



---

**Research article****Error-control B-spline Gaussian collocation PDE software with event detection****Jack Pew, Connor Tannahill and Paul Muir\***

Saint Mary's University, Halifax, NS, Canada, B3H 3C3

\* **Correspondence:** Email: [muir@smu.ca](mailto:muir@smu.ca).

**Abstract:** This paper introduces BACOLIKR, a new software package for the *error-controlled* numerical solution of systems of one-dimensional time-dependent partial differential equations (PDEs). A novel feature of this package is that it allows the user to specify a solution dependent condition, called an *event*, and then the software will determine the point in time at which the specified event occurs. This *event detection capability* can be used to provide an efficient and accurate means for dealing with time-dependent discontinuities in the PDEs or the boundary conditions.

BACOLIKR employs adaptive B-spline Gaussian collocation for the spatial discretization of the PDE system within a spatial error control algorithm. The event detection capability in BACOLIKR is based on its use of a modified version of the time integrator, DASKR, which implements event detection for time-dependent differential-algebraic equations. BACOLIKR was developed through modifications of an earlier error control PDE solver, BACOLI.

In this paper, we provide an overview of the BACOLI and DASKR packages and then describe the modifications that were made in order to develop BACOLIKR. We then show how BACOLIKR can be used for the effective solution of a number of PDE-based event detection problems including solution layer-boundary intersection detection and solution layer merge detection in a fluid mechanics model, critical tumor mass detection in a brain tumor model, steady state detection in the Gierer-Meinhardt model, and boundary condition event detection in a discontinuous heat flow model.

**Keywords:** B-splines; collocation; interpolation; error estimation; error control; event detection; partial differential equations (PDEs)

**Mathematics Subject Classification:** 65L06, 65L10, 65L80, 65M20, 65M70

---

**1. Introduction**

The numerical solution of partial differential equations (PDEs) arises as an important step in a wide variety of applications including investigations in microbial biogeochemistry [28], electrons in a

non-thermal plasma [27], pulsed microwave discharge in nitrogen [5], and one-dimensional, dynamic, catalytic reactor models [4]. This paper introduces a new software package, called BACOLIKR, for the computation of an *error-controlled* continuous numerical solution to a system of time-dependent PDEs in one space dimension.

Error control in this context means that for every time step taken by the solver, a high quality estimate of the error of the continuous numerical solution is computed, and the solution computed on the time step is not accepted unless this error estimate satisfies a user tolerance. If the error estimate does not satisfy the user tolerance, the computed solution on the step is rejected and the step is repeated using a smaller time step and/or a finer and more well adapted, spatial mesh. Advantages of computing an error-controlled numerical solution include the facts that the user can have reasonable confidence that the numerical solution has an error that is within a small multiple of the requested tolerance and that the cost of the computation will be proportional to the requested tolerance.

The key new feature of BACOLIKR is that it allows the user to specify a solution-dependent condition, called an *event*, and then the software will determine the point in time at which this event occurs. This feature is a generalization of the well-known event detection capability commonly available in state-of-the-art software for the numerical solution of initial value ordinary differential equations (ODEs). To our knowledge, BACOLIKR is the only available error-control PDE solver that offers event detection.

A simple example of an event detection task would be to ask when the solution to a PDE, at a given point in the spatial domain, takes on a specified value. However, as we will see in this paper, much more complex events that depend on, for example, a derivative of the solution, or on the spatial integral of the solution, can be treated. Examples discussed in this paper include solution layer-boundary intersection detection and solution layer merge detection in a fluid mechanics model, critical tumor mass detection in a brain tumor model, steady state detection in the Gierer-Meinhardt model, and boundary condition event detection in a discontinuous heat flow model.

*An essential point regarding event detection is that it can be used to provide an efficient and accurate means for dealing with time-dependent discontinuities.* It is well-known that the presence of such discontinuities can lead to substantial difficulties, with respect to both efficiency and accuracy, for a time integration method. See, e.g., [14] and [12], where multi-step and Runge-Kutta methods, respectively, are embedded in algorithms for the treatment of systems of time-dependent ODEs with discontinuities. See also, e.g., [20] and references within, for work on the detection and handling of discontinuities for systems of time-dependent differential-algebraic equations (DAEs). We demonstrate in this paper how the event detection capability of BACOLIKR can be used to provide efficient discontinuity detection for a heat flow problem with discontinuous boundary conditions.

It should be mentioned that all time integration algorithms that provide event detection can face issues when, e.g., there are an even number of events within a time-step or if two events are very close together. See, e.g., [25] and references within for further discussion of this point.

The advantage of software with an event detection capability is that the software itself, to within the accuracy with which the numerical solution is computed, can determine when a specified condition arises. This is in contrast to a standard PDE solver where the user must explicitly specify when the solver should finish, and in this case there is no straightforward way for the user to determine the point in time when a solution-dependent event of interest occurs.

*Accurate event detection requires that the numerical solution of the PDE be computed using error*

*control*. Since the event itself will depend on the value of the solution and/or its derivatives and/or its spatial integral, if the numerical solution is not computed accurately, then it will be impossible to determine the event time accurately. On the other hand, when an error-control solver has an event detection capability, the point in time when the event happens will also be determined to within the requested tolerance. The capability for event detection depends heavily on the fact that the approximate solution computed by BACOLIKR is represented as a continuous function of time and space. We will discuss this aspect of the approximate solution later in this paper.

The problem class we consider in this paper is a PDE system of size *NPDE* of the form,

$$\underline{u}_t(x, t) = \underline{f}\left(t, x, \underline{u}(x, t), \underline{u}_x(x, t), \underline{u}_{xx}(x, t)\right), \quad a \leq x \leq b, \quad t \geq t_0, \quad (1.1)$$

with separated boundary conditions,

$$\underline{b}_L\left(t, \underline{u}(a, t), \underline{u}_x(a, t)\right) = \underline{0}, \quad \underline{b}_R\left(t, \underline{u}(b, t), \underline{u}_x(b, t)\right) = \underline{0}, \quad t \geq t_0, \quad (1.2)$$

and initial conditions,

$$\underline{u}(x, t_0) = \underline{u}_0(x), \quad a \leq x \leq b. \quad (1.3)$$

In addition, there is a vector function of size *NRT*, called the *gstop function*, which has the form,

$$\underline{g}\left(t, x, \underline{u}(x, t), \underline{u}_x(x, t), \underline{u}_{xx}(x, t), \underline{u}_t(x, t)\right). \quad (1.4)$$

Each component of the *gstop* function defines an event and must be set to an expression whose value changes sign at the point in time where the event occurs. That is, *each event is defined to occur at a root of one of the components of the gstop function, and the task of locating the time at which an event occurs is implemented by finding roots of the components of the gstop function*.

The definition above for the *gstop* function is sufficiently general to allow a wide variety of events to be defined. It is the responsibility of the user to define an event in a way that make sense for the given application. If the user defines an event that cannot be satisfied by the solution of the PDE then this event will never be detected during the computation of the solution. See [25] for further discussion of well-posed events for time-dependent ODEs.

The BACOLIKR package has been developed through a modification of the BACOLI package [22], which is based on the earlier BACOL package [29]. Either BACOLI or BACOL was used to perform the numerical computations in the applications literature cited in the first paragraph of this section. The BACOLIKR package employs B-spline Gaussian collocation for the spatial discretization of the PDEs; this discretization process yields a system of time-dependent ODEs. These ODEs, together with the boundary conditions, represent a system of DAEs that are solved using a modification of the DAE solver, DASKR [6, 7], which is based on a family of Backward Differentiation Formulas (BDFs). An important feature of DASKR is that it provides event detection for systems of DAEs, and it is this capability that BACOLIKR builds upon to provide event detection for PDEs. (See Section 2.2 for further details.) For each time step, DASKR uses adaptive time stepping and/or BDF order selection to compute an approximate solution to the DAEs such that a corresponding error estimate is less than the user tolerance. For each time step accepted by DASKR, BACOLIKR then computes an estimate of the spatial error of the numerical solution. If this spatial error estimate satisfies the user tolerance, then the solution computed on the current time step is accepted. Otherwise, BACOLIKR will adapt the spatial mesh and then restart the computation at the beginning of the current time step.

The source code for BACOLIKR and the examples considered in this paper are available at [http://cs.smu.ca/~muir/BACOLI-3\\_Webpage.htm](http://cs.smu.ca/~muir/BACOLI-3_Webpage.htm).

This paper is organized as follows. In Section 2, we provide an overview of BACOLI and DASKR, the packages that were modified in order to obtain BACOLIKR. In Section 3, we describe the modifications that were made to the BACOLI and DASKR packages in order to develop BACOLIKR. Section 4 discusses the application of BACOLIKR to a collection of problems involving event detection where the event depends on the solution, a derivative of the solution, or the spatial integral of one of these. We close in Section 5 with our summary, conclusions, and an identification of future work.

## 2. Overview of BACOLI and DASKR

As mentioned above, the new software, BACOLIKR, has been developed through modifications of the BACOLI and DASKR software packages. We begin this section with a description of these packages.

### 2.1. BACOLI

For the problem, (1.1)-(1.3), as mentioned earlier, BACOLI employs B-spline Gaussian collocation for the spatial discretization. The numerical solution,  $\underline{U}(x, t)$ , is represented in terms of a B-spline basis of  $C^1$ -continuous piecewise polynomials in  $x$ , of degree  $p$  on each subinterval of a spatial mesh,  $\{x_i\}_{i=0}^{NINT}$ , that partitions  $[a, b]$ . Here  $NINT$  is the number of subintervals of the spatial mesh.  $\underline{U}(x, t)$  has the form,

$$\underline{U}(x, t) = \sum_{i=1}^{NC_p} y_{-p,i}(t) B_{p,i}(x), \quad (2.1)$$

where  $y_{-p,i}(t)$  is the unknown time-dependent vector coefficient of the  $i$ -th B-spline basis function,  $B_{p,i}(x)$ , and  $NC_p = NINT(p - 1) + 2$ . The B-spline basis is implemented in BACOLI using the de Boor B-spline package [9]. The use of Gaussian collocation for the spatial discretization of the PDE means that  $\underline{U}(x, t)$  is required to satisfy (1.1) at  $p-1$  collocation points on each spatial mesh subinterval, where the collocation points are chosen to be the images of the Gauss points,  $\{\rho_j\}_{j=1}^{p-1}$ , mapped on to each subinterval. Letting  $\eta_j$  be the  $j$ th collocation point, the corresponding collocation condition has the form,

$$\underline{U}_t(\eta_j, t) - \underline{f}(t, \eta_j, \underline{U}(\eta_j, t), \underline{U}_x(\eta_j, t), \underline{U}_{xx}(\eta_j, t)) = \underline{0}, \quad (2.2)$$

$j = 2, \dots, NC_p - 1$ .  $\underline{U}(x, t)$  is also required to satisfy the boundary conditions at  $\eta_1 = a$  and  $\eta_{NC_p} = b$ ; these have the form,

$$\underline{b}_L(t, \underline{U}(a, t), \underline{U}_x(a, t)) = \underline{0}, \quad \underline{b}_R(t, \underline{U}(b, t), \underline{U}_x(b, t)) = \underline{0}. \quad (2.3)$$

The Equations (2.2) and (2.3), represent an index-1 system of DAEs and the solution of this system gives the B-spline coefficients,  $y_{-p,i}(t)$ . Once these coefficients are available, the numerical solution to the PDE can be obtained from (2.1).

This collocation discretization yields a numerical solution that, for an arbitrary point in the spatial domain, has a spatial error that is  $O(h^{p+1})$ , where  $h$  is the maximum subinterval size [8, 11]. The numerical solution is said to be of spatial order  $p + 1$ .

In BACOLI, this DAE system is solved using a modified version of the DASSL package [21], which computes *error-controlled* approximations to the B-spline coefficients using a variable time step/variable order algorithm based on a family of BDFs of orders 1 to 5. One of the most significant modifications made to DASSL was the introduction of a new option for the solution of the linear systems that arise during the computation of the B-spline coefficients. Due to the use of a B-spline basis, these linear systems have what is known as an almost block diagonal (ABD) structure [10], and the modified version of DASSL employed in BACOLI makes use of the COLROW package [10], which is designed to efficiently solve such systems. The tolerance employed in the modified version of DASSL is slightly sharper than the user tolerance which is employed in the spatial error control algorithm implemented in BACOLI. This means that, generally, the time error associated with the computation of the B-spline coefficients will be slightly smaller than the spatial error of the approximate solution.

The spatial error control algorithm implemented in BACOLI requires that a high quality estimate of the spatial error of the numerical solution be computed. BACOLI has two options for computing a spatial error estimate for  $\underline{U}(x, t)$ . The first option, described in [1], is based on the observation that, at certain points within the spatial domain, the spatial accuracy of  $\underline{U}(x, t)$  is at least one order higher, i.e.,  $O(h^{p+2})$ , than it is at an arbitrary point in the spatial domain; these solution values are said to be superconvergent. The points at which  $\underline{U}(x, t)$  is superconvergent include the mesh points as well as certain other points (see [1]) internal to each subinterval. The  $\underline{U}_x(x, t)$  values at the mesh points are also superconvergent. For each spatial mesh subinterval, a Hermite-Birkhoff polynomial interpolant is constructed that interpolates the superconvergent  $\underline{U}(x, t)$  and  $\underline{U}_x(x, t)$  mesh point values at each end of the subinterval as well as the superconvergent  $\underline{U}(x, t)$  values that are internal to the subinterval. Furthermore, in order to obtain an interpolant *whose interpolation error is dominated by the error of the interpolated values*, this Hermite-Birkhoff interpolant also interpolates the two closest superconvergent  $\underline{U}(x, t)$  values internal to the left and right adjacent subintervals.

This Hermite-Birkhoff interpolant, known as the SuperConvergent Interpolant (SCI), has the following form on the  $i$ th subinterval,  $[x_i, x_{i+1}]$ . Let  $s_1 = x_i$ ,  $s_2 = x_{i+1}$ , and let  $w_j, j = 1, \dots, k$ , where  $k = p - 3$ , be the non-mesh points where the superconvergent solution values are to be interpolated. At time  $t$ , the Hermite-Birkhoff interpolant has the form [13],

$$\sum_{j=1}^2 H_j(x) \underline{U}(s_j, t) + h \sum_{j=1}^2 \bar{H}_j(x) \underline{U}_x(s_j, t) + \sum_{j=1}^k G_j(x) \underline{U}(w_j, t), \quad (2.4)$$

where  $x \in [x_i, x_{i+1}]$ ,  $h = x_{i+1} - x_i$ ,

$$H_j(x) = (1 - (x - s_j)\gamma_j) \frac{\xi_j^2(x)\phi(x)}{\xi_j^2(s_j)\phi(s_j)}, \quad \bar{H}_j(x) = (x - s_j) \frac{\xi_j^2(x)\phi(x)}{\xi_j^2(s_j)\phi(s_j)},$$

$$G_j(x) = \frac{\phi_j(x)\xi^2(x)}{\phi_j(w_j)\xi^2(w_j)},$$

where,

$$\phi(x) = \prod_{r=1}^k (x - w_r), \quad \phi_j(x) = \prod_{\substack{r=1 \\ r \neq j}}^k (x - w_r),$$

$$\xi(x) = \prod_{r=1}^2 (x - s_r), \quad \xi_j(x) = \prod_{\substack{r=1 \\ r \neq j}}^2 (x - s_r),$$

and,

$$\gamma_j = \sum_{i=1}^k \frac{1}{s_j - w_i} + 2 \sum_{\substack{i=1 \\ i \neq j}}^2 \frac{1}{s_j - s_i}.$$

Since the interpolation error of the above interpolant is dominated by the spatial error of the superconvergent interpolated values, the spatial error of the Hermite-Birkhoff interpolant is the same as the error of these superconvergent interpolated values, i.e.,  $O(h^{p+2})$ . Over  $[a, b]$ , these Hermite-Birkhoff interpolants, taken together, represent a  $C^1$ -continuous piecewise polynomial solution approximation, which we refer to as  $\underline{U}(x, t)$ . Scaled differences (see (2.5) and (2.6)) of  $\underline{U}(x, t)$  and  $\underline{U}(x, t)$  are computed in order to obtain spatial error estimates for  $\underline{U}(x, t)$ . These spatial error estimates are employed in BACOLI to provide what is known as *standard (ST) spatial error control*.

The second option available in BACOLI for computing a spatial error estimate for  $\underline{U}(x, t)$  costs slightly less to implement than the SCI, does not make use of approximate solution values from outside the subinterval when constructing the interpolant for the subinterval, and provides a slight overestimate of the spatial error. This interpolant makes use of the solution and derivatives values from each end of the subinterval as well as solution values at  $k = p - 4$  points that are internal to the subinterval. On the  $i$ th subinterval, this interpolant, which we will refer to as  $\overline{U}(x, t)$ , is also a Hermite-Birkhoff interpolant of the form (2.4), but in this case the  $w_j$  values are chosen so that *the interpolation error of this Hermite-Birkhoff interpolant is asymptotically equivalent to the spatial error for a collocation solution of one order lower than  $\underline{U}(x, t)$* . We therefore refer to  $\overline{U}(x, t)$  as the Lower Order Interpolant (LOI); the number of interpolation points is chosen so that the interpolation error dominates the error of the values being interpolated, and thus the LOI has a spatial error that is  $O(h^p)$ . See [2] for further details.

Scaled differences (again see (2.5) and (2.6)) of  $\underline{U}(x, t)$  and  $\overline{U}(x, t)$  provide an estimate of the spatial error for  $\underline{U}(x, t)$ . Since  $\underline{U}(x, t)$  is returned to the user but the spatial error control is based on a spatial error estimate for a solution approximation that is of one order lower than  $\underline{U}(x, t)$ , we have what is known as Local Extrapolation (LE) error control - see, e.g., [17].

When BACOLI is called with a given value for  $p$ , it computes and returns a numerical solution based on B-splines of degree  $p$ . If the ST spatial error control mode is chosen, then the solver constructs the SCI to generate a spatial error estimate which is then used as the basis for ST spatial error control. If the LE spatial error control mode is chosen, then BACOLI constructs the LOI and uses it to generate a spatial error estimate which is then used to provide LE spatial error control. Thus the availability of the two types of interpolants corresponds to providing two options for spatial error control, ST mode or LE mode, similar to what is available when a Runge-Kutta formula pair is used to provide error control for an initial value ODE - see, e.g., [17]. See [23] for a detailed performance analysis of BACOLI using these two error control modes.

For either error control mode, two types of spatial error estimates for  $\underline{U}(x, t)$  are computed by BACOLI. The first is the set of error estimates are given in (2.5). Each of these represents a scaled spatial error estimate, over the entire spatial domain, for one component of the solution. They have the

form,

$$E_j(t) = \sqrt{\int_a^b \left( \frac{U_j(x, t) - \widehat{U}_j(x, t)}{ATOL_j + RTOL_j |U_j(x, t)|} \right)^2 dx}, \quad (2.5)$$

where  $t$  is the current time,  $ATOL_j$  and  $RTOL_j$  are the absolute and relative tolerances for the  $j$ -th approximate solution component,  $U_j(x, t)$ , and  $\widehat{U}_j(x, t)$  is the  $j$ th component of either  $\widetilde{\underline{U}}(x, t)$  or  $\overline{\underline{U}}(x, t)$ . The second set of spatial error estimates are given in (2.6). Each of these provides a scaled spatial error estimate, over all components of  $\underline{U}(x, t)$ , for the  $i$ th subinterval. These have the form,

$$\widehat{E}_i(t) = \sqrt{\sum_{j=1}^{NPDE} \int_{x_{i-1}}^{x_i} \left( \frac{U_j(x, t) - \widehat{U}_j(x, t)}{ATOL_j + RTOL_j |U_j(x, t)|} \right)^2 dx}. \quad (2.6)$$

These spatial error estimates are computed after each accepted time step taken by DASSL. A step is accepted when  $E_j(t) < 1$  for  $j = 1, \dots, NPDE$ . Otherwise the step is rejected and the  $\widehat{E}_i(t)$  values are used in a spatial mesh refinement algorithm that attempts to construct a new mesh such that (i) the numerical solution computed on that mesh will have a spatial error estimate that satisfies the tolerance and (ii) the spatial error estimates over the subintervals of that mesh will be approximately equidistributed. Both the location and number of mesh points can be changed during a remeshing in order to adapt to the size (with respect to the user tolerance) and distribution of the spatial error estimates over the spatial domain. Once a new spatial mesh is constructed, solution information from the current time step and several previous time steps is interpolated to the new mesh. See [31] for further details.

Important related work has seen the development of the software package, EBACOLI (Extended BACOLI) [16], which is a modification of BACOLI that can compute error-controlled solutions to systems of PDEs coupled with ODEs in time and/or space.

## 2.2. DASKR

The DASKR solver was obtained through an extension of the solver DASPK [6], which itself was developed from the original member of this software family, DASSL. As mentioned earlier in this paper, DASKR is based on a family of BDFs and uses both adaptive time stepping and BDF order selection to control an estimate of the temporal error. DASKR represents the approximate solution it computes in terms of a continuous piecewise polynomial interpolant. The order of the interpolant is chosen to be consistent with the order of the BDF used to obtain the discrete solution approximation at the end of the step. In addition to a choice of direct methods (dense or banded) for the treatment of the linear systems that arise during the computation of a numerical solution, DASKR also provides the user with the option of using a Krylov method, the Generalized Minimum Residual (GMRES) method, in either complete or incomplete form, with scaling and preconditioning [6], for use when the linear systems are large. As well, DASKR improves upon DASSL by providing an option for the calculation of consistent initial conditions for the DAE system to be solved, when the user is not able to provide these.

Furthermore, as mentioned earlier, DASKR has a *time-dependent event detection capability* that can be employed while the solver is computing a numerical solution to a DAE system. At the end of each accepted time step, DASKR calls the user's `gstop` function in order to monitor for sign changes in any

of the components of the  $gstop$  function. When a sign change is detected, DASKR uses the interpolant to the approximate solution, together with the search algorithm described in [18], to locate the point in time where the root of the corresponding component of the  $gstop$  function is located, and then returns to the calling program.

While, as mentioned above, DASKR has the capability for treating large DAE systems using the preconditioned GMRES algorithm, we have not yet incorporated this feature of DASKR into the BACOLIKR package; this would require a further major modification to BACOLIKR and therefore, later in the paper, we identify this as a potential project for future work. Of course, further investigation of the effectiveness of the event location strategy of DASKR for (large) real-world problems would be required. See, e.g., [26], for further discussion on this point.

Also, the algorithm provided within DASKR for computing consistent initial conditions, mentioned earlier, is not employed because BACOLIKR computes its own consistent initial conditions for the DAE system before calling DASKR.

### 3. Development of BACOLIKR

As mentioned earlier, event detection is implemented through a user defined  $gstop$  function, each component of which is used to characterize an event; this is done by writing each component of the  $gstop$  function so that it has a root at the point in time where the event occurs. It is therefore common for event detection software to be described in terms of a root finding capability where the goal of determining the time at which an event occurs is described in terms of finding a root of the  $gstop$  function. Thus, in this section, we make reference to root finding rather than event detection when describing the software modifications. Below we describe the major modifications that were made to BACOLI and DASKR in order to develop BACOLIKR. Modifications were also made to the interface of BACOLIKR. We do not include these details here but instead refer the reader to [24] for detailed descriptions of these modifications.

#### 3.1. Major modifications

The overall structure and user interface for BACOLIKR are similar to that of BACOLI and we therefore refer the reader to [22] for additional details.

- As mentioned earlier in this paper, BACOLI makes use of a modified version of DASSL. See Section 3 of [29] for a detailed description of the changes that were made to DASSL. In order to use DASKR within BACOLIKR, it was necessary to make similar changes. One major change involved modifying DASKR to provide an option for it to use the ABD linear system solver, LAMPAK [19], to solve the ABD linear systems that arise. (BACOLIKR uses LAMPAK, rather than the COLROW package used by BACOLI, in order to remove any proprietary dependencies. COLROW package was published in the Collected Algorithms of the ACM and is therefore subject to the ACM Software Copyright and License Agreement The use of LAMPAK rather than COLROW does not lead to any significant impact on performance.)
- A subroutine called BACRT was added. This subroutine calls the user's  $gstop$  subroutine, which we will refer to as  $RT$ . For a given time,  $t$ ,  $RT$  evaluates the components of the user's  $gstop$  function (1.4). These components can depend on the current solution approximation,  $\underline{U}(x, t)$



and/or its derivatives,  $\underline{U}_t(x, t)$ ,  $\underline{U}_x(x, t)$ ,  $\underline{U}_{xx}(x, t)$ , and spatial integrals of any of these. The BACRT routine provides an interface between the rest of BACOLIKR and the user's RT routine in order to simplify the argument list for the RT routine and hide a number of implementation details.

- A number of small changes were made to allow communication between the DASKR, BACOLIKR, and BACRT routines, in order to implement the root finding capability.
- When an event is located, BACOLIKR will return to the calling program. There are three common scenarios that arise regarding what should happen next. In the simplest case, the user is interested only in locating the event and the computation can simply terminate. In the second case, the user wishes to continue the computation but does not wish to make any changes to the problem; in this case BACOLIKR should simply continue the computation; this is what is known as a *warm start* and this is the default mode when BACOLIKR is called again after it has returned having located an event. The third case is when the user wishes to change the problem (either the PDEs or the boundary conditions) after the event has been located. This introduces a discontinuity into the problem, and in this case, BACOLIKR should be called using a *cold start*. This means that BACOLIKR will restart DASKR with a low order BDF and a small step size. Use of a cold start leads to a much more efficient computation than if a warm start were attempted. In order to accommodate this last case, a capability was added to allow the user to force BACOLIKR to restart DASKR for the next time step using a cold start.

#### 4. Application of BACOLIKR to event detection problems

In this section, we demonstrate the capabilities of the BACOLIKR package by showing how to apply it to a number of event detection problems that are based on the solution of a PDE or a system of PDEs. Each subsection will consider one problem and will outline how to write an appropriate root finding routine and corresponding main program in order to implement the specific type of event detection required for each problem. These examples can be easily modified to be adapted to whatever application the user is interested in.

In order to apply BACOLIKR to an event detection problem, the user needs to modify a module called ROOTFINDING. The primary component of this module is the subroutine called RT, within which the characterization of a single event or multiple events can be specified. As mentioned earlier, the specification of a event may require, within RT, an auxiliary computation involving, for example, the evaluation of the numerical solution, its derivatives in time or space, or the integral of one of these over the spatial domain. The evaluation of the numerical solution or its time or spatial derivatives is performed through a call to the VALUES routine. Within the RT routine, assignments are made to the gstop vector, RVAL, in order to define the conditions that characterize each event as a root of one component of the gstop function. In addition to the RT routine, the ROOTFINDING module includes the SETSOL routine. The latter dynamically allocates a work array that is used within the RT routine, and, for those applications in which a spatial integral must be computed, it also computes the Gaussian quadrature points that will later be used to obtain a numerical approximation to the integral.

The other primary software component associated with the use of BACOLIKR for event detection is the main program that will make calls to the solver routine, BACOLI95, as appropriate for whatever event or events are to be detected. The other software components of BACOLIKR that are used in the main program are the BACOLI95\_INIT routine, which initializes the

computation, the BACOLI95\_VALS routine, which is used to evaluate the numerical solution and its first and second spatial derivatives, the BACOLI95\_SOL data structure which contains a number of fields where information associated with the approximate solution is stored, and the BACOLI95\_SOL\_TEARDOWN routine which must be called at the end of the computation to release the dynamic memory that is allocated during the computation.

In addition to modifying the ROOTFINDING module and the main program, the user must also provide problem definition routines that define the PDE(s), the boundary conditions, and the initial condition(s). We refer the reader to, [http://cs.smu.ca/~muir/BACOLI-3\\_Webpage.htm](http://cs.smu.ca/~muir/BACOLI-3_Webpage.htm), to see complete source code for the main program, the ROOTFINDING module, and the problem definition routines for each problem we consider.

#### 4.1. Solution layer crossing detection for the One Layer Burgers Equation

Burgers equation is a standard model in fluid mechanics. Here we consider an instance of this problem which we call the One Layer Burgers Equation (**OLBE**); it has the form,

$$u_t = \epsilon u_{xx} - uu_x, \quad (4.1)$$

with boundary conditions at  $x = 0$  and  $x = 1$  ( $t > 0$ ) and an initial condition at  $t_0 = 0$  ( $0 \leq x \leq 1$ ) taken from the exact solution,

$$u(x, t) = \frac{1}{2} - \frac{1}{2} \tanh\left(\frac{x - \frac{t}{2} - \frac{1}{4}}{4\epsilon}\right), \quad (4.2)$$

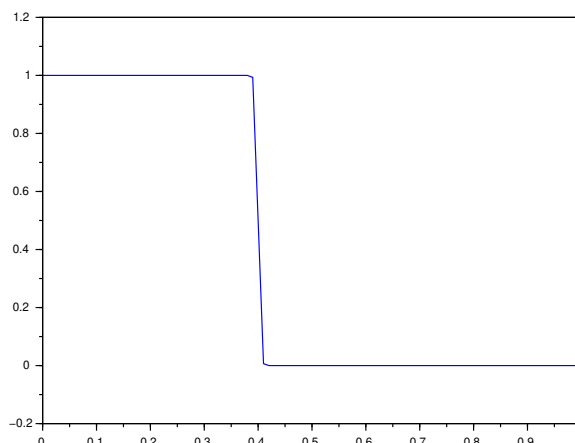
where  $\epsilon$  is a problem-dependent parameter. We will choose  $\epsilon = 10^{-3}$  for this example. We solve this problem from  $t_0 = 0$  to  $t_f = 2$  and choose a tolerance of  $10^{-6}$ . For  $t_0 = 0$ , the solution has a sharp layer located at  $x \approx 0.25$ ; the solution to the left of the layer is approximately equal to 1; the solution to the right of the layer is approximately equal to 0. As  $t$  goes from 0 to 1, the layer moves to the right and is located at  $x \approx 0.75$  for  $t = 1$ . See Figure 4 of [30] for a plot of the solution to this problem.

In order to demonstrate the event detection capability of BACOLIKR in a simple form, we will define a very basic event for this problem. The task will be to determine the time at which the approximate solution,  $U(x, t)$ , satisfies the condition that  $U(0.4, t) = 0.5$ . (If we define the layer to be the point at which the solution is halfway between the transition from a value of 1 to a value of 0, then setting the event condition to  $U(0.4, t) = 0.5$  is equivalent to asking when the layer crosses the point  $x = 0.4$ .) Thus there will be one root and the gstop function will be  $U(0.4, t) - 0.5$ .

Within the RT routine, we make one call to the VALUES routine which computes the value of the approximate solution at the current time, for  $x = 0.4$ . We pass the vector of B-spline coefficients to the VALUES routine so that the solution itself is computed. The difference between the returned solution approximation and the target value, 0.5, is then assigned to the RVAL output argument of RT in order to define the event for this example.

The corresponding main program first initializes the computation with a call to the BACOLI95\_INIT routine. This is followed by a call to the SETSOL routine and then a call to the BACOLI95 solver routine. The value of the gstop function at the end of each step is monitored to look for a change in its sign. When such a step is identified, a search is performed to determine the specific point in time when the root of the gstop occurs, and then BACOLI95 returns to the main program with an indication that the root has been found.

For this example, we find that the event occurs when  $t \approx 0.3$ . The main program writes out the solution at this time and then calls the solver again in order to complete the computation through to  $t_f = 2$ . Since the problem has not changed, a warm start can be employed. The solver returns with an indication that  $t_f$  has been reached and the solution at  $t_f$  is printed out. See Figure 1 for a plot of the solution at the time of the event. We do not provide a plot of the solution at  $t_f = 2$  since by then the layer has moved beyond the spatial domain,  $[0, 1]$ , and thus the solution is approximately equal to 1 across the entire spatial domain.



**Figure 1.** Numerical solution to **OLBE** when  $U(0.4, t) \approx 0.5$ ; we determine, using the event detection capability of BACOLIKR, that this occurs when  $t \approx 0.3$ . It can be seen from the above plot that the solution at  $x = 0.4$  has a value of  $\approx 0.5$ .

#### 4.2. Solution value detection involving multiple events for a Catalytic Surface Reaction Model

The Catalytic Surface Reaction Model (**CSRM**) [32] considered here involves a PDE system of the form,

$$\begin{aligned}(u_1)_t &= -(u_1)_x + n(D_1 u_3 - A_1 u_1 \gamma) + (u_1)_{xx}/Pe_1, \\(u_2)_t &= -(u_2)_x + n(D_2 u_4 - A_2 u_2 \gamma) + (u_2)_{xx}/Pe_1, \\(u_3)_t &= A_1 u_1 \gamma - D_1 u_3 - R u_3 u_4 \gamma^2 + (u_3)_{xx}/Pe_2, \\(u_4)_t &= A_2 u_2 \gamma - D_2 u_4 - R u_3 u_4 \gamma^2 + (u_4)_{xx}/Pe_2,\end{aligned}\tag{4.3}$$

where  $\gamma = 1 - u_3 - u_4$ , and  $n, r, Pe_1, Pe_2, D_1, D_2, R, A_1$ , and  $A_2$  are problem dependent parameters. The initial conditions at  $t_0 = 0$  ( $0 \leq x \leq 1$ ) are,

$$u_1(x, 0) = 2 - r, \quad u_2(x, 0) = r, \quad u_3(x, 0) = u_4(x, 0) = 0,$$

and the boundary conditions at  $x = 0$  and  $x = 1$  ( $t > 0$ ) are,

$$\begin{aligned}(u_1)_x(0, t) &= -Pe_1(2 - r - u_1(0, t)), \quad (u_2)_x(0, t) = -Pe_1(r - u_2(0, t)), \\(u_3)_x(0, t) &= (u_4)_x(0, t) = 0,\end{aligned}$$

$$(u_1)_x(1, t) = (u_2)_x(1, t) = (u_3)_x(1, t) = (u_4)_x(1, t) = 0.$$

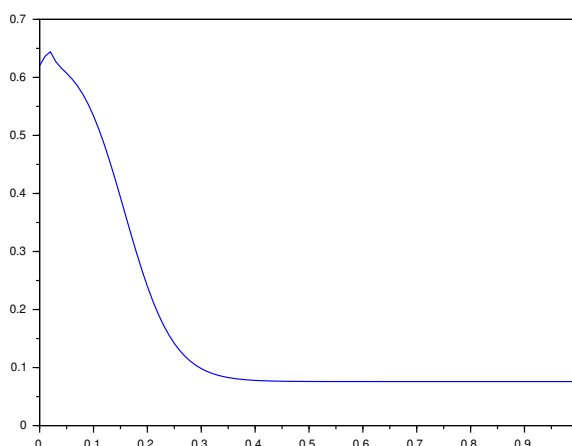
See Figures 12-15 of [30] for plots of the solution components for  $Pe_1 = Pe_2 = 100$ ,  $D_1 = 1.5$ ,  $D_2 = 1.2$ ,  $R = 1000$ ,  $r = 0.96$ ,  $n = 1$ , and  $A_1 = A_2 = 30$ . Here we choose the problem dependent parameters as above except, in order to make the problem more challenging, we choose  $Pe_1 = Pe_2 = 10000$ . We choose a tolerance of  $10^{-5}$ .

We define two simple events. We wish to find the time at which  $U_3(0.2, t) = 0.24$  and the time at which  $U_4(0.2, t) = 0.24$ . Thus there are two roots and the gstop vector function is,

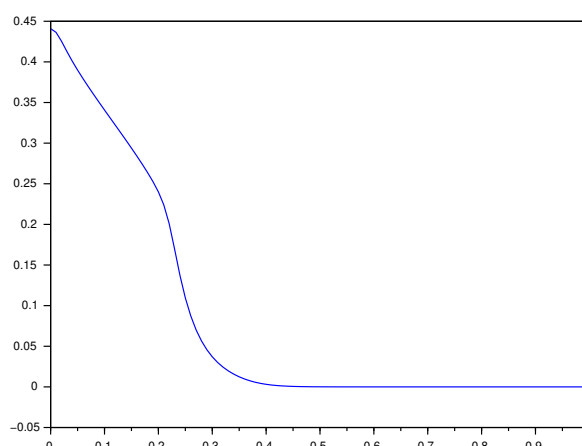
$$\begin{pmatrix} U_3(0.2, t) - 0.24 \\ U_4(0.2, t) - 0.24 \end{pmatrix}.$$

Inside RT, we make one call to the VALUES routine with  $x = 0.2$ . The VALUES routine returns a vector of the four solution component values at  $x = 0.2$  and at the current time. The differences between the third and fourth components of the returned solution approximation and the target value, 0.24, are then assigned to the first and second components of RVAL. DASKR checks, on every accepted step, to see if either component of RVAL changes sign, thereby monitoring both of the events over the duration of the computation.

The main program first initializes the computation with calls to the BACOLI95\_INIT and SETSOL and then calls BACOLI95. The solver returns with an indication of which root has been found. For this example, the first root that is found corresponds to the RVAL component equal to  $U_3(0.2, t) - 0.24$ , and this root is found at  $t \approx 1.2$ . The main program writes out the solution for this event time. See Figure 2 for a plot of the solution at the time of this first event. The main program then calls BACOLI95 again in order to continue the computation. Since the problem has not changed, a warm start can be employed. BACOLI95 returns a second time with an indication of which root has been found. In this case the root of the gstop function component,  $U_4(0.2, t) - 0.24$ , is found at  $t \approx 1.9$ . The solution is printed out and the computation is terminated. See Figure 3 for a plot of the solution at the time of the second event.



**Figure 2.** Third component of the numerical solution to **CSRM** when  $U_3(0.2, t) \approx 0.24$ ; this occurs when  $t \approx 1.2$ . It can be seen from the above plot that this solution component, at  $x = 0.2$ , has a value of  $\approx 0.24$ .



**Figure 3.** Fourth component of the numerical solution to **CSRM** when  $U_4(0.2, t) \approx 0.24$ ; this occurs when  $t \approx 1.9$ . It can be seen from the above plot that this solution component, at  $x = 0.2$ , has a value of  $\approx 0.24$ .

#### 4.3. Layer merge detection for the Two Layer Burgers Equation

The Two Layer Burgers Equation (**TLBE**) is based on the PDE, (4.1), but with boundary conditions at  $x = 0$  and  $x = 1$  ( $t > 0$ ) and an initial condition at  $t_0 = 0$  ( $0 \leq x \leq 1$ ) taken from the exact solution,

$$u(x, t) = \frac{0.1e^{-A} + 0.5e^{-B} + e^{-C}}{e^{-A} + e^{-B} + e^{-C}},$$

where,

$$A = \frac{0.05}{\epsilon}(x - 0.5 + 4.95t), \quad B = \frac{0.25}{\epsilon}(x - 0.5 + 0.75t), \quad C = \frac{0.5}{\epsilon}(x - 0.375),$$

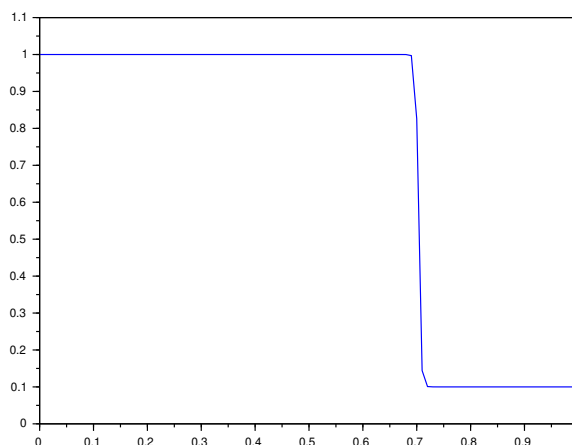
where  $\epsilon$  is a problem-dependent parameter. For this example, we choose  $\epsilon = 10^{-3}$ . When  $t_0 = 0$ , the solution has two sharp layers, one at  $x = 0.25$ , where the solution transitions rapidly from approximately 1 to approximately 0.5, and one at  $x = 0.5$ , where the solution transitions rapidly from approximately 0.5 to approximately 0.1. As  $t$  increases, these layers move to the right and eventually merge, forming a single layer. See Figure 1 of [30] for a plot of the solution to this problem.

Our goal in this example is to determine the time at which the two layers merge. Since the left layer corresponds to a sharp transition in the solution value from approximately 1 to approximately 0.5, we will define the location of this layer to be at the point,  $x_L$ , where the solution has the value 0.75, halfway through the transition in the solution values that characterizes this layer. Similarly, since the right layer corresponds to a transition in the solution value from approximately 0.5 to approximately 0.1, we will define the location of this layer to be at the point,  $x_R$ , where the solution has the value 0.3, halfway through the transition in the solution values that characterizes this layer. Even after the layers merge,  $x_L$  will always be slightly to the left of  $x_R$ . An examination of these values after the layers have merged shows that, for the above choice of  $\epsilon$ ,  $x_L$  and  $x_R$  remain a distance of about  $0.0049 \approx 5\epsilon$  apart for the rest of the computation. We will therefore define the layers to have merged when  $x_L$  and  $x_R$  are

within a distance of  $5\epsilon$ . Thus there will be one root and the gstop function will be  $|x_L - x_R| - 5\epsilon$ . We will solve this problem using a tolerance of  $10^{-6}$ .

Within the RT routine, we begin by evaluating the approximate solution at the endpoints of the spatial domain. These values are then used as input to a bisection search that is used to determine the point,  $x_L$ , in the spatial domain where  $U(x_L, t) = 0.75$ , the location of the left layer for the current time,  $t$ . Similarly, we use a second bisection search to determine the point,  $x_R$ , in the spatial domain where  $U(x_R, t) = 0.3$ , the location of the right layer for the current time,  $t$ . And then, as indicated above,  $|x_L - x_R| - 5\epsilon$  is assigned to the RVAL output argument of RT in order to define the event for this example.

The corresponding main program first initializes the computation with a call to the BACOLI95\_INIT routine and then makes a call to the SETSOL routine. This is followed by a call to the BACOLI95 solver. The solver returns when the root has been found. For this example it turns out that the two layers are determined to have merged for  $t \approx 0.62$ . The main program writes out the solution at this time and then terminates. See Figure 4 for a plot of the solution at the time of the event.



**Figure 4.** Numerical solution,  $U(x, t)$ , to TLBE when the two layers merge; this occurs when  $t \approx 0.62$ .

#### 4.4. Critical tumor mass detection in a Brain Tumor Model

The Brain Tumor Model (BTM) models the growth of a brain tumor within a region of the brain that includes three consecutive regions involving gray-white-gray matter [3] with a corresponding discontinuous diffusion coefficient. We consider a modification of this model in which the discontinuous diffusion coefficient is replaced with a continuous diffusion coefficient that has sharp layer regions corresponding to the transitions between the brain matter regions.

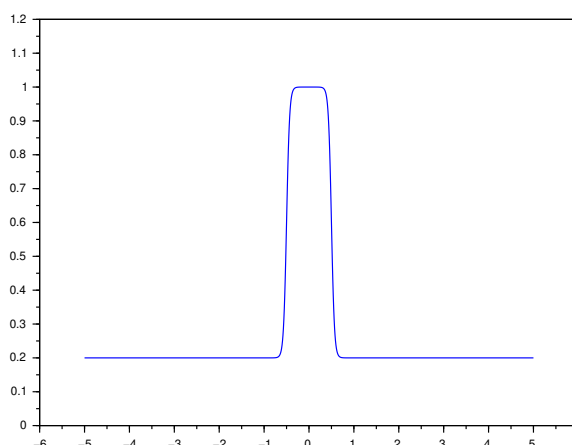
The PDE for this problem is,

$$u_t(x, t) = (D(x)u_x(x, t))_x = D_x(x)u_x(x, t) + D(x)u_{xx}(x, t), \quad (4.4)$$

where,

$$D(x) = \left( \left( \frac{1}{e^{-l(x-w_1)} + 1} \right) + \left( \frac{1}{e^{l(x-w_2)} + 1} \right) - 1 \right) (1 - \gamma) + \gamma, \quad (4.5)$$

approximates a step function whose value to the left and right of the region  $[w_1, w_2]$  (a subregion of the spatial domain,  $[a, b]$ ) is  $\gamma$ , and whose value within the region  $[w_1, w_2]$  is 1. The parameter,  $\mathfrak{l}$ , controls the sharpness of the transition layers between  $[w_1, w_2]$  and the rest of the spatial domain. We choose  $w_1 = -0.5$ ,  $w_2 = 0.5$ ,  $\gamma = 0.2$ , and  $\mathfrak{l} = 30$ . In Figure 5, we provide a plot of  $D(x)$  for the above choice of parameters on  $[-5, 5]$ .



**Figure 5.** Diffusion coefficient,  $D(x)$  (4.5), with  $w_1 = -0.5$ ,  $w_2 = 0.5$ ,  $\gamma = 0.2$ , and  $\mathfrak{l} = 30$ .

The boundary conditions are,

$$u_x(a, t) = 0, \quad u_x(b, t) = 0,$$

where  $a = -5$ ,  $b = 5$ , and the initial condition (at  $t_0 = 0$ ) is,

$$u(x, 0) = \frac{e^{-\frac{(x-\xi)^2}{\eta^2}}}{\eta \sqrt{\pi}},$$

where  $\xi = -2$  and  $\eta = 0.2$ . This gives an initial solution that has a spike of height approximately 2.8 centered at  $x = -2$ . We choose a tolerance of  $10^{-6}$ .

The tumor concentration,  $c(x, t)$ , is obtained from the solution,  $u(x, t)$ , of the above PDE from the equation,

$$c(x, t) = e^t u(x, t).$$

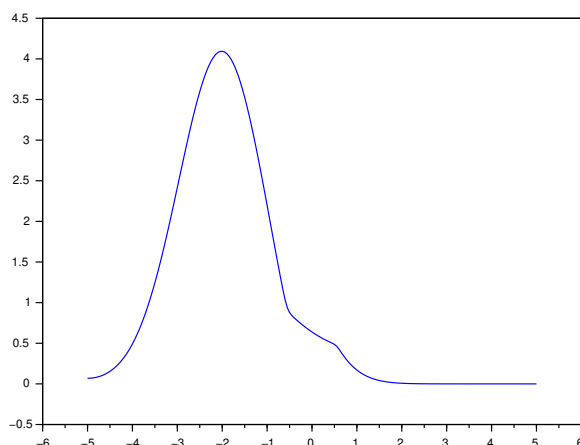
The tumor grows both in size and width across the spatial domain as time progresses, with different growth rates in the grey and white matter regions, as determined by the different values of the diffusion coefficient across the spatial domain.

For this problem we wish to find the time at which the total mass of the tumor reaches the critical value of 10. The mass of the tumor at a given point in time is obtained by integrating the tumor concentration,  $c(x, t)$ , over the spatial domain  $[a, b]$ . Thus, each time the RT routine is called, we need to compute an approximation to the integral of the approximate solution,  $U(x, t)$ , over  $[a, b]$ . To do this, we will use a Gaussian quadrature rule on each subinterval. (It may be possible to obtain a more efficient computation by using an adaptive quadrature routine to approximate the integral of  $U(x, t)$  over  $[a, b]$  to an accuracy that is consistent with the tolerance employed in the computation of  $U(x, t)$ ).

but we do not consider this here since the specific way in which the integral is computed is not central to our discussion.)

Inside the RT subroutine, we first compute the Gauss points and weights for use on each subinterval of the current spatial mesh. We then call the VALUES routine with the Gauss points as input to (simultaneously) evaluate the approximate solution at all Gauss points on all subintervals. We then multiply these values by the appropriate Gauss weights and sum over all subintervals to get an approximation to the integral of  $U(x, t)$  over  $[a, b]$ . Finally, we multiply this integral approximation by  $e^t$  to obtain an approximation to  $c(x, t)$ , the mass of the tumor over  $[a, b]$  at the current time  $t$ . The gstop function in this case is the difference between the approximate mass of the tumor and the target value of 10.

The corresponding main program, after calling BACOLI95\_INIT and SETSOL, calls BACOLI95 to determine the point in time where the approximate tumor mass reaches the critical value of 10. We find that this happens when  $t \approx 2.3$ . The main program prints out the solution and then terminates. See Figure 6 for a plot of the approximate concentration,  $c(x, t)$ , at the time of the event.



**Figure 6.** Approximate concentration,  $c(x, t)$ , for **BTM** when the mass of the tumor reaches the critical value of 10; this occurs when  $t \approx 2.3$ .

#### 4.5. Steady state detection via layer-boundary tracking for the One Layer Burgers Equation

This problem also considers the OLBE and, for this example, the event detection task is to discover when the solution reaches steady state. For a typical problem, the detection of steady state requires a more elaborate computation (see the next example) but for the OLBE, the behavior of the solution makes obtaining the answer to this question straightforward. Recall that the OLBE has a solution that, initially, has a sharp layer region at  $x \approx 0.25$ , and, as  $t$  goes from 0 to 1, the layer moves to the right and is located at  $x \approx 0.75$  for  $t = 1$ . At any point in time, to the left of the layer, the solution value is approximately 1, while to the right of the layer, the solution is approximately 0.

At some point in time after  $t = 1$ , the layer moves past the right boundary, and from that point in time onward, the solution is approximately equal to 1 across the entire spatial domain  $[0, 1]$ , i.e., the solution has reached a steady state. Thus the determination of the time at which the solution reaches



steady state is, for this problem, equal to the time at which the moving layer passes beyond the right boundary. This will be the time at which the solution at the right boundary is approximately equal to 1, since, prior to this time, the solution at the right boundary will always be less than 1. Thus, for this problem, the determination of the time when the solution reaches steady state simplifies to determining when the approximate solution,  $U(x, t)$ , satisfies the condition  $U(1, t) \approx 1$ , and we can use event detection to determine when this happens. For this example, we choose a tolerance of  $10^{-6}$ .

The RT subroutine and the main program are quite similar to those for the first example involving the OLBE. The computation terminates after the event has been found. For this example, BACOLIKR determines that steady state has been reached, i.e.,  $U(1, t) \approx 1$ , for  $t \approx 1.6$ . (We do not provide a figure showing the solution at this point in time since it is approximately equal to 1 across the entire spatial domain.)

#### 4.6. Steady state detection for the Gierer-Meinhardt Model

The Gierer-Meinhardt Model (**GMM**) [15] is an activator-inhibitor system associated with the modeling pattern formation in biological systems. A key phenomenon in such systems is that for certain parameter choices it is possible to observe spontaneous pattern formation from an initially (almost) homogeneous initial state. The form of the GMM we will consider here involves a PDE system of the form,

$$a_t(x, t) = \epsilon^2 a_{xx}(x, t) - a(x, t) + a(x, t)^2/h(x, t),$$

$$\tau h_t(x, t) = D h_{xx}(x, t) - \mu h(x, t) + a(x, t)^2,$$

where  $\epsilon$ ,  $\tau$ ,  $D$ , and  $\mu$  are problem dependent parameters. For this investigation, we choose  $\epsilon = 0.1$ ,  $\tau = 1$ ,  $D = 0.1$ , and  $\mu = 1$ . The boundary conditions are,

$$a_x(a, t) = h_x(a, t) = 0, \quad a_x(b, t) = h_x(b, t) = 0,$$

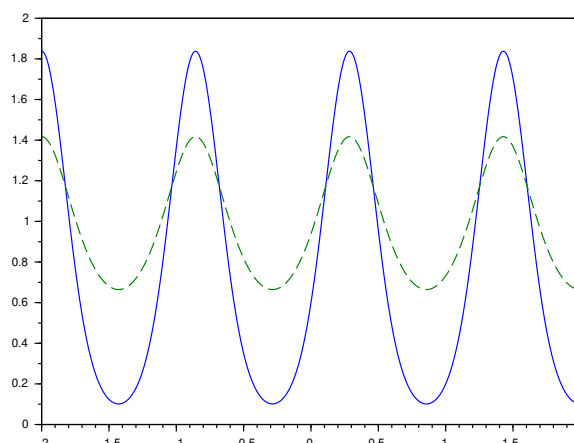
where the spatial domain is  $[a, b]$ . For this example, we choose  $a = -2, b = 2$ . The initial solutions are,

$$a(x, 0) = h(x, 0) = 0.1564 + 0.01 \sin(10\pi x), \quad h(x, 0) = 0.1564 + 0.01 \sin(10\pi x),$$

which correspond to setting each solution component to a constant plus a small amount of high frequency noise. Over time, the first component develops spikes, reaching a maximum of approximately 1.8, with troughs in between of magnitude approximately 0.1. The second solution component spikes of magnitude approximately 1.4 and troughs of magnitude approximately 0.7. See Figure 7.

Our goal in this example is to determine the time at which the first solution component reaches a steady state. We choose a tolerance of  $10^{-6}$ . We will define the time derivative of the first solution component to be sufficiently small when the integral of the absolute value of the time derivative of this solution component over  $[a, b]$  is approximately equal to the user tolerance. The `gstop` function will therefore be the difference between this integral approximation and the tolerance.

Within the RT routine, we first compute the Gauss points and weights on each subinterval of the spatial mesh. We then call the `VALUES` routine with the Gauss points as input to evaluate the time derivative of the approximate solution at the Gauss points on all subintervals. *In order to obtain values of the time derivative of the approximate solution, we pass the vector of time derivatives of the B-spline*



**Figure 7.** Numerical solution components,  $a(x, t)$  (solid curve) and  $h(t)$  (dashed curve), to **GMM** when the first solution component reaches steady state; this occurs when  $t \approx 746.8$ .

coefficients to the *VALUES* routine. We then multiply the absolute values of these time derivatives by the appropriate Gauss weights and sum over all subintervals to get an approximation to the integral of  $|U_t(x, t)|$  over  $[0, 1]$ , for the current time. The *gstop* function in this case is the difference between this integral approximation and the tolerance to which the numerical solution is computed.

The main program, after calls *BACOLI95\_INIT* and *SETSOL*, calls *BACOLI95* to determine the point in time where the first component of the approximate solution reaches steady state. We find that the solution reaches steady state when  $t \approx 746.8$ . The main program then prints out the solution and terminates. See Figure 7 for a plot of the two solution components at steady state.

#### 4.7. Event detection in the heat equation with changes in the boundary conditions at unknown times

(This problem was communicated privately to us by Sandeep Chatterjee and was the primary motivation for the development of *BACOLIKR*.) For this problem, the PDE is the heat equation,

$$u_t(x, t) = \kappa u_{xx}(x, t),$$

where  $\kappa$  is a problem dependent parameter. For this example, we set  $\kappa = 0.1$ . Initially, the boundary conditions are

$$u_x(a, t) = \alpha, \quad u_x(b, t) = \alpha.$$

They stay in this form until some unknown point in time at which either  $u(a, t)$  or  $u(b, t)$  reaches a critical value,  $u_{crit}$ . This is the first event time. If  $u(a, t) = u_{crit}$ , the left boundary condition suddenly changes to  $u_x(a, t) = -\gamma u(a, t)$ . If  $u(b, t) = u_{crit}$ , the right boundary condition suddenly changes to  $u_x(b, t) = -\gamma u(b, t)$ . As time proceeds, the solution at the other boundary also reaches the critical value,  $u_{crit}$ . This is the second event time. If the second event happens at  $a$  then the left boundary condition suddenly changes to  $u_x(a, t) = -\gamma u(a, t)$ . Otherwise, the right boundary condition suddenly changes to  $u_x(b, t) = -\gamma u(b, t)$ . Both boundary conditions remain in this form until  $t_f$ , the final time.

The change in each boundary condition will be discontinuous unless the boundary continuity condition,

$$\alpha = -\gamma u_{crit}, \quad (4.6)$$

is satisfied.

For this example, we choose the initial solution at  $t_0 = 0$  to be,

$$u(x, 0) = 10e^{-100(x-0.75)^2} + \alpha x.$$

This initial solution has a peak equal to  $10+0.75\alpha$  at  $x = 0.75$  and the solution on either side of the peak rapidly decays to  $\alpha x$ . This makes the spatial derivative of the initial solution (essentially) consistent with the boundary conditions. We choose the tolerance to be  $10^{-6}$ . The terminal time is  $t_f = 5$ .

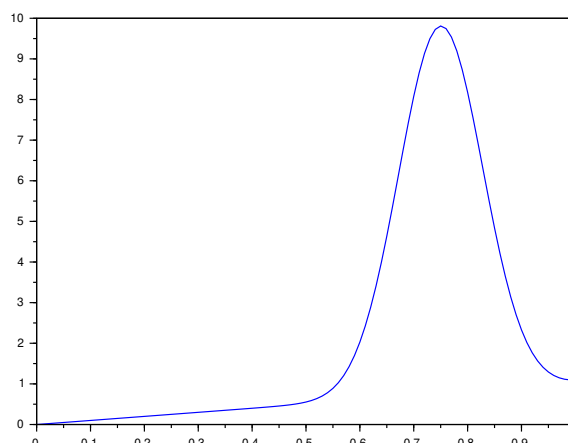
#### 4.7.1. Continuous boundary conditions

In this subsection, we consider a version of this problem in which the parameters,  $\alpha$ ,  $\gamma$ , and  $u_{crit}$  are chosen to satisfy the above boundary continuity condition (4.6). We will refer to this instance of this problem as the **Heat equation with Boundary Conditions: Continuous case (HBCC)**. We will choose  $u_{crit}$  to be 1.1. Then the boundary continuity condition (4.6) will force  $\alpha = -\gamma \times 1.1$ . For this example, we choose  $\alpha = 1.0$  and then  $\gamma$  becomes  $-1.0/1.1 \approx -0.90909$ .

Since the boundary conditions correspond to the algebraic conditions in the DAE system that is solved by DASKR, the above choice of parameters also means that the algebraic equations change in a continuous way at the time of each event, which in turn means that the B-spline coefficients that appear in the algebraic equations change in a continuous way. *However, an examination of the time derivatives of the algebraic equations shows that they do not change continuously at the time of an event. (To see this, note that the time derivative of the algebraic constraint corresponding to the left boundary condition prior to the event is simply  $u_{xt}(a, t)$  but after the event, it is  $u_{xt}(a, t) + \gamma u_t(a, t)$ . A similar situation holds for the right boundary condition.) Consequently, the time derivative of at least one of B-spline coefficients that appears in each of the boundary conditions/algebraic constraints must change discontinuously at the time of the event, in order to satisfy the new boundary condition that is imposed immediately after the event.* Due to this discontinuity in the time derivative of the B-spline coefficient(s), a cold start of DASKR should be performed after each event. See below for further discussion on this point.

The organization of the RT routine in this case is straightforward. We call the VALUES routine to get solution values at the endpoints of the spatial domain and then the gstop vector function is  $[U(0, t) - u_{crit}, U(1, t) - u_{crit}]^T$ . In the main program, we first call BACOLI95\_INIT with NRT = 2 since there are two events to be tracked. We also call SETSOL as usual. Then BACOLI95 is called at  $t = 0$  with the boundary routines BNDXA1 and BNDXB1, corresponding to the  $u_x(a, t) = \alpha$  and  $u_x(b, t) = \alpha$  boundary conditions. The main program is setup to handle whichever boundary event happens first.

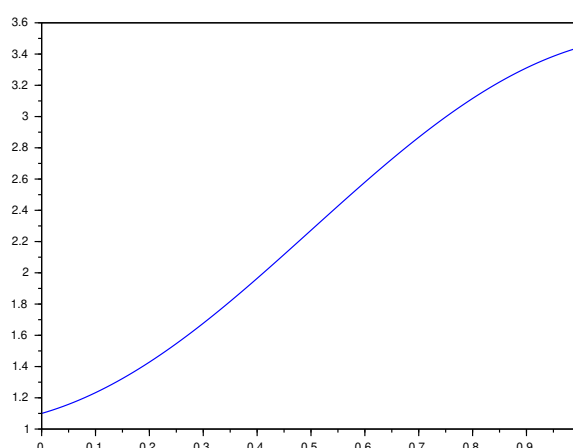
At the end of this first return from BACOLI95, the solver indicates which of the two event conditions has been satisfied and the main program writes out the time and corresponding solution values across the spatial domain. Since one of the boundary conditions will be changed, BACOLI95 is then called *with a cold start*. This call uses the BNDXA2 routine instead of the BNDXA1 routine, if the left boundary event is detected, or the BNDXB2 routine instead of the BNDXB1 routine, if the right boundary event is detected. The BNDXA2 routine imposes the  $u_x(a, t) = -\gamma u(a, t)$  boundary condition while the BNDXB2 routine imposes the  $u_x(b, t) = -\gamma u(b, t)$  boundary condition.



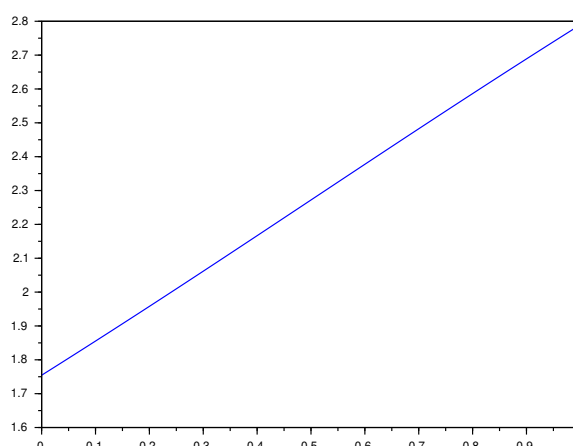
**Figure 8.** Numerical solution,  $U(x, t)$ , to **HBCC**, when the solution at the right boundary reaches the critical value,  $u_{crit} = 1.1$ ; this occurs when  $t \approx 5.5 \times 10^{-3}$ .

BACOLI95 returns when the second event is found, again indicating which event has been found, and the main program writes out the time and corresponding solution values across the spatial domain. Since the other boundary condition will now be changed, BACOLI95 is called again *with a cold start* and with input of either BNDXA2 or BNDXB2, depending on which event has been found. For this last call to BACOLIKR, both of the new boundary conditions,  $u_x(a, t) = -\gamma u(a, t)$  and  $u_x(b, t) = -\gamma u(b, t)$ , are imposed. BACOLI95 returns at  $t_f$  and the main program outputs the final solution.

For this example, with the parameter values chosen as indicated above, we find that the first event occurs at the right boundary when  $t \approx 5.5 \times 10^{-3}$ . The solution at this point in time is shown in Figure 8. We find that the second event occurs at the left boundary when  $t \approx 1.3$ . The solution at this point in time is shown in Figure 9. The integration terminates at  $t_f = 5$ . The solution at this point in time is shown in Figure 10.



**Figure 9.** Numerical solution,  $U(x, t)$ , to **HBCC**, when the solution at the left boundary reaches the critical value,  $u_{crit} = 1.1$ ; this occurs when  $t \approx 1.3$ .



**Figure 10.** Numerical solution,  $U(x, t)$ , to **HBCC**, when  $t = 5$ .

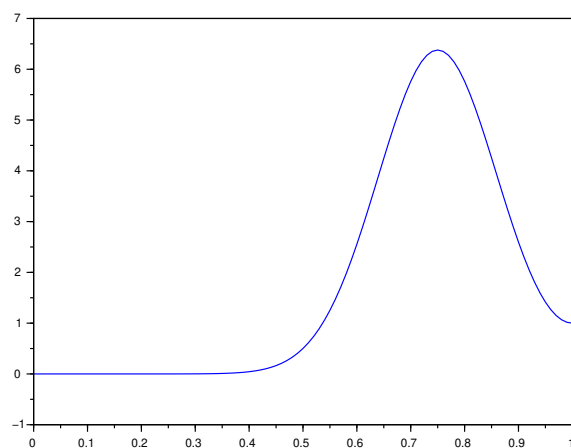
For the computation described above, where we determine the locations of the events and then employ *cold starts* to restart DASKR after each event, we find that the time integration requires a total 316 accepted time steps. When we repeat the above computation, except that we perform a *warm start* after each event, we find that the time integration requires 498 accepted time steps. As expected, due to the discontinuity in the time derivative of several of the B-spline coefficients immediately after each event, DASKR has substantial difficulty in stepping past the discontinuity, *unless a cold start is employed after the discontinuity has been located*.

#### 4.7.2. Discontinuous boundary conditions

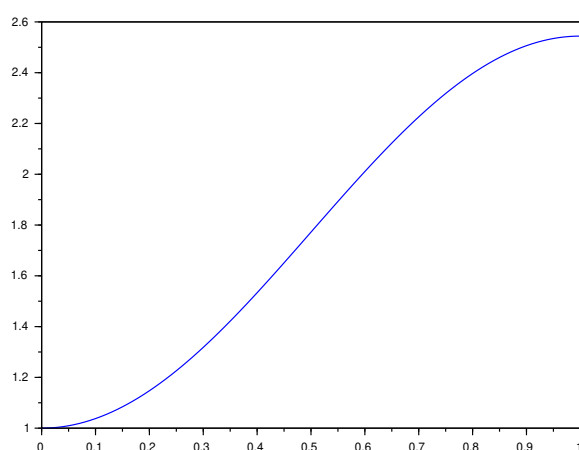
In this subsection we choose the parameters so that there are discontinuous changes in the boundary conditions after each event. We will refer to this instance of this problem as the **Heat equation with Boundary Conditions: Discontinuous case (HBCCD)**. We choose  $\alpha = 0$ ,  $u_{crit} = 1$ , and  $\gamma = 0.5$ . This means that there will be discontinuities, of magnitude 0.5, imposed on the first spatial derivative of the solution at the boundaries. The boundary continuity condition (4.6) is not satisfied.

As discussed earlier, the boundary conditions correspond to the algebraic equations in the DAE system that is given to DASKR. In this instance of the problem, since there is a discontinuity in one of the boundary conditions after each event, there is a discontinuity in the corresponding algebraic equation, which in turn forces a discontinuity in several of the B-spline coefficients and their time derivatives at the time of each event. We expect that these discontinuities will lead to substantial inefficiencies in the time integration as DASKR attempts to step past them, unless a cold start is performed after each event.

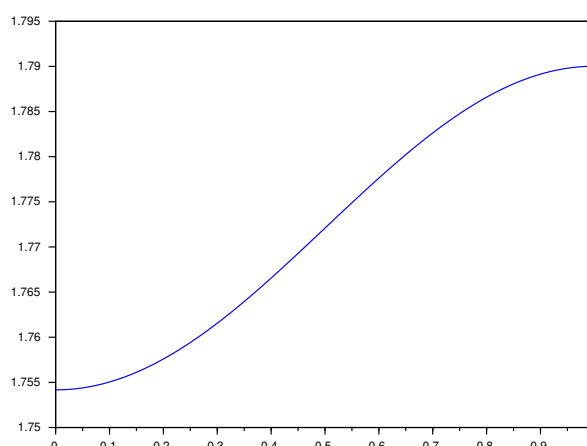
For this case we find that the first event occurs at the right boundary for  $t \approx 3.6 \times 10^{-2}$ . The solution at this point in time is shown in Figure 11. BACOLI95 is then restarted with a cold start and the new right boundary condition is imposed. We find that the second event occurs at the left boundary for  $t \approx 1.2$ . The solution at this point in time is shown in Figure 12. We then restart BACOLI95 with a cold start and impose the new left boundary condition. The solver then integrates to  $t_f = 5$ . The solution at this point in time is shown in Figure 13.



**Figure 11.** Numerical solution,  $U(x, t)$ , to **HBCD**, when the solution at the right boundary reaches the critical value,  $u_{crit} = 1$ ; this occurs when  $t \approx 3.6 \times 10^{-2}$ .



**Figure 12.** Numerical solution,  $U(x, t)$ , to **HBCD**, when the solution at the left boundary reaches the critical value,  $u_{crit} = 1$ ; this occurs when  $t \approx 1.2$ .



**Figure 13.** Numerical solution,  $U(x, t)$ , to **HBCD**, when  $t = 5$ .

For the computation described above, DASKR requires 291 accepted time steps to perform the time integration. When we repeat the above computation, except that we perform a warm start after each event, DASKR requires 568 accepted time steps to perform the time integration. As expected, due to the discontinuity after each event, *DASKR has substantial difficulty in stepping past the two event times when warm starts are employed.*

## 5. Summary, conclusions, and future work

This paper introduced, BACOLIKR, a new error-control PDE solver that features time-dependent event detection. To our knowledge, this is the only error-control PDE solver with this capability. We described the modifications to the previously developed error control PDE solver, BACOLI, and the time integrator, DASKR, that were required in order to obtain BACOLIKR. We also provided a number of examples to demonstrate how a variety of event detection problems can be handled by the new solver. We found that, with relatively straightforward modifications of the main program and the rootfinding routine that characterizes a given event or events, it was possible to use BACOLIKR to solve a collection of PDE-based event detection problems. We observed that when a problem is altered after an event and there is an expectation that the computation should continue past the event, it is important that BACOLIKR be restarted with a cold start due to the well-known difficulties that arise for error control time integration solvers in the presence of discontinuities in the DAE system obtained from the discretization of the PDE system.

Regarding future work, as mentioned earlier, since DASKR also has a feature that allows it to efficiently treat large DAE systems using a Krylov method, it may be worthwhile to modify BACOLIKR to take advantage of this capability in order to improve the efficiency of the solver for problems which lead to large DAE systems. Another direction for future work could involve the development of an updated version of the EBACOLI solver, mentioned earlier in this paper, so that an event detection capability could be available within that solver.

## Author contributions

Paul Muir: Conceptualization, funding acquisition, investigation, methodology, project administration, supervision, validation, writing - original draft, writing - review and editing; Connor Tannahill: Investigation, software; Jack Pew: Investigation, software.

## Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

## Acknowledgments

The authors wish to thank the referees for their helpful suggestions. This work was supported by the Natural Sciences and Engineering Research Council of Canada and Saint Mary's University.

## Conflict of interest

The authors declare no conflict of interest in this paper.

## References

1. T. Arsenault, T. Smith, P. H. Muir, Superconvergent interpolants for efficient spatial error estimation in 1D PDE collocation solvers, *Can. Appl. Math. Q.*, **17** (2009), 409–431.
2. T. Arsenault, T. Smith, P. H. Muir, J. Pew, Asymptotically correct interpolation-based spatial error estimation for 1D PDE solvers, *Can. Appl. Math. Q.*, 20:307–328, 2012.
3. I. E. Athanasakis, M. G. Papadomanolaki, E. P. Papadopoulou, Y. G. Saridakis, Discontinuous Hermite collocation and diagonally implicit RK3 for a brain tumour invasion model, *Proceedings of the World Congress on Engineering*, **1** (2013).
4. T. R. Boehme, C. H. Onder, L. Guzzella, Code-generator-based software package for defining and solving one-dimensional, dynamic, catalytic reactor models, *Comput. Chem. Eng.*, **32** (2008), 2445–2454. <https://doi.org/10.1016/j.compchemeng.2008.01.003>
5. Z. Bonaventura, D. Trunec, M. Meško, P. Vašina, V. Kudrle, Theoretical study of pulsed microwave discharge in nitrogen, *Plasma Sci. Technol.*, **14** (2005), 751. <https://doi.org/10.1088/0963-0252/14/4/014>
6. P. N. Brown, A. C. Hindmarsh, L. R. Petzold, Using Krylov methods in the solution of large-scale differential-algebraic systems, *SIAM J. Sci. Comp.*, **15** (1994), 1467–1488. <https://doi.org/10.1137/0915088>
7. P. N. Brown, A. C. Hindmarsh, L. R. Petzold, Consistent initial condition calculation for differential-algebraic systems, *SIAM J. Sci. Comp.*, **19** (1998), 1495–1512. <https://doi.org/10.1137/S1064827595289996>
8. J. H. Cerutti, S. V. Parter, Collocation methods for parabolic partial differential equations in one space dimension, *Numer. Math.*, **26** (1976), 227–254. <https://doi.org/10.1007/BF01395944>



9. C. de Boor, *A Practical Guide to Splines*, volume 27 of *Applied Mathematical Sciences*, Springer-Verlag, New York, revised edition, 2001.
10. J. C. Díaz, G. Fairweather, P. Keast, Algorithm 603. COLROW and ARCECO: FORTRAN packages for solving certain almost block diagonal linear systems by modified alternate row and column elimination, *ACM Trans. Math. Software*, **9** (1983), 376–380. <https://doi.org/10.1145/356044.356054>
11. J. Douglas Jr., T. Dupont, *Collocation Methods for Parabolic Equations in a Single Space Variable*, Lecture Notes in Mathematics, Vol. 385, Springer-Verlag, Berlin, 1974. <https://doi.org/10.1007/BFb0057337>
12. W. H. Enright, K. R. Jackson, S. P. Nørsett, P. G. Thomsen, Effective solution of discontinuous IVPs using a Runge-Kutta formula pair with interpolants, *Appl. Math. Comp.*, **27** (1988), 313–355. [https://doi.org/10.1016/0096-3003\(88\)90030-6](https://doi.org/10.1016/0096-3003(88)90030-6)
13. W. F. Finden, An error term and uniqueness for Hermite-Birkhoff interpolation involving only function values and/or first derivative values, *J. Comput. Appl. Math.*, **212** (2008), 1–15. <https://doi.org/10.1016/j.cam.2006.11.022>
14. C. W. Gear, O. Osterby, Solving ordinary differential equations with discontinuities, *ACM Trans. Math. Softw.*, **10** (1984), 23–44. <https://doi.org/10.1145/356068.356071>
15. A. Gierer, H. Meinhardt, A theory of biological pattern formation, *Kybernetik*, **12** (1972), 30–39. <https://doi.org/10.1007/BF00289234>
16. K. R. Green, R. J. Spiteri, Extended BACOLI: solving one-dimensional multiscale parabolic PDE systems with error control, *ACM Trans. Math. Softw.*, **45** (2019), 1–19. <https://doi.org/10.1145/3301320>
17. E. Hairer, S. P. Nørsett, G. Wanner, *Solving Ordinary Differential Equations. I*, volume 8 of *Springer Series in Computational Mathematics*, Springer-Verlag, Berlin, second edition, 1993.
18. K. L. Hiebert, L. F. Shampine, Implicitly defined output points for solutions of ODEs, *Sandia Report SAND80-0180*, 1980.
19. P. Keast, LAMPK: a Fortran package for solving certain almost block diagonal matrices, *Unpublished software*, 1982.
20. G. Mao, L. R. Petzold, Efficient integration over discontinuities for differential-algebraic systems, *Comput. Math. Applic.*, **43** (2002), 65–79. [https://doi.org/10.1016/S0898-1221\(01\)00272-3](https://doi.org/10.1016/S0898-1221(01)00272-3)
21. L. R. Petzold, *A Description of DASSL: A Differential/Algebraic System Solver*, Sandia Labs, Livermore, CA, 1982.
22. J. Pew, Z. Li, P. H. Muir, Algorithm 962: BACOLI: B-spline adaptive collocation software for PDEs with interpolation-based spatial error control, *ACM Trans. Math. Softw.*, **42** (2016), 1–17. <https://doi.org/10.1145/2818312>
23. J. Pew, Z. Li, C. Tannahill, P. H. Muir, G. Fairweather, Performance analysis of error-control B-spline Gaussian collocation software for PDEs, *Comput. Math. Appl.*, **77** (2019), 1888–1901. <https://doi.org/10.1016/j.camwa.2018.11.025>

24. J. Pew, C. Tannahill, P. H. Muir, A numerical study of an error control B-spline Gaussian collocation PDE solver with event detection, *Saint Mary's University, Mathematics and Computing Science Technical Report Series, Technical Report 2024\_001*, [http://cs.smu.ca/tech\\_reports](http://cs.smu.ca/tech_reports), 2024.
25. L. F. Shampine, S. Thompson, Event location for ordinary differential equations, *Comput. Math. Appl.*, **39** (2000), 43–54. [https://doi.org/10.1016/S0898-1221\(00\)00045-6](https://doi.org/10.1016/S0898-1221(00)00045-6)
26. A. Sommer, H. G. Bock, E. Kostina, IFDIFF - a matlab toolkit for ODEs with state-dependent switches, 2024. Available from: <https://andreassommer.github.io/ifdiff/>.
27. D. Trunec, Z. Bonaventura, D. Nečas, Solution of time-dependent Boltzmann equation for electrons in non-thermal plasma, *J. Phys. D Appl. Phys.*, **39** (2006), 2544. <https://doi.org/10.1088/0022-3727/39/12/012>
28. J. J. Vallino, J. A. Huber, Using maximum entropy production to describe microbial biogeochemistry over time and space in a meromictic pond, *Front. Environ. Sci.*, **6** (2018), 100. <https://doi.org/10.3389/fenvs.2018.00100>
29. R. Wang, P. Keast, P. H. Muir, BACOL: B-spline Adaptive COLlocation software for 1D parabolic PDEs, *ACM Trans. Math. Software*, **30** (2004), 454–470. <https://doi.org/10.1145/1039813.1039817>
30. R. Wang, P. Keast, P. H. Muir, A comparison of adaptive software for 1D parabolic PDEs, *J. Comput. Appl. Math.*, **169** (2004), 127–150. <https://doi.org/10.1016/j.cam.2003.12.016>
31. R. Wang, P. Keast, P. H. Muir, A high-order global spatially adaptive collocation method for 1D parabolic PDEs, *Appl. Numer. Math.*, **50** (2004), 239–260. <https://doi.org/10.1016/j.apnum.2003.12.023>
32. W. Zhang, Diffusive effects on a catalytic surface reaction: an initial boundary value problem in reaction-diffusion-convection equations, *J. Bifur. Chaos*, **3** (1993), 79–95. <https://doi.org/10.1142/S0218127493000052>



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)