Journal of Approximation Theory and Applied Mathematics

An Approach for a Parameter Estimation with a Wavelet Collocation Method

M. Schuchmann and M. Rasguljajew from the Darmstadt University of Applied Sciences

Abstract

In this article we describe a parameter identification method for an ODE with an error approximation. We use a wavelet collocation method which can be applied to various problems like higher order ODEs, boundary value problems or PDEs. In the example we use the Shannon wavelet. In an example we apply this method on the pyridine problem form the reaction kinetics.

Introduction

We use the following approximation function which is a function in the subspace V_j of the $L^2(R)$:

$$y_j(t) := \sum_{k=k_{min}}^{k_{max}} c_k \phi_{j,k}(t)$$

V_j is defined in the MSA (multi scale analysis). $\{\phi_{j,k}(t)\}_{k\in Z}$ is an orthonormal basis of V_j with $\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k)$. k_{max} and k_{min} depend on the approximation interval $[t_0,t_{end}]$. For an ODE of order r the scaling function ϕ should be in $C^r(R)$. If the ODE is a System, we use one approximation function for every component.

With the parameter identification we want to get the solution of

$$y'(t) = f(y(t),t,p)$$

for $p = p_{exact}$. The solution with the unknown parameter p_{exact} we call $y_{p_{exact}}$. So the ODE depends on a parameter (vector) p and we have additionally measurements $\tilde{y}_i$ of the function values $y_{p_{exact}}(\tilde{t}_i)$ at the measurement points $\tilde{t}_i$.

A Classical Approach

A classical least square approach is to solve the following minimum problem

$$min \sum_{i=1,\dots,\tilde{m}} \left\| \tilde{y}_i - y_p(\tilde{t}_i) \right\|^2 . \quad (1)$$

Here and later ||.|| is the Euklid norm. One possibility for a parameter identification with the wavelet collocation method is to start with a parameter p_0 and solve

$$y_j'(t_i) = f(y_j(t_i),t_i,p_0), \text{ with } i = 1, 2, \dots, m$$
$$\text{and } y_j(t_0) = y_0, m = |k_{max} - k_{min}|.$$

Then we could apply a Gauss Newton step to solve numerically the minimum problem (1).

There we need additionally the numerically calculated derivates $\partial y_p(\tilde{t}_i)/\partial p$ at $p = p_0$ and we replace $y_p(\tilde{t}_i)$ through its linear approximation in the minimum problem (1). So we get $\hat{p}_1$ as a point at the minimum of the quadratic problem (because of the linearization), where we should additionally apply a descent test to get $p_1 = p_0 + \alpha(\hat{p}_1 - p_0)$.

A Minimum Residual Approach with an Error Estimation

Another approach is the direct minimization of

$$Q_{\alpha,\beta}(p,c) = \alpha \cdot \sum_{i=0,\dots m} \left\| y_j'(t_i) - f(y_j(t_i), t_i, p) \right\|^2 + \beta \cdot \sum_{i=1,\dots \tilde{m}} \left\| \tilde{y}_i - y_j(\tilde{t}_i) \right\|^2 + \left\| y_0 - y_j(t_0) \right\|^2 .$$

$y_j(t_0) = y_0$ can be also used as a constraint. The numerical calculated value of the minimum from $Q_{\alpha,\beta}$ we call later Q_{min} and we set $\alpha = \beta = 1$. The estimator of c und p we call $\hat{c}$ and $\hat{p}$.

To see, if the approximation function y_j has also small residuals on other points than the collocation points, we compare Q_{min} with

$$(2)\quad Q_{\alpha,\beta,a}(\hat{p},\hat{c}) = \alpha \cdot \sum_{i=0,\dots m_a} \left\| y_j'(\tau_i) - f(y_j(\tau_i), \tau_i, \hat{p}) \right\|^2 + \beta \cdot \sum_{i=1,\dots \tilde{m}} \left\| \tilde{y}_i - y_j(\tilde{t}_i) \right\|^2 + \left\| y_0 - y_j(t_0) \right\|^2$$

with $\tau_i = t_0 + i \cdot h/a$. $m_a = a \cdot m$ and an inter $a > 1$.

For the case $\alpha = \beta = 1$ we write short Q_a (like Q for $Q_{\alpha,\beta}$ in that case).

We already found further theoretical substantiations for Q_a for the error estimation.

An estimation with less cost is, when we estimate in a first step c and minimize $Q_{0,1}$:

$$(3a)\qquad Q_{0,1}(p,\hat{c}) = \min_c Q_{0,1}(p,c)$$

In a second step we estimate p:

$$(3b)\qquad Q_{1,0}(\hat{p},\hat{c}) = \min_p Q_{1,0}(p,\hat{c})$$

(3a) is only a quadratic problem and if f is linear in p (like in examples from the reaction kinetics) then (3b) is a quadratic problem, too. So we must only solve two times the normal equations. The cost for problem (3a) could be reduced even more if we chose special measurement points (see remarks 1 number 3) in the case that we have a small measurement error.

Remark 1:
y_j depends on c. For easier notation we write short y_j. If we write y_j outside a minimization problem depending on c, we mean that y_j was calculated from (3a) and $c = \hat{c}$.

The estimation with less cost means:
In step 1 we minimize (for c)

$$\sum_{i=1,\dots,\bar{m}} \left\| \tilde{y}_i - y_j(\tilde{t}_i) \right\|^2 + \left\| y_0 - y_j(t_0) \right\|^2 = \sum_{i=0,\dots,\bar{m}} \left\| \tilde{y}_i - y_j(\tilde{t}_i) \right\|^2$$

with $\tilde{y}_0 = y_0$ and $\tilde{t}_0 = t_0$.

If

(4) $$\min_c \sum_{i=0,\dots,\bar{m}} \left\| y_{p_{exakt}}(\tilde{t}_i) - y_j(\tilde{t}_i) \right\|^2 = 0$$

(or we must consider the approximation error, see remarks 2) and (5) $\tilde{y}_i = y_{p_{exakt}}(\tilde{t}_i) + \varepsilon_i$, then

$$\sqrt{\min_c \sum_{i=0,\dots,\bar{m}} \left\| \tilde{y}_i - y_j(\tilde{t}_i) \right\|^2} \le \|\varepsilon\|$$

$\|\varepsilon\|$ can be estimated with the square root of $Q_{0,1}(p, \hat{c})$ if the measurement errors are independent and identically normal distributed (with mean 0).

In the second step we minimize (for p)

$$\sum_{i=0,\dots,m} \left\| y_j'(t_i) - f(y_j(t_i), t_i, p) \right\|^2 .$$

With the Gronwall Lemma we know, if

$y_{\hat{p}}'(t) = f(y_{\hat{p}}(t), t, \hat{p})$ with $y_{\hat{p}}(t_0) = y_0$ and

$$||y_j(t_0) - y_{\hat{p}}(t_0)|| \le \delta ,$$

$$||y_j'(t) - f(y_j(t), t, \hat{p})|| \le M$$

and

$$||f(y_{\hat{p}}(t), t, \hat{p}) - f(y_j(t), t, \hat{p})|| \le L \cdot ||y_{\hat{p}}(t) - y_j(t)|| .$$

Then for $t \ge t_0$ we get:

$$||y_j(t) - y_{\hat{p}}(t)|| \le \delta \cdot e^{L(t-t_0)} + M / L \cdot (e^{L(t-t_0)} - 1)$$

In many simulations $y_j(t_0)$ is near y_0. For easier notation we assume $y_j(t_0) = y_0$, so we get

(6) $$\sum_{i=0,\dots,\bar{m}} ||y_j(\tilde{t}_i) - y_{\hat{p}}(\tilde{t}_i)||^2 \le (M / L)^2 \cdot \underbrace{\sum_{i=0,\dots,\bar{m}} (e^{L(\tilde{t}_i - t_0)} - 1)^2}_{:= C_L} .$$

Generally we can use $\hat{M}_a^2 = \max \; ||y_j'(\tau_i) - f(y_j(\tau_i), \tau_i, \hat{p})||^2$, $\tau_i = t_0 + i \cdot h/a$ and $i = 0, 1, ..., a \cdot m$ (with an integer $a > 1$) for an approximation of M^2 (compare with (2)). Here we know the following relation:

$$\hat{M}_a^2 \leq Q_a$$

With (4), (5) and (6) we get:

(7) $$\sqrt{\sum_{i=0,...,\tilde{m}} || \, y_{p_{exact}}(\tilde{t}_i) - y_{\hat{p}}(\tilde{t}_i) \, ||^2} \leq \|\varepsilon\| + M / L \cdot C_L$$

Remarks 2:
1) Form (4) and (5) follows

$$\sqrt{\sum_{i=0,...,\tilde{m}} \left\| y_{p_{exakt}}(\tilde{t}_i) - y_j(\tilde{t}_i) \right\|^2} \leq \|\varepsilon\|$$

what we used for (7).

2) The assumption (4) is fulfilled for example,

a) if $\tilde{m} = | \, k_{max} - k_{min} \, |$. With that choice for the parameters $\tilde{m}$ we interpolate the measurements. This is only advisable if the measurement error ε is small.

b) if ϕ has compact support, $y_{p_{exakt}}$ is in V_j and if k_{min} is small and k_{max} is big enough so that $\phi_{j,k}(t) = 0$ for $t \notin [\tilde{t}_0, \tilde{t}_{\tilde{m}}]$ and $k < k_{min}$ or $k > k_{max}$.

c) if we use the Shannon wavelet and $\tilde{t}_i = 2^{-j} \cdot i$. Here:

$$\phi_{j,k}(2^{-j} \cdot i) = 2^{j/2} \cdot \phi(i - k) = \begin{cases} 2^{j/2} & if \; i = k \\ 0 & else \end{cases}$$

So $c_i = 2^{-j/2} \cdot y(2^{-j} \cdot i)$ and $y(\tilde{t}_i) = y_j(\tilde{t}_i)$ with $y = y_{p_{exakt}}$. This is an interpolation property of the Shannon wavelet. If the measurement error ε is small, we can use $\tilde{y}_i$ for an approximation of c_i. If y is in V_j then $c_i = 2^{-j/2} \cdot y(2^{-j} \cdot i)$ are the exact bases parameters (follows from Shannon's theorem).

3) The assumption (4) can be not fulfilled if we use arbitrary $\tilde{t}_i$ and y is not in V_j or if the scaling function ϕ has no compact support and we choose a too less basis elements $\phi_{j,k}$ for the calculation for y_j. In the first case we have to consider the aliasing error and in the second case we have to consider the error depending on k_{min} and k_{max} caused through the missing bases coefficients (for a too big k_{min} or a too small k_{max}).

In both cases we have to consider the approximation error $\tilde{\varepsilon}$ (i.e. the term $\|\tilde{\varepsilon}\|$) in the equation (7), if

$$y_{p_{exakt}}(\tilde{t}_i) - y_j(\tilde{t}_i) = \tilde{\varepsilon} \quad \text{with } c = \hat{c} \quad \text{from } \min_c \sum_{i=0,...,\tilde{m}} \left\| y_{p_{exakt}}(\tilde{t}_i) - y_j(\tilde{t}_i) \right\|^2 = \|\tilde{\varepsilon}\|^2 .$$

Parameter Identification with the Pyridine Problem

The reaction scheme for the denitrogenation of pyridine is as follows:

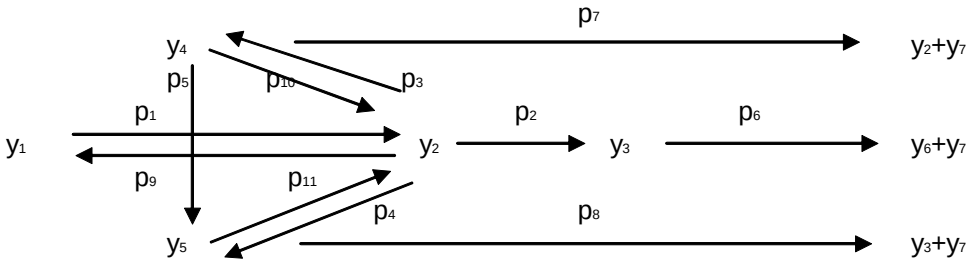


Figure 1. Reaction scheme pyridine

This yields to the following system of differential equations for the pyridine problem:

$$
\begin{aligned}
y_1' &= -p_1 y_1 + p_9 y_2 \\
y_2' &= p_1 y_1 - p_2 y_2 - p_3 y_2 y_3 + p_7 y_4 - p_9 y_2 + p_{10} y_4 y_6 \\
y_3' &= p_2 y_2 - p_3 y_2 y_3 - 2 p_4 y_3 y_3 - p_6 y_3 + p_8 y_5 + p_{10} y_4 y_6 + 2 p_{11} y_5 y_6 \\
y_4' &= p_3 y_2 y_3 - p_5 y_4 - p_7 y_4 - p_{10} y_4 y_6 \\
y_5' &= p_4 y_3 y_3 + p_5 y_4 - p_8 y_5 - p_{11} y_5 y_6 \\
y_6' &= p_3 y_2 y_3 + p_4 y_3 y_3 + p_6 y_3 - p_{10} y_4 y_6 - p_{11} y_5 y_6 \\
y_7' &= p_6 y_3 + p_7 y_4 + p_8 y_5 \qquad .
\end{aligned}
$$

The following results are related to the initial vector

$$y(0) = (1,0,0,0,0,0,0)^T \quad .$$

In this reaction, pyridine is transferred with the help of three catalysts in ammonia and pentane (see [3]), which takes about 5.5 seconds. This reaction takes place isothermally at 350° K under a pressure atm of 100.

It was used the parameter vector

$$p = (1.81,\ 0.894,\ 29.4,\ 9.21,\ 0.058,\ 2.43,\ 0.0644,\ 5.55,\ 0.0201,\ 0.577,\ 2.15)^T .$$

Now we come to the parameter estimation using the approximation y_j (k_{max} = 20 and k_{min} = -5).

We set j = 1, choose $I = [0, 5]$ (d.h. t_{end} = 5) and as collocation points we use

$$t_i = 1/16 \cdot i \text{, with } i = 1,...,80$$

(m = 80). As measurement points we used $\tilde{t}_i = 0.1 \cdot i$, with $i = 1,...,50$ ($\tilde{m} = 50$).

So:

$$Q_{\alpha,\beta}(p,c)=\alpha\cdot\sum_{i=1,\ldots,80}\left\|y_j'(t_i)-f(y_j(t_i),t_i,p)\right\|^2+\beta\cdot\sum_{i=1,\ldots,50}\left\|\tilde{y}_i-y_j(\tilde{t}_i)\right\|^2+\left\|y(t_0)-y_0\right\|^2$$

At first we do an estimation in two steps as described as follows

First Step:

$$Q_{0,1}(p,\hat{c})=\min_{c} Q_{0,1}(p,c)$$

Second Step:

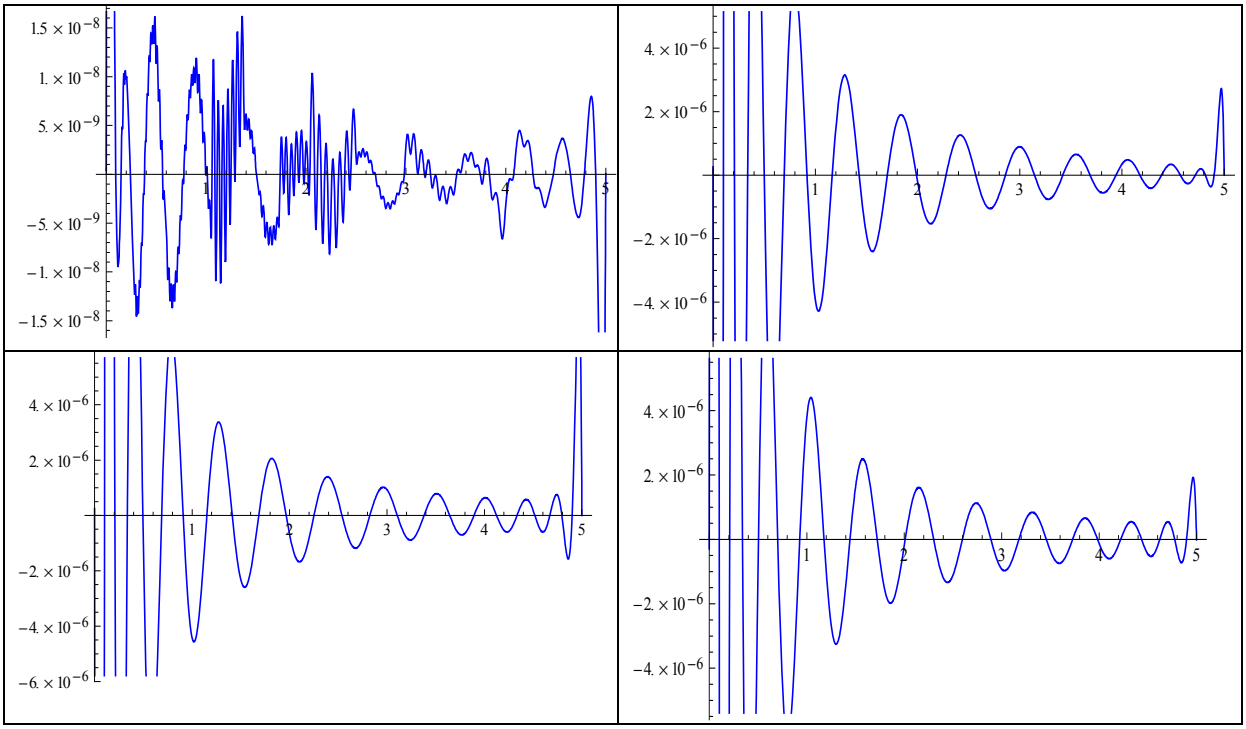
$$Q_{1,0}(\hat{p},\hat{c})=\min_{p} Q_{1,0}(p,\hat{c})$$

We had to solve two times a quadratic problem.

Here are the results of the first step:

$$Q_{0,1}(p,\hat{c})=\min_{c} Q_{0,1}(p,c)\approx 2.99612\cdot 10^{-9}$$

Here are the graphs of $y_1^{(i)} - y^{(i)}$ beginning with $i = 1$ (y is here the numerically calculated function using the Mathematica function NDSolve):



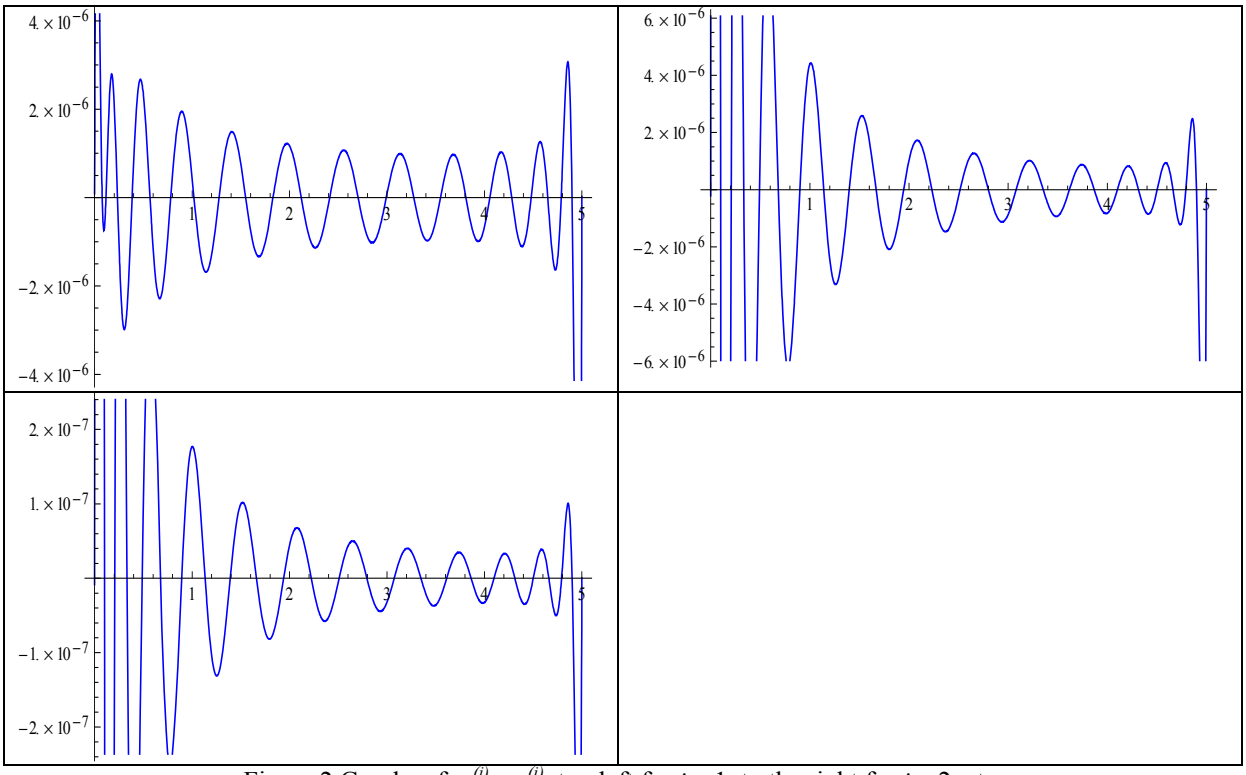


Figure 2.Graphs of $y_1^{(i)} - y^{(i)}$, top left for $i = 1$, to the right for $i = 2$, etc.

Here are the in the second step estimated parameters:

p_i	$\hat{p}_i$	$\lvert(\hat{p}_i - p_i)/p_i\rvert$
1.81	1.81125	0.00069058
0.894	0.894442	0.000494432
29.4	29.4016	0.0000557119
9.21	9.38508	0.0190099
0.058	0.0576356	0.00628269
2.43	2.42631	0.00151957
0.0644	0.0651214	0.0112015
5.55	5.53804	0.00215583
0.0201	0.0209635	0.0429597
0.577	0.576463	0.000929825
2.15	2.22964	0.0370417

We got: $Q_{1,0}(\hat{p},\hat{c}) = \min\limits_{p} Q_{1,0}(p,\hat{c}) \approx 1{,}18486 \cdot 10^{-4}$

Now again both parameters were estimated together (with $\alpha = \beta = 1$). The following was found:

$$\min_{c,p} Q_{1,1}(p,c) \approx 6.87561 \cdot 10^{-6}$$

For a comparison, $Q_2 \approx 4.92186 \cdot 10^{-4}$.

Here are the graphs of $y_1^{(i)} - y^{(i)}$ beginning with $i = 1$:

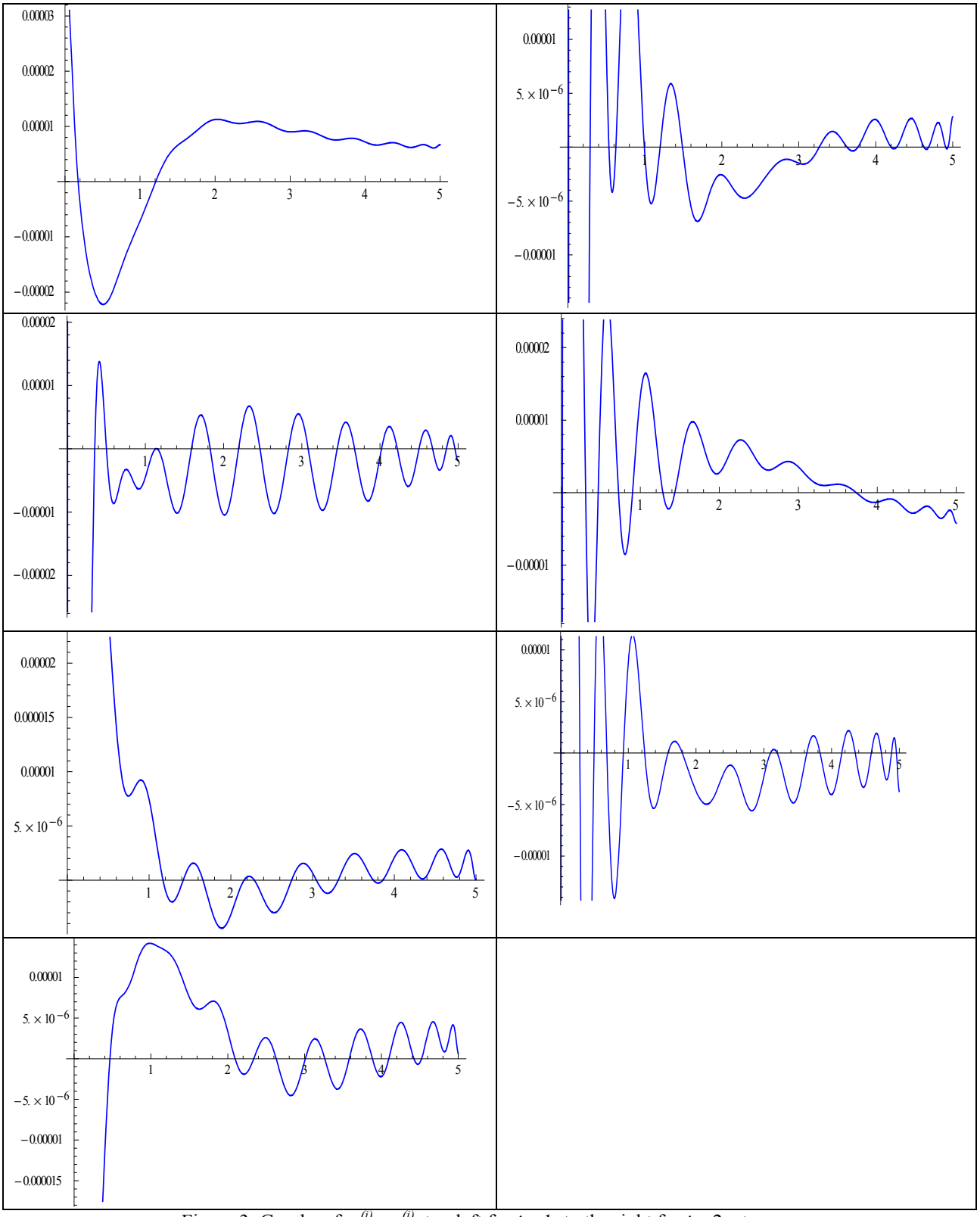


Figure 3: Graphs of $y_1^{(i)} - y^{(i)}$, top left for $i = 1$, to the right for $i = 2$, etc.

Here is the parameter estimation for p:

p_i	$\hat{p}_i$	$\lvert(\hat{p}_i - p_i)/p_i\rvert$
1.81	1.81033	0.000180172
0.894	0.893921	0.0000879803
29.4	29.4036	0.000124069
9.21	9.36569	0.0169047
0.058	0.0580298	0.000514097
2.43	2.42799	0.000827474
0.0644	0.0642593	0.00218449
5.55	5.57824	0.00508905
0.0201	0.020333	0.0115899
0.577	0.577234	0.000405281
2.15	2.1874	0.0173975

References

[1] S. Bertoluzza. *Adaptive wavelet collocation method for the solution of Burgers equation.* Transport Theory and Statistical Physics, Vol. 25, (1996).

[2] H. G. Bock. *Numerical Treatment of Inverse Problems in Chemical Reaction Systems.* Springer Series in Chem. Phys., Vol. 18, (1981).

[3] H. G. Bock. *Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen*. Bonner Mathematische Schriften Nr. 183, (1987).

[4] T. S. Carlson, J. Dockery, J. Lund. *A Sinc-Collocation Method for Initial Value Problems.* Mathematics and Computation, Vol. 66, No. 217, (1997).

[5] D. L. Donoho. *Interpolating Wavelet Transforms*. Tech. Rept. 408. Department of Statistics, Stanford University, Stanford, (1992).

[6] E. Hairer, G. Wanner. *Vol. 1 : Nonstiff Problems.* Springer, 2. Ed., (1993).

[7] E. Hairer, G. Wanner. *Vol. 2 : Stiff and Differential-Algebraic Problems.* Springer, 2. Ed., (1996).

[8] S. Kameswaran, L. T. Biegler. *Simultaneous Dynamic Optimization Strategies: Recent Advances and Challenges.* Computers and Chemical Engineering, Vol. 30, (2006).

[9] T. Lohmann, H. G. Bock, J. P. Schlöder. *Numerical Methods for Parameter Estimation and Optimal Experiment Design in Chemical Reaction Systems.* Ind. Eng. Chem. Res., Vol. 31, (1992).

[10] L. Qian. *On the Regularized Whittaker-Koltel'nikov-Shannon Sampling Theorem.* Proceedings of the Amarican Mathematical Society, Vol. 131, No. 4, (2002).

[11] M. Schuchmann. *Approximation and Collocation with Wavelets. Approximations and Numerical Solving of ODEs, PDEs and IEs*. Osnabrück, DAV, (2012).

[12] G. Strang. *Wavelets and Dilation Equations: A Brief Introduction.* SIAM Review Vol. 31, No. 4, (1989).

[13] M. Unser, T. Blu. *Comparison of Wavelets from the Point of View of their Approximation Error.* Proc. of SPIE Vol. 3458, Wavelet Applications in Signal and Image Processing, (1998).

[14] M. Unser. *Vanishing moments and the approximation power of wavelet expansions.* Proceedings of the 1996 IEEE International Conference on Image Processing, (1996).

[15] R. Vuduc. *A Wavelet Collocation Method for Solving PDEs*. J. Comp. Phys., Vol. 165 (2000).