A control theoretical approach to designing optimal experiments in systems biology

Bence Mélykúti

Dissertation

November 2007

Supervisor: Antonis Papachristodoulou
(Department of Engineering Science, Control Group)

University of Oxford
Life Sciences Interface Doctoral Training Centre

Abstract

This work contributes to theoretically optimal experiment design for the discrimination of two structurally different ordinary differential equation models of the same systems biological phenomenon that fit a set of experimental data equally well. Optimal experiment here means the external stimulus profile which gives the maximal achievable difference between outputs of the two models in L_2-norm, in other words, the two models' predictions of measured values are as different as possible. Such experiments can then be used for model invalidation.

Linear model approximations are studied in the theoretical part. A frequency domain based solution and a time domain based solution are presented. A numerical implementation is also considered in which a more rudimentary experiment design is applied to two models of the slug forming chemotaxis of starving *Dictyostelium* amœbæ.

Although more theoretical work is needed until these results can be put to practical use, the progress is very encouraging, and we expect our methods to lead to substantial savings of time, working hours and expenses in experimental biology.

Contents

1 Introduction

'*During the past 30 years biology has become a discipline for people who want to do science without learning mathematics.*'
Marvin Cassman *et al.* (2005)

Mathematical modelling has become an indispensable tool of modern systems biology. The meaning of *understanding* molecular biological phenomena, protein interactions, metabolic pathways, or the regulation of gene expression is changing: the community is less satisfied with qualitative descriptions, and increasingly searches for quantitative, mathematical descriptions.

Given observed data from experimental measurements, scientists propose hypothetical models that describe the behaviour of a biological system in terms of mathematical equations. (In this work we investigate solely ordinary differential equation models.) Of course a newly proposed model will fit the data at hand, bar small deviations that will have been said to be caused by 'noise': external effects, the inherent discrete stochastic nature of chemical reactions, and inaccuracies of measurements.

Other experiments carried out on the same system will unavoidably lead to new data, and will potentially inspire the formulation of new models. Consequently, myriads of models may be proposed to describe the same biological system.

The emergence of new data will certainly lead to the invalidation of some models if they no longer fit all the data that has been produced. Still, several seemingly equally good, but structurally different models may explain the data. If they were all intended to describe the same phenomenon, then all but at most one of them is incorrect.

One cannot ever *validate* a model. A model can fit all the data we have, it can be tested against its predictions, but still one can never be sure that it is the correct model. For validation some evidence would be needed that the model will fit all possible experimental measurements. This seems impossible: we have no idea how such evidence could be found.

However, one can prove a hypothetical model incorrect by an appropriate experiment (if it is incorrect indeed).

Hence the progress of systems biology follows a regular pattern: model identification (model fitting) is followed by model discrimination and model invalidation. As the proposal of new models may continue indefinitely, these steps close into a circle.

A basic form of model invalidation is the comparison of two competing models in an experiment, when one wishes to reject the incorrect one and (at least temporarily) accept the other one (which seems correct at the time of the experiment).

Traditionally these experiments are designed using heuristic approaches: experience, expert knowledge, intuition, or simple analyses (Feng and Rabitz, 2004). As the development of new measurement techniques and the conducting of experiments themselves are often laborious and costly, it is well worth designing them more rigorously, such that they are most informative, most discriminating.

In this work we describe an approach to defining and designing optimal experiments, that is, experiments that are the best (in some mathematically defined but practically meaningful way) at discriminating between two competing models. The key idea, in a control theoretical terminology, is to find an input profile which maximises the difference between the outputs of the two models. This remark immediately shows that our methods require that the cellular system can be stimulated from outside.

As the reader could have gathered, our approach will be an abstract mathematical one. We will try to contradict the hidden claim of the introductory quote, and prove that mathematics can indeed give a lot of insight into biology. We believe that only mathematical abstraction offers the generality we will be able to develop in this work: our results apply to various settings, not only to one certain biological system as in a case study.

The structure of this work is the following. Section 2 gives a quick review of the related scientific literature. Section 3 gives our definition of optimal experiment. Then in Section 4 we develop approaches to finding the optimal input profile. All our theoretical results are presented in this section. Computer implementation of these results are considered in Section 5 on a case study. Conclusions are drawn in the final section, where directions for further investigations are also given (Section 6).

2 Previous work on optimal experiment design

2.1 Model identification

Model identification is a central topic in experiment design literature. It is concerned with the question of what experiment is most informative about the model structure. We do not consider this question in our work which focuses on model discrimination, thus advise the interested reader to find references in the work by Chen and Asprey (2003).

Kremling *et al.* (2004) identified three different problems in (metabolic or genetic) network identification. *Identifiability* is concerned with the question whether model parameters are uniquely determined by the particular model and an input-output experiment. *Parameter estimation* means the determination of certain parameters such that the difference between measured and predicted output is minimal. *The accuracy of parameters* is usually derived from confidence intervals of the estimated parameters.

The third problem is sometimes also called *practical identifiability* to emphasise that identifiability in itself does not necessarily imply that accurate estimates can be gained from noisy measurements (Gunawan *et al.*, 2006).

The standard method to solve this third problem uses the *Fisher information matrix (FIM)*. The FIM measures the informativeness of noisy measurement data for estimating the model parameters. Its inverse provides a lower bound for the variances of parameter estimates (or equivalently, the upper bound for accuracy) through the Cramer-Rao inequality.

The two most successful FIM-based optimality criteria are called D-optimality and A-optimality. D-optimality aims to maximise the informativeness of data by maximising the determinant of the FIM, which corresponds to the volume of the information hyperellipsoid. A-optimal design aims to reduce the hyperellipsoid of uncertainty in parameter estimates, which is measured by the sum of parameter variances. (Gunawan *et al.*, 2006)

Barrett and Palsson (2006) studied a specific experiment design question: what transcriptional factors one should knock out in an organism such as Escherichia coli to learn the most of the connections in the transcriptional regulatory network (TRN). The increased interest in some automation in experiment design in TRN reconstruction (or generally, biochemical network reconstruction) stems from the recent emergence of metagenomics and environmental sequencing, which provides unprecedented amount of raw data about such, so far unknown networks.

After skimming the surface of the model identification problem, we turn to previous studies looking into the problem of having more than one model. This is the issue we intend to study in the current work.

2.2 Model discrimination

In a relatively early work Bardsley, Wood, and Melikhova (1996) investigated the problem how one should space measurements through time to perform an optimal discriminating experiment. They compared uniform spacing to geometric progression of points in

time, and to uniform spacing on the y-axis (and whatever time distribution this implies).

Kremling *et al.* (2004) presented three methods for optimal discriminating experiment design, and compared them on a test example.

Their first method uses the largest possible change in system input. This rudimentary approach compares the effects of abruptly changing one of several inputs at one time. The optimal experiment is the one in which the change in one input channel leads to the largest difference between output functions.

The second method replaces models with their linearised counterparts. It uses a sinusoid input with a frequency that discriminates the phase shifts of the two models the most, according to their Bode plots. We will consider very similar approaches in Section 5.3.

The third one is based on a method developed by Chen and Asprey (2003). It aims to bring the states of the two models as apart as possible, but in a way that if the measurement error of a state variable is large, then the difference of these states contribute less to the objective function.

They concluded that no single method can be recommended: the best choice strongly depends on the circumstances (possibilities to stimulate the system and to make high quality measurements).

Chen and Asprey (2003) reviewed different approaches to model discrimination before introducing theirs. The Bayesian approach assigns prior probabilities to each model (such that they sum to one), and updates these after each experiment. When one probability becomes sufficiently large compared to others, then the corresponding model is judged to be the best. A frequentist approach and a dynamic discrimination problem are also discussed.

These methods are aimed at inferring the most from noisy measurements. Our investigation follows a slightly different direction. Our models are deterministic, that is, we do not take account of measurement noise directly. Instead, we try to make the outputs of the two models as different as possible to ensure that even a noisy measurement has a good chance of discriminating between them.

3 A control theoretical definition of optimal experiment

3.1 The control theory framework

Systems and control theory is an interesting and challenging application driven field of mathematics, which builds on various other areas, including linear algebra, functional analysis, complex analysis, Fourier integrals and ordinary differential equations (ODEs).

Control theory looks at systems as a triple of functions: the state of the system, the input and the output. All of them are functions of time, and are dependent on one another.

Throughout this work we assume continuous time (as opposed to discrete time), and a fixed starting point (we do not allow the input function to be non-zero before this starting time). We label the starting point with 0. The notation for the three functions are

$$\begin{array}{rcll} u\colon & [0,\infty[& \to & \mathbb{R}^q \quad \text{for the input,} \\ x\colon & [0,\infty[& \to & \mathbb{R}^n \quad \text{for the state, and} \\ y\colon & [0,\infty[& \to & \mathbb{R}^p \quad \text{for the output.} \end{array}$$

For each t we write the values as column vectors.

In our investigations the abstract mathematical notion *system* is the model of the biological system, and the coordinates of the state function are some of the quantitative characteristics of the biological system. The output function represents the measurements the experimentalist would carry out. The stimuli or perturbations the experimentalist can introduce to the system during the experiment (e.g. influxes of nutrients) are modelled by the input function.

The relations of these three functions are typically described by systems of ODEs, which is in correspondence with many biological models.

In our simplest case these ODEs are time invariant linear systems, that is, for some matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times q}$ and $C \in \mathbb{R}^{p\times n}$,

$$\begin{array}{rcl} \dot{x}(t) & = & Ax(t) + Bu(t), \\ y(t) & = & Cx(t). \end{array} \tag{1}$$

For mathematical simplicity we assume that the input does not affect the output directly through the usual $+Du(t)$ term ($y^*(t) = Cx(t) + Du(t)$), because even if this was the case in reality, such a measurement would inform us about y: with known u and D,

$$y(t) = Cx(t) = y^*(t) - Du(t)$$

can easily be determined. The argument in other words is that even if the input influx contains the same chemical agents that are measured as output, we can subtract the known input from the measurement so we can infer the actual system output.

We will need the *2-norm* (or $L_2([0,\infty[)$-*norm*) of functions u and y which is defined as the square root of the square integrals of Euclidean lengths of function values:

$$\begin{aligned} \|u\|_2 &= \left(\int_0^\infty \sum_{i=1}^{q} u_i(t)^2 \, \mathrm{d}t \right)^{\frac{1}{2}}, \\ \|y\|_2 &= \left(\int_0^\infty \sum_{i=1}^{p} y_i(t)^2 \, \mathrm{d}t \right)^{\frac{1}{2}}. \end{aligned}$$

3.2 A definition of optimal experiment

The competing model discrimination problem can be formulated in the following way. We have two system descriptions (the two models), the first one with matrices $A_1 \in \mathbb{R}^{n_1 \times n_1}$, $B_1 \in \mathbb{R}^{n_1 \times q}$ and $C_1 \in \mathbb{R}^{p \times n_1}$, the other one with $A_2 \in \mathbb{R}^{n_2 \times n_2}$, $B_2 \in \mathbb{R}^{n_2 \times q}$ and $C_2 \in \mathbb{R}^{p \times n_2}$:

$$\begin{aligned} \dot{x}_1(t) &= A_1 x_1(t) + B_1 u(t), \qquad x_1(0) = 0, \\ y_1(t) &= C_1 x_1(t), \end{aligned}$$

and

$$\begin{aligned} \dot{x}_2(t) &= A_2 x_2(t) + B_2 u(t), \qquad x_2(0) = 0, \\ y_2(t) &= C_2 x_2(t). \end{aligned}$$

Most importantly, since we have only one experimental setup in reality, the input u is identical in the two systems. Coordinates of the output are the experimental measurements.

We assume that both A_1 and A_2 matrices are *Hurwitz*, that is, all their eigenvalues have negative real part, hence they define asymptotically stable systems.

Note that in biological applications one typically cannot assume linearity; biological systems are usually described by nonlinear ODEs. Keeping this in mind, we will treat the linear model as an entry level to the really relevant nonlinear case.

Optimality of the experiment is defined by the criterion of greatest achievable difference between the outputs y_1 and y_2. More precisely, we assume that the two systems have a common steady state. In the linear case this always hold with the steady state zero. We aim to maximise the difference of y_1 and y_2 in 2-norm ($\|y_1 - y_2\|_2$) over the set of inputs u with some positive, fixed 2-norm. One can simply assume $\|u\|_2 = 1$ by appropriate scaling.

The rationale behind this definition is that we assume measurements can be made periodically during the experiment thus they yield a measurement curve by interpolation.

Our aim will be to derive a method to find such an optimal u input function given the two system descriptions.

Experiments can only be done on a bounded time horizon, so u should ideally have compact support, but we sometimes ease this requirement, and only require that u is 'very small' after a positive time bound.

Our assumption of identical steady states incorporates that we do not intend to observe differences between the two potential descriptions on the long run, after they have reached their respective (in the nonlinear case potentially new, and different) steady states, but during the transient period before their return to the common steady state.

An important observation is that one can fuse the two alternative models to get one system of the form (1), of which the state is the concatenation of x_1 and x_2, the input is the common input, and the output is the difference of the two outputs:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \qquad C = [C_1 \ -C_2]. \tag{2}$$

Trivially, if A_1 and A_2 are Hurwitz matrices, so is A.

4 Theoretical results

We start our investigations with two simpler problems. First, we discuss optimal input for discrete time linear systems. Then we move on to continuous time, but initially will only look for optimal initial state with zero input. The second one is a good starting point from a theoretical viewpoint, however one has to be aware that it is not easily applicable in practice.

4.1 Optimal input for discrete time linear systems

Discrete time description can be used as an approximation to continuous time models. It is in standard textbooks how one can rewrite a continuous time linear system in discrete time. For example, the work by Franklin, Powell, and Emami-Naeini (1994) explains this technique (p 634).

The linear system in discrete time takes the form

$$\begin{array}{rcll} x(k+1) & = & Ax(k)+Bu(k), & x(0)=0, \\ y(k) & = & Cx(k), & \end{array} \tag{3}$$

with the hidden internal structure given by (2). Here the output can be explicitly written down as a function of the input: for any positive integer N,

$$\begin{pmatrix} y(N) \\ y(N-1) \\ \vdots \\ y(2) \\ y(1) \end{pmatrix} = \begin{bmatrix} CB & CAB & \dots & CA^{N-2}B & CA^{N-1}B \\ & CB & & & CA^{N-2}B \\ & & \ddots & & \vdots \\ & & & CB & CAB \\ 0 & & & & CB \end{bmatrix} \begin{pmatrix} u(N-1) \\ u(N-2) \\ \vdots \\ u(1) \\ u(0) \end{pmatrix},$$

or

$$y = Gu$$

for short (here we do not indicate dependency on N). The optimal input problem has to be formulated on a finite time horizon, and it becomes a finite dimensional optimisation problem,

$$\max_u y^T y,$$

$$u^T u = 1,$$

and obviously,

$$y^T y = u^T G^T G u.$$

It will turn out that finding the optimal initial state in continuous time involves a very similar optimisation. It will be shown in the forthcoming section how this optimum can be found, therefore we spare the discussion here.

4.2 Optimal initial state

Now we wish to find the best common initial state for the two systems, that is the one which differentiates the outputs of the two models as much as possible in case of zero input ($u = 0$).

This is rather straightforward using standard theory (see e.g. Dullerud and Paganini (2000)). We use the formulation of (2): for some $x_0 \in \mathbb{R}^n$ initial value

$$\begin{array}{rcll} \dot{x}(t) & = & Ax(t), & x(0) = \binom{x_0}{x_0} \\ y(t) & = & Cx(t). & \end{array} \tag{4}$$

Here we have to assume that the sizes of A_1 and A_2, and C_1 and C_2 are equal.

One defines the *observability gramian* of (C, A) by

$$Y_o := \int_0^\infty e^{A^T\tau} C^T C e^{A\tau} \, \mathrm{d}\tau.$$

It is almost trivial that Y_o is positive semi-definite, and the 2-norm of the output y (the so-called *energy*) is given by

$$\|y\|_2^2 = \binom{x_0}{x_0}^T Y_o \binom{x_0}{x_0}.$$

Hermitian complex (or symmetric real) matrices can be diagonalised with base transformations with unitary (respectively, orthogonal) matrices. The diagonalisation paves the way to the definition of the square root of such matrices.

Now let us consider the *observability ellipsoid*

$$\mathcal{E} := \left\{ Y_o^{\frac{1}{2}} \binom{x_0}{x_0} \;\middle|\; x_0 \in \mathbb{R}^n, \; |x_0| = 1 \right\}.$$

Since Y_o is positive semi-definite, this set is an ellipsoid indeed. Let

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{2n} \geq 0$$

be the eigenvalues of $Y_o^{\frac{1}{2}}$, and

$$v_1, v_2, \ldots, v_{2n}$$

be their corresponding unit-length eigenvectors. Then the v_i give the directions of the principal axes of the ellipsoid, and the μ_i the length of each axis.

Ideally, the eigenvector corresponding to the biggest non-zero eigenvalue (v_1) would give the optimal initial state. In our case we have to restrict the observability ellipsoid $\mathcal{E}$ to the n-dimensional subspace $\{\binom{y}{y} \mid y \in \mathbb{R}^n\}$, which restriction gives an ellipsoid again.

Although we do not give further details, this idea can be followed to derive the solution (Papachristodoulou and El-Samad, 2007).

Finding the optimal initial state is an incomparably easier task than finding the optimal input. This optimisation is done over $\mathbb{R}^n$, while the input problem requires optimisation over an infinite dimensional function space.

4.3 Optimal input for linear systems based on frequency domain investigations

It is now time to turn to our main concern, finding the input which maximises

$$\frac{\|y\|_{L_2([0,\infty[)}}{\|u\|_{L_2([0,\infty[)}},$$

for the system outlined in (2). (Note that the results of this section apply to any system of form (1).)

This is the *system gain* (more specifically, the 2-norm/2-norm system gain) which is defined by

$$\sup\Big\{\|y\|_2 \;\Big|\; \|u\|_2 \le 1\Big\} = \sup\Big\{\|y\|_2 \;\Big|\; \|u\|_2 = 1\Big\}.$$

For this section, where we build on frequency domain characterisation, we assume that systems are single-input/single-output type ($q = 1$, $p = 1$).

From a state space description one goes to the frequency domain by taking the Laplace transform of the state, input and output functions: for a function f in the time domain its *Laplace transform* is

$$\hat{f}(s) = \int_0^\infty f(t)e^{-st}\,\mathrm{d}t.$$

If $x(0) = 0$, for the transformed functions the input-output relationship becomes

$$\hat{y}(s) = \hat{G}(s)\hat{u}(s),$$

where

$$\hat{G}(s) = C(sI - A)^{-1}B$$

is called the *transfer function* of the system. (Doyle, Francis, and Tannenbaum (1990), Dullerud and Paganini (2000))

The ∞-norm of the transfer function $\hat{G}$ is defined by

$$\|\hat{G}\|_\infty := \sup_{\omega\in\mathbb{R}} |\hat{G}(i\omega)|,$$

where $i \in \mathbb{C}$ is the imaginary unit.

The *Bode magnitude plot* is the standard tool to visualise the function $\omega \mapsto |\hat{G}(i\omega)|$. In the most common form it uses decibel units on the y-axis, hence its definition is

$$\omega \mapsto 20\,\log_{10}|\hat{G}(i\omega)| \;=\; 20\,\log_{10}|C(i\omega I - A)^{-1}B|$$

for each $\omega \in \mathbb{R}$.

A related definition is the *Bode phase plot*, which plots the argument of the complex number $\hat{G}(i\omega)$ as a function of $\omega \in \mathbb{R}$. This is the *phase-shift* between input and output.

The *Bode plot* is the pair of the Bode magnitude and phase plots. The x-axis is usually logarithmical in both plots.

The importance of this is that the system gain is the ∞-norm of the transfer function:

$$\sup\left\{\|y\|_2 \;\middle|\; \|u\|_2 = 1\right\} = \|\hat{G}\|_\infty.$$

The proof can be found in Doyle *et al.* (1990).

They also give the Laplace transform of an input function for which $\|y\|_2/\|u\|_2$ arbitrarily approximates the maximum: if $\omega_0 \in \mathbb{R}$ maximises $|\hat{G}(i\omega)|$, that is,

$$\|\hat{G}\|_\infty = |\hat{G}(i\omega_0)|,$$

then for $\omega \in \mathbb{R}$ and an appropriately small $\varepsilon > 0$,

$$\hat{u}(i\omega) := \begin{cases} \sqrt{\pi/2\varepsilon}, & \text{if } |\omega - \omega_0| < \varepsilon \text{ or } |\omega - (-\omega_0)| < \varepsilon, \\ 0, & \text{otherwise,} \end{cases}$$

is the sought input function in the frequency domain.[1]

By taking the inverse, the *Fourier* transform of $\hat{u}$ one can calculate the corresponding time domain form of this almost optimal input: with A normalising constant

$$u(t) = \frac{A}{t}\cos(\omega_0 t)\sin(\varepsilon t). \tag{5}$$

If $\varepsilon \ll \omega_0$, then this formula gives a cosine wave ($\cos(\omega_0 t)$) which is multiplied by a sinc = sin/id function whose main role is to ensure the 2-norm is finite by making the amplitude of the cosine wave tiny outside a zero-centred interval (Figure 1).

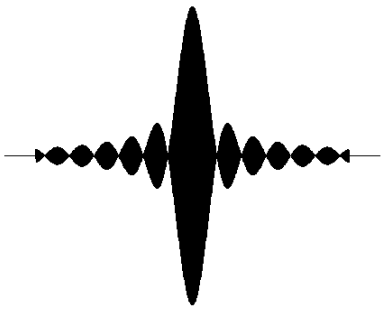

Figure 1: A plot of the almost optimal input u and the x-axis. The solid black areas are filled with the very high frequency cosine wave.

There are obvious problems with this input: 1) the frequency might be too high to be implemented in an experimental setting, 2) negative input values might be meaningless

[1]The real optimal solution would be the mean of two Dirac-deltas at $\pm\omega_0$, but this would give a cosine wave of which the 2-norm is infinity, what we do not allow.

in reality, 3) the support of the input function is the whole real line (the function is even).

Without the genuine biological systems at hand we do not know whether any of the first two problems arises. We cannot tell how dense the cosine wave is. On a practically reasonable time span one may need to finish the input at a time within the support of the central black patch. Even the dense wave might not be felt, but the input could seem to be a roughly constant function (the cosine in a very small neighbourhood of zero).

There are mathematical results which can help with the third problem. Theorem 4.3.1 would ensure that the input is zero in the past, while Theorem 4.3.2 would provide an input with compact support. A version of the first theorem is included in Dullerud and Paganini (2000), but one rather finds these theorems in standard complex analysis textbooks (e.g. Halász (2001)). (The original work is *R. Paley and N. Wiener. Fourier Transforms in the Complex Domain. Amer. Math. Soc., New York, 1934.*)

Theorem 4.3.1. *If Φ is holomorphic in the right open complex half plane, and there exists some $M > 0$ such that for all $\sigma > 0$*

$$\int_{-\infty}^{+\infty} |\Phi(\sigma + it)|^2 \, \mathrm{d}t \leq M,$$

then there exists an $f \in L_2(\mathbb{R})$, for which $f(x) = 0$ if $x < 0$ almost everywhere, and

$$\Phi(s) = \int_0^{\infty} e^{-sx} f(x) \, \mathrm{d}x.$$

Theorem 4.3.2 (Paley–Wiener)**.** *Assume that there exists some $A > 0$ and $\Delta > 0$ such that for a Φ entire function*

$$|\Phi(s)| \leq A e^{\Delta |s|}$$

holds. Then there exists an $f \in L_2(\mathbb{R})$, for which $f(x) = 0$ if $|x| > \Delta$ almost everywhere, and

$$\Phi(s) = \int_{-\Delta}^{\Delta} e^{sx} f(x) \, \mathrm{d}x.$$

These theorems give directions how one should construct a holomorphic input function $\hat{u}$ in the frequency domain which has strong peaks at $\pm i\omega_0$, thus approximates the optimal input well. If $\|\hat{G}\|_\infty < \infty$, then the multiplication of $\hat{u}$ by $\hat{G}$ is a continuous operator in L_2, so inputs that are close in L_2-norm to the optimal input in the frequency domain will be mapped to outputs that are close in L_2-norm to the optimal output in the frequency domain. As the Fourier transform is continuous in L_2 (even isometric, if appropriately normalised), an output which is close in the L_2 sense to the optimal output in the frequency domain will be close to the optimal output in the time domain, too. We did not pursue this direction any further because of the apparent numerical difficulties.

The ∞-norm of the transfer function $\hat{G}$ can be computed by a search procedure based on its definition, or based on the following theorem (Doyle *et al.*, 1990).

Theorem 4.3.3. *For a $\gamma > 0$,*

$$\|\hat{G}\|_\infty < \gamma$$

if and only if

$$H_\gamma := \begin{bmatrix} A & \frac{1}{\gamma^2} BB^T \\ -C^T C & -A^T \end{bmatrix}$$

has no eigenvalues on the imaginary axis.

We implemented the latter search method in MATLAB, but the built-in `norm(G,inf)` routine proved superior: it was faster and numerically more stable than our script.

In the next section the KYP Lemma will give a third alternative way of computing $\|\hat{G}\|_\infty$.

4.4 A bridge between frequency and time domain characterisations

A *linear matrix inequality (LMI)* in the variable X is an inequality of the form $F(X) < Q$ (in the usual definiteness sense), where X is a real matrix, F is a linear (or affine) mapping to the set of symmetric matrices (or to the real vector space of complex Hermitian matrices), and Q is a symmetric or Hermitian matrix.

Their importance stems from two facts: many problems can be reduced to LMIs, and they can be solved efficiently with semidefinite programming.

The seminal *KYP Lemma* made a link between system descriptions in the frequency and time domains. The following version is proved in e.g. Dullerud and Paganini (2000). The proof uses the deepest relationships of linear systems theory: the dissipation of storage functions, Hamiltonian matrices and Riccati equations.

Lemma 4.4.1 (Kalman–Yakubovich–Popov)**.** *If $\hat{G}(s) = C(sI - A)^{-1}B$, then the following are equivalent conditions.*

(I) The matrix A is Hurwitz and

$$\|\hat{G}\|_\infty < \gamma.$$

(II) There exists a matrix $X > 0$ such that

$$\begin{bmatrix} A^T X + XA + C^T C & \frac{1}{\gamma} XB \\ \frac{1}{\gamma} B^T X & -I \end{bmatrix} < 0.$$

We mention that the applications of this fundamental result go far beyond the computation of $\|\hat{G}\|_\infty$.

In the next section we reformulate the optimal discriminating experiment problem as an optimal control problem in order to use this developed theory.

4.5 Optimal experiment as an optimal control problem

As we have just suggested we now formulate our optimal experiment problem in the well-known LQ framework (*l*inear state equation, *q*uadratic cost functional). With x, A, B and C given in the usual way by our competing models in (2), the problem is

$$\max_u \int_0^\infty y^T(t)y(t)\,\mathrm{d}t,$$

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t), \\ x(0) &= 0, \end{aligned} \tag{6}$$

$$\int_0^\infty u^T(t)u(t)\,\mathrm{d}t = 1.$$

Our cost functional to be *minimised* is

$$J := \frac{1}{2}\int_0^\infty -y^T(t)y(t)\,\mathrm{d}t$$

in the standard formulation (see e.g. Bryson and Ho (1975)). We impose no penalty on the limiting state $\lim_\infty x$ (or indeed on any of the states).

By adjoining the system ODEs and the constraint on the input to J, it becomes

$$\begin{aligned} \bar{J} &= \int_0^\infty -\frac{1}{2}y^T(t)y(t) + \lambda^T(t)\Big(Ax(t) + Bu(t) - \dot{x}(t)\Big)\,\mathrm{d}t + \frac{\lambda_0}{2}\left(\int_0^\infty u^T(t)u(t)\,\mathrm{d}t - 1\right) \\ &= \int_0^\infty -\frac{1}{2}y^T(t)y(t) + \lambda^T(t)\Big(Ax(t) + Bu(t) - \dot{x}(t)\Big) + \frac{\lambda_0}{2}u^T(t)u(t)\,\mathrm{d}t - \frac{\lambda_0}{2}, \end{aligned}$$

with some $\lambda : [0,\infty[\to \mathbb{R}^n$ continuously differentiable function and $\lambda_0 > 0$ scalar. Now the corresponding *Hamiltonian* is

$$H(x(t), u(t), \lambda(t), \lambda_0, t) = -\frac{1}{2}y^T(t)y(t) + \lambda^T(t)\Big(Ax(t) + Bu(t)\Big) + \frac{\lambda_0}{2}u^T(t)u(t).$$

For optimality the *Euler-Lagrange equations* must hold:

$$\begin{aligned} \dot{\lambda}^T &= -\partial_x H, \\ \lim_\infty \lambda &= 0, \\ \partial_u H &= 0, \end{aligned}$$

which in our case means

$$\begin{aligned} \dot{\lambda} &= C^T Cx - A^T\lambda, \\ \lim_\infty \lambda &= 0, \\ B^T\lambda + \lambda_0 u &= 0. \end{aligned}$$

From the last one we get the optimal input as a function of λ:

$$u = -\frac{1}{\lambda_0} B^T \lambda.$$

The constraint $\|u\|_2 = 1$ yields the value of λ_0 once we have λ:

$$\lambda_0{}^2 = \int_0^\infty \lambda^T(t) B B^T \lambda(t)\,\mathrm{d}t.$$

At the end we have this system of ODEs to solve:

$$\begin{pmatrix} \dot{x} \\ \dot{\lambda} \end{pmatrix} = \begin{bmatrix} A & -\frac{1}{\lambda_0} B B^T \\ C^T C & -A^T \end{bmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix}, \qquad \begin{aligned} x(0) &= 0, \\ \lim_\infty \lambda &= 0. \end{aligned} \tag{7}$$

We could have derived the same following a slightly different way. One may say they want to maximise $\|y\|_2$ and minimise $\|u\|_2$ simultaneously: with an appropriate $\lambda_0 > 0$ scaling factor let J be

$$J := \frac{1}{2} \int_0^\infty -y^T(t) y(t) + \lambda_0 u^T(t) u(t)\,\mathrm{d}t.$$

Then

$$\bar{J} = \int_0^\infty -\frac{1}{2} y^T(t) y(t) + \frac{\lambda_0}{2} u^T(t) u(t) + \lambda^T(t) \Big(A x(t) + B u(t) - \dot{x}(t) \Big)\,\mathrm{d}t,$$

which is the same $\bar{J}$ what we had previously, apart from the missing constant term at the end.

The system (7) is not easy to solve even numerically, because it has boundary conditions at the initial point for half of the coordinates, and at the end for the other half. Moreover, we do not know the value of λ_0 in the system matrix beforehand.

One can use a trial and error (so-called *shooting*) algorithm with some $\lambda(0)$ values, which are then iteratively updated according to the resulting $\lim_\infty \lambda$. (We shoot from $\lambda(0)$, and adjust after each shot until we manage to hit $\lim_\infty \lambda = 0$.) Unfortunately the value of λ_0 is unknown in (7) beforehand. (We may have merely missed to note a detail which would tell us the value.) Therefore a further search among different λ_0 values is also needed in the solver algorithm.

We remark that the similarity between the matrices H_γ of Theorem 4.3.3 and the one in (7) is apparent.

If $\gamma^2 = \lambda_0$, then the eigenvalues of the two matrices are identical. Simply consider that under this assumption a change of basis (which conserves the spectrum) with transformation matrix

$$\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$$

transforms one matrix into the other.

In addition to its applicability in the linear case, the optimisation framework has further strengths.

First, this is the obvious way to follow in the nonlinear case. No doubt the algebra and the algorithmic implementation will become more difficult compared to the linear case, but there is no theoretical obstacle to progression.

Second, the optimisation framework gives means to incorporate additional limitations, such as constraints on the input function, into the experiment design. For instance, we understand that demanding a sinusoid input may be unrealistic in a laboratory setting, and experimentalists may want to find the optimal input from a smaller input function space. If such constraints can be expressed in an appropriate mathematical formulation, then they can be added to the optimisation criteria.

5 Applications

To illustrate our methods with applications, two biological systems were considered.

5.1 A biochemical engineering application

The first one is from Kremling *et al.* (2004), and is an invented bioreactor where some organism grows in a well-mixed chemostat.

This benchmark problem seemed very attractive at the beginning, partly because its authors developed an online simulation environment for it, where users can specify a piecewise constant input profile and see how good it is at discriminating between the two proposed models.

However, as it turned out — unfortunately relatively late — this pair of models has serious drawbacks as an example. The most prominent is that the two models have different steady states, so they violate one of our starting assumptions. This does not mean that we cannot differentiate between them. On the contrary, it is easier than what we assumed in our theoretical work.

Although this system will not be cited in the section on numerical simulations, we introduce it in order to give one more example of the type of systems we are interested in.

Inputs of this system are flow rates q_{in} and q_{out}, and feed concentration c_{in}. Typically the volume is held constant, so $q_{\text{in}} = q_{\text{out}}$.

State variables are concentrations of biomass B, substrate S and intracellular concentrations of metabolites M_1, M_2 and M_3. Metabolite 1 is the first substance synthesised after uptake, with a Michaelis–Menten mechanism. Metabolite 3 acts as an enzyme (hence we will call it E instead of M_3), and converts Metabolite 1 to Metabolite 2 irreversibly. Metabolite 2 degrades under the Michaelis–Menten reaction law. It is assumed that the flux from Metabolite 2 is responsible for the entire biomass.

Both proposed models of the biochemical network include dilution of intracellular components by growth rate $\mu = Y_{\text{XS}}\, r_{1\text{max}}\, S \,/\, (K_S + S)$.

Under Model A the conversion of Metabolite 1 to 2 is given by noncompetitive inhibition of the enzyme by Metabolite 2. Enzyme synthesis is supposed to proceed with a constant velocity.

$$
\begin{aligned}
\dot{B} &= \mu B - \frac{q_{\text{in}}}{V} B, \\
\dot{S} &= -r_{1\text{max}} \frac{S}{K_S + S}\, m_w\, B + q_{\text{in}} c_{\text{in}} - q_{\text{in}} S, \\
\dot{M}_1 &= r_{1\text{max}} \frac{S}{K_S + S} - k_{2\text{A}} E\, \frac{M_1}{K_{M_1} + M_1}\, \frac{K_{\text{IA}}}{K_{\text{IA}} + M_2} - \mu\, M_1, \\
\dot{M}_2 &= k_{2\text{A}} E\, \frac{M_1}{K_{M_1} + M_1}\, \frac{K_{\text{IA}}}{K_{\text{IA}} + M_2} - r_{3\text{max}}\, \frac{M_2}{K_{M_2} + M_2} - \mu\, M_2, \\
\dot{E} &= k_{\text{synmaxA}} - \mu\, E.
\end{aligned}
$$

In Model B the enzyme behaviour is different. Here the enzymatic conversion of Metabolite 1 follows a Michaelis–Menten kinetic rate law. For the enzyme synthesis, a formal kinetic rate law representing an inhibition is used. The equations for $\dot{B}$ and $\dot{S}$ are the same, but the other three change:

$$\begin{aligned}
\dot{M}_1 &= r_{1\max}\frac{S}{K_S+S} - k_{2\mathrm{B}}E\,\frac{M_1}{K_{M_1}+M_1} - \mu\,M_1,\\
\dot{M}_2 &= k_{2\mathrm{A}}E\,\frac{M_1}{K_{M_1}+M_1} - r_{3\max}\,\frac{M_2}{K_{M_2}+M_2} - \mu\,M_2,\\
\dot{E} &= k_{\mathrm{synmaxB}}\frac{K_{\mathrm{IB}}}{K_{\mathrm{IB}}+M_2} - \mu\,E.
\end{aligned}$$

We do not detail any further what the different parameters mean, because with this example we only intended to give a flavour of a pair of competing biochemical ODE models.

As we have already mentioned, in this example model discrimination can be based on the difference of steady states. If one measures a state of which the values in the two models differ at equilibrium, then without any stimulation (that is, with zero input) arbitrarily big difference of outputs in L_2-norm is achievable on a finite time horizon. So even if the steady states are close to each other, with fine enough measuring techniques and multiple sampling one can differentiate between the models.

5.2 Chemotaxis of starving *Dictyostelium discoideum*

The example of Papachristodoulou and El-Samad (2007) is based on the chemotactic social amœba *Dictyostelium discoideum*. Under starvation these amœbæ secrete cAMP, thus attract other Dictyostelium amœbæ to aggregate and form a multicellular slug, then a fruiting body. This then produces spores, that is, inactive cells, capable of starting their lives when food is abundant again.

The authors compared two models of the adaptation mechanism observed when amœbæ encounter a step function input of chemoattractant cAMP. Both proposed models have three variables. (See Figure 2.)

The response regulator is either in active (R^*) or in inactive (R) state. The sum of these two is a constant: $R_T = R^*(t) + R(t)$. The regulator is activated through the action of enzyme A, and inactivated through the action of enzyme I. In both models $\dot{R}^*$ is given by the equation

$$\begin{aligned}
\dot{R}^* &= -k_{-r}\,I\,R^* + k_r\,A\,R\\
&= -\big(k_{-r}\,I + k_r\,A\big)R^* + k_r\,A\,R_T,
\end{aligned}$$

with forward and backward rate constants k_r and k_{-r}. (The parameters are given in Table 1.)

In Model A both enzymes are regulated by an external signal S (the input, starting from the reference value $S(0) = S_0$), which is proportional to cAMP concentration. With

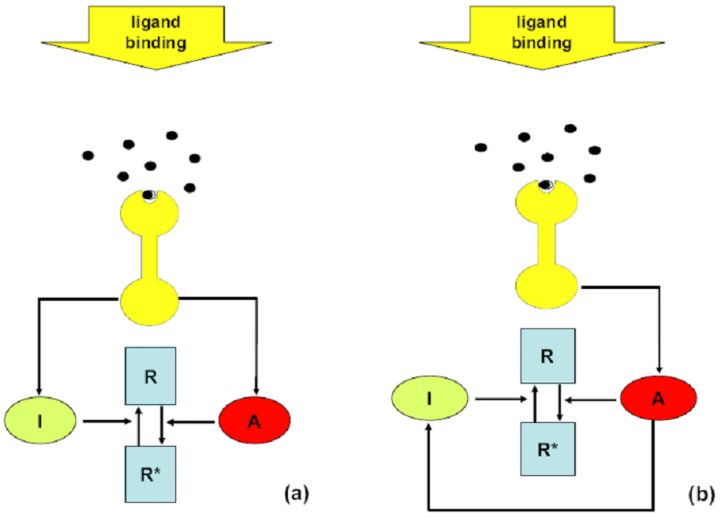

Figure 2: Two models of the chemotaxis system of Dictyostelium amœbæ from Papachristodoulou and El-Samad (2007).

$k_r = 1$	$k_{-r} = 1$
$k_a = 3$	$k_{-a} = 2$
$k_{i_1} = 1$	$k_{-i} = 0.1$
$k_{i_2} = 2/3$	
$S_0 = 0.2$	$R_T = 0.7667$

Table 1: Parameter values of the two models in Papachristodoulou and El-Samad (2007).

rate constants k_a, k_{-a}, k_{-i} and k_{i_1},

$$\begin{aligned} \dot{A} &= -k_{-a}\, A + k_a\, S, \\ \dot{I} &= -k_{-i}\, I + k_{i_1}\, S. \end{aligned}$$

In Model B the inhibitory molecule I is activated through the indirect action of activator A instead of direct activation by ligand binding: with some rate constant k_{i_2},

$$\begin{aligned} \dot{A} &= -k_{-a}\, A + k_a\, S, \\ \dot{I} &= -k_{-i}\, I + k_{i_2}\, A. \end{aligned}$$

A step input of chemoattractant triggers a transient response, after which the chemosensory network described by either model returns to its pre-stimulus values (to its steady state), such as was observed in experiments with Dictyostelium.

In this example both models have the same steady state.

5.3 Numerical simulations

Two different numerical solutions were developed and implemented in the programming environment MATLAB using the SIMULINK package.

As all our theoretical work was concerned with linear systems, both numerical methods will approximate nonlinear systems with their numerical linearisations around their steady states, which are chosen as initial states.

The main problem with linearisation is that its accuracy degrades as the system evolves and departs from the state around which the linearisation was done. Two potential solutions can be easily proposed.

First, to keep the system close to its steady state by applying low energy input. Although this will result in low energy output, the gain (the ratio of the two energies) will not necessarily be much smaller than the nonlinear system gain.

Second, as the two systems evolve, if they are too far from their initial states, then to linearise them again around the points where they are. And to do this repeatedly.

Our most refined solution was derived in Section 4.5, namely the ODE given in (7). As it was pointed out there, this ODE can be solved by a shooting method, but implementing it would have probably taken too long, and was not sensible under the given time constraints.

We rewrote the systems of Papachristodoulou and El-Samad (2007) to single-input/ single-output form to adhere to the format which was assumed during our earlier investigations. This only means the assumption that only one input function is allowed to be changed, the others (if there are any) have to be kept constant. The other change is that only one state variable could be measured (and it could be measured directly).

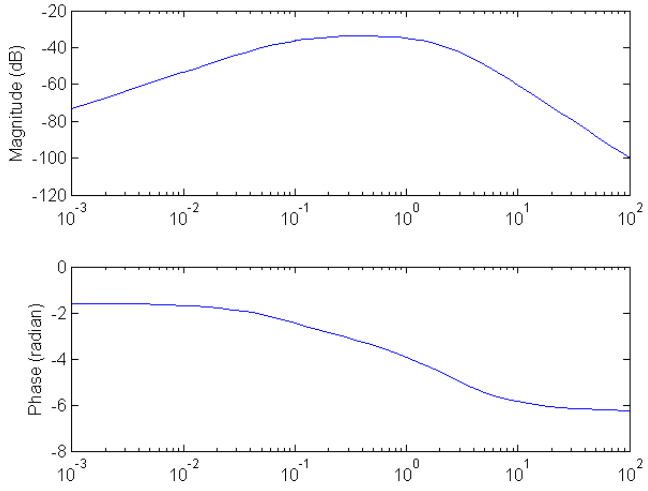

Figure 3: Bode plot of the linearised difference system of the two models with output R^* from Papachristodoulou and El-Samad (2007). The shape of the Bode plot for output I is very similar, but magnitudes are bigger.

The first program which we implemented applies an input which can be chosen from a number of options, such as constant, cosine, cosine with exponential decay, sinc, or the function given by Equation (5) of Section 4.3. Where needed, the frequency is determined by finding the frequency corresponding to the maximum amplification using the Bode magnitude plot of the linearised difference system of the two models (Figure 3). This input is fed to the systems between time zero and a user-defined future time, then zero input is applied until the end of the simulation to allow the two systems to return to steady state. The amplitude is defined only after the shape and time span of the input has been fixed, to give $\|u\|_2 = 1$ (or some fixed small value).[2] This program linearises the nonlinear difference system to find the optimal frequency only once at the beginning.

Table 2 shows gains for different input functions, input energies and output variables.

$\|u\|_2 = 1$	R^*	I
Cosine or (5)	0.0216	0.473
Sinc	0.0191	0.451
Cosine w. exp. decay	0.0214	0.452
Constant	0.0070	0.197
$\|u\|_2 = 0.01$	R^*	I
Cosine or (5)	0.0208	0.484
Sinc	0.0195	0.457
Cosine w. exp. decay	0.0193	0.456
Constant	0.0085	0.198

Table 2: Numerical estimates of $\max_u \|y_1 - y_2\|_2/\|u\|_2$ with different input profiles for the models of Papachristodoulou and El-Samad (2007). Here maximisation is over different frequencies, for inputs where frequency makes sense. The first table shows values for $\|u\|_2 = 1$, the second for $\|u\|_2 = 0.01$. In the first column the output variable is R^*, while in the second column it is I. (A values are always equal for the two models.) Input is fed between 0 and 60 time units, the whole simulation time spans 180 time units. At the steady state $\|\hat{G}\|_\infty = 0.0204$ for output R^*, and $\|\hat{G}\|_\infty = 0.476$ for output I. ($\hat{G}$ is the transfer function of the linearised difference system.)

One can notice that the cosine input gives the same gains as the function (5) of Section 4.3. This is easily explained by the fact that they were almost the same: with very small ε one can hardly distinguish (5) from cosine in the [0, 60] interval, because in this case the interval lies well within the central black patch of Figure 1.

The table clearly shows that the cosine (or (5)) input is far superior to a constant input, and is somewhat better than the sinc or the exponentially decaying cosine input.

One can also see that the linear system gain can be closely approximated by $\|y_1 - y_2\|_2/\|u\|_2$, and sometimes it can even be surpassed as an effect of steering the systems to a region where the linear approximation is no longer accurate enough.

Figure 4 compares differences between state variables with output R^*. (Note that discriminating between the two models based on output I is much easier.) In the first

[2]The input L_2-norm $\|u\|_2$ is defined as the L_2-distance from a constant input with the reference value S_0.

plot the input between 0 and 60 time units is a constant function, in the second one it is cosine. The cosine input yields a visibly bigger difference between R^* values of the two models. One has to note that one measurement may not be enough to the discrimination, but a series of measurements is needed. This is in correspondence with the founding of our definition of optimality on L_2-distance.

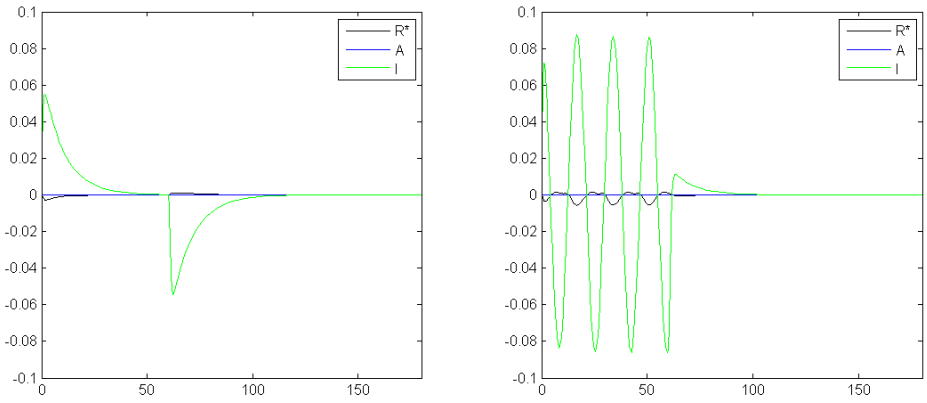

Figure 4: Differences between state variables of the two competing models applying constant input (left) or cosine input (right) until time 60 (states of Model A minus states of Model B). Here the output variable (on which the choice of frequency depends) is R^*, and $\|u\|_2 = 1$.

In the introductory Section 2.2 we have already mentioned the work of Kremling *et al.* (2004), who (among other methods) used the algorithm that we call here first program. Both our and their methods use the Bode plots. The only difference between the two is that when they maximise difference between phase shifts, we maximise difference between gains. We think that our choice is more sensible, because measuring differences between gains is easier than between phase shifts.

The second program has all the features of the first one except that there is no choice of input profile. This second program linearises the systems around where they are sequentially, that is, every time they leave a user-specified spherical neighbourhood of the point around which the last linearisation happened.

After each linearisation the optimal frequency is determined again by the maximum amplification according to the Bode magnitude plot.

In the implemented code a cosine wave input is fed starting from phase zero, with the amplitude that gives unit energy input if this input is kept until the end and no further linearisations happen. When there is a new linearisation, a new cosine is started with phase zero and with the amplitude that will give unit total input energy if no new linearisations happen.

This method results in a broken input profile (Figure 5), where jumps often occur more frequently than what the wavelength is. These inputs give far inferior gains than the first program, which linearises only at the beginning (Table 3).

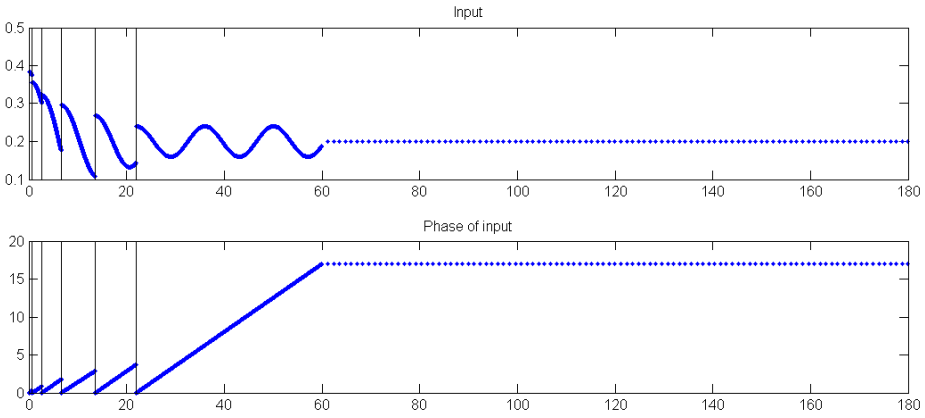

Figure 5: Input profile generated by the second program, and the respective phase of the cosine. Vertical lines mark linearisations, where the input function is modified. Here $\|u\|_2 = 1$, the output variable is I, and square distance threshold for new linearisation between current state and state around which linearisation was done is 0.1.

As the remaining input energy is consumed, the remaining time is also decreasing, but still the amplitude of the input decreases at each linearisation. (This happens because the cosine wave starts from its maximal value, hence uses more energy per time unit in the beginning than what the average is for the whole time interval in which the stimulus is non-zero.)

An enhanced version of the second program saves the previous phase of the input at each linearisation, and starts the new input with this phase, changing the frequency and amplitude only (Figure 6). Table 3 shows that this method gives better, but still dissatisfactory gains.

$\|u\|_2 = 1$	R^*	I
Remembers phase of input	0.0176	0.457
Always starts from phase 0	0.0159	0.410
$\|u\|_2 = 0.01$	R^*	I
Remembers phase of input	0.0178	0.421
Always starts from phase 0	0.0153	0.381

Table 3: Numerical estimates of $\|y_1 - y_2\|_2/\|u\|_2$ achieved by the second program for the models of Papachristodoulou and El-Samad (2007). The first table shows values for $\|u\|_2 = 1$, the second for $\|u\|_2 = 0.01$. The criterion for new linearisation was that the squared distance between (R^*, A, I) and the point around which the last linearisation happened reached 0.1 for the first table, and 0.00001 for the second one. In the first column the output variable is R^*, while in the second column it is I. (A values are always equal for the two models.) Input is fed between 0 and 60 time units, the whole simulation time spans 180 time units. At the steady state $\|\hat{G}\|_\infty = 0.0204$ for output R^*, and $\|\hat{G}\|_\infty = 0.476$ for output I. ($\hat{G}$ is the transfer function of the linearised difference system.)

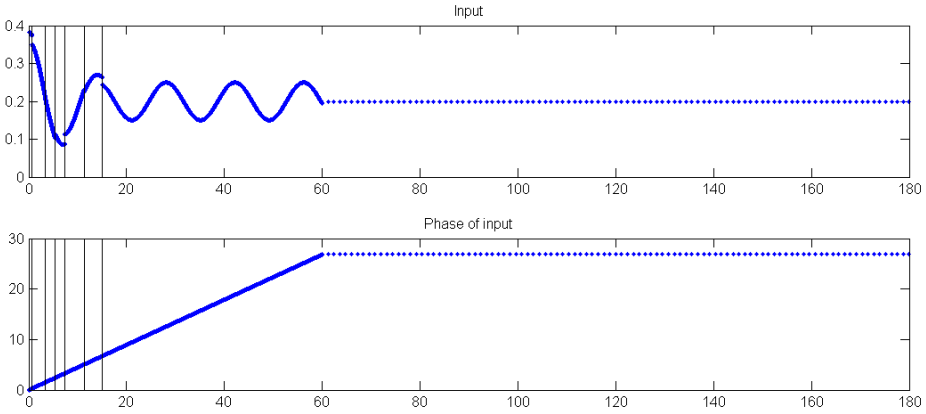

Figure 6: Input profile and the respective phase of the cosine generated by the enhanced version of the second program, which saves and reloads pre-linearisation phases. Vertical lines mark linearisations where the input function is modified. Here $\|u\|_2 = 1$, the output variable is I, and square distance threshold for new linearisation between current state and state around which linearisation was done is 0.1.

It seems probable that low gain is a result of the loss of periodicity of the cosine during the many new linearisations. The jumps of the discontinuous, broken input profile may have an effect which acts contrarily to that of the wave, and thus practically annihilate its beneficial effect.

The second program could be improved upon in at least two ways.

The less important modification would be that linearisations should happen every time when the actual linearisation of the system is too far from its last fixed linearised version, instead of when the state is too far from the state around which the last linearisation happened. So, with temporary but self-explanatory notation, instead of having a relinearisation criterion for $|x(t) - x_{\text{last lin.}}|$, one should have a criterion for $\|\hat{G}_t - \hat{G}_{\text{last lin.}}\|_\infty$.

A more important improvement is expected from keeping the amplitude (and also the phase) of the cosine input at new linearisations, and only changing its frequency. This would give a continuous input profile without the problematic jumps. Here the stopping of the stimulation would be triggered by the input energy reaching the pre-specified value.

6 Discussion and directions for future research

As we have discussed, the currently available version of the second program gives results which are inferior to those of the simpler first program.

If low energy inputs are allowed, then a single linearisation may well be justified. This gives way to the application of inputs based on frequency domain investigations (Section 4.3) or on the most refined and sophisticated solution we have developed, the ODE (7) of Section 4.5. Some pondering is needed how the search for the parameter λ_0 could be included in the shooting method.

Sequential linearisation along the trajectory might lead to a good input profile if the proposed modifications to the second program prove useful.

6.1 The nonlinear case

The treatment of the nonlinear case would be the ultimate task, but this is substantially more difficult than the linear case. Here we formulate the problem without studying it. Given more time, the optimal control formulation of Section 4.5 would be the way to start its investigation.

Let us assume that each system is of the form

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + g(x(t))\,u(t), \\ y(t) &= h(x(t)), \end{aligned} \tag{8}$$

with functions f and g that ensure the existence of a unique solution. (It is enough that f is locally Lipschitz, g and h are continuous.) Note that the system is nonlinear in x, but affine dependency on the input u is assumed.

As one can expect, we define *optimal* experiment as an input u with $\|u\|_2 \le 1$ which meets the requirement that for the two models of the form (8), $\|y_1 - y_2\|_2$ is maximal.

The problem can be formulated in more detail as follows. Given locally Lipschitz functions f_1 and f_2, continuous functions g_1, g_2, h_1 and h_2, and a common steady state $x_0 \in \mathbb{R}^n$, we are looking for the input function u that solves the following optimisation problem:

$$\max_u \int_0^\infty \Big(y_1(t) - y_2(t)\Big)^T \Big(y_1(t) - y_2(t)\Big)\,\mathrm{d}t,$$

$$\int_0^\infty u(t)^T u(t)\,\mathrm{d}t \le 1,$$

$$\begin{aligned} \dot{x}_1(t) &= f_1(x_1(t)) + g_1(x_1(t))\,u(t), \\ y_1(t) &= h_1(x_1(t)), \\ \dot{x}_2(t) &= f_2(x_2(t)) + g_2(x_2(t))\,u(t), \\ y_2(t) &= h_2(x_2(t)), \\ x_1(0) &= x_2(0) = x_0. \end{aligned}$$

6.2 A different problem formulation with the Hankel operator

So far we have been looking at and been stimulating systems from time zero, applying input and measuring output simultaneously. In fact, one could consider a different approach, in which input is applied before a fixed time, say 0, no input is applied afterwards, and output is taken only once this fixed time has passed.

In addition to the state space description one can use the input-output operator $G\colon\ L_2(]-\infty,\infty[)\ \to\ L_2(]-\infty,\infty[)$, for which $y = Gu$.

The *Hankel operator* $\Gamma_G\colon\ L_2(]-\infty,0])\ \to\ L_2([0,\infty[)$ of G is defined by

$$\Gamma_G = P_+ G|_{L_2(]-\infty,0])}\,,$$

where $P_+\colon\ L_2(]-\infty,\infty[)\ \to\ L_2([0,\infty[)$ is the projection by truncation. So the Hankel operator connects input

$$u\colon\ ]-\infty,\ 0]\ \to\ \mathbb{R}^q$$

with output

$$y\colon\ [0,\ \infty[\ \to\ \mathbb{R}^p.$$

Equivalently, $\Gamma_G = \Psi_o \Psi_c$, where

- Ψ_c is the *controllability operator*, which maps u to $x(0)$,
- Ψ_o is the *observability operator*, which maps $x(0)$ to y with no input after time 0.

Indeed, as the support of input u is in the past, it only affects future output y through the state at time 0, $x(0)$.

The introductory results of this theory are about the norm of the Hankel operator, which marks a good starting point for potential investigations of optimal experiment design (Dullerud and Paganini, 2000).

6.3 Conclusion

After setting a definition of optimal experiment, this piece of work engaged with the design problem from multiple directions. The assumptions, for instance on the dimension of input and output, were sometimes changed in order to enable the derivation of strongest results.

During the project a new and promising idea formed, the Hankel formulation, which we could not track down because of the lack of time. This would be an interesting direction for future research.

The optimal control approach applied to the nonlinear case is worth just as much consideration. Its flexibility in incorporating constraints makes it a very promising tool in experiment design.

The mathematical apparatus of our designs may be challenging to biologists, so often collaborations will become necessary between experimentalists and people with mathematical sciences background to implement these and future experiment design techniques.

If we turn to the motivating practical biological problems once again, we can see that our results are not mature yet, but very importantly, applications are not beyond reach.

Acknowledgements

The author is thankful to Antonis Papachristodoulou, who helped his work with many ideas, and whose literacy of the subject proved crucial to a better understanding, and helped collate an informative list of references. This work could not have been carried out without financial support from the EPSRC through the Life Sciences Interface Doctoral Training Centre, University of Oxford.

References

W. G. Bardsley, R. M. W. Wood, and E. M. Melikhova. Optimal design: a computer program to study the best possible spacing of design points. *Computers Chem.*, 20 (1996) 145–157.

C. L. Barrett and B. O. Palsson. Iterative Reconstruction of Transcriptional Regulatory Networks: An Algorithmic Approach. *PLoS Computational Biology*, 2 (2006) 429–438.

A. E. Bryson and Y.-C. Ho. Applied Optimal Control. Optimization, Estimation and Control. Taylor & Francis, 1975.

M. Cassman, A. Arkin, F. Katagiri, D. Lauffenburger, F. J. Doyle III, and C. L. Stokes. Barriers to progress in systems biology. *Nature*, 438 (2005) 1079.

B. H. Chen and S. P. Asprey. On the Design of Optimally Informative Dynamic Experiments for Model Discrimination in Multiresponse Nonlinear Situations. *Ind. Eng. Chem. Res.*, 42 (2003) 1379–1390.

J. Doyle, B. Francis, and A. Tannenbaum. Feedback Control Theory, chapter 2. Macmillan Publishing Co., 1990. 11–26.

G. E. Dullerud and F. G. Paganini. A Course in Robust Control Theory, A convex approach. Springer, 2000.

X.-j. Feng and H. Rabitz. Optimal Identification of Biochemical Reaction Networks. *Biophysical Journal*, 86 (2004) 1270–1281.

G. F. Franklin, J. D. Powell, and A. Emami-Naeini. Feedback control of dynamic systems. Addison-Wesley, third edition, 1994.

R. Gunawan, K. G. Gadkar, and F. J. Doyle III. System modelling in cellular biology: from concepts to nuts and bolts, chapter 11 in Methods to Identify Cellular Architecture and Dynamics from Experimental Data. MIT Press, 2006. 221–242.

G. Halász. Fourier integrál. Komplex függvénytani füzetek I. Eötvös Kiadó, 2. edition, 2001. In Hungarian.

A. Kremling, S. Fischer, K. Gadkar, F. J. Doyle, T. Sauter, E. Bullinger, F. Allgöwer, and E. D. Gilles. A Benchmark for Methods in Reverse Engineering and Model Discrimination: Problem Formulation and Solutions. *Genome Res.*, 14 (2004) 1773–1785.

A. Papachristodoulou and H. El-Samad. Algorithms for Discriminating Between Biochemical Reaction Network Models: Towards Systematic Experimental Design. *Proceedings of the 2007 American Control Conference*, (2007) 2714–2719.