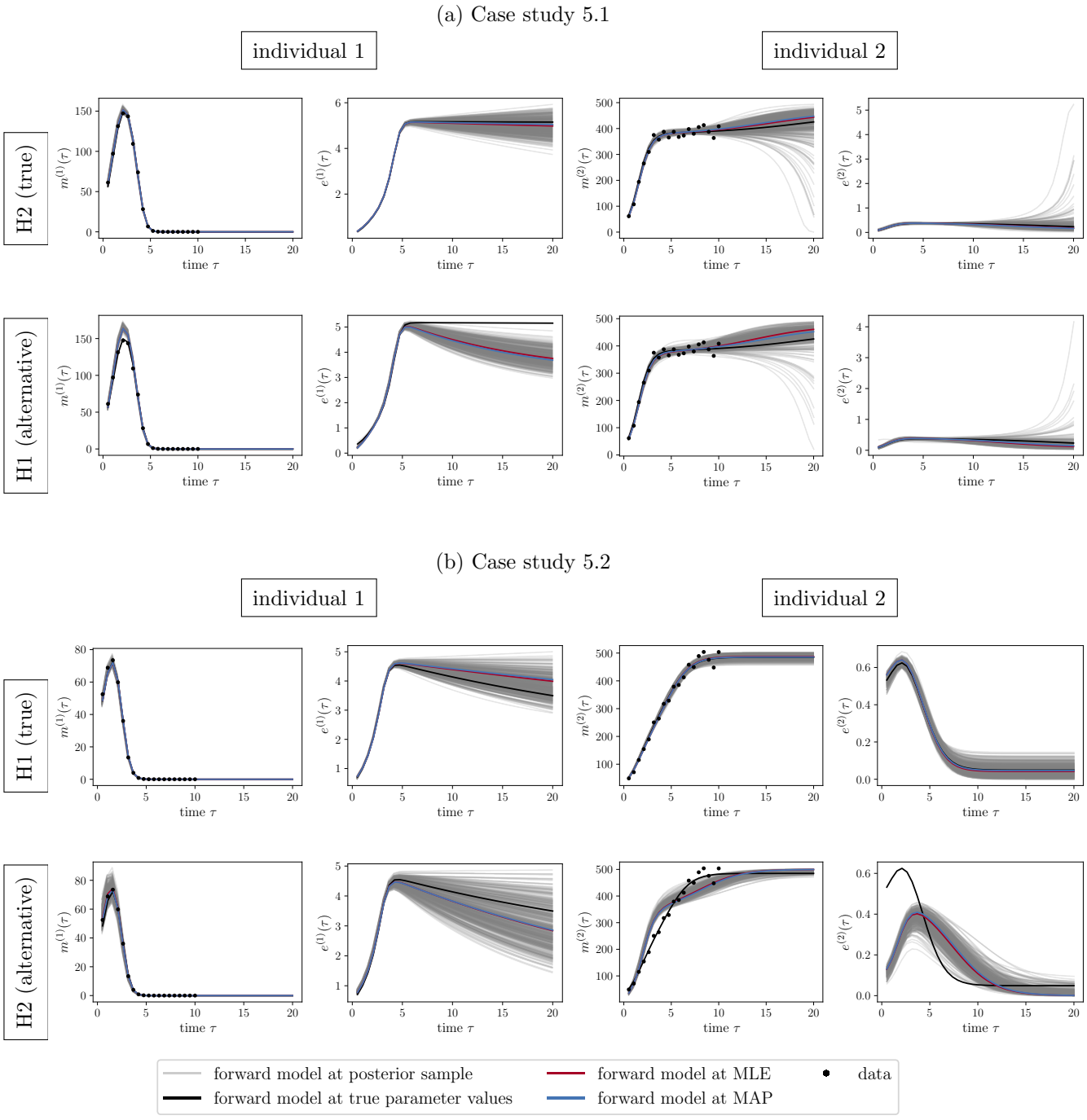
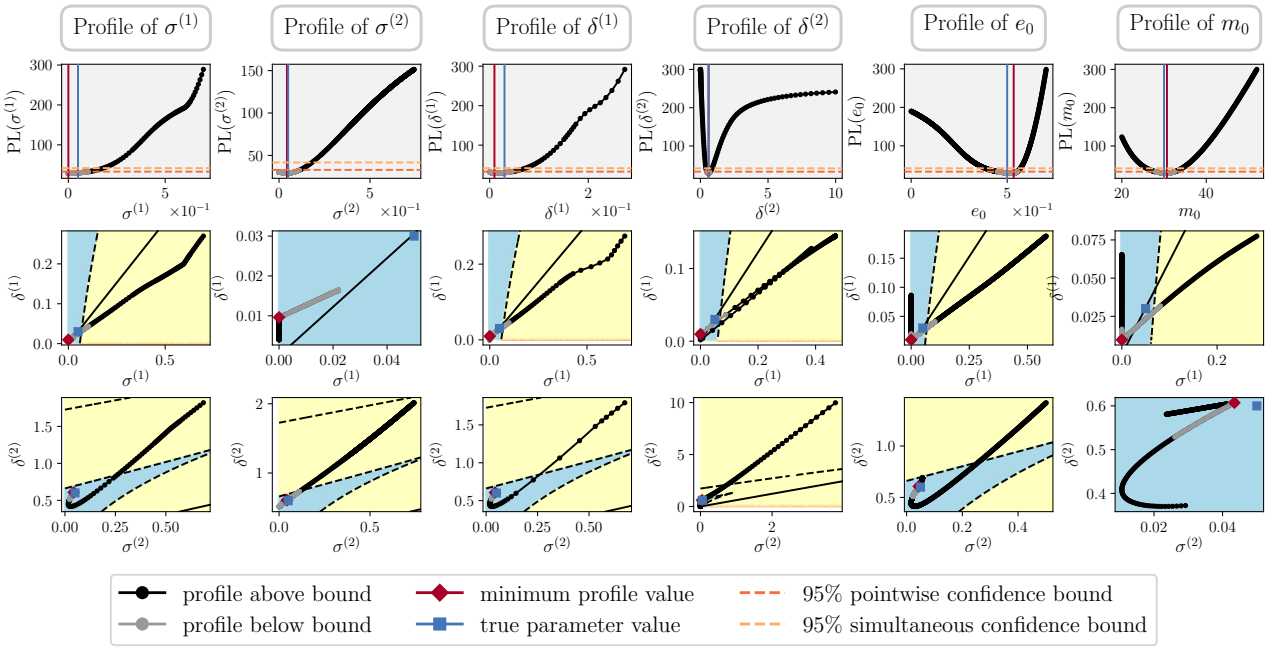


Figure 5.14: Series of plots showing the sensitivity of forward model solutions for parameters sampled from the posterior given data generated from (a) a **bistable** regime (**case study 5.1**), or (b) a **bifurcating** regime (**case study 5.2**), with observation model $h_1(\cdot)$, standard deviation $\xi = 0.05$ and under prior 3. In subplots (a) and (b), the top and bottom panels correspond to fits according to true and alternative hypotheses, respectively. In the time window of the data, fits under the true hypotheses produce solutions that correspond well to tumour cell data. Under the alternative hypothesis, solutions correspond well to tumour cell data for one individual, but exhibit small discrepancies from the data from the other individual and the corresponding effector cell data-generating dynamics.

(a) Hypothesis 1 (true)



(b) Hypothesis 2 (alternative)

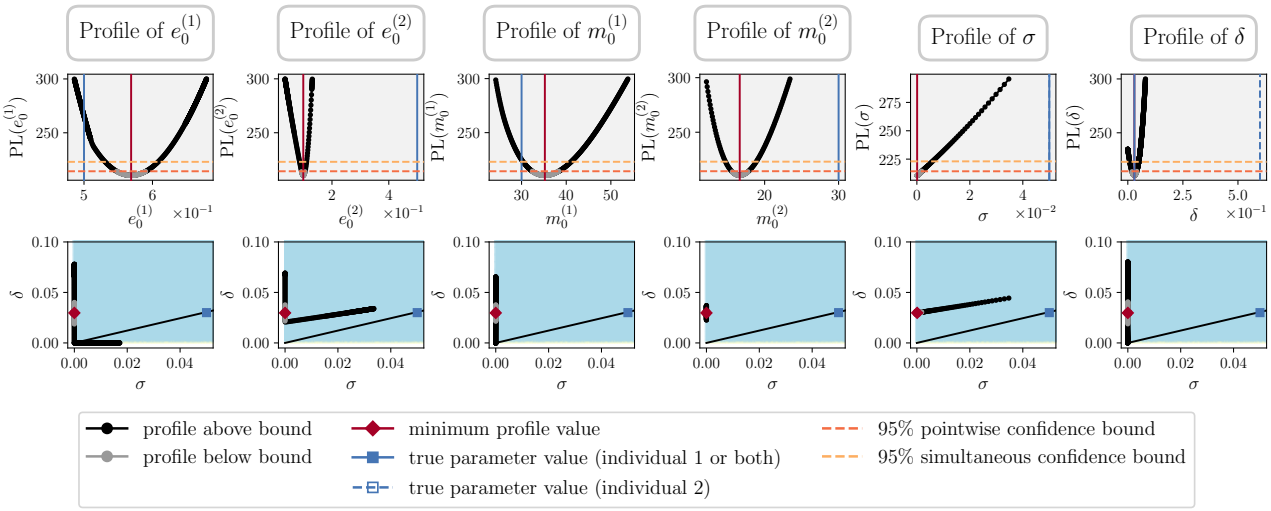


Figure 5.15: Series of plots showing profile likelihoods for a dataset generated from two regimes with different **bifurcating** behaviour (**case study 5.2**), with observation standard deviation $\xi = 0.05$ and for observation model $h_1(\cdot)$, fitted according to (a) H1 (true) and (b) H2 (alternative) hypotheses. Bottom panels show which bifurcation regions the profiles explore. Maximum likelihood estimates of H1 fit correspond well to the true parameter values. The confidence region (grey) in (σ, δ) space of H2 fit lies close to the confidence region in $(\sigma^{(1)}, \delta^{(1)})$ of H1 fit, and is better adjusted for individual 1 (elimination).

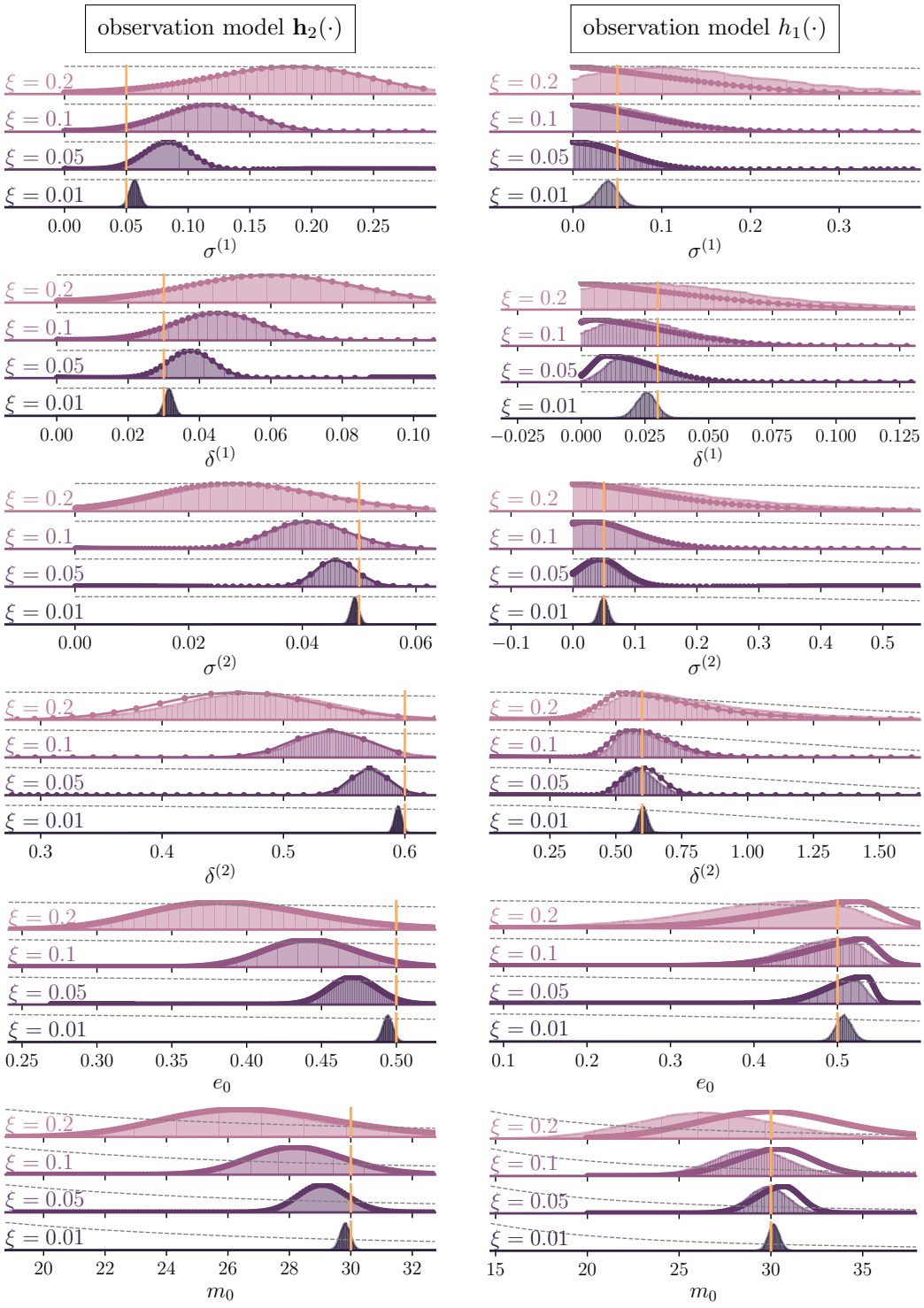


Figure 5.16: Series of plots showing marginal posterior distributions and profile likelihoods (PLs) given datasets generated from two regimes with different **bifurcating** behaviour (**case study 5.2**) with different noise levels, and given **prior 3** and **H1** (true). Marginal posteriors and PLs agree for observation model $\mathbf{h}_2(\cdot)$, but show slight discrepancies for observation model $h_1(\cdot)$. Note that PLs and priors are scaled so their maxima in the interval of interest correspond with the maximum of the marginal posterior density. Legend: light shaded histogram = marginal posterior, solid curve with points = PL, dashed curve = prior, orange vertical line = true parameter value (for the population parameters, i.e. those fixed over the population during fitting, the true data-generating values of the corresponding individual parameters of both individuals are marked).

state (see Fig. 5.15a). For observation model $h_1(\cdot)$ and $\xi \geq 0.03$, parameter $\sigma^{(1)}$ is pushed to the edge of the bistable region, where $\sigma^{(1)} \approx 0$. For parameter values within the CI of $\sigma^{(1)}$ and close to $\sigma^{(1)} \approx 0$, the elimination curve matches equilibrium-type model trajectories that oscillate towards the small-tumour steady state, but have long-periods of oscillation as they pass the neighbourhood of the zero-tumour steady state. These trajectories provide a good match to elimination-type behaviour in the initial time period for which data are available (see Fig. 5.14b), but eventually diverge from the data-generating trajectory (not shown). We speculate that this is since the data-generating parameter values lie near the bifurcation curve, $\delta = \delta^{\diamond}(\sigma, \alpha, \kappa_2)$, on which the equilibrium steady state emerges via a transcritical bifurcation from the zero-tumour steady state.

For H2, fits for both individual's data are found on the edge of the bistable region of equilibrium and escape, where $\sigma \approx 0$. The CI for (σ, δ) thus lies on the edge, and is close to, or at least partly covers, the CI for $(\sigma^{(1)}, \delta^{(1)})$ under H1. Again, the elimination data for individual 1 match well with equilibrium-type model trajectories, with long-periods of oscillations via $m \approx 0$. By comparison, the model fit to the escape data for individual 2 in this parameter region is poorer. While all solutions capture the correct qualitative behaviour, they do not cover the range of escape data fully for tumour and effector measurements (see Fig. 5.14b). Moreover, the edge-bound confidence region explains the small CIs of σ and correlated parameters for this fit, and demonstrates how a fit to a misspecified model may result in small uncertainties.

We also note that marginal posteriors for H1 match the PLs very well (see Fig. 5.16), while, for H2 (alternative), there are discrepancies (see Fig. 5.17), e.g. shorter tails in PLs and differences between MAP and MLE estimates, particularly for $e_0^{(2)}$ and $m_0^{(2)}$ for observation model $h_1(\cdot)$. A likely explanation for these discrepancies is that in the Bayesian problem the standard deviation $\boldsymbol{\xi}$ is an unknown parameter inferred from the data, whereas in the MLE problem it is

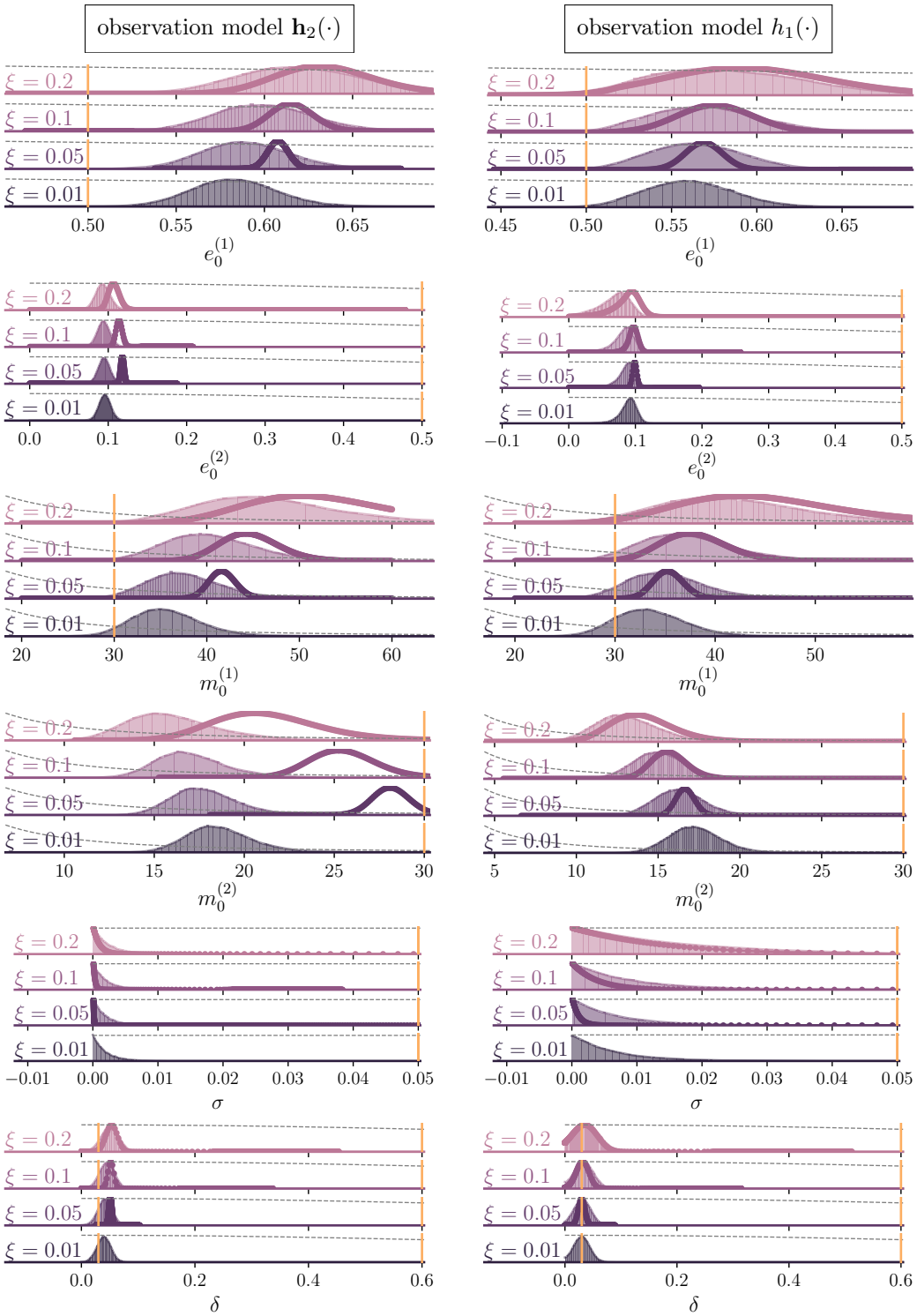


Figure 5.17: Series of plots showing marginal posterior distributions and profile likelihoods (PLs) given datasets generated from two regimes with different **bifurcating** behaviour (**case study 5.2**) with different noise levels, and given **prior 3** and **H2** (alternative). Marginal posteriors and PLs exhibit large discrepancies (e.g. modes and width of regions of high density) for observation model $\mathbf{h}_2(\cdot)$, but these are smaller for $h_1(\cdot)$. Note that PLs and priors are scaled so their maxima in the interval of interest correspond with the maximum of the marginal posterior density. Legend: light shaded histogram = marginal posterior, solid curve with points = PL, dashed curve = prior, orange vertical line = true parameter value (for the population parameters, i.e. those fixed over the population during fitting, the true data-generating values of the corresponding individual parameters of both individuals are marked).

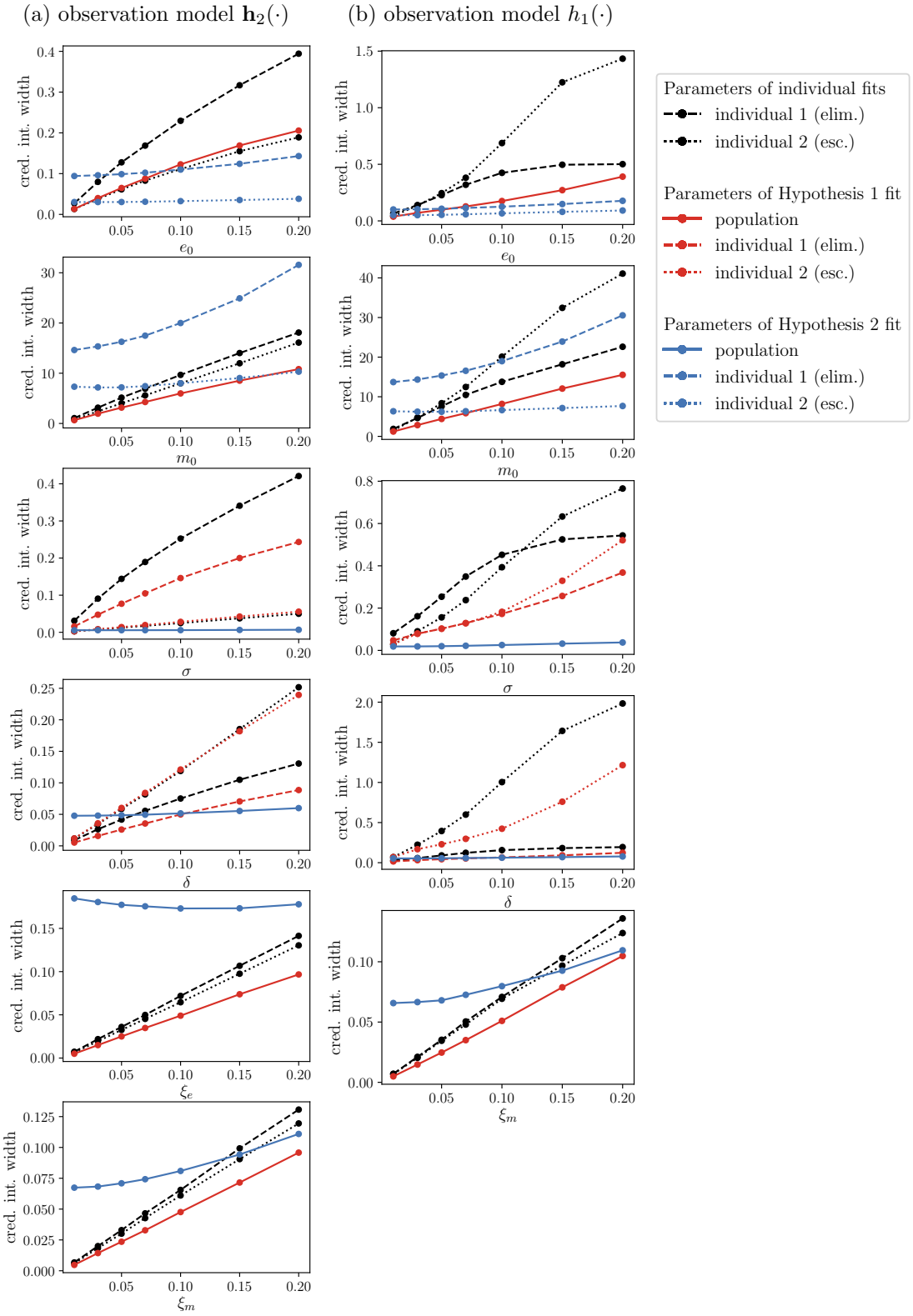


Figure 5.18: Series of plots showing the width of credible intervals for model fits to the pooled dataset for **case study 5.2** under H1 and H2 hypotheses, and for separate fits to each individual's data in this dataset. Credible intervals decrease for population fits under the true hypothesis (H1) in comparison to individual fits, whereas the decrease in uncertainty is not definitive for fits under the alternative hypothesis (H2).

assumed known. Treating $\boldsymbol{\xi}$ as unknown increases the number of degrees of freedom and parameter uncertainty. This should be investigated further in future.

5.3.2 Comparison of fits

For pooled datasets with two curves, we used MCMC to compute samples from the posterior distribution given Hypotheses 1 and 2. We used these samples to compute metrics of predictive accuracy of the fits (see Section 5.2.4). For every dataset of varying noise levels and observation models, a hypothesis test was performed to determine whether the estimated metrics are statistically different between model

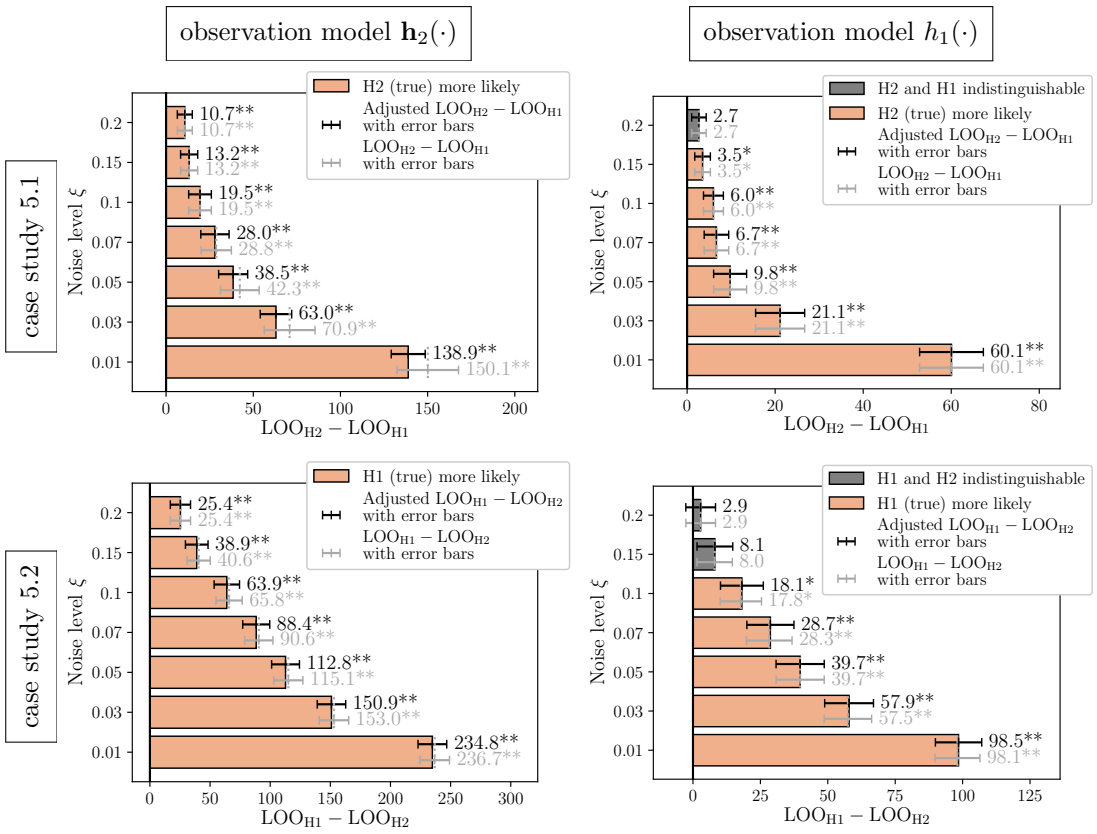


Figure 5.19: Series of bar plots showing differences in predictive accuracy (LOO and adjusted LOO) for models fitted according to the true and alternative hypotheses, and results of hypothesis tests for the differences being positive. The top and bottom rows show results for case studies 5.1 and 5.2, respectively. $*$ and $**$ denote when p-values of hypothesis tests are smaller than significance levels $\alpha_1 = 0.05$ and $\alpha_2 = 0.01$, respectively. Datasets with effector cells (observation model $\mathbf{h}_2(\cdot)$) are more likely to have been generated according to the true hypothesis for all noise levels. This also holds for datasets without effector cells (observation model $h_1(\cdot)$) and most noise levels, whereas for the highest noise levels, there is not enough evidence to distinguish the hypotheses.

fits according to different hypotheses.

Case study 5.1. Datasets of interest were generated according to H2, from a bistable parameter regime. Results indicate that almost all of these datasets are more likely to have been generated from H2 than H1, based on higher values of the non-adjusted and adjusted LOO metrics (see Fig. 5.19), and also WAIC (results for WAIC not shown as they match those for LOO, unless otherwise stated below). The difference is significant, as the computed p-values of hypothesis tests are all below significance levels $\alpha_1 = 0.05$ and $\alpha_2 = 0.01$. The true data-generating hypothesis (H2) can be distinguished with good confidence from the alternative hypothesis (H1) for a dataset from this parameter regime and the chosen prior. For observation model $h_1(\cdot)$ and $\xi = 0.2$, while the difference in WAIC is significant, it is not so for LOO metrics. Using the latter metric, the null hypothesis cannot be rejected with significance level $\alpha_1 = 0.05$, and the two fits are indistinguishable.

Case study 5.2. Here datasets are generated according to H1, from two different parameter regimes, with identical initial conditions. Results largely agree with those of case study 5.1. For most cases, hypothesis tests clearly reject the null hypothesis that H2 (alternative) is a more likely explanation for the dataset than H1 (true) (see Fig. 5.19). In the absence of effector data, we cannot distinguish between the two hypotheses for $\xi = 0.15$ and $\xi = 0.2$ using LOO and WAIC metrics.

In both case studies, we detect potential instabilities in the PSIS procedure for some posterior distributions (mainly alternative fits), which may affect the accuracy of the LOO estimates. The problematic points are some measurements taken at $\tau = \tau_1$, for which the estimated value of the parameter $\hat{k}_i$ of the Pareto distribution, fitted to the largest 20% of raw importance weights, $r_i^{(n)} = 1/p(s_i \mid \boldsymbol{\phi})$, exceeds the threshold of 0.7 recommended by Vehtari et al. [169]. The LOO predictive accuracies for these data points, $p(s_i \mid \mathbf{s}_{-i})$, are estimated by resampling the posterior distribution for a dataset without each problematic i-th data point

separately, and the LOO metric is adjusted accordingly. The differences between adjusted and non-adjusted LOO values are minimal (see Fig. 5.19), and do not alter the overall conclusions. We also note that due to high variance of pointwise log-likelihoods for some points (above 0.4), there is reason to doubt the reliability of the WAIC metric.

While both case studies demonstrated that for synthetic data the true data-generating processes, driven either by model bistability or bifurcations, can be identified, we noticed that, at low observation noise, the level of mismatch (i.e. the difference in the metric) is higher for bifurcating datasets (case study 5.2) than for bistable ones (case study 5.1). This may indicate that the assumption of fixed kinetic parameters over the population constrains model behaviour considerably more, than the analogous assumption for initial conditions. So while the number of unknown parameters for both hypotheses is the same, the richness of behaviour of the population model may be reduced under H2.

5.4 Discussion and conclusions

In this chapter we extended the work presented in Chapter 4 by using a population approach to fit the QSSA model of tumour-immune interactions by Kuznetsov et al. [88] to time series data. We focused on pooled synthetic datasets with dichotomous tumour growth curves from two individuals, each representative of one of the three stages of immunoediting – elimination, equilibrium or escape [47].

The key results are based on two case studies with synthetic data generated from Kuznetsov's model. We generated the first dataset from a parameter regime with bistability of escape and elimination steady states, and achieved different qualitative behaviour of individuals by choosing initial conditions that lie on different sides of the phase-space separatrix. The second dataset was generated such that both individuals have the same initial conditions, but their kinetic parameters lie in

disjoint regions of the bifurcation diagram that exhibit stability of different steady states, elimination versus escape. We applied Bayesian inference-based model selection to each dataset in order to investigate whether model fits under different assumptions about the underlying cause of dichotomous tumour responses may be distinguished based on their predictive performance across different experimental conditions (noise and availability of effector cell measurements).

We sampled posterior distributions given pooled datasets with the popular Hamiltonian Monte Carlo (HMC) algorithm [160]. We considered several cases, varying which parameters are "individualised" (those that vary between individuals) and which are fixed for the entire "population", as in *mixed effects modelling* (MEM) [93]. We assumed that either kinetic parameters or initial conditions vary between the two individuals to explain their dichotomous tumour growth dynamics using Kuznetsov's model; these assumptions represent different model-fitting hypotheses.

Data-generating versus alternative hypothesis. Our selection of the hypothesis that is most likely to explain the data was based on the approximate leave-one-out (LOO) and Watanabe-Akaike information criteria (WAIC) metrics using an approach from [169]. For both datasets the fit to the true data-generating hypothesis had a significantly higher estimated predictive accuracy for both metrics. The difference in accuracy decreased when effector data were withdrawn and/or observation noise was increased. The hypotheses became indistinguishable for data with high noise levels ($\xi \geq 0.15$) and only tumour cell observables. We concluded that for synthetic data with two individuals and under additional assumptions (e.g. the prior was chosen so that only one posterior mode was considered due to the difficulties of sampling from multiple modes) it is possible to determine the source of variability in tumour responses, i.e. variability in initial cell numbers versus variability in the functioning of the immune system. We discuss below the simplifications made in our study, and highlight the lessons learned with regard to

parameter inference, and also other considerations to be made when extrapolating these findings to other contexts and datasets.

Inference for individual curves. By first using HMC to fit individual curves from Chapter 4 under different priors, and by comparing results with profile likelihoods, we established that priors with decreasing probability as the parameter approaches zero (e.g. lognormal) can bias our results for Kuznetsov's model, and artificially constrain the posterior in a way that appears more identifiable than it is. Zero-mean Gaussian or Cauchy priors, truncated to positive real numbers, had a more neutral effect for this model, and allowed efficient and accurate sampling from unimodal posteriors for elimination and equilibrium data from Chapter 4.

For less identifiable cases, and particularly for multimodal likelihoods, we found that combinations of the likelihood, the prior and the sampler can significantly impact the posteriors sampled. For escape data from Chapter 4, two disjoint regions of high posterior density are separated by an energetic barrier, which HMC may be unable to breach. The sampler can thus get stuck in regions of low probability or bias sampling towards one of the modes when both are similarly probable. A poorly informative prior may increase the likelihood of the former, whereas a more informative prior may artificially constrain the problem so that only the mode of interest is sampled. We showed that even a mildly informative prior and a poorly informative multimodal likelihood may interact to bias our sampling.

Individual versus population approach. Under the hypothesis used to generate the pooled dataset, parameter uncertainties were smaller than for separate fits of each individual's data, and the confidence region for the pooled data was contained within confidence regions associated with data for an individual. The decrease in uncertainty was most pronounced for the population parameters, but the less identifiable individual parameters associated with the escape curve also benefited

from the extra information provided by the elimination curve. Multimodality of the posterior, seen for individual fits of the escape curve, was not observed.

Under the alternative hypothesis, we had no knowledge of the ground truth, and sought the closest alternative fit. We found that parameter uncertainties depended largely on the size and shape of the parameter region in which the best fit was found, only marginally increasing with noise in some cases. The alternative confidence regions lay close to the parameter values of one individual, whereas the parameters of the other individual adjusted and lay in other parameter regions that match the qualitative behaviour of their data. The shapes of the marginal posteriors for individual kinetic parameters remained largely unchanged compared to individual fits. Multimodality, energetic barriers and flat regions persisted from individual fits, and became more pronounced as the fit adjusted better for one individual's data, increasing the difficulty of parameter identification.

Tackling multimodality. To overcome the difficulties associated with fitting multimodal distributions, we artificially imposed a prior that restricts sampling to the mode that is closer to the origin. As a result, our comparisons of fits were limited to a subset of forward model solutions that can reasonably fit the data. To extend results to solutions from both modes, sampling methods that can jump between the modes and overcome energetic barriers are needed. As this is an active area of research, a multitude of methods that are better suited to exploring multimodality have been proposed. Many involve sampling from flattened posteriors to lower any energetic barriers and allow movement between modes [48, 76, 187]. In future work it would be interesting to apply these methods to our problem.

Overestimation of predictive performance for time series data. In our study we used the LOO and WAIC metrics due to their low computational cost and suitability for posterior distributions of any shape and for small datasets. However,

because of the relatively high variance of pointwise log-likelihoods, WAIC may be unreliable in our case and overestimate predictive accuracy. The LOO metric may also exhibit inaccuracies for time series data, because the posterior tends to be less sensitive to removal of single data points, which may lead to overestimation of predictive accuracy. By contrast, the posterior may be very sensitive to removal of points at the start of the observation period. The Pareto smoothed importance sampling (PSIS) LOO procedure [169] can detect the sensitive points, and correct the LOO estimate by refitting the model with these points held out one at a time. However, alternative approaches are needed to remedy overestimation due to insensitive points.

Ideally, to use the model in a predictive setting, its performance should be tested on new, unseen data. Non-temporal and independent measurements are normally partitioned into training and test sets, used for model fitting and evaluation of model performance, respectively. For small datasets, where this is infeasible, cross-validation can be used, where the full dataset is repeatedly randomly partitioned, each time leaving out a different test set. For multiple time series from different individuals, as considered in this chapter, a more appropriate validation method involves fitting the model to data up to a chosen time point (training set), and then assessing model performance on later data points (test set); this approach is termed leave-future-out (LFO) in [21]. For estimation of model performance, the model can be refit to different training sets, where each set temporally precedes the corresponding test set, but this is computationally demanding for Bayesian models.

Bürkner et al. [21] recently proposed the PSIS-LFO method, an extension of the efficient PSIS-LOO method [169], for temporal dataset partitions. While PSIS-LFO can correct for overestimation and possesses the advantages of PSIS-LOO, a level of temporal resolution of data is still required, but difficult to achieve in the tumour-immune context of interest. We view this is as a challenge for performing model selection in this context. We also highlight that model performance may

vary between different temporal dataset partitions because of dynamical differences in solutions of Kuznetsov et al.'s model that may impact identifiability in different partitions. These questions are interesting avenues for future work.

Implications for real datasets and model discrepancy. Our limited understanding of the mechanisms generating real experimental data means that the model will always be an imperfect representation of reality. This difference is typically called "model discrepancy" or "model mismatch" [177]. Model calibration to synthetic datasets under the alternative hypothesis in our study represents a situation where the data-generating process and the model do not match. We argue that distinguishing hypotheses from real data would correspond to comparison of two alternative hypotheses, and speculate that such hypotheses would be less distinguishable than in our study. Lei et al. [95] highlighted that ignoring model discrepancy can result in inaccuracies in quantification of uncertainty of parameter estimates and hence model predictions. To remedy this, model discrepancy could be considered in our modelling, which would also bring our synthetic studies closer to real datasets. How to include discrepancy depends on the context and aim, and remains an open question.

Summary and extensions. In this chapter we have shown that the correct "population" model, with regards to variability of either initial conditions or kinetic parameters, can be distinguished from a misspecified one given synthetic data of two curves generated from the model. While our findings cannot be easily generalised as we examined only two types of datasets and made several simplifications, they suggest that dichotomous solutions generated from Kuznetsov's model for two individuals are sufficiently distinct to be distinguishable from the proposed alternatives. We also show that while HMC is an efficient method for exploring unimodal posteriors, even for misspecified models, multimodality is a concern for complex models and requires refined inferential methods.

There are several related questions that could be explored in future. For example, whether or not bistable datasets are easier to fit to a bifurcating hypothesis with good accuracy than vice versa. Due to the model's complex bifurcation structure, population constraints for some parameters may lead to greater reduction in richness of model behaviour than for others. There is also a question of how hypothesis distinguishability extends to datasets with larger numbers of individuals, and whether good alternative fits are more difficult to find for such datasets. It would be interesting to investigate how the classic MEM approach, where individual parameters have prescribed distributions, with "population" hyperparameters that are estimated from the data, compares to the approach adopted in this chapter, particularly for datasets with larger numbers of individuals.

Chapter 6

Discussion

6.1 Summary and discussion

In this thesis we investigated ODE models describing tumour-immune interactions. Our motivation throughout was preclinical tumour data from AstraZeneca (AZ) which showed a dichotomous response to treatment with an immune checkpoint inhibitor; some tumours shrank following treatment, while others grew uncontrollably. We studied mechanistic models that can, through model bifurcations and/or bistability, capture the observed heterogeneity. The thesis comprised two parts: in Chapters 2 and 3 we characterised a well-known model from the literature using a dynamical systems approach, while in Chapters 4 and 5 we addressed questions relating to model calibration and hypothesis testing.

In Chapter 2 we examined a classical model of tumour-immune interactions proposed by Kuznetsov et al. [88], in which tumour and effector cells bind (reversibly) to form conjugates, and may become inactivated/damaged as a result of interactions in the conjugate. We identified that a separation of timescales arises in the model under the assumption that conjugate formation and dissociation without any cells in the conjugate being damaged occur on a shorter timescale than other processes. We used the method of matched asymptotics to derive a new two-dimensional approximation of the original model, which we validated against the full model. The dynamics of conjugates saturate on the short-timescale, and vary parametrically

with total tumour and immune cell numbers on the long-timescale.

With a phase-plane analysis we established that the reduced model exhibits excitable dynamics (i.e. following small perturbations the system undergoes large phase-plane excursions before relaxing to a steady state). They arise similarly as in the Fitzhugh-Nagumo model of the dynamics of a spiking neuron [116], due to a combination of the nullcline structure and a large ratio of timescales associated with tumour and immune dynamics, the latter of which are slower. The assumption of tumour cell logistic growth also contributes to excitability in our reduced model as it enables tumour cells to respond quickly to changes in immune cell numbers.

In Chapter 3 we investigated the long-term dynamics of our reduced model. We identified and characterised its steady state solutions, their linear stability and bifurcation structure. We showed that the model admits steady states which correspond to the three stages of immunoediting: elimination (i.e. no tumour), equilibrium (i.e. a dormant tumour controlled by the immune system), escape (i.e. a large tumour). We also identified limit cycle solutions with large periods and amplitudes that are typical of excitable systems and arise due to a homoclinic orbit.

By combining analytical results with numerical parameter sweeps, we identified disjoint regions of a two-dimensional parameter subspace in which the long-term model dynamics are distinct, with stationary or periodic solutions that are mono- or bistable. A complex bifurcation structure emerges and shows how tumour-immune outcomes change as parameters associated with the strength of the immune system, and/or initial conditions, vary. For patients with strong immune systems, tumours are eliminated, whereas for patients with weaker immune systems, outcomes such as tumour equilibrium, periodic tumour recurrence and tumour escape are dominant. By varying parameters associated with the immune response to the tumour we explored the potential effects of immunotherapy. We found that the tumours of patients with weak immune systems could be controlled at equilibrium and that

patient specific combination treatments that also increase the patient's immunocompetence could eliminate the tumour. These conclusions are consistent with ongoing efforts to identify effective clinical treatment combinations for patients, and ultimately to personalise treatment for each patient. Mechanistic models of cancer will play a key role in achieving this goal, but only if individual patient data can be utilised for their calibration [80, 140].

Asymptotic analysis enabled us to reduce the full model to two dimensions. We showed that, in the limit of fast conjugate dynamics, our reduced model more accurately represents the full model by Kuznetsov et al. [88] than its quasi-steady-state approximation (QSSA). While the long-term behaviour of the two models agrees in some parameter regimes, discrepancies of the QSSA from the full model can be noticed in numerical solutions, in the bifurcation structure and in the absence of excitable dynamics. QSSAs are often used for model simplification without careful checks of their validity. Our study highlights the need to carefully justify their use.

To our knowledge, our study was the first to identify and characterise excitable properties in models of tumour-immune interactions. Several phenomena seen in cancer and its response to therapy were found to be qualitatively consistent with the excitable behaviour. The model thus suggested that excitability may underpin tumour-immune interactions. The question is whether this hypothesis is supported by other models, and if it could be observed experimentally.

While mechanistic models can provide insights into tumour-immune interactions, they are often too complex to validate against limited data, which typically capture only tumour volume time series. Model excitability, highlighted in Chapter 2, increases the complexity further. While fitting of excitable models has been carried out in cardiac modelling, in immuno-oncology, challenges with calibration of simple models have not yet been sufficiently investigated. In Chapters 4 and 5 we studied the problem of calibrating a bifurcating model of tumour-immune interactions to

dichotomous datasets. Our focus was the QSSA model by Kuznetsov et al. [88] as it is mathematically and dynamically simpler that the original model, but still captures key features, such as the three Es of immunoediting.

In Chapter 4 we investigated whether the parameters and dynamics of the QSSA model can be recovered uniquely, with small uncertainty, from synthetic data generated from the model. Of interest was whether the parameter estimation problem was structurally and practically identifiable, as that would mean increased robustness of parameter estimation from real data. We found that given perfect measurements of tumour numbers, noise-free and continuous in time, model parameters (including the initial system state) can be uniquely determined, i.e. the problem is structurally identifiable.

We then examined practical identifiability, considering also the noise and spacing of measurements in time. We formulated a maximum likelihood parameter estimation problem for a subset of model parameters and initial conditions, chosen so that the model exhibits key features, such as immunoediting, bifurcations and bistability. We characterised the problem using profile likelihoods [125, 134]. Roesch and Stumpf [142] previously investigated parameter identifiability of simple, one-equation models with one parameter and co-dimension 1 bifurcations. Our work extended their analysis to a system which undergoes multiple bifurcations (mainly co-dimension 1), has multiple unknown parameters and is partially observed.

We performed case studies for synthetic datasets that we generated from parameter regimes that exhibit distinct model behaviour, with observables of only tumour cell numbers, or also immune cell numbers, and with different levels of measurement noise. We found that parameters related to the immune system are highly correlated in the model, and lead to large parameter confidence intervals for some datasets. Uncertainties in parameter estimates also appeared to be affected by the imposed parameter constraints and the model's bifurcation structure. We observed differences in parameter identifiability and its robustness for qualitatively

different datasets. Our results suggest that model calibration is more robust and less sensitive to the absence of immune cell data when the tumour cell dynamics exhibited by the data are rich (i.e. when they are affected by the immune system and logistic tumour growth is not dominant). For data showing tumour escape, sensitivity was large, and multiple disjoint regions of parameter space could fit some noisy datasets almost equally well. We concluded that while we may be able to recover well-fitting dynamics, low confidence in parameter estimates, or parameter sets which are close to phase-plane separatrices (potentially also bifurcations, but this requires further investigation), can make extrapolation of model predictions highly uncertain. Our recommendation for model calibration using dichotomous data was to either adopt a population approach or a simpler model.

Population approaches to data fitting, such as mixed effects modelling, are often used in pharmacology to explain drug uptake and effects, and how this affects disease progression [93]. Repeated measurements for different patients/animals are used simultaneously to fit models and account for inter-individual variability by assuming a statistical distribution of some parameters over the population. Motivated by this, the QSSA model's bifurcating/bistable behaviour, and AZ's dichotomous data, in Chapter 5 we proposed two distinct approaches/hypotheses to calibrating such models using population data. Inter-individual variability in tumour responses may be explained either by variability of initial conditions for a parameter regime with model bistability, or by variability in kinetic model parameters due to the bifurcations the model undergoes as these parameters vary (here initial conditions are assumed fixed over the population). We generated synthetic datasets from the model for two individuals ("population"). We investigated whether we may distinguish which of the two hypotheses has generated a synthetic dataset, and thus provide a mechanistic explanation for the dichotomy in tumour responses. Comparison of Bayesian parameter fits via different measures of predictive accuracy revealed that, for each dataset considered, the true data-generating hypothesis can

be distinguished from the alternative one. For fits according to the true hypothesis, we found that parameter uncertainty decreased compared to fitting those curves individually, while this was not necessarily the case for the alternative hypothesis.

Chapters 4 and 5 stand as a cautionary tale highlighting the possible pitfalls of calibration/uncertainty quantification for a model that exhibits bifurcations and/or bistability. Identifying potential challenges required a systematic approach with gradually increasing complexity, starting from perfect synthetic data, moving towards partial data that are noisy and time-spaced, and then adding Bayesian prior assumptions and adopting a population approach, all with the prospect of best utilising real data in future. Our study represents an attempt to relate features of the parameter estimation problem to the dynamical properties of the model and the data. While using experimental/clinical data to fit the model was outside the scope of this thesis, the work presented in Chapter 5 showed how a misspecified model may be fitted to data. Studies of this type are especially relevant when using real datasets, where there may be large uncertainties relating to the data-generating process. We showed that, for bistable/bifurcating models, misspecification may stem from assumptions of the population approach. As it is unclear how to identify and tackle such misspecification, we presented initial results of a Bayesian approach that may allow us to distinguish between different explanations for the data, related to the model's bifurcation structure.

6.2 Future directions

Our industrial partner AZ is particularly interested in explaining variable responses to anti-PD-L1 immune checkpoint therapy. Interest in modelling these therapies has also recently increased across the field [59, 82, 112, 118, 130, 153]. Immune checkpoint therapies aim to target T cell exhaustion by inhibiting signalling through co-inhibitory pathways that may impair effector T cell activities. The model

by Kuznetsov et al. [88] focuses on cellular interactions between tumour cells, activated immune cells, their complexes, programmed-for-death tumour cells and inactivated or exhausted immune cells. A natural extension would be to relate the cellular dynamics to sub-cellular processes, for example by accounting for signalling through the PD-1/PD-L1 axis that anti-PD-L1 therapy targets. We could follow Nikolopoulou et al. [118], where a constant proportion of tumour and effector cells are assumed to express PD-1 or PD-L1. Under the assumption that interactions through PD-1/PD-L1 are quasi-steady, the product of their densities inhibits T cell recruitment and proliferation, and decreases when anti-PD-1 treatment is applied.

Existing models view T cells as either activated or inactivated/exhausted [45, 88], but this is an oversimplification. Exhaustion is not a binary property of T cells; rather it develops gradually, with the level of T cell exhaustion correlating with the frequency of T cell expression of PD-1 receptors [166]. Progressive T cell exhaustion has been previously modelled by Dritschel [45], who assumed that T cell functions diminish with prolonged tumour exposure. Alternatively, one could account for this in Kuznetsov et al.'s model by assuming that the population of effector cells is structured by a continuous phenotypic variable that represents the level of exhaustion of a T cell, and, within this framework, immune checkpoint therapy would alter the distribution of this variable. In the context of tumour-immune interactions, phenotype-structured models, formulated in terms of coupled integro-differential equations, have been previously used to describe heterogeneity of tumour cell antigenicity [37] and phenotype polarisation of macrophages [49]. Of our interest would be to explore whether excitability can be detected in phenotype-structured versions of Kuznetsov et al.'s original model.

In deterministic ODE models, the outcome is fully determined by initial conditions and kinetic parameters. However, randomness of biological processes and variability of the tumour microenvironment impact cell interactions in ways not considered in such models. To account for intrinsic and/or extrinsic stochastic

effects, components that randomly fluctuate according to a stochastic process could be introduced in Kuznetsov et al.'s model. Stochasticity could have an especially large impact at small cell numbers, where it may contribute to development of small tumours or to their eradication, and affect probabilities of tumour recurrence and eradication. For stochastic systems, we may also derive the steady state distribution of tumour size, modes of which are states that the system spends most time in at large times. This distribution can undergo stochastic bifurcations and can greatly depend on the properties of extrinsic stochastic effects. An interesting question that follows on from our research is how different types of stochastic effects interact with excitability in our reduced model and affect its behaviour.

The proposed extensions increase model complexity in order to better describe the biology, but such models are more difficult to analyse and validate. With the limitations of data in mind, we next propose extensions that could advance analysis and calibration of simple models that are more fit for this purpose.

We studied identifiability of an existing model of tumour-immune interactions, but there are many others. The importance of identifiability analyses in this field is gaining recognition as being vital for robust model parametrisation [148]. An extension of our work would be to compare the structural and practical identifiability of existing ODE models of tumour-immune interactions, focusing on simple, two- or three-equation models (see, for example, models reviewed in Chapter 1 and [50]). This would allow us to tease out differences in identifiability, and highlight those which could be calibrated to tumour volume data alone.

In this thesis we explored whether the availability of immune data increase parameter identifiability and how this varies in qualitatively different synthetic datasets. We could explore other avenues related to experiment design. For example, we could investigate different time-spacing of measurements so that the captured dynamics in the data are more sensitive to changes in the least important parameters. Again, we would expect the optimal spacing to vary for qualitatively

different datasets. We could also consider additional observables, such as measuring the ratio of tumour versus immune cells.

Another research direction would be investigating indistinguishability of existing models based on different observables. In this thesis we investigated whether two curves with different qualitative behaviour, and generated from a bistable regime, can be distinguished from curves generated from parameter regimes separated by bifurcations, and vice versa. A broader question would be whether different models, rather than different formulations of the same model, can generate similar, indistinguishable observables. In this thesis we used a Bayesian inference-based approach, but alternative approaches that rely on differential algebra, linear algebra and statistics, could also be adopted (see, for example, [66]).

When characterising parameter estimation problems, the greatest challenge was finding an appropriate method for our problem, and recognising difficulties in the absence of the ground truth. Methods for structural identifiability analysis differ in the complexity of analytical expressions they can cope with [28]. A range of software packages [12, 27, 110] implemented within different programming languages exist, but this means that switching the tool used requires the user to reimplement the model within a different package and familiarise themselves anew with how to use it. Similar issues arise for practical identifiability and parameter estimation, for which the use of optimisation or inference algorithms is required. Different algorithms are suited for objective functions of specific formulations and shapes. Posterior multimodality is particularly challenging and new efficient methods that can tackle such problems are being continuously developed [e.g. 60, 187].

Since the user has poor knowledge of the properties of the parameter estimation problem *a priori*, their analyses would be facilitated by availability of different methods within a single software package, in order to easily trial different methods. Widely applicable packages for parameter estimation that implement different algorithms, or provide an interface to other algorithm implementations, have

started to appear [30, 161]. We see these as key to increasing our understanding of inverse problems posed by tumour-immune interactions models. When it comes to structural identifiability, there is also potential to increase the ease of applying methods for such analyses, and hence their utilisation. New methods that can deal with a wider range of models are also desired, with extensions to more complex models that cannot be expressed as differential polynomials, and those formulated statistically such as mixed effects models.

6.3 Final remarks

In this thesis we studied models of tumour-immune interactions with a rich bifurcation structure in order to increase understanding of their predictions and their calibration. We provided a new perspective on how qualitative differences between tumour growth dynamics of responding and non-responding individuals may be explained by a bifurcating and/or bistable model. We showcased how integration of mathematical and statistical methods can facilitate development of simplified models, their analysis and calibration. Modelling has enormous potential to support experimental cancer research and clinical applications, but its full potential is yet to be unleashed. Obtaining systematically validated models is key, but will require a multidisciplinary approach. Models must be developed with experimental data in mind. Our work highlighted a gap between understanding of the behaviour of mathematical models and how this affects parameter estimation in relation to available synthetic and real data. We showed that systematic approaches, utilising analytical and numerical methods, can be adopted to aid this. Further gaps exist in using insights from such analyses to inform experimental design and enable model validation. Bridging these gaps requires iterative model development and data collection, and close collaborations with experimentalists and clinicians.

Appendices

A Asymptotic approximation and conditions for physicality of the nonzero-tumour steady states of the asymptotically reduced model

Here we detail the derivation of asymptotic approximations of the nonzero-tumour steady states presented in Section 3.2.1, and discuss where they are physically realistic (real and nonnegative).

The nonzero-tumour steady states are defined as $(e^*, m^*) = (\frac{\alpha}{\kappa_2}(1-\beta m), m)$, where m is a root of the cubic (3.3a). We assume $0 < \xi \ll 1$, $\mu = \xi\hat{\mu}$ and $\beta = \xi\hat{\beta}$. Under this rescaling, (3.3a) becomes

$$p(m) = \xi^2\hat{a}_3 m^3 + \xi\hat{a}_2 m^2 + \hat{a}_1 m + a_0 + \xi^2\hat{a}_{21}m^2 + \xi\hat{a}_{11}m = 0, \tag{A.1a}$$

where

$$\begin{aligned} a_0 &= \eta(\tfrac{\sigma\kappa_2}{\alpha} - \delta), & \hat{a}_{11} &= \eta(\hat{\beta}\delta - \hat{\mu}), & \hat{a}_{21} &= \eta\hat{\mu}\hat{\beta}, \\ \hat{a}_1 &= \tfrac{\sigma\kappa_2}{\alpha} + \rho - \delta, & \hat{a}_2 &= -\hat{\mu} + \hat{\beta}(\delta - \rho), & \hat{a}_3 &= \hat{\mu}\hat{\beta}. \end{aligned} \tag{A.1b}$$

We seek approximate solutions to (A.1a) of the form $m = \xi^{-\gamma}\hat{m}$ with γ a constant and $\hat{m} = \mathcal{O}(1)$, giving

$$p(\xi^\gamma\hat{m}) = \xi^{2-3\gamma}\hat{a}_3\hat{m}^3 + \xi^{1-2\gamma}\hat{a}_2\hat{m}^2 + \xi^{-\gamma}\hat{a}_1\hat{m} + \hat{a}_0 + \xi^{2-2\gamma}\hat{a}_{21}\hat{m}^2 + \xi^{1-\gamma}\hat{a}_{11}\hat{m} = 0. \tag{A.2}$$

We find pairwise dominant balances among terms in the cubic to simplify it.

First, balancing the third and fourth terms gives $\gamma = 0$. With $\gamma = 0$, we obtain a leading-order solution $\hat{m}_2 = \mathcal{O}(1)$ to (A.2) of the form

$$\hat{m}_2 := -\frac{a_0}{\hat{a}_1} = \frac{\eta(\alpha\delta - \sigma\kappa_2)}{\sigma\kappa_2 - \alpha\delta + \alpha\rho}, \tag{A.3a}$$

which gives

$$m_2^* \sim m_2 \sim \hat{m}_2 \quad \text{and} \quad e_2^* \sim \frac{\alpha}{\kappa_2}(1 - \xi\hat{\beta}\hat{m}_2) \sim \frac{\alpha}{\kappa_2} =: \hat{e}_2. \tag{A.3b}$$

These approximate solutions correspond to an intermediate-sized-tumour steady state given by $(e_2^*, m_2^*) \sim (\hat{e}_2, \hat{m}_2)$. Noting $\rho > 0$, it is straightforward to show that this steady state is physically realistic if and only if $\frac{\sigma\kappa_2}{\alpha} \leq \delta < \rho + \frac{\sigma\kappa_2}{\alpha}$ (or equivalently $\frac{a_0}{\eta} \leq 0 < \hat{a}_1$).

We recover the other roots of the cubic by balancing the first, second and third terms of (A.2), by setting $\gamma = 1$. This gives leading-order solutions $\hat{m}_{3,4} = \mathcal{O}(1)$ to (A.2) of the form

$$\hat{m}_{3,4} := \frac{-\hat{a}_2 \pm \sqrt{\hat{a}_2^2 - 4\hat{a}_3\hat{a}_1}}{2\hat{a}_3}, \tag{A.4a}$$

so that

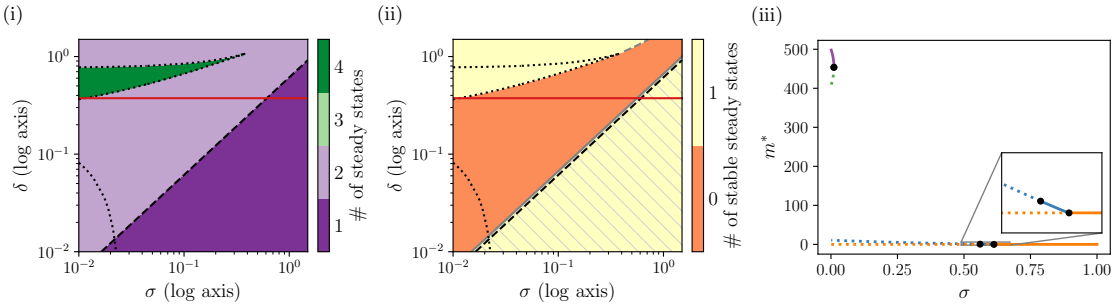
$$m_{3,4}^* \sim m_{3,4} \sim \frac{1}{\xi}\hat{m}_{3,4} \quad \text{and} \quad e_{3,4}^* \sim \frac{\alpha}{\kappa_2}(1 - \hat{\beta}\hat{m}_{3,4}) =: \hat{e}_{3,4}. \tag{A.4b}$$

These approximate solutions define two large-tumour steady states given by $(e_{3,4}^*, m_{3,4}^*) \sim (\hat{e}_{3,4}, \hat{m}_{3,4}/\xi)$. We note that $\hat{e}_{3,4}$ and $\hat{m}_{3,4}$ are both nonnegative if $0 \leq \hat{m}_{3,4} \leq \frac{1}{\hat{\beta}}$. Given $\hat{a}_3 > 0$, we identify regions of $(\hat{a}_1, \hat{a}_2)$ coefficient space with different numbers of real and nonnegative roots $\hat{m}_{3,4}$ (see Table A.1). This allows us to identify regions in which there are zero, one or two physically realistic steady states with large tumours. In summary, provided $\hat{m}_{3,4} \leq \frac{1}{\hat{\beta}}$, there is a single large-tumour steady state solution when $\hat{a}_1 < 0$, two such solutions when $0 < \hat{a}_1 < \hat{a}_2^2/4\hat{a}_3 =: \hat{a}^{\#}$ and $\hat{a}_2 < 0$, and none otherwise.

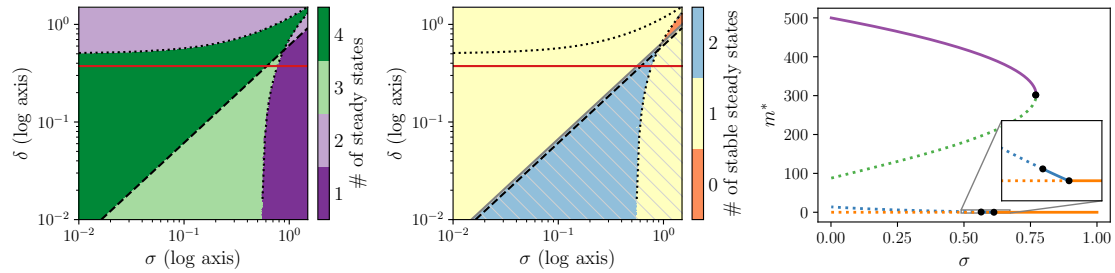
$\hat{a}_2$ \ $\hat{a}_1$	$\hat{a}_1 < 0$	$0 < \hat{a}_1 < \hat{a}_1^{\#}$	$\hat{a}_1^{\#} < \hat{a}_1$
$\hat{a}_2 < 0$	$\hat{m}_4 < 0 < \hat{m}_3$	$\hat{m}_{3,4} > 0$	$\hat{m}_{3,4}$ complex
$\hat{a}_2 > 0$	$\hat{m}_4 < 0 < \hat{m}_3$	$\hat{m}_{3,4} < 0$	$\hat{m}_{3,4}$ complex

Table A.1: Table showing where in $(\hat{a}_1, \hat{a}_2)$ coefficient space asymptotic approximations (A.4) of the large roots $m_{3,4}^*$ of the cubic (3.3a) are real and nonnegative. Note that $\hat{a}_1^{\#} := \hat{a}_2^2/4\hat{a}_3$.

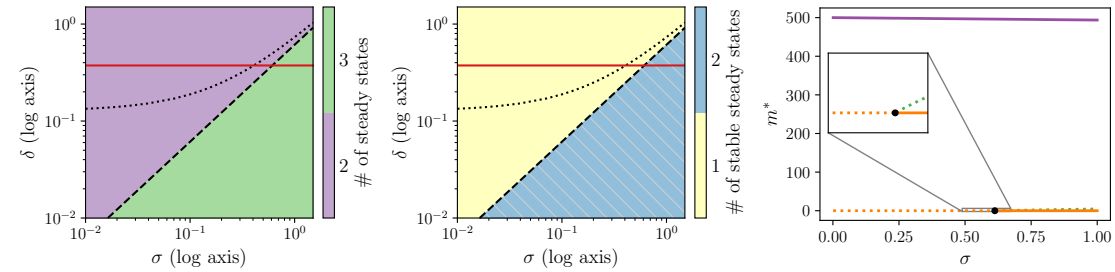
B Bifurcation structure of the asymptotically reduced model as parameter μ varies



(a) $\mu = \mu_0/1.8 = 0.0017$



(b) $\mu = 2\mu_0 = 0.006$



(c) $\mu = 32.2\mu_0 = 0.1$

Figure B.1: Series of plots (a)–(c) showing how the bifurcation structure of the long-timescale system (2.17) changes as μ is varied from its default value $\mu_0 = 0.003$. The effects of increasing μ on the system's bifurcation structure are similar to the effects of decreasing ρ (see Fig. 3.7). For each value of μ, we present three bifurcation diagrams: plots of (σ, δ) parameter space showing (i) regions with different numbers of physically realistic steady states (see Fig. 3.1 for legend and details), and (ii) regions with different numbers of stable steady states (see Fig. 3.6 for legend and details); (iii) bifurcation diagrams of tumour steady state numbers, m^*, as σ is varied and δ is fixed at its default value (see Fig. 3.2 for legend and details).

C Optimisation bounds

parameter ϕ_i associated with	σ	δ	e_0	m_0
lower bound	0	0	0	0
upper bound	50	50	200	200
scale factor	1	1	1	10

Table C.1: Optimisation bounds used for computation of profile likelihoods (Algorithm 1) in Chapters 4 and 5

D Profile likelihoods for case study 4.2b

See Fig. D.1.

E Sensitivity-based identifiability

Here we present preliminary results of sensitivity-based identifiability analysis [111] of Kuznetsov et al.'s QSSA model [88] in parameter regimes used for case studies 4.1-4.3 (see Table 4.1).

Let $\tilde{\mathbf{s}} := \mathcal{O}(\mathcal{F}(\boldsymbol{\theta})) = [\tilde{\mathbf{y}}(\tau_{k_1}, \boldsymbol{\theta})^T, \tilde{\mathbf{y}}(\tau_{k_2}, \boldsymbol{\theta})^T, \ldots, \tilde{\mathbf{y}}(\tau_{k_N}, \boldsymbol{\theta})^T]^T$ be model outputs. Then the normalised sensitivity matrix is $S(\boldsymbol{\theta}) := \left(\frac{\mathrm{d}\tilde{\mathbf{s}}_i}{\mathrm{d}\theta_j}\theta_j\right)_{ij}$. For each parameter θ_j, we can compute the overall sensitivity, $\mathrm{OS}_j = \sum_{i=1}^{N_s}(S(\boldsymbol{\theta}))_{ij}^2$. The lower the overall sensitivity for a particular parameter, the less locally sensitive model outputs are to it. Parameters ranked the lowest according to their overall sensitivity are considered unidentifiable. Alternatively, we can rank parameters via a *principal component analysis* (PCA) of the matrix $S(\boldsymbol{\theta})^T S(\boldsymbol{\theta})$, the eigenvalues and eigenvectors of which are computed using Singular Value Decomposition (SVD). Small eigenvalues indicate unidentifiable parameters or parameter combinations, depending on the associated eigenvectors (not shown due to limited space, but see caption of Fig. E.2). Large eigenvalue spectra have been related to model "sloppiness" [62].

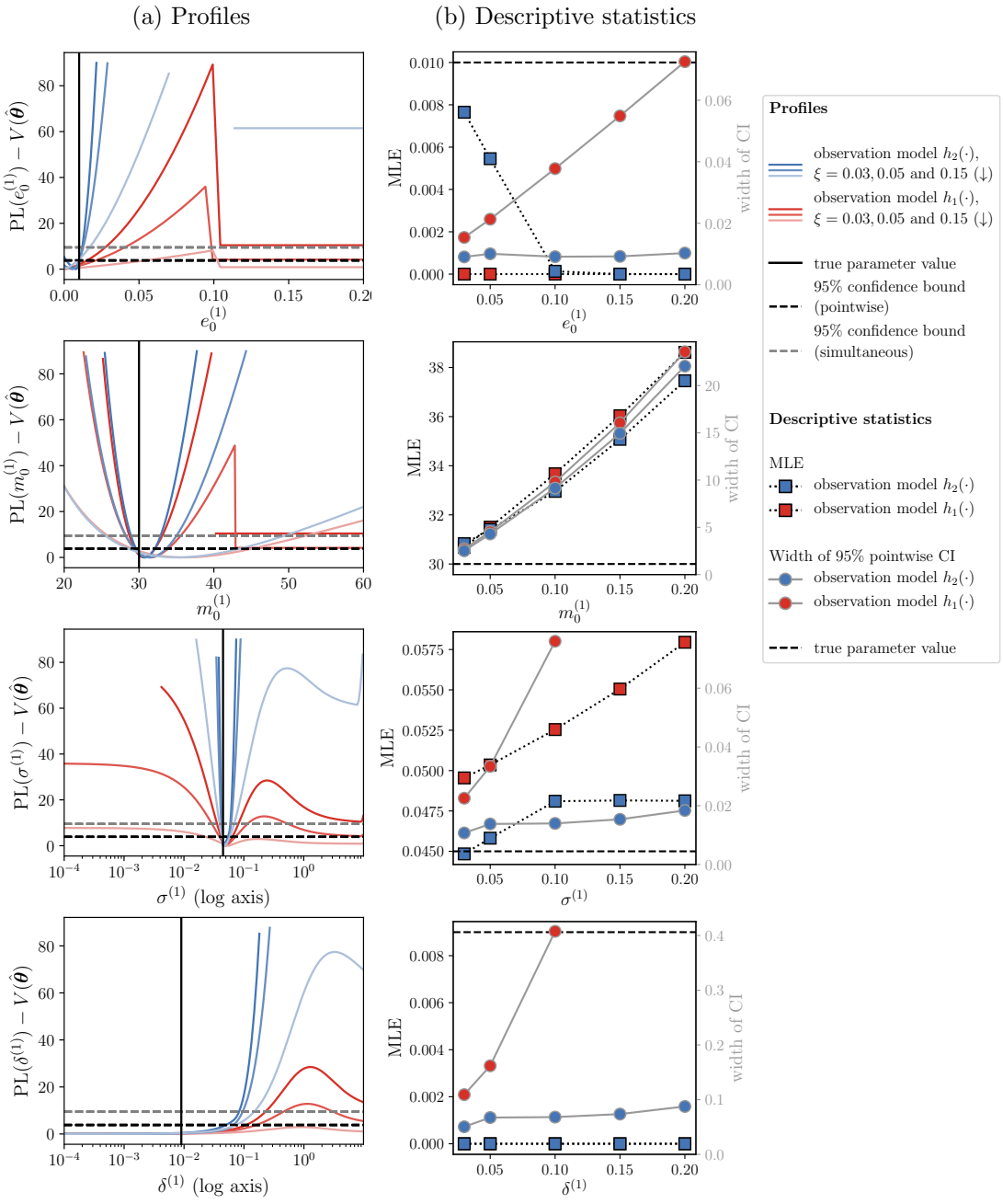


Figure D.1: Series of plots showing the impact of observation noise standard deviation ($\xi \in [0.03, 0.2]$) and observation models with and without effector cell measurements ($\mathbf{h}_2(\cdot)$ and $h_1(\cdot)$) on profile likelihoods for data generated from the **escape** regime (**case study 4.2b**). Panel (a) shows profile values translated by $V(\hat{\boldsymbol{\theta}})$ as parameters of interest in $\boldsymbol{\theta}_{\text{sub}} = [e_0, m_0, \sigma, \delta]^T$ individually vary, while panel (b) shows how the corresponding maximum likelihood estimates (MLE) and the size of confidence intervals (CIs) vary with increasing ξ, split by different observation models. Curves and points plotted in blue correspond to observation model $\mathbf{h}_2(\cdot)$, while those in red correspond to $h_1(\cdot)$. Profiles are similar to case study 4.2a (see Fig. 4.8b), but flatter. We note that, for $\xi = 0.1$ and $h_1(\cdot)$, where CIs are non-unique (disjoint) for each parameter, one of the CIs is unbounded and not considered in the computed CI widths. For $\xi > 0.1$ and $h_1(\cdot)$, the CIs are unbounded and their widths are not shown.

For Kuznetsov et al.'s QSSA model, overall sensitivities show that, in all four considered regimes, δ is the least sensitive parameter, followed by e_0 (see Fig. E.1). The sensitivity of outputs to individual parameters decreases considerably when effector data are not considered in escape regimes (4.2a-b), whereas it only slightly decreases in elimination and equilibrium regimes (4.1 and 4.3). This is visible from the overall sensitivities and eigenvalues of the sensitivity matrix and their range (Fig. E.2), and is in line with our findings in Chapter 4. Moreover, we highlight that the eigenvalue spectra are smallest for the equilibrium regime, and largest for escape regimes, which indicates that model "sloppiness" varies with different model outputs.

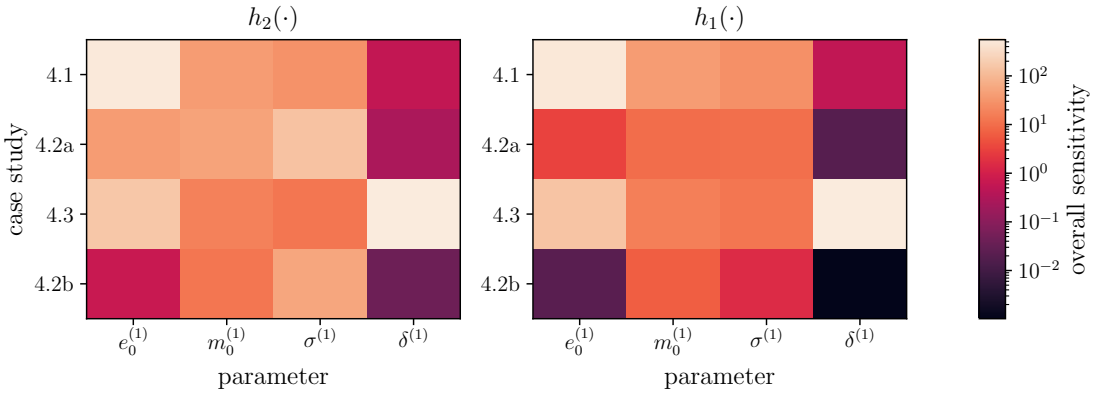
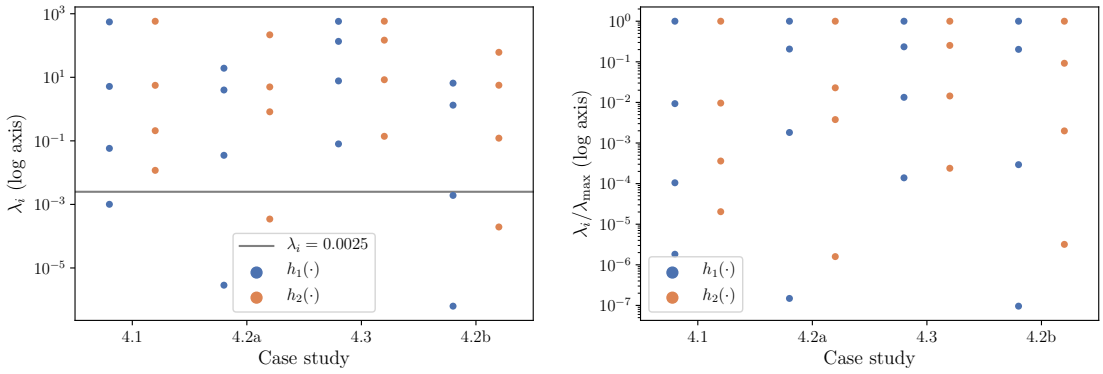


Figure E.1: Overall sensitivities of model outputs to model parameters for **case studies 4.1-4.3** and observation models $\mathbf{h}_2(\cdot)$ and $h_1(\cdot)$.



(a) non-normalised (b) normalised

Figure E.2: Eigenvalues, λ_i, of $S(\boldsymbol{\theta})^T S(\boldsymbol{\theta})$ obtained via SVD. The eigenvalue spectra are smallest in regime 4.3, and increase when effector outputs are not considered. In regimes 4.2a and 4.2b, the smallest eigenvalues are very small (below the arbitrarily chosen threshold), and the eigenvector associated with the smallest eigenvalue lies along the δ direction (not shown). In regime 4.1, the eigenvector associated with the smallest eigenvalue lies along δ and σ directions (not shown).

Bibliography

[1] M. Al-Tameemi, M. Chaplain, and A. D'Onofrio. "Evasion of tumours from the control of the immune system: Consequences of brief encounters". *Biol. Direct* 7 (2012), pp. 1–22.

[2] P. M. Altrock, L. L. Liu, and F. Michor. "The mathematics of cancer: Integrating quantitative models". *Nat. Rev. Cancer* 15.12 (2015), pp. 730–745.

[3] A. R. A. Anderson and P. K. Maini. "Mathematical oncology". *Bull. Math. Biol.* 80.5 (2018), pp. 945–953.

[4] M. Baar et al. "A stochastic model for immunotherapy of cancer". *Sci. Rep.* 6.1 (2016), p. 24169.

[5] B. Ballnus et al. "Comprehensive benchmarking of Markov chain Monte Carlo methods for dynamical systems". *BMC Syst. Biol.* 11.1 (2017), p. 63.

[6] E. N. Baruch, A. L. Berg, M. J. Besser, J. Schachter, and G. Markel. "Adoptive T cell therapy: An overview of obstacles and opportunities". *Cancer* 123.S11 (2017), pp. 2154–2162.

[7] D. J. Bearup, N. D. Evans, and M. J. Chappell. "The input–output relationship approach to structural identifiability analysis". *Comput. Methods Programs Biomed.* 109.2 (2013), pp. 171–181.

[8] G. L. Beatty and W. L. Gladney. "Immune escape mechanisms as a guide for cancer immunotherapy." *Clin. Cancer Res.* 21.4 (2015), pp. 687–92.

[9] S. Bekisz and L. Geris. "Cancer modeling: From mechanistic to data-driven approaches, and from fundamental insights to clinical applications". *J. Comput. Sci.* 46 (2020), p. 101198.

[10] N. Bellomo and M. Delitala. "From the mathematical kinetic, and stochastic game theory to modelling mutations, onset, progression and immune competition of cancer cells". *Phys. Life Rev.* 5.4 (2008), pp. 183–206.

[11] N. Bellomo, N. K. Li, and P. K. Maini. "On the foundations of cancer modelling: selected topics, speculations, and perspectives". *Math. Model. Methods Appl. Sci.* 18.04 (2008), pp. 593–646.

[12] G. Bellu, M. P. Saccomani, S. Audoly, and L. D'Angiò. "DAISY: A new software tool to test global identifiability of biological and physiological systems". *Comput. Methods Programs Biomed.* 88.1 (2007), pp. 52–61.

[13] A. Besse et al. “Stability analysis of a model of interaction between the immune system and cancer cells in chronic myelogenous leukemia”. *Bull. Math. Biol.* 80.5 (2018), pp. 1084–1110.

[14] M. Betancourt. *A conceptual introduction to Hamiltonian Monte Carlo.* Preprint. 2017. arXiv: `1701.02434`.

[15] M. Binnewies et al. “Understanding the tumor immune microenvironment (TIME) for effective therapy”. *Nat. Med.* 24.5 (2018), pp. 541–550.

[16] T. Bose and S. Trimper. “Noise-assisted interactions of tumor and immune cells”. *Phys. Rev. E* 84.2 (2011), p. 021927.

[17] T. Bose and S. Trimper. “Stochastic model for tumor growth with immunization”. *Phys. Rev. E* 79.5 (2009), p. 051903.

[18] F. Boulier. “Differential elimination and biological modelling”. *Gröbner Bases Symb. Anal.* 2 (2007), pp. 109–137.

[19] C. M. Boyer et al. “Heterogeneity of antigen expression in benign and malignant breast and ovarian epithelial cells”. *Int. J. Cancer* 43.1 (1989), pp. 55–60.

[20] F. Bray et al. “Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries”. *CA Cancer J. Clin.* 68.6 (2018), pp. 394–424.

[21] P.-C. Bürkner, J. Gabry, and A. Vehtari. “Approximate leave-future-out cross-validation for Bayesian time series models”. *J. Stat. Comput. Simul.* 90.14 (2020), pp. 2499–2523.

[22] A. Cappuccio, M. Elishmereni, and Z. Agur. “Cancer immunotherapy by interleukin-21: Potential treatment strategies evaluated in a mathematical model”. *Cancer Res.* 66.14 (2006), pp. 7293–7300.

[23] G. Caravagna, A. d’Onofrio, P. Milazzo, and R. Barbuti. “Tumour suppression by immune system through stochastic oscillations”. *J. Theor. Biol.* 265.3 (2010), pp. 336–345.

[24] C. Cartis and L. Roberts. “A derivative-free Gauss–Newton method”. *Math. Prog. Comp.* 11.4 (2019), pp. 631–674.

[25] M. Chappell et al. “Mathematical modelling for combinations of immuno-oncology and anti-cancer therapies”. 2015.

[26] D. S. Chen and I. Mellman. “Oncology meets immunology: The cancer-immunity cycle”. *Immunity* 39.1 (2013), pp. 1–10.

[27] O. Chiş, J. R. Banga, and E. Balsa-Canto. “GenSSI: a software toolbox for structural identifiability analysis of biological models”. *Bioinformatics* 27.18 (2011), pp. 2610–2611.

[28] O.-T. Chis, J. R. Banga, and E. Balsa-Canto. “Structural identifiability of systems biology models: A critical comparison of methods”. *PLoS ONE* 6.11 (2011), e27755.

[29] G. D. Clapp et al. “Implication of the autologous immune system in BCR–ABL transcript variations in chronic myelogenous leukemia patients treated with imatinib”. *Cancer Res.* 75.19 (2015), pp. 4053–4062.

[30] M. Clerx et al. “Probabilistic Inference on Noisy Time Series (PINTS)”. *J. Open Res. Softw.* 7.1 (2019), p. 23.

[31] A. C. Daly, D. Gavaghan, J. Cooper, and S. Tavener. “Inference-based assessment of parameter identifiability in nonlinear biological models”. *J. R. Soc. Interface.* 15.144 (2018), p. 20180318.

[32] P. De Maesschalck and M. Wechselberger. “Neural excitability and singular bifurcations”. *J. Math. Neurosci.* 5.1 (2015).

[33] L. G. de Pillis and A. E. Radunskaya. “A mathematical tumor model with immune resistance and drug therapy: An optimal control approach”. *J. Theor. Med.* 3.2 (2001), pp. 79–100.

[34] L. G. de Pillis, A. E. Radunskaya, and C. L. Wiseman. “A validated mathematical model of cell-mediated immune response to tumor growth”. *Cancer Res.* 65.17 (2005), pp. 7950–7958.

[35] H. P. de Vladar and J. A. González. “Dynamic response of cancer under the influence of immunological activity and therapy”. *J. Theor. Biol.* 227.3 (2004), pp. 335–348.

[36] C. DeLisi and A. Rescigno. “Immune surveillance and neoplasia-1 a minimal mathematical model”. *Bull. Math. Biol.* 39.2 (1977), pp. 201–221.

[37] M. Delitala, T. Lorenzi, and M. Melensi. “A structured population model of competition between cancer cells and T cells under immunotherapy”. In: *Math. Model. Tumor-Immune Syst. Dyn.* Ed. by A. Eladdadi, P. Kim, and D. Mallet. Vol. 107. Springer Proceedings in Mathematics & Statistics. New York, NY: Springer New York, 2014, pp. 47–58.

[38] L. Denis-Vidal, G. Joly-Blanchard, and C. Noiret. “Some effective approaches to check the identifiability of uncontrolled nonlinear systems”. *Math. Comput. Simul.* 57.1-2 (2001), pp. 35–44.

[39] A. d’Onofrio. “A general framework for modeling tumor-immune system competition and immunotherapy: Mathematical analysis and biomedical inferences”. *Phys. D Nonlinear Phenom.* 208.3-4 (2005), pp. 220–235.

[40] A. d’Onofrio. “Bounded-noise-induced transitions in a tumor-immune system interplay”. *Phys. Rev. E* 81.2 (2010), p. 021923.

[41] A. d’Onofrio, G. Caravagna, and S. de Franciscis. “Bounded noise induced first-order phase transitions in a baseline non-spatial model of gene transcription”. *Phys. A Stat. Mech. Its Appl.* 492 (2018), pp. 2056–2068.

[42] A. d’Onofrio and A. Ciancio. “Simple biophysical model of tumor evasion from immune system control”. *Phys. Rev. E* 84.3 (2011), p. 031910.

[43] A. d'Onofrio and A. Gandolfi. "Resistance to antitumor chemotherapy due to bounded-noise-induced transitions". *Phys. Rev. E* 82.6 (2010), p. 061901.

[44] S. J. Dovedi et al. "Fractionated radiation therapy stimulates antitumor immunity mediated by both resident and infiltrating polyclonal T-cell populations when combined with PD-1 blockade". *Clin. Cancer Res.* 23.18 (2017), pp. 5514–5526.

[45] H. Dritschel. "Mathematical modelling of heterogeneity in tumour-immune cell interactions". PhD thesis. University of Oxford, 2018.

[46] H. Dritschel, S. L. Waters, A. Roller, and H. M. Byrne. "A mathematical model of cytotoxic and helper T cell interactions in a tumour microenvironment". *Lett. Biomath.* 5.2 (2018), S36–S68.

[47] G. P. Dunn, L. J. Old, and R. D. Schreiber. "The three Es of cancer immunoediting". *Annu. Rev. Immunol.* 22.1 (2004), pp. 329–360.

[48] D. J. Earl and M. W. Deem. "Parallel tempering: Theory, applications, and new perspectives". *Phys. Chem. Chem. Phys.* 7.23 (2005), p. 3910.

[49] R. Eftimie. "Investigation into the role of macrophages heterogeneity on solid tumour aggregations". *Math. Biosci.* 322 (2020), p. 108325.

[50] R. Eftimie, J. L. Bramson, and D. J. D. Earn. "Interactions between the immune system and cancer: A brief review of non-spatial mathematical models". *Bull. Math. Biol.* 73.1 (2011), pp. 2–32.

[51] L. A. Emens et al. "Cancer immunotherapy: Opportunities and challenges in the rapidly evolving clinical landscape". *Eur. J. Cancer* 81 (2017), pp. 116–129.

[52] S. Farkona, E. P. Diamandis, and I. M. Blasutig. "Cancer immunotherapy: the beginning of the end of cancer?" *BMC Med.* 14.1 (2016), p. 73.

[53] W. H. Fridman, F. Pagès, C. Sautès-Fridman, and J. Galon. "The immune contexture in human tumours: impact on clinical outcome". *Nat. Rev. Cancer* 12.4 (2012), pp. 298–306.

[54] A. Friedman and X. Lai. "Combination therapy for cancer with oncolytic virus and checkpoint inhibitor: A mathematical model". *PLoS One* 13.2 (2018), pp. 1–21.

[55] M Gałach. "Dynamics of the tumor-immune system competition—the effect of time delay". *Int. J. Appl. Math. Comput. Sci.* 13.3 (2003), pp. 395–406.

[56] R. A. Gatti et al. "Cyclic leukocytosis in chronic myelogenous leukemia: new perspectives on pathogenesis and therapy." *Blood* 41.6 (1973), pp. 771–82.

[57] A. E. Gelfand, D. K. Dey, and H. Chang. "Model determination using predictive distributions with implementation via sampling-based methods". In: *Bayesian Stat. 4*. Ed. by J. Bernardo, J. Berger, A. Dawid, and A. Smith. Oxford: Oxford University Press, 1992, pp. 147–167.

[58] A. Gelman, J. Hwang, and A. Vehtari. "Understanding predictive information criteria for Bayesian models". *Stat. Comput.* 24.6 (2014), pp. 997–1016.

[59] C. Gong et al. "A computational multiscale agent-based model for simulating spatio-temporal tumour immune response to PD1 and PDL1 inhibition". *J. R. Soc. Interface* 14.134 (2017), p. 20170320.

[60] M. M. Graham and A. J. Storkey. "Continuously tempered Hamiltonian Monte Carlo". In: *Proceedings of the 33rd Conference on Uncertainty in Artificial Intelligence*. 2017. arXiv: 1704.03338. URL: http://auai.org/uai2017/proceedings/papers/289.pdf.

[61] E. Grywalska, M. Pasiarski, S. Góźdź, and J. Roliński. "Immune-checkpoint inhibitors for combating T-cell dysfunction in cancer". *OncoTargets Ther.* 11 (2018), pp. 6505–6524.

[62] R. N. Gutenkunst et al. "Universally sloppy parameter sensitivities in systems biology models". *PLoS Comput. Biol.* 3.10 (2007), pp. 1871–1878.

[63] S. P. Haen, M. W. Löffler, H. G. Rammensee, and P. Brossart. "Towards new horizons: characterization, classification and implications of the tumour antigenic repertoire". *Nat. Rev. Clin. Oncol.* 17.10 (2020), pp. 595–610.

[64] A. E. Hammerstrom, D. H. Cauley, B. J. Atkinson, and P. Sharma. "Cancer immunotherapy: Sipuleucel-T and beyond". *Pharmacotherapy* 31.8 (2011), pp. 813–828.

[65] D. Hanahan and R. A. Weinberg. "Hallmarks of cancer: The next generation". *Cell* 144.5 (2011), pp. 646–674.

[66] H. A. Harrington, K. L. Ho, and N. Meshkat. "A parameter-free model comparison test using differential algebra". *Complexity* 2019 (2019), pp. 1–15.

[67] M. Hashimoto et al. "CD8 T cell exhaustion in chronic infection and cancer: Opportunities for interventions". *Annu. Rev. Med.* 69.1 (2018), pp. 301–318.

[68] J. He, Y. Hu, M. Hu, and B. Li. "Development of PD-1/PD-L1 pathway in tumor immune microenvironment and treatment for non-small cell lung cancer". *Sci. Rep.* 5.1 (2015), p. 13110.

[69] A. Hilfinger and J. Paulsson. "Separating intrinsic from extrinsic fluctuations in dynamic biological systems". *Proc. Natl. Acad. Sci. U. S. A.* 108.29 (2011), pp. 12167–12172.

[70] E. J Hinch. *Perturbation Methods*. Cambridge: Cambridge University Press, 1991, pp. 52–67.

[71] K. E. Hines, T. R. Middendorf, and R. W. Aldrich. "Determination of parameter identifiability in nonlinear biophysical models: A Bayesian approach". *J. Gen. Physiol.* 143.3 (2014), pp. 401–416.

[72] F. Hoffman et al. “A mathematical model of antibody-dependent cellular cytotoxicity (ADCC)”. *J. Theor. Biol.* 436 (2018), pp. 39–50.

[73] M. D. Hoffman and A. Gelman. “The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo”. *J. Mach. Learn. Res.* 15.1 (2014), pp. 1593–1623. arXiv: `1111.4246`.

[74] S. Hug et al. “High-dimensional Bayesian parameter estimation: Case study for a model of JAK2/STAT5 signaling”. *Math. Biosci.* 246.2 (2013), pp. 293–304.

[75] M. Itik and S. P. Banks. “Chaos in a three-dimensional cancer model”. *Int. J. Bifurc. Chaos* 20.01 (2010), pp. 71–79.

[76] A. Jasra, D. A. Stephens, and C. C. Holmes. “On population-based simulation for static inference”. *Stat. Comput.* 17.3 (2007), pp. 263–279.

[77] B. Joshi et al. “On immunotherapies and cancer vaccination protocols: A mathematical modelling approach”. *J. Theor. Biol.* 259.4 (2009), pp. 820–827.

[78] S. Khajanchi and S. Banerjee. “Stability and bifurcation analysis of delay induced tumor immune interaction model”. *Appl. Math. Comput.* 248 (2014), pp. 652–671.

[79] D. Kirschner and J. C. Panetta. “Modeling immunotherapy of the tumor - immune interaction”. *J. Math. Biol.* 37.3 (1998), pp. 235–252.

[80] A. Konstorum, A. T. Vella, A. J. Adler, and R. C. Laubenbacher. “Addressing current challenges in cancer immunotherapy with mathematical and computational modelling”. *J. R. Soc. Interface* 14.131 (2017), p. 20170150.

[81] E. Köse et al. “Immuno-kinetics of immunotherapy: dosing with DCs”. *Lett. Biomath.* 4.1 (2017), pp. 39–58.

[82] Y. Kosinsky et al. “Radiation and PD-(L)1 treatment combinations: immune response and dose optimization via a predictive systems model”. *J. Immunother. Cancer* 6.1 (2018), p. 17.

[83] T. Kostova, R. Ravindran, and M. Schonbek. “Fitzhugh-Nagumo revisited: Types of bifurcations, periodical forcing and stability regions by a Lyapunov functional”. *Int. J. Bifurcation Chaos* 14.03 (2004), pp. 913–925.

[84] N. Kronik, Y. Kogan, V. Vainstein, and Z. Agur. “Improving alloreactive CTL immunotherapy for malignant gliomas using a simulation model of their interactive dynamics”. *Cancer Immunol. Immunother.* 57.3 (2008), pp. 425–439.

[85] N. Kronik et al. “Predicting outcomes of prostate cancer immunotherapy by personalized mathematical models”. *PLoS One* 5.12 (2010), e15482.

[86] R. Kumar, C. Carroll, A. Hartikainen, and O. Martin. “ArviZ a unified library for exploratory analysis of Bayesian models in Python”. *J. Open Source Softw.* 4.33 (2019), p. 1143.

[87] V. A. Kuznetsov and G. D. Knott. "Modeling tumor regrowth and immunotherapy". *Math. Comput. Model.* 33.12-13 (2001), pp. 1275–1287.

[88] V. A. Kuznetsov, I. A. Makalkin, M. A. Taylor, and A. S. Perelson. "Nonlinear dynamics of immunogenic tumors: Parameter estimation and global bifurcation analysis". *Bull. Math. Biol.* 56.2 (1994), pp. 295–321.

[89] X. Lai and A. Friedman. "Combination therapy for melanoma with BRAF/MEK inhibitor and immune checkpoint inhibitor: A mathematical model". *BMC Syst. Biol.* 11.1 (2017), pp. 1–18.

[90] X. Lai and A. Friedman. "Combination therapy of cancer with cancer vaccine and immune checkpoint inhibitors: A mathematical model". *PLoS One* 12.5 (2017), e0178479.

[91] B. Lambert. *A Student's Guide to Bayesian Statistics*. Los Angeles: SAGE Publications Ltd, 2018.

[92] B. Lambert, D. Gavaghan, and S. Tavener. "Inverse sensitivity analysis of mathematical models avoiding the curse of dimensionality". Preprint. 2018. DOI: 10.1101/432393.

[93] M. Lavielle. *Mixed Effects Models for the Population Approach*. New York: Chapman and Hall/CRC, 2014.

[94] R. Lefever and W. Horsthemke. "Bistability in fluctuating environments. Implications in tumor immunology". *Bull. Math. Biol.* 41.4 (1979), pp. 469–490.

[95] C. L. Lei et al. "Considering discrepancy when calibrating a mechanistic electrophysiology model". *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 378.2173 (2020), p. 20190349.

[96] C. Letellier, F. Denis, and L. A. Aguirre. "What can be learned from a chaotic cancer model?" *J. Theor. Biol.* 322 (2013), pp. 7–16.

[97] K.-L. Liao, X.-F. Bai, and A. Friedman. "Mathematical modeling of interleukin-27 induction of anti-tumor T cells response". *PLoS One* 9.3 (2014), e91844.

[98] K.-L. Liao, X.-F. Bai, and A. Friedman. "Mathematical modeling of interleukin-35 promoting tumor growth and angiogenesis". *PLoS One* 9.10 (2014), e110126.

[99] S. W. X. Lim. "Bayesian inverse problems and seismic inversion". PhD thesis. University of Oxford, 2016.

[100] E. J. Lipson et al. "Antagonists of PD-1 and PD-L1 in cancer treatment". *Semin. Oncol.* 42.4 (2015), pp. 587–600.

[101] X. Liu, Q. Li, and J. Pan. "A deterministic and stochastic model for the system dynamics of tumor–immune responses to chemotherapy". *Phys Stat Mech Its Appl* 500 (2018), pp. 162–176.

[102] L. Ljung. *System Identification: Theory for the User*. 2nd edition. Upper Saddle River, NJ 07458: Prentice Hall PTR, 1999.

[103] L. Ljung and T. Glad. “On global identifiability for arbitrary model parametrizations”. *Automatica* 30.2 (1994), pp. 265–276.

[104] D. C. López C., T. Barz, S. Körkel, and G. Wozny. “Nonlinear ill-posed problem analysis in model-based parameter estimation and experimental design”. *Comput. Chem. Eng.* 77 (2015), pp. 24–42.

[105] S. Z. Makaryan, C. G. Cess, and S. D. Finley. “Modeling immune cell behavior across scales in cancer”. *WIREs Syst. Biol. Med.* 12.4 (2020), e1484.

[106] *Maple*. Version 2018.0. Waterloo, Ontario: Maplesoft, a division of Waterloo Maple Inc.

[107] A. Matzavinos, M. A. Chaplain, and V. A. Kuznetsov. “Mathematical modelling of the spatio-temporal response of cytotoxic T-lymphocytes to a solid tumour”. *Math. Med. Biol.* 21.1 (2004), pp. 1–34.

[108] B. Mehta and M. Agarwal. “Cyclic oscillations in leukocyte count in chronic myeloid leukemia”. *Acta Haematol.* 63.2 (1980), pp. 68–70.

[109] I. Melero et al. “Evolving synergistic combinations of targeted immunotherapies to combat cancer”. *Nat. Rev. Cancer* 15.8 (2015), pp. 457–472.

[110] N. Meshkat, C. Er-zhen Kuo, and J. DiStefano. “On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and COMBOS: A novel web implementation”. *PLoS ONE* 9.10 (2014), e110261.

[111] H. Miao, X. Xia, A. S. Perelson, and H. Wu. “On identifiability of nonlinear ODE models and applications in viral dynamics”. *SIAM Rev.* 53.1 (2011), pp. 3–39.

[112] O. Milberg et al. “A QSP model for predicting clinical responses to monotherapy, combination and sequential therapy following CTLA-4, PD-1, and PD-L1 checkpoint blockade”. *Sci. Rep.* 9.1 (2019), pp. 1–17.

[113] H. Mohammad, M. Y. Waziri, and S. A. Santos. “A brief survey of methods for solving nonlinear least-squares problems”. *Numer. Algebr. Control Optim.* 9.1 (2019), pp. 1–13.

[114] J. J. Moré. “The Levenberg-Marquardt algorithm: Implementation and theory”. In: *Numerical Analysis*. Ed. by G. A. Watson. Vol. 630. Berlin, Heidelberg: Springer Berlin Heidelberg, 1978, pp. 105–116.

[115] K. Murphy et al. *Janeway’s Immunobiology*. 9th edition. New York, NY: Garland Science, 2016.

[116] J. D. Murray. *Mathematical Biology*. Second, corrected edition. Berlin ; London: Springer-Verlag, 1993.

[117] V. I. Nekorkin. *Introduction to Nonlinear Oscillations*. Weinheim: Wiley-VCH, 2015.

[118] E. Nikolopoulou et al. “Tumour-immune dynamics with an immune checkpoint inhibitor”. *Lett. Biomath.* 5.2 (2018), S137–S159.

[119] K. A. Norton, C. Gong, S. Jamalian, and A. S. Popel. “Multiscale agent-based and hybrid modeling of the tumor immune microenvironment”. *Processes* 7.1 (2019), pp. 1–23.

[120] A. Osojnik, E. A. Gaffney, M. Davies, J. W. Yates, and H. M. Byrne. “Identifying and characterising the impact of excitability in a mathematical model of tumour-immune interactions”. *J. Theor. Biol.* 501 (2020), p. 110250.

[121] P. A. Ott, F. S. Hodi, and C. Robert. “CTLA-4 and PD-1/PD-L1 blockade: New immunotherapeutic modalities with durable clinical benefit in melanoma patients”. *Clin. Cancer Res.* 19.19 (2013), pp. 5300–5309.

[122] D. M. Pardoll. “The blockade of immune checkpoints in cancer immunotherapy”. *Nat. Rev. Cancer* 12.4 (2012), pp. 252–264.

[123] D. S. Park et al. “The Goldilocks window of personalized chemotherapy: Getting the immune response just right”. *Cancer Res.* 79.20 (2019), pp. 5302–5315.

[124] P. Pathria, T. L. Louis, and J. A. Varner. “Targeting tumor-associated macrophages in cancer”. *Trends Immunol.* 40.4 (2019), pp. 310–327.

[125] Y. Pawitan. *In All Likelihood : Statistical Modelling and Inference Using Likelihood*. Oxford: Clarendon Press, 2001.

[126] K. Perica, J. C. Varela, M. Oelke, and J. Schneck. “Adoptive T cell immunotherapy for cancer”. *Rambam Maimonides Med. J.* 6.1 (2015), e0004.

[127] K. Peskov et al. “Quantitative mechanistic modeling in support of pharmacological therapeutics development in immuno-oncology”. *Front. Immunol.* 10 (2019), p. 924.

[128] H. Pohjanpalo. “System identifiability based on the power series expansion of the solution”. *Math. Biosci.* 41.1-2 (1978), pp. 21–33.

[129] S. A. Quezada, K. S. Peggs, T. R. Simpson, and J. P. Allison. “Shifting the equilibrium in cancer immunoediting: from tumor tolerance to eradication”. *Immunol. Rev.* 241.1 (2011), pp. 104–118.

[130] A. Radunskaya, R. Kim, and T. Woods II. “Mathematical modeling of tumor immune interactions: a closer look at the role of a PD-L1 inhibitor in cancer immunotherapy”. *SPORA J. Biomath.* 4.1 (2018), pp. 25–41.

[131] M. Rajasagi et al. “Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia”. *Blood* 124.3 (2014), pp. 453–462.

[132] A. Raue, V. Becker, U. Klingmüller, and J. Timmer. “Identifiability and observability analysis for experimental design in nonlinear dynamical models”. *Chaos* 20.4 (2010).

[133] A. Raue, J. Karlsson, M. P. Saccomani, M. Jirstrand, and J. Timmer. "Comparison of approaches for parameter identifiability analysis of biological systems". *Bioinformatics* 30.10 (2014), pp. 1440–1448.

[134] A. Raue et al. "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood". *Bioinformatics* 25.15 (2009), pp. 1923–1929.

[135] A. Raue, C. Kreutz, F. J. Theis, and J. Timmer. "Joining forces of Bayesian and frequentist methodology: A study for inference in the presence of non-identifiability". *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 371.1984 (2013).

[136] A. Raue et al. "Lessons learned from quantitative dynamical modeling in systems biology". *PLoS One* 8.9 (2013).

[137] F. A. Rihan, D. H. Abdel Rahman, S. Lakshmanan, and A. S. Alkhajeh. "A time delay model of tumour-immune system interactions: Global dynamics, parameter estimation, sensitivity analysis". *Appl. Math. Comput.* 232 (2014), pp. 606–623.

[138] M. Robertson-Tessi, A. El-Kareh, and A. Goriely. "A mathematical model of tumor–immune interactions". *J. Theor. Biol.* 294.10 (2012), pp. 56–73.

[139] M. Robertson-Tessi, A. El-Kareh, and A. Goriely. "A model for effects of adaptive immunity on tumor response to chemotherapy and chemoimmunotherapy". *J. Theor. Biol.* 380 (2015), pp. 569–584.

[140] R. C. Rockne et al. "The 2019 mathematical oncology roadmap". *Phys. Biol.* 16.4 (2019).

[141] A. R. Rodriguez and C. L. Lutcher. "Marked cyclic leukocytosis-leukopenia in chronic myelogenous leukemia". *Am. J. Med.* 60.7 (1976), pp. 1041–1047.

[142] E. Roesch and M. P. H. Stumpf. "Parameter inference in dynamical systems with co-dimension 1 bifurcations". *R. Soc. Open Sci.* 6.10 (2019), p. 190747.

[143] K. Roesch, D. Hasenclever, and M. Scholz. "Modelling lymphoma therapy and outcome". *Bull. Math. Biol.* 76.2 (2014), pp. 401–430.

[144] J. J. Rojas, P. Sampath, W. Hou, and S. H. Thorne. "Defining effective combinations of immune checkpoint blockade and oncolytic virotherapy". *Clin. Cancer Res.* 21.24 (2015), pp. 5543–5551.

[145] R. Rosenthal et al. "Neoantigen-directed immune escape in lung cancer evolution". *Nature* 567.7749 (2019), pp. 479–485.

[146] M. P. Saccomani. "Structural vs practical identifiability in system biology". In: *IWBBIO 2013 Proceedings*. Ed. by I. Rojas and F. M. Ortuño Guzman. Granada, Spain: Copicentro Editorial, 2013, pp. 305–313. URL: `http://iwbbio.ugr.es/papers/iwbbio_055.pdf`.

[147] M. P. Saccomani and K. Thomaseth. "Calculating all multiple parameter solutions of ODE models to avoid biological misinterpretations". *Math. Biosci. Eng.* 16.6 (2019), pp. 6438–6453.

[148] M. P. Saccomani and K. Thomaseth. "The union between structural and practical identifiability makes strength in reducing oncological model complexity: A case study". *Complexity* 2018 (2018), pp. 1–10.

[149] I. Sagiv-Barfi et al. "Therapeutic antitumor immunity by checkpoint blockade is enhanced by ibrutinib, an inhibitor of both BTK and ITK". *Proc. Natl. Acad. Sci. USA* 112.9 (2015), E966–E972.

[150] H. Schättler, U. Ledzewicz, and B. Amini. "Dynamical properties of a minimally parameterized mathematical model for metronomic chemotherapy". *J. Math. Biol.* 72.5 (2016), pp. 1255–1280.

[151] R. D. Schreiber, L. J. Old, and M. J. Smyth. "Cancer immunoediting: Integrating immunity's roles in cancer suppression and promotion". *Science* 331.6024 (2011), pp. 1565–1570.

[152] S. Seabold and J. Perktold. "statsmodels: Econometric and statistical modeling with Python". In: *Proceedings of the 9th Python in Science Conference*. 2010. URL: `http://conference.scipy.org/proceedings/scipy2010/pdfs/seabold.pdf`.

[153] R. Serre et al. "Mathematical modeling of cancer immunotherapy and its synergy with radiotherapy". *Cancer Res.* 76.17 (2016), pp. 4931–4940.

[154] P. Sharma, S. Hu-Lieskovan, J. A. Wargo, and A. Ribas. "Primary, adaptive, and acquired resistance to cancer immunotherapy". *Cell* 168.4 (2017), pp. 707–723.

[155] M. J. Simpson, R. E. Baker, S. T. Vittadello, and O. J. Maclaren. "Practical parameter identifiability for spatio-temporal models of cell invasion". *J. R. Soc. Interface* 17.164 (2020), p. 20200055.

[156] H Siu, E. S. Vitetta, R. D. May, and J. W. Uhr. "Tumor dormancy. I. Regression of BCL1 tumor and induction of a dormant tumor state in mice chimeric at the major histocompatibility complex." *J. Immunol.* 137.4 (1986), pp. 1376–82.

[157] L. M. Sompayrac. *How the Immune System Works*. 4th edition. John Wiley & Sons, Ltd., 2011.

[158] O. Sotolongo-Costa, L. Morales Molina, D. Rodríguez Perez, J. Antoranz, and M. Chacón Reyes. "Behavior of tumors under nonstationary therapy". *Phys. D Nonlinear Phenom.* 178.3-4 (2003), pp. 242–253.

[159] Stan Development Team. *Brief guide to Stan's warnings*. 2020. URL: `https://mc-stan.org/misc/warnings.html` (visited on 03/28/2020).

[160] Stan Development Team. *CmdStan: the command-line interface to Stan*. Version 2.21.0. 2018. URL: `http://mc-stan.org`.

[161] P. Stapor et al. “PESTO: Parameter EStimation TOolbox”. *Bioinformatics* 34.4 (2018), pp. 705–707.

[162] H. J. Steer, R. A. Lake, A. K. Nowak, and B. W. Robinson. “Harnessing the immune response to treat cancer”. *Oncogene* 29.48 (2010), pp. 6301–6313.

[163] N. V. Stepanova. “Course of the immune reaction during the development of a malignant tumor”. *Biophys.* 24 (1980), pp. 917–923.

[164] A. M. Stuart. “Inverse problems: A Bayesian perspective”. *Acta Numerica* 19 (2010), pp. 451–559.

[165] J. B. Swann and M. J. Smyth. “Immune surveillance of tumors”. *J. Clin. Invest.* 117.5 (2007), pp. 1137–46.

[166] D. S. Thommen et al. “Progression of lung cancer is associated with increased dysfunction of T cells defined by coexpression of multiple inhibitory receptors”. *Cancer Immunol. Res.* 3.12 (2015), pp. 1344–1354.

[167] T. Toni, D. Welch, N. Strelkowa, A. Ipsen, and M. P. H. Stumpf. “Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems”. *J. R. Soc. Interface.* 6.31 (2009), pp. 187–202.

[168] J. Vanlier, C. A. Tiemann, P. A. Hilbers, and N. A. van Riel. “An integrated strategy for prediction uncertainty analysis”. *Bioinformatics* 28.8 (2012), pp. 1130–1135.

[169] A. Vehtari, A. Gelman, and J. Gabry. “Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC”. *Stat. Comput.* 27.5 (2017), pp. 1413–1432.

[170] A. Vehtari, A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner. “Rank-Normalization, Folding, and Localization: An Improved $\widehat{R}$ for Assessing Convergence of MCMC (with Discussion)”. *Bayesian Anal.* 16.2 (2021).

[171] A. F. Villaverde, A. Barreiro, and A. Papachristodoulou. “Structural identifiability of dynamic systems biology models”. *PLoS Comput. Biol.* 12.10 (2016), e1005153.

[172] Y. Vodovotz et al. “Solving immunology?” *Trends Immunol.* 38.2 (2017), pp. 116–127.

[173] T. A. Waldmann. “Cytokines in cancer immunotherapy”. *Cold Spring Harb. Perspect. Biol.* 3.4 (2017), a028472.

[174] R. Walker and H. Enderling. “From concept to clinic: Mathematically informed immunotherapy”. *Curr. Probl. Cancer* 40.1 (2016), pp. 68–83.

[175] E. Walter and L. Pronzato. *Identification of Parametric Models from Experimental Data*. London: Springer-Verlag, 1997.

[176] R. A. Weinberg. *The Biology of Cancer*. Second edition. New York: W.W. Norton & Company, 2014.

[177] D. G. Whittaker, M. Clerx, C. L. Lei, D. J. Christini, and G. R. Mirams. "Calibration of ionic and cellular cardiac electrophysiology models". *WIREs Syst. Biol. Med.* 12.4 (2020), e1482.

[178] K. P. Wilkie. "A review of mathematical models of cancer–immune interactions in the context of tumor dormancy". In: *Systems Biology of Tumor Dormancy*. Ed. by H. Enderling, N. Almog, and L. Hlatky. Vol. 734. New York, NY: Springer New York, 2013, pp. 201–234.

[179] K. P. Wilkie and P. Hahnfeldt. "Modeling the dichotomy of the immune response to cancer: Cytotoxic effects and tumor-promoting inflammation". *Bull. Math. Biol.* 79.6 (2017), pp. 1426–1448.

[180] K. P. Wilkie and P. Hahnfeldt. "Tumor–immune dynamics regulated in the microenvironment inform the transient nature of immune-induced tumor dormancy". *Cancer Res.* 73.12 (2013), pp. 3534–3544.

[181] Wolfram Research. *Mathematica*. Version 11.3. Champaign, IL, 2018.

[182] World Health Organization. *Cancer*. URL: `https://www.who.int/newsroom/fact-sheets/detail/cancer` (visited on 05/31/2021).

[183] Y. Xu, J. Feng, J. Li, and H. Zhang. "Stochastic bifurcation for a tumor–immune system with symmetric Lévy noise". *Phys. A Stat. Mech. Its Appl.* 392.20 (2013), pp. 4739–4748.

[184] L. Yang, A. Li, Q. Lei, and Y. Zhang. "Tumor-intrinsic signaling pathways: Key roles in the regulation of the immunosuppressive tumor microenvironment". *J. Hematol. Oncol.* 12.1 (2019), pp. 1–14.

[185] Y. Yang, C. Li, T. Liu, X. Dai, and A. V. Bazhin. "Myeloid-derived suppressor cells in tumors: From mechanisms to antigen specificity and microenvironmental regulation". *Front. Immunol.* 11 (2020), p. 1371.

[186] T. E. Yankeelov et al. "Multi-scale modeling in clinical oncology: Opportunities and barriers to success". *Ann. Biomed. Eng.* 44.9 (2016), pp. 2626–2641.

[187] Y. Yao, C. Cademartori, A. Vehtari, and A. Gelman. *Adaptive path sampling in metastable posterior distributions*. Preprint. 2020. arXiv: `2009.00471`.