5-2019

Identification of Parameters in Systems Biology

Roby Poteau

Identification of Parameters in Systems Biology

by

Roby Poteau

Bachelor of Science
Department of Mathematical Sciences
Florida Institute of Technology
2012

A dissertation
submitted to Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Doctorate of Philosophy
in
Operations Research

Melbourne, Florida
May, 2019

We the undersigned committee
hereby approve the attached dissertation

Identification of Parameters in Systems Biology by Roby Poteau

---

Ugur G. Abdulla, Ph.D.,Dr.Sci.,Dr.rer.nat.habil.
Professor
Department of Mathematical Sciences
Committee Chair

Robert van Woesik, Ph.D.
Professor
Department of Ocean Engineering and Marine Sciences
Outside Committee Member

Jian Du, Ph.D.
Associate Professor
Department of Mathematical Sciences
Committee Member

Vladislav Bukshtynov, Ph.D.
Assistant Professor
Department of Mathematical Sciences
Committee Member

Munevver Mine Subasi, Ph.D.
Associate Professor and Department Head
Department of Mathematical Sciences
Committee Member

ABSTRACT

Title:

Identification of Parameters in Systems Biology

Author:

Roby Poteau

Major Advisor:

Ugur G. Abdulla, Ph.D.,Dr.Sci.,Dr.rer.nat.habil.

Systems Biology is an actively emerging interdisciplinary area between biology and applied mathematics, based on the idea of treating biological systems as a whole entity which is more than the sum of its interrelated components. One of the major goals of systems biology is to reveal, understand, and predict such properties through the development of mathematical models based on experimental data. In many cases, predictive models of systems biology are described by large systems of nonlinear differential equations. Quantitative identification of such systems requires the solution of inverse problems on the identification of parameters of the system. This dissertation explores the inverse problem for the identification of the finite dimensional set of parameters for systems of nonlinear ordinary differential equations (ODEs) arising in systems biology. Two numerical methods are implemented. The first method combines the ideas of Pontryagin optimization or Bellman's quasilinearization with sensitivity analysis and Tikhonov's regularization. The method is applied to various biological models such as the classical Lotka-Volterra system, bistable switch model in genetic regulatory networks, gene regulation and repressilator models from synthetic biology. The numerical results and application to real data demonstrate the superlinear convergence with convergence rate close to quadratic. The method proved to be extremely effective in moderate

scale models of systems biology. The results are published in a recent paper *Mathematical Biosciences, 305(2018), 133-145*. To address adaptation and scalability of the method for large-scale models of systems biology the modification of the method is developed by embedding a method of staggered corrector for sensitivity analysis and by enhancing multi-objective optimization which enables application of the new method to large-scale models with practically non-identifiable parameters based on multiple data sets, possibly with partial and noisy measurements. The new method is applied to benchmark model of three-step pathway modelled by 8 nonlinear ODEs with 36 unknown parameters and two control input parameters. The numerical results demonstrate the superlinear convergence with minimum five data sets and with minimum measurements per data set. The method is extremely robust with respect to partial and noisy measurements, and in terms of required number of measurements for each component of the system. Optimal choice of the Tikhonov regularization parameter significantly improves convergence rate, precision and convergence range of the algorithm. Software package *qlopt* is developed for both methods and posted in GitHub. MATLAB package AMIGO2 is used to demonstrate advantage of *qlopt* over most popular methods/software such as *lsqnonlin*, *fmincon* and *nl2sol* in an equivalent setting.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like state my sincere gratitude to my Ph.D. advisor, Dr. Ugur Abdulla, who pushed me to accomplish greater things than I would have imagined for myself. I would like to say thank you to all the professors at Florida Tech who played a part in my academic and personal developlment, especially Drs. Subasi, Dshalalow and Koksal. I would also like to give warm thank you to my managers, Erika Williams, Tracy Spaulding, Andrea Pinzon-Marquardt at Christian Care Ministry. They allowed me the flexibility at work so I could finish my Ph.D. Last, I would like to say acknowledge the many people that believed in me even when I did not.

Dedication

I would like to dedicate this dissertation to my grandparents, Reneau and Romaine, who worked hard to enable my parents to immigrate to the United States. Their sacrifice opened a new world of opportunties for all their children. The affect is still being felt by us today. I would like to dedicate this to my parents, Seige and Dinette Poteau, who raised me. Their intentionality in raising my sister and I made a significant difference in our trajectory for the best. We have beaten all the statistics. I would like to dedicate my dissertation to my older sister Dina Poteau. She set the bar for me academically and she set it high. As I sought to follow in her steps I pursued greater academic achievements than I would have without her. She also challenged my way of thinking, thereby teaching how to think critically about the world around me. Last, I would like to dedicate this to my extended family. This marks another milestone in the Poteau family legacy.

Chapter 1

Introduction

Systems Biology is an actively emerging interdisciplinary field between biology and mathematics, based on the idea of treating biological systems as a whole entity which is more than the sum of its interrelated components. These systems are networks with emerging properties generated by complex interaction of a large number of cells or organisms. The mission is to reveal and understand the global properties of biological or bioengineering systems that occur through complex interactions on a microscopic level. One of the major goals of systems biology is to reveal, understand, and predict such properties through the development of mathematical models based on experimental data. In many cases, predictive models of systems biology are described by large systems of nonlinear differential equations. Quantitative identification of such systems requires the solution of inverse problems on the identification of parameters of the system. Inverse problems are ill-posed, meaning that a solution may not exist, may not be unique, and most importantly, may not be continuously dependent on the measurements. Since measurements always contain some error, it is impossible to solve inverse problems

without regularization techniques [1, 2].

Ill-posedness of the inverse problems is strongly associated with the problem of *non-identifiability* and correlation of parameters. This creates an essential obstacle in solving the inverse problems for many Systems Biology models [3–6]. The major difficulty arises due to *structural non-identifiability* which is an intrinsic property of the model and cannot be resolved with additional or more accurate measurements. On the other hand, structurally identifiable parameter set in many models could be *practically non-identifiable* meaning that it can be resolved with improved and/or additional measurement or through implementation of various regularization methods [1,3–8]. Besides the difficulties associated with ill-posedness of inverse problems, nonlinearity of the system causes extreme sensitivity with respect to the parameters. Therefore, delicate sensitivity analysis is of high demand [9]. Inverse problems on the identification of parameters in systems biology are an actively growing research area [2, 10–22]. We refer to survey articles [11, 19, 23, 24] for the extensive list of references. Standard popular approach to parameter identification problem in systems biology is its formulation as nonlinear optimization problem with objective to find unknown parameters via minimization of the mismatch or residual between experimental data and model dynamics. Ultimate goal is to develop a global optimization method with least computational cost, which is robust with respect to nonlinearities and scales well with problem size [24]. Currently such an ideal method does not exist. Local methods can be classified as gradient-based methods ([25]) and gradient-free ones [26, 27]. Gradient based methods are very effective, but they only provide convergence to local optima. Gradient-free methods are less efficient, and has slow convergence rate [27]. Global optimization methods can be classified as stochastic ([28]) and deterministic ones [29]. Stochas-

tic global optimization algorithms are implementing pseudorandom sequences to determine search directions toward the global optimum. Most effective class of stochastic methods are pure random search and adaptive sequential methods, clustering methods, population-based methods and nature inspired methods [19, 30]. Advantage of popular global stochastic methods such as simulated annealing, particle swarm optimization, or genetic algorithms is their scalability, but the downside is high computational cost [31]. Deterministic local optimization methods can be used as global optimization method by embedding "Multi-start" strategy into it which facilitates many optimization runs from randomly selected initial parameter guesses [19,23]. Latin hypercube sampling [32] for partition of the parameter space is used to guarantee that each parameter estimation iteration starts with initial guess from different region in parameter space.

Some of the most popular local optimization methods available as open source software ([19, 33, 34]) are the following:

- Levenberg-Marquardt algorithm and trust-region-reflective method (function *lsqnonlin* in MatLab) [10]
- Sequential Quadratic Programming (function *fmincon* in MatLab Optimization toolbox)
- An adaptive non-linear least-squares algorithm (function *nl2sol* in MatLab) [20]

One can also consider so called *hybrid optimization algorithms* as a combination of stochastic and deterministic algorithms: first candidate set of parameter values are generated by stochastic algorithm, followed by deterministic iterative algorithms by choosing candidate set elements as initial guesses [19,23]. In [23] comprehensive

comparison of 15 optimization algorithms from three groups is pursued. 12 stochastic optimization algorithms ([35]) were compared with a deterministic trust region algorithm ([36]) in combination with Latin hypecube sampling and two different approaches for calculating derivatives of the objective function: finite difference approximation and analytically derived sensitivity equations. The results show that multi-start deterministic optimization using the sensitivity equations for the calculation of derivatives significantly outperforms all other tested algorithms. The performance of stochastic optimization algorithms was surprisingly low compared to the hybrid and fully deterministic optimization algorithms [23] .

In another recent paper [24] collection of benchmark problems were selected to evaluate the performance of two families of optimization methods:

- **MS**: multi-start local optimization.
- **eSS**: enhanced scatter search metaheuristic,

the latter may be combined with deterministic local searches, leading to hybrid methods. Selected local optimization methods were

- *nl2sol*-FWD: the nonlinear least-squares algorithm *nl2sol*, using forward sensitivity analysis for evaluating the gradients of the residuals.
- *fmincon*-ADJ: the interior point algorithm included in *fmincon*, using adjoint sensitivities for evaluating the gradient of the objective function.
- *dhc*-a gradient-free dynamic hill climbing algorithm.
- **eSS** without any local methods **NOLOC** and **PSO** - particle swarm optimization [37].

Comprehensive evaluation of [24] clearly shows that high-quality sensitivity calculation methods provide a competitive advantage to local methods that exploit them. Optimization using adjoint and forward sensitivity analysis $fmincon - ADJ$ and $nl2sol - FWD$ usually outperform the gradient-free alternative *dhc*. The combinations of **eSS** with gradient-based methods, **eSS**-*fmincon*-ADJ and **eSS**-*nl2sol*-FWD, clearly outperform the gradient-free alternatives, **eSS**-*dhc* and **eSS-NOLOC**, as well as the gradient-free **PSO**.

The comparison analysis performed in [23, 24] demonstrates that robust deterministic local optimization methods embedded with **MS** or **eSS** strategy, and with sharp sensitivity analysis platform are the best candidates for creation of powerful global optimization methods for large-scale models of systems biology.

The goal of the dissertation is to develop effective local optimization method for solving inverse problems on the identification of finite-dimensional parameter sets for the systems of nonlinear ODEs arising in Systems Biology, and to develop an effective software which is competitive with currently known popular methods/software such as *lsqnonlin*, *fmincon* and *nl2sol*.

In Chapter 2 we implement the numerical method introduced originally in [38, 39]. The iterative method combines ideas of Pontryagin's optimization or Bellman's quasilinearization with sensitivity analysis and Tikhonov regularization. A description of the method is given in Section 2.1, including a discussion on regularization and identifiability. The results are presented in Section 2.2 with a summary of the experimental design. Subsequently, Section 2.2.1 presents the results for the Lotka-Volterra model with both simulated and real data. Section 2.2.2 presents the results for genetic circuit model with bistable switch. Continuing on to Section 2.2.3 are the results for Lorenz's famous atmospheric model, which is known for

its chaotic behavior. Synthetic biology toggle switch model which describes gene transcription regulation by repressors is explored in Section 2.2.4. Lastly, in Section 2.2.5, we analyze the model for the genetic network for three repressor-protein concentrations and their corresponding *mRNA* concentrations. Extensive computational analysis pursued in Sections 2.2.1-2.2.5 demonstrated that the method is very well adapted to canonical models of system biology with moderate size parameter sets and has quadratic convergence. Software package *qlopt* was developed and posted in GitHub [40]. MATLAB package AMIGO2 [19] was used to demonstrate the competitiveness and advantage of *qlopt* with other most popular local search methods like *lsqnonlin, fmincon, nl2sol*. The comparison of our method and developed software *qlopt* with the most advance methods/software is demonstrated in Section 2.3. The results of the Chapter 2 is published in [41].

However, direct adaptation and scalability of the method implemented in Chapter 2 to inverse problems with significantly larger size was not as effective. In Chapter 3 we introduce a modification of the method which is effective to solve the inverse problem on the identification of parameters for large scale models in systems biology. In Section 3.1 we introduce the new modified method accompanied with two types of Tikhonov regularization algorithms. The modification is twofold.

- Method of staggered corrector [42] is embedded into the step of calculation of the sensitivity vectors. Precisely, instead of solving linearized system and associated sensitivity system, we first solve original system through quasi-linearization [43], and then use its solution to solve linear sensitivity system corresponding to the original nonlinear system. We use software package CVODES [44] to implement the method of staggered corrector into our al-

gorithm.

- Multi-objective optimization is added to the method which enables the application of the method to large-scale models that are practically non-identifiable due to parameter correlation.

In Section 3.2 we present results and analysis of the application modified method to benchmark model of a biological network for a three-step pathway modelled by 8 nonlinear ODEs describing 8 metabolic concentrations with 36 unknown parameters and two control input parameters specifying the experimental design. Section 3.2.1 presents the results for noise-free simulated data. It is demonstrated that with 5 to 16 data sets are satisfactory to uniquely identify all 36 parameters with high precision. This is followed by Section 3.2.2 where we demonstrate that the delicate implementation of the Tikhonov regularization with optimal choice of the regularization parameter significantly affects the convergence rate and precision of the algorithm. In the following Section 3.2.3 we demonstrate the effect of the number of time measurements for each component of the system on the convergence and accuracy of the method. Subsequently in Section 3.2.4 we explore in depth the major question of identifiability of the parameters with minimum number of data sets. It is demonstrated that minimum five data sets are required to identify uniquely all parameters with high precision. Section 3.2.5 demonstrates that the careful implementation of the type 2 Tikhonov regularization significantly improves the convergence range of the algorithm. In Section 3.2.6 we apply the method to simulated noisy data and demonstrate its robustness. Section 3.2.7 presents numerical analysis of the rate of convergence of the method. Section 3.2.8 demonstrates the robustness of the method with respect to partial measurements.

Finally, in he last Section 3.2.9, we pursue comparison and demonstrate the competitiveness and the advantages of our method and the associated software package *qlopt* against most advance methods/software like *lsqnonlin, fmincon, nl2sol*. The results of the Chapter 3 are submitted for publication and the preprint is posted in Math arXiv [45].

Chapter 2

Identification of constant parameters in systems biology

The results of this chapter were published in *Mathematical Biosciences, 305(2018), 133-145.*

2.1 Description of the problem and numerical method

Consider a dynamical system:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(t, \mathbf{x}, \mathbf{u}) \tag{2.1}$$

$$\mathbf{x}(t_0) = \mathbf{x}^0 \in \mathbb{R}^n, \tag{2.2}$$

where

$$\mathbf{x} = \mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$$

is the state vector,

$$\mathbf{u} = (u_1, u_2, \ldots, u_m) \in \mathbb{R}^m$$

is the parameter vector, and

$$\mathbf{f} = (f_1(t, \mathbf{x}, \mathbf{u}), f_2(t, \mathbf{x}, \mathbf{u}), \ldots, f_n(t, \mathbf{x}, \mathbf{u})) :$$

$$[t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$$

is differentiable vector function with continuous derivatives

$$\frac{\partial f_i}{\partial x_j}, \frac{\partial f_i}{\partial u_j}, \frac{\partial^2 f_i}{\partial u_k \partial x_j}.$$

Consider *inverse problem of finding the parameter* $\mathbf{u}$ *given measurements for* $\mathbf{x} = \mathbf{x}(t) = \mathbf{x}(t, \mathbf{u})$ *on an interval* $t_0 \leq t \leq t_1$, *where* $\mathbf{x}(t_0) = \mathbf{x}^0$.

Having chosen the initial vector function $\mathbf{x}_0$ (say, $\mathbf{x}_0 = \mathbf{x}(t)$), we implement quasilinearization of (2.1) ([43]) and at each iteration $N = 1, 2, ...$ consider the linear system of ODEs for the vector function $\mathbf{x}_N(t, u)$:

$$\frac{\mathrm{d}\mathbf{x}_{_N}}{\mathrm{d}t} = \mathbf{f}(t, \mathbf{x}_{_{N-1}}, \mathbf{u}) + J(t, \mathbf{x}_{_{N-1}}, \mathbf{u})(\mathbf{x}_{_N} - \mathbf{x}_{_{N-1}}) \tag{2.3}$$

$$\mathbf{x}_{_N}(t_0) = \mathbf{x}^0, \tag{2.4}$$

where

$$J(t, \mathbf{x}_{_{N-1}}, \mathbf{u}) = \left[\frac{\partial f_i(t, \mathbf{x}_{_{N-1}}, \mathbf{u})}{\partial x_j}\right]$$

is the n×n Jacobian matrix evaluted at $\mathbf{x}_{_{N-1}}$.

Given the initial guess $\mathbf{u}_0$ of the unknown parameter $\mathbf{u}$, at every step of the

iteration we identify a new approximation

$$\mathbf{u}_N = \mathbf{u}_{N-1} + \Delta\mathbf{u}, \tag{2.5}$$

which minimizes the residue $\mathcal{R} = \mathbf{x}(t, \mathbf{u}) - \mathbf{x}_{_N}(t, \mathbf{u}_{_N})$. We have

$$\mathcal{R} = \Delta\mathbf{x}_N(t) - U_N\Delta\mathbf{u} + o(|\Delta\mathbf{u}|), \quad \text{as } |\Delta\mathbf{u}| \to 0, \tag{2.6}$$

where

$$\Delta\mathbf{x}_{_N}(t) = \mathbf{x}(t, \mathbf{u}) - \mathbf{x}_N(t, \mathbf{u}_{N-1}),$$

U_N is an $n \times m$ sensitivity matrix with columns

$$U_N^j = \Big(\frac{\partial\mathbf{x}_N(t, \mathbf{u}_{N-1})}{\partial u^j}\Big), \ j = 1, ..., m.$$

U_N solves the matrix differential system

$$\frac{dU_N}{dt} = \frac{\partial}{\partial\mathbf{u}}\mathbf{f}(t, \mathbf{x}_{N-1}, \mathbf{u}_{N-1}) + \Phi_N + J(t, \mathbf{x}_{_{N-1}}, \mathbf{u}_{N-1})U_N. \tag{2.7}$$

$$U_N(t_0) = 0 \tag{2.8}$$

where Φ_N is an $n \times m$ matrix with elements

$$\phi_N^{ik} = \sum_{j=1}^{n} \frac{\partial^2 f^{\,i}(t, \mathbf{x}_{_{N-1}}, \mathbf{u}_{_{N-1}})}{\partial u^{\,k}\partial x^j}(x^j_{_N} - x^j_{_{N-1}}),$$

To find $\Delta\mathbf{u}$, we minimize the quadratic function

$$\mathcal{J}_N(\Delta\mathbf{u}) = \int_{t_0}^{t_1} (\Delta\mathbf{x}_N - U_N\Delta\mathbf{u})^T(\Delta\mathbf{x}_N - U_N\Delta\mathbf{u})\mathrm{d}t \tag{2.9}$$

We have

$$\mathcal{J}'_N(\Delta\mathbf{u}) = 2\int_{t_0}^{t_1} (U_N^T U_N\Delta\mathbf{u} - U_N^T\Delta\mathbf{x}_N)\mathrm{d}t,$$

$$\mathcal{J}''_N(\Delta\mathbf{u}) = 2\int_{t_0}^{t_1} U_N^T U_N\mathrm{d}t.$$

Therefore, minimum $\Delta\mathbf{u}$ satisfies the following system of linear algebraic equations

$$A_N\Delta\mathbf{u} = P_N, \tag{2.10}$$

where

$$A_N = \int_{t_0}^{t_1} U_N^T U_N\mathrm{d}t = \left(a_N^{ij}\right)_{i,j=1}^{m}$$

is an $m \times m$ symmetric matrix with elements

$$a_N^{ij} = \sum_{k=1}^{n}\int_{t_0}^{t_1} \frac{\partial x_N^k(t,\mathbf{u}_{N-1})}{\partial u^i}\frac{\partial x_N^k(t,\mathbf{u}_{N-1})}{\partial u^j}dt,$$

and

$$P_N = \int_{t_0}^{t_1} (U_N^T\Delta\mathbf{x}_N)\mathrm{d}t = \left(p_N^j\right)_{j=1}^{m}$$

is an m-vector with elements

$$p_N^j = \sum_{k=1}^{n}\int_{t_0}^{t_1} \frac{\partial x_N^k(t,\mathbf{u}_{N-1})}{\partial u^j}\left(x^k(t,\mathbf{u}) - x_N^k(t,\mathbf{u})\right)dt$$

In fact, A_N is a Gram matrix of vectors U_N^j, and

$$a_N^{ij} = (U_N^i, U_N^j)_{L_2^n(t_0,t_1)}$$

where $L_2^n(t_0, t_1)$ is a Hilbert space of vector functions $g : (t_0, t_1) \to \mathbb{R}^n$ with inner product

$$(g, h)_{L_2^n(t_0,t_1)} = \int_{t_0}^{t_1} g^T h dt.$$

It is known [46] that

$$det(A_N) = \Gamma(U_N^1, ..., U_N^m) \geq 0$$

and it is positive, that is to say, A_N is non-singular if and only if the vectors $U_N^j, j = 1, ..., m$ are linearly independent.

The following numerical algorithm is suggested in [38].

2.1.1 Algorithm

1. Initialize $\mathbf{u}_0$, $\mathbf{x}_0(t)$ and set $N = 1$.

2. Find $\mathbf{x}_N(t, \mathbf{u}_{N-1})$ and sensitivity matrix U_N by solving linear problems (2.3), (2.4), (2.7), (2.8).

3. Find $\Delta\mathbf{u}$ by solving linear algebraic equations system (2.10) and update the new value $\mathbf{u}_N$ of the parameter using (2.5).

4. If satisfactory accuracy is achieved, then terminate the process, otherwise replace N with $N + 1$ and go move back to Step 2. As termination criteria,

the smallness of either of the expressions

$$|\mathbf{u}_{N-1} - \mathbf{u}_N|,\ \mathcal{J}_N(\Delta\mathbf{u}),\ ||\mathbf{x}(\cdot) - \mathbf{x}_N(\cdot, \mathbf{u}_N)||_{L_2^n}$$

can be used.

2.1.2 Regularization

In general, the matrix A_N may be ill-conditioned, and to solve the ill-conditioned system (2.10) we implement two types of Tikhonov regularization. Type I regularization is performed by replacing the function (2.9) with

$$||\Delta\mathbf{x}_N - U_N\Delta\mathbf{u}||_{L_2}^2 + \alpha|\Delta\mathbf{u}|^2. \tag{2.11}$$

This yields the following linear system instead of (2.10)

$$(A_N + \alpha I)\Delta\mathbf{u} = P_N \tag{2.12}$$

where I is the identity matrix and α is a regularization parameter. Type II regularization is performed by replacing the function (2.9) with

$$||\Delta\mathbf{x}_N - U_N\Delta\mathbf{u}||_{L_2}^2 + \alpha|\mathbf{u}_{N-1} + \Delta\mathbf{u} - \mathbf{u}^*|^2 \tag{2.13}$$

where $\mathbf{u}^*$ is a known vector with the expectation to be close to the true value of the unknown parameter. The idea of Type II regularization is that the unknown parameter is searched in the neighborhood of $\mathbf{u}^*$. This yields the following linear

system instead of (2.10):

$$(A_N + \alpha I)\Delta \mathbf{u} = P_N + \alpha(\mathbf{u}^* - \mathbf{u}_{N-1}). \tag{2.14}$$

2.1.3 Identifiability

Implementation and convergence of the method is directly connected to identifiability of unknown parameters. In fact, the Gram matrix A_N in (2.10) is the so called Fisher information matrix (FIM) for the linearized ODE system and it characterizes the information content of the experimental data. Singularity of A_N is equivalent to linear dependence of the sensitivity vectors $U_N^j, j = 1, ..., m$, which is the indication of the presence of the non-identifiable parameters . On the contrary, non-singularity of A_N is equivalent to *identifiability* of parameters. If A_N is non-singular, but $detA_N$ is sufficiently small then for computer simulation A_N is treated as a singular matrix [8, 17, 47]. Our two regularization algorithms are developed to address such *practical non-identifiability* cases.

2.2 Results and Discussions

In this section, we present the results of numerical experiments on several canonical models of system biology. Software package *qlopt* is developed to implement the method described in Section 2.1 and posted in GitHub under the GNU General Public License v3.0 [40]. All the model examples considered in this section and basic instructions for their execution are included in *qlopt*. In all examples, we generated simulated measurements by solving the system of nonlinear ODEs using the Runge-Kutta 4 (RK4) method. The computational cost of each iteration consists

of solving a system of $n(1+m)$ linear differential equations for finding $\mathbf{x}_N(t, \mathbf{u})$, U_N, and consequently by solving a system of m linear algebraic equations for $\Delta\mathbf{u}$ in order to update $\mathbf{u}$. We also did experiments with simulated noisy data that was generated using the following formula [14]

$$\mathbf{y} = \mathbf{x}(t, \mathbf{u}) + \sqrt{\rho}\nu \tag{2.15}$$

where the measurement $\mathbf{x}(t, \mathbf{u})$ is generated by using RK4 as explained above, ρ is the noise level and ν is an n-dimensional random variable with standard normal distribution:

$$\nu_i \sim N(0, 1), \text{ for } i = 1, \ldots, n. \tag{2.16}$$

We pursued denoising with algorithm TSQR from [49]. We used the Simpson method for the calculation of the L_2-norm of $\mathbf{x}(t) - \mathbf{x}_N(t, u)$ as the measure of error.

For each noise level we ran 100 experiments and created a box plot for the discrete L_2-norm error distribution. Numerical method was applied to real dataset for the Lotka-Volterra system. For our calculations we used the C/C++, *GSL* [48] and *Eigen* [50] in a Linux environment. *Eigen* was used as our matrix library. *GSL* was used for adding noise to the simulated data. *GSL* and *Eigen* were both used for solving linear algebraic systems.

2.2.1 Lotka-Volterra Model

Consider the classical Lotka-Volterra model describing predator-prey interaction. Given some measurements $\{(x(t_i), y(t_i))\}_{i=0}^{p}$, it is required to estimate parameters

u_1, u_2, u_3 for the system

$$\dot{x} = u_1 x - u_2 xy,\ x(t_0) = x_0$$

$$\dot{y} = u_2 xy - u_3 y,\ y(t_0) = y_0$$

Quaslinearization and differentiation with respect to parameters provide the following system of 8 linear differential equations instead of (2.3),(2.4), (2.7), (2.8).

$$\begin{aligned}
\dot{x}_N =& (u_1 - u_2 y_{N-1}) x_N - u_2 x_{N-1} y_N + u_2 x_{N-1} y_{N-1} \\
\dot{y}_N =& u_2 y_{N-1} x_N - (u_3 - u_2 x_{N-1}) y_N - u_2 x_{N-1} y_{N-1} \\
\dot{x}_{1N} =& (u_1 - u_2 y_{N-1}) x_{1N} - u_2 x_{N-1} y_{1N} + x_N \\
\dot{y}_{1N} =& u_2 y_{N-1} x_{1N} - (u_3 - u_2 x_{N-1}) y_{1N} \\
\dot{x}_{2N} =& (u_1 - u_2 y_{N-1}) x_{2N} + y_{N-1}(x_{N-1} - x_N) - x_{N-1} y_N - u_2 x_{N-1} y_{2N} \\
\dot{y}_{2N} =& u_2 y_{N-1} x_{2N} + (u_2 x_{N-1} - u_3) y_{2N} + y_{N-1}(x_N - x_{N-1}) + x_{N-1} y_N \\
\dot{x}_{3N} =& (u_1 - u_2 y_{N-1}) x_{3N} - u_2 x_{N-1} y_{3N} \\
\dot{y}_{3N} =& u_2 y_{N-1} x_{3N} + (u_2 x_{N-1} - u_3) y_{3N} - y_N
\end{aligned}$$

where $x_{_N}(t_0) = x_0$, $y_{_N}(t_0) = y_0$, $x_{_{iN}}(t_0) = 0$, $y_{_{iN}}(t_0) = 0$, $\frac{\partial x_{_N}}{\partial u_i} = x_{iN}$ and $\frac{\partial y_{_N}}{\partial u_i} = y_{iN}$ for $i = 1, 2, 3$. The typical sample result is demonstrated in Table 2.1: the true value of the parameter is $\mathbf{u} = (1, 2, 1)$, the time interval is $[0, 1]$ is discretized with step-size 0.001, and the initial guess is chosen as $\mathbf{u}_0 = (6, 7, 6)$. Convergence is achieved in 9 steps, and the columns of the Table 2.1 show the values of $\mathbf{x}_N$ and $\mathbf{u}_N$ for $N = 3, 6, 9$. Next we added noise to the measurement according to (2.15). Table 2.2 demonstrates the typical case of convergence with noisy data: the second column is the actual measurement, the third column is a state vector after adding

time	x_1	x_3	x_6	x_9	x
0	1	1	1	1	1
0.2	0.768 257	0.807 069	0.778 768	0.789 416	0.789 416
0.4	0.669 305	0.561 748	0.574 733	0.590 780	0.590 780
0.6	0.638 116	0.471 259	0.414 677	0.433 923	0.433 923
0.8	0.591 143	0.691 000	0.300 779	0.322 706	0.322 706
1	0.553 546	1.138 543	0.223 664	0.248 121	0.248 121
time	y_1	y_3	y_6	y_9	y
0	1	1	1	1	1
0.2	1.034 744	1.216 615	1.177 104	1.171 492	1.171 492
0.4	0.847 512	1.145 686	1.276 765	1.262 684	1.262 684
0.6	0.729 035	0.877 258	1.292 257	1.266 932	1.266 932
0.8	0.674 171	0.619 188	1.242 411	1.205 046	1.205 046
1	0.619 427	0.665 298	1.152 692	1.104 786	1.104 786
	$\mathbf{u}_1$	$\mathbf{u}_3$	$\mathbf{u}_6$	$\mathbf{u}_9$	$\mathbf{u}$
u_1	8.471 601	2.359 123	0.993 071	1	1
u_2	7.532 462	2.437 976	1.991 796	2	2
u_3	5.151 895	1.373 598	0.993 176	1	1

Table 2.1: Numerical results for Lotka-Volterra using the parameters: $\mathbf{u}_0 = (6, 7, 6)$, $\mathbf{u} = (1, 2, 1)$, $\mathbf{x}_0 = (1, 1)$, time interval $[0, 1]$, with step-size 0.001, and tolerance $|\Delta\mathbf{u}| < 0.00001$.

a noise and subsequently removing it using least squares. In Table 2.3, we show the mean and standard deviation of the parameters between 100 convergent runs of the algorithm when the noise level is 5. In Figure 2.1, we constructed a box plot based on 100 simulations for several noise levels. We used the standard box plot where the whiskers are the max and min, the top and bottom of the box are the 75th and 25th percentile, the middle is the median, and the marks outside the box plot are outliers. The box plots demonstrate that the error distribution for each noise level is skewed towards zero. The minimum and median errors grow modestly, while in contrast the maximum error grows much more rapidly.

Convergence of the method without regularization is restrictive in terms of the size of the convergence range for the initial guess $\mathbf{u}_0$. For example, in the model

time	**Actual**	**Noisy**	x_4	x_8	x_{12}
0	35	34.439 600	34.439 680	34.439 680	34.439 680
0.4	40.671 100	40.577 800	36.454 060	40.289 750	40.240 960
0.8	47.155 870	47.224 300	37.818 960	47.025 440	46.909 730
1.2	54.368 860	54.379 000	37.781 280	54.540 020	54.334 350
1.6	62.018 050	62.042 100	35.228 700	62.463 470	62.140 490
2	69.406 730	70.213 400	22.929 640	69.892 940	69.429 100
time	**Actual**	**Noisy**	y_4	y_8	y_{12}
0	4	3.898 269	3.898 269	3.898 269	3.898 269
0.4	4.083 992	3.835 040	7.330 974	3.958 642	3.960 196
0.8	4.441 914	4.309 099	13.866 830	4.318 066	4.320 769
1.2	5.188 602	5.320 444	26.267 370	5.107 948	5.111 061
1.6	6.550 325	6.869 077	49.367 430	6.598 891	6.600 013
2	8.949 686	8.954 998	97.414 010	9.314 478	9.302 844
	u		$\mathbf{u}_4$	$\mathbf{u}_8$	$\mathbf{u}_{12}$
u_1	0.480 000		0.470 200	0.501 210	0.500 970
u_2	0.026 000		0.007 583	0.028 677	0.028 610
u_3	0.930 000		−0.747 350	1.028 940	1.026 720

Table 2.2: Numerical results for Lotka-Volterra using the parameters: $\mathbf{u}_0 = (0.05, 0.05, 0.05)$, $\mathbf{u} = (0.48, 0.026, 0.93)$, $\mathbf{x}_0 = (35, 4)$, time interval $[0, 2]$, with step-size 0.005, with $\rho = 5^2$ and tolerance $|\Delta\mathbf{u}| < 0.00001$.

	Actual	Mean	Deviation
u_1	0.48	0.4784493	0.00784333
u_2	0.026	0.02603669	0.000462745
u_3	0.93	0.9199701	0.05152869

Table 2.3: Mean and standard deviation of 100 convergent runs of Lotka-Volterra with noise level 5.

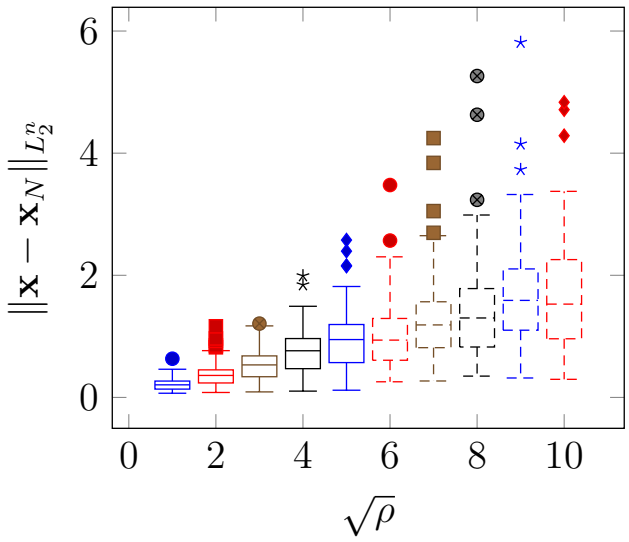

Figure 2.1: Distribution of $\|\mathbf{x}-\mathbf{x}_N\|_{L_2^n}$ for several noise levels of the Lotka-Volterra model.

example with true parameter value $(1, 2, 1)$, the farthest range of the initial guess $\mathbf{u}_0$ was the small neighborhood of $(9, 8, 9)$. To address this deficiency we applied Type II regularization according to (2.14). The typical numerical simulation is demonstrated in Table 2.4, where the initial guess is chosen as $(10, 50, 120)$, and the reference point is $\mathbf{u}^* = (5, 6, 5)$. The regularization parameter is chosen according to the value $\alpha = \lambda|\mathbf{u}^* - \mathbf{u}_{N-1}|$, where λ is a small positive constant.

Next, we analyze the dependence of the accuracy of the method on the number of available measurements. We fix p equidistant time instances $t_i, i = 1, ..., p$ and choose our p measurements as $\mathbf{x}(t_i)$, where $\mathbf{x}$ is a solution of the ODE system with true parameters calculated using RK4. We generate simulated measurement by replacing $\mathbf{x}$ with piecewise linear interpolation of the points $(t_i, \mathbf{x}(t_i)), i = 0, 1, ..., p$ and pursue our numerical method until termination step N. Typical result is demonstrated in Figure 2.2 and Table A.1. Dependence of error $||\mathbf{x} - \mathbf{x}_N||_{L_2^n}$ vs. p forms L-shape curve where edge point corresponds to number of measurements

time	x_3	x_6	x_9	x
0	1	1	1	1
0.2	0.876 344	0.795 198	0.795 004	0.789 416
0.4	0.535 542	0.594 250	0.594 326	0.590 780
0.6	0.291 609	0.435 457	0.435 729	0.433 923
0.8	0.217 213	0.325 082	0.325 341	0.322 706
1	0.240 489	0.253 327	0.253 431	0.248 121
time	y_3	y_6	y_9	y
0	1	1	1	1
0.2	1.465 193	1.183 012	1.182 105	1.171 492
0.4	1.697 322	1.277 765	1.276 432	1.262 684
0.6	1.516 675	1.274 472	1.273 428	1.266 932
0.8	1.159 809	1.198 557	1.198 238	1.205 046
1	0.857 457	1.083 544	1.084 022	1.104 786
	$\mathbf{u}_3$	$\mathbf{u}_6$	$\mathbf{u}_9$	$\mathbf{u}$
u_1	2.151 610	1.245 434	1.246 370	1
u_2	3.209 686	2.181 702	2.182 409	2
u_3	1.744 912	1.126 209	1.126 675	1
α	$2.576\,325 \times 10^{-5}$	$5.931\,209 \times 10^{-5}$	$5.947\,253 \times 10^{-5}$	0

Table 2.4: Numerical results for Lotka-Volterra with regularization using the parameters: $\mathbf{u}_0 = (10, 50, 120)$, $\mathbf{u} = (1, 2, 1)$, $\mathbf{u}^* = (5, 6, 5)$, $\mathbf{x}_0 = (1, 1)$, time interval $[0, 1]$, with step-size 0.01, and tolerance $|\Delta \mathbf{u}| < 0.00001$.

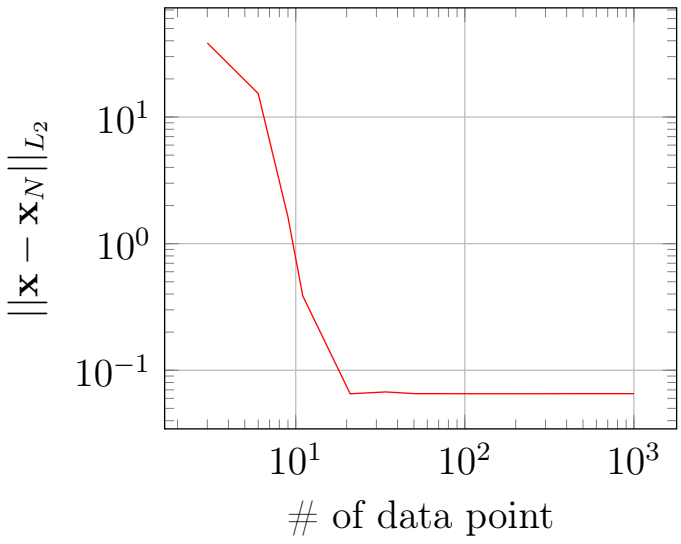

Figure 2.2: Graph of the number of data points vs. $||\mathbf{x} - \mathbf{x}_N||_{L_2}$ for the Lotka Volterra model. The parameters are $\mathbf{u}_0 = (0.05, 0.05, 0.05)$, $\mathbf{u} = (0.48, 0.026, 0.93)$, $\mathbf{y}_0 = (35, 4)$, time interval $[0, 2]$ and respective tolerances were set at 1×10^{-7}.

$p = 21$ with error of order 10^{-2}. The error almost doesn't change for $p > 21$, and it starts increasing when $p < 21$. Algorithm still converges for $p = 11$ or $p = 9$, but with very rough error. There is no convergence for $p < 9$.

Lastly, we considered the classic lynx-hare dataset to test the algorithm on real data. The data is the population of the Canadian lynx and hare from Hudson Bay company over twenty years given in Table A.2, where the population size is in the unit of 1000s. We apply the method in three successive steps. In the first step, we apply the method in the interval $[0, 1]$ by applying linear interpolation to the data values for the first two years. The result is demonstrated in Table A.3. We choose the final value of the parameter as the initial guess and apply the method to the whole set of data, first in the interval $[0, 5]$ (Table A.4), and then in the interval $[0, 20]$ (Table A.5). The method achieves excellent matching: in Figure 2.3 we compare the actual data with the curve generated by solving the original system

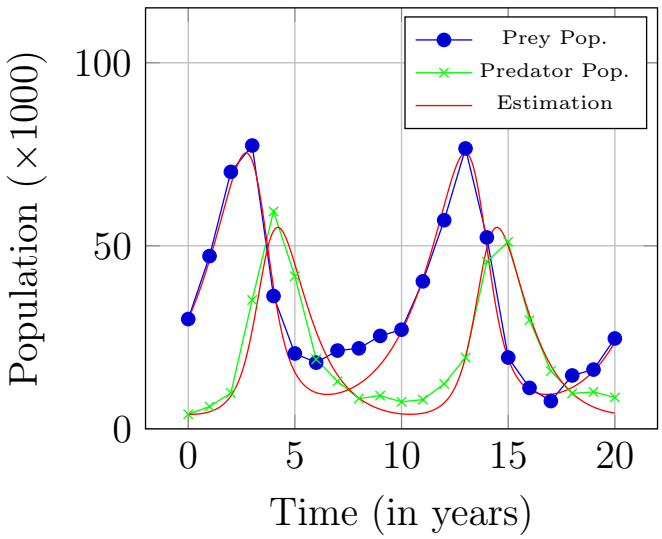

Figure 2.3: Graph of the data and the solution of the ODE with approximated parameters of $\mathbf{u} = (0.5520907, 0.02990494, 0.02602782, 0.8700797)$.

of ODEs with the parameters discovered using our algorithm. We also split up the time interval into a training and test set. For training we considered the interval $t \in [0, 5]$. The comparison to the entire interval including the test data, i.e. $t \in [6, 20]$ can be found in Figure 2.4. The estimation projected from the test set is skewed to the right when compared to figure 2.3, but still follows the general trend. Also the projection along the test set appears more accurate for the prey population than for the predator population.

The algorithm described in section 2.1.1 demonstrates superlinear convergence close to quadratic. Typical result for the estimation of the numerical convergence rate is demonstrated in Figure 2.5. The estimation of the rate of convergence is $r = 1.9149$. This shows that we are almost retaining theoretical quadratic convergence rate in our computations.

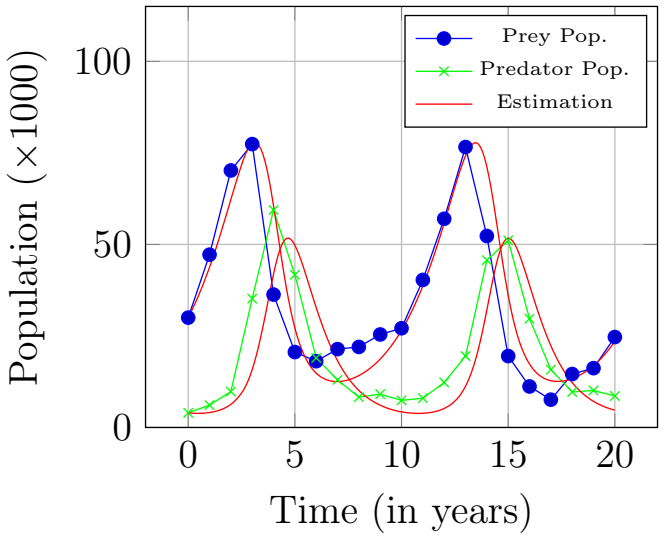

Figure 2.4: Graph of the data and the solution of the ODE. Data from the time interval $[0, 5]$ was used for determining the parameters which were approximated to be $\mathbf{u} = (0.5661793, 0.02739231, 0.02908555, 1.007009)$.

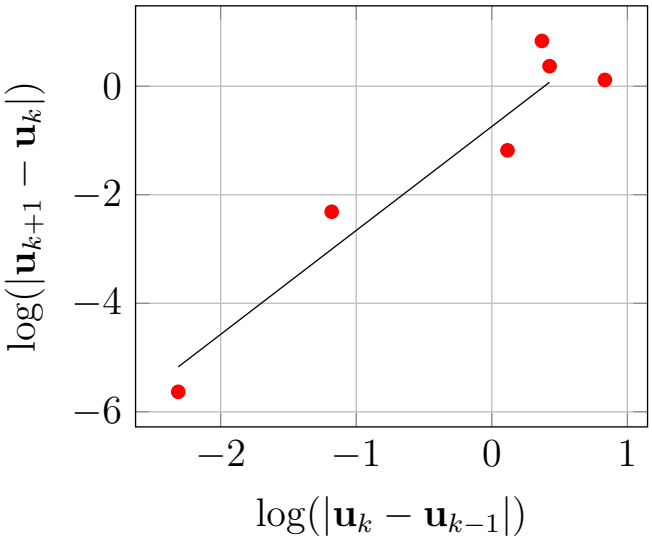

Figure 2.5: A convergence rate graph for Lotka Volterra using the parameters: $\mathbf{u}_0 = (6, 7, 6)$, $\mathbf{u} = (1, 2, 1)$, $\mathbf{x}_0 = (1, 1)$, time interval $[0, 1]$, with step-size 0.001, and tolerance $|\Delta\mathbf{u}| < 0.00001$. The convergence rate is $r = 1.9149$ and $C = .18077$.

2.2.2 Genetic Circuit Model with Bistable Switch

In this section we consider the simplest model for genetic circuit with a bistable switch introduced in [51]. Proteins A and B repress the production of each other, and their competitive dynamic has two stable equilibrium states corresponding to low and high states, similar to electric circuits. By adding an inducer I one can suppress the activity of the protein B on behalf of protein A, and potentially produce a biological circuit which inherits bistable switch mechanisms of electrical circuits. This mechanism was implemented in [51] where synthetic bistable switch was demonstrated experimentally in Escherichia coli.

The mathematical model of the bistable switch in the genetic network is given by [14, 51]

$$\begin{aligned} \dot{p}_A &= \frac{\alpha}{1+(up_B)^n} - p_A & p_A(t_0) &= p_{A_0} \\ \dot{p}_B &= \frac{\alpha}{1+p_A^n} - p_B & p_B(t_0) &= p_{B_0} \end{aligned}$$

where p_A, p_B are concentrations of proteins A and B, α is effective production rate, n is the cooperativity and u is an input. The bistable switch is controlled by an inducer I, which represses the activity of the protein B, by decreasing the parameter u according to the formula

$$u = \frac{\beta}{1+\left(\frac{I}{K}\right)^m}$$

where β, K, m are known parameters depending on the type of inducer.

Our goal is to apply the method to identify the parameters (α, u, n) given some measurements for the concentrations of p_A and p_B. By implementing linearization

and differentiation with respect to the parameters, we derive the following system of eight linear differential equations instead of (2.3),(2.4) and (2.7),(2.8):

$$
\begin{aligned}
\dot{p}_{A_N} &= -p_{A_N} - \frac{\alpha n u^n p_{B_{N-1}}^{n-1}}{g^2} p_{B_N} + \frac{\alpha n u^n p_{B_{N-1}}^n + g}{g^2} \\
\dot{p}_{B_N} &= -\frac{\alpha n p_{A_{N-1}}^{n-1}}{d^2} p_{A_N} - p_{B_N} + \frac{\alpha n p_{A_{N-1}}^n + d}{d^2} \\
\dot{p}_{A_{1N}} &= -p_{A_{1N}} - \frac{n u^n p_{B_{N-1}}^{n-1}}{g^2} p_{B_N} - \frac{\alpha n u^n p_{B_{N-1}}^{n-1}}{g^2} p_{B_{1N}} + \frac{n u p_{B_{N-1}}^n + g}{g^2} \\
\dot{p}_{B_{1N}} &= -\frac{n p_{A_{N-1}} n - 1}{d^2} p_{A_N} - \frac{\alpha n p_{A_{N-1}}^{n-1}}{d^2} p_{A_{1N}} - p_{B_{1N}} + \frac{n p_{A_{N-1}}^n + d}{d^2} \\
\dot{p}_{A_{2N}} &= -p_{A_{2N}} - \left(\frac{\alpha n^2 \left(u p_{B_{N-1}}\right)^{n-1}}{g^2} - \frac{2\alpha n^2 \left(u p_{B_{N-1}}\right)^{2n-1}}{g^3} \right) p_{B_N} - \\
&\quad \frac{\alpha n u^n p_{B_{N-1}}^{n-1}}{g^2} p_{B_{2N}} + \frac{\alpha u^{n-1} p_{B_{N-1}}^n (n^2 - n)}{g^2} - \frac{2\alpha n^2 u^{2n-1} p_{B_{N-1}}^{2n}}{g^3} \\
\dot{p}_{B_{2N}} &= -\frac{\alpha n p_{A_{N-1}}^{n-1}}{d^2} p_{A_{2N}} - p_{B_{2N}} \\
\dot{p}_{A_{3N}} &= -p_{A_{3N}} - \frac{\alpha u p_{B_{N-1}}^n}{p_B g^2} \left(\left(1 - \frac{2u p_{B_{N-1}}^n}{g^2} \right) \log(u p_{B_{N-1}}) p_{B_N} + n p_{B_{3N}} \right) \\
&\quad - \frac{\alpha u p_{B_{N-1}}^{n-1}}{g^2} p_{B_N} + \frac{\alpha u p_{B_{N-1}}}{g^2} \left(1 + n \log(u p_B) - \frac{2n \log(u p_{B_{N-1}})}{g u p_{B_{N-1}}^n} \right) \\
&\quad - \frac{\alpha u p_{B_{N-1}}^n \log(u p_{B_{N-1}})}{g^2} \\
\dot{p}_{B_{3N}} &= -\frac{\alpha p_{A_{N-1}}^{n-1}}{d^2} \left(\left(1 + n \log(p_A) - \frac{2n p_{A_{N-1}} n \log(p_A)}{d} \right) p_{A_N} + n p_{A_{3N}} \right) \\
&\quad - p_{B_{3N}} + \frac{\alpha p_{A_{N-1}} n}{d^2} \left(1 + (n-1) \log(p_A) - \frac{2n p_{A_{N-1}} n \log(p_A)}{d} \right) \qquad (2.17)
\end{aligned}
$$

with initial conditions $p_{A_N}(t_0) = p_{A_0}$, $p_{B_N}(t_0) = p_{B_0}$, $p_{A_{iN}}(t_0) = 0$, and $p_{B_{iN}}(t_0) = 0$ for $i = 1, 2, 3$, where $p_{A_{1N}} = \frac{\partial p_{A_N}}{\partial \alpha}$, $p_{A_{2N}} = \frac{\partial p_{A_N}}{\partial u}$, $p_{A_{3N}} = \frac{\partial p_{A_N}}{\partial n}$, similarly for $p_{B_{iN}}$, and $g = 1 + (u p_B)^n$ and $d = 1 + {p_A}^n$.

Tables 2.5 and 2.6 demonstrate typical results on the convergence of the algo-

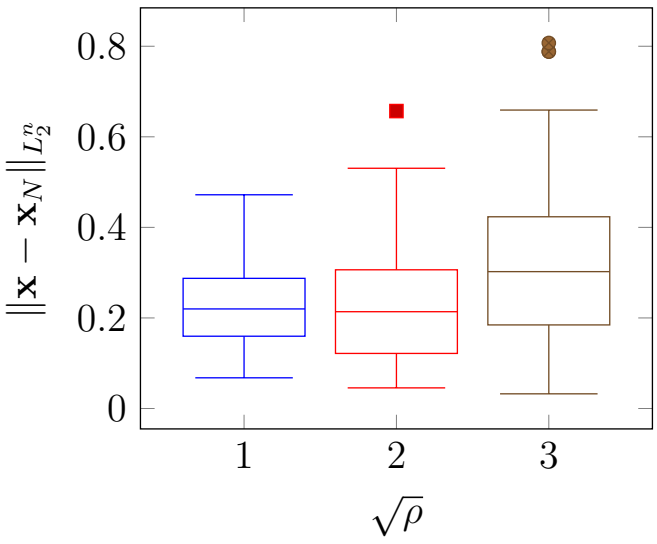

Figure 2.6: Distribution of $\|\mathbf{x}-\mathbf{x}_N\|_{L_2^n}$ for several noise levels of the bistable switch model.

rithm with and without noise, respectively. In Table 2.7, we show the mean and standard deviation of the parameters between 100 convergent runs of the algorithm when the noise level is 3. In Figure 2.6, we constructed a box plot based on 100 simulations for several noise levels We summarized the results of 100 runs when the noise level was 5 in Table 2.7. In Figure 2.6, there is a box plot to represent the distribution of L_2-norm errors for several noise levels. As in the case of Lotka-Volterra model, convergence of the method without regularization is restrictive in terms of the size of the convergence range for the initial guess $\mathbf{u}_0$. To address this issue, we applied Type II regularization according to (2.14). The typical numerical simulation is demonstrated in Table 2.8, where initial guess is chosen as $(200, 1, 1)$, and the reference point is $\mathbf{u}^* = (135, 2, 1.5)$ with time interval $[0, 3]$.

Lastly, we analyze the dependence of the accuracy of the method on the number of available measurements. Figure 2.7 and Table A.6 present L-shape curve expressing dependence of the error $||\mathbf{x}-\mathbf{x}_N||_{L_2^n}$ vs. p for one typical example. Edge

time	p_{A4}	p_{A7}	p_{A11}	p_A
0	14	14	14	14
0.2	11.512 000	11.482 000	11.481 000	11.481 000
0.4	9.498 500	9.428 200	9.427 400	9.427 400
0.6	7.881 000	7.757 400	7.756 200	7.756 200
0.8	6.592 900	6.401 400	6.399 600	6.399 600
1	5.572 000	5.301 600	5.299 100	5.299 100
time	p_{B4}	p_{B7}	p_{B11}	p_B
0	13	13	13	13
0.2	10.797 000	10.814 000	10.814 000	10.814 000
0.4	9.072 100	9.105 700	9.105 800	9.105 800
0.6	7.775 600	7.826 700	7.826 800	7.826 800
0.8	6.877 000	6.951 600	6.951 800	6.951 800
1	6.359 700	6.479 400	6.479 900	6.479 900
	$\mathbf{u}_4$	$\mathbf{u}_7$	$\mathbf{u}_{11}$	$\mathbf{u}$
α	159.740 000	150.010 000	150	150
u	2.433 100	3.199 600	3.200 000	3.200 000
n	2.037 500	2	2	2

Table 2.5: Numerical results for the Bistable Switch model using the paramters: $\mathbf{u}_0 = (125, 1.85, 1)$, $\mathbf{u} = (150, 3.2, 2)$, $\mathbf{x}_0 = (14, 13)$, time interval $[0, 1]$, with step-size 0.001, and tolerance $|\Delta\mathbf{u}| < .00001$.

point corresponds to number of measurement $p = 15$ with error of order 10^{-3}. The error almost doesn't change for $p > 15$, and it starts slowly increasing when $p < 15$. For a number of measurements $p = 7$ there is still convergence, but the error is increased to 0.008.

Nuumerical convergence rate of the algorithm is superlinear and close to theoretical quadratic convergence. Typical result for the calculation of the convergence rate is demonstrated in Figure 2.8. Having approximated the rate of convergence for the numerical results we get $r = 1.8341$.

time	Actual	Noisy	p_{A3}	p_{A5}	p_{A7}
0	25	24.645 870	24.645 870	24.645 870	24.645 870
0.6	13.741 590	14.929 890	13.578 490	13.554 950	13.554 890
1.2	7.605 100	9.793 819	7.593 733	7.525 757	7.525 555
1.8	4.302 712	6.243 021	4.424 316	4.304 100	4.303 672
2.4	2.465 823	4.362 626	2.638 810	2.501 320	2.500 772
3	1.384 659	4.258 046	1.527 525	1.414 724	1.414 286
time	Actual	Noisy	p_{B3}	p_{B5}	p_{B7}
0	25	24.735 540	24.735 540	24.735 540	24.735 540
0.6	13.940 850	16.474 830	13.883 640	13.802 190	13.802 090
1.2	8.369 862	13.323 380	8.536 980	8.316 489	8.316 302
1.8	6.824 716	13.106 440	7.212 268	6.845 344	6.845 621
2.4	10.095 050	14.829 430	10.182 590	10.123 050	10.127 020
3	21.865 080	10.111 580	19.846 010	21.819 480	21.837 120
	$\mathbf{u}$		$\mathbf{u}_3$	$\mathbf{u}_5$	$\mathbf{u}_7$
α	150		154.806 800	154.648 500	154.652 600
u	3.200 000		2.674 066	2.759 770	2.759 640
n	2		2.007 254	2.010 426	2.010 442

Table 2.6: Numerical results for the Bistable Switch model using the paramters: $\mathbf{u}_0 = (125, 1.85, 1)$, $\mathbf{u} = (150, 3.2, 2)$, $\mathbf{x}_0 = (25, 25)$, time interval $[0, 3]$, with step-size 0.005, with $\rho = 3^2$ and tolerance $|\Delta\mathbf{u}| < 0.00001$.

	Actual	Mean	Deviation
α	150	151.0282	2.648399
u	3.2	2.712192	0.1770636
n	2	1.995046	0.01826305

Table 2.7: Mean and standard deviation of 100 convergent runs of the Bistable Switch model with noise level 3.

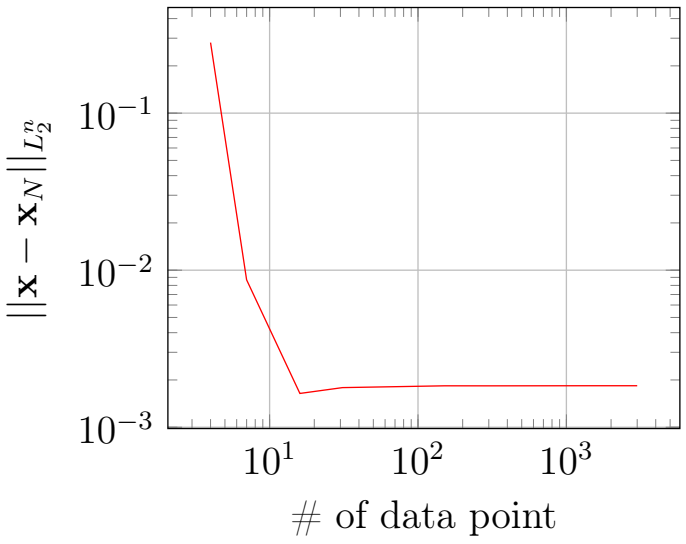

Figure 2.7: Graph of the number of data points vs. $||\mathbf{x} - \mathbf{x}_N||_{L_2^n}$ for the bistable switch model. The parameters were $\mathbf{u} = (150, 3.2, 2.0)$, $\mathbf{u}_0 = (125, 2.2, 1.0)$, $\mathbf{x}_0 = (14, 13)$, $t \in (0, 3)$ and respective tolerances were set at 1×10^{-7}.

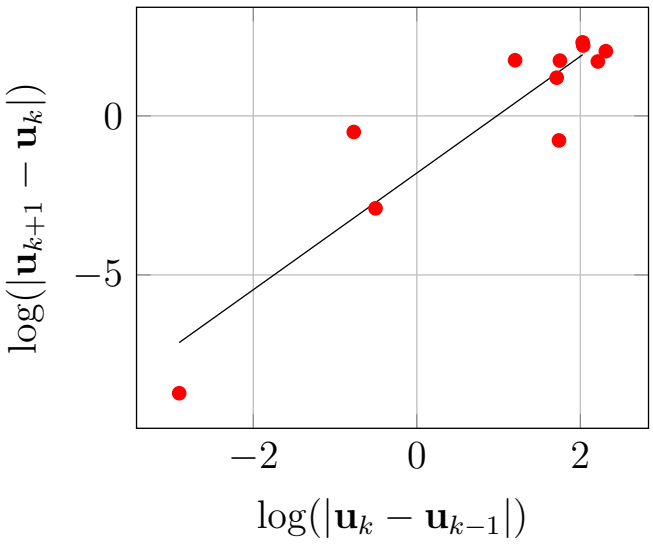

Figure 2.8: A convergence rate graph for the Bistable Switch model using the paramters: $\mathbf{u}_0 = (125, 1.85, 1)$, $\mathbf{u} = (150, 3.2, 2)$, $\mathbf{x}_0 = (14, 13)$, time interval $[0, 1]$, with step-size 0.001, and tolerance $|\Delta \mathbf{u}| < .00001$. The convergence rate is $r = 1.8341$ and $C = .015981$.

time	p_{A4}	p_{A8}	p_{A12}	p_A
0	25	25	25	25
0.6	14.122 785	13.744 919	13.737 542	13.741 591
1.2	8.593 628	7.615 866	7.590 242	7.605 100
1.8	5.722 691	4.326 141	4.268 262	4.302 712
2.4	4.450 640	2.496 073	2.425 694	2.465 823
3	3.941 657	1.408 635	1.356 555	1.384 659
time	p_{B4}	p_{B8}	p_{B12}	p_B
0	25	25	25	25
0.6	14.532 557	13.945 219	13.943 890	13.940 851
1.2	9.926 217	8.375 125	8.374 407	8.369 862
1.8	8.973 079	6.793 261	6.827 833	6.824 716
2.4	7.117 085	9.829 680	10.109 903	10.095 046
3	−17.368 235	20.778 787	21.793 320	21.865 075
	$\mathbf{u}_4$	$\mathbf{u}_8$	$\mathbf{u}_{12}$	$\mathbf{u}$
α	277.632 432	143.704 237	144.053 210	150
u	1.568 518	3.501 743	3.624 186	3.200 000
n	2.057 920	1.979 541	1.980 935	2

Table 2.8: Numerical results for the Bistable Switch model using the parameters: $\mathbf{u}_0 = (200, 1, 1)$, $\mathbf{u} = (150, 3.2, 2)$, $\mathbf{x}_0 = (25, 25)$, $\mathbf{u}^* = (135, 2.0, 1.5)$, time interval $[0, 3]$, with $\alpha = .00005$ with step-size 0.01, and tolerance $|\Delta\mathbf{u}| < .00001$.

2.2.3 Lorenz System

The following model is referred to as the the Lorenz model. It was initially constructed to model atmospheric dynamics and has been integral in the study of chaos [52].

$$\begin{aligned}\dot{x} &= \sigma(y - x)\\ \dot{y} &= rx - xz - y\\ \dot{z} &= xz - bz\end{aligned}$$

The parameter choice was $(\sigma, r, b) = (10, 28, 2.66)$, because in the region the system is sensitive to perturbation of the initial condition. Since this system has three parameters and three equations after linearizing and deriving the sensitivity equations we end up with a system of 12 linear equations:

$$
\begin{aligned}
\dot{x}_N &= \sigma(y_N - x_N) \\
\dot{y}_N &= (r - z_{N-1})x_N - x_{N-1}(z_N - z_{N-1}) - y_N \\
\dot{z}_N &= y_{N-1}x_N + x_{N-1}y_N - bz_N - y_{N-1}x_{N-1} \\
\dot{x}_{1N} &= \sigma(y_{1N} - x_{1N}) + y_N - x_N \\
\dot{y}_{1N} &= (r - z_{N-1})x_{1N} - x_{N-1}z_{1N} - y_{1N} \\
\dot{z}_{1N} &= y_{N-1}x_{1N} + x_{N-1}y_{1N} - bz_{1N} \\
\dot{x}_{2N} &= \sigma(y_{2N} - x_{2N}) \\
\dot{y}_{2N} &= (r - z_{N-1})x_{2N} + x_N - x_{N-1}z_{2N} - y_{2N} \\
\dot{z}_{2N} &= y_{N-1}x_{2N} + x_{N-1}y_{2N} - bz_{2N} \\
\dot{x}_{3N} &= \sigma(y_{3N} - x_{3N}) \\
\dot{y}_{3N} &= (r - z_{N-1})x_{3N} - x_{N-1}z_{3N} - y_{3N} \\
\dot{z}_{3N} &= y_{N-1}x_{3N} + x_{N-1}y_{3N} - bz_{3N} - z_N
\end{aligned}
$$

where $x_{_N}(t_0) = x_0$, $y_{_N}(t_0) = y_0$, $x_{_{iN}}(t_0) = 0$, $y_{_{iN}}(t_0) = 0$, and x_{iN}, y_{iN} for $i = 1, 2, 3$ are the partials with respect to σ, r, b respectively.

In Table A.7 we see the convergence of algorithm without noise, like previous result it converges with quickly to the true solution. Next we add noise and run the algorithm again. We converge but with noise in final parameter, Table A.8.

	Actual	Mean	Deviation
σ	10	10.32386	6.250791
r	28	27.2093	0.787697
b	2.666	2.646076	0.3809348

Table 2.9: Mean and standard deviation of 100 convergent runs of the Lorenz system with noise level 1. Numerical results used the following parameters: $\mathbf{u}_0 = (5, 15, 5)$, $\mathbf{u} = (10, 28, 2.666)$, $\mathbf{y}_0 = (1, 1, 1)$, time interval $[0, 2]$, with step-size 0.005, and tolerance $|\Delta\mathbf{u}| < 0.00001$.

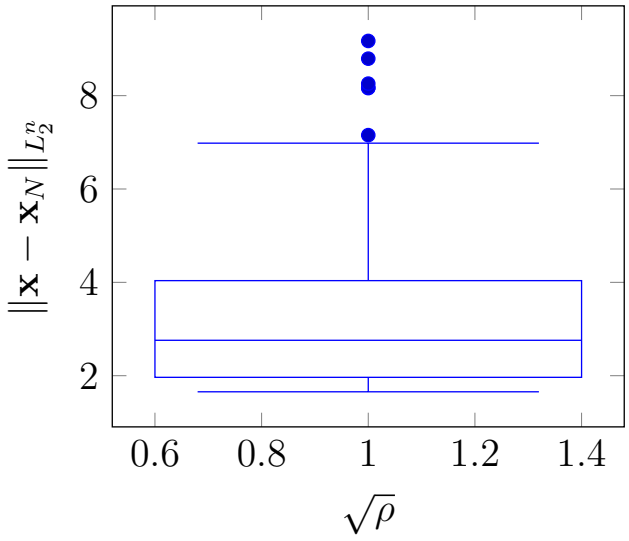

Figure 2.9: Distribution of $\|\mathbf{x} - \mathbf{x}_N\|_{L_2^n}$ for the noisy Lorenz runs. These results used the following parameters: $\mathbf{u}_0 = (5, 15, 5)$, $\mathbf{u} = (10, 28, 2.666)$, $\mathbf{y}_0 = (1, 1, 1)$, time interval $[0, 2]$, with step-size 0.005, and tolerance $|\Delta\mathbf{u}| < 0.00001$

In Table 2.9 we show the distribution of the mean and standard deviation for the problem when $\rho = 1$ and we have the boxplot to show the error distribution 2.9. The system is in the chaotic region, noise is added to initial condition and the method converges regardless.

2.2.4 Synthetic Biology Toggle Switch Model

In this section, we consider the synthetic biology toggle switch model, which describes gene transcription regulation by repressors [16]:

$$\begin{aligned}
\dot{m}_A &= c_A \frac{K_B^{n_B}}{K_B^{n_B} + p_B^{n_B}} - d_A m_A \\
\dot{p}_A &= f_A m_A - g_A p_A \\
\dot{m}_B &= c_B \frac{K_A^{n_A}}{K_A^{n_A} + p_A^{n_A}} - d_B m_B \\
\dot{p}_B &= f_B m_B - g_B p_B
\end{aligned}$$

where m_A and m_B represent the concentration of messenger RNA (mRNA) obtained through transcription of the considered gene; p_A, p_B are the concentrations of proteins A and B obtained through translation of the mRNA, at time t; the parameters of the model are the maximal transcription rates with respect to each protein c_A, c_B, protein degradation rates g_A, g_B, the repression coefficients K_A, K_B, Hill coefficients - cooperativity constants n_A, n_B, the mRNA degradation rates d_A, d_B and protein translation rates f_A, f_B.

We consider several algebraic variations of this system. Note that the parameters n_A, n_B are correlated with K_A, K_B. For the first set of experiments, we assume that the parameters $n_A, n_B = 2$ are fixed and we identify 10 parameters. Linearization and differentiation with respect to 10 parameters imply a system of 44 linear differential equations. We implement both Type I and Type II regularization algorithms in different configuration. Typical simulation results with two regularization methods according to (2.11),(2.12) and (2.13),(2.14) with a constant regularization

time	p_{A40}	p_{A82}	p_{A124}	p_{A166}	p_A
0	20	20	20	20	20
1.2	4.599 706	4.600 191	4.600 215	4.600 216	4.600 216
2.4	1.794 548	1.792 042	1.791 918	1.791 910	1.791 909
3.6	1.196 525	1.199 149	1.199 269	1.199 278	1.199 278
4.8	0.934 380	0.934 907	0.934 924	0.934 925	0.934 925
6	0.762 078	0.758 429	0.758 245	0.758 232	0.758 231
time	p_{B40}	p_{B82}	p_{B124}	p_{B166}	p_B
0	20	20	20	20	20
1.2	6.207 289	6.206 958	6.206 943	6.206 941	6.206 941
2.4	2.984 701	2.985 839	2.985 896	2.985 900	2.985 900
3.6	3.063 985	3.064 418	3.064 444	3.064 446	3.064 446
4.8	3.523 094	3.521 537	3.521 469	3.521 465	3.521 464
6	3.895 586	3.897 012	3.897 087	3.897 093	3.897 093
$\mathbf{u}$	$\mathbf{u}_{40}$	$\mathbf{u}_{82}$	$\mathbf{u}_{124}$	$\mathbf{u}_{166}$	$\mathbf{u}_{true}$
c_A	4.517 128	4.972 712	4.998 045	4.999 858	5
d_A	3.000 186	3.000 011	3.000 001	3	3
K_B	2.150 777	2.007 698	2.000 549	2.000 040	2
f_A	1.000 832	1.000 042	1.000 003	1	1
g_A	4.999 763	4.999 990	4.999 999	5	5
c_B	5.998 470	5.999 934	5.999 995	6	6
d_B	3.999 548	3.999 979	3.999 998	4	4
K_A	−2.000 173	−2.000 003	−2	−2	2
f_B	0.999 934	0.999 997	1.000 000	1	1
g_B	4.999 604	4.999 982	4.999 999	5	5

Table 2.10: Numerical results for the synthetic biology toggle switch model using the paramters: $\mathbf{u}_0 = (2, 1, 1, 1, 1, 2, 1, 1, 1, 2)$, $\mathbf{u} = (5, 3, 2, 1, 5, 6, 4, 2, 1, 5)$, $\mathbf{x}_0 = (20, 20, 20, 20)$, time interval $[0, 6]$, with step-size 0.005, using regularization constant parameters $\alpha = 0.005$ and tolerance $|\Delta\mathbf{u}| < .00001$.

parameter are demonstrated in Table 2.10 and Table A.9. Note that Type II regularization did not increase the range of convergence, but it drastically reduced the number of iterations. Notice that we have an example of non-uniqueness in the parameter solution, since the parameter $n_A = 2$ allows for both of $K_A = 2, -2$ to be possible solutions.

For the second set of experiments, we combined the K^n terms into one term

and simplified the Hill function terms in mRNA equations as

$$c\frac{K^n}{K^n + p^n} = \frac{c}{1 + \frac{p^n}{K^n}}.$$

Note that this transformation changes the sensitivity matrix system (2.7), which is obtained by differentiating the linearized system (2.3) with respect to the parameters. Table A.10 presents the results obtained by implementing Type I regularization (2.11),(2.12) with constant regularization parameter for the identification of 10 parameters in new configuration, and by assigning $n_A = n_B = 2$. Table 2.11 demonstrates the typical case of convergence with noisy data: the second column is the actual measurement; third column is a state vector after adding noise and subsequently removing it using the least squares method. Finally, we consider the same configuration of parameters, but without fixing parameters n_A and n_B. Therefore, we are considering the problem of identification of 12 parameters. Linearization and differentiation with respect to 12 parameters imply a system of 52 linear differential equations in place of (2.3),(2.7). Table A.11 demonstrates the typical result by implementing Type I regularization with constant regularization parameter. Note that for each table we only show the concentrations of the proteins.

In first two sets of experiments, we implemented Tikhonov regularization with constant regularization parameter. In the third set of experiments, we aim to demonstrate that the optimal choice of the regularization parameter α can significantly improve the convergence rate and computational cost of the method. We make the following adjustment to the first experiment (Table 2.10). We ran the algorithm for $N - 1$ iterations with constant α as before, and then plot the

time	**Actual**	**Noisy**	p_{A210}	p_{A420}	p_{A630}
0	20	19.767 600	19.767 600	19.767 600	19.767 600
1.2	4.600 216	4.657 116	4.632 734	4.632 850	4.632 877
2.4	1.791 909	1.813 596	1.795 661	1.795 078	1.794 942
3.6	1.199 278	1.243 556	1.185 361	1.185 916	1.186 046
4.8	0.934 925	0.737 109	0.920 608	0.920 702	0.920 724
6	0.758 231	1.041 570	0.744 255	0.743 378	0.743 172
time	**Actual**	**Noisy**	p_{B210}	p_{B420}	x_{630}
0	20	19.955 270	19.955 270	19.955 270	19.955 270
1.2	6.206 941	6.229 592	6.245 389	6.245 323	6.245 307
2.4	2.985 900	3.008 225	3.002 313	3.002 485	3.002 525
3.6	3.064 446	3.125 816	3.061 767	3.061 865	3.061 888
4.8	3.521 464	3.310 081	3.509 469	3.509 230	3.509 174
6	3.897 093	4.240 228	3.882 091	3.882 603	3.882 722
$\mathbf{u}$	$\mathbf{u}$		$\mathbf{u}_{210}$	$\mathbf{u}_{420}$	$\mathbf{u}_{630}$
c_A	5		4.952 746	5.085 514	5.117 724
d_A	3		2.819 124	2.819 180	2.819 192
$K_B^{n_B}$	4		3.935 764	3.793 802	3.760 892
f_A	1		0.990 649	0.990 462	0.990 419
g_A	5		4.756 149	4.756 322	4.756 361
c_B	6		5.854 101	5.854 821	5.854 987
d_B	4		3.996 556	3.996 934	3.997 022
$K_A^{n_A}$	4		4.117 262	4.115 939	4.115 637
f_B	1		0.992 053	0.992 057	0.992 058
g_B	5		4.992 462	4.992 844	4.992 933

Table 2.11: Numerical results for the synthetic biology toggle switch model using the paramters: $\mathbf{u}_0 = (2,2,2,2,2,2,2,2,2,2)$, $\mathbf{u} = (5,3,4,1,5,6,4,4,1,5)$, $\mathbf{x}_0 = (20,20,20,20)$, $\alpha = 0.005$, time interval $[0,6]$, with step-size 0.005, $\alpha = 0.00001$, $\rho = 1$, and tolerance $|\Delta\mathbf{u}| < .0001$.

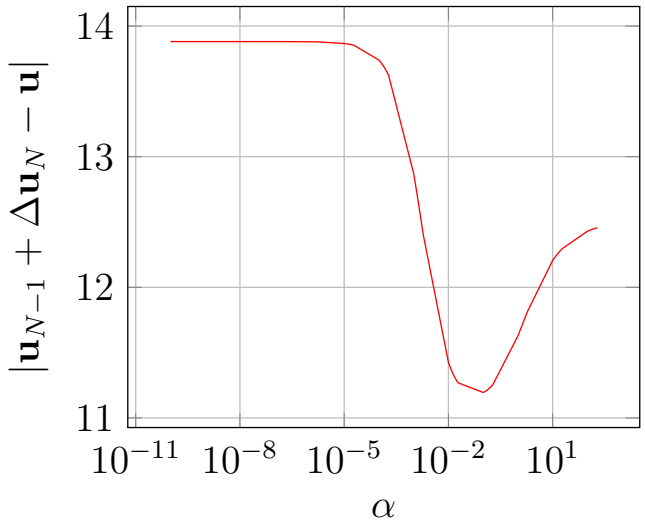

Figure 2.10: Graph of the α vs. $|\mathbf{u}_{N-1} + \Delta\mathbf{u}_N - \mathbf{u}|$ for Toggle Switch at iteration 6.

α vs $|\mathbf{u}_{N-1} + \Delta\mathbf{u}_N - \mathbf{u}|$ graphs within Nth step for the purpose of finding the best possible value of α. Typical graphs are demonstrated in Figure 2.10 for the case $N = 6$, and in Figure 2.11 for the case $N = 20$. Both curves have a clear minimum, which is the best choice of the regularization parameter α. Next, we apply modified Type I regularization, when regularization parameter α is chosen as a minimum in α vs $|\mathbf{u}_{N-1} + \Delta\mathbf{u}_N - \mathbf{u}|$ dependence. Convergence is achieved in only 9 steps as demonstrated in Table 2.12. Since the cost of calculating α is high, we also tested another modified algorithm, where we chose the optimal value of α every 5th iteration. As demonstrated in Table A.12, convergence is achieved in 15 iterations instead of 9.

Lastly, we analyze the dependence of the accuracy of the method on the number of available measurements. Figure 2.12 and Table A.13 present L-shape curve expressing dependence of error $||\mathbf{x} - \mathbf{x}_N||_{L_2^n}$ vs. p for one typical example. Edge point corresponds to number of measurements $p = 31$ with error of order 10^{-4}.

time	x_1	x_3	x_6	x_9	x
0	20	20	20	20	20
3.6	2.747 729	0.916 567	1.143 766	1.199 278	1.199 278
6	0.754 578	0.573 116	0.690 266	0.758 231	0.758 231
time	x_1	x_3	x_6	x_9	x
0	20	20	20	20	20
3.6	0.130 388	3.173 090	3.100 629	3.064 446	3.064 446
6	0.249 408	3.338 705	4.006 661	3.897 093	3.897 093
$\mathbf{u}$	$\mathbf{u}_1$	$\mathbf{u}_3$	$\mathbf{u}_6$	$\mathbf{u}_9$	$\mathbf{u}_{true}$
c_A	2.088 324	3.299 465	4.946 468	5	5
d_A	0.174 888	2.948 731	3.000 079	3	3
K_B	1.308 159	2.798 753	2.016 641	2	2
f_A	0.917 620	1.001 334	1.000 166	1	1
g_A	1.294 533	4.847 020	4.999 957	5	5
c_B	1.238 391	5.148 717	6.005 194	6	6
d_B	0.856 369	3.990 632	3.999 617	4	4
K_A	1.656 336	1.008 029	1.998 252	2	2
f_B	0.797 221	0.975 804	0.999 970	1	1
g_B	1.872 435	5.000 549	4.999 609	5	5
α	1.891 000	1×10^{-17}	$1.300\,000 \times 10^{-15}$	1×10^{-17}	

Table 2.12: Numerical results for the synthetic biology toggle switch model using the paramters: $\mathbf{u}_0 = (2, 1, 1, 1, 1, 2, 1, 1, 1, 2)$, $\mathbf{u} = (5, 3, 2, 1, 5, 6, 4, 2, 1, 5)$, $\mathbf{x}_0 = (20, 20, 20, 20)$, on the interval $[0, 6]$, with step-size 0.005, with α chosen optimally at each iteration and with tolerance $|\Delta\mathbf{u}| < .00001$. Note 1×10^{-17} is effectively zero.

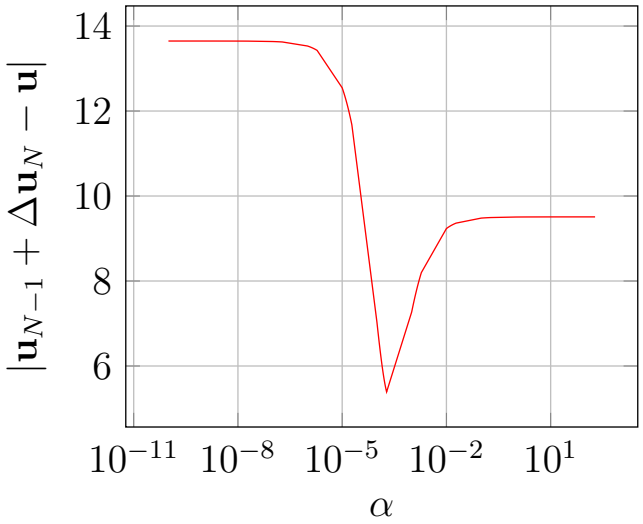

Figure 2.11: Graph of the α vs. $|\mathbf{u}_{N-1} + \Delta\mathbf{u}_N - \mathbf{u}|$ for Toggle Switch at iteration 20.

The error almost doesn't change for $p > 31$, and it increases when $p < 31$. The method converges even with a number of measurements $p = 13$ or $p = 11$, with error at 0.029 and 0.084 respectively.

Optimal choice of the regularization parameter can increase the convergence rate. In Figure 2.13 we see that with a constant choice for α it takes 166 iterations with a sublinear convergence rate of $r = 0.96427$. When we instead chose α optimally the convergence rate jumps to $r = 1.8508$.

2.2.5 Synthetic Biology Repressilator Model

In this section we consider the repressilator model introduced in [18] as a genetic network for three repressor-protein concentrations p_i, and their corresponding mRNA concentrations, m_i, where three components index $i = A, B, C$ stands

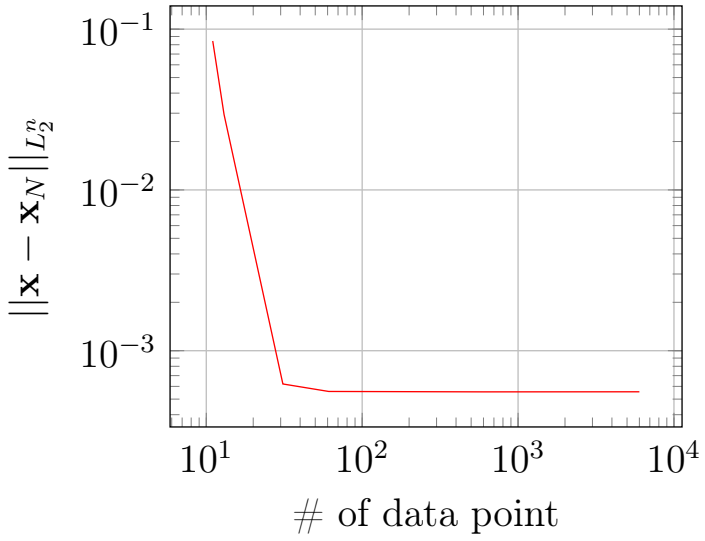

Figure 2.12: Graph of the number of data points vs. $||\mathbf{x} - \mathbf{x}_N||_{L_2^n}$ for the toggle switch model. The parameters were $\mathbf{u} = [1, 2, 1, 1, 1, 2, 1, 1, 10, 10]$, $\mathbf{u}_0 = [.5, 1, .5, .5, .5, 1, .5, .5, 5, 5]$, $\mathbf{x}_0 = [5, 2, 5, 2]$, $t \in [0, 6]$ and respective tolerances were set at 1×10^{-7}.

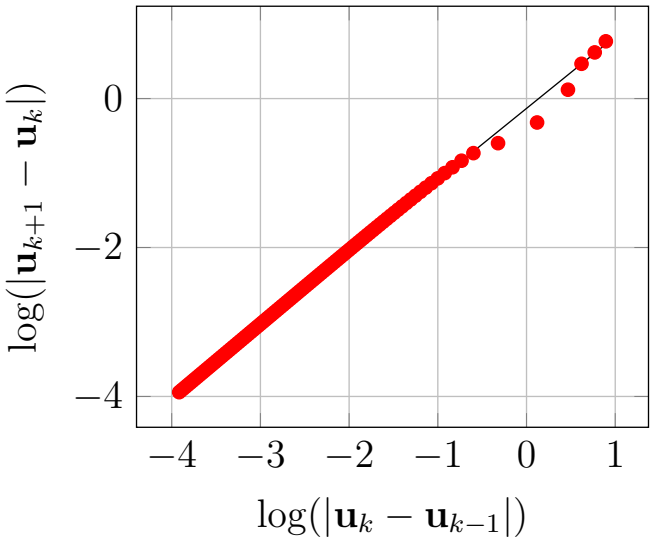

Figure 2.13: A convergence rate graph for the synthetic biology toggle switch model using the paramters: $\mathbf{u}_0 = (2, 2, 2, 2, 2, 2, 2, 2, 2, 2)$, $\mathbf{u} = (5, 3, 2, 1, 5, 6, 4, 2, 1, 5)$, $\mathbf{x}_0 = (20, 20, 20, 20)$, time interval $[0, 6]$, with step-size 0.005, regularization parameters is chosen to be α, and tolerance $|\Delta\mathbf{u}| < .00001$. The convergence rate is $r = 0.96427$ and $C = .73785$.

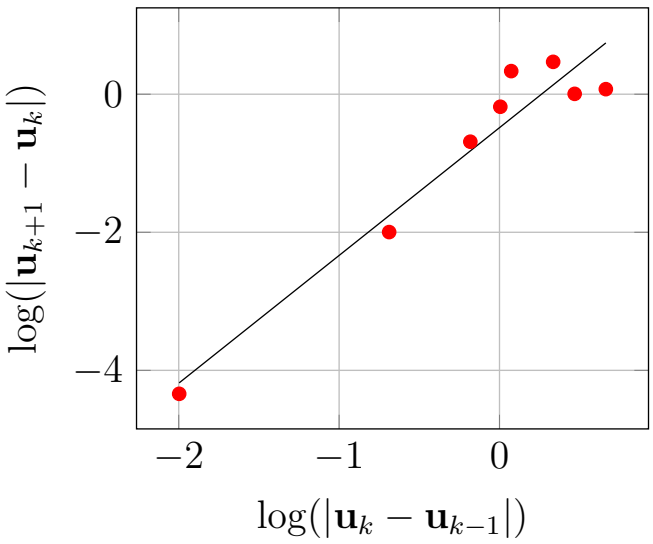

Figure 2.14: A convergence rate graph for the synthetic biology toggle switch model using the paramters: $\mathbf{u}_0 = (2, 2, 2, 2, 2, 2, 2, 2, 2, 2)$, $\mathbf{u} = (5, 3, 2, 1, 5, 6, 4, 2, 1, 5)$, $\mathbf{x}_0 = (20, 20, 20, 20)$, time interval $[0, 6]$, with step-size 0.005, regularization parameters is chosen optimally and tolerance $|\Delta\mathbf{u}| < .00001$. The convergence rate is $r = 1.8508$ and $C = .32758$.

for *lacl*, *tetR*, *cl*:

$$\dot{m}_A = -m_A + \frac{\alpha_A}{1 + p_C^{n_A}} + \beta_A$$

$$\dot{p}_A = -\gamma_A p_A + \delta_A m_A$$

$$\dot{m}_B = -m_B + \frac{\alpha_B}{1 + p_A^{n_B}} + \beta_B$$

$$\dot{p}_B = -\gamma_B p_B + \delta_B m_B$$

$$\dot{m}_C = -m_C + \frac{\alpha_C}{1 + p_B^{n_C}} + \beta_C$$

$$\dot{p}_C = -\gamma_C p_C + \delta_C m_C$$

with $\alpha_i, \beta_i, \gamma_i, \delta_i, i = A, B, C$ being parameters of the system. We consider the inverse problem on the identification of these parameters with given measurements

of the components of the system. In the paper [18] the authors considered the case where $\alpha_i = \alpha_j$, $\beta_i = \beta_j$, $n_i = n_j$, $\gamma_i = \gamma_j$ and $\gamma_i = \delta_i$ for all $i, j = A, B, C$, and therefore there are only 4 parameters of the system. We apply our method to two different cases of identification of 4 or all 15 parameters based on the measurement of the components of the system. The case of 4 (respectively 15) parameters leads to a system of 30 (respectively 96) linear differential equations instead of (2.3), (2.7). Table A.14 presents the typical result on the convergence of the method without regularization for the repressilator model with 4 unknown parameters. The true value of the parameter is $(216, 0.216, 2, 0.2)$, and with the initial choice of the parameter as $(300, 0.5, 1, 0.5)$, the method converges to the true value in 15 steps. To improve the convergence range of the initial guess, we implemented the Type II regularization method. Table 2.13 demonstrates the typical convergence result with the Type II regularization with constant regularization parameter α. With the reference point chosen as $(250, 0.4, 3, 0.4)$,the initial value $(100, 0.75, 0.75, 0.75)$ of the parameter implies convergence of Type II regularization with constant regularization parameter α to the true value of the parameter in 9 steps. Without regularization however, the latter initial choice did not provide a convergent sequence as in Table A.14.

In the case of 15-parameters, our method did not provide a convergent algorithm without regularization. Table 2.14 presents the typical result on the convergence of the method with Type II regularization. Note that we have chosen the same model example as for the 4-parameter case demonstrated in Table 2.13. Although we have chosen the same reference point $\mathbf{u}^*$, the convergence range of $\mathbf{u}^0$ is reduced in the 15-parameter case. Finally, Table A.15 demonstrates the typical convergence result for the 4-parameter repressilator model with noise.

time	p_{A3}	p_{A6}	p_{A9}	p_A
0	1	1	1	1
2	−1.909 402	5.022 548	5.618 842	5.616 314
4	−3.946 382	5.358 636	5.815 716	5.814 707
6	−6.764 595	5.226 210	5.590 746	5.590 611
8	−11.936 782	5.462 840	5.890 827	5.890 989
10	−21.340 461	6.068 294	6.652 376	6.652 593
time	p_{B3}	p_{B6}	p_{B9}	p_B
0	1	1	1	1
2	−3.370 038	7.122 780	8.014 053	8.009 791
4	−6.830 577	7.862 081	8.546 051	8.544 433
6	−11.185 145	7.544 163	8.028 404	8.028 295
8	−17.639 805	7.240 887	7.627 051	7.627 376
10	−26.769 947	6.844 977	7.127 921	7.128 320
time	p_{C3}	p_{C6}	p_{C9}	p_C
0	1	1	1	1
2	−3.415 119	6.534 050	7.278 035	7.274 865
4	−5.872 782	6.440 608	6.938 601	6.937 266
6	−8.609 888	5.567 087	5.849 038	5.848 559
8	−12.660 731	4.910 608	5.103 273	5.102 954
10	−18.887 143	4.522 007	4.709 491	4.709 068
	$\mathbf{u}_3$	$\mathbf{u}_6$	$\mathbf{u}_9$	$\mathbf{u}$
α	94.071 162	213.333 295	216.352 630	216
β	−1.376 693	0.224 982	0.222 439	0.216 000
n	1.938 660	2.004 965	2.001 578	2
γ, δ	−0.011 369	0.201 585	0.200 105	0.200 000

Table 2.13: Numerical results for the synthetic biology repressilator model using the parameters: $\mathbf{u}_0 = (100, .75, .75, .75)$, $\mathbf{u}^* = (250, .4, 3, .4)$, $\mathbf{u} = (216, 0.216, 2, 0.2)$, $\mathbf{x}_0 = (5, 1, 10, 1, 15, 1)$, time interval $[0, 10]$, with step-size 0.01, $\alpha = 0.00005$ and tolerance $|\Delta\mathbf{u}| < .00001$.

time	p_{A3}	p_{A6}	p_{A9}	p_A
0	1	1	1	1
2	6.776 085	5.625 591	5.616 732	5.616 314
6	5.893 548	5.622 450	5.590 985	5.590 611
10	5.132 980	6.594 276	6.652 293	6.652 593
time	p_{B3}	p_{B6}	p_{B9}	p_B
0	1	1	1	1
2	9.336 061	8.007 041	8.010 520	8.009 791
6	7.019 081	8.065 853	8.028 267	8.028 295
10	8.269 912	7.200 896	7.128 233	7.128 320
time	p_{C3}	p_{C6}	p_{C9}	p_C
0	1	1	1	1
2	7.439 256	7.289 729	7.277 208	7.274 865
6	6.379 881	5.911 191	5.847 817	5.848 559
10	6.165 000	4.773 809	4.707 173	4.709 068
	$\mathbf{u}_3$	$\mathbf{u}_6$	$\mathbf{u}_9$	$\mathbf{u}$
α_A	228.982 871	216.442 948	216.460 894	216
β_A	2.830 946	0.239 973	0.238 368	0.216 000
n_A	2.350 906	2.003 270	2.003 545	2
γ_A	0.261 627	0.200 002	0.200 017	0.200 000
δ_A	0.238 305	0.199 993	0.200 002	0.200 000
α_B	201.004 049	216.204 270	216.208 511	216
β_B	5.129 999	0.222 161	0.220 634	0.216 000
n_B	1.541 097	2.001 066	2.000 946	2
γ_B	0.223 425	0.200 004	0.200 022	0.200 000
δ_B	0.219 662	0.200 002	0.200 013	0.200 000
α_C	208.287 351	216.314 969	216.327 521	216
β_C	0.943 553	0.232 950	0.232 513	0.216 000
n_C	1.789 792	2.003 089	2.003 110	2
γ_C	0.180 625	0.200 226	0.200 188	0.200 000
δ_C	0.191 181	0.200 133	0.200 110	0.200 000

Table 2.14: Numerical results for the synthetic biology repressilator model using the parameters: $\mathbf{u}_0 = (175, .5, 1, .5, .5, 175, .5, 1, .5, .5, 175, .5, 1, .5, .5)$, $\mathbf{u}^* = (250, .4, 3, .4, .4, 250, .4, 3, .4, .4, 250, .4, 3, .4, .4)$, $\mathbf{u} = (216, .216, 2, .2, .2, 216, .216, 2, .2, .2, 216, 0.216, 2, .2, .2)$, $\mathbf{x}_0 = (5, 1, 10, 1, 15, 1)$, $\alpha = 0.00001$, time interval $[0, 10]$, with step-size 0.01 and tolerance $|\Delta\mathbf{u}| < .00001$.

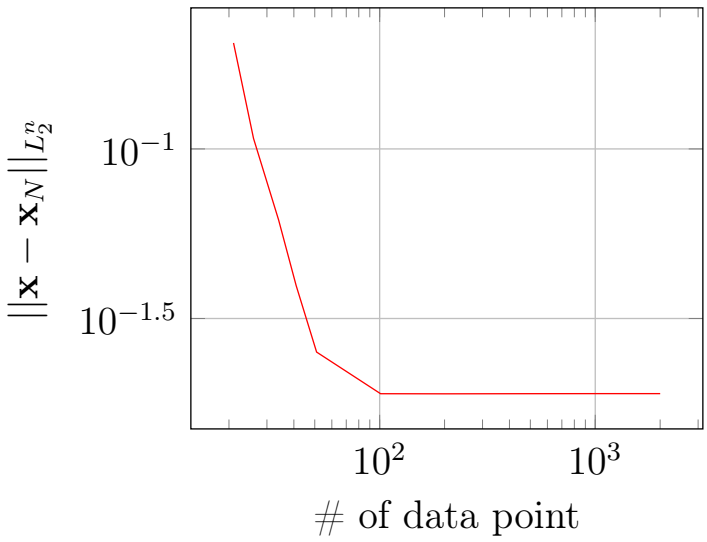

Figure 2.15: Graph of the number of data points vs. $||\mathbf{x} - \mathbf{x}_N||_{L_2^n}$ for the repressilator model. The parameters were $\mathbf{u}_0 = (175, .5, 1, .5, .5, 175, .5, 1, .5, .5, 175, .5, 1, .5, .5)$, $\mathbf{u}^* = (250, .4, 3, .4, .4, 250, .4, 3, .4, .4, 250, .4, 3, .4, .4)$, $\mathbf{u} = (216, .216, 2, .2, .2, 216, 0.216, 2, .2, .2, 216, 0.216, 2, .2, .2)$, $\mathbf{x}_0 = (5, 1, 10, 1, 15, 1)$, $\alpha = 0.00001$, time interval $[0, 10]$ and respective tolerances were set at 1×10^{-7}.

Lastly, we analyze the dependence of the accuracy of the method on the number of available measurements. Figure 2.15 and Table A.16 present L-shape curve expressing dependence of error $||\mathbf{x} - \mathbf{x}_N||_{L_2^n}$ vs. p for a typical example. Edge point corresponds to number of measurements $p = 101$ with error of order 10^{-2}. The error almost doesn't change for $p > 101$, and it increases when $p < 101$. For $30 \leq p \leq 50$, the error is between 0.06 and 0.025. The method converges with a number of measurements at $p = 26$ or $p = 21$, with error at 0.1 and 0.2 respectively.

Here we see that the choice of regularization method also make a difference in convergence rate. In particular having a good guess for where the true solution lies. Consider figure 2.16 where we use type I regularization, we get a convergence rate of $r = 1.7006$. When we switch to type II while leaving the other settings

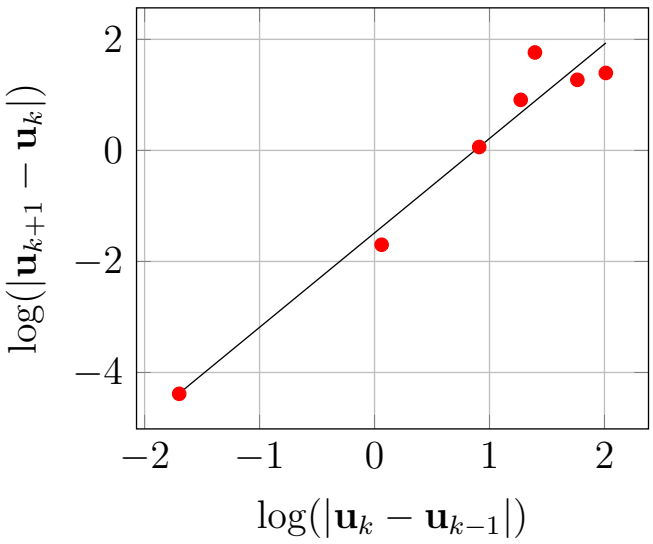

Figure 2.16: A convergence rate graph for the synthetic biology repressilator model using type I regularization and the parameters: $\mathbf{u}_0 = (175, .5, 1, .5, .5, 175, .5, 1, .5, .5, 175, .5, 1, .5, .5)$, $\mathbf{u} = (216, .216, 2, .2, .2, 216, .216, 2, .2, .2, 216, 0.216, 2, .2, .2)$, $\mathbf{x}_0 = (5, 1, 10, 1, 15, 1)$, $\alpha = 0.000001$, time interval $[0, 10]$, with step-size 0.02 and tolerance $|\Delta\mathbf{u}| < .00001$. The convergence rate is $r = 1.7006$ and $C = .03267$.

the same, given a sufficiently close guess we get a convergence rate of $r = 2.0407$. Figure 2.16 demostrates this.

2.3 Comparison with *lsqnonlin*, *fmincon* and *nl2sol*

In this section we compare our method against some of the popular methods used in open source software [19, 33, 34]. Competitive methods include

- Levenberg-Marquardt algorithm and trust-region-reflective method (function *lsqnonlin* in MatLab) [10]
- Sequential Quadratic Programming (function *fmincon* in MatLab Optimization toolbox)

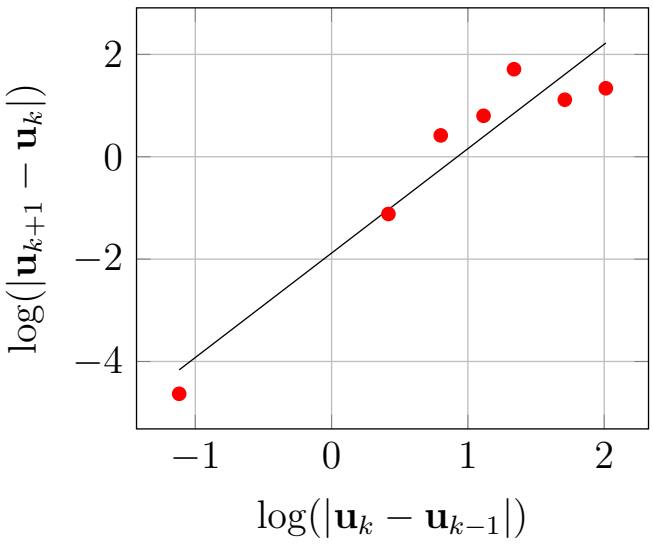

Figure 2.17: A convergence rate graph for the synthetic biology repressilator model using type II regularization and the parameters: $\mathbf{u}_0 = (175, .5, 1, .5, .5, 175, .5, 1, .5, .5, 175, .5, 1, .5, .5)$, $\mathbf{u}^* = (250, .4, 3, .4, .4, 250, .4, 3, .4, .4, 250, .4, 3, .4, .4)$, $\mathbf{u} = (216, .216, 2, .2, .2, 216, .216, 2, .2, .2, 216, 0.216, 2, .2, .2)$, $\mathbf{x}_0 = (5, 1, 10, 1, 15, 1)$, $\alpha = 0.00001$, time interval $[0, 10]$, with step-size 0.02 and tolerance $|\Delta\mathbf{u}| < .00001$. The convergence rate is $r = 2.0407$ and $C = .013117$.

- An adaptive non-linear least-squares algorithm (function *nl2sol* implemented in FORTRAN ported into MatLab) [20]

We called our method *qlopt* and used MatLab package AMIGO2 [19] in our comparison. We matched the default settings from AMIGO2 to the equivalent settings in *qlopt*. So we set our tolerance for our termination conditions to 1×10^{-7}. Additionally our method was implemented in C/C++ so we enabled the two features in AMIGO2 that used C implementations to gain a speed-up. In particular the models were set to *charmodelC* mode and we used CVODES for solving the systems of ODEs. By default the examples in AMIGO2 are setup to run in the "enhanced C mode". The cost function type was set to *lsq* (i.e. weighted least squares) with the weights being the identity matrix.

For *qlopt* we solved the system of ODEs and acquired the sensitivity functions by considering the finite difference approximation of (2.3) and (2.7). Then we solved both systems simultaneously using a simple textbook implementation of RK4 [54, 55].

In both frameworks we chose a common starting value for $\mathbf{u}_0$, we used the same set of simulated data for all the models.

We generate simulated noisy data according to the following percent noise model:

$$y_i = x_i(t, u) + px_i(t, u)\nu_i \tag{2.18}$$

for $i = 1, \ldots, n$ where $\mathbf{x}(t, u)$ represents the measurement, p is a percentage and ν is an n-dimensional random variable with standard normal distribution as in (2.16).

For every model example we ran each algorithm 25 times and recorded the *average relative error of the parameter values*, *the median number of objective function evaluations*, *the average computational time*, and *the median number of iterations*. Tables 2.15 & 2.16 demonstrate the comparison results for two typical examples on bistable switch and synthetic toggle switch models. All four methods have a comparable relative error. Our method has a clear advantage over all three methods on computational time and number of function evaluations. In terms of the required number of iterations, our method has a clear advantage over *lsqnonlin* and *fmincon*, and still has a slight advantage over *nl2sol*. It should be noted that our software package *qlopt* is implemented using C++ and Eigen, which gives advantage over Matlab-based methods in terms of computational time.

metric	qlopt	lsqnonlin	fmincon	nl2sol
r.e.	0.0291	0.0260	0.0264	0.0257
n.i.	7	18	30	8
c.t. (s)	0.0128	0.2304	0.4844	0.2180
f.e.	7	73	128	38

Table 2.15: Comparison of several local optimization method against the presented method for the bistable switch model. We took the average of 25 runs. The parameters were $\mathbf{u} = (150, 3.2, 2.0)$, $\mathbf{u}_0 = (125, 2.2, 1.0)$, $\mathbf{x}_0 = (14, 13)$, $t \in [0, 3]$, $\Delta t = .1$, 1% noise and respective tolerances were set at 1×10^{-7}. Mean of the relative errors (r.e.), median of the number of iterations (n.i.), mean of the computational time (c.t.), and the median number of function evaluations (f.e.).

metric	qlopt	lsqnonlin	fmincon	nl2sol
r.e.	0.017173	0.011348	0.011796	0.0110759
n.i.	7	13	69	9
c.t. (s)	0.128976	0.30799	1.64119	0.55999
f.e.	7	144	794	113

Table 2.16: Comparison of several local optimization method against the presented method for the synthetic toggle switch model. We took the average of 25 runs. The parameters were$\mathbf{u} = [1, 2, 1, 1, 1, 2, 1, 1, 10, 10]$, $\mathbf{u}_0 = [.5, 1, .5, .5, .5, 1, .5, .5, 5, 5]$, $\mathbf{x}_0 = [5, 2, 5, 2]$, $t \in [0, 6]$, $\Delta t = .1$, 1% noise and respective tolerances were set at 1×10^{-7}. Mean of the relative errors (r.e.), median of the number of iterations (n.i.), mean of the computational time (c.t.), and the median number of function evaluations (f.e.).

Chapter 3

Identification of Parameters for Large-scale Models in Systems Biology

The results in this chapter were submitted for publication in the journal of *Mathematical Biosciences*.

3.1 Description of the Method

Consider a dynamical system:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(t, \mathbf{x}, \mathbf{u}, \mathbf{v}),\ t_0 \leq t \leq t_1 \tag{3.1}$$

$$\mathbf{x}(t_0) = \mathbf{x}^0 \in \mathbb{R}^n, \tag{3.2}$$

where

$$\mathbf{x} = \mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t)) : [t_0, t_1] \to \mathbb{R}^n$$

is the state vector,

$$\mathbf{u} = (u_1, u_2, \ldots, u_m) \in \mathbb{R}^m$$

is the unknown parameter vector,

$$\mathbf{v} = (v_1, v_2, \ldots, v_p) \in \mathbb{R}^p$$

is the control input parameter vector, and

$$\mathbf{f} = (f_1(t, \mathbf{x}, \mathbf{u}, \mathbf{v}), f_2(t, \mathbf{x}, \mathbf{u}, \mathbf{v}), \ldots, f_n(t, \mathbf{x}, \mathbf{u}, \mathbf{v})) :$$
$$[t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}^n$$

is a continuous vector function with continuous derivatives

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}}, \ \frac{\partial \mathbf{f}}{\partial \mathbf{u}}.$$

Consider *inverse problem of finding the parameter* $\mathbf{u}$ *given* D *measurements for the state vector* $\mathbf{x}$ *corresponding to* D *fixed values of the control vector* $\mathbf{v}$*:*

$$\mathbf{x} = \mathbf{x}^d(t) = \mathbf{x}^d(t; \mathbf{u}) := \mathbf{x}(t, \mathbf{u}, \mathbf{v}^d), \ d = 1, ..., D$$

on an interval $t_0 \le t \le t_1$*, where* $\mathbf{x}^d(t_0) = \mathbf{x}^0$.

Having chosen the initial vector function $\mathbf{x}^d_{N,0}$ (say, $\mathbf{x}^d_{N,0} = \mathbf{x}^d(t)$), and initial approximation $\mathbf{u} = \mathbf{u}_0$, we implement quasilinearization of (3.1) ([43]) and at each

fixed iteration $N = 1, 2, ...$ we find the solution as a limit

$$\mathbf{x}_N^d(t) = \lim_{p\to\infty} \mathbf{x}_{N,p}^d(t),\ t_0 \le t \le t_1, \tag{3.3}$$

where $\mathbf{x}_{N,p}^d$ solves the linear system of ODEs in $[t_0, t_1]$ with $\mathbf{u} = \mathbf{u}_{N-1}$:

$$\frac{\mathrm{d}\mathbf{x}_{N,p}^d}{\mathrm{d}t} = \mathbf{f}(t, \mathbf{x}_{N,p-1}^d, \mathbf{u}, \mathbf{v}^d) + J(t, \mathbf{x}_{N,p-1}^d, \mathbf{u}, \mathbf{v}^d)(\mathbf{x}_{N,p}^d - \mathbf{x}_{N,p-1}^d), \tag{3.4a}$$

$$\mathbf{x}_{N,p}^d(t_0) = \mathbf{x}^0, \tag{3.4b}$$

where

$$J(t, \mathbf{x}, \mathbf{u}, \mathbf{v}) = \left[\frac{\partial f_i(t, \mathbf{x}, \mathbf{u}, \mathbf{v})}{\partial x_j}\right]$$

is the n×n Jacobian matrix. It is well known that the convergence (3.3) has a quadratic rate [43]. Given the initial guess $\mathbf{u}_0$ of the unknown parameter $\mathbf{u}$, at every step of the iteration we identify a new approximation

$$\mathbf{u}_N = \mathbf{u}_{N-1} + \Delta\mathbf{u}, \tag{3.5}$$

which minimizes the L_2-norm of the residues

$$\mathcal{R} = \mathbf{x}^d(t, \mathbf{u}) - \mathbf{x}_N^d(t, \mathbf{u}_N).$$

We have

$$\mathcal{R} = \Delta\mathbf{x}_N^d(t) - U_N^d\Delta\mathbf{u} + o(|\Delta\mathbf{u}|), \quad \text{as } |\Delta\mathbf{u}| \to 0, \tag{3.6}$$

where

$$\Delta\mathbf{x}_N^d(t) = \mathbf{x}^d(t, \mathbf{u}) - \mathbf{x}_N^d(t, \mathbf{u}_{N-1}),$$

U_N^d is an $n \times m$ sensitivity matrix with columns

$$U_N^{d,j} = \Big(\frac{\partial \mathbf{x}_N^d(t, \mathbf{u}_{N-1})}{\partial u^j}\Big),\ j = 1, ..., m.$$

U_N^d solves the matrix differential system

$$\frac{dU_N^d}{dt} = \frac{\partial}{\partial \mathbf{u}}\mathbf{f}(t, \mathbf{x}_N^d, \mathbf{u}_{N-1}, \mathbf{v}^d) + J(t, \mathbf{x}_N^d, \mathbf{u}_{N-1}, \mathbf{v}^d)U_N^d,\ t_0 \le t \le t_1 \tag{3.7a}$$

$$U_N^d(t_0) = 0, \tag{3.7b}$$

where $\mathbf{x}_N^d$ is the solution of (3.1),(3.2) with $\mathbf{u} = \mathbf{u}_{N-1}, \mathbf{v} = \mathbf{v}^d$ as it is constructed in (3.3). Finding $\mathbf{x}_N^d, U_N^d$ from (3.3), (3.4), (3.7) form the method of staggered corrector [42].

To find $\Delta\mathbf{u}$, we minimize the multi-objective function

$$\mathcal{J}(\Delta\mathbf{u}) = \sum_{d=1}^{D} ||\Delta\mathbf{x}_N^d - U_N^d\Delta\mathbf{u}||^2_{L_2(t_0,t_1)}. \tag{3.8}$$

We have

$$\mathcal{J}_N'(\Delta\mathbf{u}) = 2\sum_{d=1}^{D}\int_{t_0}^{t_1}\Big[(U_N^d)^T U_N^d\Delta\mathbf{u} - (U_N^d)^T\Delta\mathbf{x}_N^d\Big]\mathrm{d}t,$$

$$\mathcal{J}_N''(\Delta\mathbf{u}) = 2\sum_{d=1}^{D}\int_{t_0}^{t_1}(U_N^d)^T U_N^d\mathrm{d}t.$$

Therefore, minimum $\Delta\mathbf{u}$ satisfies the following system of linear algebraic equations

$$A_N\Delta\mathbf{u} = P_N, \tag{3.9}$$

where

$$A_N = \sum_{d=1}^{D} \int_{t_0}^{t_1} (U_N^d)^T U_N^d \mathrm{d}t = \left(a_N^{ij}\right)_{i,j=1}^{m}$$

is an $m \times m$ symmetric matrix with elements

$$a_N^{ij} = \sum_{d=1}^{D} \int_{t_0}^{t_1} \Big(\frac{\partial x_N^d(t, \mathbf{u}_{N-1})}{\partial u^i}\Big)^T \frac{\partial x_N^d(t, \mathbf{u}_{N-1})}{\partial u^j} dt,$$

and

$$P_N = \sum_{d=1}^{D} \int_{t_0}^{t_1} (U_N^d)^T \Delta \mathbf{x}_N \mathrm{d}t = \left(p_N^j\right)_{j=1}^{m}$$

is an m-vector with elements

$$p_N^j = \sum_{d=1}^{D} \int_{t_0}^{t_1} \Big(\frac{\partial x_N^d(t, \mathbf{u}_{N-1})}{\partial u^j}\Big)^T (x^d(t, \mathbf{u}) - x_N^d(t, \mathbf{u}_{N-1}))dt.$$

In fact, A_N is a sum of Gram matrices of vectors $U_N^{d,j}$, and

$$a_N^{ij} = \sum_{d=1}^{D} (U_N^{d,i}, U_N^{d,j})_{L_2^n(t_0,t_1)}$$

where $L_2^n(t_0, t_1)$ is a Hilbert space of vector functions $g : (t_0, t_1) \to \mathbb{R}^n$ with inner product

$$(g, h)_{L_2^n(t_0,t_1)} = \int_{t_0}^{t_1} g^T h dt.$$

It is known [46] that

$$det(A_N) = \sum_{d=1}^{D} \Gamma(U_N^{d,1}, ..., U_N^{d,m}) \geq 0$$

and it is positive, that is to say, A_N is non-singular if and only if the vectors

$U_N^{d,j}, j = 1, ..., m$ are linearly independent at least for one $d = 1, ..., D$.

Hence, we suggest the following modfication of the numerical algorithm from [38].

3.1.1 Algorithm

1. Initialize $\mathbf{u}_0$ and set $N = 1$.

2. Set $\mathbf{x}_{N,0}^d(t)$ and find $\mathbf{x}_N^d(t, \mathbf{u}_{N-1})$ via quasilinearization from (3.3),(3.4).

3. Having $\mathbf{x}_N^d(\cdot, \mathbf{u}_{N-1})$ find sensitivity matrices U_N^d by solving linear ODE systems (3.7).

4. Find $\Delta\mathbf{u}$ by solving linear algebraic equations system (3.9) and update the new value $\mathbf{u}_N$ of the parameter using (3.5).

5. If satisfactory accuracy is achieved, then terminate the process, otherwise replace N with $N + 1$ and go move back to Step 2. As termination criteria, the smallness of either of the expressions

$$|\mathbf{u}_{N-1} - \mathbf{u}_N|, \ \mathcal{J}_N(\Delta\mathbf{u}), \ \sum_{d=1}^{D} ||\mathbf{x}^d(\cdot) - \mathbf{x}_N^d(\cdot, \mathbf{u}_N)||_{L_2^n}$$

can be used.

3.1.2 Regularization

As in a previous chapter we implement two types of Tikhonov regularization. Type I regularization is performed by replacing the function (3.8) with

$$\sum_{d=1}^{D} \left(||\Delta\mathbf{x}_N^d - U_N^d \Delta\mathbf{u}||_{L_2}^2 \right) + \alpha |\Delta\mathbf{u}|^2. \tag{3.10}$$

This yields the following linear system instead of (3.9)

$$(A_N + \alpha I)\Delta\mathbf{u} = P_N \tag{3.11}$$

where I is the identity matrix and α is a regularization parameter. Type II regularization is performed by replacing the function (3.8) with

$$\sum_{d=1}^{D} \left(||\Delta\mathbf{x}_N^d - U_N^d \Delta\mathbf{u}||_{L_2}^2 \right) + \alpha |\mathbf{u}_{N-1} + \Delta\mathbf{u} - \mathbf{u}^*|^2 \tag{3.12}$$

where $\mathbf{u}^*$ is a known vector with the expectation to be close to the true value of the unknown parameter. This implies the following linear system instead of (3.9):

$$(A_N + \alpha I)\Delta\mathbf{u} = P_N + \alpha(\mathbf{u}^* - \mathbf{u}_{N-1}). \tag{3.13}$$

3.1.3 Identifiability

Convergence of the algorithm is connected to identifiability of unknown parameters. In fact, dth Gram matrix summand of A_N in (3.9) is the so called Fisher information matrix (FIM) for the ODE system (3.1) and it characterizes the information content of the experimental measurement in dth data set. Singularity of A_N is equivalent to linear dependence of the sensitivity vectors $U_N^{d,j}, j = 1, ..., m$ for all $d = 1, ..., D$, which is the indication of the presence of the non-identifiable parameters . On the contrary, non-singularity of A_N is equivalent to *identifiability* of parameters. If A_N is non-singular, but $detA_N$ is sufficiently small then for computer simulation A_N is treated as a singular matrix [8, 17, 47]. Our two regularization algorithms are developed to address such *practical non-identifiability*

cases. Major factor for the convergence of the algorithm for the identification of *practically non-identifiable* parameters is the increase of number of data sets D. Precisely, there is a minimum number of data sets with different inputs of control parameters for experimental design are needed to relieve the parameter correlations and acquire suitable measurement data for unique parameter estimation [4].

3.2 Results and Discussions

We tested the method on a benchmark model of a biological network for a three-step pathway modeled by 8 nonlinear ODEs describing 8 metabolic concentrations and 36 parameters $p_i, i = 1, ..., 36$ [56]:

$$
\begin{aligned}
\dot{x}_1 &= \frac{p_1}{1 + \left(\frac{P}{p_2}\right)^{p_3} + \left(\frac{p_4}{S}\right)^{p_5}} - p_6 x_1 \\
\dot{x}_2 &= \frac{p_7}{1 + \left(\frac{P}{p_8}\right)^{p_9} + \left(\frac{p_{10}}{x_7}\right)^{p_{11}}} - p_{12} x_2 \\
\dot{x}_3 &= \frac{p_{13}}{1 + \left(\frac{P}{p_{14}}\right)^{p_{15}} + \left(\frac{p_{16}}{x_8}\right)^{p_{17}}} - p_{18} x_3 \\
\dot{x}_4 &= \frac{p_{19} x_1}{p_{20} + x_1} - p_{21} x_4 \\
\dot{x}_5 &= \frac{p_{22} x_2}{p_{23} + x_2} - p_{24} x_5 \\
\dot{x}_6 &= \frac{p_{25} x_3}{p_{26} + x_3} - p_{27} x_6 \\
\dot{x}_7 &= \frac{p_{28} x_4 \left(S - x_7\right)}{p_{29}\left(1 + \frac{S}{p_{29}} + \frac{x_7}{p_{30}}\right)} - \frac{p_{31} x_5 \left(x_7 - x_8\right)}{p_{32}\left(1 + \frac{x_7}{p_{32}} + \frac{x_8}{p_{33}}\right)} \\
\dot{x}_8 &= \frac{p_{31} x_5 \left(x_7 - x_8\right)}{p_{32}\left(1 + \frac{x_7}{p_{32}} + \frac{x_8}{p_{33}}\right)} - \frac{p_{34} x_6 \left(x_8 - P\right)}{p_{35}\left(1 + \frac{x_8}{p_{35}} + \frac{P}{p_{36}}\right)}
\end{aligned}
$$

Two parameters P and S are control input parameters specified by the experimental design. The unknown parameters p_i are correlated, but their functional relationship with one another is dependent on the input parameters P and S, and in general parameters are practically identifiable with multiple data sets. In [8], the inverse problem was analyzed with 16 noise-free data sets, and in [31] with 16 both noise-free and noisy data sets. The results demonstrated strong parameter correlations in several groups, with accurate parameter values identified in [31]. Parameter correlations were analyzed in [4]. It is demonstrated that correlated parameters are practically non-identifiable for a single data set and at least 5 data sets with different control inputs are required to uniquely estimate the 36 parameters of this model.

For our experiments we used the common values for the initial conditions $(6.6667e{-}1, 5.7254e{-}1, 4.1758e{-}1, 4.0e{-}1, 3.6409e{-}1, 2.9457e{-}1, 1.419, 9.3464e{-}1)$, with $t_0 = 0$ and $t_1 = 120$. We implemented the 16 input parameters given in AMIGO2 [19] and 5 input parameters given in [4] for our experiments. We chose the regularization parameter α as a function of the residual:

$$\sum_{d=1}^{D} C||\mathbf{x}^d - \mathbf{x}_N^d||_{L_2}^{\gamma} \tag{3.14}$$

where $C, \gamma > 0$ are chosen experimentally.

3.2.1 Numerical Results with Noise-free Data Sets

We applied the numerical method to identify the 36 parameters with 16 and 5 data sets. We generated simulated measurements for each data set by solving the system of 8 nonlinear ODEs with true values of 36 parameters. We chose the number of

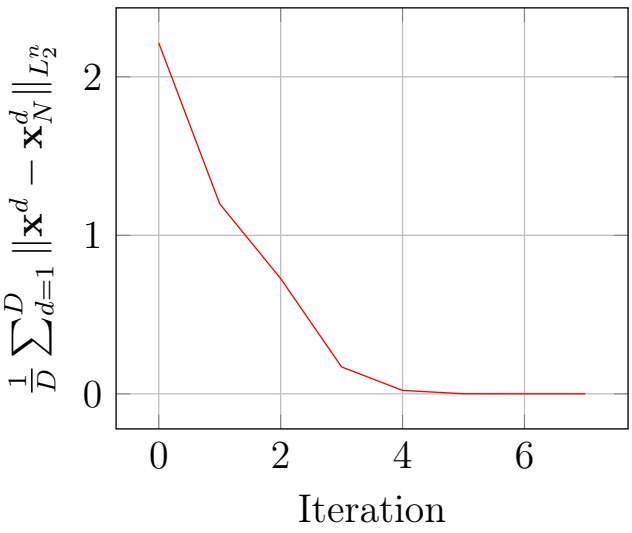

Figure 3.1: The average error at each iteration corresponding to Table 3.1.

time data points for each of the 8 components of the system either at 240 or at 20 uniformly distributed time grid points in the segment $[0, 120]$. Computational cost of each iteration per one data set consists of iterative solution of the system of 8 nonlinear ODEs through quasilinearization; solving a system of 288 linear ODEs to identify sensitivity matrix-function; calculation of 1332 integrals for entries of the matrix A_N and vector P_N; and finally solvinga system of 36 linear algebraic equations to find the increment of the parameters. In Table A.17 and Figure 3.1 we demonstrate the results for 16 data sets with 240 time points. Rapid convergence to the true solution happens in only 7 iterations. Next we applied the method with 5 data sets. Though it required 3 extra iterations, Table A.18 and Figure 3.2 demonstrate the rapid convergence of the method with reduced error. Next we applied the method by choosing measurements at 20 time grid points for each of the 8 components. The results are demonstrated for 16 and 5 data sets in Tables A.19 and A.20, respectively. The algorithm converges in the same number of iterations with respect to the number of data sets, while maintaining around the same level

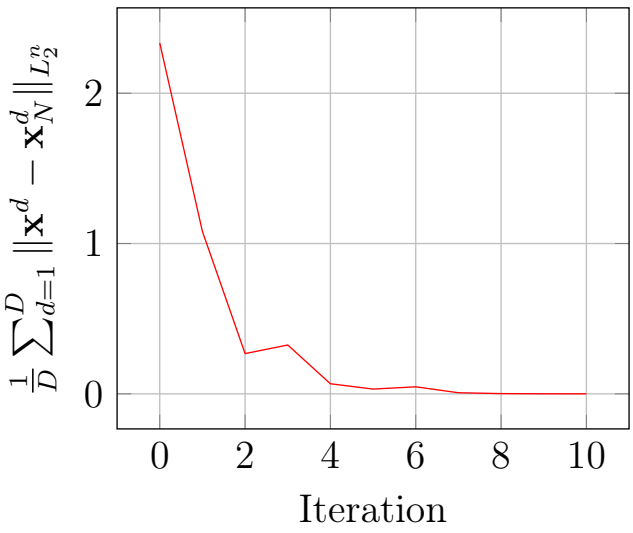

Figure 3.2: The average error at each iteration corresponding to Table 3.2.

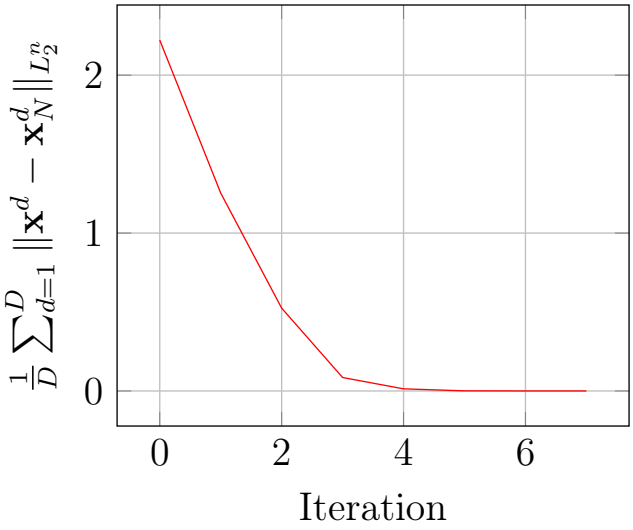

Figure 3.3: The average error at each iteration corresponding to Table 3.3.

of accuracy, as demonstrated in Figures 3.3 and 3.4.

3.2.2 Effect of the Regularization Parameter α

The choice of the regularization parameter α is an important factor which significantly improves the convergence rate and computational cost of the algorithm.

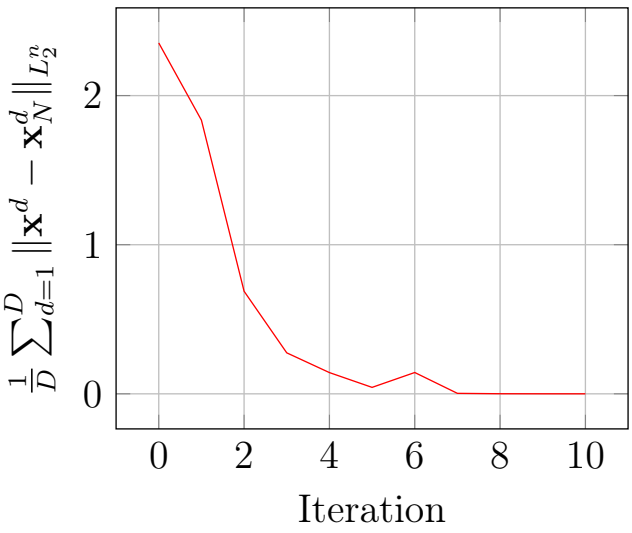

Figure 3.4: The average error at each iteration corresponding to Table 3.4.

To demonstrate the existence of the optimal non-trivial value of α at every fixed step N, we considered profiles of α vs $|\mathbf{u}_{N-1} + \Delta\mathbf{u} - \mathbf{u}^*|$, where $\mathbf{u}^*$ is the true solution. Figures 3.5 and 3.6 correspond to iterations 2 and 4 respectively for the results demonstrated in Table A.17. Similarly, Figures 3.7 and 3.8 correspond to iterations 3 and 6 respectively for the results from Table A.18. In each example there is a clear minimum which is the best choice of the regularization parameter. The bullets on the graph corresponds to our choice of the regularization parameter according to the residual method (3.14). In fact, optimal or nearly optimal choice of the regularization parameter significantly increases convergence rate of the method from geometric to be close to quadratic convergence (see Section 3.2.7). The residual method provides a close, but not necessarily optimal value of α. This analysis demonstrated that there is room for improvement of the convergence rate of the algorithm through implementation of a more effective method for the search of regularization parameter α without significantly affecting computational cost.

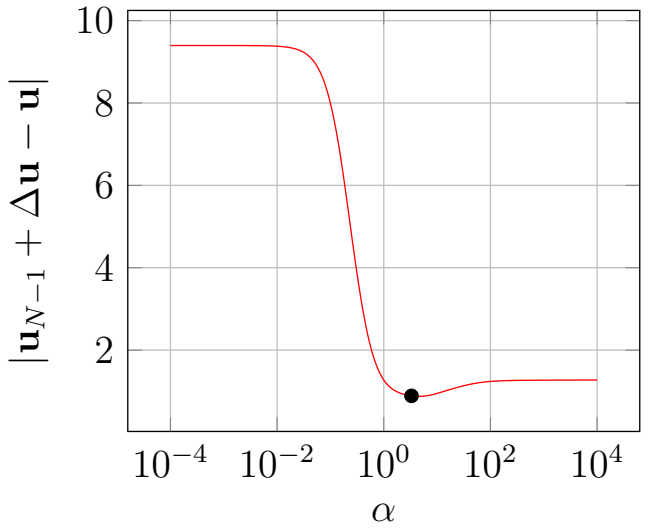

Figure 3.5: Profile of α at iteration 2. Corresponding to Table 3.1.

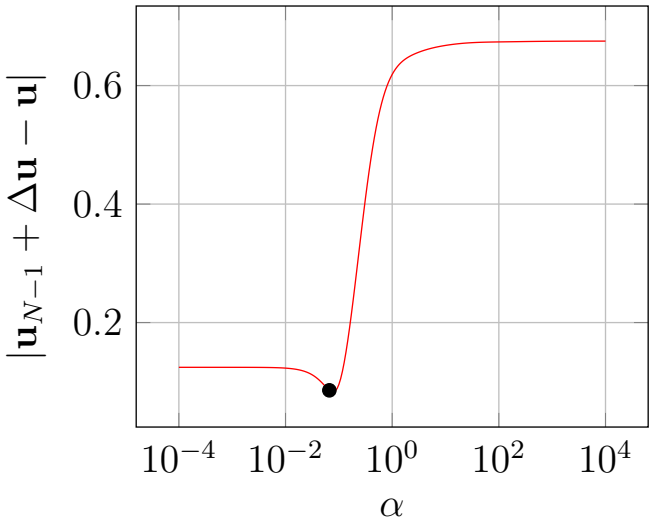

Figure 3.6: Profile of α at iteration 4. Corresponding to Table 3.1.

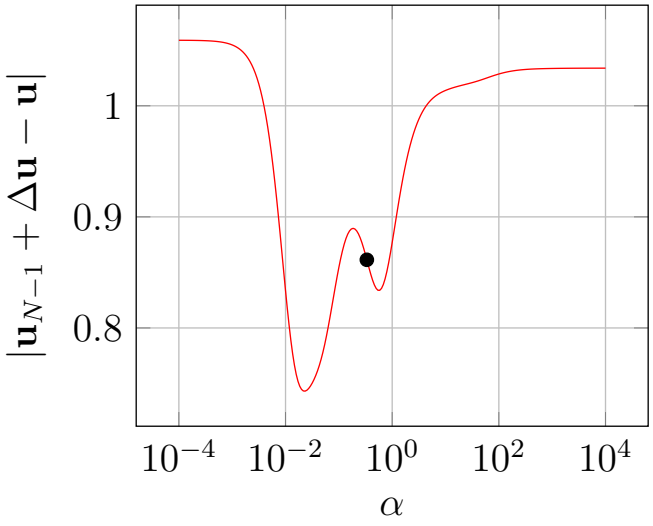

Figure 3.7: Profile of α at iteration 3. Corresponding to Table 3.2.

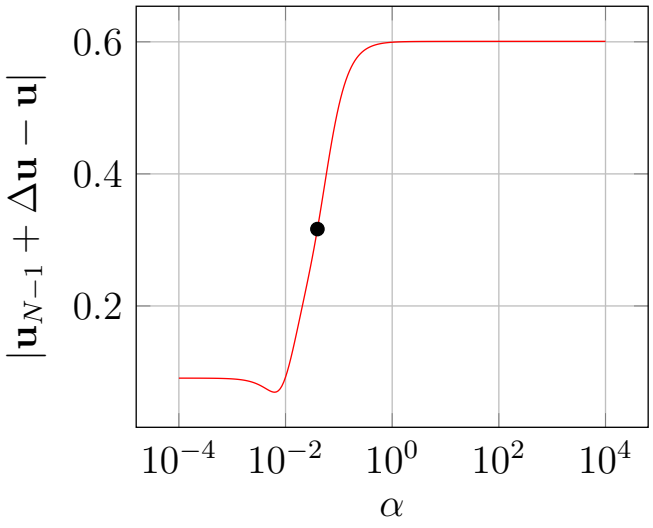

Figure 3.8: Profile of α at iteration 6. Corresponding to Table 3.2.

3.2.3 Convergence vs. Number of Data Points

The method is very robust and convergence is still the case if the number of data points is reduced to a single time measurement at the end of the time interval for each of the 8 components of the system. Figure 3.9 demonstrates the dependence of the number of time measurements for each component on the average error

$$\frac{1}{D}\sum_{d=1}^{D}||\mathbf{x}^d - \mathbf{x}_N^d||_{L_2^n}$$

calculated at the final iteration in the experiment with $D = 5$ data sets. Three graphs correspond to three different settings of the relative and absolute tolerances for CVODES. Decrease of the latter increases the overall accuracy of the result. Similar dependence in the experiment with 16 data sets and with CVODES tolerance being set up at $1E-6$ is demonstrated in Figure 3.10. Some of the variation in the chart can be attributed to error accumulation and noise. Table 3.1 demonstrates the final values of 36 parameters in a numerical experiment with D=5 data sets with 1, 20 and 240 time measurements for each of the 8 components.

3.2.4 Convergence vs. Number of Data Sets

Our numerical analysis confirms the result of [4] that at least 5 data sets with different control inputs are required to uniquely estimate the 36 parameters of this model. Tables 3.2 and A.21 demonstrate the results of the numerical experiments when the number of data sets vary from 1 to 5, and time measurements for each of the 8 components of the system is 240 and 20 respectively. Table 3.2 demonstrates that when the number of data sets increases from 1 to 5 with accuracy 10^{-3}, the number of identified parameters increases as 22, 27, 32, 34 and 36 accordingly,

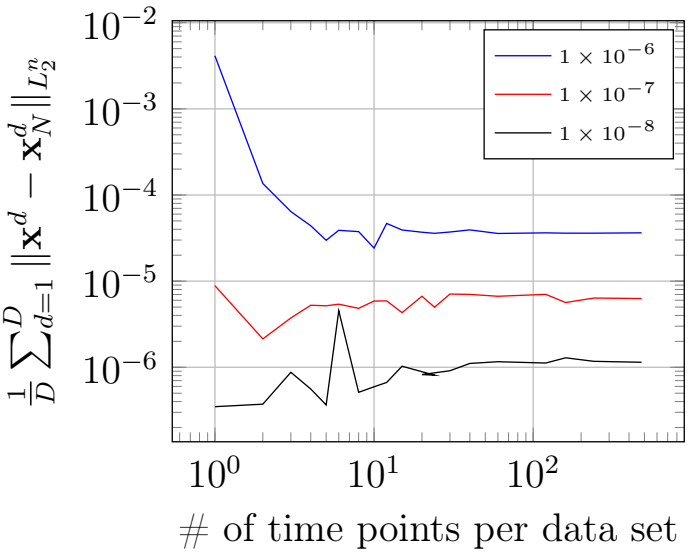

Figure 3.9: Graph of the number time points vs. $\frac{1}{D}\sum_{d=1}^{D}\|\mathbf{x}^d - \mathbf{x}_N^d\|_{L_2^n}$ using 5 data sets. With the relative and absolute tolerance for CVODES set to 1×10^{-6}, 1×10^{-7}, and 1×10^{-8}. Where $\mathbf{u}_0 = \mathbf{u} + 0.25\mathbf{u}$.

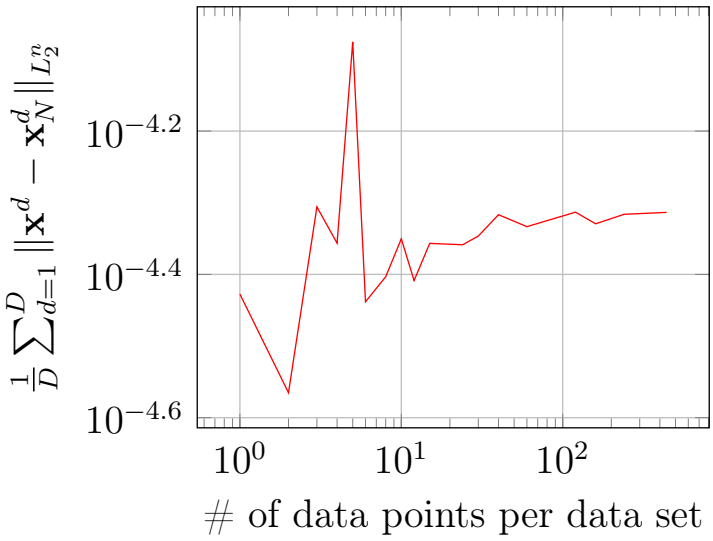

Figure 3.10: Graph of the number time points vs. $\frac{1}{D}\sum_{d=1}^{D}\|\mathbf{x}^d - \mathbf{x}_N^d\|_{L_2^n}$ using 16 data sets. With the relative and absolute tolerance for CVODES set to $1E-6$. Where $\mathbf{u}_0 = \mathbf{u} + 0.25\mathbf{u}$.

$\mathbf{u}$	1	20	240	$\mathbf{u}_{true}$
p_1	0.939	1.000	1.000	1
p_2	1.000	1.000	1.000	1
p_3	2.000	2.000	2.000	2
p_4	1.000	1.000	1.000	1
p_5	2.000	2.000	2.000	2
p_6	0.939	1.000	1.000	1
p_7	1.279	1.000	1.000	1
p_8	1.000	1.000	1.000	1
p_9	2.000	2.000	2.000	2
p_{10}	1.000	1.000	1.000	1
p_{11}	2.000	2.000	2.000	2
p_{12}	1.279	1.000	1.000	1
p_{13}	1.117	1.000	1.000	1
p_{14}	1.000	1.000	1.000	1
p_{15}	2.000	2.000	2.000	2
p_{16}	1.000	1.000	1.000	1
p_{17}	2.000	2.000	2.000	2
p_{18}	1.117	1.000	1.000	1
p_{19}	0.080	0.100	0.100	0.1
p_{20}	1.000	1.000	1.000	1
p_{21}	0.080	0.100	0.100	0.1
p_{22}	0.116	0.100	0.100	0.1
p_{23}	1.000	1.000	1.000	1
p_{24}	0.116	0.100	0.100	0.1
p_{25}	0.116	0.100	0.100	0.1
p_{26}	1.000	1.000	1.000	1
p_{27}	0.116	0.100	0.100	0.1
p_{28}	1.167	1.000	1.000	1
p_{29}	1.012	1.000	1.000	1
p_{30}	1.019	1.000	1.000	1
p_{31}	1.170	1.000	1.000	1
p_{32}	1.008	1.000	1.000	1
p_{33}	1.007	1.000	1.000	1
p_{34}	1.173	1.000	1.000	1
p_{35}	1.009	1.000	1.000	1
p_{36}	1.001	1.000	1.000	1

Table 3.1: The evolution of the parameters as the number of time points increase, from 1 to 5. Where 5 data sets were considered, $\mathbf{u}_0 = 1.25\mathbf{u}$. In each case regularization parameter α is chosen optimally.

provided that 240 time measurements are given. Table A.21 demonstrates that with 20 time measurements the same number increases as 11, 24, 32, 33, 36.

3.2.5 Range of convergence

We define the range of convergence as a neighborhood of the true solution $\mathbf{u}$ in $\mathbb{R}^{36}$ such that for any $\mathbf{u}_0$ chosen from it, the sequence $\mathbf{u}_N$ constructed according to our algorithm converges to $\mathbf{u}$. Consider the rectangular prism neighborhood of $\mathbf{u}$:

$$\mathcal{P}_{\tau}^{\omega} = \{p \in \mathbb{R}^{36} : \tau u_i \leq p_i \leq \omega u_i,\ i = 1, ..., 36\}$$

where τ and ω are two positive real numbers satisfying $\tau < 1 < \omega$. Numerical analysis demonstrates that for our model example, $\mathcal{P}_{0.5}^{1.65}$ is the largest rectangular prism contained in the convergence range according to the algorithm accompanied by Type I regularization. Tables A.22 and A.23 demonstrate the convergence with initial iteration $\mathbf{u}_0$ chosen at extremes of $\mathcal{P}_{0.5}^{1.65}$, namely $0.5\mathbf{u}$ and $1.65\mathbf{u}$ respectively.

Careful implementation of Type II regularization allows significant expansion of the convergence range. In fact, by selecting $\mathbf{u}^*$ at the extremes of $\mathcal{P}_{0.5}^{1.65}$, namely $\mathbf{u}^* = 0.5\mathbf{u}$ and $\mathbf{u}^* = 1.65\mathbf{u}$ we increased the convergence range to $\mathcal{P}_{0.03}^{\infty}$ according to the algorithm accompanied by Type II regularization. Table A.24 demonstrates the results of convergence of the method with Type II regularization when $\mathbf{u}^* = 0.5\mathbf{u}$, and initial iteration is chosen as $0.03\mathbf{u}$. Table A.25 demonstrates the results when $\mathbf{u}^* = 1.65\mathbf{u}$, and initial iteration is chosen as $1001\mathbf{u}$.

$\mathbf{u}$	1	2	3	4	5	$\mathbf{u}_{true}$
p_1	0.994	0.995	1.004	0.999	1.000	1
p_2	1.250	1.238	1.200	0.874	1.000	1
p_3	2.500	2.493	2.470	2.343	2.000	2
p_4	1.254	1.317	1.146	1.003	1.000	1
p_5	2.496	2.480	1.891	1.995	2.000	2
p_6	1.000	1.000	1.000	1.000	1.000	1
p_7	1.000	0.998	0.999	1.000	1.000	1
p_8	1.250	0.878	1.000	1.000	1.000	1
p_9	2.500	2.310	1.999	2.000	2.000	2
p_{10}	1.000	0.999	1.000	1.000	1.000	1
p_{11}	2.001	1.999	2.000	2.000	2.000	2
p_{12}	1.002	0.999	0.999	1.000	1.000	1
p_{13}	0.999	1.004	0.998	1.001	1.002	1
p_{14}	1.250	0.881	1.000	1.000	1.000	1
p_{15}	2.500	2.301	2.000	1.999	2.000	2
p_{16}	0.999	1.001	1.000	1.000	1.000	1
p_{17}	2.000	1.996	2.000	2.001	2.000	2
p_{18}	1.001	1.003	0.998	1.002	1.002	1
p_{19}	0.100	0.100	0.100	0.100	0.100	0.1
p_{20}	1.000	1.000	1.000	1.000	1.000	1
p_{21}	0.100	0.100	0.100	0.100	0.100	0.1
p_{22}	0.100	0.100	0.100	0.100	0.100	0.1
p_{23}	1.000	1.000	1.000	1.000	1.000	1
p_{24}	0.100	0.100	0.100	0.100	0.100	0.1
p_{25}	0.100	0.100	0.100	0.100	0.100	0.1
p_{26}	1.000	1.000	1.000	1.000	1.000	1
p_{27}	0.100	0.100	0.100	0.100	0.100	0.1
p_{28}	1.025	1.000	1.000	1.000	1.000	1
p_{29}	1.279	1.000	1.000	1.000	1.000	1
p_{30}	1.247	1.000	0.999	0.999	1.000	1
p_{31}	1.000	1.000	0.999	1.000	1.000	1
p_{32}	1.001	1.000	1.001	1.000	1.000	1
p_{33}	1.000	0.998	1.004	1.000	1.000	1
p_{34}	1.000	1.000	1.000	1.000	1.000	1
p_{35}	1.011	1.000	1.001	1.000	1.000	1
p_{36}	1.259	0.999	1.001	1.000	1.000	1

Table 3.2: The evolution of the parameters against the number of data sets, from 1 to 5. Where $\mathbf{u}_0 = 1.25\mathbf{u}$, $t_0 = 0$, $t_1 = 120$, $\Delta t = 0.5$, giving us 240 time points. In each case α was determined using (3.14).

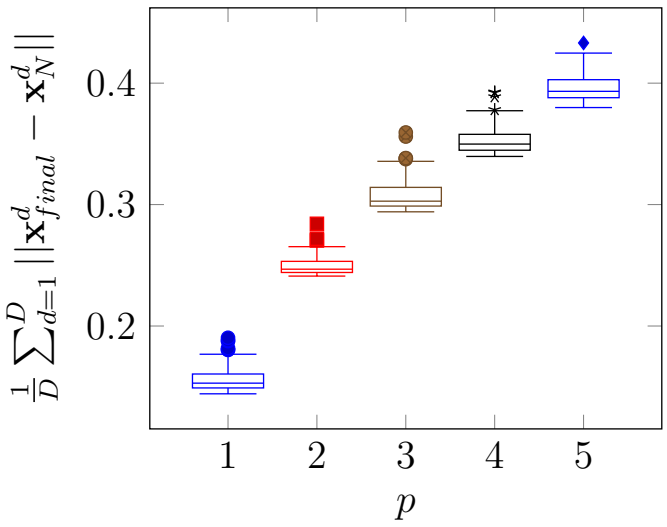

Figure 3.11: Distribution of the average residual error for several noise levels. Each run had 240 time points.

3.2.6 Convergence with Noisy Measurements

We pursued numerical experiments with simulated noisy data with Gaussian noise

$$y_i = x_i^d(t;\mathbf{u}) + px_i^d(t;\mathbf{u})\nu_i,\ \ i = 1, ..., n \tag{3.15}$$

where p is a percentage and ν_i is a random variable with standard normal distribution:

$$\nu_i \sim N(0,1).$$

Tables A.26, A.27, and A.28 demonstrate the convergence in the experiment with 5 data sets and 240 noisy time measurements with $p = 1, 2$ and 5 respectively. In Figures 3.11 and 3.12 we show the box plot based on 100 simulations for the residual and parameter vector error dependence on the noise percentage p. Similar results with 20 noisy time measurements are given in Tables A.29, A.30, and A.31.

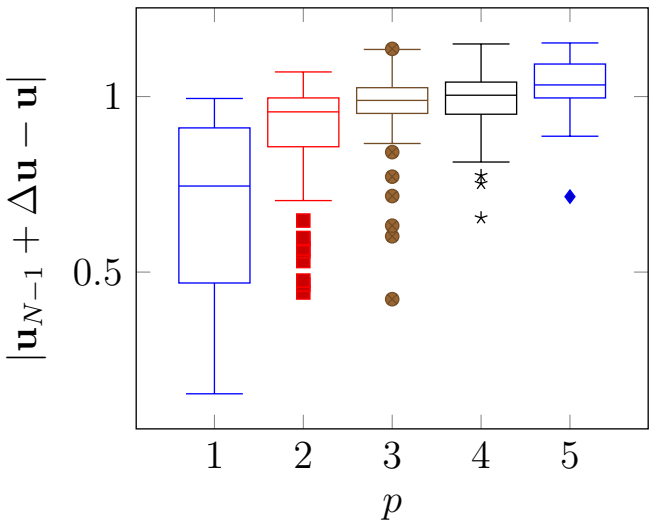

Figure 3.12: Distribution of the parameter error for the three step metabolic network at several noise levels. Each run had 240 time steps.

3.2.7 Rate of convergence

To estimate the convergence rate γ from the relation

$$|\mathbf{u}_{k+1} - \mathbf{u}_k| \sim C|\mathbf{u}_k - \mathbf{u}_{k-1}|^\gamma$$

we plot $\log|\mathbf{u}_{k+1} - \mathbf{u}_k|$ vs. $\log|\mathbf{u}_k - \mathbf{u}_{k-1}|$ and find a line of best fit to identify γ and C. Figures 3.13 and 3.14 demonstrate the outcome. For the numerical experiment from Table A.17 we have $\gamma = 1.6104, C = 3.1622E-3$, and for results from Table A.18 we have $\gamma = 1.1674$, $C = 6.3271E-1$. The difference in convergence rate of two examples is in particular due to choice of the regularization parameter α. Almost optimal choice of α, as it is demonstrated in Figures 3.5 and 3.6 vs. Figures 3.7 and 3.8, causes higher convergence rate for the numerical experiment in Table A.17 vs. Table A.18. We expect theoretical convergence rate of the method is quadratic [38].

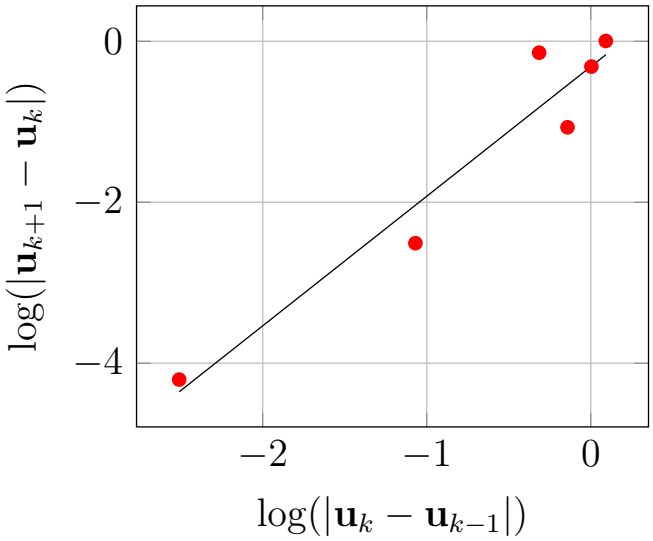

Figure 3.13: The convergence rate graph corresponding to Table 3.1, where $r = 1.6104$ and $C = 3.1622E - 3$.

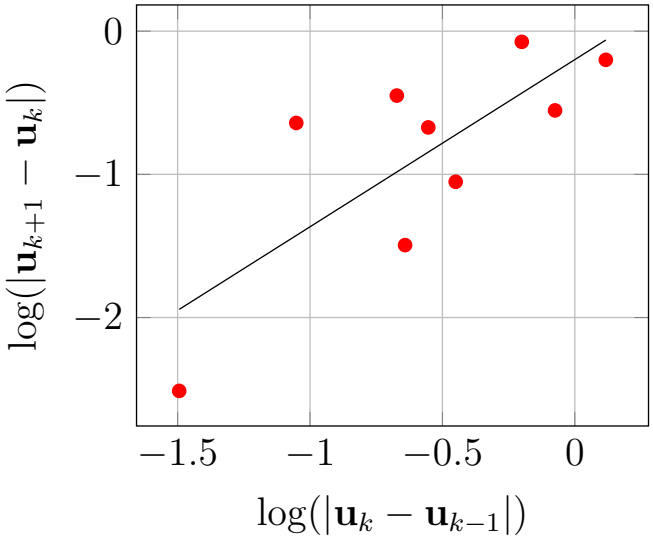

Figure 3.14: The convergence rate graph corresponding to Table 3.2, where $r = 1.1674$ and $C = 6.3271E - 1$.

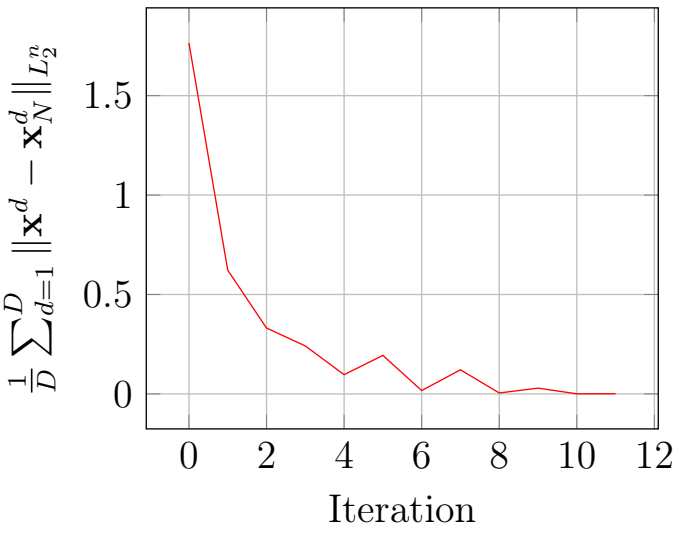

Figure 3.15: The average error at each iteration corresponding to Table A.32.

3.2.8 Convergence with Partial Measurements

We tested the convergence of the method when only some of the components of the system have available measurements or partial measurements. In this case the inverse problem must be solved with partial observations. A typical result is demonstrated In Table A.32 and Figure 3.15. We considered our numerical experiment with 5 data sets, and with 20 time measurements of only components 3, 4, 5, and 7. As can be seen from Table A.32 and Figure 3.15, the iteration converges to the true solution, but some small error is present in the final value of the parameter vector.

3.2.9 Comparison with *lsqnonlin* and *nl2sol*

As in chapter 2 we are comparing our method *qlopt* with the most popular methods available as open software [19, 33, 34]. We consider the two methods, *lsqnonlin*, *nl2sol*. We used a model example provided by AMIGO2, which had 16 data sets

with each component evaluated at 21 time points giving a total of 2688 data points. In both AMIGO2 and *qlopt* we used CVODES for acquiring the solution to the systems (3.4) and (3.7). We implemented CVODES by copying the implementation in AMIGO2, though we used a later version. We also matched the settings in our software to the example file *mendes3step_PE.m* in AMIGO2, in particular we added the weights on the objective function. The termination condition included both objective function and parameter search direction absolute and relative errors. All these tolerances were set to 1×10^{-7}.

We ran each algorithm 20 times and recorded the *average relative error of the parameter values (r.e.), the median number of objective function evaluations (f.e.), the average computational time (c.t.), and the median number of iterations (n.i.)*. The results are demonstrated in Table 3.3. All three methods have a comparable relative error. In terms of required number of iterations, our method is comparable to *nl2sol*, and both have a clear advantage over *lsqnonlin*. In terms of *computational time* and *function evaluations* our method has an enormous advantage over both methods. The algorithm *fmincon* did not converge for the settings considered. It should be noted that our software package *qlopt* is using entirely in C/C++, which gives it an advantage over MatLab-based methods with respect to computational time.

metric	qlopt	lsqnonlin	nl2sol
r.e.	4.271×10^{-4}	3.236×10^{-5}	4.354×10^{-4}
n.i.	8	16	7
c.t. (s)	1.389	7.656	5.043
f.e.	8	593	299

Table 3.3: Comparison of several local optimization method against the presented method for the three step metabolic network. The initial guess was $\mathbf{u}_0 = 1.25\mathbf{u}$ the relative errors (r.e.), the number of iterations (n.i.), mean of the computational time (c.t.) of 20 runs, and the number of function evaluations (f.e.). We considered the data sets provided by AMIGO2 which contained 16 datasets evaluated a 21 time values giving us a total of 2688 data points.

Chapter 4

Conclusions

Dissertation implements the numerical method introduced in [38] and develops a new significantly modified method for the solution of the inverse problem on the identification of the finite dimensional parameter set for nonlinear dynamical systems arising in systems biology. The idea of the method is based on Pontryagin optimization, Bellman quasilinearization and Tikhonov regularization. Inverse problems are highly ill-posed. The additional major difficulty of solving inverse problems for canonical models of systems biology is nonlinearity of the system both with respect to state vectors, and with respect to parameters. This dissertation demonstrates that the developed method is extremely well-suited for solving inverse problems for such systems. The main idea of the method is the combination of the ideas of linearization of the system with respect to state vector and subsequent application of the sensitivity analysis with respect to parameters (or equivalently, linearization with respect to parameters) and Tikhonov regularization. Two types of regularization is implemented. Type I regularization is implemented for the last step of the algorithm where parameter increments are found from the sys-

tem with an ill-conditioned matrix. Type II regularization is meant to address the non-uniqueness of the solution to inverse problem, and to increase the convergence range with respect to initial iteration. The method is applied the method to the classical Lotka-Volterra model describing predator-prey style competition, the model of a genetic circuit with bistable switch mechanism, the synthetic biology toggle switch model which describes gene transcription regulation by repressors, and the synthetic biology repressilator model describing genetic network for three repressor-protein and their corresponding mRNA concentrations. The method is very effective for the moderate size parameter sets. The main conclusions of the implementation of the method can be summarized as follows:

- If the size of the *identifiable* parameter vector is relatively small then the method has a local quadratic convergence rate with or without regularization.
- the algorithm is robust with respect to Gaussian noise added to the simulated data
- regularization significantly increases the convergence range with respect to initial guess.
- if the size of the parameter vector is high, regularization is essential to have quadratic convergence.
- optimal choice of the regularization parameter significantly increases the convergence rate and precision.
- the well-posedness and convergence of the method is directly associated with the identifiability of the parameters

- accompanied with regularization the method presents a robust algorithm to address practical non-identifiability cases.

- the method has an advantage over popular methods such as *lsqnonlin, fmincon, nl2sol* in terms of computational time, number of iterations and function evaluations.

Direct adaptation of the method to inverse problems with significantly large size of the parameter set is not as effective. For the adaptation and scalability of the method and developed software package *qlopt* to inverse problems with significantly large size of the parameter set, dissertation suggested a significant modification of the method. Modification is twofold: method of staggered corrector is embedded into the step for sensitivity analysis, and the software package CVODES [44] is connected with the developed software package *qlopt*; multi-objective optimization is enhanced into the method which allows for the application of the method to large-scale models with a practically non-identifiable set of parameters based on multiple data sets, possibly with partial and noisy measurements. The modified method is applied to a benchmark model of a biological network for a three-step pathway modeled by 8 nonlinear ODEs describing 8 metabolic concentrations with 36 unknown parameters, and two control input parameters specified by the experimental design. Extensive analysis demonstrates that the modified method is extremely well adapted to large scale problems. The main conclusions of the second part of the dissertation can be summarized as follows:

- There is a minimum number of data sets with different control parameter inputs required to achieve geometric convergence and unique identifiability of parameters for large-scale problems. The method has a geometric conver-

gence and high accuracy for the benchmark model if at least five data sets with different control parameter inputs are used.

- Increase of data sets beyond the minimum doesn't significantly affect convergence rate and accuracy, but possibly affects the computational cost.

- The method is extremely robust in terms of required number of time measurements for components of the system for every data set. For the benchmark model, high accuracy is achieved if the number of time measurements is between 1 and 240 in a segment [0,120].

- Optimal choice of the Tikhonov regularization parameter significantly increases the convergence rate and precision.

- The method is robust with respect to noisy measurements. Simulating up to 5% Gaussian noise in a benchmark model does not affect the convergence rate, but only adds some additional error to final output in accordance with the noise level.

- Implementation of the Type II Tikhonov regularization significantly increases the convergence range of the algorithm.

- Method is robust with respect to partial measurements. Application to the benchmark model with measurements of only four components instead of eight demonstrates convergence with slightly reduced but still quite high accuracy.

- The method is highly competitive and has an advantage over popular methods such as *lsqnonlin, fmincon, nl2sol* in terms of computational time, number of iterations and function evaluations.

Bibliography

[1] Andreĭ N Tikhonov and Vasiliĭ Y. Arsenin. *Methods for Solving Ill-posed Problems*. John Wiley & Sons, 1977.

[2] Attila Gábor and Julio R Banga. Robust and efficient parameter estimation in dynamic models of biological systems. *BMC systems biology*, 9(1):74, 2015.

[3] Stefania Audoly, Giuseppina Bellu, Leontina D'Angio, Maria Pia Saccomani, and Claudio Cobelli. Global identifiability of nonlinear models of biological systems. *IEEE Transactions on Biomedical Engineering*, 48(1):55–65, 2001.

[4] Pu Li and Quoc Dong Vu. Identification of parameter correlations for parameter estimation in dynamic biological models. *BMC Systems Biology*, 7(1):91, Sep 2013.

[5] Michael J Chappell, Keith R Godfrey, and Sandor Vajda. Global identifiability of the parameters of nonlinear systems with specified inputs: A comparison of methods. *Mathematical Biosciences*, 102(1):41–73, 1990.

[6] Andreas Raue, Clemens Kreutz, T Maiwald, U Klingmüller, and Jens Timmer. Addressing parameter identifiability by model-based experimentation. *IET Systems Biology*, 5(2):120–130, 2011.

[7] Oana-Teodora Chis, Julio R Banga, and Eva Balsa-Canto. Structural identifiability of systems biology models: A critical comparison of methods. *PLOS ONE*, 6(11):1–16, 11 2011.

[8] Maria Rodriguez-Fernandez, Pedro Mendes, and Julio R Banga. A hybrid approach for efficient and robust parameter estimation in biochemical pathways. *Biosystems*, 83(2):248 – 265, 2006. 5th International Conference on Systems Biology.

[9] Rajko Tomović. *Sensitivity analysis of dynamic systems*. McGraw-Hill electronic sciences series. McGraw-Hill, New York, NY, 1963.

[10] Martin Hanke. A regularizing levenberg - marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Problems*, 13(1):79–95, 02 1997.

[11] Heinz W Engl, Christoph Flamm, Philipp Kügler, James Lu, Stefan Müller, and Peter Schuster. Inverse problems in systems biology. *Inverse Problems*, 25(12):123014, dec 2009.

[12] Eberhard Voit and I-Chun Chou. Parameter estimation in canonical biological systems models. *International Journal of Systems and Synthetic Biology*, 1(1):1–19, 2010.

[13] I-Chun Chou and Eberhard O Voit. Recent developments in parameter estimation and structure identification of biochemical and genomic systems. *Mathematical Biosciences*, 219(2):57 – 83, 2009.

[14] Mary J Dunlop and Richard M Murray. Towards biological system identification: fast and accurate estimates of parameters in genetic regulatory networks. *Parameters*, page 4, 2006.

[15] Yanping Qiu, Tao Hu, Baosheng Liang, and Hengjian Cui. Robust estimation of parameters in nonlinear ordinary differential equation models. *Journal of Systems Science and Complexity*, 29(1):41–60, 2016.

[16] James T MacDonald, Chris Barnes, Richard I Kitney, Paul S Freemont, and Guy-Bart V. Stan. Computational design approaches and tools for synthetic biology. *Integrative Biology*, 3(2):97–108, 01 2011.

[17] Daniel E Zak, Gregory E Gonye, James S Schwaber, and Francis J Doyle. Importance of input perturbations and stochastic gene expression in the reverse engineering of genetic regulatory networks: insights from an identifiability analysis of an in silico network. *Genome Research*, 13(11):2396–2405, 2003.

[18] Michael B Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, 2000.

[19] Eva Balsa-Canto, David Henriques, Attila Gábor, and Julio R. Banga. Amigo2, a toolbox for dynamic modeling, optimization and control in systems biology. *Bioinformatics*, 32(21):3357–3359, 2016.

[20] John E Dennis Jr, David M Gay, and Roy E Welsch. An adaptive nonlinear least-squares algorithm. *ACM Transactions on Mathematical Software*, 7(3):369–383, 1981.

[21] Fabian Fröhlich, Barbara Kaltenbacher, Fabian J Theis, and Jan Hasenauer. Scalable parameter estimation for genome-scale biochemical reaction networks. *PLOS Computational Biology*, 13(1):1–18, 01 2017.

[22] Ernesto Andrianantoandro, Subhayu Basu, David K Karig, and Ron Weiss. Synthetic biology: new engineering rules for an emerging discipline. *Molecular Systems Biology*, 2(1), 05 2006.

[23] Andreas Raue, Marcel Schilling, Julie Bachmann, Andrew Matteson, Max Schelke, Daniel Kaschek, Sabine Hug, Clemens Kreutz, Brian D Harms, Fabian J Theis, Ursula Klingmüller, and Jens Timmer. Lessons learned from quantitative dynamical modeling in systems biology. *PLOS ONE*, 8(9):1–17, 09 2013.

[24] Alejandro F Villaverde, Fabian Fröhlich, Daniel Weindl, Jan Hasenauer, and Julio R Banga. Benchmarking optimization methods for parameter estimation in large kinetic models. *Bioinformatics*, 35(5):830–838, 03 2019.

[25] K Schittkowski. *Numerical Data Fiting in Dynamical Systems: A Practical Introduction with Applications and Software*, volume 77. Springer Science and Business Media, 2013.

[26] MH Wright. *Direct Search Methods: once scorned, now respectable*, volume 344. Pittman Res. Notes Math. Ser., 1996.

[27] AR Conn and et al. *Introduction to Derivative-Free Optimization*. SIAM, Philadelphia, 2009.

[28] Anatoly Anatoly Zhigljavsky and Antanas Žilinskas. *Stochastic Global Optimization*. Springer US, 1 edition, 2008.

[29] William R Esposito and Christodoulos A Floudas. Global optimization for the parameter estimation of differential-algebraic systems. *Industrial & Engineering Chemistry Research*, 39(5):1291–1310, 2000.

[30] Johann Dréo, Alan Pétrowski, Patrick Siarry, and Eric Taillard. *Metaheuristics for Hard Optimization: Methods and Case Studies*. Springer, 2006.

[31] Carmen G Moles, Pedro Mendes, and Julio R Banga. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome research*, 13(11):2467–2474, 2003.

[32] Art B. Owen. A central limit theorem for latin hypercube sampling. *Journal of the Royal Statistical Society, Series B*, 54(2):541–551, 1992.

[33] Andreas Raue, Bernhard Steiert, Max Schelker, Clemens Kreutz, Tim Maiwald, Helge Hass, Joep Vanlier, C Tönsing, L Adlung, R Engesser, Wolfgang Mader, Tim Heinemann, Jan Hasenauer, Marcel Schilling, Thomas Höfer, Edda Klipp, Fabian J Theis, Ursula Klingmüller, Birgit Schoeberl, and Jens Timmer. Data2Dynamics: a modeling environment tailored to parameter estimation in dynamical systems. *Bioinformatics*, 31(21):3558–3560, 07 2015.

[34] Paul Stapor, Daniel Weindl, Benjamin Ballnus, Sabine Hug, Carolin Loos, Anna Fiedler, Sabrina Krause, Sabrina Hroß, Fabian Fröhlich, and Jan Hasenauer. PESTO: Parameter EStimation TOolbox. *Bioinformatics*, 34(4):705–707, 10 2017.

[35] M Kronfeld, H Planatscher, and A Zell. The eva2 optimization framework. *Learning and Intelligent Optimization*, 6073:247–250, 2010.

[36] T Coleman and Y Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6:418–445, 1996.

[37] J Kennedy and R Eberhart. Particla swatm optimization. *Proceedings of IEEE*, pages 1942–1945, 1995.

[38] Ugur G. Abdullaev. Quasilinearization and inverse problems of nonlinear dynamics. *Journal of Optimization Theory and Applications*, 85(3):509–526, 1995.

[39] Ugur G. Abdullaev. Quasilinearization and inverse problems for nonlinear control systems. *Journal of Optimization Theory and Applications*, 85(3):527–543, 1995.

[40] Ugur G Abdulla and Roby Poteau. QLOPT. https://github.com/robypoteau/qlopt, 2018.

[41] Ugur G Abdulla and Roby Poteau. Identification of parameters in systems biology. *Mathematical Biosciences*, 305:133–145, 2018.

[42] William F Feehery, John E Tolsma, and Paul I Barton. Efficient sensitivity analysis of large-scale differential-algebraic systems. *Applied Numerical Mathematics*, 25(1):41 – 54, 1997.

[43] Richard E Bellman and Robert E. Kalaba. *Quasilinearization and nonlinear boundary-value problems*. RAND Corporation, Santa Monica, CA, 1965.

[44] Alan C Hindmarsh, Peter N Brown, Keith E Grant, Steven L Lee, Radu Serban, Dan E Shumaker, and Carol S. Woodward. Sundials: Suite of nonlinear

and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software*, 31(3):363–396, 2005.

[45] Ugur G Abdulla and Roby Poteau. Identification of Parameters for Large-scale Models in Systems Biology. *arXiv e-prints*, page arXiv:1904.07197, 04 2019.

[46] Edwin F Beckenbach and Richard E. Bellman. *Inequalities*. Springer Verlag, Berlin, New York, NY, 1961.

[47] Lennart Ljung. *System Identification: Theory for the User*. Prentice Hall, Englewood Cliffs, NJ, 1999.

[48] Mark Galassi, Jim Davies, James Theiler, Brian Gough, Gerard Jungman, Patrick Alken, Michael Booth, Fabrice Rossi, and Rhys Ulerich. Gnu scientific library reference manual, 2009.

[49] James Demmel, Laura Grigori, Mark Hoemmen, and Julien Langou. Communication-optimal parallel and sequential qr and lu factorizations. *SIAM J. Sci. Comput.*, 34(1):206–239, February 2012.

[50] Gaël Guennebaud, Benoît Jacob, Philip Avery, et al. Eigen v3. http://eigen.tuxfamily.org, 2010.

[51] Timothy S Gardner, Charles R Cantor, and James J. Collins. Construction of a genetic toggle switch in escherichia coli. *Nature*, 403(6767):339–342, 2000.

[52] Edward N Lorenz. Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141, 1963.

[53] Arnold Neumaier, Oleg Shcherbina, Waltraud Huyer, and Tamás Vinkó. A comparison of complete global optimization solvers. *Mathematical Programming*, 103:355, 2005.

[54] Kendall E Atkinson. *An Introduction to Numerical Analysis*. John Wiley & Sons, 2 edition, 1989.

[55] Richard L Burden and J Douglas Faires. *Numerical Analysis*. Thomas Brooks/Cole, Belmont, CA, 8 edition, 2005.

[56] Pedro Mendes. *Foundations of Systems Biology*, chapter 8, pages 163–186. MIT Press, Cambridge, MA, 2001.

Appendix A

Tables with Numerical Results

# of data point	$\|\|\mathbf{x} - \mathbf{x}_N\|\|_{L_2^n}$
1001	0.065 31
501	0.065 30
201	0.065 24
101	0.065 26
51	0.065 33
34	0.067 44
21	0.065 18
11	0.386 93
9	1.619 55
6	15.328 30
3	38.450 80

Table A.1: Table of the number of data points vs. $||\mathbf{x} - \mathbf{x}_N||_{L_2^n}$ for the Lotka Volterra model. The parameters are $\mathbf{u}_0 = (0.05, 0.05, 0.05)$, $\mathbf{u} = (0.48, 0.026, 0.93)$, $\mathbf{x}_0 = (35, 4)$, time interval $[0, 2]$ and respective tolerances were set at 1×10^{-7}.

time	0	1	2	3	4
hare	30	47.2	70.2	77.4	36.3
lynx	4	6.1	9.8	35.2	59.4
time	5	6	7	8	9
hare	20.6	18.1	21.4	22	25.4
lynx	41.7	19	13	8.3	9.1
time	10	11	12	13	14
hare	27.1	40.3	57	76.6	52.3
lynx	7.4	8	12.3	19.5	45.7
time	15	16	17	18	19
hare	19.5	11.2	7.6	14.6	16.2
lynx	51.1	29.7	15.8	9.7	10.1

Table A.2: Sample of the population for Canadian lynx and hare (in 1000s) for 20 years.

time	Prey	x_1	x_5	x_{10}
0	30	30	30	30
0.2	33.124 340	29.814 330	32.557 110	33.080 870
0.4	36.327 590	29.620 120	35.371 990	36.317 840
0.6	39.688 670	29.420 790	39.132 550	39.717 810
0.8	43.286 500	29.217 630	45.538 690	43.306 690
1	47.200 000	29.008 190	59.577 800	47.135 550
time	Pred. Pop.	y_1	y_5	y_{10}
0	4	4	4	4
0.2	4.602 303	4.237 255	3.886 941	4.606 703
0.4	5.159 030	4.482 176	4.092 876	5.163 934
0.6	5.624 606	4.731 336	4.775 209	5.626 564
0.8	5.953 455	4.983 421	6.550 910	5.950 192
1	6.100 000	5.240 870	11.442 560	6.096 148
	Guess	$\mathbf{u}_1$	$\mathbf{u}_5$	$\mathbf{u}_{10}$
u_1	0.010 000	−0.969 868	0.047 577	0.650 685
u_2	0.010 000	−0.360 320	−0.082 554	0.037 601
u_3	0.010 000	0.213 387	0.049 103	−0.042 786
u_4	0.010 000	5.656 212	1.274 734	−2.055 026

Table A.3: First of three successive runs of the method on the lynx-hare data set in the time interval $[0, 1]$

time	Prey Pop.	x_1	x_7	x_{14}
0	30	30	30	30
1	36.300 000	47.135 550	49.610 990	47.327 440
2	22	73.847 010	70.194 710	70.622 250
3	57	124.371 200	70.483 990	77.057 860
4	11.200 000	191.751 300	35.813 680	35.223 780
5	24.700 000	511.096 800	−5.320 904	14.344 960
time	Pred. Pop.	y_1	y_7	y_{14}
0	4	4	4	4
1	59.400 000	6.096 146	7.854 519	4.413 269
2	8.300 000	3.694 961	15.924 430	8.920 003
3	12.300 000	−14.702 590	32.984 970	31.114 480
4	29.700 000	−229.176 800	66.400 460	60.509 260
5	8.600 000	−1573.827 000	126.696 000	42.256 530
	Guess	$\mathbf{u}_1$	$\mathbf{u}_7$	$\mathbf{u}_{14}$
u_1	0.650 685	0.759 647	0.580 134	0.566 179
u_2	0.037 601	0.031 269	0.026 274	0.027 392
u_3	−0.042 786	−0.006 681	0.019 486	0.029 086
u_4	−2.055 026	−1.991 581	0.428 631	1.007 009

Table A.4: Second of three successive runs of the method on the lynx-hare data set in the time interval $[0, 5]$.

time	Prey Pop.	x_1	x_6	x_{11}
0	30	30	30	30
4	36.300 000	35.223 790	41.655 430	41.722 230
8	22	17.034 390	10.301 940	10.314 690
12	57	77.020 510	66.868 110	66.872 440
16	11.200 000	15.247 280	10.357 930	10.374 030
20	24.700 000	55.183 360	26.373 980	26.402 330
time	Pred. Pop.	y_1	y_6	y_{11}
0	4	4	4	4
4	59.400 000	60.509 290	51.093 790	51.132 860
8	8.300 000	5.626 455	7.331 523	7.328 381
12	12.300 000	20.375 010	6.137 608	6.138 105
16	29.700 000	−2.733 334	27.348 860	27.363 980
20	8.600 000	10.804 060	3.869 743	3.872 246
	Guess	$\mathbf{u}_1$	$\mathbf{u}_6$	$\mathbf{u}_{11}$
u_1	0.566 179	0.547 074	0.552 239	0.552 091
u_2	0.027 392	0.029 881	0.029 920	0.029 905
u_3	0.029 086	0.027 183	0.026 024	0.026 028
u_4	1.007 009	0.892 991	0.869 701	0.870 080

Table A.5: Third of three successive runs of the method on the lynx-hare data set in the time interval $[0, 20]$.

# of data point	$\|\|\mathbf{x} - \mathbf{x}_N\|\|$
3001	0.001 837 7
1501	0.001 837 7
301	0.001 834 6
151	0.001 835 7
31	0.001 787 8
16	0.001 641 2
7	0.008 680 5
4	0.281 322 5

Table A.6: Table of the number of data points vs. $||\mathbf{x}-\mathbf{x}_N||$ for the bistable switch model. The parameters were $\mathbf{u} = (150, 3.2, 2.0)$, $\mathbf{u}_0 = (125, 2.2, 1.0)$, $\mathbf{x}_0 = (14, 13)$, $t \in (0, 3)$ and respective tolerances were set at 1×10^{-7}.

time	x_{10}	x_{25}	x_{40}	x
0	1	1	1	1
2	−696.785 900	−16.733 580	−8.170 151	−8.170 151
4	-18198.47	−53.288 750	−10.084 270	−10.084 270
6	$1.109\,448 \times 10^{7}$	308.272 200	−9.749 939	−9.749 939
8	$1.391\,787 \times 10^{9}$	4918.989	−7.028 439	−7.028 439
10	$-3.456\,365 \times 10^{11}$	-60566.51	−4.913 545	−4.913 545
time	y_{10}	y_{25}	y_{40}	y
0	1	1	1	1
2	-1011.31	−13.839 500	−9.555 516	−9.555 516
4	-146694.5	−150.075 800	−10.929 550	−10.929 550
6	4484442	−453.638 300	−7.725 792	−7.725 792
8	$3.086\,984 \times 10^{9}$	9918.452	−4.594 378	−4.594 378
10	$-4.514\,722 \times 10^{11}$	-75741.9	−3.740 552	−3.740 552
time	z_{10}	z_{25}	z_{40}	z
0	1	1	1	1
2	897.333 800	54.148 460	24.622 170	24.622 170
4	-174652.6	10.245 470	27.957 490	27.957 490
6	$-2.779\,453 \times 10^{7}$	-1251.444	30.884 680	30.884 680
8	$5.154\,351 \times 10^{8}$	-2682.492	28.411 410	28.411 410
10	$4.110\,991 \times 10^{11}$	79370.87	24.724 430	24.724 430
	$\mathbf{u}_{10}$	$\mathbf{u}_{25}$	$\mathbf{u}_{40}$	$\mathbf{u}_{true}$
σ	8.672 728	7.117 864	10	10
r	34.136 870	32.151 950	28	28
b	0.010 142	1.887 447	2.666 000	2.666 000

Table A.7: Numerical results for Lorenz system using the parameters: $\mathbf{u}_0 = (5, 15, 5)$, $\mathbf{u} = (10, 28, 2.666)$, $\mathbf{y}_0 = (1, 1, 1)$, time interval $[0, 10]$, with step-size 0.005, and tolerance $|\Delta\mathbf{u}| < 0.00001$.

time	**Actual**	**Noisy**	x_4	x_7	x_{10}
0	1	−1.229 254	−1.229 254	−1.229 254	−1.229 254
0.4	15.365 610	14.696 940	16.141 140	15.369 170	15.368 720
0.8	−8.632 609	−8.004 626	−9.178 356	−8.611 755	−8.611 613
1.2	−7.175 225	−7.670 839	−7.802 890	−7.525 111	−7.524 957
1.6	−9.642 507	−9.967 407	−9.501 689	−9.812 583	−9.812 718
2	−8.170 152	−8.233 643	−9.672 875	−8.805 806	−8.805 618
time	**Actual**	**Noisy**	y_4	y_7	y_{10}
0	1	2.613 948	2.613 948	2.613 948	2.613 948
0.4	1.112 024	6.438 095	−0.687 615	−1.545 745	−1.546 393
0.8	−10.090 680	−9.684 571	−10.347 710	−9.935 833	−9.935 800
1.2	−6.784 048	−8.116 521	−7.731 765	−7.055 237	−7.054 975
1.6	−8.811 513	−9.187 876	−8.294 911	−8.981 821	−8.981 972
2	−9.555 516	−8.044 922	−10.932 450	−10.107 720	−10.107 620
time	**Actual**	**Noisy**	z_4	z_7	z_{10}
0	1	−5.432 259	−5.432 259	−5.432 259	−5.432 259
0.4	46.758 880	43.027 260	48.631 990	47.640 090	47.639 810
0.8	25.647 230	26.788 310	25.224 820	24.076 130	24.075 800
1.2	25.981 190	25.530 140	25.126 420	25.282 880	25.282 970
1.6	29.435 660	29.209 520	28.584 650	28.344 660	28.344 850
2	24.622 170	24.343 510	25.560 900	24.155 080	24.154 750
	$\mathbf{u}_0$		$\mathbf{u}_4$	$\mathbf{u}_7$	$\mathbf{u}_{10}$
σ	10		11.022 290	11.049 860	11.049 870
r	28		27.058 150	27.059 440	27.059 440
b	2.666 000		2.963 891	2.944 664	2.944 662

Table A.8: Numerical results for Lorenz system using the parameters: $\mathbf{u}_0 = (5, 15, 5)$, $\mathbf{u} = (10, 28, 2.666)$, $\mathbf{y}_0 = (1, 1, 1)$, time interval $[0, 2]$, with step-size 0.005, $\rho = 1$, and tolerance $|\Delta\mathbf{u}| < 0.00001$.

time	p_{A13}	p_{A21}	p_{A29}	p_A
0	20	20	20	20
1.2	4.525 584	4.605 430	4.586 361	4.600 216
2.4	1.730 465	1.807 860	1.786 334	1.791 909
3.6	0.932 633	1.150 268	1.179 931	1.199 278
4.8	0.422 593	1.060 460	0.924 043	0.934 925
6	−0.953 727	1.114 969	0.761 619	0.758 231
time	p_{B13}	p_{B21}	p_{B29}	p_B
0	20	20	20	20
1.2	6.947 640	6.661 002	6.211 452	6.206 941
2.4	4.485 700	1.880 450	2.980 094	2.985 900
3.6	9.938 839	−0.590 146	3.057 195	3.064 446
4.8	35.344 958	−2.313 971	3.511 337	3.521 464
6	134.471 653	−3.477 956	3.867 927	3.897 093
$\mathbf{u}$	$\mathbf{u}_{13}$	$\mathbf{u}_{21}$	$\mathbf{u}_{29}$	$\mathbf{u}_{true}$
c_A	1.534 990	1.484 411	3.878 841	5
d_A	1.561 212	2.812 863	2.836 690	3
K_B	11.446 937	5.270 093	2.411 962	2
f_A	1.039 785	0.999 691	1.001 051	1
g_A	3.246 963	4.773 537	4.797 418	5
c_B	4.263 347	7.099 480	5.958 913	6
d_B	−2.149 411	0.217 913	3.861 982	4
K_A	0.777 123	1.423 297	2.019 354	2
f_B	0.923 806	0.984 908	1.001 304	1
g_B	−1.235 480	1.132 451	4.853 971	5

Table A.9: Numerical results for the synthetic biology toggle switch model using the paramters: $\mathbf{u}_0 = (2, 1, 1, 1, 1, 2, 1, 1, 1, 2)$, $\mathbf{u} = (5, 3, 2, 1, 5, 6, 4, 2, 1, 5)$, $\mathbf{u}^* = (4, 2, 3, .5, 4, 5, 3, 3, .5, 4)$, $\mathbf{x}_0 = (20, 20, 20, 20)$, $\alpha = 0.005$, time interval $[0, 6]$, with step-size 0.005 and tolerance $|\Delta\mathbf{u}| < .00001$.

time	p_{15}	p_{30}	p_{45}	p_{60}	p
0	20	20	20	20	20
1.2	4.595 684	4.600 148	4.600 215	4.600 216	4.600 216
2.4	1.805 337	1.792 210	1.791 914	1.791 909	1.791 909
3.6	1.176 095	1.198 955	1.199 273	1.199 278	1.199 278
4.8	0.924 571	0.934 860	0.934 925	0.934 925	0.934 925
6	0.775 879	0.758 675	0.758 239	0.758 231	0.758 231

time	p_{15}	p_{30}	p_{45}	p_{60}	p
0	20	20	20	20	20
1.2	6.209 291	6.206 982	6.206 942	6.206 941	6.206 941
2.4	2.980 981	2.985 765	2.985 898	2.985 900	2.985 900
3.6	3.067 975	3.064 406	3.064 445	3.064 446	3.064 446
4.8	3.540 113	3.521 675	3.521 467	3.521 464	3.521 464
6	3.895 746	3.896 946	3.897 090	3.897 093	3.897 093

$\mathbf{u}$	$\mathbf{u}_{15}$	$\mathbf{u}_{30}$	$\mathbf{u}_{45}$	$\mathbf{u}_{60}$	$\mathbf{u}_{true}$
c_A	3.115 775	4.948 059	4.999 069	4.999 983	5
d_A	2.997 089	3.000 005	3	3	3
$K_B^{n_B}$	8.368 858	4.059 118	4.001 046	4.000 019	4
f_A	1.005 510	1.000 081	1.000 001	1	1
g_A	4.993 412	4.999 960	4.999 999	5	5
c_B	5.985 880	5.999 883	5.999 998	6	6
d_B	3.999 884	3.999 987	4	4	4
$K_A^{n_A}$	4.017 560	4.000 013	4	4	4
f_B	0.999 716	0.999 993	1.000 000	1	1
g_B	5.000 089	4.999 992	5	5	5

Table A.10: Numerical results for the synthetic biology toggle switch model using the paramters: $\mathbf{u}_0 = (2, 1, 1, 2, 2, 2, 1, 1, 2, 2)$, $\mathbf{u} = (5, 3, 4, 1, 5, 6, 4, 4, 1, 5)$, $\mathbf{x}_0 = (20, 20, 20, 20)$, $\alpha = 0.0005$, time interval $[0, 6]$, with step-size 0.005 and tolerance $|\Delta\mathbf{u}| < .00001$.

time	p_1	p_{173}	p_{347}	p_{521}	p
0	20	20	20	20	20
1.2	4.325 154	4.600 220	4.600 217	4.600 216	4.600 216
2.4	1.819 917	1.791 906	1.791 908	1.791 909	1.791 909
3.6	1.244 393	1.199 288	1.199 280	1.199 279	1.199 278
4.8	1.103 041	0.934 920	0.934 924	0.934 925	0.934 925
6	1.050 674	0.758 221	0.758 229	0.758 231	0.758 231

time	p_1	p_{173}	p_{347}	p_{521}	p
0	20	20	20	20	20
1.2	6.626 534	6.206 943	6.206 942	6.206 941	6.206 941
2.4	3.307 127	2.985 901	2.985 900	2.985 900	2.985 900
3.6	2.727 441	3.064 447	3.064 447	3.064 446	3.064 446
4.8	2.700 541	3.521 459	3.521 463	3.521 464	3.521 464
6	2.733 321	3.897 093	3.897 093	3.897 093	3.897 093

	$\mathbf{u}_1$	$\mathbf{u}_{173}$	$\mathbf{u}_{347}$	$\mathbf{u}_{521}$	$\mathbf{u}_{true}$
c_A	3.669 616	5.016 703	5.002 855	5.000 490	5
d_A	2.561 619	2.999 971	2.999 995	2.999 999	3
$K_B^{n_B}$	3.254 075	3.971 374	3.995 080	3.999 155	4
n_B	1.439 900	1.997 797	1.999 622	1.999 935	2
f_A	0.996 317	1.000 019	1.000 003	1.000 001	1
g_A	4.500 948	4.999 958	4.999 993	4.999 999	5
c_B	6.995 826	5.999 953	5.999 992	5.999 999	6
d_B	4.027 254	3.999 991	3.999 998	4	4
$K_A^{n_A}$	1.240 480	4.000 137	4.000 023	4.000 004	4
n_B	2.035 982	2.000 014	2.000 002	2	2
f_B	0.909 997	1.000 000	1	1	1
g_B	5.050 438	4.999 990	4.999 998	5	5

Table A.11: Numerical results for the synthetic biology toggle switch model using the paramters: $\mathbf{u}_0 = (4, 2, 3, 1, 1, 4, 5, 3, 3, 1, 1, 4)$, $\mathbf{u} = (5, 3, 4, 2, 1, 5, 6, 4, 4, 2, 1, 5)$, $\mathbf{x}_0 = (20, 20, 20, 20)$, $\alpha = 0.00025$, time interval $[0, 6]$, with step-size 0.005 and tolerance $|\Delta\mathbf{u}| < .00001$.

time	x_5	x_{10}	x_{15}	x
0	20	20	20	20
1.2	4.082 539	4.592 630	4.600 216	4.600 216
2.4	1.038 880	1.797 838	1.791 909	1.791 909
3.6	0.345 113	1.181 174	1.199 278	1.199 278
4.8	0.164 807	0.926 630	0.934 925	0.934 925
6	0.105 735	0.769 626	0.758 231	0.758 231
time	x_5	x_{10}	x_{15}	x
0	20	20	20	20
1.2	7.013 937	6.209 101	6.206 941	6.206 941
2.4	2.957 726	2.980 605	2.985 900	2.985 900
3.6	1.978 922	3.063 640	3.064 447	3.064 446
4.8	1.897 203	3.529 736	3.521 464	3.521 464
6	1.953 988	3.890 444	3.897 093	3.897 093
	$\mathbf{u}_5$	$\mathbf{u}_{10}$	$\mathbf{u}_{15}$	$\mathbf{u}_{true}$
c_A	2.106 303	3.480 371	5	5
d_A	0.284 665	2.979 783	3	3
K_B	1.323 878	2.636 124	2	2
f_A	0.911 122	1.003 604	1	1
g_A	1.695 411	4.973 076	5	5
c_B	2.920 784	5.988 788	6	6
d_B	1.245 924	3.993 330	4	4
K_A	1.754 553	2.001 866	2	2
f_B	0.816 337	0.999 766	1	1
g_B	2.190 457	4.993 264	5	5
α	19	0.010 000	1×10^{-17}	

Table A.12: Numerical results for the synthetic biology toggle switch model using the paramters: $\mathbf{u}_0 = (2, 1, 1, 1, 1, 2, 1, 1, 1, 2)$, $\mathbf{u} = (5, 3, 2, 1, 5, 6, 4, 2, 1, 5)$, $\mathbf{x}_0 = (20, 20, 20, 20)$, with α chosen as a piecewise constant function, on the interval $[0, 6]$, with step-size 0.005 and tolerance $|\Delta\mathbf{u}| < .00001$. Note 1×10^{-17} is effectively zero.

# of data point	$\|\|\mathbf{x} - \mathbf{x}_N\|\|_{L_2^n}$
6001	0.000 555 3
601	0.000 554 7
61	0.000 558 0
31	0.000 621 1
13	0.029 457 0
11	0.084 374 0

Table A.13: Table of the number of data points vs. $||\mathbf{x} - \mathbf{x}_N||_{L_2^n}$ for the toggle switch model. The parameters were $\mathbf{u} = [1, 2, 1, 1, 1, 2, 1, 1, 10, 10]$, $\mathbf{u}_0 = [.5, 1, .5, .5, .5, 1, .5, .5, 5, 5]$, $\mathbf{x}_0 = [5, 2, 5, 2]$, $t \in [0, 6]$ and respective tolerances were set at 1×10^{-7}.

time	p_{A7}	p_{A11}	p_{A15}	p_A
0	1	1	1	1
4	15.525 460	6.773 822	5.814 707	5.814 707
8	−6.226 472	6.395 880	5.890 989	5.890 989
12	−5.243 080	6.102 894	7.562 618	7.562 618
16	334.536 400	6.945 404	8.355 485	8.355 485
20	76.317 700	9.087 267	6.871 589	6.871 589
time	p_{B7}	p_{B11}	p_{B15}	p_B
0	1	1	1	1
4	14.469 640	9.401 776	8.544 433	8.544 433
8	21.571 780	8.129 293	7.627 376	7.627 376
12	82.520 820	7.908 639	6.413 707	6.413 707
16	−43.801 770	7.594 950	4.888 034	4.888 034
20	−473.355 200	6.709 543	4.189 657	4.189 657
time	p_{C7}	p_{C11}	p_{C15}	p_C
0	1	1	1	1
4	19.990 370	8.002 936	6.937 266	6.937 266
8	24.806 300	6.555 242	5.102 954	5.102 954
12	−28.630 660	6.133 007	4.631 667	4.631 667
16	−129.888 100	5.143 035	5.604 858	5.604 858
20	969.244 800	2.391 992	8.072 784	8.072 784
u	$\mathbf{u}_7$	$\mathbf{u}_{11}$	$\mathbf{u}_{15}$	**u**
α	272.435 300	217.551 500	216	216
β	−7.363 514	0.407 159	0.216 000	0.216 000
n	1.315 480	1.976 396	2	2
γ	0.198 156	0.209 163	0.200 000	0.200 000

Table A.14: Numerical results for the synthetic biology repressilator model using the parameters: $\mathbf{u}_0 = (300, .5, 1, .5)$, $\mathbf{u} = (216, 0.216, 2, 0.2)$, $\mathbf{x}_0 = (5, 1, 10, 1, 15, 1)$, time interval $[0, 20]$, with step-size 0.01 and tolerance $|\Delta\mathbf{u}| < .00001$.

time	*Actual*	*Noisy*	p_{A1}	p_{A4}	p_{A8}
0	1	1.123 343	1.123 343	1.123 343	1.123 343
2	5.616 314	5.406 332	8.615 364	5.673 826	5.665 617
4	5.814 707	5.632 575	9.541 272	5.934 601	5.830 233
6	5.590 611	5.556 056	10.426 970	5.717 984	5.582 472
8	5.890 989	6.005 701	9.548 942	5.980 661	5.870 152
10	6.652 593	6.686 275	2.886 326	6.683 259	6.642 046
time	*Actual*	*Noisy*	p_{B1}	p_{B4}	p_{B8}
0	1	1.133 155	1.133 155	1.133 155	1.133 155
2	8.009 791	7.799 504	12.894 330	7.862 575	7.906 440
4	8.544 433	8.362 017	13.447 940	8.509 079	8.430 596
6	8.028 295	7.993 548	11.192 920	8.070 208	7.953 468
8	7.627 376	7.741 861	8.846 023	7.719 969	7.599 999
10	7.128 320	7.161 175	12.319 810	7.276 924	7.130 436
time	*Actual*	*Noisy*	p_{C1}	p_{C4}	p_{C8}
0	1	1.126 293	1.126 293	1.126 293	1.126 293
2	7.274 865	7.064 747	11.754 590	7.269 409	7.284 520
4	6.937 266	6.755 123	10.465 870	7.049 872	6.955 619
6	5.848 559	5.814 005	9.480 804	6.019 408	5.876 539
8	5.102 954	5.217 706	11.092 080	5.288 247	5.127 194
10	4.709 068	4.742 966	12.830 930	4.882 396	4.717 291
$\mathbf{u}$	$\mathbf{u}_0$		$\mathbf{u}_1$	$\mathbf{u}_4$	$\mathbf{u}_8$
α	216		251.751 700	247.660 100	248.309 700
β	0.216 000		14.093 120	0.238 159	0.216 057
n	2		2.183 837	2.079 370	2.078 818
γ	0.200 000		0.114 682	0.196 338	0.196 483

Table A.15: Numerical results for the synthetic biology repressilator model using the parameters: $\mathbf{u}_0 = (150, .5, 1, .5)$, $\mathbf{u} = (216, 0.216, 2, 0.2,)$, $\mathbf{x}_0 = (5, 1, 10, 1, 15, 1)$, time interval $[0, 10]$, with step-size 0.01, with $\rho = 1$ $\alpha = 0.00001$, and tolerance $|\Delta\mathbf{u}| < .0001$.

# of data point	$\|\|\mathbf{x}-\mathbf{x}_N\|\|_{L_2^n}$
2001	0.018 992 2
1001	0.018 992 1
501	0.018 983 9
201	0.018 965 8
101	0.018 979 8
51	0.025 178 1
41	0.039 534 2
34	0.061 705 7
26	0.107 374 7
21	0.205 325 7

Table A.16: Table of the number of data points vs. $||\mathbf{x}-\mathbf{x}_N||_{L_2^n}$ for the repressilator model. The parameters were $\mathbf{u}_0 = (175, .5, 1, .5, .5, 175, .5, 1, .5, .5, 175, .5, 1, .5, .5)$, $\mathbf{u}^* = (250, .4, 3, .4, .4, 250, .4, 3, .4, .4, 250, .4, 3, .4, .4)$, $\mathbf{u} = (216, .216, 2, .2, 0.2, 216, .216, 2, .2, .2, 216, .216, 2, .2, .2)$, $\mathbf{x}_0 = (5, 1, 10, 1, 15, 1)$, $\alpha = 0.00001$, time interval $[0, 10]$ and respective tolerances were set at 1×10^{-7}.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_2$	$\mathbf{u}_4$	$\mathbf{u}_7$	$\mathbf{u}_{true}$
p_1	1.250	1.246 385	0.952 770	1.001 145	1
p_2	1.250	1.001 447	1.000 117	1.000 012	1
p_3	2.500	2.012 506	2.000 115	1.999 906	2
p_4	1.250	0.987 740	1.000 119	0.999 990	1
p_5	2.500	2.004 162	1.999 713	1.999 961	2
p_6	1.250	1.250 303	0.952 722	1.001 144	1
p_7	1.250	1.285 206	0.969 968	1.000 598	1
p_8	1.250	0.991 183	1.000 320	1.000 013	1
p_9	2.500	2.009 607	1.999 472	2.000 098	2
p_{10}	1.250	1.011 872	0.999 948	1.000 001	1
p_{11}	2.500	1.921 661	1.997 797	1.999 965	2
p_{12}	1.250	1.280 106	0.970 008	1.000 600	1
p_{13}	1.250	1.247 893	1.021 281	1.000 462	1
p_{14}	1.250	1.039 075	1.000 543	1.000 066	1
p_{15}	2.500	2.023 756	2.000 256	2.000 424	2
p_{16}	1.250	0.956 822	0.999 215	0.999 925	1
p_{17}	2.500	2.032 207	2.000 047	2.000 053	2
p_{18}	1.250	1.307 969	1.021 354	1.000 546	1
p_{19}	0.125	0.083 562	0.099 729	0.100 006	0.1
p_{20}	1.250	0.937 926	1.000 100	1.000 171	1
p_{21}	0.125	0.080 522	0.099 725	0.099 996	0.1
p_{22}	0.125	0.052 330	0.095 843	0.099 995	0.1
p_{23}	1.250	0.797 974	0.996 369	0.999 968	1
p_{24}	0.125	0.051 189	0.096 116	0.099 998	0.1
p_{25}	0.125	0.096 901	0.099 411	0.099 998	0.1
p_{26}	1.250	0.958 029	1.001 007	1.000 050	1
p_{27}	0.125	0.098 487	0.099 319	0.099 994	0.1
p_{28}	1.250	0.949 329	1.000 958	0.999 982	1
p_{29}	1.250	1.170 614	1.000 228	0.999 773	1
p_{30}	1.250	1.445 018	0.991 610	0.999 712	1
p_{31}	1.250	1.003 148	1.000 466	1.000 045	1
p_{32}	1.250	1.134 347	1.003 180	0.999 761	1
p_{33}	1.250	1.119 954	0.995 898	0.999 374	1
p_{34}	1.250	1.075 493	1.000 044	0.999 913	1
p_{35}	1.250	1.139 160	1.002 067	0.999 828	1
p_{36}	1.250	1.021 508	0.999 322	0.999 867	1
α		3.312 629	0.066 380	$1.910\,267 \times 10^{-7}$	

Table A.17: The evolution of the parameters at select iterations, with 16 data sets, $t_0 = 0$, $t_1 = 120$, $\Delta t = 0.5$, i.e. 240 time points. Regularization parameter α was determined using (3.14) where $C = 0.009$ and $\gamma = 2$.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_3$	$\mathbf{u}_6$	$\mathbf{u}_{10}$	$\mathbf{u}_{true}$
p_1	1.250	1.234 031	1.049 707	1.000 204	1
p_2	1.250	0.979 737	0.999 533	1.000 222	1
p_3	2.500	2.023 799	2.001 545	1.999 513	2
p_4	1.250	0.912 416	0.999 166	0.999 987	1
p_5	2.500	1.663 102	1.997 803	1.999 993	2
p_6	1.250	1.232 310	1.049 798	1.000 203	1
p_7	1.250	1.246 832	1.080 485	1.000 488	1
p_8	1.250	0.980 262	1.002 019	0.999 969	1
p_9	2.500	2.015 112	1.996 192	2.000 094	2
p_{10}	1.250	0.999 491	0.999 706	0.999 999	1
p_{11}	2.500	1.983 552	2.001 728	1.999 892	2
p_{12}	1.250	1.246 227	1.080 867	1.000 479	1
p_{13}	1.250	1.237 325	1.196 263	1.001 719	1
p_{14}	1.250	0.986 985	0.997 799	1.000 055	1
p_{15}	2.500	2.011 167	2.008 516	1.999 883	2
p_{16}	1.250	0.982 420	1.006 480	0.999 873	1
p_{17}	2.500	2.016 849	1.985 511	2.000 196	2
p_{18}	1.250	1.256 953	1.189 653	1.001 841	1
p_{19}	0.125	0.091 758	0.098 676	0.099 998	0.1
p_{20}	1.250	1.007 986	0.999 589	1.000 062	1
p_{21}	0.125	0.090 385	0.098 689	0.099 994	0.1
p_{22}	0.125	0.091 384	0.099 158	0.100 001	0.1
p_{23}	1.250	0.983 998	1.003 812	1.000 029	1
p_{24}	0.125	0.092 232	0.098 924	0.100 000	0.1
p_{25}	0.125	0.101 217	0.098 939	0.099 998	0.1
p_{26}	1.250	1.017 513	1.002 232	0.999 961	1
p_{27}	0.125	0.100 320	0.098 792	0.100 001	0.1
p_{28}	1.250	0.991 738	0.997 033	0.999 994	1
p_{29}	1.250	1.177 392	1.031 585	1.000 105	1
p_{30}	1.250	1.385 904	1.060 467	1.000 251	1
p_{31}	1.250	1.052 826	1.006 240	1.000 036	1
p_{32}	1.250	1.157 744	1.020 309	1.000 006	1
p_{33}	1.250	1.154 407	1.036 050	0.999 941	1
p_{34}	1.250	1.096 913	1.016 500	1.000 070	1
p_{35}	1.250	1.151 202	1.022 090	1.000 101	1
p_{36}	1.250	1.014 709	1.010 373	1.000 084	1
α		0.334 784	0.039 635	$2.517\,840 \times 10^{-4}$	

Table A.18: The evolution of the parameters at select iterations, with 5 data sets, $t_0 = 0$, $t_1 = 120$, $\Delta t = 0.5$, i.e. 240 time points. Regularization parameter α was determined using (3.14) where $C = 0.25$ and $\gamma = 1$.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_2$	$\mathbf{u}_4$	$\mathbf{u}_7$	$\mathbf{u}_{true}$
p_1	1.250	1.255 125	0.927 122	1.000 857	1
p_2	1.250	0.999 863	1.000 026	1.000 007	1
p_3	2.500	1.971 788	1.999 917	1.999 912	2
p_4	1.250	0.990 144	0.999 896	0.999 995	1
p_5	2.500	1.958 539	1.999 858	1.999 969	2
p_6	1.250	1.252 678	0.927 182	1.000 854	1
p_7	1.250	1.295 469	0.921 503	1.000 501	1
p_8	1.250	0.998 113	1.000 023	1.000 011	1
p_9	2.500	1.987 854	2.000 079	2.000 101	2
p_{10}	1.250	1.013 728	0.999 979	1.000 002	1
p_{11}	2.500	1.997 037	1.999 888	1.999 972	2
p_{12}	1.250	1.294 275	0.921 415	1.000 504	1
p_{13}	1.250	1.260 932	0.930 609	1.000 512	1
p_{14}	1.250	1.013 309	0.999 998	1.000 062	1
p_{15}	2.500	1.946 763	1.999 888	2.000 396	2
p_{16}	1.250	0.985 547	1.000 000	0.999 930	1
p_{17}	2.500	1.973 271	2.000 031	2.000 043	2
p_{18}	1.250	1.276 165	0.930 141	1.000 589	1
p_{19}	0.125	0.084 312	0.099 709	0.100 006	0.1
p_{20}	1.250	0.981 822	1.000 110	1.000 137	1
p_{21}	0.125	0.077 621	0.099 642	0.099 999	0.1
p_{22}	0.125	0.086 556	0.099 973	0.099 995	0.1
p_{23}	1.250	0.956 214	0.999 842	0.999 961	1
p_{24}	0.125	0.085 489	0.099 979	0.099 998	0.1
p_{25}	0.125	0.085 249	0.099 988	0.099 998	0.1
p_{26}	1.250	0.984 254	0.999 697	1.000 072	1
p_{27}	0.125	0.082 208	0.099 987	0.099 993	0.1
p_{28}	1.250	0.907 849	0.999 259	0.999 988	1
p_{29}	1.250	1.105 865	0.999 881	0.999 783	1
p_{30}	1.250	1.343 387	1.001 044	0.999 683	1
p_{31}	1.250	1.007 168	0.999 722	1.000 042	1
p_{32}	1.250	1.094 810	0.999 454	0.999 775	1
p_{33}	1.250	0.986 461	0.999 900	0.999 385	1
p_{34}	1.250	0.991 172	0.999 655	0.999 901	1
p_{35}	1.250	1.047 460	0.999 342	0.999 823	1
p_{36}	1.250	0.974 966	1.000 300	0.999 875	1
α		2.011 031	0.009 319	$2.688\,093 \times 10^{-8}$	

Table A.19: The evolution of the parameters at select iterations, with 16 data sets, $t_0 = 0$, $t_1 = 120$, $\Delta t = 6.0$, i.e. 20 time points. Regularization parameter α was determined using (3.14) where $C = 0.005$ and $\gamma = 2$.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_3$	$\mathbf{u}_6$	$\mathbf{u}_{10}$	$\mathbf{u}_{true}$
p_1	1.250	1.246 880	0.999 212	1.001 162	1
p_2	1.250	0.929 354	0.999 789	1.000 210	1
p_3	2.500	2.283 276	2.000 626	1.999 535	2
p_4	1.250	1.178 553	0.999 607	0.999 982	1
p_5	2.500	2.544 557	1.998 669	1.999 983	2
p_6	1.250	1.256 205	0.999 315	1.001 161	1
p_7	1.250	1.244 915	1.086 246	1.002 152	1
p_8	1.250	0.951 579	1.000 077	0.999 961	1
p_9	2.500	2.080 143	2.000 190	2.000 105	2
p_{10}	1.250	0.985 077	0.999 654	0.999 993	1
p_{11}	2.500	2.016 335	1.999 381	1.999 881	2
p_{12}	1.250	1.259 644	1.086 447	1.002 146	1
p_{13}	1.250	1.213 525	1.056 664	1.000 136	1
p_{14}	1.250	1.003 145	0.999 433	0.999 997	1
p_{15}	2.500	2.292 804	2.003 081	2.000 044	2
p_{16}	1.250	0.937 601	1.001 599	1.000 011	1
p_{17}	2.500	2.103 381	1.995 184	1.999 980	2
p_{18}	1.250	1.293 584	1.057 255	1.000 128	1
p_{19}	0.125	0.084 872	0.099 164	0.099 981	0.1
p_{20}	1.250	1.146 473	0.999 841	1.000 054	1
p_{21}	0.125	0.075 808	0.099 175	0.099 978	0.1
p_{22}	0.125	0.071 769	0.099 028	0.099 986	0.1
p_{23}	1.250	1.006 124	1.002 401	1.000 057	1
p_{24}	0.125	0.069 550	0.098 876	0.099 982	0.1
p_{25}	0.125	0.101 170	0.099 493	0.100 016	0.1
p_{26}	1.250	1.024 531	1.000 536	0.999 917	1
p_{27}	0.125	0.101 246	0.099 453	0.100 022	0.1
p_{28}	1.250	1.016 807	0.998 093	0.999 949	1
p_{29}	1.250	1.264 646	1.016 599	1.000 392	1
p_{30}	1.250	1.345 099	0.979 134	1.001 097	1
p_{31}	1.250	1.114 168	1.002 479	1.000 099	1
p_{32}	1.250	1.244 859	1.008 953	1.000 172	1
p_{33}	1.250	1.115 354	1.009 549	1.000 316	1
p_{34}	1.250	1.214 632	1.005 088	1.000 225	1
p_{35}	1.250	1.317 899	1.007 395	1.000 271	1
p_{36}	1.250	0.974 291	1.004 802	1.000 178	1
α		2.958 286	0.011 684	$4.004\,639 \times 10^{-7}$	

Table A.20: The evolution of the parameters at select iterations, with 5 data sets, $t_0 = 0$, $t_1 = 120$, $\Delta t = 6.0$, i.e. 20 time points. Regularization parameter α was determined using (3.14) where $C = 0.25$ and $\gamma = 2$.

$\mathbf{u}$	1	2	3	4	5	$\mathbf{u}_{true}$
p_1	0.964	0.994	1.004	1.001	1.001	1
p_2	1.250	1.246	1.200	0.837	1.000	1
p_3	2.500	2.494	2.470	2.475	2.000	2
p_4	1.255	1.318	1.146	1.004	1.000	1
p_5	2.495	2.478	1.891	1.994	2.000	2
p_6	0.970	0.999	0.999	1.002	1.001	1
p_7	0.963	0.997	0.998	1.001	1.002	1
p_8	1.250	0.880	1.000	1.000	1.000	1
p_9	2.499	2.302	1.999	2.000	2.000	2
p_{10}	0.990	0.999	1.000	1.000	1.000	1
p_{11}	1.981	2.000	2.000	2.000	2.000	2
p_{12}	0.964	0.998	0.998	1.001	1.002	1
p_{13}	1.002	1.013	0.997	0.998	1.000	1
p_{14}	1.250	0.877	1.000	1.000	1.000	1
p_{15}	2.500	2.313	1.999	1.999	2.000	2
p_{16}	1.001	1.004	1.000	0.999	1.000	1
p_{17}	1.996	1.991	2.001	2.001	2.000	2
p_{18}	1.002	1.010	0.997	0.999	1.000	1
p_{19}	0.092	0.100	0.100	0.100	0.100	0.1
p_{20}	0.847	1.000	1.000	1.000	1.000	1
p_{21}	0.100	0.100	0.100	0.100	0.100	0.1
p_{22}	0.099	0.100	0.100	0.100	0.100	0.1
p_{23}	0.990	1.000	1.000	1.000	1.000	1
p_{24}	0.100	0.100	0.100	0.100	0.100	0.1
p_{25}	0.100	0.100	0.100	0.100	0.100	0.1
p_{26}	0.999	1.000	1.000	1.000	1.000	1
p_{27}	0.100	0.100	0.100	0.100	0.100	0.1
p_{28}	1.021	1.000	1.000	1.000	1.000	1
p_{29}	1.272	1.002	1.000	1.000	1.000	1
p_{30}	1.262	1.002	0.999	1.000	1.001	1
p_{31}	1.003	1.000	1.000	1.000	1.000	1
p_{32}	1.021	1.001	1.000	1.000	1.000	1
p_{33}	1.026	1.000	1.001	1.000	1.000	1
p_{34}	1.008	1.000	1.000	1.000	1.000	1
p_{35}	1.027	1.001	1.000	1.000	1.000	1
p_{36}	1.258	0.999	1.000	1.000	1.000	1

Table A.21: The evolution of the parameters against the number of data sets, from 1 to 5. Where $\mathbf{u}_0 = 1.25\mathbf{u}$, $t_0 = 0$, $t_1 = 120$, $\Delta t = 6.0$, giving us 20 time points. In each case α was determined using (3.14).

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_3$	$\mathbf{u}_6$	$\mathbf{u}_9$	$\mathbf{u}_{true}$
p_1	0.50	0.626 169	0.977 716	0.999 467	1
p_2	0.50	0.845 634	1.001 652	1.000 320	1
p_3	1	2.076 520	1.995 797	1.999 288	2
p_4	0.50	0.982 500	0.999 144	1.000 009	1
p_5	1	1.932 977	1.997 472	2.000 070	2
p_6	0.50	0.626 183	0.977 573	0.999 464	1
p_7	0.50	0.671 519	0.953 503	0.999 078	1
p_8	0.50	1.030 579	1.003 004	0.999 622	1
p_9	1	1.924 555	1.991 375	2.000 945	2
p_{10}	0.50	1.016 320	0.999 801	0.999 996	1
p_{11}	1	1.840 233	2.002 840	1.999 174	2
p_{12}	0.50	0.670 292	0.952 888	0.999 007	1
p_{13}	0.50	0.516 680	0.707 436	1.000 203	1
p_{14}	0.50	0.473 022	1.005 274	0.999 925	1
p_{15}	1	1.960 049	1.881 358	2.000 360	2
p_{16}	0.50	0.700 850	0.999 723	1.000 257	1
p_{17}	1	2.056 845	1.993 584	1.999 494	2
p_{18}	0.50	0.568 753	0.714 151	1.000 016	1
p_{19}	0.05	0.089 717	0.101 100	0.100 017	0.1
p_{20}	0.50	0.833 641	1.003 210	1.000 043	1
p_{21}	0.05	0.099 511	0.100 882	0.100 008	0.1
p_{22}	0.05	0.108 314	0.096 627	0.100 042	0.1
p_{23}	0.50	0.936 750	0.992 308	1.000 001	1
p_{24}	0.05	0.109 960	0.097 338	0.100 046	0.1
p_{25}	0.05	0.115 358	0.098 594	0.099 977	0.1
p_{26}	0.50	0.960 169	1.005 821	0.999 941	1
p_{27}	0.05	0.119 978	0.098 423	0.099 983	0.1
p_{28}	0.50	0.936 870	1.004 243	1.000 046	1
p_{29}	0.50	1.079 540	0.937 324	0.999 449	1
p_{30}	0.50	1.038 933	0.819 932	0.998 214	1
p_{31}	0.50	0.884 892	0.983 656	0.999 804	1
p_{32}	0.50	1.167 248	0.955 376	0.999 531	1
p_{33}	0.50	1.340 602	0.919 584	0.999 417	1
p_{34}	0.50	0.969 342	0.960 231	0.999 696	1
p_{35}	0.50	1.118 842	0.943 111	0.999 615	1
p_{36}	0.50	1.000 133	0.974 238	0.999 934	1
α		1.165 914	0.015 849	10^4	

Table A.22: The evolution of the parameters at select iterations, with 5 data sets, $\mathbf{u}_0 = 0.5\mathbf{u}$, $t_0 = 0$, $t_1 = 120$, $\Delta t = 0.5$, i.e. 240 time points. Regularization parameter α was chosen optimally.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_4$	$\mathbf{u}_8$	$\mathbf{u}_{13}$	$\mathbf{u}_{true}$
p_1	1.650	1.651 545	1.105 821	0.999 941	1
p_2	1.650	0.911 544	1.000 604	1.000 238	1
p_3	3.300	2.116 065	1.980 435	1.999 476	2
p_4	1.650	1.378 180	0.996 264	0.999 998	1
p_5	3.300	3.344 761	1.988 214	2.000 026	2
p_6	1.650	1.662 437	1.107 683	0.999 940	1
p_7	1.650	1.627 982	1.040 038	1.000 101	1
p_8	1.650	0.589 713	1.007 239	0.999 866	1
p_9	3.300	2.593 956	1.983 755	2.000 348	2
p_{10}	1.650	1.001 204	0.999 993	0.999 998	1
p_{11}	3.300	2.196 277	2.009 800	1.999 675	2
p_{12}	1.650	1.680 336	1.043 087	1.000 074	1
p_{13}	1.650	1.608 928	1.025 381	1.004 219	1
p_{14}	1.650	0.711 963	1.047 664	0.999 943	1
p_{15}	3.300	2.903 146	2.006 762	2.000 284	2
p_{16}	1.650	0.894 844	0.690 989	1.000 212	1
p_{17}	3.300	1.592 314	1.976 131	1.999 577	2
p_{18}	1.650	1.758 230	1.334 716	1.004 029	1
p_{19}	0.165	0.076 345	0.096 245	0.100 004	0.1
p_{20}	1.650	1.331 635	1.001 496	1.000 057	1
p_{21}	0.165	0.061 656	0.096 285	0.100 001	0.1
p_{22}	0.165	0.036 063	0.096 632	0.100 013	0.1
p_{23}	1.650	0.805 532	1.003 137	1.000 019	1
p_{24}	0.165	0.048 024	0.096 479	0.100 012	0.1
p_{25}	0.165	0.149 164	0.097 287	0.099 983	0.1
p_{26}	1.650	1.856 892	0.990 514	0.999 995	1
p_{27}	0.165	0.088 907	0.097 226	0.099 984	0.1
p_{28}	1.650	0.993 899	0.989 121	1.000 026	1
p_{29}	1.650	1.672 354	1.073 179	0.999 770	1
p_{30}	1.650	1.761 728	0.681 674	0.999 428	1
p_{31}	1.650	1.268 865	1.016 222	0.999 941	1
p_{32}	1.650	1.644 150	1.036 446	0.999 790	1
p_{33}	1.650	0.790 454	1.050 411	0.999 668	1
p_{34}	1.650	1.599 149	1.023 033	0.999 887	1
p_{35}	1.650	2.159 075	1.016 680	0.999 849	1
p_{36}	1.650	1.348 863	1.027 738	0.999 979	1
α		2.511 886	0.015 849	10^4	

Table A.23: The evolution of the parameters at select iterations, with 5 data sets, $\mathbf{u}_0 = 1.65\mathbf{u}$, $t_0 = 0$, $t_1 = 120$, $\Delta t = 0.5$, i.e. 240 time points. Regularization parameter α was chosen optimally.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_4$	$\mathbf{u}_8$	$\mathbf{u}_{14}$	$\mathbf{u}_{true}$
p_1	0.030	0.737 221	0.963 898	0.999 975	1
p_2	0.030	1.151 142	1.200 726	1.000 243	1
p_3	0.060	1.521 362	1.600 840	1.999 459	2
p_4	0.030	0.875 456	0.959 865	0.999 993	1
p_5	0.060	1.714 217	1.929 222	2.000 016	2
p_6	0.030	0.721 198	0.960 669	0.999 973	1
p_7	0.030	0.618 657	0.873 705	0.999 926	1
p_8	0.030	0.772 749	1.037 865	0.999 995	1
p_9	0.060	2.005 887	1.748 461	2.000 016	2
p_{10}	0.030	0.927 440	0.994 638	1.000 002	1
p_{11}	0.060	1.362 255	2.003 130	1.999 979	2
p_{12}	0.030	0.559 493	0.857 172	0.999 923	1
p_{13}	0.030	0.536 878	0.942 788	1.000 198	1
p_{14}	0.030	0.487 655	1.146 855	0.999 985	1
p_{15}	0.060	1.538 615	1.306 044	2.000 074	2
p_{16}	0.030	1.094 387	0.963 919	1.000 051	1
p_{17}	0.060	1.722 927	2.038 109	1.999 918	2
p_{18}	0.030	0.462 456	1.075 056	1.000 153	1
p_{19}	0.003	0.083 855	0.089 038	0.100 002	0.1
p_{20}	0.030	0.982 138	0.836 519	1.000 004	1
p_{21}	0.003	0.080 360	0.096 229	0.100 002	0.1
p_{22}	0.003	0.140 839	0.051 691	0.100 005	0.1
p_{23}	0.030	0.806 694	0.622 619	1.000 001	1
p_{24}	0.003	0.156 229	0.065 420	0.100 005	0.1
p_{25}	0.003	0.141 181	0.030 620	0.100 012	0.1
p_{26}	0.030	0.654 611	0.184 156	0.999 959	1
p_{27}	0.003	0.177 855	0.047 645	0.100 015	0.1
p_{28}	0.030	0.521 966	0.936 343	1.000 000	1
p_{29}	0.030	0.813 260	0.545 625	0.999 891	1
p_{30}	0.030	0.267 789	0.658 219	0.999 763	1
p_{31}	0.030	0.420 732	0.889 447	0.999 953	1
p_{32}	0.030	−0.207 562	0.796 603	0.999 869	1
p_{33}	0.030	0.654 627	0.733 923	0.999 843	1
p_{34}	0.030	0.475 655	1.047 390	0.999 898	1
p_{35}	0.030	0.064 479	1.204 619	0.999 846	1
p_{36}	0.030	0.894 143	1.134 247	0.999 963	1
α		0.100 000	0.012 589	10^{-6}	

Table A.24: The evolution of the parameters at select iterations, with 5 data sets, type 2 regularization, $\mathbf{u}_* = 0.5\mathbf{u}$, $\mathbf{u}_0 = 0.03\mathbf{u}$, $t_0 = 0$, $t_1 = 120$, $\Delta t = 0.5$, i.e. 240 time points. Regularization parameter α was chosen optimally.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_4$	$\mathbf{u}_7$	$\mathbf{u}_{true}$
p_1	1001	0.999 208	0.999 985	1
p_2	1001	0.900 593	0.999 814	1
p_3	2002	1.929 048	2.000 441	2
p_4	1001	0.985 492	1.000 008	1
p_5	2002	1.960 852	1.999 985	2
p_6	1001	1.000 290	0.999 988	1
p_7	1001	0.764 393	0.999 994	1
p_8	1001	0.989 047	0.999 998	1
p_9	2002	2.025 275	2.000 021	2
p_{10}	1001	0.997 283	1.000 003	1
p_{11}	2002	1.974 810	1.999 987	2
p_{12}	1001	0.775 715	0.999 993	1
p_{13}	1001	1.364 719	1.000 442	1
p_{14}	1001	0.970 241	0.999 993	1
p_{15}	2002	1.781 646	2.000 128	2
p_{16}	1001	1.105 672	0.999 983	1
p_{17}	2002	1.846 120	1.999 888	2
p_{18}	1001	1.209 495	1.000 452	1
p_{19}	100.1	0.099 303	0.100 000	0.1
p_{20}	1001	0.945 424	1.000 023	1
p_{21}	100.1	0.096 006	0.099 998	0.1
p_{22}	100.1	0.086 554	0.100 004	0.1
p_{23}	1001	0.950 801	1.000 004	1
p_{24}	100.1	0.089 322	0.100 004	0.1
p_{25}	100.1	0.086 878	0.100 010	0.1
p_{26}	1001	0.931 344	0.999 950	1
p_{27}	100.1	0.090 673	0.100 014	0.1
p_{28}	1001	0.994 722	0.999 990	1
p_{29}	1001	1.031 879	1.000 038	1
p_{30}	1001	0.962 557	1.000 087	1
p_{31}	1001	1.019 049	0.999 970	1
p_{32}	1001	0.955 442	1.000 012	1
p_{33}	1001	0.929 475	1.000 142	1
p_{34}	1001	0.920 078	0.999 982	1
p_{35}	1001	0.762 935	0.999 975	1
p_{36}	1001	1.012 456	1.000 007	1
α		$7.943\,282 \times 10^{-6}$	10^{-6}	

Table A.25: The evolution of the parameters at select iterations, with 5 data sets, type 2 regularization, $\mathbf{u}_* = 1.65\mathbf{u}$, $\mathbf{u}_0 = 1001\mathbf{u}$, $t_0 = 0$, $t_1 = 120$, $\Delta t = 0.5$, i.e. 240 time points. Regularization parameter α was chosen optimally.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_2$	$\mathbf{u}_4$	$\mathbf{u}_7$	$\mathbf{u}_{true}$
p_1	1.250	0.789 043	1.040 876	1.040 975	1
p_2	1.250	0.993 026	1.032 342	1.031 973	1
p_3	2.500	1.988 623	1.933 009	1.932 790	2
p_4	1.250	1.046 233	1.002 354	1.004 118	1
p_5	2.500	2.114 532	2.003 682	2.003 388	2
p_6	1.250	0.789 084	1.039 669	1.039 570	1
p_7	1.250	0.824 965	0.988 019	0.986 739	1
p_8	1.250	0.919 250	0.993 260	0.993 247	1
p_9	2.500	2.027 455	2.029 580	2.029 619	2
p_{10}	1.250	1.009 966	0.999 024	0.998 881	1
p_{11}	2.500	1.911 154	1.990 230	1.990 226	2
p_{12}	1.250	0.816 637	0.986 670	0.987 960	1
p_{13}	1.250	0.707 889	0.459 160	0.469 475	1
p_{14}	1.250	0.894 429	1.017 099	1.018 386	1
p_{15}	2.500	2.201 228	1.808 078	1.808 328	2
p_{16}	1.250	1.253 066	0.954 101	0.946 776	1
p_{17}	2.500	1.744 857	2.153 859	2.152 663	2
p_{18}	1.250	0.583 646	0.505 584	0.495 462	1
p_{19}	0.125	0.088 688	0.099 789	0.100 295	0.1
p_{20}	1.250	0.956 177	0.997 900	0.998 500	1
p_{21}	0.125	0.090 584	0.099 909	0.100 345	0.1
p_{22}	0.125	0.090 717	0.099 083	0.099 218	0.1
p_{23}	1.250	1.018 703	0.998 596	0.998 034	1
p_{24}	0.125	0.088 609	0.099 156	0.099 281	0.1
p_{25}	0.125	0.096 327	0.102 087	0.100 910	0.1
p_{26}	1.250	0.985 426	0.978 906	0.977 709	1
p_{27}	0.125	0.096 141	0.103 664	0.102 364	0.1
p_{28}	1.250	0.994 565	1.006 542	1.009 386	1
p_{29}	1.250	1.265 777	0.966 401	0.967 308	1
p_{30}	1.250	1.380 330	0.904 214	0.904 649	1
p_{31}	1.250	1.047 865	1.016 387	1.015 679	1
p_{32}	1.250	1.190 241	0.955 824	0.955 515	1
p_{33}	1.250	1.064 943	0.845 881	0.845 803	1
p_{34}	1.250	1.114 115	0.990 735	0.988 495	1
p_{35}	1.250	1.260 312	0.983 263	0.983 132	1
p_{36}	1.250	1.025 835	0.976 957	0.977 586	1
α		1.847 850	0.002 326	1×10^6	

Table A.26: The evolution of the parameters at select iterations, with 1% noise, 5 data sets, $t_0 = 0$, $t_1 = 120$, $\Delta t = 0.5$, i.e. 240 time points.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_2$	$\mathbf{u}_5$	$\mathbf{u}_{true}$
p_1	1.250	0.836 336	1.221 648	1
p_2	1.250	0.923 864	0.910 236	1
p_3	2.500	2.250 142	2.299 440	2
p_4	1.250	1.004 638	1.012 229	1
p_5	2.500	1.987 762	1.993 987	2
p_6	1.250	0.835 990	1.221 867	1
p_7	1.250	0.880 299	0.899 171	1
p_8	1.250	0.903 396	0.965 961	1
p_9	2.500	2.063 017	2.105 599	2
p_{10}	1.250	1.007 379	0.997 799	1
p_{11}	2.500	1.785 982	1.931 839	2
p_{12}	1.250	0.861 207	0.894 629	1
p_{13}	1.250	0.775 512	1.035 516	1
p_{14}	1.250	0.998 741	1.029 283	1
p_{15}	2.500	2.001 591	1.936 591	2
p_{16}	1.250	0.906 727	0.936 962	1
p_{17}	2.500	1.911 488	2.162 696	2
p_{18}	1.250	0.803 303	1.102 851	1
p_{19}	0.125	0.098 278	0.098 710	0.1
p_{20}	1.250	1.028 637	1.066 927	1
p_{21}	0.125	0.096 697	0.095 021	0.1
p_{22}	0.125	0.112 589	0.109 834	0.1
p_{23}	1.250	0.976 820	0.970 523	1
p_{24}	0.125	0.113 772	0.111 934	0.1
p_{25}	0.125	0.095 979	0.097 351	0.1
p_{26}	1.250	0.948 764	0.956 709	1
p_{27}	0.125	0.097 716	0.100 637	0.1
p_{28}	1.250	0.993 412	1.004 231	1
p_{29}	1.250	1.183 310	1.103 212	1
p_{30}	1.250	1.403 603	1.198 700	1
p_{31}	1.250	1.037 517	1.042 222	1
p_{32}	1.250	1.106 275	1.047 794	1
p_{33}	1.250	1.099 180	0.984 786	1
p_{34}	1.250	1.123 892	1.078 495	1
p_{35}	1.250	1.208 693	1.112 372	1
p_{36}	1.250	1.031 405	1.004 588	1
α		0.398 107	1×10^6	

Table A.27: The evolution of the parameters at select iterations, with 3% noise, 5 data sets, $t_0 = 0$, $t_1 = 120$, $\Delta t = 0.5$, i.e. 240 time points.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_2$	$\mathbf{u}_5$	$\mathbf{u}_{true}$
p_1	1.250	0.936 203	0.936 105	1
p_2	1.250	0.944 607	0.945 127	1
p_3	2.500	2.064 182	2.064 389	2
p_4	1.250	1.136 179	1.134 881	1
p_5	2.500	2.330 238	2.330 258	2
p_6	1.250	0.931 285	0.930 997	1
p_7	1.250	0.867 592	0.868 217	1
p_8	1.250	0.852 489	0.850 891	1
p_9	2.500	2.139 890	2.140 027	2
p_{10}	1.250	1.043 828	1.045 171	1
p_{11}	2.500	2.168 586	2.169 290	2
p_{12}	1.250	0.871 243	0.871 135	1
p_{13}	1.250	0.763 352	0.764 159	1
p_{14}	1.250	0.910 831	0.910 980	1
p_{15}	2.500	2.283 576	2.283 546	2
p_{16}	1.250	1.293 854	1.293 517	1
p_{17}	2.500	1.888 273	1.888 349	2
p_{18}	1.250	0.569 716	0.568 710	1
p_{19}	0.125	0.077 526	0.070 025	0.1
p_{20}	1.250	1.008 313	1.008 065	1
p_{21}	0.125	0.075 648	0.068 536	0.1
p_{22}	0.125	0.017 230	0.034 802	0.1
p_{23}	1.250	1.314 924	1.315 341	1
p_{24}	0.125	0.014 088	0.029 054	0.1
p_{25}	0.125	0.118 236	0.117 677	0.1
p_{26}	1.250	0.899 598	0.899 771	1
p_{27}	0.125	0.132 364	0.132 993	0.1
p_{28}	1.250	0.926 567	0.923 786	1
p_{29}	1.250	1.278 543	1.278 197	1
p_{30}	1.250	1.277 169	1.276 818	1
p_{31}	1.250	0.996 671	0.999 609	1
p_{32}	1.250	1.257 486	1.257 011	1
p_{33}	1.250	1.148 582	1.148 995	1
p_{34}	1.250	1.027 455	1.026 670	1
p_{35}	1.250	1.196 092	1.196 718	1
p_{36}	1.250	1.002 668	1.002 318	1
α		89.125 094	10^6	

Table A.28: The evolution of the parameters at select iterations, with 5% noise, 5 data sets, $t_0 = 0$, $t_1 = 120$, $\Delta t = 0.5$, i.e. 240 time points.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_2$	$\mathbf{u}_4$	$\mathbf{u}_6$	$\mathbf{u}_{true}$
p_1	1.250	1.123 565	1.129 082	1.129 074	1
p_2	1.250	0.912 056	0.893 121	0.893 122	1
p_3	2.500	2.361 377	2.370 120	2.370 120	2
p_4	1.250	1.153 810	1.154 428	1.154 421	1
p_5	2.500	2.481 353	2.485 822	2.485 821	2
p_6	1.250	1.133 771	1.138 468	1.138 471	1
p_7	1.250	0.917 569	0.912 968	0.912 965	1
p_8	1.250	0.853 647	0.883 403	0.883 393	1
p_9	2.500	2.150 992	2.110 381	2.110 382	2
p_{10}	1.250	0.986 181	0.991 176	0.991 186	1
p_{11}	2.500	2.106 475	2.063 963	2.063 967	2
p_{12}	1.250	0.928 245	0.918 050	0.918 055	1
p_{13}	1.250	0.867 920	0.856 573	0.856 566	1
p_{14}	1.250	0.954 252	0.951 146	0.951 136	1
p_{15}	2.500	2.379 816	2.369 630	2.369 630	2
p_{16}	1.250	1.146 560	1.138 749	1.138 753	1
p_{17}	2.500	1.841 376	1.801 377	1.801 380	2
p_{18}	1.250	0.771 342	0.768 998	0.769 006	1
p_{19}	0.125	0.089 383	0.086 190	0.086 016	0.1
p_{20}	1.250	1.082 972	1.063 479	1.063 483	1
p_{21}	0.125	0.084 046	0.083 046	0.083 049	0.1
p_{22}	0.125	0.036 810	0.048 034	0.048 071	0.1
p_{23}	1.250	1.159 678	1.123 859	1.123 861	1
p_{24}	0.125	0.034 240	0.045 081	0.045 092	0.1
p_{25}	0.125	0.061 910	0.073 884	0.073 832	0.1
p_{26}	1.250	1.123 065	1.093 392	1.093 397	1
p_{27}	0.125	0.057 153	0.068 235	0.068 307	0.1
p_{28}	1.250	0.956 116	0.957 381	0.957 364	1
p_{29}	1.250	1.305 799	1.312 115	1.312 108	1
p_{30}	1.250	1.316 101	1.325 849	1.325 848	1
p_{31}	1.250	1.067 063	1.049 698	1.049 709	1
p_{32}	1.250	1.229 932	1.232 791	1.232 790	1
p_{33}	1.250	1.000 748	0.974 776	0.974 777	1
p_{34}	1.250	1.123 945	1.093 784	1.093 785	1
p_{35}	1.250	1.371 616	1.378 169	1.378 171	1
p_{36}	1.250	0.935 573	0.912 489	0.912 484	1
α		10	39.810 717	10^4	

Table A.29: The evolution of the parameters at select iterations, with 1% noise, 5 data sets, $t_0 = 0$, $t_1 = 120$, $\Delta t = 6.0$, i.e. 20 time points.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_2$	$\mathbf{u}_5$	$\mathbf{u}_{true}$
p_1	1.250	1.005 973	1.000 968	1
p_2	1.250	0.887 295	0.938 806	1
p_3	2.500	2.226 314	2.232 817	2
p_4	1.250	1.169 731	1.153 280	1
p_5	2.500	2.450 097	2.455 687	2
p_6	1.250	1.002 284	1.005 881	1
p_7	1.250	0.920 877	0.915 860	1
p_8	1.250	0.817 403	0.860 835	1
p_9	2.500	2.243 152	2.262 637	2
p_{10}	1.250	1.048 164	0.998 359	1
p_{11}	2.500	2.307 717	2.324 448	2
p_{12}	1.250	0.933 040	0.937 631	1
p_{13}	1.250	0.917 997	0.908 126	1
p_{14}	1.250	0.905 250	0.982 276	1
p_{15}	2.500	2.390 156	2.381 588	2
p_{16}	1.250	1.148 383	1.140 192	1
p_{17}	2.500	1.933 012	1.900 712	2
p_{18}	1.250	0.802 913	0.800 103	1
p_{19}	0.125	0.074 700	0.087 037	0.1
p_{20}	1.250	1.144 662	1.121 854	1
p_{21}	0.125	0.066 571	0.080 444	0.1
p_{22}	0.125	0.035 924	0.034 487	0.1
p_{23}	1.250	1.361 505	1.355 227	1
p_{24}	0.125	0.029 140	0.028 465	0.1
p_{25}	0.125	0.169 802	0.189 212	0.1
p_{26}	1.250	0.933 386	0.919 063	1
p_{27}	0.125	0.183 288	0.206 007	0.1
p_{28}	1.250	1.008 911	1.001 976	1
p_{29}	1.250	1.333 563	1.316 589	1
p_{30}	1.250	1.277 131	1.283 331	1
p_{31}	1.250	1.154 324	1.130 494	1
p_{32}	1.250	1.330 916	1.337 431	1
p_{33}	1.250	1.109 166	1.104 105	1
p_{34}	1.250	1.146 498	1.121 970	1
p_{35}	1.250	1.218 652	1.246 917	1
p_{36}	1.250	1.095 104	1.070 377	1
α		232.630 507	10^4	

Table A.30: The evolution of the parameters at select iterations, with 3% noise, 5 data sets, $t_0 = 0$, $t_1 = 120$, $\Delta t = 6.0$, i.e. 20 time points.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_3$	$\mathbf{u}_6$	$\mathbf{u}_9$	$\mathbf{u}_{true}$
p_1	1.250	1.232 740	1.233 644	1.233 644	1
p_2	1.250	0.960 106	0.960 484	0.960 502	1
p_3	2.500	2.291 210	2.291 328	2.291 334	2
p_4	1.250	1.168 495	1.168 060	1.168 026	1
p_5	2.500	2.545 187	2.545 566	2.545 573	2
p_6	1.250	1.242 534	1.242 307	1.242 306	1
p_7	1.250	1.066 921	1.068 467	1.068 491	1
p_8	1.250	1.048 514	1.049 396	1.049 392	1
p_9	2.500	2.360 057	2.360 057	2.360 057	2
p_{10}	1.250	1.014 869	1.012 557	1.012 542	1
p_{11}	2.500	2.228 983	2.228 568	2.228 572	2
p_{12}	1.250	1.090 283	1.088 637	1.088 615	1
p_{13}	1.250	1.087 528	1.088 659	1.088 660	1
p_{14}	1.250	1.020 463	1.021 787	1.021 775	1
p_{15}	2.500	2.448 987	2.448 964	2.448 963	2
p_{16}	1.250	1.060 048	1.059 286	1.059 304	1
p_{17}	2.500	2.200 585	2.200 557	2.200 569	2
p_{18}	1.250	0.966 110	0.964 862	0.964 861	1
p_{19}	0.125	0.064 520	0.078 592	0.078 436	0.1
p_{20}	1.250	1.222 884	1.222 517	1.222 512	1
p_{21}	0.125	0.054 670	0.067 783	0.067 638	0.1
p_{22}	0.125	0.058 904	0.057 035	0.057 059	0.1
p_{23}	1.250	1.183 592	1.183 388	1.183 389	1
p_{24}	0.125	0.052 988	0.051 569	0.051 589	0.1
p_{25}	0.125	0.052 577	0.052 560	0.052 659	0.1
p_{26}	1.250	1.327 037	1.327 063	1.327 071	1
p_{27}	0.125	0.049 758	0.048 977	0.049 069	0.1
p_{28}	1.250	1.041 488	1.041 407	1.041 367	1
p_{29}	1.250	1.456 771	1.457 208	1.457 195	1
p_{30}	1.250	1.216 048	1.216 017	1.216 013	1
p_{31}	1.250	1.022 266	1.022 291	1.022 293	1
p_{32}	1.250	1.205 809	1.205 586	1.205 589	1
p_{33}	1.250	1.226 593	1.226 659	1.226 658	1
p_{34}	1.250	1.296 873	1.297 056	1.297 074	1
p_{35}	1.250	1.366 907	1.366 817	1.366 824	1
p_{36}	1.250	1.280 307	1.280 612	1.280 601	1
α		92.611 873	10^4	10^4	

Table A.31: The evolution of the parameters at select iterations, with 5% noise, 5 data sets, $t_0 = 0$, $t_1 = 120$, $\Delta t = 6.0$, i.e. 20 time points.

$\mathbf{u}$	$\mathbf{u}_0$	$\mathbf{u}_3$	$\mathbf{u}_6$	$\mathbf{u}_{11}$	$\mathbf{u}_{true}$
p_1	1.250	1.230 250	1.225 924	1.000 683	1
p_2	1.250	0.991 905	1.057 139	0.998 554	1
p_3	2.500	2.276 342	1.989 146	1.999 998	2
p_4	1.250	1.097 817	0.941 104	1.001 338	1
p_5	2.500	2.350 048	1.988 150	2.000 040	2
p_6	1.250	1.293 943	1.254 393	0.999 529	1
p_7	1.250	1.256 998	1.310 132	0.965 145	1
p_8	1.250	0.998 805	1.028 752	1.033 829	1
p_9	2.500	1.980 252	2.019 452	2.000 300	2
p_{10}	1.250	0.936 127	0.967 154	0.967 096	1
p_{11}	2.500	2.069 476	1.986 336	1.999 605	2
p_{12}	1.250	1.247 677	1.195 329	0.993 356	1
p_{13}	1.250	1.147 725	1.137 171	1.001 718	1
p_{14}	1.250	1.155 333	1.082 797	0.999 560	1
p_{15}	2.500	2.473 027	2.092 534	1.998 726	2
p_{16}	1.250	0.934 641	0.922 533	1.000 001	1
p_{17}	2.500	2.328 476	2.109 298	2.000 955	2
p_{18}	1.250	1.336 110	1.266 135	1.001 417	1
p_{19}	0.125	0.092 350	0.108 095	0.099 688	0.1
p_{20}	1.250	1.278 255	1.228 581	0.994 818	1
p_{21}	0.125	0.078 654	0.095 766	0.100 013	0.1
p_{22}	0.125	0.086 130	0.105 161	0.107 449	0.1
p_{23}	1.250	1.253 270	1.266 847	1.115 164	1
p_{24}	0.125	0.076 544	0.097 403	0.100 055	0.1
p_{25}	0.125	0.070 298	0.091 737	0.099 872	0.1
p_{26}	1.250	1.280 153	0.931 293	0.998 805	1
p_{27}	0.125	0.060 804	0.094 306	0.099 967	0.1
p_{28}	1.250	1.086 576	0.999 404	0.999 888	1
p_{29}	1.250	1.389 865	1.122 434	0.999 518	1
p_{30}	1.250	1.202 391	1.386 692	0.998 857	1
p_{31}	1.250	1.144 167	1.112 450	0.999 479	1
p_{32}	1.250	1.241 765	1.181 591	0.997 380	1
p_{33}	1.250	1.199 009	1.192 942	0.994 904	1
p_{34}	1.250	1.213 849	1.029 189	1.001 242	1
p_{35}	1.250	1.398 013	1.062 881	1.002 326	1
p_{36}	1.250	0.881 366	1.023 785	1.001 360	1
α		1.372 353	0.469 269	$3.118\,463 \times 10^{-6}$	

Table A.32: The evolution of the parameters at select iterations, with 5 data sets, 4 out of 8 states, $t_0 = 0$, $t_1 = 120$, $\Delta t = 6.0$, i.e. 20 time points. Regularization parameter α was determined using (3.14) where $C = 0.5$ and $\gamma = 2$.