

A Homotopy Method for Parameter Estimation of Nonlinear Differential Equations with Multiple Optima

Wenrui Hao¹

Received: 23 March 2017 / Revised: 19 June 2017 / Accepted: 25 July 2017 /
Published online: 19 August 2017
© Springer Science+Business Media, LLC 2017

Abstract A numerical method for estimating multiple parameter values of nonlinear systems arising from biology is presented. The uncertain parameters are modeled as random variables. Then the solutions are expressed as convergent series of orthogonal polynomial expansions in terms of the input random parameters. Homotopy continuation method is employed to solve the resulting polynomial system, and more importantly, to compute the multiple optimal parameter values. Several numerical examples, from a single equation to problems with relatively complicated forms of governing equations, are used to demonstrate the robustness and effectiveness of this numerical method.

Keywords Biological systems · Parameter estimation · Homotopy continuation method

1 Introduction

The biological systems are often described by nonlinear ordinary differential equations (ODEs) or partial differential equations (PDEs) with uncertain and unknown parameters [5, 9, 10, 15, 18, 29–31, 35]. The estimations of these unknown parameters are required for accurate descriptions of the biological processes. Biologically speaking, how to treat the uncertainty involved in these unknown parameters reflects measurement error, non-stringent experimental design, the flexibility of the processes themselves, etc. Mathematically speaking, the parameter estimation also involves some advanced computational and statistical methods in optimization problems. Different methodologies have been developed to estimate parameters such as the Bayesian framework [3, 19, 20], the combination of the polynomial chaos theory and the Extended Kalman Filter theory [2], and the inverse problem theory [28].

The polynomial chaos approach has been shown to be an efficient method for quantifying the effects of such uncertainties on nonlinear systems of differential equations [22, 32–34].

✉ Wenrui Hao
wxh64@psu.edu

¹ Department of Mathematics, The Penn State University, University Park, PA 16802, USA

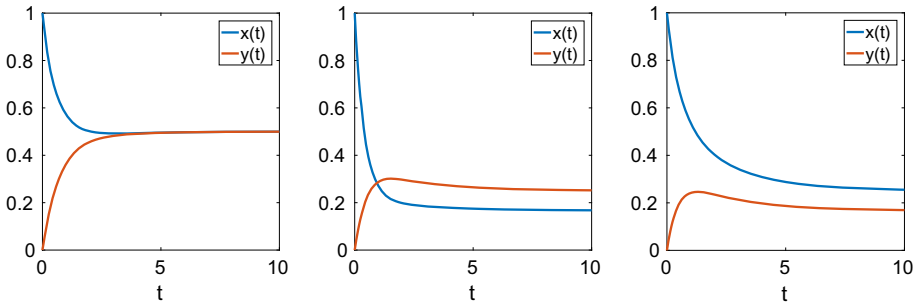


Fig. 1 Solutions for system (1) with $K_X = K_Y = \frac{1}{2}$, $X(0) = 1$ and $Y(0) = 0$: *Left* X and Y coexist with $d_X = d_Y = 1$; *Middle* Y dominates with $d_X = 2$ and $d_Y = 1$; *Right*: X dominates with $d_X = 1$ and $d_Y = 2$;

Thus it is a valuable approach by applying polynomial chaos approach to problems with expensive computation such as parameter estimation [24]. However, the polynomial chaos approach assumes that the solution is a smooth function of the parameter set [8, 34], and is hard to generalize to the non-smooth case of the parameter space such as, i.e., bifurcation, multiple solutions in piecewise-smooth parameter space [4, 6, 36]. Indeed, multiple solutions and bifurcations are highly related to the parameter estimation. Thus it is very important to know the relationship between solution structures and non-smooth parameter space.

Here is a simple example of multiple solutions of parameter values for a given dataset. Let’s consider a model with two cells interaction. There are two cell type: X and Y . They promote each other and degrade by themselves. The following ODE system is used to model the growth of X and Y :

$$\begin{cases} \frac{dX}{dt} = \underbrace{\frac{Y}{Y + K_Y}}_{\text{production}} \underbrace{-d_X X}_{\text{degradation}} \\ \frac{dY}{dt} = \underbrace{\frac{X}{X + K_X}}_{\text{production}} \underbrace{-d_Y Y}_{\text{degradation}} \end{cases}, \tag{1}$$

where K_X, K_Y, d_X and d_Y are capacities and degradation rates for X and Y respectively. On the one hand, the solutions have different behaviors by choosing different parameter values (see Fig. 1): X and Y could co-exist; either X or Y could dominate by choosing different parameter values. On the other hand, a same solution behavior may have different parameter values. For example, we consider the following “biological data”: X and Y coexist eventually i.e., $X_f = Y_f$ (X_f and Y_f are the equilibriums), as well as the decreasing rates of X and Y are $\frac{1}{6}$ and $\frac{1}{7}$ respectively when X and Y are half of the equilibriums, i.e., when $X = Y = \frac{X_f}{2} = \frac{Y_f}{2}$, $\frac{dX}{dt} = -\frac{1}{6}$ and $\frac{dY}{dt} = -\frac{1}{7}$. After normalization, i.e., $X_f = Y_f = 1$, we have the following nonlinear system with the parameters as the variables

$$\begin{cases} \frac{1}{1+K_Y} - d_X = 0 \\ \frac{1}{1+K_X} - d_Y = 0 \\ \frac{1}{1+2K_X} - \frac{d_Y}{2} = -\frac{1}{6} \\ \frac{1}{1+2K_Y} - \frac{d_X}{2} = -\frac{1}{7} \end{cases}, \tag{2}$$

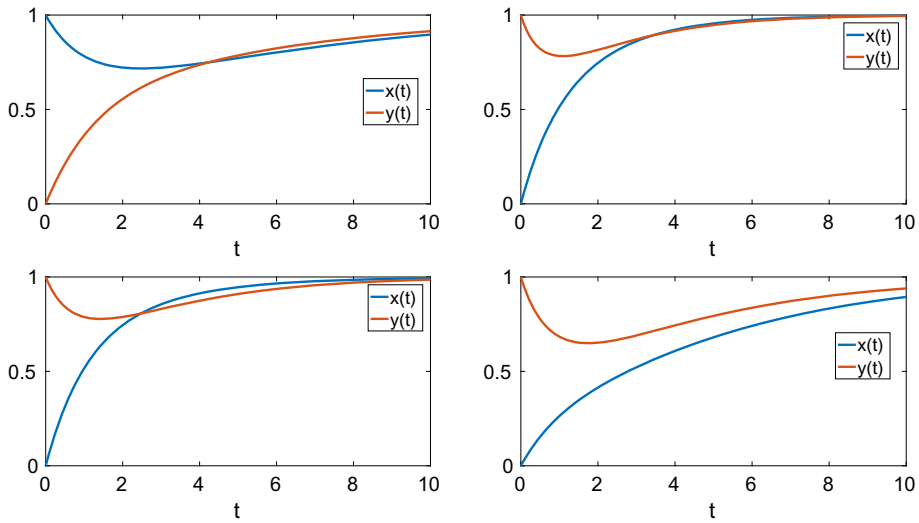


Fig. 2 Solutions for system (1) with the parameter value pairs shown in (3)

which implies that we have four solutions:

$$\left\{ \begin{array}{llll} K_X = 1 & K_X = \frac{1}{2} & K_X = 1 & K_X = \frac{1}{2} \\ K_Y = 1 + \frac{\sqrt{2}}{2} & K_Y = 1 - \frac{\sqrt{2}}{2} & K_Y = 1 - \frac{\sqrt{2}}{2} & K_Y = 1 + \frac{\sqrt{2}}{2} \\ d_X = \frac{2}{4+\sqrt{2}} & d_X = \frac{2}{4-\sqrt{2}} & d_X = \frac{2}{4-\sqrt{2}} & d_X = \frac{2}{4+\sqrt{2}} \\ d_Y = \frac{1}{2} & d_Y = \frac{2}{3} & d_Y = \frac{1}{2} & d_Y = \frac{2}{3} \end{array} \right. \quad (3)$$

The evolutions of $X(t)$ and $Y(t)$ with these four parameter values are shown in Fig. 2.

This simple example illustrates that computation of multiple parameter values is essential for parameter estimation with given datasets. In order to handle the multiple solutions of parameter values, we employ the techniques from Numerical Algebraic Geometry (NAG) [25], whose main topic is to solve the systems of polynomial equations. The key method in NAG is the homotopy continuation method: it can be shown that for a given system of polynomial equations to be solved, a homotopy between the system and a start system (which is easier to solve and shares many features with the former system) can be constructed (see [1, p. 24] and [26, p. 16] for detailed descriptions). Then tracking each solution path from the start system towards the original system along the homotopy finally obtains all the solutions of the original system. The homotopy continuation method has been applied to compute the solutions of nonlinear systems arising from discretized nonlinear PDEs [13, 14], physical phenomenology [16, 17], biology [11, 12], and etc.

In this paper, we will use the polynomial chaos expansion (PCE) method to take into account variations of model parameters and generate a cluster of sampling points which converge to different optimal values of parameters. For each group of sampling points, we apply the homotopy continuation method to compute all the solutions of polynomial systems generated by the PCE method. Several numerical examples are used to demonstrate the feasibility and efficiency of this new method. The outline of this paper is as follows: In Sect. 2, we introduce the problem and numerical methods; In Sect. 3, we provide the convergence error estimation; In Sect. 4, we use several examples to test our numerical methods.

2 A Numerical Method for Estimating Parameter with Multiple Optima

In general, the discretized nonlinear system of differential equations for each time step can be written as

$$\mathbf{F}(\mathbf{u}, \mathbf{p}) = \mathbf{0},$$

where \mathbf{u} is a vector representing numerical solutions, and \mathbf{p} is a also vector containing all the parameters. The dimension of \mathbf{u} depends on the number of grid points of discretization and the dimension of \mathbf{p} depends on the number of parameters in biological models. The observations of \mathbf{u} could be from clinical or biological data. Considering N pairs of observations (for example, different time measurement), $\mathbf{v}_1, \dots, \mathbf{v}_N$, typically, the points \mathbf{v}_i will not exactly equal to \mathbf{u}_i due to measurement and numerical errors. Therefore, we need to find the value of the parameter vector \mathbf{p} that makes the function \mathbf{u} provide the best fit to the data points, namely,

$$\min_{\mathbf{p}} G(\mathbf{p}) \doteq \sum_i \|\mathbf{u}_i - \mathbf{v}_i\|. \tag{4}$$

2.1 Polynomial Chaos Expansion

The unknown parameters, \mathbf{p} , are viewed as random variables which are independently identically distributed (IID). Therefore \mathbf{u} is a function of the random variables \mathbf{p} , i.e., $\mathbf{u} = \mathbf{u}(\mathbf{p})$. Following the generalized polynomial chaos (gPC) method, the numerical solution $\mathbf{u}(\mathbf{p})$ is expanded in terms of polynomial basis functions

$$\mathbf{u}(\mathbf{p}) \approx \sum_{|\alpha|=0}^S \mathbf{u}_\alpha \phi_\alpha(\mathbf{p}),$$

where \mathbf{u}_α is a vector. By denoting the dimensionality of \mathbf{p} as n_p , $\alpha = [\alpha_1, \dots, \alpha_{n_p}]$ is an n_p -dimensional multi-index, and $|\alpha|$ is the sum of the vector elements, i.e., $|\alpha| := \alpha_1 + \dots + \alpha_{n_p}$. Each element α_i of α can take on a non-negative integer value between 0 and S . The expansion of \mathbf{p} is according to the known parameter prior distribution, and hence \mathbf{p}_i is known for all $1 \leq i \leq n_p$. For small size systems, parameter estimation can be easily done by using the optimization and the generalized polynomial chaos framework. But for systems involving large-scale computation, it is very expensive to choose enough sampling points to get a good approximation of $\mathbf{u}(\mathbf{p})$. Moreover, it is very unclear of relationship between parameter space and solution structure for the large-scale systems. In another word, it may have multiple optimal parameter values to fit the given real data. In this case, we need to design an automatical way to compute the useful sampling points and design a new numerical method to compute the multiple optimal parameter values.

Suppose that we already have n sample points, i.e., $\mathbf{p}_1, \dots, \mathbf{p}_n$, $\mathbf{u}(\mathbf{p})$ can be expanded on the parameter space by using polynomial chaos expansion $\mathbf{u}(\mathbf{p}) = \sum_{j=1}^n \mathbf{u}_{\alpha_j} \phi_{\alpha_j}(\mathbf{p})$. Polynomial chaos theory then solves for the coefficients \mathbf{u}_α of the polynomial chaos state expansion using the collocation approach: For each sampling point \mathbf{p}_j , the numerical solution can be solved by solving $\mathbf{F}(\mathbf{u}, \mathbf{p}_j) = \mathbf{0}$ and obtain $\mathbf{u} = \mathbf{u}(\mathbf{p}_j)$. The \mathbf{u}_α can be solved by the following linear system

$$\begin{pmatrix} \phi_{\alpha_1}(\mathbf{p}_1) & \cdots & \phi_{\alpha_n}(\mathbf{p}_1) \\ \vdots & & \vdots \\ \phi_{\alpha_1}(\mathbf{p}_n) & \cdots & \phi_{\alpha_n}(\mathbf{p}_n) \end{pmatrix} \begin{pmatrix} \mathbf{u}_{\alpha_1} \\ \vdots \\ \mathbf{u}_{\alpha_n} \end{pmatrix} = \begin{pmatrix} \mathbf{u}(\mathbf{p}_1) \\ \vdots \\ \mathbf{u}(\mathbf{p}_n) \end{pmatrix} \tag{5}$$

Then the optimization problem (4) can be rewritten as

$$\min_{\mathbf{p}} G(\mathbf{p}) \doteq \left(\sum_{j=1}^n \mathbf{u}_{\alpha_j} \phi_{\alpha_j}(\mathbf{p}) - \mathbf{v} \right)^T \left(\sum_{j=1}^n \mathbf{u}_{\alpha_j} \phi_{\alpha_j}(\mathbf{p}) - \mathbf{v} \right), \tag{6}$$

which is equivalent to

$$f_s := \left(\sum_{j=1}^n \mathbf{u}_{\alpha_j} \phi_{\alpha_j}(\mathbf{p}) - \mathbf{v} \right)^T \sum_{j=1}^n \mathbf{u}_{\alpha_j} \partial_{\mathbf{p}_s} \phi_{\alpha_j}(\mathbf{p}) = 0. \tag{7}$$

Here $\partial_{\mathbf{p}_s}$ is the partial derivative with respect to \mathbf{p}_s ($1 \leq s \leq n_{\mathbf{p}}$). Since $\phi_{\alpha_j}(\mathbf{p})$ is a polynomial basis on parameter space, the nonlinear equation f_s , $1 \leq s \leq n_{\mathbf{p}}$ forms a polynomial system.

2.2 Homotopy Continuation Method

In order to solve the polynomial system, we employed the homotopy continuation method to solve the polynomial system:

$$H(\mathbf{p}, t) = (1 - t) \begin{pmatrix} f_1 \\ \vdots \\ f_{n_{\mathbf{p}}} \end{pmatrix} + \gamma t \begin{pmatrix} \mathbf{p}_1^d - 1 \\ \vdots \\ \mathbf{p}_{n_{\mathbf{p}}}^d - 1 \end{pmatrix}, \tag{8}$$

where γ is a random complex number and d is highest degree of $\phi_{\alpha_j}(\mathbf{p})$ for $1 \leq j \leq n$. Moreover, $t \in [0, 1]$ is a homotopy parameter: When $t = 1$, we have known solutions to $H(\mathbf{p}, 1) = 0$. Specifically, the solutions of $\mathbf{p}_s^d - 1 = 0$ are $\mathbf{p}_s = e^{2\pi i s/d}$, $s = 0, \dots, d - 1$, which forms the solutions of $H(\mathbf{p}, 1) = 0$. The known solutions are called start points, and the system $H(\mathbf{p}, 1) = 0$ is called the start system. Such a start system with the degree d equal to the degree of f_s for all s is called a total degree start system. Choosing a total degree start system and a random complex number γ guarantees finding all the solutions. The use of the random γ , which is called the γ -trick, was introduced in [1, p. 26,32] and [26, p. 18]. Therefore, γ is randomly chosen once for solving the polynomial systems (8). A good discussion of a more general version of this trick is given in [1, 26].

The parameter estimation algorithm for the nonlinear system $\mathbf{F}(\mathbf{u}, \mathbf{p}) = \mathbf{0}$ consists of the following steps:

1. Initialize sampling points on the parameter space $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_n$;
2. Solve the nonlinear system $\mathbf{F}(\mathbf{u}, \mathbf{p}) = \mathbf{0}$ and get $\mathbf{u}(\mathbf{p}_i)$ for $1 \leq i \leq n$;
3. Compute \mathbf{u}_{α_j} by solving the linear system (5);
4. Compute the next sampling point \mathbf{p}_{n+1} by using homotopy method to solve (8);
5. Cluster sampling points, update $n = n + 1$ and go to Step 1.

Remark • The initialization of this algorithm could be a very rough estimate of parameter space [23], i.e. normally we give an upper bound and a lower bound for each parameter value.

- The major computational efforts will likely be on solving the polynomial system by using homotopy methods. In this case, we will get cluster of sampling points near the different optimal parameter values. The stopping criteria for this algorithm is $\|\mathbf{p}_{n+1} - \mathbf{p}_n\| < Tol$.
- Although the numerical method used for solving the nonlinear system $\mathbf{F}(\mathbf{u}, \mathbf{p}) = \mathbf{0}$ does not affect parameter estimation, it will definitely affect the efficiency of the whole process. In another word, an efficient nonlinear solver for $\mathbf{F}(\mathbf{u}, \mathbf{p}) = \mathbf{0}$ would speed up the parameter estimation process.

- The rationale of using gPC representation is that the chosen basis $\phi(\mathbf{p})$ is optimal (in L_2) with respect to the probability density distribution of \mathbf{p} . Accordingly, the choice of $\phi(\mathbf{p})$ depends on the probability density function of \mathbf{p} . In numerical simulations, we use the Legendre polynomial to construct the gPC representation since that we assume \mathbf{p} be uniform within bounded domain.

3 Proof of Convergence

Theorem 1 (Existence) *There exists a sequence of \mathbf{p}_n which is convergent to the exact solution of (6).*

Proof According to definition of $G(\mathbf{p})$, it is clear that

$$G(\mathbf{p}_0) \geq G(\mathbf{p}_1) \geq G(\mathbf{p}_2) \geq \dots \geq G(\mathbf{p}_n) \geq \dots$$

Then there exists a subsequence of $G(\mathbf{p}_n)$ which is strictly decreasing. Moreover $\mathbf{p} \in P$ which is a bounded closed set. Then there exists a limit of the sequence \mathbf{p}_n . \square

However (6) may have multiple optima, namely, one subsequence could converge to one optimum while another subsequence may converge to another optimum. Therefore, we consider one converging sequence and denote \mathbf{p}^* is one of multiple optima. Let us assume that an accurate and stable numerical scheme is employed for discretized nonlinear system of differential equations with fixed parameter \mathbf{p}_n and the error of its numerical solution $u(\mathbf{p}_n)$ is

$$\varepsilon_n = \|\mathbf{u}(\mathbf{p}_n) - \mathbf{u}(\mathbf{p}^*)\|. \tag{9}$$

Theorem 2 (Error superposition). *Let ε_Δ be the error induced by solving the discretized nonlinear system of differential equations, and ε_Q be the aliasing error of approximating the polynomial chaos expansion via given polynomial basis as defined in (5), then $\varepsilon_n \leq C(\varepsilon_\Delta + \varepsilon_Q + \|\mathbf{p}_n - \mathbf{p}^*\|)$, where C depends on $\|\mathbf{u}_\alpha\|$ and $\|\nabla_{\mathbf{p}}\phi_\alpha\|$.*

Proof

$$\begin{aligned} \|\mathbf{u}(\mathbf{p}_n) - \mathbf{u}(\mathbf{p}^*)\| &= \|\mathbf{u}(\mathbf{p}_n) - \tilde{\mathbf{u}}(\mathbf{p}_n) - (\mathbf{u}(\mathbf{p}^*) - \tilde{\mathbf{u}}(\mathbf{p}^*)) + \tilde{\mathbf{u}}(\mathbf{p}_n) - \tilde{\mathbf{u}}(\mathbf{p}^*)\| \\ &\leq \|\mathbf{u}(\mathbf{p}_n) - \tilde{\mathbf{u}}(\mathbf{p}_n)\| + \|\mathbf{u}(\mathbf{p}^*) - \tilde{\mathbf{u}}(\mathbf{p}^*)\| + \|\tilde{\mathbf{u}}(\mathbf{p}_n) - \tilde{\mathbf{u}}(\mathbf{p}^*)\| \\ &\leq 2\varepsilon_\Delta + \|\tilde{\mathbf{u}}(\mathbf{p}_n) - \tilde{\mathbf{u}}(\mathbf{p}^*)\| \\ &\leq 2\varepsilon_\Delta + 2\varepsilon_Q + \left\| \sum_{|\alpha|=0}^n \mathbf{u}_\alpha(\phi_\alpha(\mathbf{p}_n) - \phi_\alpha(\mathbf{p}^*)) \right\| \\ &\leq 2\varepsilon_\Delta + 2\varepsilon_Q + \max_\alpha \|\mathbf{u}_\alpha\| \|\nabla_{\mathbf{p}}\phi_\alpha\| \|\mathbf{p}_n - \mathbf{p}^*\|. \end{aligned} \tag{10}$$

\square

In numerical computation, we can check $\|\mathbf{p}^n - \mathbf{p}^*\| \rightarrow 0$ which implies that $\|\mathbf{u}(\mathbf{p}_n) - \mathbf{u}(\mathbf{p}^*)\| \rightarrow 0$.

4 Numerical Results

In this section, we will present some numerical examples in order to demonstrate the performance and efficiency of the proposed method.

Example 4.1 The first example is a parametric ordinary differential equation defined as

$$\begin{cases} u_{xx} = pu & \text{on } [0, 1] \\ u(0) = 1 \text{ and } u(1) = \frac{e^{\sqrt{p}} + e^{-\sqrt{p}}}{2} \end{cases} \quad (11)$$

It is clear that the exact solution of (11) is $u(x) = \frac{e^{\sqrt{p}x} + e^{-\sqrt{p}x}}{2}$, where p is a parameter. The finite difference method is employed to discretize this ODE (11) as follows

$$F(\mathbf{U}, p) := \begin{cases} U_{i+1} + U_{i-1} - 2U_i = ph^2U_i & \text{for } i = 1, 2, \dots, N - 1 \\ U_0 = 1 & \text{and } U_N = \frac{e^{\sqrt{p}} + e^{-\sqrt{p}}}{2} \end{cases} = 0. \quad (12)$$

We choose the data points V_i by solving $F(\mathbf{V}, p) = 0$ for $p = 1$. Then this example is considered as one dimensional parameter with a single value the exact value $p^* = 1$. The region of the estimated parameter p is given by $[0.1, 2]$. We choose the initial sampling point as $p_0 = 2$. After 3 iterations, the numerical results of p_n are $p_1 = 0.7034$, $p_2 = 1.0417$ and $p_3 = 0.9999$. Therefore we obtained our numerical estimated value of the parameter p while the objective function value $G(p_3) = 1.121 \times 10^{-4}$. The computing time for this example is 2 s.

Example 4.2 We change the equation in 4.1 as follows

$$\begin{cases} u_{xx} = p^2u & \text{on } [0, 1] \\ u(0) = 1 & \text{and } u(1) = \frac{e^1 + e^{-1}}{2} \end{cases} \quad (13)$$

The system corresponding to numerical solution U_i , $F(\mathbf{U}, p) = 0$, is also changed accordingly. We still choose the data points V_i by solving $F(\mathbf{V}, p) = 0$ for $p = 1$. However, there should be two exact solutions for p^* : 1 and -1 . After 8 iterations, we found two solutions $p = 0.9999$ and $p = -0.9999$. The sequence of p_i for each iteration is shown in Fig. 3. The number of p_i is changing as i changes, but eventually converges to two points 1 and -1 . The computing time for this example is 3 s.

Example 4.3 Next we consider an example with two parameters by changing 4.1 to

$$\begin{cases} u_{xx} = p_1^2u & \text{on } [0, 1] \\ u(0) = 1 & \text{and } u(1) = p_2(e^1 + e^{-1}) \end{cases} \quad (14)$$

The system corresponding to numerical solution U_i , $F(\mathbf{U}, p) = 0$, is also changed accordingly. We still choose the data points V_i by solving $F(\mathbf{V}, \mathbf{p}) = 0$ for $\mathbf{p} = (1, 1/2)^T$. However, there should be two exact solutions for \mathbf{p}^* : $(1, 1/2)^T$ and $(-1, 1/2)^T$. After 8 iterations, we found these two solutions. The errors of \mathbf{p}_i for each iteration are shown in Table 1: We define

$$Err_1 = \min_j \|\mathbf{p}_i^j - (1, 1/2)^T\| \text{ and } Err_2 = \min_j \|\mathbf{p}_i^j - (-1, 1/2)^T\|,$$

where \mathbf{p}_i^j is the j -th solution for i -th iteration. The computing time for this example is 10 s.

Example 4.4 We consider a model for a genetic toggle switch in *Escherichia coli*, which was constructed in [7]. It is composed of two repressors and two constitutive promoters, where each promoter is inhibited by the repressor that is transcribed by the opposing promoter.

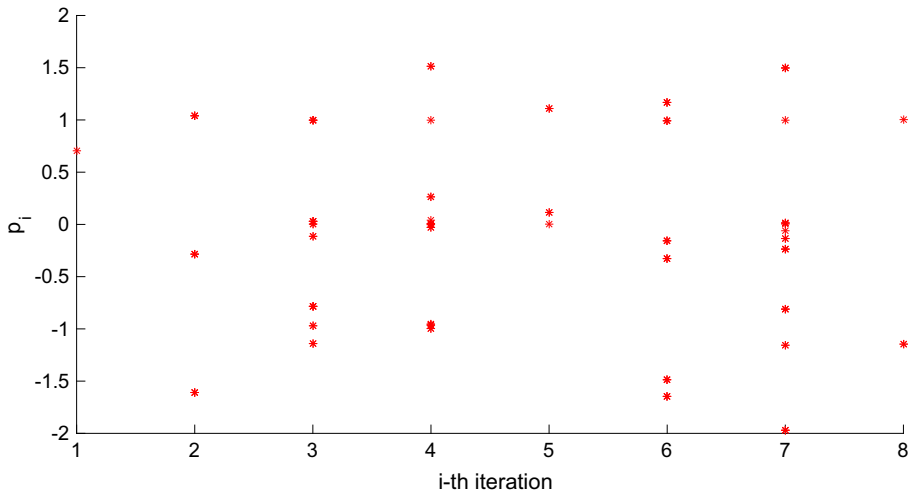


Fig. 3 The parameter sequence p_i versus i -th iteration: the number of solutions is changing as i goes from 1 to 8. Eventually the parameter sequence converges to two points 1 and -1

Table 1 The errors for each iteration

Iteration	Err_1	Err_2
1	0.71081	1.2619
2	0.73239	1.2619
3	0.42954	1.2619
4	0.30881	1.2619
5	0.11124	0.33258
6	0.0038304	0.33258
7	0.0013875	0.013271
8	$1.1e-7$	$2.64e-5$

Details of experimental measurement can be found in [7]. The following dimensionless model derived from a biochemical rate equation formulation of gene expression

$$\frac{du}{dt} = \frac{\alpha_1}{1 + v^2\beta_1} - u, \tag{15}$$

$$\frac{dv}{dt} = \frac{\alpha_2}{1 + u^3\beta_2} - v. \tag{16}$$

This is an ODE system, with four parameters $\mathbf{p} = (\alpha_1, \alpha_2, \beta_1, \beta_2)^T$. The values of these parameters are estimated in [7]. Here we set the solution of $p = (156.25, 2.5, 15.6, 1)^T$. We used Legendre polynomials as the gPC basis. The parameter space is four-dimensional ($N = 4$). The initial guess we chosen is $p_0 = (100, 1, 10, 1)^T$. After three iterations, the numerical solution we found is $\mathbf{p}_3 = (156.25, 2.5, 15.599, 0.999)^T$. The computing time for this example is around 2 min.

Example 4.5 We consider the following simplified model of a small glioma growing in the brain which elicits a response from the host immune system [27]. This model consists of four variables denoted by T, σ_{brain}, I and σ_{serum} which represent the concentration of glioma

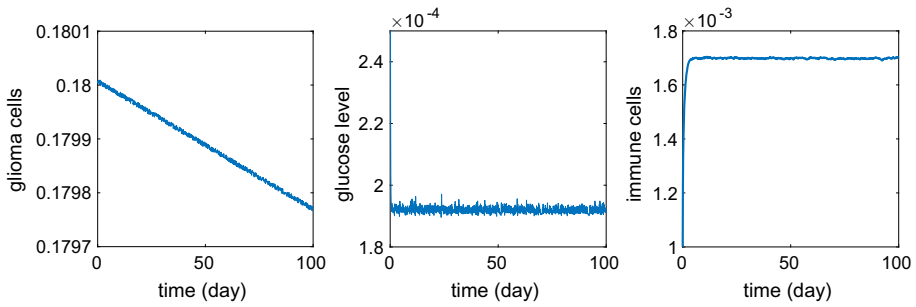


Fig. 4 Clinical data for glioma cells, glucose level and immune cells versus time

Table 2 The estimated parameter values for the system (17–20)

Parameter	Values		
α_T	0.65304	0.2743	0.5085
α_S	31.2803	28.1577	10.1253
d_{TT}	0.9570	0.3021	0.7125

cells, the concentration of glucose in the brain, the concentration of immune system cells and the concentration of serum glucose levels respectively. The model is described by the following system of differential equations

$$\frac{dT}{dt} = \alpha_T \sigma_{brain} T (1 - T/K_T) - d_T T - d_{TI} T I, \tag{17}$$

$$\frac{d\sigma_{brain}}{dt} = \alpha_\sigma (\sigma_{serum} + T + I) - d_{T\sigma} T \sigma_{brain} - (d_\sigma + \alpha_s) \sigma_{brain}, \tag{18}$$

$$\frac{dI}{dt} = \alpha_s \sigma_{brain} + d_{TI} T I - d_I I - d_{TT} T I, \tag{19}$$

$$\sigma_{serum} + \sigma_{brain} = \max\{\sigma_{min}, \sigma_0 \sin(6\pi t)\}. \tag{20}$$

We choose the following initial conditions:

$$T(0) = 0.18, \quad \sigma_{brain}(0) = 4.43 \times 10^{-3}, \quad I(0) = 10^{-3}.$$

In [27], the sensitivity analysis of parameter values has been studied and shown that α_T , α_s and d_{TT} are the most sensitive parameters with respect to the growth of glioma. Then we fix the other parameters as follows:

$$K_T = 2, \quad d_T = 10^{-4}, \quad d_{TI} = 0.072, \quad \alpha_\sigma = 20, \quad d_{T\sigma} = 1, \\ d_\sigma = 0.01, \quad d_I = 0.01, \quad \sigma_{min} = 8 \times 10^{-4}, \quad \sigma_0 = 1.6 \times 10^{-3}.$$

In order to accurately estimate these sensitive parameter values, we employ the clinical data from a public database, SAGEmap [21]. The data for period of 100 days is shown in Fig. 4. With respect to this clinical data, we choose the initial guess as $\alpha_T = 0.5$, $\alpha_s = 10$ and $d_{TT} = 0.7$. After 12 iterations, we found three group of parameter values which are shown in Table 2. The computing time for this example is around 5 min.

5 Conclusion

In this paper we introduced a numerical method which is based on polynomial chaos and homotopy continuation methods for estimating the parameter values of nonlinear systems of differential equations. The resulting numerical method is able to compute the multiple optimal parameter values with respect to a given dataset, and generates a cluster of sampling points which converge to different estimated parameter values. Several numerical examples are presented to test the presented numerical method and show that it works from single equations to systems with complex form.

Acknowledgements The authors have been supported by the American Heart Association under Grant 17SDG33660722.

References

1. Bates, D., Hauenstein, J., Sommese, A., Wampler, C.: Numerically Solving Polynomial Systems with Bertini, vol. 25. SIAM, Philadelphia (2013)
2. Blanchard, E.D., Sandu, A., Sandu, C.: A polynomial chaos-based Kalman filter approach for parameter estimation of mechanical systems. *J. Dyn. Syst. Meas. Control* **132**(6), 061404 (2010)
3. Coelho, F.C., Codeco, C.T., Gomes, M.G.: A Bayesian framework for parameter estimation in dynamical models. *PLoS ONE* **6**(5), e19616 (2011)
4. Dobson, I.: Computing a closest bifurcation instability in multidimensional parameter space. *J. Nonlinear Sci.* **3**(1), 307–327 (1993)
5. Friedman, A., Hao, W.: Mathematical modeling of liver fibrosis. *Math. Biosci. Eng.* **14**(1), 143–164 (2017)
6. Gallas, J.: Structure of the parameter space of the Hénon map. *Phys. Rev. Lett.* **70**(18), 2714 (1993)
7. Gardner, T.S., Cantor, C.R., Collins, J.J.: Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**(6767), 339–342 (2000)
8. Gomes, W., Beck, A., da Silva Jr, C.: Modeling random corrosion processes via polynomial chaos expansion. *J. Braz. Soc. Mech. Sci. Eng.* **34**(SPE2), 561–568 (2012)
9. Hao, W., Crouser, E.D., Friedman, A.: Mathematical model of sarcoidosis. *Proc. Natl. Acad. Sci. U.S.A.* **111**(45), 16065–16070 (2014)
10. Hao, W., Friedman, A.: Mathematical model on Alzheimer’s disease. *BMC Syst. Biol.* **10**(1), 108 (2016)
11. Hao, W., Hauenstein, J., Hu, B., Liu, Y., Sommese, A., Zhang, Y.-T.: Bifurcation for a free boundary problem modeling the growth of a tumor with a necrotic core. *Nonlinear Anal. Real World Appl.* **13**(2), 694–709 (2012)
12. Hao, W., Hauenstein, J., Hu, B., Sommese, A.: A three-dimensional steady-state tumor system. *Appl. Math. Comput.* **218**(6), 2661–2669 (2011)
13. Hao, W., Hauenstein, J., Hu, B., Sommese, A.: A bootstrapping approach for computing multiple solutions of differential equations. *J. Comput. Appl. Math.* **258**, 181–190 (2014)
14. Hao, W., Hauenstein, J., Shu, C.-W., Sommese, A., Xu, Z., Zhang, Y.-T.: A homotopy method based on weno schemes for solving steady state problems of hyperbolic conservation laws. *J. Comput. Phys.* **250**, 332–346 (2013)
15. Hao, W., Marsh, C., Friedman, A.: A mathematical model of idiopathic pulmonary fibrosis. *PLoS ONE* **10**(9), e0135097 (2015)
16. Hao, W., Nepomechie, R., Sommese, A.: Completeness of solutions of bethe’s equations. *Phys. Rev. E* **88**(5), 052113 (2013)
17. Hao, W., Nepomechie, R., Sommese, A.: Singular solutions, repeated roots and completeness for higher-spin chains. *J. Stat. Mech. Theory Exp.* **2014**(3), P03024 (2014)
18. Hao, W., Schlesinger, L.S., Friedman, A.: Modeling granulomas in response to infection in the lung. *PLoS ONE* **11**(3), e0148738 (2016)
19. Jaakkola, T.S., Jordan, M.I.: Bayesian parameter estimation via variational methods. *Stat. Comput.* **10**(1), 25–37 (2000)
20. Kramer, S.C., Sorenson, H.W.: Bayesian parameter estimation. *IEEE Trans. Autom. Control* **33**(2), 217–222 (1988)

21. Lal, A., Lash, A.E., Altschul, S.F., Velculescu, V., Zhang, L., McLendon, R.E., Marra, M.A., Prange, C., Morin, P.J., Polyak, K., Papadopoulos, N., Vogelstein, B., Kinzler, K.W., Strausberg, R.L., Riggins, G.J.: A public database for gene expression in human cancers. *Cancer Res.* **59**(21), 5403–5407 (1999)
22. Li, J., Xiu, D.: A generalized polynomial chaos based ensemble Kalman filter with high accuracy. *J. Comput. phys.* **228**(15), 5454–5469 (2009)
23. Lillacci, G., Khammash, M.: Parameter estimation and model selection in computational biology. *PLoS Comput. Biol.* **6**(3), e1000696 (2010)
24. Pence, B.L., Fathy, H.K., Stein, J.L.: A maximum likelihood approach to recursive polynomial chaos parameter estimation. In: American Control Conference (ACC), pp. 2144–2151. IEEE (2010)
25. Sommese, A., Verschelde, J., Wampler, C.: Numerical algebraic geometry. In: *The Mathematics of Numerical Analysis*, vol. 32 of Lectures in Applied Mathematics. Citeseer (1996)
26. Sommese, A., Wampler, C.: *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*, vol. 99. World Scientific, Singapore (2005)
27. Sturrock, M., Hao, W., Schwartzbaum, J., Rempala, G.A.: A mathematical model of pre-diagnostic glioma growth. *J. Theor. Biol.* **380**, 299–308 (2015)
28. Tarantola, A.: *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia (2005)
29. Vanlier, J., Tiemann, C., Hilbers, P., van Riel, N.: Parameter uncertainty in biochemical models described by ordinary differential equations. *Math. Biosci.* **246**(2), 305–314 (2013)
30. Voit, E., Chou, I.: Parameter estimation in canonical biological systems models. *Int. J. Syst. Synth. Biol.* **1**, 1–19 (2010)
31. Wu, H., Wang, F., Chang, M.: Dynamic sensitivity analysis of biological systems. *BMC Bioinf.* **9**(12), S17 (2008)
32. Xiu, D., Karniadakis, G.E.: Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. *Comput. Methods Appl. Mech. Eng.* **191**(43), 4927–4948 (2002)
33. Xiu, D., Karniadakis, G.E.: The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.* **24**(2), 619–644 (2002)
34. Xiu, D., Karniadakis, G.E.: Modeling uncertainty in flow simulations via generalized polynomial chaos. *J. Comput. Phys.* **187**(1), 137–167 (2003)
35. Zhan, C., Yeung, L.: Parameter estimation in systems biology models using spline approximation. *BMC Syst. Biol.* **5**(1), 14 (2011)
36. Zhusubaliyev, Z., Mosekilde, E.: *Bifurcations and Chaos in Piecewise-Smooth Dynamical Systems: Applications to Power Converters, Relay and Pulse-Width Modulated Control Systems, and Human Decision-Making Behavior*, vol. 44. World Scientific, Singapore (2003)