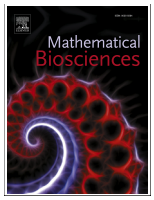
Mathematical Biosciences 399 (2026) 109710

Contents lists available at ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs



Original research article

Optimal experiment design for practical parameter identifiability and model discrimination ⋆

Yue Liu [a,b,*], Philip K. Maini [a], Ruth E. Baker [a]

[a] Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Road, Oxford, OX2 6GG, UK
[b] Department of Mathematics, Purdue University, 150 N. University St, West Lafayette, 47906, Indiana, USA

ARTICLE INFO

2000 MSC:
92B05
49N90
62F99

Keywords:
Experimental design
Parameter identifiability
Profile likelihood
Optimal control

ABSTRACT

Mechanistic mathematical models of biological systems usually contain a number of unknown parameters whose values need to be estimated from available experimental data in order for the models to be validated and used to make quantitative predictions. This requires that the models are practically identifiable, that is, the values of the parameters can be confidently determined, given available data. A well-designed experiment can produce data that are much more informative for the purpose of inferring parameter values than a poorly designed experiment. It is, therefore, of great interest to optimally design experiments such that the resulting data maximise the practical identifiability of a chosen model. Experimental design is also useful for model discrimination, where we seek to distinguish between multiple distinct, competing models of the same biological system in order to determine which model better reveals insight into the underlying biological mechanisms. In many cases, an external stimulus can be used as a control input to probe the behaviour of the system. In this paper, we will explore techniques for optimally designing such a control for a given experiment, in order to maximise parameter identifiability and model discrimination, and demonstrate these techniques in the context of commonly applied ordinary differential equation models. We use a profile likelihood-based approach to assess parameter identifiability. We then show how the problem of optimal experimental design for model discrimination can be formulated as an optimal control problem, which can be solved efficiently by applying Pontryagin's Maximum Principle.

1. Introduction

Mechanistic mathematical models that describe the dynamics of biological systems usually contain a number of parameters, whose values are often difficult to measure directly, and are instead determined by fitting the model to available experimental data. Certain experimental data can be more informative than others for the purpose of inferring model parameters. This naturally raises the question of how to optimally design an experiment such that the data it produces are the most informative. This idea dates back over a century, as discussed in [1], with [2] serving as an early example. In mathematical biology, the ideas of experimental design have been applied to models of signalling pathways [3,4], gene regulatory networks [5], yeast fermentation [6], and PKPD models to study the effects of drugs [7], and also in physics to models of crystallisation [8] and polymerisation [9]. These studies illustrate that a few well-designed experiments have the potential to provide as much, or even more, insight into the system being modelled as many naively-designed ones. Therefore, optimal experimental design is of great interest, as it can reduce costs by requiring fewer experiments, and improves the effectiveness of experiments in biological research.

In this paper, we discuss optimal experimental design in the context of ordinary differential equation (ODE) models. Two important tasks in a modelling study are parameter estimation and model selection. These two objectives are inherently intertwined, and can both benefit from careful experimental design. The ability to accurately estimate parameter values is referred to as practical parameter identifiability, and can be understood as the inverse of the uncertainty in parameter estimates. Therefore, designing an experiment to optimise parameter identifiability aids parameter estimation. Alternatively, we can optimise the experiment for model discrimination, aiming to maximise the discrepancy between the predictions made by competing models: this discrepancy in predictions makes it easier to distinguish between the models given data, and to select the model that most accurately represents the biological system. Although these objectives share a connection, they are nonethe-

⋆ Given her role as Deputy Editor, Ruth Baker had no involvement in the peer-review of this article and has no access to information regarding its peer-review. Full responsibility for the editorial process for this article was delegated to another journal editor.

* Corresponding author.
E-mail address: liu4194@purdue.edu (Y. Liu).

https://doi.org/10.1016/j.mbs.2026.109710
Received 12 June 2025; Received in revised form 2 March 2026; Accepted 27 April 2026
Available online 22 May 2026
0025-5564/© 2026 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

less distinct. Previous authors recognised this connection, as reflected in numerous papers that discuss both simultaneously [1,6]. However, it is worthwhile to elaborate on the similarities and distinctions between them.

1.1. Parameter identifiability

Practical parameter identifiability refers to the ability to confidently identify the values of model parameters given available data, which are often noisy, and limited in resolution [10]. Practical identifiability should be distinguished from structural identifiability, which refers to the ability to uniquely determine parameter values in an idealised situation where the data are free from observational noise, and of arbitrarily high precision [11,12]. We focus on practical, rather than structural, identifiability in this paper, since practical identifiability accounts for the noisy and limited nature of realistic experimental datasets. A well-designed experiment will provide sufficient data to reduce the uncertainty in the parameter estimates, thereby improving practical identifiability.

There are several approaches for measuring practical identifiability. Many earlier works on experiment design use criteria based on the Fisher Information Matrix (FIM), which depends only on the local likelihood landscape at the point of maximum likelihood [1,13]. Bayesian approaches also give rise to measures of uncertainty based on the posterior distribution [14]. In this paper, we will use a profile likelihood-based approach [5,15], where the widths of the 95% confidence intervals for parameter values are used as a measure of practical identifiability. This approach incorporates global information about the likelihood landscape, and therefore provides a more accurate view of identifiability than the FIM, while being computationally cheaper than the Bayesian approach [16,17].

1.2. Model discrimination

In contrast to the problem of improving parameter identifiability, model discrimination involves two or more models, which may incorporate alternative mechanisms, or different parameter values for the same underlying dynamics. In order to select the most appropriate model, we aim to design an experiment in such a way that the predictions of the models differ as much as possible.

The notions of model discrimination and parameter identifiability are distinct, but related. Consider a scenario where the practical identifiability of a model with undetermined parameters is poor given existing data, and model fitting yields multiple parameter sets that reproduce the data well and represent distinct hypotheses about the underlying biological mechanisms. We can recast the problem of overcoming non-identifiability as a model discrimination problem, by selecting two or more parameter sets obtained this way, and designing a new experiment to maximise the difference in model output under the selected parameter sets. This allows us to iteratively improve our understanding of the biological system. Such an iterative approach is common in biological research [1,13,18]. The approach described in Section 1.1 would be more suitable for the first iteration of model parameterisation, while the model discrimination approach for experimental design would be more suitable for later iterations, since it assumes more knowledge of the parameter values.

1.3. Design of experimental inputs

There are many aspects of an experiment that can be varied. For a dynamical model in mathematical biology, these aspects include the initial conditions of the experiment, the time span, the measurement or observation times and their frequency, the observation procedure, which variables to measure, any external stimuli or inputs to the experiment, and more. Each of these has received some attention in the literature [1,10,19,20].

In this work, we choose to focus on optimally designing an external stimulus, or control input, to the experiment. In existing works that discuss the design of controls for experiments, the control input is often restricted to take simple forms, such as step functions or low degree polynomials [4,13,21,22]. These restrictions simplify analysis, and can be sufficient in experiments where the ability to precisely manipulate the stimulation is limited. However, in settings where advances in experimental technology permit more precise manipulation of experimental inputs, it is natural to relax these limitations. In this work, we therefore consider both continuous controls, where the control can be arbitrary piecewise continuous functions of time, and “bang-bang" controls [23], where the control is restricted to two discrete values, and compare their performance.

In order to constrain the design space, all other aspects of the experiment will be considered fixed, and cannot be changed in the design. Notably, we assume that the observations are made at evenly spaced times $t_i = i\Delta t$, $i = 0, \dots, n_t$, with $T = n_t \Delta t$ being the time span of the experiment. We consider the experiment parameters T, n_t and Δt to be fixed. Experiment design approaches based on varying these parameters are discussed in [24].

1.4. Overview

This paper seeks to achieve two related but distinct objectives which are, (i) improving practical identifiability of a single model that contains undetermined parameters, and (ii) discriminating between a pair of models with known parameters, by optimally designing a control input. For each of these objectives, we state the mathematical formulation of the problem, then demonstrate the experimental design procedure using simple ODE models.

In Section 2, we review the profile likelihood-based approach for identifiability analysis and quantification of parameter uncertainty, and the construction of confidence intervals for parameter values using profile likelihoods. We then demonstrate how to optimally design a control input to an experiment to improve the practical identifiability of a model, using the logistic and Richards growth laws for population growth as examples. The main contribution of this section is to demonstrate how to use a profile likelihood-based metric for practical identifiability in the context of optimal experimental design, and how to deal with the associated computational challenges.

In Section 3, we formulate the model discrimination problem as a open-loop optimal control problem. The control problem, which has an unusual form, is then solved by applying Pontryagin’s Maximum Principle (PMP). We use the same population growth models to illustrate the process of optimal experimental design. This section introduces a new way to formulate the model discrimination problem, and implements an efficient algorithm to solve the optimisation problem. The main novelty of this section, compared to earlier works, is to allow the control function to take a general form, providing greater flexibility in experiment design. Finally, the results are summarised, and their significance is addressed, in Section 4.

2. Optimal experimental design for parameter identifiability

In this section, we consider using a control to improve the identifiability of a model. We quantify the practical identifiability of a parameter in a model as the width of the 95% confidence interval for that parameter, which is found via its univariate profile likelihood. The method of profile likelihoods is one of the main techniques for practical identifiability analysis, and has been used in many studies [10,16,17,25–27]. Here we introduce the approach briefly.

Consider a general ODE model given by Eq. (1), and an observation model, Eq. (2),

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, t, \boldsymbol{u}; \boldsymbol{\theta}), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{1}$$

$$\boldsymbol{y}_i(\boldsymbol{\theta}, \boldsymbol{u}) = g(\boldsymbol{x}(t_i; \boldsymbol{\theta}, \boldsymbol{u}); \boldsymbol{\theta}) + \boldsymbol{\xi}_i, \quad i = 0, \dots, n_t, \quad \boldsymbol{\xi}_i \sim \mathcal{N}(0, \sigma^2), \tag{2}$$

where $\boldsymbol{x}(t)$ denotes the state variables, $\boldsymbol{u}(t)$ the control variables, $\boldsymbol{\theta}$ the parameters, $\boldsymbol{y}$ the observables, which are measured at time-points $t_i, i = 0, \dots, n_t$, and $\boldsymbol{\xi}$ the observational noise. Let $\boldsymbol{y}_{\text{model},i}(\boldsymbol{\theta}, \boldsymbol{u})$ denote the output of the model, and $\boldsymbol{y}_{\text{data},i}$ denote the experimental data. We use the notation $\boldsymbol{\theta}_{-i}$ to denote the parameter vector with θ_i, the i^{th} parameter, removed. Let $L(\boldsymbol{y}_{\text{data}}|\boldsymbol{\theta}, \sigma)$ denote the likelihood function, and $\boldsymbol{\theta}^*, \sigma^*$ be the maximum likelihood estimates (MLEs) generated using available data. The normalised univariate profile likelihood function $l_i(\theta_i')$ for parameter θ_i is defined to be

$$l_i(\theta_i') = \max_{\boldsymbol{\theta}_{-i}, \sigma} \left[\log L(\boldsymbol{y}_{\text{data}}|\boldsymbol{\theta}, \sigma)|_{\theta_i = \theta_i'}\right] - \log L(\boldsymbol{y}_{\text{data}}|\boldsymbol{\theta}^*, \sigma^*), \tag{3}$$

which will henceforth be referred to as the profile likelihood function for brevity. The profile likelihood function can be used to construct an approximate 95% confidence region (which is usually an interval, but can possibly be a union of disjoint intervals if the profile likelihood is multi-modal) for the value of θ_i [26,28],

$$\{\theta_i | l_i(\theta_i) > -\chi^2(0.95; 1)/2 \approx -1.92\}. \tag{4}$$

We use $\Delta\theta_i$ to denote the size of the confidence interval, which depends on the control $\boldsymbol{u}$. For a given model parameter θ_i that we would like to estimate, the goal of experimental design is to minimise $\Delta\theta_i$ by optimally choosing a control function, $\boldsymbol{u}(t)$. This problem is rather difficult to formulate as a traditional optimal control problem, as the objective, $\Delta\theta_i$, cannot be easily expressed as a function of the state and control variables. To make the problem easier, we will restrict our attention to the case of a scalar control taking the form of a window function,

$$u(t) = \begin{cases} u_{\max}, & \tau_0 < t < \tau_0 + \tau, \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

where $u_{\max} \geq 0$ is the height of the window function, $0 \leq \tau_0 \leq T$ is the time when the control is turned on, and $\tau \leq T - \tau_0$ is the duration for which the control is turned on. We have effectively parameterised the control with three control parameters, $u_{\max}$, τ_0, and τ. We use MATLAB's *fmincon* optimisation routine to perform the optimisation in the calculations for the profile likelihoods (Eq. (3)), and use the *fzero* root-finding routine to find the lower and upper bounds of the confidence intervals, as defined in Eq. (4).

2.1. Logistic growth model

We now illustrate the process for designing an experiment to optimise parameter identifiability with the logistic growth model. Here, we use logistic growth to describe the dynamics of a population of cells in a cell proliferation assay. The model equations are

$$\frac{\mathrm{d}C}{\mathrm{d}t} = rC\left(1 - \frac{C}{K}\right) - \delta C := f_L(C; (r, \delta, K)), \quad C(0) = C_0, \tag{6}$$

$$y_i = C(t_i) + \xi_i, \quad \xi_i \sim N(0, \sigma^2), \quad i = 1, \dots, n_t, \tag{7}$$

where $C(t)$ is the number density of cells at time t, y_i is the noisy measurement of cell density at time t_i, $r > 0$ the proliferation rate, $K > 0$ the carrying capacity, and $\delta \geq 0$ the rate of cell death. The model parameters are $\boldsymbol{\theta} = (r, \delta, K)$. The units for the quantities in the logistic model and the control, used throughout the rest of the paper, are

$$[C] = [K] = [\sigma] = \text{cell/mm}^2, \quad [t] = \text{h}, \quad [r] = [\delta] = \text{h}^{-1}, [\tau] = [\tau_0] = \text{h}. \tag{8}$$

The logistic model, as usually given in the literature, does not explicitly include the death term, since it can be absorbed into the linear proliferation term. If we define the effective growth rate and carrying capacity as

$$r_{\text{eff}} = r - \delta, \quad K_{\text{eff}} = K\left(1 - \frac{\delta}{r}\right) = K\left(\frac{r_{\text{eff}}}{r}\right), \tag{9}$$

then

$$f_L(C; (r, \delta, K)) = f_L(C; (r_{\text{eff}}, 0, K_{\text{eff}})) \text{ for all } C \geq 0.$$

There are biologically relevant reasons why we might introduce an explicit death term. For example, it is plausible that the proliferation rate is density-dependent while the death rate is density-independent, or we might be interested in estimating both parameters separately, such as in [29], where the authors used a linear decay term to represent the necrosis of tumour cells, separate from the logistic growth term. However, the addition of the death term means that the parameters r, δ and K in both models are structurally non-identifiable, since a change in the value of one of the three can be compensated by a corresponding change in the value of the other two, by holding r_{eff} and K_{eff} constant. It can be shown that r_{eff} and K_{eff} are, however, structurally identifiable parameter combinations (see Supplementary Materials S.3).

Now, suppose we have the following parameter set as "ground truth", representing a best educated "guess" of the parameter values before the experiment:

$$\hat{r} = 0.45 \text{ h}^{-1}, \quad \hat{\delta} = 0.15 \text{ h}^{-1},$$
$$\hat{K} = 3900 \text{ cell/mm}^2, \quad \sigma = 20 \text{ cell/mm}^2. \tag{10}$$

Such a "guess" can be informed by existing knowledge of biological systems similar to the one being studied. These values are chosen so that the effective parameter values, r_{eff} and K_{eff} as defined in Eq. (9), approximate the parameter values inferred using an experimental dataset from a proliferation assay with MDCK cells in [16]. The other experimental parameters are set to be

$$\Delta t = 0.25 \text{ h}, \quad n_t = 100, \quad T = 25 \text{ h}, \quad C_0 = 100 \text{ cell/mm}^2,$$

which are realistic for such an experiment. We now consider three possible ways to introduce a control into the experiment.

2.2. Using u_K as the control variable

The first way to introduce a control is to assume that we can additively modulate the carrying capacity, K. This can be done by, for example, changing the amount of nutrients provided to the cell culture. The model equation (Eq. (6)) can be modified to

$$\frac{\mathrm{d}C}{\mathrm{d}t} = rC\left[1 - \frac{C}{K - u_K(t)}\right] - \delta C, \tag{11}$$

where $0 \leq u_K(t) < K$ has units cell/mm^2, and takes the form of a window function, as in Eq. (5). For this demonstration, we restrict our attention to the case $u_K \geq 0$, where the intervention reduces carrying capacity, although the $u_K < 0$ case is also biologically sensible. Applying this control allows us to effectively observe the behaviour of the system in two different scenarios: when $u_K = 0$, and when $u_K = u_{\max}$. If another parameter set $\boldsymbol{\theta} = (r, \delta, K)$ yields the same model output as $\hat{\boldsymbol{\theta}} = (\hat{r}, \hat{\delta}, \hat{K})$ in the absence of observational error, then the parameters must satisfy the following three equations:

$$r_{\text{eff}} = r - \delta = \hat{r} - \hat{\delta}, \tag{12a}$$
$$K_{\text{eff}} = K(1 - \delta/r) = \hat{K}(1 - \hat{\delta}/\hat{r}), \tag{12b}$$
$$(K - u_{\max})(1 - \delta/r) = (\hat{K} - u_{\max})(1 - \hat{\delta}/\hat{r}). \tag{12c}$$

It can be shown that this set of equations imply $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, so in the presence of the control, the model is at least structurally identifiable.

In Fig. 1, we demonstrate how the profile likelihood curves change as the control parameters ($u_{\max}$, τ_0, and τ) are varied. First, in Fig. 1(a), we show the profile likelihoods with no control applied. Since the model is known to be structurally non-identifiable, so that a change in one parameter can be perfectly compensated for by changes in other parameters, we obtain a flat top for all three profile likelihood functions, as expected. Nonetheless, we can identify lower bounds for r and K. This is because we restrict all parameters to be non-negative, based on the biological interpretations of the model. A value of r below r_{eff}, or a value of K below K_{eff}, would require a negative δ to compensate, which is not allowed, hence the profile likelihood functions for r and K drop sharply at these values.

Fig. 1(b) shows the results of applying a very weak control, of magnitude $u_{\max} = 50$ cells/mm^2 (compared to $\hat{K} = 3900$ cells/mm^2), in a modestly sized time window. This results in only a marginal improvement in parameter identifiability, as the flat top in the profile likelihoods in Fig. 1(a) changes to a curved one, yet the parameters remain practically non-identifiable (recall that we consider a parameter to be practically identifiable if and only if the confidence interval is finite, and the profile likelihood has a unique global maximum). Increasing the magnitude of the control to $u_{\max} = 200$ cells/mm^2 while maintaining the same values for τ_0 and τ (Fig. 1(c)) renders all parameters identifiable, and increasing it to $u_{\max} = 400$ cells/mm^2 (Fig. 1(d)) further improves the accuracy of the parameter estimates. In Fig. 1(e), the control is applied at the beginning of the experiment, rather than in the middle. Since K has relatively little impact on the dynamics in the early phase of the experiment, placing the window during which u_K is active early on is less effective for improving parameter identifiability than at a later time, and the parameters remain non-identifiable. Finally, Fig. 1(f) shows that non-identifiability will also persist if the control is applied too briefly, since there are too little data within the control window to observe the dynamics of the system. On the other hand, applying the control for too long (Fig. 1(g)) also results in non-identifiability, but for the opposite reason. In this case, we have too little data outside the window.

Furthermore, we observe that the shapes of the profile likelihood curves of the three parameters tend to be similar, and the widths of the corresponding confidence intervals are highly correlated, as can be seen in the bivariate likelihood plots. In each of those plots, the confidence region, represented by the yellow pixels, forms a very narrow ridge, indicating that the combination of any two parameters is identifiable, even if individually they are not. This is a consequence of the fact that r_{eff} and K_{eff} are both structurally identifiable parameter combinations. For this reason, it is sufficient to look at the identifiability of r to understand the overall identifiability of the whole model. Therefore, the rest of this section will focus on r.

Next, we explore the effects of the control on identifiability more systematically by plotting Δr, the width of the confidence interval of r, as a function of the control variables. First, in Fig. 2(a–c), we take $u_{\max} = 200$ cells/mm^2, $\tau_0 = 10$ h, $\tau = 10$ h which, as shown in Fig. 1, makes the model parameters identifiable. Then, we vary one control parameter at a time while leaving the other two fixed, and plot Δr as a function of the varying parameter. We observe that Δr decreases monotonically as $u_{\max}$ increases, confirming our earlier observations, while the effects of τ_0 and τ on Δr are more nuanced. There seem to be two good choices for τ_0: one around $\tau_0 = 9$ h and another around $\tau_0 = 15$ h, which are the local minima of Δr as a function of τ_0. We will later provide an intuitive explanation for this using local sensitivity.

In order to further look into the effects of the location of the control window on parameter identifiability, we fix $u_{\max}$ and plot Δr as both τ_0 and τ vary in Fig. 2(d), which contains Fig. 2(b,c) as cross-sections. We can ignore the part of the plot above the red diagonal dashed line as it represents the region where $\tau_0 + \tau > T$, in which case the control stays on past the end of the experiment, which is no different than the case of $\tau_0 + \tau = T$. We observe that there are two diagonal valleys (darkest blue) representing the combinations of (τ_0, τ) that give the highest degree of identifiability. One of these two valleys represents $\tau_0 < 20$ h and $\tau_0 + \tau \approx 19$ h, and the other represents $\tau_0 > 13$ h and $\tau_0 + \tau \approx 25$ h.

Altogether, the results show that in order to achieve identifiability, the magnitude of the control must be high enough, and the timing and duration of the control window $[\tau_0, \tau_0 + \tau]$ must be chosen carefully, so that the span of time where the system behaviour depends strongly on the parameter being acted upon by the control (K in this case), overlaps significantly with both the time span inside the window when the control is turned on, which is $[\tau_0, \tau_0 + \tau]$, and the time span outside of it.

We can understand this intuitively by thinking about the source of the non-identifiability. For any set of parameter values, we need sufficient data on the violation of at least one condition in Eq. (12) in order to confidently conclude that it does not reflect the data. If $u_{\max}$ is not sufficiently high, then any set of parameters that satisfy Eq. (12b) would also approximately satisfy Eq. (12c), leaving us with effectively only two conditions. If the control window is too short, then we will not have enough data on the violation of Eq. (12c). Similarly, if the control window is too large, leaving us little data on system behaviour outside of the window, then we would not be able to conclude a violation of Eq. (12b). In all of these cases, we have two conditions and three unknown parameters, preventing unique identification of parameter values.

This idea provides a heuristic approach to approximate the optimal controls in terms of parameter sensitivity. Let $\phi_{\theta_j}(t)$ be the local sensitivity of the solution with respect to parameter θ_j, defined as

$$\phi_{\theta_j}(t) = \frac{\partial C_{\text{model}}(t; \boldsymbol{\theta}, C_0)}{\partial \theta_j}, \tag{13}$$

where C_{model} is the output of the model, Eq. (6). The sensitivities are calculated using the following ODE, which can be derived from Eq. (6) using the chain rule:

$$\frac{\mathrm{d}\phi_{\theta_j}}{\mathrm{d}t} = \frac{\partial f(C;\boldsymbol{\theta})}{\partial C}\phi_{\theta_j} + \frac{\partial f(C;\boldsymbol{\theta})}{\partial \theta_j}, \quad \phi_{\theta_j}(0) = 0.$$

We plot the sensitivities with respect to r, δ and K in Fig. 3. A larger value of $|\phi_\theta(t)|$ means the solution is more sensitive to a change in the parameter θ at time t, while holding the other parameter values constant. We have $\phi_\theta(0) = 0$, since the initial condition is fixed. For the sake of this heuristic, let us define the sensitive interval for parameter θ as

$$\mathcal{T}_{\theta_j} = \left\{ t \in [0,T] : |\phi_{\theta_j}(t)| \geq \frac{1}{2} \max_{t' \in [0,T]} |\phi_{\theta_j}(t')| \right\}. \tag{14}$$

This interval represents the time span during which the solution is relatively more sensitive to the parameter. For K, we have $\mathcal{T}_K \approx [13\,\text{h}, 25\,\text{h}]$. As a heuristic for optimal control, we should ensure that the control window $[\tau_0, \tau_0 + \tau]$ covers roughly half of $\mathcal{T}_K$, so that we have sufficient time to observe the system both when the control is on, and when it is off. This condition gives us two sensible options: one with $\tau_0 < 13.5$ h and $\tau_0 + \tau \approx 19$ h, and another with $\tau_0 \approx 19$ and $\tau_0 + \tau = 25$. These two options each correspond to a local minimum of Δr with respect to τ_0 and τ as we found in Fig. 2(d).

Since the local sensitivities ϕ_θ, and therefore $\mathcal{T}_\theta$, are much cheaper to compute compared to $\Delta\theta$, this heuristic provides a cheaper alternative to the profile likelihood-based approach. A major disadvantage of it, however, is that it only uses local information in the parameter space, which is the same major weakness of using the FIM to perform identifiability analysis, hence all criticisms of FIM also apply here. Furthermore, the heuristic method does not take into account the ways the parameters compensate for each other (which is encoded in the second mixed partial derivatives of the likelihood function with respect to the parameters, that is, the off-diagonal elements of the FIM). Furthermore, local sensitivity analysis by itself can be misleading—this is illustrated in Section 2.3.

2.3. Using u_δ as the control variable

We now consider applying a different control, u_δ, to the experiment, where we additively modulate the death rate, δ. A plausible way to apply this control is to add a toxin that kills cells. The model equations (Eq. (6)) in the presence of u_δ can be written as

$$\frac{\mathrm{d}C}{\mathrm{d}t} = rC\left[1 - \frac{C}{K}\right] - (\delta + u_\delta(t))C, \tag{15}$$

where u_δ has units h^{-1}. Applying the same reasoning from the beginning of Section 2.2, it can be shown that any parameter set $\boldsymbol{\theta}$ can produce the same model output as $\hat{\boldsymbol{\theta}}$ if the following two equations are satisfied:

$$r_{\text{eff}} = r - \delta = \hat{r} - \hat{\delta},$$
$$K_{\text{eff}} = K(1 - \delta/r) = \hat{K}(1 - \hat{\delta}/\hat{r}).$$

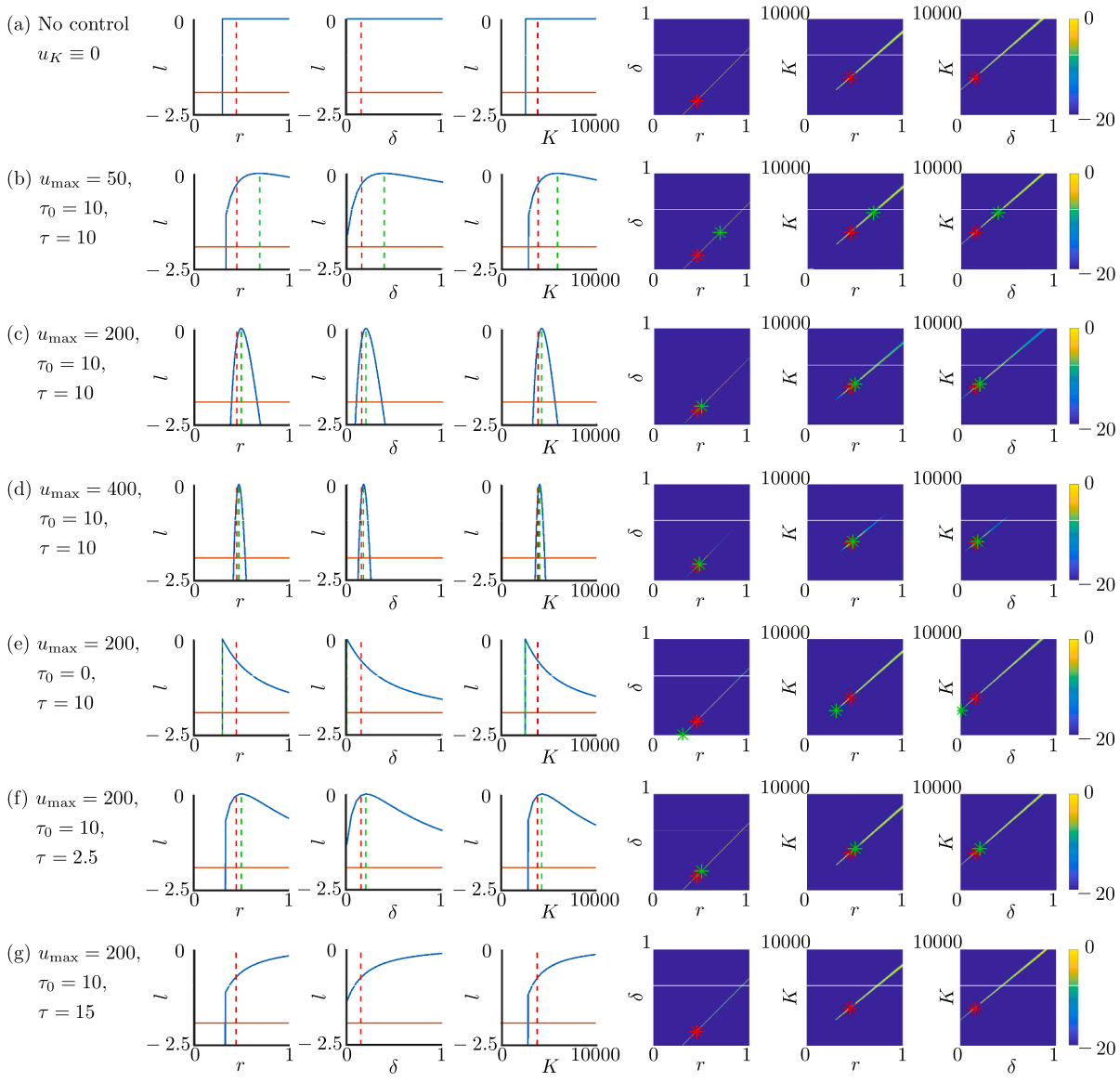


Fig. 1. The profile likelihood functions for the parameters in the logistic model with u_K as an additive control (Eq. (11)), using a synthetic dataset generated with the parameter values in Eq. (10), with $u_K(t)$ being a window function as in Eq. (5) with height $u_{\max}$, turning on at $t = \tau_0$, with a width of τ, and u_δ and u_r set to zero. Starting from the left, the first three columns show the univariate profile likelihoods for r, δ and K, respectively, and the next three columns show the bivariate profile likelihoods for r–δ, r–K, and δ–K, respectively. The green dashed lines in the univariate likelihood plots denote the MLEs, and the red dashed lines denote the parameter values used to generate the data. These are denoted with green and red stars in the bivariate plots. The parameter units are as in Eq. (8).

These two equations are not sufficient to uniquely determine (r, δ, K), and therefore the model remains structurally non-identifiable in the presence of u_δ. This means that applying u_δ as an additive control will not improve parameter identifiability. However, this is not revealed by local sensitivity analysis (Fig. 3(b)), which suggests that applying u_δ in the early stages of the experiment might be useful for improving parameter identifiability. This demonstrates that local sensitivity analysis, while a useful heuristic, cannot fully replace identifiability analysis for experimental design.

2.4. Using u_r as the control variable

We consider one last control input to the experiment, u_r, which additively modulates the proliferation rate of the cells, r. This can be done by e.g. applying a certain signalling protein that accelerates the cell cycle, therefore speeding up growth, corresponding to a temporarily elevated r. The model equations (Eq. (6)) in the presence of u_δ can be written as

$$\frac{\mathrm{d}C}{\mathrm{d}t} = (r + u_r(t))C\left[1 - \frac{C}{K}\right] - \delta C, \tag{16}$$

where $u_r(t)$ has units h^{-1}. With u_r as the control variable, the conditions for two parameter sets $\boldsymbol{\theta} = (r, \delta, K)$ and $\hat{\boldsymbol{\theta}} = (\hat{r}, \hat{\delta}, \hat{K})$ to produce the same model output in the absence of observational noise is

$$r_{\text{eff}} = r - \delta = \hat{r} - \hat{\delta}, \tag{17a}$$

$$K_{\text{eff}} = K(1 - \delta/r) = \hat{K}(1 - \hat{\delta}/\hat{r}), \tag{17b}$$

$$K\left(1 - \frac{\delta}{r + u_{\max}}\right) = \hat{K}\left(1 - \frac{\hat{\delta}}{\hat{r} + u_{\max}}\right), \tag{17c}$$

which are sufficient to ensure that $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, and therefore the addition of u_r makes the model parameters structurally identifiable.

As in Section 2.2, we explore how Δr changes as the control parameters change. We centre our explorations around $u_{\max} = 0.02\ \mathrm{h}^{-1}$, $\tau = 10$

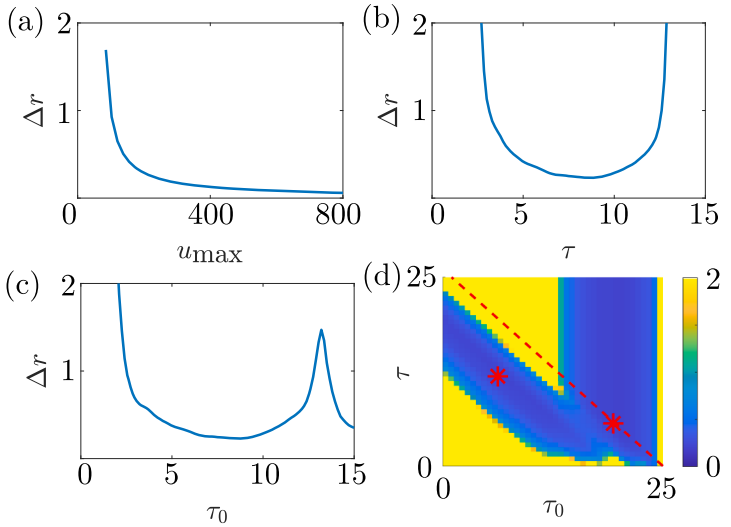


Fig. 2. The widths of the confidence intervals for r, Δr, in the logistic model with additive control u_K (Eq. (11)), using the synthetic dataset generated with the parameter values in Eq. (10), as a function of the control parameters for u_K defined in Eq. (5). (a–c) Δr as a function of a single varying control parameter. The control parameter being varied is labelled on the x-axis in each plot, while the other control parameters are held fixed at $u_{\max} = 200$ cells/mm^2, $\tau_0 = 10$ h, $\tau = 10$ h. (d) Δr shown as a function of τ_0 and τ, represented as a heat-map, with $u_{\max} = 1200$ cells/mm^2 fixed. The red dashed line marks the line $\tau_0 + \tau = 25$ h. The two red stars denote local minima of Δr at $(\tau_0, \tau) = (6.3\text{ h}, 12.4\text{ h})$, $(19.5\text{ h}, 5.5\text{ h})$.

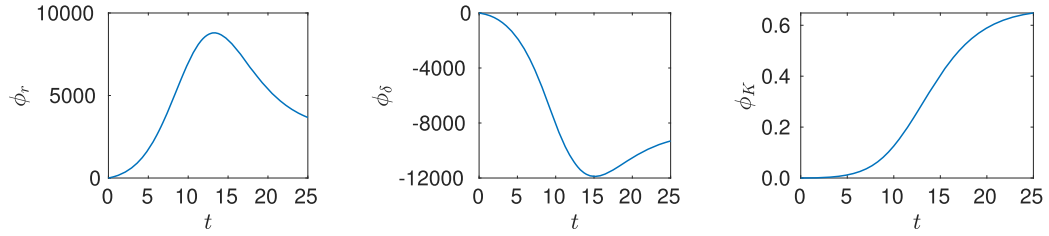


Fig. 3. Local sensitivities of the logistic model (Eq. (6)) with parameter values from Eq. (10), with respect to the model parameters r, δ, K, as defined in Eq. (13). Units are $[t] = \text{h}, [\phi_r] = [\phi_\delta] = \text{cells} \cdot \text{h/mm}^2, [\phi_K] = 1$.

h, $\tau_0 = 10$ h, which makes the model identifiable, and vary one or two of these parameters at a time. The results are plotted in Fig. 4.

We observe that the overall shape of Δr in the u_r case shown in Fig. 4 is similar to that for u_K, shown in Fig. 2. A higher $u_{\max}$, the magnitude of the control, is always helpful for improving identifiability. The combinations of (τ_0, τ) that give the highest degree of identifiability satisfy either $\tau_0 < 16$ h and $\tau_0 + \tau \approx 18$ h, or $\tau_0 > 16$ h and $\tau_0 + \tau \approx 25$ h. The sensitive interval (Eq. (14)) for r is $\mathcal{T}_r \approx [8, 22]$, and the above combinations represent windows that roughly overlap with half of this interval, so our sensitivity-based heuristic is still valid in this case.

The results in this section show that an intelligent choice of the control variable, along with a strategic choice of the control, can be effective at removing structural non-identifiability and improving practical identifiability. This choice can be informed by using profile likelihoods to inspect the dependence of uncertainty in parameter estimates on the control. We have also shown a heuristic way to select the control, which involves computing local sensitivities of the parameters and some simple calculations. This heuristic is intuitive to understand, and is much less computationally expensive, but should be used only in conjunction with structural identifiability analysis.

3. Optimal experimental design for model discrimination

In the previous section, we explored how to improve the identifiability of model parameters by designing an experiment based on a single hypothesis about the underlying dynamics, represented by the "ground truth" parameter set. We now consider a slightly different scenario, where we would like to decide between multiple competing hypotheses, presented by models with different equations and parameter values. This scenario can arise if there are multiple proposed mechanisms that can, in theory, lead to similar observations, or if parameter fitting using existing data resulted in a multi-modal posterior distribution for the parameter values, where each mode can be seen as a distinct model. In this scenario, we seek to optimally distinguish two models with fixed parameter values by carefully designing a control input. Suppose that the two models can be written as

$$\frac{\mathrm{d}\boldsymbol{x}_i}{\mathrm{d}t} = f_i(\boldsymbol{x}_i, t, \boldsymbol{u}), \quad \boldsymbol{y}_i = g_i(\boldsymbol{x}) + \boldsymbol{\xi}_i, \quad \boldsymbol{x}_i(0) = \boldsymbol{x}_0, \quad i = 1, 2, \tag{18}$$

and $\boldsymbol{x}_i(t; \boldsymbol{u}(t))$ is the solution to these models given the control input $\boldsymbol{u}(t)$. Since the two models describe the same system, we assume that they share the same initial condition. We wish to maximise the squared difference between model measurements over time, that is, we want to find $\boldsymbol{u}(t)$ to maximise the cost functional

$$\begin{aligned} \underset{\boldsymbol{u}\in\mathcal{U}}{\operatorname{argmax}}\, J[\boldsymbol{u}] &= \underset{\boldsymbol{u}\in\mathcal{U}}{\operatorname{argmin}} \int_0^T \mathcal{L}(\boldsymbol{x}, \boldsymbol{u})\mathrm{d}t, \\ \mathcal{L}(\boldsymbol{x}, \boldsymbol{u}) &= -\left[g_1(\boldsymbol{x}_1(t; \boldsymbol{u}(t))) - g_2(\boldsymbol{x}_2(t; \boldsymbol{u}(t)))\right]^2 + \boldsymbol{\alpha} \cdot \boldsymbol{u}^2(t), \end{aligned} \tag{19}$$

where $\mathcal{U}$ denotes the set of values the control variable is allowed to take, and $\boldsymbol{\alpha}$ encodes the cost of applying the control. We note that, in practice, the measurements are noisy, and made at discrete time points (e.g. $\boldsymbol{y}_{i,j} = g(\boldsymbol{x}_i(t_j))$), and the integrals should be replaced with summation. However, for simplicity, we assume that we have access to noiseless measurements in continuous time. The integrand $\mathcal{L}$ is known as the Lagrangian, which is assumed to be continuous in all its inputs, and all of its partial derivatives are assumed to be continuous. We consider both

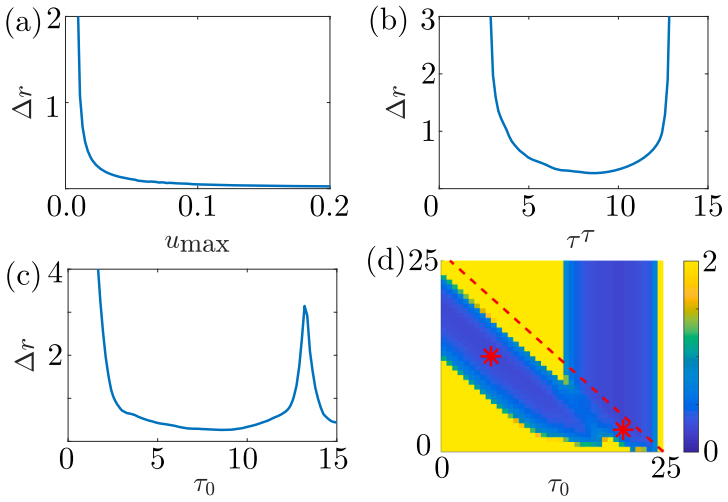


Fig. 4. Similar to Fig. 2, except the applied control variable is u_r instead of u_K. (a-c) The widths of the confidence intervals of r, Δr are shown as a function of one of the control parameters for u_r as defined in Eq. (5), with the other two held fixed at $u_{\max} = 0.02, \tau_0 = 10, \tau = 10$. The units are given in Eq. (8). (d) Δr shown as a function of τ_0 and τ, represented as a heat-map, with $u_{\max} = 0.02$ fixed. The red dashed line marks the line $\tau_0 + \tau = 25$. The two red stars denote the local minima of Δr at (τ_0, τ) =(5.6 h, 13.0 h), (20.5 h, 4.5 h).

the case of continuous control, and bang-bang control. In continuous control, each of the control variables may take any value between zero and an imposed upper bound. The upper bounds represent what is realistically achievable in an experimental setting, or a "safe" level beyond which we might alter the system dynamics too much for the model to be valid. In bang-bang control, the control variable is a piecewise constant function that may only take the value of zero or the upper bound. While continuous control allows more flexibility and greater room for optimisation, in an experiment it is often difficult to fine-tune the precise value of the control, so bang-bang control may be more applicable. For this reason, bang-bang control has seen wide-ranging applications, such as cancer chemotherapy [30] and viral treatment [31].

To solve this optimal control problem, we apply Pontryagin's Maximum Principle (PMP), using the Forward-Backward Sweep (FBS) algorithm to solve the resulting systems of equations numerically [32–35]. The details of the FBS algorithm are given in Supplementary Materials S.1. We will now illustrate the aforementioned approach for experimental design by applying PMP with the FBS algorithm to optimally discriminate between models of cell proliferation.

3.1. Logistic model with additive controls

First, consider discriminating between two instances of the logistic model with the following two parameter sets:

$$\frac{dC_1}{dt} = 0.45C_1\left(1 - \frac{C_1}{3900}\right) - 0.15C_1 = f_L(0.45, 0.15, 3900), \tag{20a}$$

$$\frac{dC_2}{dt} = 0.3C_2\left(1 - \frac{C_2}{2600}\right) = f_L(0.3, 0, 2600), \tag{20b}$$

where f_L is from Eq. (6). The timespan of the experiment is taken to be $T = 25$ h, and the initial condition is $C(0) = 100$ cells/mm^2. The first parameter set was chosen to coincide with those in Eq. (10), and the second parameter set has the same effective parameter values r_{eff} and K_{eff} (Eq. (9)) as the first parameter set, which means they give rise to identical solutions without control. We might be interested in discriminating between these two models from a biological perspective to determine whether the mechanism represented by the linear death term genuinely exists within the system.

We first consider the model with additive controls, as we have done in Section 2. As discussed in Section 2.3, applying u_δ additively does not remove structural identifiability, as both models yield the exact same solution for any u_δ. Therefore, we consider only applying u_K or u_r.

When applying u_K additively, the model equations become

$$\frac{dC_1}{dt} = 0.45C_1\left(1 - \frac{C_1}{3900 - u_K(t)}\right) - 0.15C_1, \tag{21a}$$

$$\frac{dC_2}{dt} = 0.3C_2\left(1 - \frac{C_2}{2600 - u_K(t)}\right). \tag{21b}$$

We take $u_{K,\max} = 1200$ cells/mm^2 as the upper bound for the control u_K, and $\alpha_K = 0.03$ as the cost of applying control. A much larger α_K (e.g., $\alpha_K = 0.1$) results in the optimal control $u_K^*(t) \equiv 0$, as the cost of the control is too great to be compensated by the reward of being able to discriminate the models. On the other hand, a much smaller α_K (e.g., $\alpha_K = 0.005$) results in $u_K^*(t) \equiv u_{K,\max}$, as a larger u_K is always helpful in discriminating between the two models in this case, and the cost of turning u_K on is negligible. Therefore, there is a range for α_K to make the problem non-trivial.

In Fig. 5, we show the result of applying the FBS algorithm to the optimal control problem, with u_K the control variable. The strategy implied by the resulting optimal control can be summarised as allowing both solutions to reach roughly $K_{\text{eff}}/2$ without much intervention, then abruptly turning on the control. In the early stages of the experiment (for small t), the carrying capacity K (and therefore the control variable u_K) has relatively little impact on the solutions of both models, so the control remains low. From $t \approx 9.5$ h, the impact of the control on discriminating between the two model solutions is large enough to overcome the cost of applying the control, and u_K is turned on to the maximum, where it remains until $t \approx 23.3$ h. After this point, the control is almost turned off completely until $t = T$, when it is exactly zero, as we know it must be. We can relate this control strategy to the local sensitivity, ϕ_K. In Fig. 3(c), we can observe that ϕ_K is monotonically increasing with time. The period of time where $u_K^* = u_{K,\max}$ roughly coincides with the time window where ϕ_K is large. This is expected, since the control is likely to be most effective in distinguishing between the model solutions at points where the models are sensitive to the parameter associated with the control. Interestingly, the optimal control has a small "bump" near $t \approx 24$ h. The reason for its presence is not immediately apparent, however removing the "bump" by setting $u_K = 0$ past $t = 24$ h results in a higher cost functional, suggesting that it is not simply a numerical artifact.

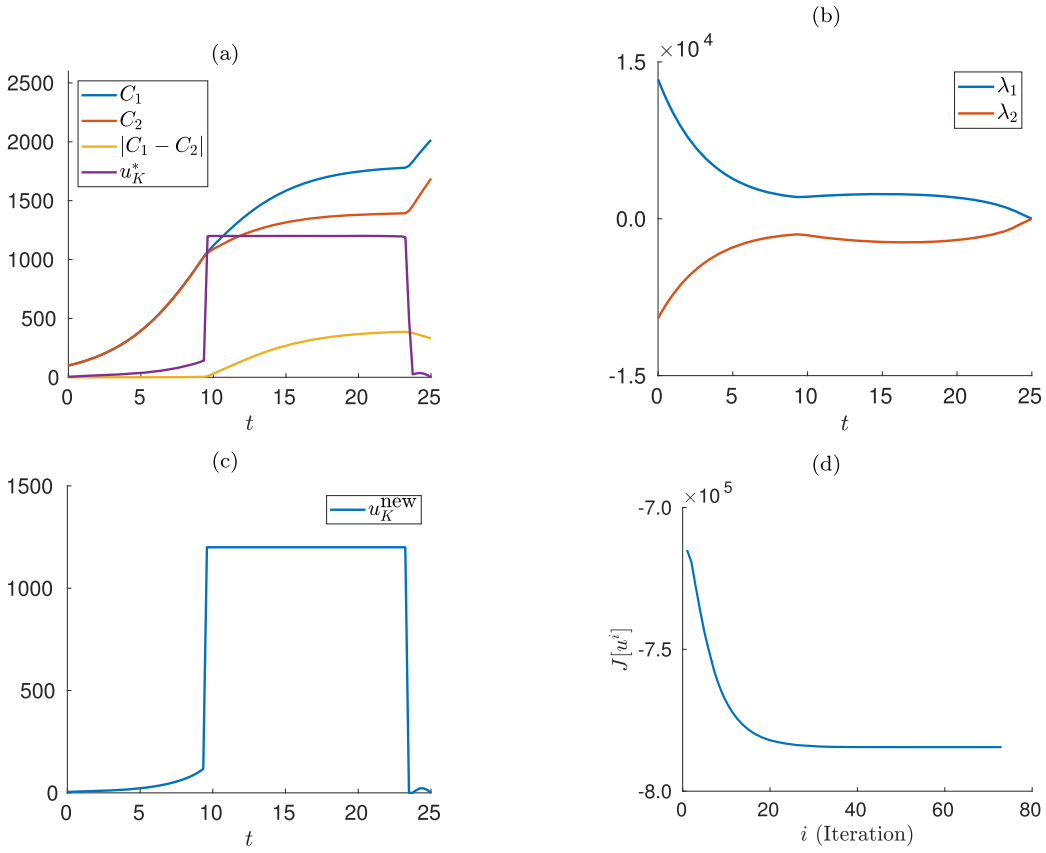


Fig. 5. Results for discriminating between logistic models with additive u_K as control (Eqs. (19), (21)), $u_{K,\max} = 1200$ cells/mm^2, and $\alpha_K = 0.03$. (a) The optimal control $u_K^*(t)$, as found by the algorithm, and the corresponding trajectories of the state variables C_1, C_2, as well as their difference over time. (b) The trajectories of the adjoint variables λ_1, λ_2 (auxillary variables for solving the control problem, see Supplementary Materials S.1) corresponding to the optimal control. (c) u_K^{new}, the update to the control, computed at the final iteration before termination of the algorithm. It is very close to u_K^*, suggesting that further iterations will not lead to significant changes in the control. (d) The cost functional $J[u]$ over the iterations, showing its convergence near the final iterations. The units $[u_K]$ =cells/mm^2, $[J] = (\text{cells/mm}^2)^2$.

Next, we consider a few variations of the optimal control problem, to determine how various factors impact the optimal control. Changing the value of α_K does not result in a qualitative change in the optimal control. Fig. 6 shows that with a higher value of $\alpha_k = 0.05$, the control is turned on over a shorter window between $t \approx 12$ h and $t \approx 22.3$ h, but retains the same shape as in Fig. 5. The small bump near $t = T$ is still present. The optimal bang-bang control for the same α, as found by direct optimisation, is turned on at $t \approx 9.4$ h, almost the same time as when the optimal continuous control is ramped up sharply to its maximum, and turned off at $t \approx 23.4$ h, very close to the time when the optimal continuous control decreases sharply to near zero. The bang-bang optimal control achieves an objective value of $J = -779401$ $(\text{cells/mm}^2)^2$, only slightly worse than the continuous optimal control, which is $J = -784470$ $(\text{cells/mm}^2)^2$. These results show that in this case, we do not gain much from allowing the control to take continuous values, and a bang-bang optimal control is both easier to compute, and experimentally easier to implement.

The optimal control for model discrimination is quite different from the optimal control for improving parameter identifiability, discussed in Section 2.2. This highlights that model discrimination and improving identifiability are distinct problems. In the model discrimination problem, we are essentially proposing a new experiment in addition to the experiment with which we calibrated our two models. For the new experiment to be most useful, we want to observe the system behaviour in an environment as different to the first experiment as possible, i.e., with the control turned to a maximum, whereas in the problem of identifiability improvement, we are essentially conducting two experiments at once, so the data should be split between when the control is on or off.

Next, we consider the case of applying u_r as the control variable, with the following model equations:

$$\frac{dC_1}{dt} = (0.45 + u_r(t))C_1\left(1 - \frac{C_1}{3900}\right) - 0.15C_1, \tag{22a}$$

$$\frac{dC_2}{dt} = (0.3 + u_r(t))C_2\left(1 - \frac{C_2}{2600}\right). \tag{22b}$$

The optimal control found by the FBS algorithm is shown in Fig. 7. The optimal control is roughly constant, at $u_r \approx 0.7$ h^{-1} over most of the domain, with a noticeable bump near the inflection point of the solutions, and ramps down smoothly to zero as $t \to T$. The strategy seems to be to turn on the control early so that both solutions reach their carrying capacity quickly, then maintain the control in order to observe the difference in K_{eff} for the two models.

The reason for a local maximum in u_r^* near the inflection point of the two model solutions is not immediately clear, but a possible explanation can be given using local sensitivity. Examining the plot of local sensitivity ϕ_r (Fig. 3(a)) and the associated model solution (yellow curve in Fig. 9), we can observe that the inflection point of the model solution is near the global maximum of ϕ_r. Therefore, the local maximum of the control, u_r^*, is close to where both model solutions are most sensitive to the parameter r, which is sensible for the objective of distinguishing the two model solutions. The window in time where the optimal bang-bang control is turned on covers most of the time span of the simulation, but it is turned off a short time before $t = T$, around the same time as

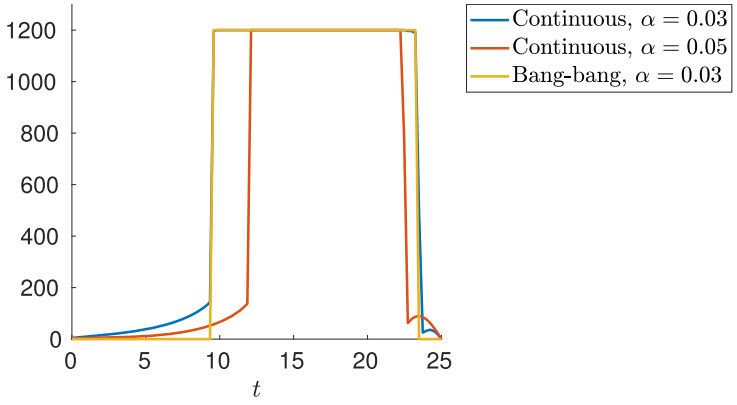


Fig. 6. Comparison of the solutions of variations of the optimal control problem. The blue curve is the same control as in Fig. 5. The orange curve is the optimal continuous control with a different value for the weight of control cost, α, and the yellow curve is the optimal bang-bang control using the same value of α as the blue curve. The units are $[t] = \text{h}, [u_K] = \text{cells/mm}^2$.

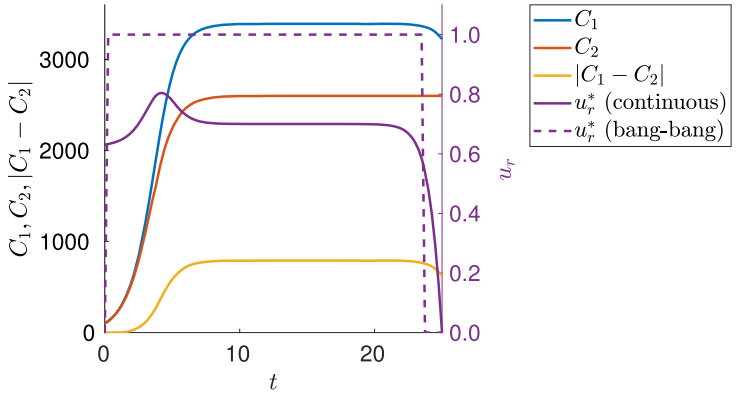


Fig. 7. Results of FBS (Algorithm 1) applied to the optimal control problem for discriminating between logistic models with additive u_r as control (Eqs. (19), (22)), $u_{r,\max} = 1\ \text{h}^{-1}$ as the control upper bound, and $\alpha_r = 500000\ \text{h}^2(\text{cells/mm}^2)^2$ as control cost, showing the optimal control $u_K^*(t)$ found by the algorithm, the corresponding trajectories of the state variables C_1, C_2, and their difference over time. The optimal bang-bang control is also shown for comparison. The units are $[C] = \text{cells/mm}^2, [t] = \text{h}, [u_r] = \text{h}^{-1}$.

the optimal continuous control is ramped down. The bang-bang control achieved an objective value of $J = -4618503\ (\text{cells/mm}^2)^2$, compared to $J = -6388826\ (\text{cells/mm}^2)^2$ for the optimal continuous control (dashed purple line in Fig. 7), which is a noticeably worse performance.

3.2. Logistic model with multiplicative controls

Next, we again consider the case for distinguishing between two instances of the logistic model, but now the controls are assumed to affect the associated parameters multiplicatively, rather than additively. While this is a small conceptual difference compared to the previous problem in Section 3.1, it nonetheless results in significant changes in the outcome. The model equations (Eq. (20)) in the presence of multiplicative controls can be written as

$$\frac{\mathrm{d}C_1}{\mathrm{d}t} = 0.45(1 + u_r(t))C_1\left(1 - \frac{C_1}{3900(1 - u_K(t))}\right) - 0.15(1 + u_\delta(t))C_1, \quad (23a)$$

$$\frac{\mathrm{d}C_2}{\mathrm{d}t} = 0.3(1 + u_r(t))C_2\left(1 - \frac{C_2}{2600(1 - u_K(t))}\right). \quad (23b)$$

Note that Eq. (23b) does not contain u_δ because $\delta = 0$ in that model, so a multiplicative u_δ has no effect. The control variables u_r, u_δ, u_K are dimensionless in this formulation. As before, we only consider applying one control variable at a time, while setting the other two to zero.

In contrast to the case of additive controls, where u_δ is ineffective in removing structural non-identifiability and u_r and u_K are effective, as shown in Section 2, in the multiplicative case u_K is ineffective in removing structural non-identifiability, whereas u_r, u_δ are effective (Fig. 8). For u_r, the strategy implied by the optimal control seems to be to turn the control on during most of the time span of the simulation, then ramp it down smoothly to zero as $t \to T$. This is because u_r has an impact on the difference between the two solutions throughout the time span of the simulation, but towards the end of the simulation the cost of keeping it on outweighs the gain in the difference between model solutions. Comparing the shape of u_r^* in Fig. 8(a) to that in the additive control case shown in Fig. 7, we can observe that in both cases, u_r^* has a local maximum near the inflection points of the two model solutions but, in the multiplicative case, u_r^* decreases more rapidly after that point, whereas it stays near a constant value for a period of time in the additive case. This is because u_r has a larger effect on the steady state of the models in the additive case compared to the multiplicative case. Therefore, as the model solutions approach the steady state K_{eff}, the effect of u_r on model difference diminishes in the multiplicative case, while the effect remains significant in the additive case as the model solution approaches the steady state.

In the case of u_δ, the optimal control strategy involves setting its value to the maximum for most of the simulation time span, followed by a sharp decrease in the control as t approaches T. This can be explained by the fact that the model solution C_2 is entirely unaffected by u_δ, due to its parameter δ_2 being equal to zero, whereas C_1 is influenced by u_δ at all time. Maintaining a high value of u_δ depresses C_1 relative to C_2, which is helpful for distinguishing the two model solutions in all stages of the experiment.

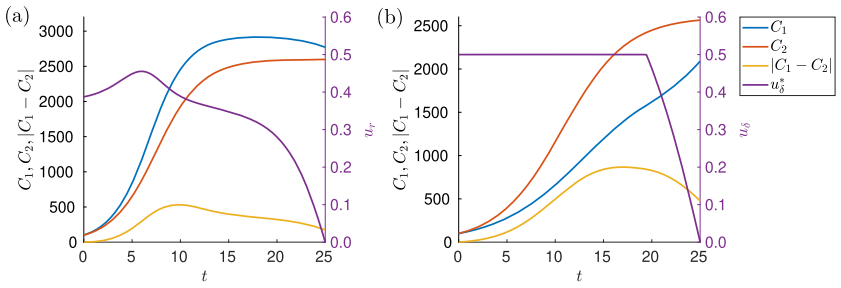


Fig. 8. Results of the FBS (Algorithm 1) applied to the optimal control problem for discriminating between logistic models with multiplicative controls (Eqs. (19), (23)). (a) The control variable is u_r, with $u_{r,\max} = 0.5$, $\alpha_r = 700000$ (cells/mm^2)^2. Here $u_\delta = u_K \equiv 0$. (b) The control variable is u_δ, with $u_{\delta,\max} = 0.5$, $\alpha_\delta = 1300000$ (cells/mm^2)^2. Here $u_r = u_K \equiv 0$. The units are $[C]$ = cells/mm^2, $[t]$ = h, and u_r and u_δ are dimensionless.

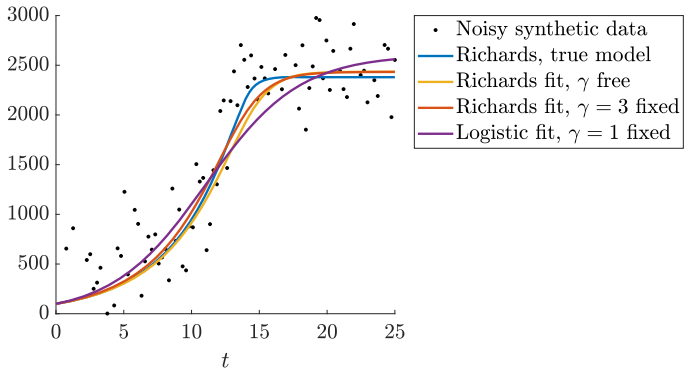


Fig. 9. Illustrations of typical solutions of the logistic and Richards models. The blue, red, yellow, and purple curves are the solution to the Richards model (Eq. (24)) with the parameters $(r, \gamma, K) = (0.225, 8, 2381)$ (true model), $(0.222, 3.709, 2435)$, $(0.235, 3, 2433)$, and $(0.291, 1, 2605)$, respectively. The units are $[r] = \text{h}^{-1}$, $[K] = \text{cell/mm}^2$, $[\gamma] = 1$.

3.3. Richards model with additive controls

We provide another example for model discrimination using the Richards model, which is an extension of the logistic model for proliferation. Richards [36] introduced an additional "shape parameter" $\gamma > 0$ to the logistic model, which provides more flexibility in describing the shape of the growth curve. The model can be written as

$$\frac{\mathrm{d}C}{\mathrm{d}t} = rC\left[1 - \left(\frac{C}{K}\right)^\gamma\right] = f_R(C; r, \gamma, K), \tag{24}$$

where the units and interpretations of the quantities in this model remain the same as in the logistic model, and setting $\gamma = 1$ reduces the Richards model to the logistic model. It has been shown in [26] that the Richards model, without control, is practically non-identifiable given a realistic amount of data, with the identifiability of the parameter γ being especially poor. We plot typical solutions to the Richards model, and demonstrate its practical non-identifiability in Fig. 9. The black dots are synthetic data generated by perturbing the exact solution to the Richards model using a parameter set with $\gamma = 8$ (blue curve), by adding Gaussian i.i.d. noise with $\sigma = 400$ cell/mm^2. This parameter set and noise magnitude are within biologically realistic ranges, as found in [16]. The other three curves are the solutions of the Richards model using parameter sets found by fitting the model to the data (i.e. MLEs), with γ free (i.e. simultaneously fitting three parameters, r, K and γ), $\gamma = 3$ or $\gamma = 1$ fixed (i.e. fitting the remaining two parameters, r and K), respectively. Notice that the differences between the curves are small, and the parameter set obtained by fitting the Richards model (corresponding to the yellow curve) is far from the true parameter values, a symptom of non-identifiability. If we mistakenly believed that the true parameter set has $\gamma = 3$, or if the logistic model ($\gamma = 1$) was the true model, the data would be insufficient to refute either belief, as the red and purple curves fit just as well.

Therefore, as an example, we seek to design an experiment to discriminate between the two instances of the Richards model that correspond to the blue and red curve in Fig. 9, using additive control inputs:

$$\frac{\mathrm{d}C_1}{\mathrm{d}t} = (0.225 + u_r(t))C_1\left[1 - \left(\frac{C_1}{2381 - u_K(t)}\right)^8\right] - u_\delta(t)C_1, \tag{25a}$$

$$\frac{\mathrm{d}C_2}{\mathrm{d}t} = (0.235 + u_r(t))C_2\left[1 - \left(\frac{C_2}{2433 - u_K(t)}\right)^3\right] - u_\delta(t)C_2. \tag{25b}$$

In Fig. 10, we present the optimal control obtained with the FBS algorithm in the case where one of the three control variables, u_r, u_δ, and u_K, is applied individually. Unlike the case of distinguishing between logistic models in Sections 3.1 and 3.2, all three controls are at least somewhat effective. Out of the three, u_r helps the least to distinguish between the two models. The optimal control u_r^* never hits the imposed upper bound $u_r^{\max}$, and it is turned on relatively briefly. The difference between the models in the presence of the control is negligibly greater than that in the absence of the control.

Both u_δ and u_K are much more effective. The control strategy implied by u_δ^* allows both models to approach their respective carrying capacity, then turns u_δ on abruptly to maximum. This is a sensible strategy—the effect of u_δ is similar for both models in the earlier phase, where both model solutions, C_1 and C_2, display exponential growth, and u_δ lowers the effective growth rate r_{eff} by the same amount for both models, so it is useless to turn on the control in that phase. The control is turned on during the later saturation phase, where both model solutions approach their respective steady states. The steady state, with u_δ treated as a constant, can be written as

$$K\left(1 - \frac{\delta + u_\delta}{r}\right)^{1/\gamma}.$$

A greater value of γ diminishes the effect of u_δ on the steady state (since the exponent $1/\gamma$ approaches zero, so the steady state approaches K, effectively independently from u_δ), therefore, the steady state of model 2, which has a lower γ, is decreased much more by a positive u_δ compared

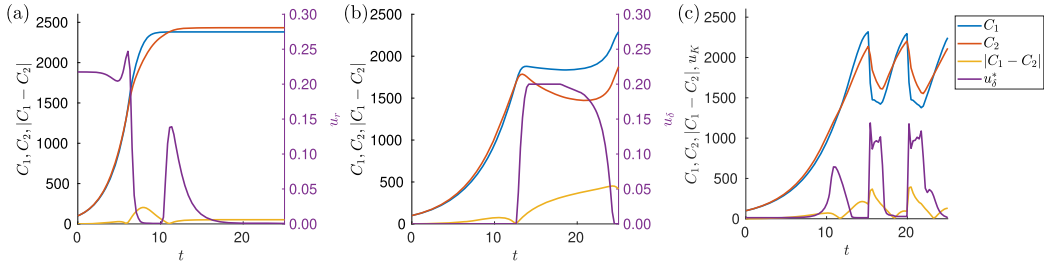


Fig. 10. Results of the FBS algorithm (Algorithm 1) applied to the optimal control problem for discriminating between Richards models with additive controls (Eqs. (19), (25)). (a) The control variable is u_r, with $u_{r,\max} = 0.4$ h^{-1}, $\alpha_r = 30000$ h^2(cells/mm^2)^2. (b) The control variable is u_δ, with $u_{\delta,\max} = 0.2$ h^{-1}, $\alpha_\delta = 500000$ h^2(cells/mm^2)^2. (c) The control variable is u_K, with $u_{K,\max} = 1200$ cells/mm^2, $\alpha_K = 0.03$. The units are $[C]$ = cells/mm^2, $[t]$ = h.

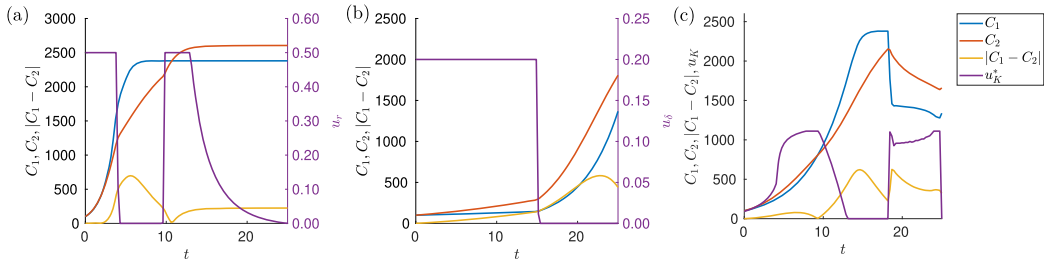


Fig. 11. Results of the FBS algorithm (Algorithm 1) applied to the optimal control problem for discriminating logistic and Richards models (Eqs. (19), (26)). (a) The control variable is u_r, with $u_{r,\max} = 0.5$ h^{-1}, $\alpha_r = 30000$ h^2(cells/mm^2)^2. (b) The control variable is u_δ, with $u_{\delta,\max} = 0.2$ h^{-1}, $\alpha_\delta = 30000$ h^2(cells/mm^2)^2. (c) The control variable is u_K, with $u_{K,\max} = 1200$ cells/mm^2, $\alpha_K = 0.03$. The units are $[C]$ = cells/mm^2, $[t]$ = h.

to that of model 1. The difference in steady states means turning the control on in the saturation phase allows the two models to be easily distinguished.

In contrast, u_K modifies the K_{eff} of both models by the same amount, which means the case of using u_K as the control variable requires a different strategy from the case of using u_δ. The strategy in the case of u_K is to first allow both models to get close, but not too close, to their respective carrying capacities, then cyclically and abruptly turn u_K on and off so that both model solutions stay near, but not too close, to their respective carrying capacities. Recall that in the absence of the control, the two model solutions are best distinguished by their behaviour just before they approach K. This control strategy maximises the amount of time spent within the regime where the behaviours of the two models are most different, and thereby maximises the difference between the two model solutions.

3.4. Logistic and Richards model with additive controls

In the final example, we consider the problem of discriminating between instances of the logistic model and the Richards model, again using additive controls:

$$\frac{dC_1}{dt} = (0.225 + u_r(t))C_1\left[1 - \left(\frac{C_1}{2381 - u_K(t)}\right)^8\right] - u_\delta(t)C_1, \tag{26a}$$

$$\frac{dC_2}{dt} = (0.291 + u_r(t))C_2\left[1 - \frac{C_2}{2605 - u_K(t)}\right] - u_\delta(t)C_2. \tag{26b}$$

In the absence of controls, the solutions to these two models are the blue and purple curves in Fig. 9, which are very similar. We plot the optimal controls for discriminating between these two models, as found by the FBS algorithm, in Fig. 11.

For the case of applying u_r, the strategy is to turn on the control in two stages. First, the control is turned on from the beginning of the experiment until both C_1 and C_2 are near their inflection point, at around $t \approx 5$ h, at which point the control is turned off. Since the Richards model approaches its carrying capacity much quicker than the logistic model after this point, we are able to observe a significant difference between the two model solutions. Later, the control is turned on again from $t \approx 10$ h, and gradually ramped down to zero. Since K_1 and K_2 are different, control maximises the difference between the two models by allowing the logistic model to approach its carrying capacity quicker, so we can observe the two different limiting behaviours for longer.

The optimal control strategy for applying u_δ is to simply turn on the control to maximum limit from the beginning of the experiment until $t \approx 15$, then abruptly turn it off. The effective growth rate for the Richards model is more heavily impacted by the control than that of the logistic model, so turning on the control depresses C_1 more than C_2, allowing us to distinguish between the two models. The control is turned off in the later stages of the experiment as the difference between C_1 and C_2 is already significant, and the cost of applying the control at that stage outweighs the additional discrepancy it can cause between the two models. Lastly, for the case of applying u_K, the strategy is similar to the one seen in Fig. 10, where we use the control to keep both C_1 and C_2 away from their equilibrium values, so we can distinguish the two models by their behaviour as they approach their carrying capacity.

Summarising the results of this section, we find that it is possible to design an optimal control to enable us to differentiate between two models that were previously indistinguishable, or produce very similar solutions without such intervention. We can understand the optimal control strategies intuitively by looking at how the control variable affects the difference between the models at various points in time, using local sensitivities of the parameters associated with the control variable as a rough guide. Whether the control variables act on model parameters additively or multiplicatively can make a big difference, as an efficient control strategy in one scenario may prove entirely ineffective in the other.

4. Discussion

This work examines how to achieve two objectives, maximising the practical parameter identifiability of a model and discriminating between two models to the greatest extent, by designing a control input. As such, it presents several major innovations over existing research [1,5,13–15]. First, we utilise profile likelihoods as a measure of practical identifiability, providing a more “global” view across the

parameter space, as opposed to the measures that rely on point estimates that were considered in prior work. In addition, we considered both general continuous controls as well as piecewise constant ("bang-bang") controls for the problem of designing optimal control inputs in an experiment—this grants a greater degree of flexibility compared to earlier work which restricted the space of allowed controls to simpler classes of functions.

We have established that the problems of model discrimination and improving identifiability, although related, represent distinct objectives and that, as a result, the optimal controls for achieving those objectives are different. This difference arises from the fact that the two objectives makes different assumptions on the availability of prior knowledge concerning the system. In tackling the problem of improving identifiability, we used an assumed set of parameter values to generate the synthetic data in order to evaluate the profile likelihoods. In contrast, for model discrimination, the model parameters are fixed at values that encode much more prior knowledge on the parameters. In the case where such knowledge is derived from an earlier experiment, in the model discrimination problem the optimal experimental design should ideally be very different from the earlier experiment to avoid collecting redundant information. On the other hand, there is less need to avoid repeating a similar experimental design if the goal is to maximise identifiability.

Our results highlight that structural non-identifiability issues can potentially be addressed by adding a control variable to the model, and incorporating the corresponding external stimulus into the experiment. The logistic model with a death term is structurally non-identifiable, as a change in δ, the death rate, can be perfectly compensated for by a corresponding change in the growth rate, r, and the carrying capacity, K. Adding the appropriate control variable allows the system to be observed in a state that otherwise would not be accessible, providing additional information on the underlying mechanisms. For example, using u_r as a control variable, with a window function for control, allows us to observe the system in two states, a state where the control is off, and another where the control is on. In this scenario, a change in δ can no longer be compensated for by changing r and K to obtain identical behaviour in both states. However, this approach can only succeed if knowledge of how the stimulus impacts the mechanisms (e.g., whether it induces an additive or multiplicative change in a parameter, and to what extent) is available. Without this knowledge, introducing the control variable introduces additional unknowns, which prevents the extraction of additional information on the parameters. For example, suppose we wish to introduce an external stimulus that modulates the growth rate, represented by the control variable u_r, but do not know how much this stimulus changes r in advance. It would be necessary to introduce another unknown parameter, say ρ, as a coefficient of u_r in the model. This ρ must be estimated in the same way as the original model parameters r, δ and K. As a result, the inference problem becomes more complicated and requires more data, preventing any gains from introducing the control variable. One possibility of circumventing this issue is to test the control mechanism on a system with known dynamics.

We also demonstrated that the optimal control for model discrimination depends on the control variable. The optimal strategy for control variables acting on different model parameters can be completely different. For example, in Fig. 10, the optimal strategies for the three control variables are qualitatively and quantitatively distinct. In some cases, the shape of the optimal continuous control is close to a window function, such as the case of an additive control u_K for distinguishing between two logistic models with death terms in Fig. 5, therefore it provides only a minor improvement upon the optimal bang-bang control. In this case, we might prefer to implement the bang-bang control in practice, since it requires relatively simple experimental techniques, whereas continuous control requires the ability to modulate the control mechanism more precisely. In other cases, the optimal continuous control significantly out-performs the optimal bang-bang control, due to its increased flexibility. This is the case for an additive u_r (Fig. 7), where the optimal continuous control has a significantly lower cost compared to the optimal bang-bang control. In these cases, the implementation of the continuous control is preferred, if it is feasible to carry out in the experiment. Note that bang-bang controls, or more generally piecewise constant controls, are more commonly used in control applications, due to the fact that they are both easier to compute and to implement compared to continuous controls.

The FBS algorithm with adaptive update rate, based on PMP, is adequate for solving continuous optimal control problems, at least for the examples considered in this work. We have investigated alternative algorithms for solving the optimisation problem, such as direct control, or a hybrid method combining FBS and direct control (see Supplementary Material Sec.S.4). In summary, the hybrid method can outperform FBS in certain cases, but its implementation is more difficult. Other methods for solving optimal control problems, such as shooting methods for PMP [37, Sec. 7], dynamic programming methods based on Hamilton-Jacobi-Bellman equations [38], and more recent machine learning-based methods [39–41], may also be appropriate.

This paper employs a frequentist, profile likelihood-based approach to measure parameter identifiability. However, it could be replaced with a Bayesian framework, which offers greater flexibility in quantifying parameter uncertainty, and mitigates the potential impact of a poorly chosen "ground truth" parameter set. In the Bayesian approach, instead of having a single "ground truth" parameter set as in Eq. (10), we have a prior distribution of parameter values, which could be simply a normal distribution centred at a "best guess", or the posterior distribution from a previous iteration of the experiment, if available. To evaluate a certain experiment design, we could sample multiple parameter sets from the prior, and simulate the model for each parameter set to create an ensemble of model outputs, which better accounts for the uncertainty in prior knowledge. Finally, we can use e.g. MCMC or approximate Bayesian computation to build a posterior distribution by fitting to the "ground truth" ensemble. The drawback is that this approach incurs a much higher computational cost, which can render it impractical in many applications.

There are many future directions in which to extend the methods considered in this paper. We can generalise the methods in Section 2 by optimising the practical identifiability of multiple parameters at the same time, using multi-objective optimisation. It would also be useful to consider the case of applying multiple control inputs at the same time. Finally, a key limitation of the approach proposed in this paper is the use of sets of fixed "ground truth" parameter values to guide the design process. A more robust approach would be to replace the "ground truth" parameter sets with a prior distribution over the parameters, which would better account for uncertainty.

CRediT authorship contribution statement

Yue Liu: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization; **Philip K. Maini:** Writing – review & editing, Supervision; **Ruth E. Baker:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Data and code availability

The code for the analysis in this paper is written in MATLAB, and will be provided on Github at https://github.com/liuyue002/experiment_design.

Declaration of competing interest

We declare no competing interests.

Acknowledgment:

YL is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Postgraduate Scholarships -

Doctoral program, reference number PGSD3-535584-2019, as well as by the Canadian Centennial Scholarship Fund (CCSF). For the purpose of open access, the authors have applied a CC BY public copyright licence to any author accepted manuscript arising from this submission.

Supplementary material

Supplementary material associated with this article can be found in the online version at 10.1016/j.mbs.2026.109710.

References

[1] C. Kreutz, J. Timmer, Systems biology: experimental design, FEBS J. 276 (4) (2009) 923–942. https://doi.org/10.1111/j.1742-4658.2008.06843.x
[2] W.G. Hunter, A.M. Reiner, Designs for discriminating between two rival models, Technometrics 7 (3) (1965) 307–323. https://doi.org/10.1080/00401706.1965.10490265
[3] K.-H. Cho, S.-Y. Shin, W. Kolch, O. Wolkenhauer, Experimental design in systems biology, based on parameter sensitivity analysis using a Monte Carlo method: a case study for the TNFα-mediated NF-κ B signal transduction pathway, Simulation 79 (12) (2003) 726–739. https://doi.org/10.1177/0037549703040943
[4] S. Bandara, J.P. Schlöder, R. Eils, H.G. Bock, T. Meyer, Optimal experimental design for parameter estimation of a cell signaling model, PLOS Comput. Biol. 5 (11) (2009) e1000558. https://doi.org/10.1371/journal.pcbi.1000558
[5] B. Steiert, A. Raue, J. Timmer, C. Kreutz, Experimental design for parameter estimation of gene regulatory networks, PLOS One 7 (7) (2012) e40052. https://doi.org/10.1371/journal.pone.0040052
[6] D. Espie, S. Macchietto, The optimal design of dynamic experiments, AIChE J. 35 (2) (1989) 223–229. https://doi.org/10.1002/aic.690350206
[7] B. Smucker, M. Krzywinski, N. Altman, Optimal experimental design, Nat. Methods 15 (8) (2018) 559–560. https://doi.org/10.1038/s41592-018-0083-2
[8] B.H. Chen, S.P. Asprey, S.K. Bermingham, A.M. Neumann, H.J.M. Kramer, On the design of optimally informative experiments for model discrimination among dynamic crystallization process models, in: 4th Int. Conf. on Foundations of Computer-Aided Process Operations, 2003, pp. 455–458.
[9] A.L. Burke, T.A. Duever, A. Penlidis, Model discrimination via designed experiments: discriminating between the terminal and penultimate models on the basis of composition data, Macromolecules 27 (2) (1994) 386–399. https://doi.org/10.1021/ma00080a011
[10] F.-G. Wieland, A.L. Hauber, M. Rosenblatt, C. Tönsing, J. Timmer, On structural and practical identifiability, Curr. Opin. Syst. Biol. 25 (2021) 60–69. https://doi.org/10.1016/j.coisb.2021.03.005
[11] R. Bellman, K.J. Åström, On structural identifiability, Math. Biosci. 7 (3) (1970) 329–339. https://doi.org/10.1016/0025-5564(70)90132-X
[12] C. Cobelli, J.J. DiStefano, Parameter and structural identifiability concepts and ambiguities: a critical review and analysis, Am. J. Physiol.-Regul. Integr. Comp. Physiol. 239 (1) (1980) R7–R24. https://doi.org/10.1152/ajpregu.1980.239.1.R7
[13] D. Faller, U. Klingmüller, J. Timmer, Simulation methods for optimal experimental design in systems biology, Simulation 79 (12) (2003) 717–725. https://doi.org/10.1177/0037549703040937
[14] J. Vanlier, C.A. Tiemann, P.A.J. Hilbers, N.A.W. van Riel, A Bayesian approach to targeted experiment design, Bioinformatics 28 (8) (2012) 1136–1142. https://doi.org/10.1093/bioinformatics/bts092
[15] T. Litwin, J. Timmer, C. Kreutz, Optimal experimental design based on two-dimensional likelihood profiles, Front. Mol. Biosci. 9 (2022) 800856.
[16] Y. Liu, K. Suh, P.K. Maini, D.J. Cohen, R.E. Baker, Parameter identifiability and model selection for partial differential equation models of cell invasion, J. R. Soc. Interface 21 (212) (2024) 20230607. https://doi.org/10.1098/rsif.2023.0607
[17] A.F. Villaverde, E. Raimúndez, J. Hasenauer, J.R. Banga, Assessment of prediction uncertainty quantification methods in systems biology, IEEE/ACM Trans. Comput. Biol. Bioinf. (2022) 1–12. https://doi.org/10.1109/TCBB.2022.3213914
[18] K.G. Gadkar, R. Gunawan, F.J. Doyle, Iterative approach to model identification of biological networks, BMC Bioinf. 6 (1) (2005) 155. https://doi.org/10.1186/1471-2105-6-155
[19] S.V. Chin, M.J. Chappell, Structural identifiability and indistinguishability analyses of the minimal model and a euglycemic hyperinsulinemic clamp model for glucose–insulin dynamics, Comput. Methods Programs Biomed. 104 (2) (2011) 120–134. https://doi.org/10.1016/j.cmpb.2010.08.012
[20] A. Raue, V. Becker, U. Klingmüller, J. Timmer, Identifiability and observability analysis for experimental design in nonlinear dynamical models, Chaos: Interdiscip. J. Nonlinear Sci. 20 (4) (2010) 045105. https://doi.org/10.1063/1.3528102
[21] M.A.J. Roberts, E. August, A. Hamadeh, P.K. Maini, P.E. McSharry, J.P. Armitage, A. Papachristodoulou, A model invalidation-based approach for elucidating biological signalling pathways, applied to the chemotaxis pathway in r. sphaeroides, BMC Syst. Biol. 3 (1) (2009) 105. https://doi.org/10.1186/1752-0509-3-105
[22] A. Hamadeh, E. August, M. Roberts, P.K. Maini, J.P. Armitage, B. Ingalls, A. Papachristodoulou, Feedback control architecture of the r. sphaeroides chemotaxis network, in: 2011 50th IEEE Conference on Decision and Control and European Control Conference, 2011, pp. 3014–3019. https://doi.org/10.1109/CDC.2011.6161020
[23] R. Bellman, I. Glicksberg, O. Gross, On the “Bang-Bang” control problem, Q. Appl. Math. 14 (1) (1956) 11–18. https://doi.org/10.1090/qam/78516
[24] J. Qi, R.E. Baker, Optimal experimental design for parameter estimation in the presence of observation noise, 2025, https://arxiv.org/abs/2504.19233.
[25] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, J. Timmer, Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, Bioinformatics 25 (15) (2009) 1923–1929. https://doi.org/10.1093/bioinformatics/btp358
[26] M.J. Simpson, A.P. Browning, D.J. Warne, O.J. Maclaren, R.E. Baker, Parameter identifiability and model selection for sigmoid population growth models, J. Theor. Biol. 535 (2022) 110998.
[27] M.J. Simpson, R.E. Baker, S.T. Vittadello, O.J. Maclaren, Practical parameter identifiability for spatio-temporal models of cell invasion, J. R. Soc. Interface 17 (164) (2020) 20200055. https://doi.org/10.1098/rsif.2020.0055
[28] S.A. Murphy, A.W. Van Der Vaart, On profile likelihood, J. Am. Stat. Assoc. 95 (450) (2000) 449–465. https://doi.org/10.1080/01621459.2000.10474219
[29] A.P. Browning, T.D. Lewin, R.E. Baker, P.K. Maini, E.G. Moros, J. Caudell, H.M. Byrne, H. Enderling, Predicting Radiotherapy Patient Outcomes with Real-Time Clinical Data Using Mathematical Modelling, Bull Math Biol. (2023). https://doi.org/10.48550/arXiv.2201.02101
[30] U. Ledzewicz, H. Schättler, Optimal bang-bang controls for a two-compartment model in cancer chemotherapy, J. Optim. Theory Appl. 114 (3) (2002) 609–637. https://doi.org/10.1023/A:1016027113579
[31] T. Mapder, S. Clifford, J. Aaskov, K. Burrage, A population of bang-bang switches of defective interfering particles makes within-host dynamics of dengue virus controllable, PLOS Comput. Biol. 15 (11) (2019) e1006668. https://doi.org/10.1371/journal.pcbi.1006668
[32] S. Lenhart, J.T. Workman, Optimal Control Applied to Biological Models, Chapman and Hall/CRC, 2007. https://doi.org/10.1201/9781420011418
[33] D. Liberzon, Calculus of Variations and Optimal Control Theory: A Concise Introduction, Princeton University Press, 2011.
[34] L.S. Pontrjagin, Mathematical Theory of Optimal Processes, Routledge, 1987. https://doi.org/10.1201/9780203749319
[35] J.A. Sharp, A.P. Browning, T. Mapder, K. Burrage, M.J. Simpson, Optimal control of acute myeloid Leukaemia, J. Theor. Biol. 470 (2019) 30–42. https://doi.org/10.1016/j.jtbi.2019.03.006
[36] F.J. Richards, A flexible growth function for empirical use, J. Exp. Bot. 10 (2) (1959) 290–301. https://doi.org/10.1093/jxb/10.2.290
[37] H.J. Pesch, A practical guide to the solution of real-life optimal control problems, Control Cybern. 23 (1) (1994) 7–60. https://doi.org/10.1.1.53.5766
[38] R. Bellmann, Dynamic Programming, Princeton University Press, 1957.
[39] D. Bertsekas, Reinforcement Learning and Optimal Control, Athena Scientific, 2019.
[40] C. Gu, H. Xiong, Y. Chen, Pontryagin optimal control via neural networks, 2024. https://doi.org/10.48550/arXiv.2212.14566
[41] X. Li, D. Verma, L. Ruthotto, A neural network approach for stochastic optimal control, SIAM J. Sci. Comput. 46 (5) (2024) C535–C556. https://doi.org/10.1137/23M155832X