

Research

Open Access

Inverse bifurcation analysis: application to simple gene systems

James Lu^{*1}, Heinz W Engl¹ and Peter Schuster²

Address: ¹Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040 Linz, Austria and ²Theoretical Biochemistry Group, Institute for Theoretical Chemistry, University of Vienna, Währingerstrasse 17, A-1090 Vienna, Austria

Email: James Lu^{*} - james.lu@oeaw.ac.at; Heinz W Engl - heinz.engl@oeaw.ac.at; Peter Schuster - pks@tbi.univie.ac.at

^{*} Corresponding author

Published: 21 July 2006

Received: 21 June 2006

Algorithms for Molecular Biology 2006, 1:11 doi:10.1186/1748-7188-1-11

Accepted: 21 July 2006

This article is available from: <http://www.almob.org/content/1/1/11>

© 2006 Lu et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Bifurcation analysis has proven to be a powerful method for understanding the qualitative behavior of gene regulatory networks. In addition to the more traditional *forward problem* of determining the mapping from parameter space to the space of model behavior, the *inverse problem* of determining model parameters to result in certain desired properties of the bifurcation diagram provides an attractive methodology for addressing important biological problems. These include understanding how the robustness of qualitative behavior arises from system design as well as providing a way to engineer biological networks with qualitative properties.

Results: We demonstrate that certain inverse bifurcation problems of biological interest may be cast as optimization problems involving minimal distances of reference parameter sets to bifurcation manifolds. This formulation allows for an iterative solution procedure based on performing a sequence of eigen-system computations and one-parameter continuations of solutions, the latter being a standard capability in existing numerical bifurcation software. As applications of the proposed method, we show that the problem of maximizing regions of a given qualitative behavior as well as the reverse engineering of bistable gene switches can be modelled and efficiently solved.

1 Background

The use of mathematical models provides tools for the analysis of complex molecular interactions aiming at an understanding of processes occurring in living cells. For many problems in cellular control, stochastic effects and time-delays can be ignored and systems of first-order ordinary differential equations (ODEs) can adequately model the underlying processes. Denoting by x and p the biochemical concentrations and parameters, respectively, the instantaneous change in x is described by the vector field f :

$$\dot{x} = f(x, p). \quad (1)$$

In the study of such systems, an important goal is to understand how the observed physiological behavior arises out of gene network topology and parameters p . Some of these questions may be studied via examining the *bifurcation manifolds* Σ of the ODE system, which partition the parameter space into regions of different qualitative behavior (see e.g., [1] for a general overview to bifurcation theory). From ODE models and measured parameters, the *forward problem* of computing the bifurcation diagram has contributed significantly towards elucidating the complex mechanisms underlying cellular processes. For instance, mathematical and symbolic bifurcation analysis has led to an understanding of the possible dynamical behaviors

that may arise out of simple gene systems (for a monograph, see [2], examples of more recent papers dealing with natural, designed, and model systems are [3-7]). For cell cycle models, bifurcation diagrams have given biologists a systems-level perspective of the roles played by the various constituent modules, as well as providing the ability to predict the behavior of mutant cells [8,9]. The desire to locate regions in parameter space exhibiting interesting dynamics has led to the development of computational tools for the automatic discovery of bifurcation sets [10].

In contrast, *inverse problems* have only recently attracted attention in biology as a way to unravel the workings of cellular mechanisms. In inverse problems one looks for causes for observed or desired effects [11]. Mathematically, such problems are typically ill-posed, in particular unstable; special mathematical techniques, called "regularization methods", have to be used to cope with this ill-posedness. Many variational and iterative regularization techniques have been developed over the years and applied to a variety of problems in science, engineering and finance (see [11] and some references quoted there, and [12]).

In the current context of cell biology, one would like to address problems such as: which parameter configurations lead to an observed qualitative behaviour of the system ("identification")? How can one introduce a certain qualitative behaviour into the system via parameter variations ("design")? We summarize such problems under the name of *inverse bifurcation problems*, where the task is to map the space of bifurcation diagrams back to the space of parameters. In particular, we consider inverse bifurcation problems of two types: *identification* and *design*. For the former, one would like to infer parameter values from observed bifurcation diagrams and hence the issue of uniqueness is typically of concern. For the latter, one is interested in parameter values that produce the desired outcome, hence uniqueness is not an issue. The notion of "inverse bifurcation" was first introduced into biology in [13]; more recently, another inverse bifurcation problem from mathematical ecology was studied in [14] by integral equation methods. These papers are concerned with the existence and uniqueness of nonlinear terms in equations realizing prescribed bifurcation diagrams and use analytical methods. In contrast, given the size and complexity of gene networks, we take a computational approach in addressing problems of inverse bifurcation. In this paper, we consider problems of moderate size for which the ill-posedness is not yet a crucial issue. Since ill-posedness increases with dimension, a major issue for larger problems will be to use appropriate regularization techniques.

For the *design* type of inverse bifurcation problems, there exists previous work in the engineering literature: The dis-

tance to bifurcation manifolds has been introduced to quantify the "parametric robustness" of system designs in [15]; in the context of design of chemical processes, optimization problems with constraints involving distance to bifurcation manifolds have been treated in [16]. For a recent review see [17]. For biological applications as we have them in mind, this issue of parametric robustness is also important. In addition, other geometric properties of the bifurcation diagram are of interest. These include the size of the parameter region resulting in bistability of solutions and the parametric distance between regions of different qualitative behavior. We will develop methodologies by which inverse bifurcation problems involving the optimization of such quantities can be solved.

In this paper, we show the applicability of our inverse bifurcation algorithms to low dimensional gene systems. We formulate the inverse problems as constrained optimization problems, whose objective function and constraints involve geometric properties of bifurcation diagrams. We demonstrate that these problems can be solved efficiently by applying gradient-based nonlinear optimization algorithms in combination with one-parameter continuation methods to locate bifurcation points. The latter is a standard capability provided by existing bifurcation analysis software (see [1] for references to state-of-the-art numerical implementations).

2 Inverse bifurcation analysis

The ODE model (1) defines a mathematical relationship between the parameters p and the time course of biochemical concentrations, $x(t)$. Of the set of all parameters, some describe the biochemical mechanisms constituting the machinery of the regulatory system, while others are parameters to whose variation the gene system should respond. In this paper, we refer to the former as the *system parameters* and the latter as the *control inputs*.

In situations where the qualitative behavior of the regulatory system changes in response to shifts in the control input, the corresponding bifurcation diagram is of importance. In particular, the proper working of the regulatory system may depend on the locations, shapes and sizes of the various regions of qualitative behavior. For instance, in the cell cycle model [8] the correct spatial relationship (in parameter space) of the transition points is crucial for ensuring the survival of cells. In fact, the model predicts that "dynamically challenged" mutants with shifted bifurcation diagrams may suffer irrecoverably. The resulting effects include difficulties exiting mitosis and decreases in cell mass after each cell division. For this particular model, the inverse bifurcation problem is to map geometric as well as topological relationships in the bifurcation diagrams to conditions on the parameters. Another application of inverse bifurcation in the cell cycle model is for the

so-called "Pinocchio effect" involving check points. This effect occur when the necessary conditions for safe progression to the next cell phase are not satisfied hence the regulatory system should delay the state transition as far as possible. Here, the inverse bifurcation problem is to find out how the system may be constructed so that, when triggered by the presence or absence of certain chemicals, the range of bistability is maximized.

In the case of system parameters, they are either regulated by other control mechanisms in the cell, or the behavior of the system is constructed to be insensitive to variation in these parameters [18]. Within an operating region in the control input space, the minimal distance in the system parameter space to bifurcation manifolds provides a quantitative measure of the robustness of the system to environmental perturbations. Here, the inverse bifurcation problem would involve finding the parameter combinations so that the system bifurcation diagram is insensitive to the unregulated parameters. Biological applications of this class of problems include homeostasis and rhythmic pacemakers. Again, the problem of interest is to map some shape property of the bifurcation diagram to conditions on the parameter space.

For inverse bifurcation problems of biological interest such as those described above, the question arises as to how to formulate them mathematically so that the solution can be obtained in a computationally tractable and stable manner. Typically, in biological applications the parameter space is of high dimension. Furthermore, in carrying out inverse analysis, the (forward) bifurcation analysis usually has to be applied a large number of times. In most cases, the former condition precludes formulations based on the use of multi-parameter continuation techniques [19]. Instead, formulations based on continuation along rays in parameter space are preferred. Below, we formulate distance to bifurcation manifolds (first introduced in [15]) in terms of a forward operator and an l_2 functional. Subsequently, we describe the sensitivity of the minimal distance with respect to parameters by adjoint methods.

Consider the splitting of m -dimensional parameter space $P \subset \mathbb{R}^m$ into input and system parameters, $p = (p_i, p_s) \in P_i \times p_s$. For an ODE system, let Σ denote a bifurcation manifold of interest, consisting of sets in parameter space P for which structural stability breaks down [1]. For a given system parameter p_s , we further define $\Sigma(p_s) \equiv \Sigma \cap \{p_s\}$ as the intersection of Σ with the p_s -plane. In Figure 2, the geometric relationship between Σ and $\Sigma(p_s)$ is illustrated. Let the forward operator $F : P \rightarrow P$ be a mapping in parameter space, taking a given point to its orthogonal projection on $\Sigma(p_s)$, assumed to be well-defined. That is,

$$F(p) \equiv (F(p)_i, F(p)_s)$$

$$= \left(\mathbb{P}_{\Sigma(p_s)} p_i, p_s \right).$$

The notation is illustrated in Figure 1.

Using $F(p)$, the inverse bifurcation problems examined in this paper are mathematically formulated as:

$$\min_{p_s} J(p) = \|F(p)_i - p_i\|$$

Subject to: $p_{low} \leq p \leq p_{upp}$

$$0 \leq c(F(p)_i), \quad (2)$$

where $\|\cdot\|$ denotes the l_2 -norm and $c : P_i \rightarrow \mathbb{R}^k$ represents k -dimensional nonlinear constraints. In the rest of this section, we discuss adjoint sensitivity analysis of $F(p)$, which is important for applying gradient-based optimization methods for solving (2) as well as for computing $F(p)$ iteratively. Denote the linearization of the above map at p as $F'(p)$. The adjoint operator $F'^*(p)$ is defined to satisfy the following: for all $\delta p, \delta \tilde{p} \in \mathbb{R}^m$,

$$\langle F'(p) \delta p, \delta \tilde{p} \rangle = \langle \delta p, F'^*(p) \delta \tilde{p} \rangle, \quad (3)$$

where the notation $\langle \cdot, \cdot \rangle$ denotes the l_2 -inner product. Suppose $\langle \cdot, l \rangle$ is a linearized functional of interest on P . Given a parameter perturbation δp , the forward functional sensitivity is then given by $\langle F'(p) \delta p, l \rangle$. The same sensitivity can be obtained via the adjoint solution, $\psi \equiv F'^*(p)l$. From the definition of the adjoint operator (3), it follows that:

$$\langle F'(p) \delta p, l \rangle = \langle \delta p, \psi \rangle. \quad (4)$$

That is, for all $\delta p \in P$ the functional sensitivities can be computed via a single linear solution for ψ rather than repeated application of the linearized forward operator, $F'(p) \delta p$.

For the case where the functional of interest is the l_2 -distance $J(p) = \|F(p)_i - p_i\|$, the adjoint solution for the linearization $J'(p)(\cdot)$ is given in terms of the vector normal to the manifold Σ . Denoting N_i and N_s as the components of the normal vector in P_i and P_s respectively, it can be shown that (up to sign),

$$\psi = \pm \frac{1}{\|N_i\|} (N_i, N_s). \quad (5)$$

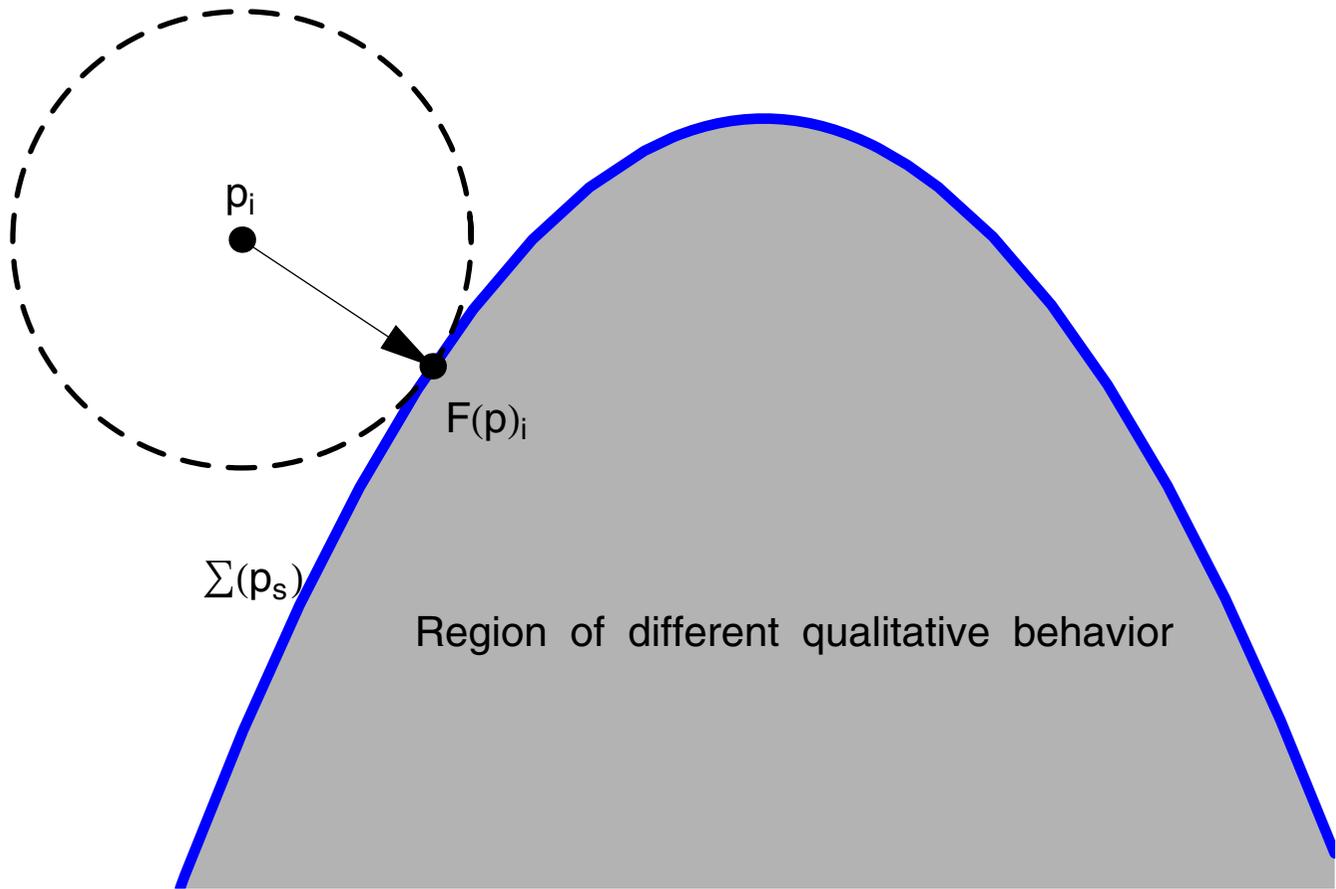


Figure 1
Forward map in the input plane, P_i .

To fix the sign of ψ , the component N_i is chosen to be $p_i - F(p)_i$. Figure 2 illustrates the components of the vector N normal to the manifold Σ .

Thus, obtaining expressions for the adjoint vector for various bifurcations of interest reduces to the problem of deriving the associated normal vectors, N_s . Under certain transversality conditions, the normal vectors for several codimension-one and higher bifurcations have been derived [16,17]. Here we consider the generic codimension-one bifurcations, namely saddle-node and Hopf bifurcations. Let the left and right critical eigenvectors of f_x at the given bifurcation point be denoted by w and v respectively. That is, w and v solve the following eigen-systems for the critical eigenvalue ω_{crit} and its conjugate $\bar{\omega}_{crit}$

$$f_x v = \omega_{crit} v$$

$$f_x^T w = \bar{\omega}_{crit} w.$$

The expressions for normal vectors are given as:

$$\begin{aligned} \text{Saddle - node} &: w^H f_p, \\ \text{Hopf} &: \text{Re} \left\{ \frac{1}{w^H v} \left(w^H f_{xp} v - \lambda^H f_p \right) \right\}, \end{aligned} \tag{6}$$

where superscript H denotes conjugate transpose and $\lambda \equiv \lambda \equiv (w^H f_{xx} v)$. These expressions above prescribe the components of the adjoint solution, thus enabling efficient gradient calculation via (4). Now, we briefly mention methods for computing the projection $F(p)$. Methods of iterative and direct type for finding the (locally) closest bifurcation point have been derived [17]. In the current work we use the former approach, based on using the component of the normal vector in the input plane, N_i . Provided certain conditions on the principal curvatures of $\Sigma(p_s)$ are met, geometric convergence is assured. The algo-

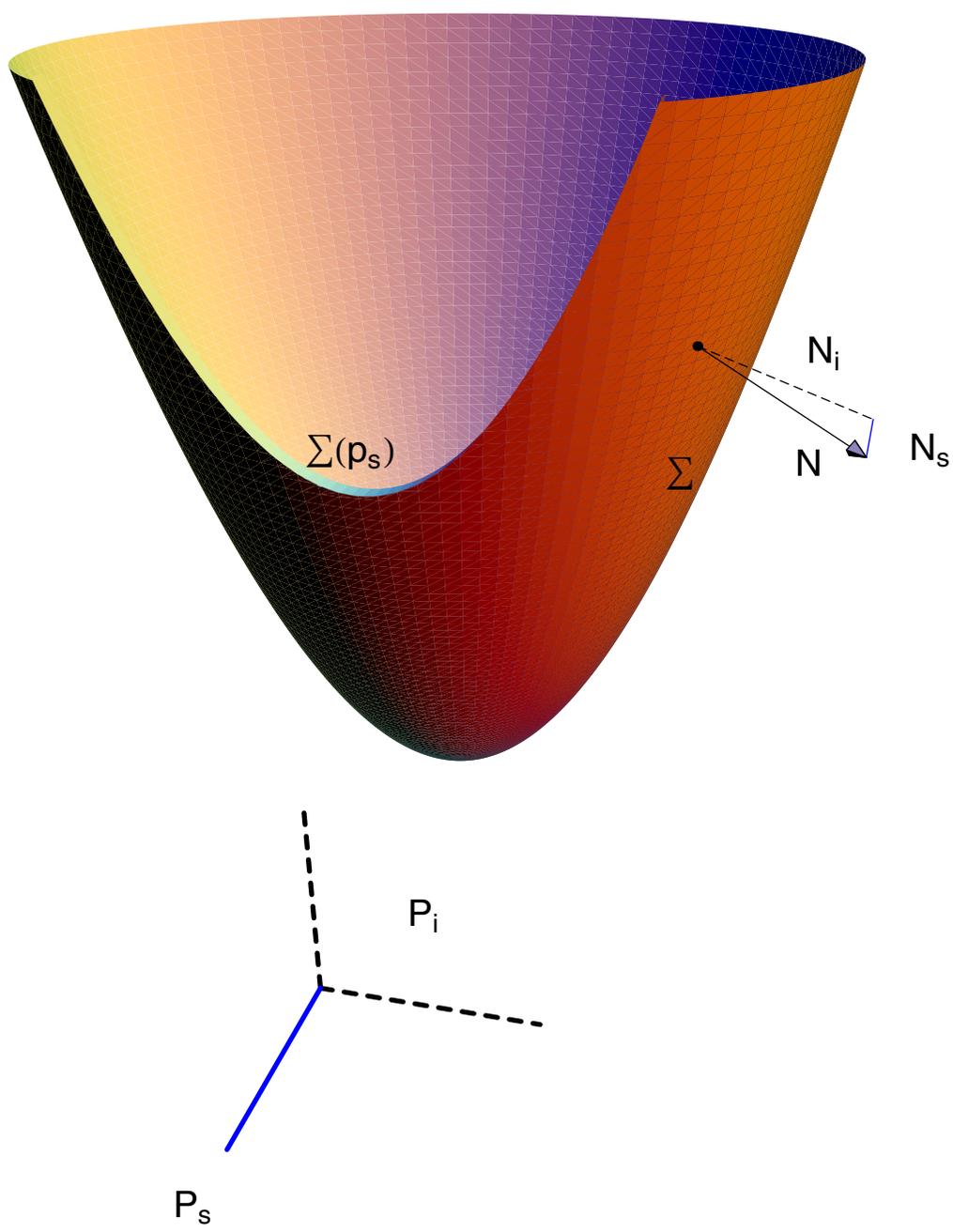


Figure 2
Illustration of normal vector N and components N_i, N_s .

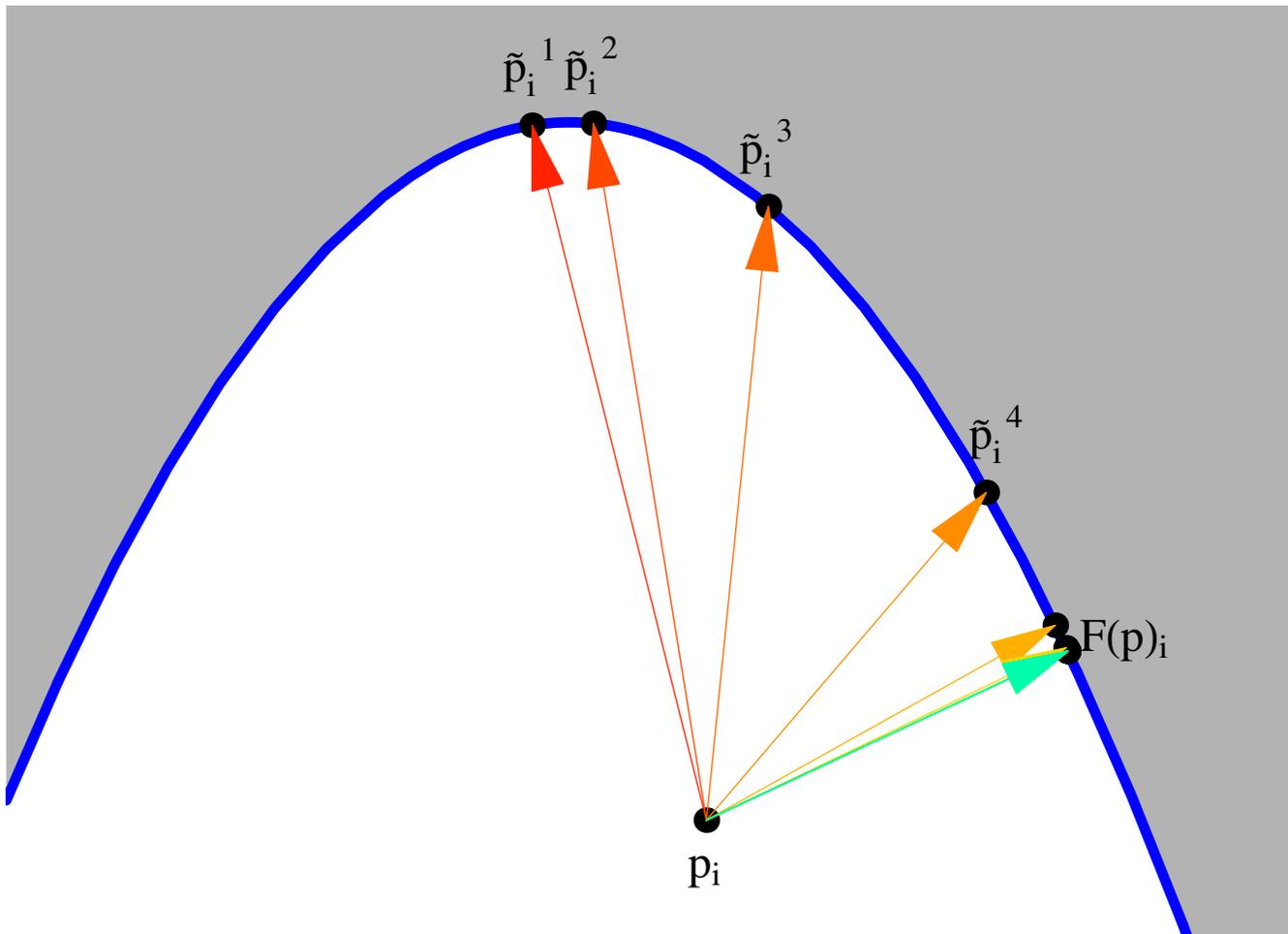


Figure 3
 Demonstration of iterative procedure for computing $F(p)$.

rithm is discussed in Section 3. Figure 3 illustrates the method in a simple example, producing a sequence of iterates (\tilde{p}_i) converging to the point $F(p)_i$ that is closest to a (non-convex) neighboring region with respect to p_i .

Finally, we mention that the minimal distance functional can be used to model many other problems of interest. For instance, it can serve as an estimate of the separation between a reference manifold Σ_{ref} and a given region of qualitative behavior, or the size of the region of a qualitative behavior via the maximum radius of the inscribed sphere.

3 Algorithm and software implementation

Here we give an outline of the algorithm for general inverse bifurcation problems. The main ingredients are applications of the projection operator $F(p)$, as well as the adjoint, $F^*(p)$. The computation of the former is denoted by the routine APPLYF (see Figure 4). Each time APPLYF

is called, corrector steps (using, for instance, Newton's method) have to be carried out on the previously computed x_{init} to find the initial solution for the current value of p_s . Once the corrected solution x_{corr} is computed, the iterative procedure LOCMINDIST (see Figure 5) is called to compute the nearest point on the bifurcation manifold of interest. This procedure is based on a series of one-parameter continuations and gradient calculations according to (6). The inputs include the following: v the initial search direction in parameter space; ϵ_{proj} the relative tolerance on the iterative procedure. In general, several search directions (denoted by $V = \{v_1, \cup, v_{max}\}$) have to be used to approximate the globally closest point. However, for the examples shown in the paper the initial search space V is only 1 dimensional. Once $F(p)$ and the corresponding solution are obtained, $F^*(p)$ is then computed. The derivative information, together with constraints ($p_{low}, p_{upp}, c(p)$) and Hessian approximation (HessA) are

ALGORITHM: LOCINDIST $(x, (p_i, p_s), v, \epsilon,)$

- Set initial parameter: $p^0 \leftarrow p, x^0 \leftarrow x$
- FOR $j = 1, \dots, j_{max}$
 1. From p and x , continue along parameter ray $\{(p_i + rv, p_s): r \in \mathbb{R}_+\}$, until bifurcation point p^b detected
 2. Compute normal vector at bifurcation point p^b : $v \leftarrow N_i(p^b)$
 3. Update: parameter iterate $p^j \leftarrow p^b$
ODE solution at bifurcation point $x^j \leftarrow x(p^b)$
 4. Terminate if $\|p^j - p^{j-1}\|/\|p^0\| < \epsilon$
- END
- Return $[(p_i^j, p_s), x^j]$

Figure 4
Algorithm APPLYF.

then used to compute an SQP step and update the Hessian. Figure 6 describes the inverse bifurcation algorithm.

Our Inverse Bifurcation Toolbox is an implementation of the above described algorithm. It combines the capability of several packages: the MATCONT package [20] for performing the underlying one-parameter continuations, the

ALGORITHM: APPLYF $(x_{init}, (p_i, p_s), \epsilon)$

- $x_{init} \leftarrow \text{INTODESOLN}(x_{init}, p)$
- Generate initial search vectors $V \leftarrow \{v_1, v_2, \dots, v_{max}\}$
- FOR $j = 1, \dots, \dim(V)$

$$\begin{aligned} [F^j, x^j] &\leftarrow \text{LOCINDIST}(x_{init}, p, v_j, \epsilon) \\ d_j &\leftarrow \|F^j - p\| \end{aligned}$$
- END
- $j_m \leftarrow \arg \min_{j=1, \dots, \dim(V)} d_j$
- Return $[F^{j_m}, x^{j_m}, x_{init}]$

Figure 5
Algorithm LOCINDIST.

ALGORITHM: INVERSE BIFURCATION

- Inputs:
 - SBML document
 - Initial parameter $p_i \in P_i$, $p_s \in P_s$, ODE solution x_{init}
 - Parameter bounds $p_{\text{low}}, p_{\text{upp}} \in \mathbb{R}^m$
 - Tolerances $\epsilon_{\text{proj}}, \epsilon_{\text{optim}} > 0$
 - Step-size constraint $\Delta p_{\text{max}} \in \mathbb{R}^m$
 - Nonlinear constraints $c : P_i \rightarrow \mathbb{R}^k$
 - Constrained optimization step:
 - FOR $j = 1, \dots, j_{\text{max}}$
 - $[F, x, x_{\text{init}}] \leftarrow \text{APPLYF}(x_{\text{init}}, (p_i, p_s), \epsilon)$
 - $F'^* \leftarrow \text{APPLYFDERIVADJ}(F, x)$
 - $c \leftarrow \text{APPLYC}(F, x)$
 - $c'^* \leftarrow \text{APPLYCDERIVADJ}(F, F'^*, x)$
 - $[p_s, \tilde{H}] \leftarrow \text{SQPSTEP}(F, F'^*, p_{\text{low}}, p_{\text{upp}}, c, c'^*, \tilde{H}, \Delta p_{\text{max}})$
 - $J_{j+1} \leftarrow \|F(p) - p\|$
 - Terminate if $|(J_{j+1} - J_j)/J_0| \leq \epsilon_{\text{optim}}$
- END

Figure 6
Algorithm INVERSE BIFURCATION.

MATLAB Optimization Toolbox [21] for performing gradient-based constrained optimization, the MathSBML package [22] for reading in biological models in the SBML format and finally the Mathematica Symbolic Toolbox for MATLAB [23] for communications. The combination of MATLAB and Mathematica has the advantage of allowing building on existing freely-available software.

After the SBML model is read in via MathSBML, the vector fields are differentiated symbolically to obtain expressions for f_x, f_p, f_{xp} and f_{xx} . Subsequently, they are numerically evaluated when called by MATCONT. Reparametrization of parameters is done automatically to allow continuation in arbitrary directions. Presently, the toolbox is able to handle bifurcations of saddle-node and

Hopf type. Augmentation to include bifurcations of limit cycles is currently underway.

Given the limitations of the current implementation, our software is suitable for handling problems involving tens of variables and parameters but possibly not suited for problem dimensionality on the order of hundreds or higher. Further algorithm development is needed to significantly upscale the method. In addition, in high dimensional applications instabilities may appear and hence appropriate regularization techniques need to be developed (e.g., stopping rules in the case of iterative methods).

4 Applications

4.1 Finding distance to oscillation in a 4 gene model

To understand design principles underlying biological systems, simple oscillatory and switch-like systems have been constructed experimentally. Several systems based on *E. coli* [24,25] have been successfully demonstrated. Recently, components of the Lac and Ntr systems have been used to construct a genetic clock that upon change in connectivity, exhibits switch behavior [26]. After ignoring the dynamics of the read-out gene, the system of three equations for the concentrations of mRNA (x_1, x_3, x_5) and proteins (x_2, x_4, x_6) is given as,

$$\begin{aligned} \frac{dx_{2k-1}}{dt} &= \beta_{2k-1} (f_{2k-1} - x_{2k-1}) \\ \frac{dx_{2k}}{dt} &= \beta_{2k} (x_{2k-1} - x_{2k}), \end{aligned}$$

where β_i are rate constants. The rates of transcription are represented by the following tri-phasic functions,

$$f_1 = \begin{cases} B & \text{for } x_2^{g_{12}} x_4^{g_{14}} \leq B \\ x_2^{g_{12}} x_4^{g_{14}} & \text{for } B < x_2^{g_{12}} x_4^{g_{14}} < M \\ M & \text{for } x_2^{g_{12}} x_4^{g_{14}} \geq M \end{cases}$$

$$f_3 = \begin{cases} B & \text{for } x_2^{g_{32}} \leq B \\ x_2^{g_{32}} & \text{for } B < x_2^{g_{32}} < M \\ M & \text{for } x_2^{g_{32}} \geq M \end{cases}$$

$$f_5 = \begin{cases} 1/M & \text{for } x_4^{g_{54}} \leq 1/M \\ x_4^{g_{54}} & \text{for } 1/M < x_4^{g_{54}} < 1/B, \\ 1/B & \text{for } x_4^{g_{54}} \geq 1/B \end{cases}$$

where g_{jk} are kinetic-order parameters, typically less than 4.

Given a stable, steady solution for a gene system with a set of nominal parameters, a relevant engineering question is how to construct a genetic oscillator. For the current system, the analytical solutions for the bifurcation manifolds are available via the Routh-Hurwitz stability criterion ([26], supplementary data). However, such analysis cannot be easily carried out for more complex models. Instead, computational methods have to be used.

Suppose the stable system has rate parameter $\beta_1 = \beta_3 = 22.2$, $\beta_2 = \beta_4 = 1.39$, kinetic-order parameters $g_{32} = -2$, $g_{12} = g_{14} = g_{54} = 1$. If in practice only the rate parameters β_i can be varied, the question is how best to construct an oscillatory system. The bifurcation diagram for the model is shown in Figure 7, where it can be seen that for the nominal values of g_{jk} the system is stable and far away from the

line of Hopf bifurcation. The inverse bifurcation question is: do there exist parameters β_i such that the system is oscillatory? We consider the problem with the following parameter constraints:

$$0.1 \leq \beta_1, \beta_3 \leq 25$$

$$0.1 \leq \beta_2, \beta_4 \leq 4.$$

By minimizing the distance from the nominal parameter point ($g_{12} = 1, g_{32}g_{14} = -2$) to the Hopf bifurcation line, the system can be made to lie on the boundary of the stable regime; see Figure 8. Table 1 shows the iteration counts required to obtain the result, using Sequential Quadratic Programming (SQP) with line-search [27]. Owing to the use of line-search, each optimization iteration entails a number of functional evaluations, this in turn requiring a number of one-parameter continuation iterations to find the (locally) closest point on $\Sigma(p_s)$. The optimality tolerance on the objective is set as 10^{-3} relative to its initial value and the relative parameter tolerance for approximating $F(p)$ is set to 10^{-4} . Clearly, the number of one-parameter continuations is significantly higher than the optimization iterations. Thus, for high dimensional examples it is important that the former can be carried out efficiently.

4.2 Maximizing regime of oscillations in repressilator system

Motivated by biological applications such as increasing the robustness of rhythmic pacemakers, we consider maximizing the region of oscillations with respect to a given operating point. We utilize a case of the generalized repressilator [7] as our test system.

After non-dimensionalizing, an ODE system for the protein and RNA concentrations x_i and y_i is obtained, with the following dimensionless parameters: α_i, β_i are the ratios of degradation rates, δ_i represents the leakiness of the gene transcription, h_i the Hill-coefficient reflecting the degree of cooperativity of gene transcription,

$$\begin{aligned} \dot{x}_i &= \beta_i (y_i - x_i) \\ \dot{y}_i &= \alpha_i \left(\frac{1 - \delta_i}{1 + x_{i-1}^{h_i \bmod n}} + \delta_i \right) - y_i, \quad i = 0, \dots, n-1. \end{aligned}$$

In this example, we take $n = 3$ and examine the symmetric case, i.e., $\alpha_i = \alpha, \beta_i = \beta$, etc. Analysis shows that for ranges of parameters α and β , the system can exhibit oscillations or be in a stable, steady-state [7]. Thus, we take $p_i = (\alpha, \beta)$ as the input parameters and seek values of $p_s = (\delta, h)$ within the constraint set,

$$(10^{-4}, 0) \leq (\delta, h) \leq (10^{-1}, 2)$$

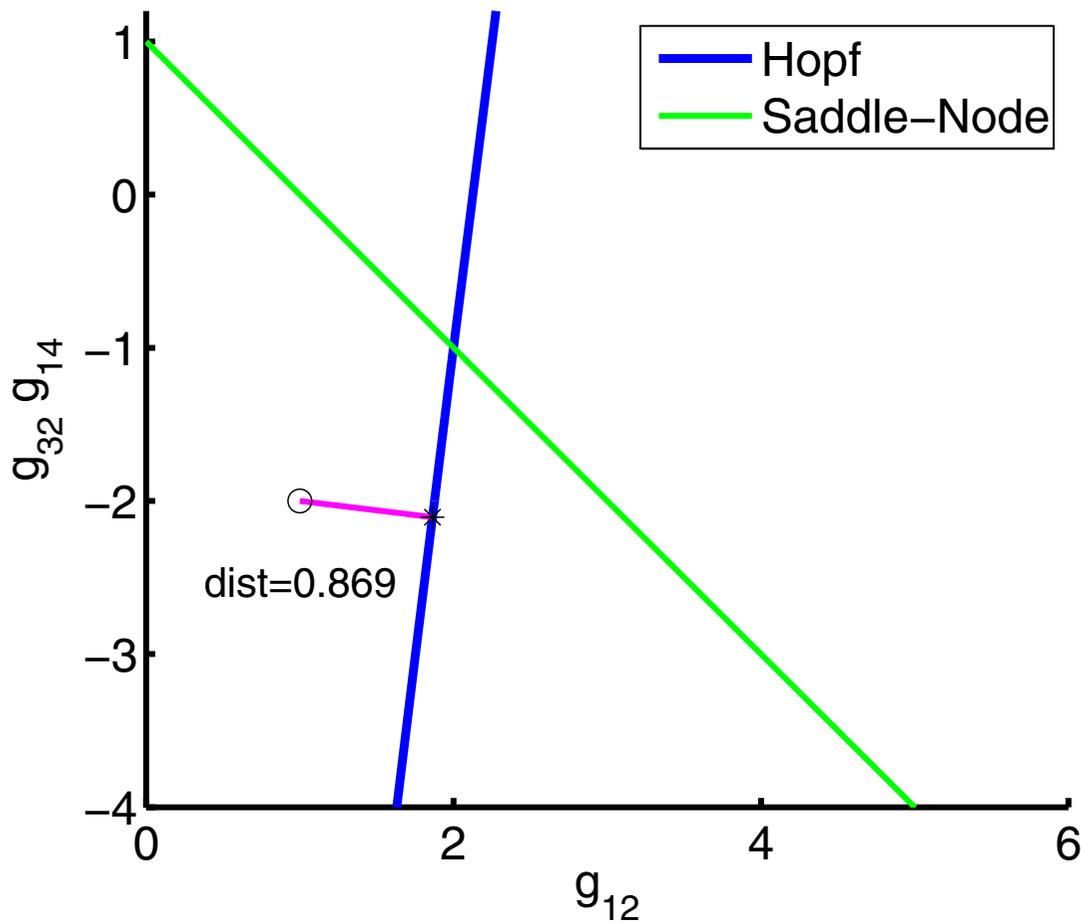


Figure 7
Initial bifurcation diagram for 4 gene model: $\beta_1 = \beta_3 = 222$, $\beta_2 = \beta_4 = 1.39$.

so as to maximize the minimal distance to a Hopf bifurcation with respect to a reference point ($\alpha = 10^{2.5}$, $\beta = 10^0$). Figure 9 shows the initial bifurcation diagram and the computed minimal distance from the reference point. The result of the inverse bifurcation analysis shows that the optimum solution is found at $(\delta, h) = (10^{-4}, 2)$; thus, the inequality constraints are active equality constraints. Moreover, the signs of the Lagrange multipliers for the inequality constraints show that the (local) optimum distance to the Hopf bifurcation manifold increases if the upper bound h_{upp} is increased and/or the lower bound δ_{low} is decreased. Figure 10 shows the bifurcation diagram with the optimized parameters.

Table 2 shows the iteration counts required to obtain the result. The tolerances for optimization are the same as those of Section 4.1.

4.3 Engineering bistable switch in G_1/S transition

We consider the reverse engineering of check points in mammalian cell phase transition. In particular, we use a 9 state, 40 parameter differential equation model [28]. In the publication describing the model, parameter scanning is done to find out parameter values that achieve certain behaviors, such as changing the threshold for the onset of synthesis (S) phase from the G_1 state and conditions ensuring irreversibility of the gene switch. Using our inverse bifurcation analysis, specifying a desired behavior on the transition points gives rise to conditions on the parameters in a more systematic and efficient manner.

Figure 11 shows the bifurcation diagram for the regulatory module of the mammalian G_1/S transition. The bifurcation parameter is the strength of mitogenic stimulation, F_m . The quantity of primary interest is the level of E2F1, which is a transcription factor targeting genes underlying

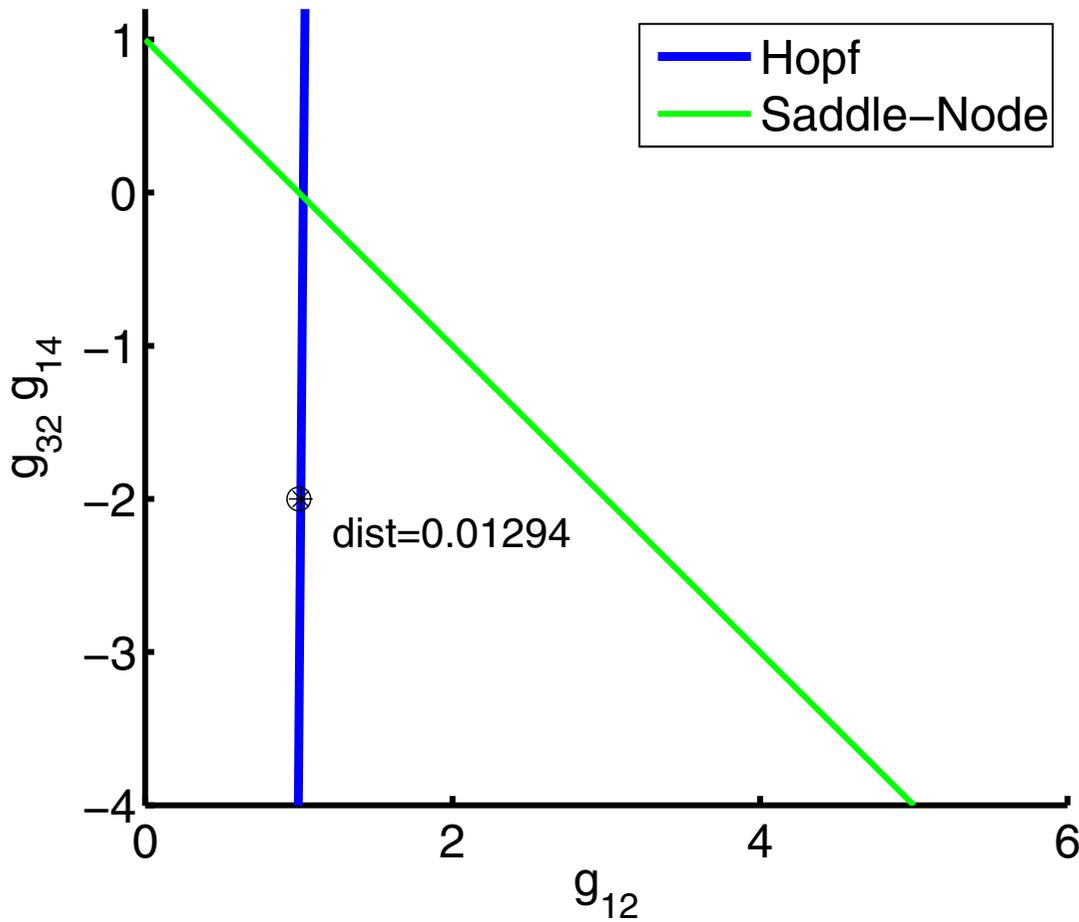


Figure 8
 Optimized bifurcation diagram for 4 gene model: $\beta_1 = 22.186, \beta_2 = 4, \beta_3 = 22.145, \beta_4 = 0.1$.

cell cycle progression. For sufficiently low level of F_m , the level of E2F1 stays at a constant low value. At a critical value of F_m , a transcritical point is reached, above which the level of E2F1 starts to increase with F_m . As the level of F_m increases further, a saddle-node (SN_1) is reached and the level of E2F1 undergoes a discontinuous jump upwards, thereby triggering the S phase of the cell cycle.

4.3.1 Example 1: engineer irreversibility

First, we mention a number of main species and their interactions. The regulatory system consists of a core double inhibitor-activator module as well as several positive-

feedback loops. The transcription factor E2F1 activates itself as well as pRB, a tumor suppressor. In turn, pRB inhibits itself as well as E2F1. The following equations describe this core module:

$$\begin{aligned} \frac{d}{dt}[pRB] &= k_1 \frac{[E2F1]}{K_{m1} + [E2F1]} \frac{J_{11}}{J_{11} + [pRB]} \\ &\quad - \phi_{pRB}[pRB], \\ \frac{d}{dt}[E2F1] &= k_p + k_2 \frac{a^2 + [E2F1]^2}{K_{m2}^2 + [E2F1]^2} \frac{J_{12}}{J_{12} + [pRB]} \\ &\quad - \phi_{E2F1}[E2F1] \end{aligned} \tag{7}$$

Table 1: Iteration summary

	Optim. iter.	Func. eval.	One-param contin.
4 gene system	7	15	60

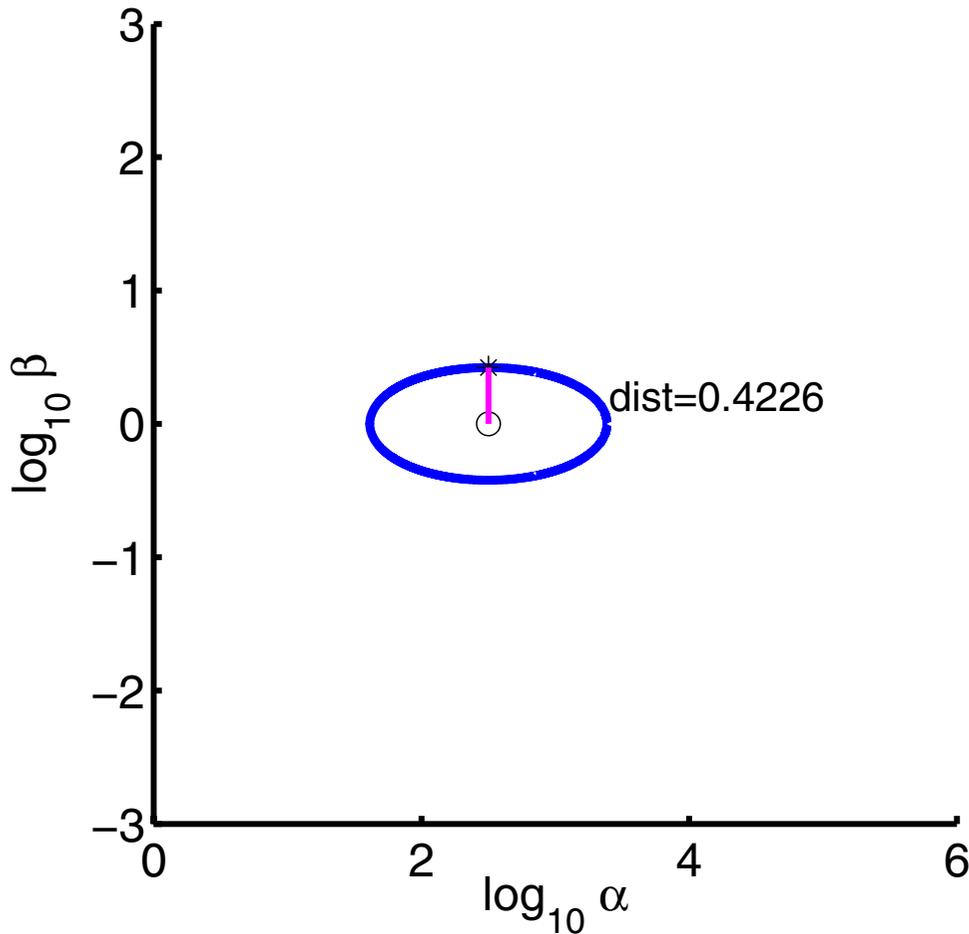


Figure 9
Initial bifurcation diagram for symmetric repressilator: $\delta = 10^{-3}$, $h = 1.5$.

Another important collection of species is AP-1, which is used to denote a family of transcription factors that mediate the mitogenic signal F_m . The family AP-1 is also activated by E2F1, with feedback strength k_{25} . The dynamics for the concentration of the family [AP-1] is modelled as:

$$\frac{d}{dt}[AP-1] = F_m + k_{25}[E2F1] \frac{J_{15}}{J_{15} + [PRB]} \frac{J_{65}}{J_{65} + [PRB_p]} - \phi_{AP-1}[AP-1]. \tag{8}$$

In the first example, we consider the construction of an irreversible gene switch via parameter changes. It has been demonstrated [28] that for some value of k_{25} in (8), the G_1/S transition becomes irreversible. That is, if F_m is increased beyond a certain point, a subsequent decrease to zero will not lead to a G_0 state with a low level of E2F1.

Geometrically, the inverse problem is to find the value of k_{25} such that the x -abscissa of the upper saddle node (SN_2) is as close to zero as possible. As inequality constraints, we take $0.1 \leq k_{25} \leq 1.5$. The result of the inverse analysis shows that the feedback strength should increase from its initial value of $k_{25} = 0.9$ to $k_{25} = 1.099$, resulting in a change of bifurcation diagram shown in Figure 12. Row 1 of Table 3 shows the number of optimization iterations as well as the number of functional evaluations (which in this case equals the number of one-parameter continuations) required to reach a tolerance of 10^{-3} on the function value.

4.3.2 Example 2: engineer irreversibility, fixed G_1/S transition

In this example, we consider an extension of the previous problem. Here, we would like to engineer an irreversible switch with the additional constraint that the x -abscissa of

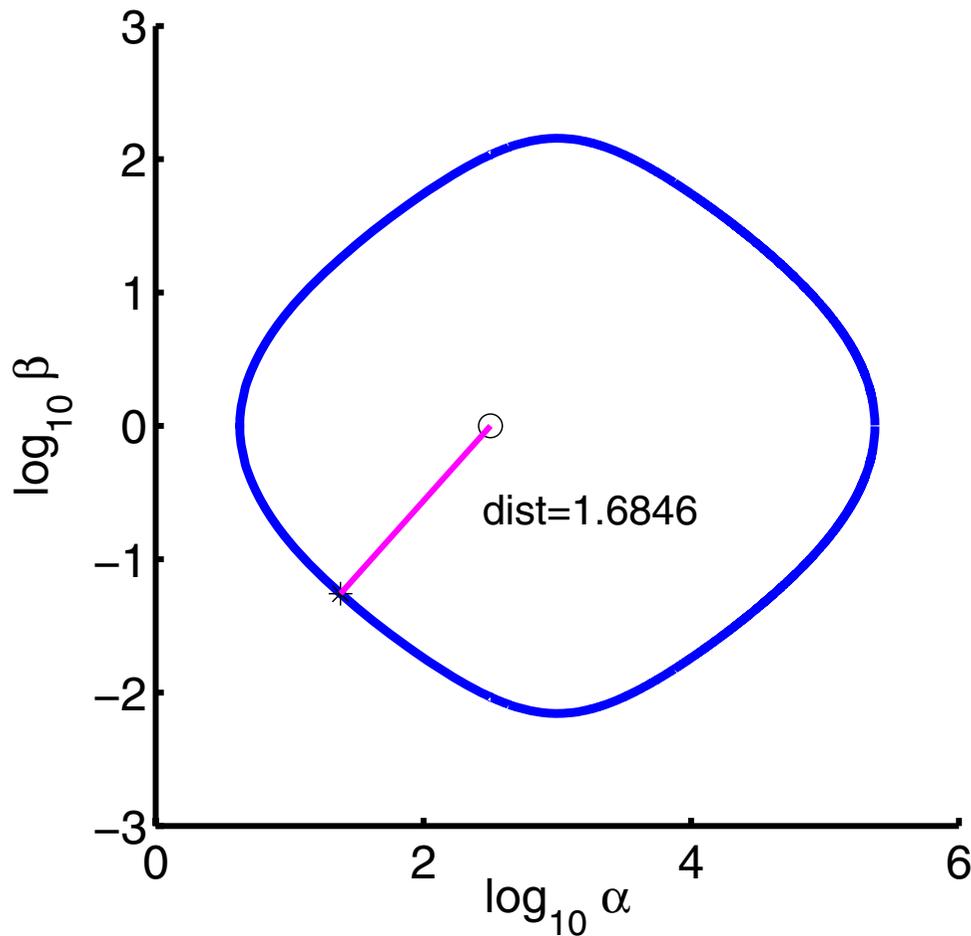


Figure 10
Optimized bifurcation diagram for symmetric repressilator: $\delta = 10^{-4}$, $h = 2$.

the G_1/S transition point SN_1 should be fixed at the value of the nominal system. In addition to allowing $0.1 \leq k_{25} \leq 1.5$ as before, we also modify the stability of E2F1 and AP-1. That is, we relax the values of ϕ_{E2F1} and ϕ_{AP-1} in (7) and (8) from their fixed nominal values 0.1 and 0.01 respectively, to allow searching within the ranges,

$$0.08 \leq \phi_{E2F1} \leq 0.12$$

$$0.008 \leq \phi_{AP-1} \leq 0.012.$$

Mathematically, the inverse problem is formulated as an nonlinear optimization problem with equality constraint on the x -abscissa of the G_1/S point, SN_1 . The bifurcation diagram corresponding to the solution ($k_{25} = 1.182$, $\phi_{E2F1} = 0.0992$, $\phi_{AP-1} = 0.0109$) is given in Figure 13. Row 2 of Table 3 summarizes the number of iterations required.

5 Conclusion

We propose an inverse problems methodology to address a class of problems arising in the study and design of gene systems. The methodology is based on the (inverse) map from the space of bifurcation diagrams to the space of bio-

Table 2: Iteration summary

	Optim. iter.	Func. eval.	One-param contin.
Symmetric repressilator	2	5	48

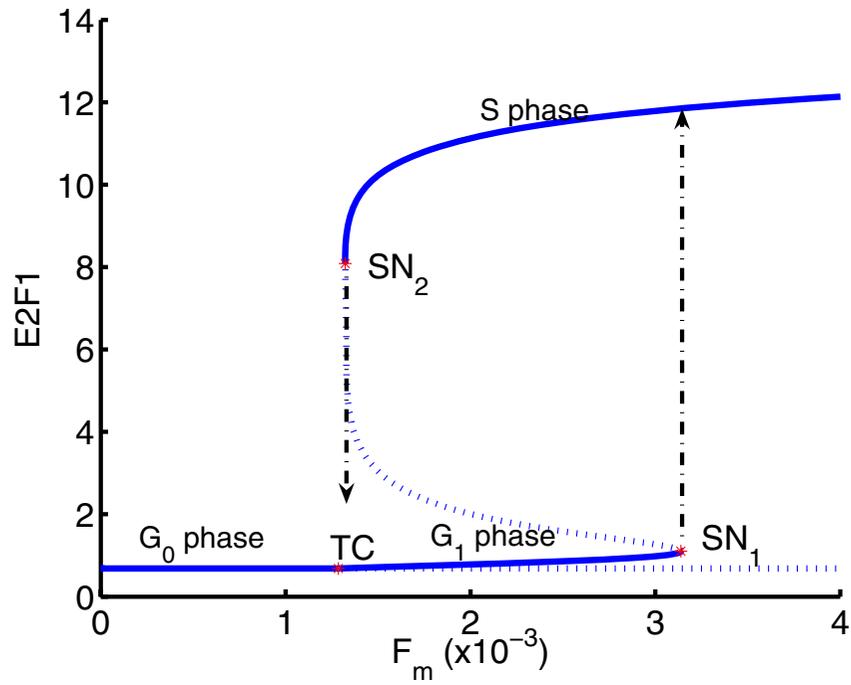


Figure 11
Bifurcation diagram of cell phase transition.

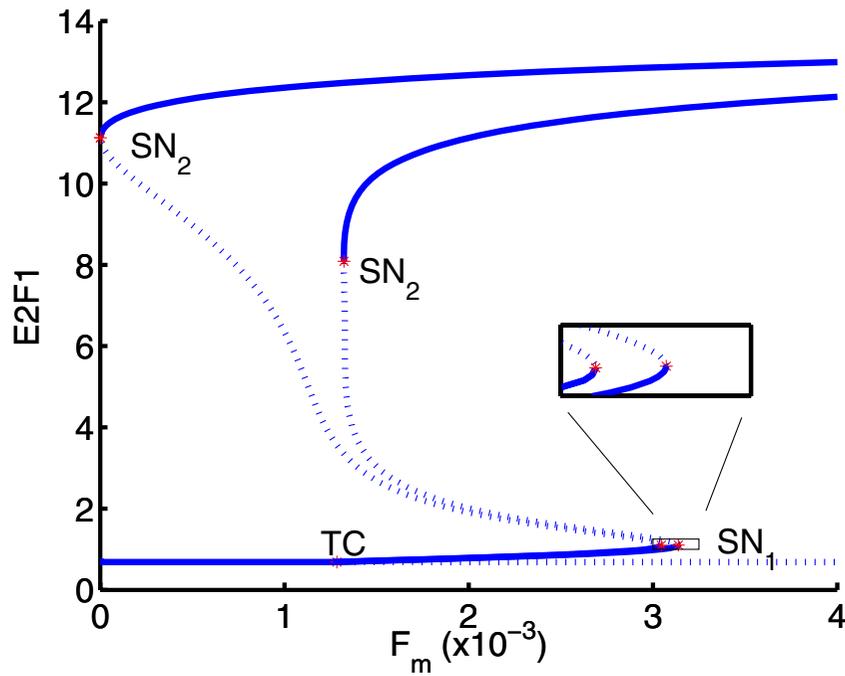


Figure 12
Example 1: initial and optimized gene switches.

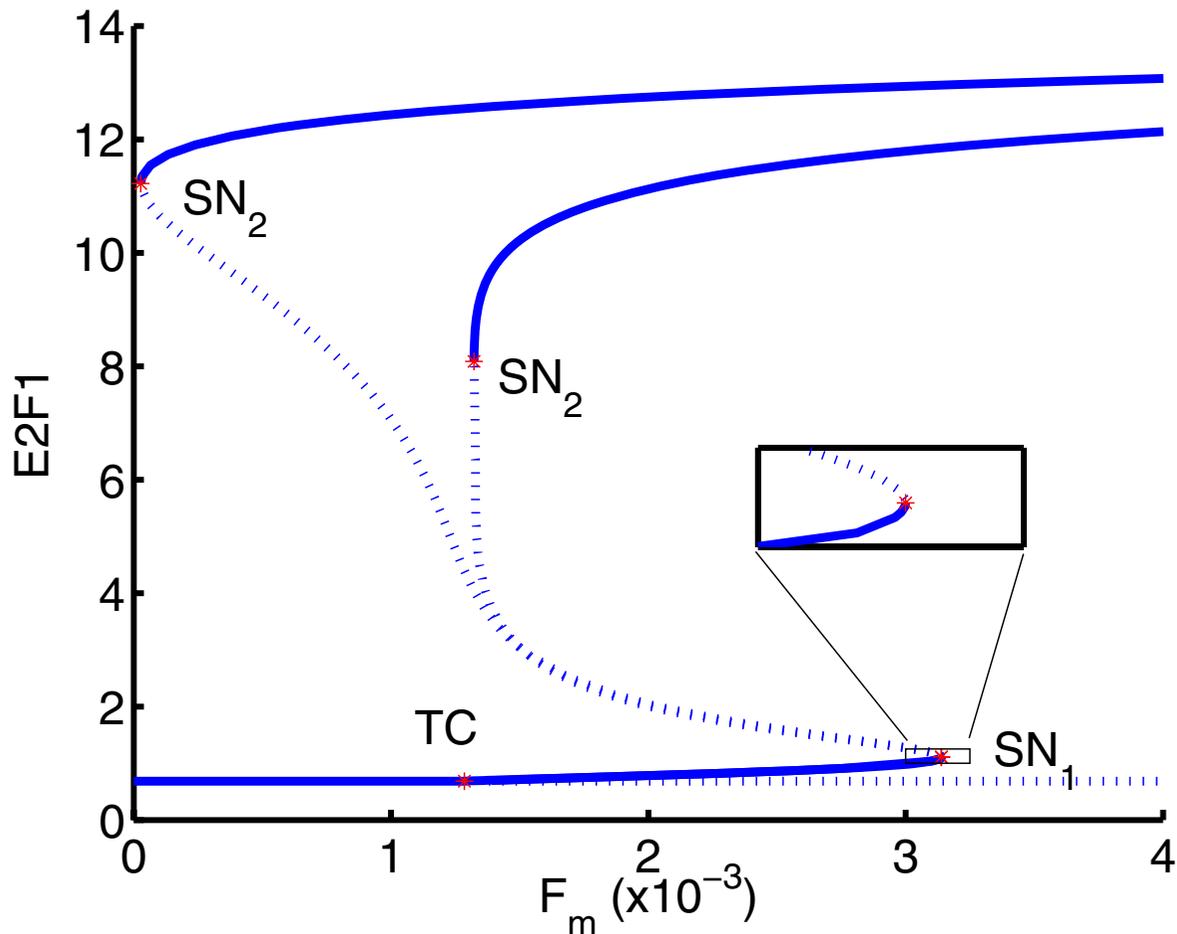


Figure 13
Example 2: initial and optimized gene switches.

chemical parameters. Within this general inverse bifurcation framework, we show that many questions of biological interest may be formulated as optimization problems involving minimal distances to bifurcation manifolds. Adjoint sensitivity analysis is carried out to allow for efficient solution using gradient-based optimization methods.

Several general questions remain for the methodology: what biological problems can be formulated in terms of geometric properties of the bifurcation diagrams? Given desired geometric properties for bifurcation diagrams, how best can the problem be formulated mathematically to allow for tractable solution? Finally, we remark that the development of appropriate numerical methods, novel

Table 3: Iteration summary

	Optim. iter.	Func. eval.
Example 1: irreversibility	6	14
Example 2: irreversibility, fixed G_1/S	9	22

computational strategies and regularization techniques are necessary to allow for the upscaling of the current computational tool to high dimensional biological applications.

Acknowledgements

We gratefully acknowledge Stefan Müller for providing the repressilator model, Lukas Endler for helpful discussions on the cell cycle model, as well as Andreas Hofinger and Philipp Kügler for comments on the paper. This work is supported by the WWTF, project number MA05.

References

1. Kuznetsov YA: *Elements of Applied Bifurcation Theory* New York, USA: Springer-Verlag; 2004.
2. Thomas R, D'Ari R: *Biological feedback* Florida, USA: CRC-Press; 1990.
3. Hasty J, Dolnik M, Rottschäfer V, Collins JJ: **Synthetic gene network for entraining and amplifying cellular oscillations.** *Phys Rev Lett* 2002, **88(14)**:148101.
4. Tyson JJ, Chen KC, Novak B: **Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell.** *Curr Opin Cell Biol* 2003, **15(2)**:221-231.
5. Angeli D, Ferrell JEJ, Sontag ED: **Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems.** *Proc Natl Acad Sci USA* 2004, **101(7)**:1822-1827.
6. Widder S, Schicho J, Schuster P: **Dynamic patterns of gene regulation I: simple two gene systems.** *J Theor Biol* in press.
7. Müller S, Hofbauer J, Endler L, Flamm C, Widder S, Schuster P: **A generalized model of the repressilator.** *J Math Biol* in press.
8. Tyson JJ, Csikasz-Nagy A, Novak B: **The dynamics of cell cycle regulation.** *Bioessays* 2002, **24(12)**:1095-1109.
9. Battogtokh D, Tyson JJ: **Bifurcation analysis of a model of the budding yeast cell cycle.** *Chaos* 2004, **14(3)**:653-661.
10. Chickarmane V, Paladugu SR, Bergmann F, Sauro HM: **Bifurcation discovery tool.** *Bioinformatics* 2005, **21(18)**:3688-3690.
11. Engl HW, Hanke M, Neubauer A: *Regularization of inverse problems, Volume 375 of Mathematics and its Applications* Dordrecht:Kluwer Academic Publishers Group; 1996.
12. Engl HW, Kügler P: **Nonlinear inverse problems: theoretical aspects and some industrial applications.** In *Multidisciplinary Methods for Analysis, Optimization and Control of Complex Systems* Edited by: V Capasso JP. New York: Springer Verlag; 2005:3-48.
13. Iwasaki K, Kamimura Y: **An inverse bifurcation problem and an integral equation of the Abel type.** *Inverse Problems* 1997, **13**:1015-1031.
14. Iwasaki K, Kamimura Y: **Inverse bifurcation problem, singular Wiener-Hopf equations, and mathematical models in ecology.** *J Math Biol* 2001, **43**:101-143.
15. Dobson I: **Computing a closest bifurcation instability in multidimensional parameter space.** *J Nonlinear Sci* 1993, **3(3)**:307-327.
16. Mönnigmann M, Marquardt W: **Normal vectors on manifolds of critical points for parametric robustness of equilibrium solutions of ODE systems.** *J Nonlinear Sci* 2002, **12(2)**:85-112.
17. Dobson I: **Distance to bifurcation in multidimensional parameter space: margin sensitivity and closest bifurcations.** In *Bifurcation control, theory and application* Edited by: Chen G, Hill DJ, Yu X. New York: Springer Verlag; 2003:49-66.
18. Morohashi M, Winn AE, Borisuk MT, Bolouri H, Doyle J, Kitano H: **Robustness as a measure of plausibility in models of biochemical networks.** *J Theor Biol* 2002, **216**:19-30.
19. Henderson ME: **Multiple parameter continuation: computing implicitly defined k-anifolds.** *J Bifur Chaos* 2002, **12**:451-476.
20. Dhooge A, Govaerts W, Kuznetsov YA: **MATCONT: a MATLAB package for numerical bifurcation analysis of ODEs.** *ACM Trans Math Software* 2003, **29(2)**:141-164.
21. **MathWorks Ltd** [<http://www.mathworks.com/>]
22. Shapiro BE, Hucka M, Finney A, Doyle J: **MathSBML: a package for manipulating SBML-based biological models.** *Bioinformatics* 2004, **20**:2829-2831.
23. **Mathematica Symbolic Toolbox for MATLAB – Version 2.0** [<http://library.wolfram.com/infocenter/MathSource/5344/>]
24. Elowitz MB, Leibler S: **A synthetic oscillatory network of transcriptional regulators.** *Nature* 2000, **403(6767)**:335-338.
25. Gardner TS, Cantor CR, Collins JJ: **Construction of a genetic toggle switch in Escherichia coli.** *Nature* 2000, **403(6767)**:339-342.
26. Atkinson MR, Savageau MA, Myers JT, Ninfa AJ: **Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli.** *Cell* 2003, **113(5)**:597-607.
27. Gill PE, Murray W, Wright MH: *Practical Optimization* London: Academic Press; 1981.
28. Swat M, Kel A, Herzog H: **Bifurcation analysis of the regulatory modules of the mammalian G1/S transition.** *Bioinformatics* 2004, **20(10)**:1506-1511.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

