Initialization Strategy for Nonlinear Systems

Problem presented by

Munawar Saudagar

Saudi Basic Industries Corporation (SABIC)

Executive Summary

The Study Group was asked to provide some hints concerning a choice
of initial values to be used for nonlinear algebraic systems. The group
has considered the available options and outlined the pros and cons of
various methods and provided some recommendations.

Version 1.0
July 13, 2011
iii+12 pages

Report coordinator

E. Coskun

Contributors

Rohanin Ahmad (Universiti Technologi Malaysia, Malaysia)
Erhan Coskun (Karadeniz Technical University, Trabzon, Turkey)
Anas Ismail (KAUST, Saudi Arabia)
Nabil Kerdid (Al-Imam university, Riyadh, Saudi Arabia)
Moetaz Mohamed (KAUST, Saudi Arabia)
Fatma-Zohra Nouri (University of Annaba, Algeria)
Amiya Pani (Indian Institute of Technology, Bombay, India)

Contents

1   Introduction

Given a nonlinear algebraic system

$$F(X) = 0 \tag{1}$$

where

$$F \colon \mathbb{R}^n \to \mathbb{R}^n$$

is a smooth function in the sense that it will justify the requirements for a chosen numerical method, the question presented to the Study Group was how to choose an initial guess $X^{(0)} \in \mathbb{R}^n$ to solve the system 1 provided that such a solution $P \in \mathbb{R}^n, F(P) = 0$, does exist.

The nonlinear systems brought to the attention of the Study Group originated from applications in Chemical Engineering. Rather than requiring a solution to a specific problem, the problem presenter has asked for a general initialization strategy. An implicit request was also to identify the right numerical method that will work with the initial guess. The presenter has indicated that common practice in his field was to use black box solvers and mentioned their dissatisfaction with such solvers, as they give no control over the problem.

In this report, we focus on giving appropriate hints for the right method and the initialization strategy for (1), though we warn that this is an active research field and that there is no superior method that will work for all the problems. We should also mention that this report is not intended to be survey of the existing methods. It just aims to provide an outline of common trends that may be helpful for people who just need a reasonable solution for a system that they met during their research. Before we go into the details we first remind the reader of the limitations of working in finite precision arithmetic. It is not just the method or the initial point to be blamed for the answer that seems not quite right. There are many issues one has to be aware of, including the coding of a nonlinear system, efficient ways the functions are evaluated and derivatives are computed, not to mention setting reasonable error tolerances and stopping criteria.

2   An overview of the methods

We mention the following methods for nonlinear algebraic systems:

- Fixed Point iterative Methods, in particular, Newton's Method and its variants.

- Continuation methods with different strategies.

- Minimization methods applied to $H(X) = F(X)^T F(X)$.

- Not so common, though, are the ODE methods in which the solution of (1) becomes a stable stationary point of a related ODE system, the so-called adjoint system for (1).

The idea behind the fixed point iterative methods is to reformulate the problem (1) as

$$X = G(X) \tag{2}$$

so that the required zero of F becomes a fixed point of G, to be determined. Although there may be countless options for the choice of G, not many of them will yield useful results. What is expected from a proper G is a convergent sequence defined by

$$X^{(k+1)} = G(X^{(k)}), k = 0, 1, \ldots \tag{3}$$

for a chosen $X^{(0)} \in \mathbb{R}^n$.

If the sequence $\left\{X^{(k)}\right\}$ converges to a point P, and G is continuous in a neighborhood of P, then P is a fixed point of the map $G : \mathbb{R}^{(n)} \to \mathbb{R}^{(n)}$.

The question is then, under what conditions the sequence will converge? A classical result from the fixed point theory is that G has to be a contraction mapping in the region, say D, containing P, with $X^{(0)} \in D$. Recall that G is called a contraction mapping on D if there exists a constant $c \in (0, 1)$ such that

$$\|G(X) - G(Y)\| \leq c\|X - Y\|, \tag{4}$$

for all $X, Y \in D$.

The difficulty, however, is to find a proper G which will yield a contractive mapping in a neighborhood of an unknown point. Even if the sequence converges, one is tempted to search for a better G, one that leads to a sequence that converges faster. At this point we recall the concept of the order of convergence of a sequence $\left\{X^{(k)}\right\}$.

We say that the sequence $\left\{X^{(k)}\right\}$ *converges* to the point P with an *order of convergence* M if there exist numbers C, M, and and a positive integer N such that

$$\|X^{(k+1)} - P\| < C\|X^{(k)} - P\|^M \tag{5}$$

for all $n \geq N$.

If $M = 1, C$ is required to be a positive number less than 1. In this case, the sequence converges linearly. If $M = 2$, the sequence is said to converge quadratically.

For any G, chosen without some proper search, the best hope would be a linearly convergent sequence, which may meet the expectations though.

2.1   Newton's Method

A particular choice of G is well known:

$$G(X) = X - J^{-1}(X)F(X) \tag{6}$$

provided that the required partial derivatives exist so that both the Jacobian, $J(X)$, and its inverse exists in a neighborhood of the unknown solution P.

In practice, one hardly ever needs J^{-1}, instead, at each iteration step, the linear system

$$J(X^{(k)})\Delta X = -F(X^{(k)}), k = 0, 1, \ldots \tag{7}$$

is solved and the sequence is updated by the iteration

$$X^{(k+1)} = X^{(k)} + \Delta X \tag{8}$$

until an appropriate termination criterion is met. The method is the celebrated Newton's Method which is still one of the first choices one considers for solving nonlinear algebraic systems.

Newton's method, as all others, has pros and cons:

- The method is locally quadratically convergent, meaning that it converges quadratically provided that the initial guess $X^{(0)}$ is close to the the solution being searched. How close is "close enough" depends on the problem. For Newton's method, between the initial point and the zero of the system, there should not be a relative extremum, because the iterations cannot proceed through the relative extremum point where the Jacobian is not invertible.

- One needs to be able to solve the required linear systems at each iteration system without much difficulty.

- One must have an efficient way of computing or approximating the partial derivatives in the Jacobian.

- Given a system (1), there is no prior knowledge of the nature of the Jacobian that will arise at each step of the iteration, so it is hard to expect what outcome it will yield unless a good initial guess, i.e. preferably the one that is closer to the zero of the system than any relative extremum, is chosen.

2.1.1   Variants of Newton's Method

Different variants of Newton's method have been developed to ease the computational load of the method. One such approach is to use the same Jacobian for all, or most or at least much of the iterations, to reduce the computational complexity of evaluating Jacobians at each step of the iteration. Another variant approximates the partial derivatives by appropriate finite difference methods, say Euler's method. Depending on a modification, different names are attached, such as inexact Newton's method, finite difference Newton's method and Quasi Newton's method where the Jacobian is updated using Broyden's type of formula.

Using difference methods to approximate the Jacobians has to be handled carefully, as the very small step size used to approximate the derivative may lead to excessive roundoff, while larger step sizes to a considerable truncation error.

However, none of the variants seems to lift the stringent restriction on the initial guess that it has to be a *good guess*.

A relatively new class of methods, the so-called global Homotopy Continuation Methods, have been proposed to ease the restriction on the initial guess.

This method, however, requires an effective implicit ODE solver.

2.2   Homotopy Continuation Methods

The idea behind the homotopy methods is to proceed from the known zeros of a smooth function, say G, to the zeros of F that we would like to determine. To this end a homotopy, a smooth function,

$$H : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$$

is defined such that

$$H(X, 0) = G(X),$$

and

$$H(X, 1) = F(X)$$

There are many ways to define H. The so-called *convex homotopy* is defined as

$$H(X, \lambda) = \lambda F(X) + (1 - \lambda)G(X), \quad \lambda \in [0, 1] \tag{9}$$

so that from the known zeros of $H(X, 0) = G(X) = 0$, one proceeds along the curve $H(X, \lambda) = 0$, till $\lambda = 1$, which gives

$$H(X, 1) = F(X) = 0,$$

the solution of the required system. Let us denote such a curve by $I(s) = (X(s), \lambda(s))$ with $I(0) = (a, 0)$ so our task is to trace the curve numerically till s_1, such that $I(s_1) = (X(s_1), 1)$. $X(s_1)$ will be a solution to (1).

For an appropriate $a \in \mathbb{R}^n$, a choice $G(X) = X - a$ is among the common choices. However, we will see below that it may not be a good choice for some problems. Another homotopy, the so-called *Newton Homotopy (or global homotopy)* is defined as

$$H(X, \lambda) = F(X) + (\lambda - 1)F(a) \tag{10}$$

We note that for the case of Newton Homotopy, for $\lambda = 0$, we may have a solution other than $X = a$ if F is not a *one-to-one* function.

The question is under what conditions (1) such a smooth curve exists; and (2) reaches the point $(X(s_1), 1)$ in finite length?

We refer to [5] for details, yet note that the first part of the question is answered by the implicit function theorem, which in terms of homotopy, translates to the assumption that $(a, 0)$ is a regular point of H, i.e., the Jacobian has full rank n. In this case, such a curve exists locally in the neighborhood of $(a, 0)$. It is also known that if all the zeros of H are regular points then such a curve is *diffeomorphic* to either a *circle* or to the *real line*.

On the other hand, having a curve that will reach the point $(X(s_1), 1)$ in finite length is another major issue, see the good reference just cited for details.

Here, we would like to focus on the problem of tracing such a curve numerically. There are several approaches, but we will concentrate on just two:

2.2.1   Case I: $I = I(\lambda)$

If the curve can be parametrized with respect to λ, to proceed from the known set of zeros, a classical imbedding method can be used: Divide the λ interval $[0, 1]$ into n subintervals of equal length with

$$\lambda(i) = \frac{i}{n}, \quad i = 0, 1, \ldots, n$$

and solve

$$H(x_i, \lambda_i) = 0, \quad i = 1, \ldots, n$$

using the solution (x_{i-1}, λ_{i-1}), the solution of the previous step, as an initial guess, with $x_0 = a$ being the solution to $H(x, 0) = G(x) = 0$. As for the solution procedure at each step several iterations of Newton's method is suggested.

This approach has pros and cons:

- Numerical implementation is relatively easy.

- The zeros are implicitly assumed to be continuous functions of the homotopy parameter, in order for the zeros of the previous step to serve as a good initial guess for the next step.

- It works only if the curve can be parametrized with λ. If the curve has turning points, then λ would not be an appropriate for parametrization, so then the method fails, since this would imply multivaluedness [5].

2.2.2   Case II: $I = I(s)$, with s being the arclength (arclength parametrization)

The curve $I(s) = (X(s), \lambda(s))$, with $H(I(s)) = 0$, $I(0) = (a, 0)$ with s being the arclength can be obtained as the solution of the following initial value problem:
Consider

$$H(I(s)) = 0 \tag{11}$$

and differentiate it with respect to s to have

$$H'(I(s))I'(s) = 0, \tag{12}$$

We choose $\|I'(s)\| = 1$ which implies that s is an arclength. Furthermore, the traverse direction has to be specified. For a positively oriented curve $I(s)$, we need

$$\det \left[\begin{array}{c} H'(I(s)) \\ I'(s)^* \end{array} \right] \quad > \quad 0 \tag{13}$$

where $H'(I(s))$ is the Jacobian of size $(n \times (n + 1))$, and $I'(s)^*$, a curve of size $1 \times (n + 1)$, is the tangent to the curve $I(s)$.
For the fixed homotopy

$$H(X(s), \lambda(s)) = \lambda(s)F(X(s)) + (1 - \lambda(s))(X(s) - a) \tag{14}$$

the equation (12) reads

$$\lambda'(s)\Big(F(X(s)) - (X(s) - a)\Big) + X'(s)\Big(\lambda(s)F'(X(s)) + (1 - \lambda(s))\Big) = 0$$

with

$$(\lambda'(s))^2 + (X'(s))^2 = 1$$

and the initial conditions

$$X(0) = a, \lambda(0) = 0$$

The condition (13) is needed for a proper orientation of the curve $I(s)$.

2.2.3   Homotopy paths for Case I and Case II

We consider each of the above mentioned approaches on very simple examples to highlight their pros and cons.

For simplicity we consider a function of one variable, $F(x) = e^x - (x + 3)$. This function has two real zeros $x_1 = 1.5052, x_2 = -2.9475$. First, we consider the global homotopy defined as

$$H_a(x, \lambda) = F(x) + (\lambda - 1)F(a)$$

for a given a.

In Figure 1(a), we draw the contour lines of the surface $H(x, \lambda)$ with contour levels at zero with $a = -4 : da : 4$ with an appropriate $da = 0.4$. We note that the homotopy paths go through the points $(x_1, 1), (x_2, 1)$, which we aimed to determine.

In Figure 1(b), we use the method of Case I, to determine the λ parametrized homotopy paths $x_a(\lambda)$ and draw the contour lines of the surface $H(x, \lambda)$ with contour levels at zero, as well as the computed homotopy paths for various a's, displayed in Figure 1(b) with marks "o". We used $\lambda_i = 0.1 \times i, i = 0, 1, \ldots, 10$ and a Newton type solver at each step of the iteration. From Figure 1(b), we see that the computed homotopy paths follow the zero level contour lines of the Homotopy, meaning that they are indeed the right paths for various a's which corresponds to the starting values of the curves on Figure 1(b).

For the same function as above, next we consider the convex homotopy

$$H_a(x, \lambda) = \lambda F(x) + (1 - \lambda)(x - a)$$

for a given a.

In Figure 2(a), we draw the contour lines of the surface $H_a(x, \lambda)$ with contour levels at zero with $a = -4 : da : 4$ with appropriate $da = 0.1$. We note that the homotopy paths go through the points $(x_1, 1), (x_2, 1)$, which we aimed to determine.

However, there exists a major distinction between the homotopy paths of Figure (1) and Figure(2), which correspond to different homotopies. In Figure(1), the homotopy paths were rather uniform, or technically speaking formed a family of *equicontinuous* curves and ended up at either zero with λ serving as a parameter of the curves for any a chosen. On the other hand, the homotopy paths corresponding
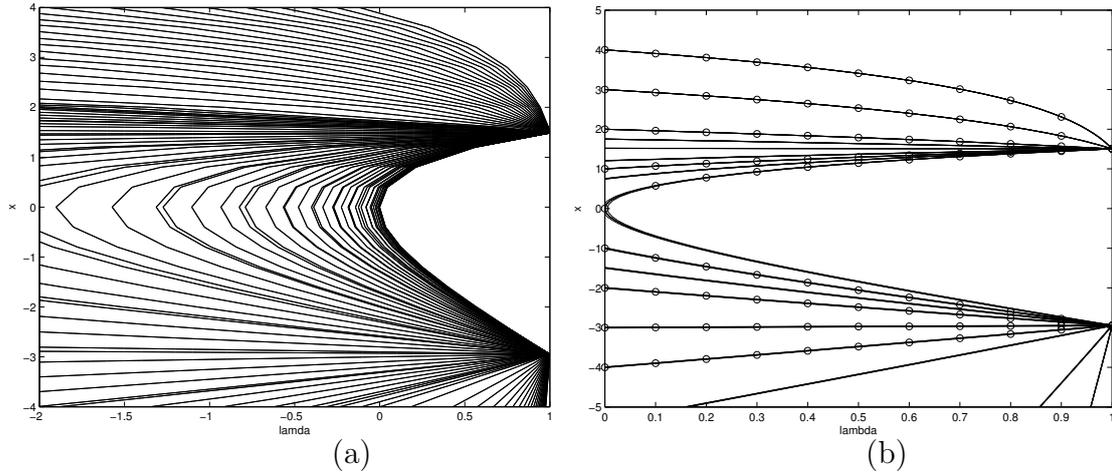
Figure 1: (a) Homotopy paths for various a's. (b) Paths computed with $\lambda_i = 0.1i$

to convex homotopy, displayed in Figure 2(a), display a different behavior around the starting point -3, which is very close to the zero $x_2 = -2.9475$. In Figure 2(b), we use the method in Case I, to determine the λ parametrized homotopy paths $x_a(\lambda)$. The contour lines of the surface $H(x, \lambda)$ with contour levels at zero, as well as the computed homotopy paths for various $a\prime s$, displayed in Figure 2(b) with marks 'o'. We note that none of the curves starting at $(a, 0)$, no matter what a is chosen, reaches the point $(x_2, 1)$. The reason is that in order for the relation $H(x, \lambda) = 0$ to determine a function $x = x(\lambda)$, in the neighborhood of a point (x_*, λ_*), we need $H_x(x_*, \lambda_*) \neq 0$. In this example, however, near the point $(-3, 0.5)$, the hypothesis of the implicit function theorem does not hold. So, one cannot guarantee the existence of such a curve in that neighborhood. Indeed, there is no such a curve as illustrated in Figure 2(b).

Next, we try the *implicit ODE system* as outlined in case II. Not to our surprise, we obtain similar results. The result is illustrated in Figure 3 for several values of a's as indicated in the illustration corresponding to the values $\lambda = 0$. We solved the system using MATLAB's implicit ODE solver `ODE15i`. We used the *decic* function to initialize $x'(0), \lambda'(0)$ values as needed by the solver. Furthermore, we checked the orientation by the inequality (13) and stopped the computation when $\lambda = 1$ using an appropriate *event* function.

Notice that although we use the homotopy method with arclength parametrization in order to avoid the difficulties with the fixed step version, we cannot reach at the zero near -3, no mather what initial values are chosen.

To remedy the situation, we propose a slightly different homotopy

$$H_a(x, \lambda) = \lambda F(x) + (1 - \lambda)(x - a)(x + a)$$

and illustrate the surface graph of $H_a(x, \lambda)$ over the domain $(x, \lambda) \in [-4, 4] \times [0, 1]$ in Figure 4(a). We also plot the surface $H = 0$ to draw attention to the intersection for a particular a. We observe that the intersection is a smooth curve that goes through the zeros of F and the homotopy paths can be easily computed using the method outlined in case I.
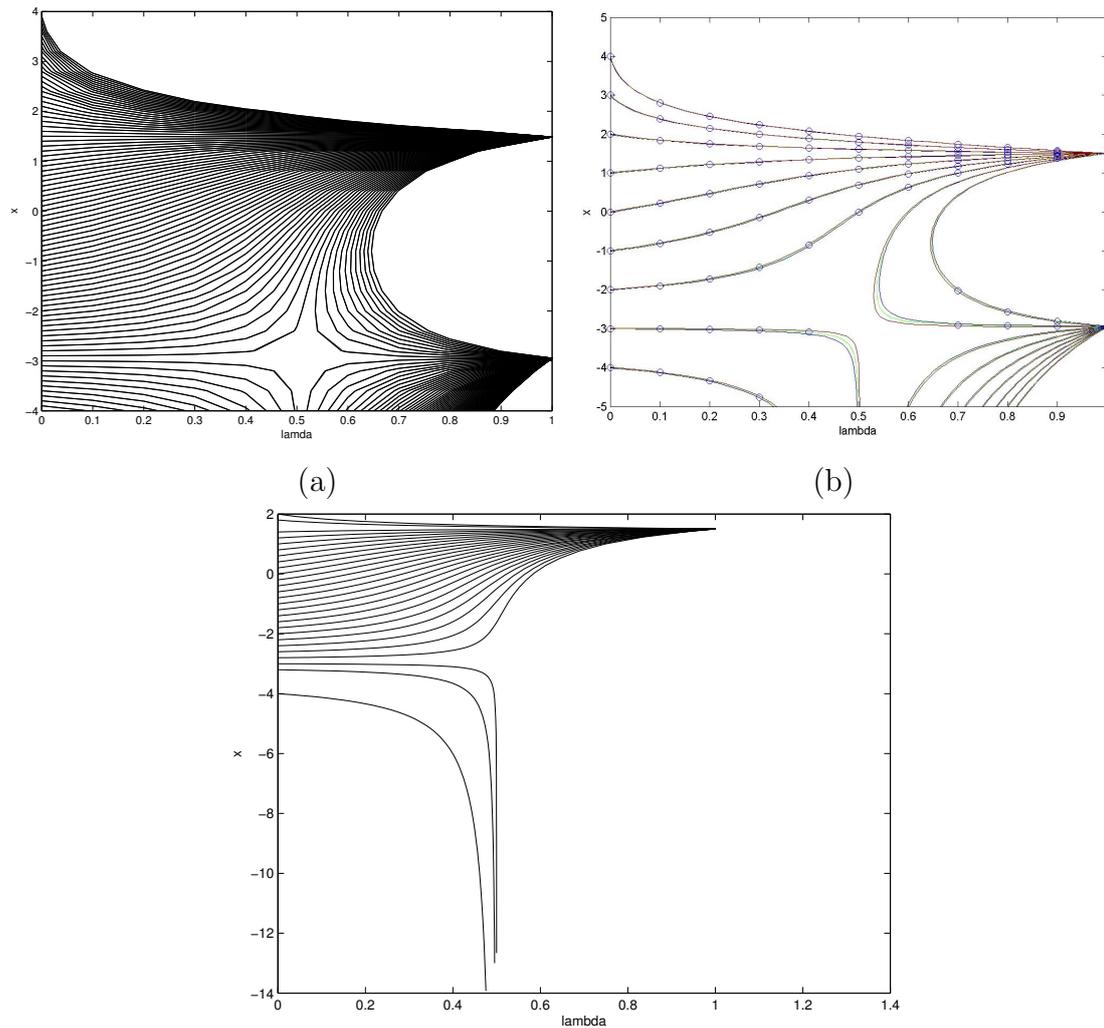
(a)   (b)

Figure 2: Homotopy paths obtained with ODE system.

As a result we *recommend that one has to be careful about choosing the right homotopy for a given problem, not just the numerical method of tracing the homotopy paths.*

3   Problem Specific Consideration

The researcher having the nonlinear system is most likely to have the best initial guess. The existing solvers also rely on the user for the initial guess. For example, the MATLAB equation solver `fzero` requires an initial guess, so does `fsolve`, MATLAB's nonlinear system solver.

The problem regarding the choice of an initialization strategy for nonlinear systems originated from having to solve differential algebraic systems where the algebraic part needs an initialization. To this end, the problem presenter brought to the attention of the Study Group the following model:
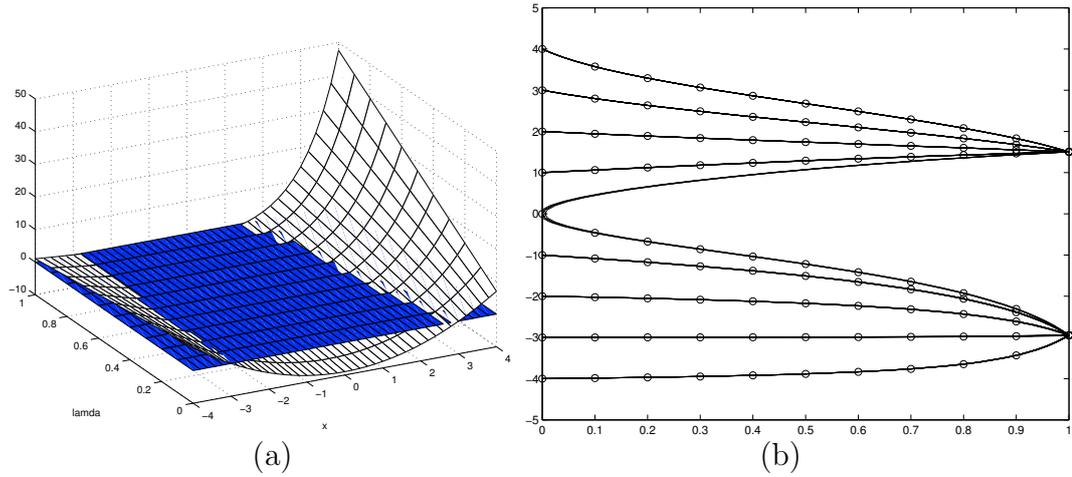
Figure 3: (a) $H(x, \lambda) = \lambda F(x) + (1 - \lambda)(x - a)(x + a)$.(b) Homotopy paths for $H(x, \lambda)$.

The model reads [4]:

$$\begin{aligned}
\frac{d[B]}{dx} &= -1442([B] - [B]_s) & (15) \\
\frac{d[C]}{dx} &= 28.8(R_1 - R_2) \\
9.88([A] - [A_s]) &= (R_1 + R_2) \\
\frac{1442}{28.8}([B] - [B]_s) &= R_1
\end{aligned}$$

where

$$\begin{aligned}
R_1 &= \frac{k_1 K_{1m} K_{2m}[A]_s [B]_s}{\left(1 + \sqrt{K_{1m}[A]_s} + K_{2m}[B]_s + K_{3m}[C]\right)^3} \, , \\
R_2 &= \frac{k_2 K_{1m} K_{3m}[A]_s [C]}{\left(1 + \sqrt{K_{1m}[A]_s} + K_{2m}[B]_s + K_{3m}[C]\right)^3}
\end{aligned}$$

In (15), A, the concentration of a chemical component is kept constant over a reactor. So the unknows are the concentrations of the other two components, namely $B(x)$ and $C(x)$ and the concentrations at the surface of the catalyst, denoted by $[A]_s$, and $[B]_s$. The $[B]_0, [C]_0$, the so-called feed concentrations of B and C, and the constant concentration A are given. The company provided 12 sets of experiments where the concentrations at the end of a tube is measured for different sets of initial concentrations and reactor lengths, L.

The system is to be solved over $(0, L]$ with given constant concentration A, and the feed concentrations $[B]_0, [C]_0$. There are two problems associated with the system (15):

- A closer look at the Nonlinear Algebraic Part of (15)

$$9.88([A] - [A_s]) - \frac{k_1 K_{1m} K_{2m}[A]_s[B]_s + k_2 K_{1m} K_{3m}[A]_s[C]}{\left(1 + \sqrt{K_{1m}[A]_s} + K_{2m}[B]_s + K_{3m}[C]\right)^3} = 0 \quad (16)$$

$$\frac{1442}{28.8}([B] - [B]_s) - \frac{k_1 K_{1m} K_{2m}[A]_s[B]_s}{\left(1 + \sqrt{K_{1m}[A]_s} + K_{2m}[B]_s + K_{3m}[C]\right)^3} = 0$$

for a fixed initial values of A, B, C, say the initial values A, $[B]_0, [C]_0$, where A is kept constant.

By a problem specific consideration we mean the reasonable choice of $[A_s]_0, [B_s]_0$ to be fed into the system to determine the right $[A_s]_0, [B_s]_0$. We than keep in mind the meaning of these quantities being the concentration of A and B at the surface of the catalyst. Then a reasonable initialization strategy would be to take those values to be very close to the concentrations in the liquid, i.e., A and $[B]_0$.

The Newton method applied to (16) then determines

$$[A_s]_0 = 10.1778$$
$$[B_s]_0 = 76.0810$$

with the initial guess for concentrations on the catalyst being equal to that of the liquid, i.e., $[A_s]_0 = A = 10.68, B_{s_0} = [B]_0 = 76.17$. The parameters were taken from the Eurokine Experiment 12 [4]. Notice that the initial guess was quite good. Therefore, Newton's method is expected to perform very well, as it does.

- **Forward problem**: Starting with intial conditions $B(0), C(0)$ with an appropriate set of parameters $k_1, k_2, k_{1m}, k_{2m}, k_{3m}$ determine the exit values $[B]_L, [C]_L$.

- **Parameter estimation problem:** Starting with an appropriate set of initial kinetic parameters, determine the right parameters to give the experimentally measured exit concentrations $[B]_L, [C]_L$.

- Once the initial values are obtained, the exit values of the unknowns in (15) can be obtained by any numerical method designed for differential algebraic systems. For example the MATLAB routine `ode15s` with a singular mass matrix gives

$$C = 3903.1$$
$$B = 34.0661$$

which are very close to the measured values of $C = 3903.0$ and $B = 34.2$.

- The result above indicates that we have used a good set of parameters
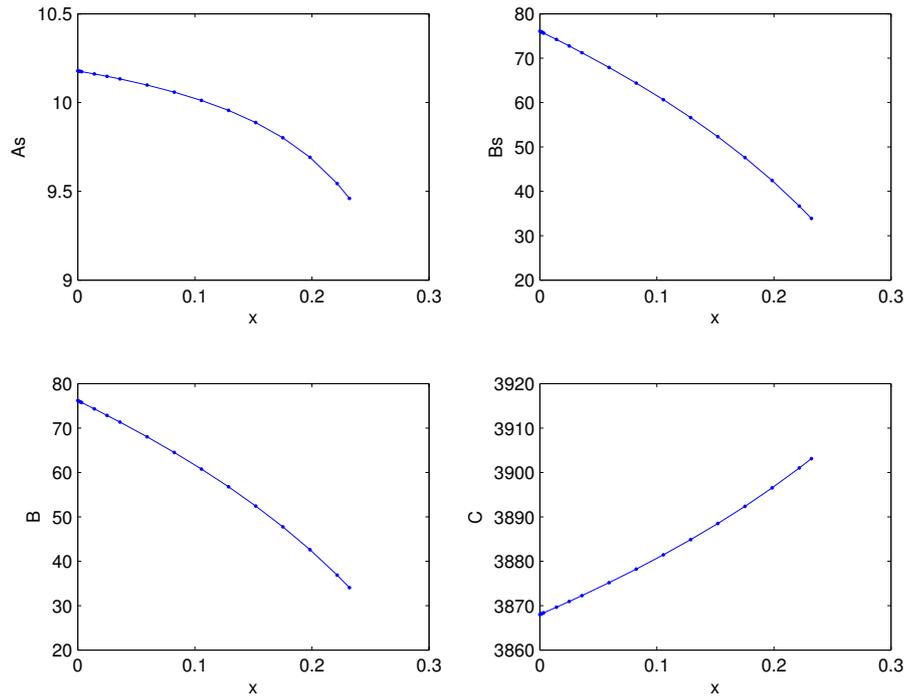
$$k_1 = 732.357582444032$$
$$k_2 = 412.503338973680$$
$$k_{1m} = 5.095730509841$$
$$k_{2m} = 0.777628071844$$
$$k_{3m} = 0.003078172740$$

The parameters are in the suggested ranges, yet we were unable to obtain them correctly by the end of the Study Group. The parameters for the experiments are obtained by a search algorithm using a Nonlinear Least Squares Method with lsqnonlin of MATLAB.

We say a good set of parameters, because there are other sets that yield the same result as shown below

k_1	732.3576	256.2028	340.9437	433.2298	551.7516
k_2	412.5033	162.5032	238.2723	310.2961	373.3151
k_{1m}	5.0957	5.9871	4.7670	3.8557	6.9613
k_{2m}	0.7776	0.4338	0.4887	0.5191	0.8117
k_{3m}	0.0031	0.0015	0.0016	0.0016	0.0027

Furthermore the above set is linearly independent, considered as a matrix with a determinant of 0.2908. Different sets are obtained through a search in different parts of the parameter domain given in [4] Although we minimize the difference between the measured values and exit values of concentrations, the solution components with

E–11

a different set of kinetic parameters listed above are in close agreement through the whole solution domain. Follow-up work on the parameter sensitivity analysis and related issues is being carried out by the report coordinator.

4   Final recommendations

- Initialization strategies by problem specific considerations work very well, as demonstrated in this report for the Eurokine test problem. Having a good estimate then suggests the use of Newton's method.

- In cases where a good initial estimate is not available, then homotopy methods can be used. The disadvantage that fixed-step-homotopy experiences when there are turning points can be avoided using homotopy methods with ar-clength parametrization. **However, the examples chosen in this report indicate that the choice of homotopy is much more important then the numerical method that traces the homotopy paths.**

- Optimization methods can also be used to minimize the associated problem, however, there is no guarantee that the methods will converge to the required zero. The MATLAB routine fsolve uses similar techniques.

- HOMEPACK: A package for solving systems by homotopy methods can be used if one decides to use such methods.

- ODE methods work if the solution being searched for is an asymptotically stable stationary point of the nonlinear system, and if the initial values can be chosen within the domain of influence of the stationary point.

Bibliography

[1] MATLAB: `http://mathworks.com`

[2] CONTIN: `http://gams.nist.gov/cgi-bin/serve.cgi/Package/CONTIN`

[3] HOMPACK: `http://gams.nist.gov/cgi-bin/serve.cgi/Package/HOMPACK`

[4] Rob J. Berger at al., Software functionality asseement for kinetic parameter estimation, model discrimination and design of experiments, EUROKIN(`http://www.eurokin.org`)

[5] Eugene L. Allgower and Kurt George, Introduction to Numerical Continuation Methods, Colorado State University, 1990.