A Numerical Study of Global Error and Defect Control Schemes for BVODEs *

Jason J. Boisvert †, Paul H. Muir ‡,
Raymond J. Spiteri §

Abstract

Boundary value ordinary differential equations (BVODEs) are systems of ODEs with boundary conditions imposed at two or more distinct points. The global error (GE) of a numerical solution to a BVODE is the difference between that numerical solution and the exact solution. The defect is the amount by which the numerical solution fails to satisfy the ODEs and boundary conditions. The BVODE solver, `BVP_SOLVER`, computes a numerical solution whose (estimated) defect satisfies a given user tolerance but it can also provide an a posteriori estimate of the GE using Richardson extrapolation (RE). Using a modified version of `BVP_SOLVER`, we present, in this report, numerical experiments comparing four strategies for a posteriori GE estimation of a defect controlled numerical solution, based on (i) RE, (ii) the direct use of a higher order (HO) discretization formula (a mono-implicit Runge–Kutta (MIRK) formula), (iii) the use of a higher order discretization formula (a MIRK formula) within a deferred correction (DC) framework, and (iv) the product of the defect estimate and an estimate of the BVODE conditioning constant (CO). We also present numerical experiments investigating a (further) modified version of `BVP_SOLVER` that provides options for (i) defect control (DefC), (ii) GE control (GEC), and combinations thereof: (iii) a sequential combination control (SCC) in which we first compute a defect controlled solution and then, using this solution, continue on to compute a GE controlled solution, and (iv) a parallel combination control (PCC) in which we simultaneously control estimates of the defect and the GE.

Subject Classification: 65L06, 65L10.
Keywords: Boundary value ordinary differential equations, conditioning, defect control, deferred correction, global error estimation, Richardson extrapolation, Runge–Kutta methods

*This work was supported by the Mathematics of Information Technology and Complex Systems Network and by the Natural Sciences and Engineering Research Council of Canada.

†Department of Computer Science, University of Saskatchewan, Saskatoon SK, Canada, S7N 5C9, jjb701@mail.usask.ca

‡Department of Mathematics and Computing Science, Saint Mary's University, Halifax, NS, Canada, B3H 3C3, muir@smu.ca

§Department of Computer Science, University of Saskatchewan, Saskatoon SK, Canada, S7N 5C9, spiteri@cs.usask.ca

1 Introduction

In this report, we consider software for the numerical solution of boundary value ordinary differential equations (BVODEs) having the form

$$\mathbf{y}'(x) = \mathbf{f}(x, \mathbf{y}(x)), \quad a \le x \le b, \qquad \mathbf{g}_a(\mathbf{y}(a)) = \mathbf{0}, \quad \mathbf{g}_b(\mathbf{y}(b)) = \mathbf{0}, \qquad (1)$$

where $\mathbf{y}$, $\mathbf{f}$, and $[\mathbf{g}_a^T, \mathbf{g}_b^T]^T$ are vectors of length n.

There are two common approaches to controlling solution accuracy in software for BVODEs: global error (GE) control and defect control. The GE is the difference between the numerical solution and the exact solution. The defect of a continuous numerical solution is the amount by which the solution fails to satisfy the ODEs and boundary conditions. Typically the GE or defect estimate is scaled to accommodate a blend of absolute and relative tolerances, based on the numerical solution or its derivative. This estimate is used to adapt the computation to return a solution for which the estimate is less than a user-provided tolerance. We refer to such solvers as providing GE control or defect control, respectively. Although control of the GE is often more familiar to users, control of the defect has an interesting backward error interpretation: the defect controlled numerical solution is the exact solution to a perturbation (on the order of the tolerance) of the original BVODE.

This report describes numerical experiments involving the study of defect and GE control. We employ modified versions of the defect control solver, `BVP_SOLVER` [15]. The original version of `BVP_SOLVER` returns a defect controlled numerical solution but provides an option for an a posteriori estimate of the GE of the numerical solution, based on Richardson extrapolation (RE), see, e.g., [2]. We first consider, within the `BVP_SOLVER` framework, implementations of three well-known GE estimation schemes as alternatives for the a posteriori estimation of the GE. These schemes are based on (i) the direct use of a higher order (HO) discretization formula (a mono-implicit Runge–Kutta (MIRK) formula - see, e.g. [12] and references within), (ii) the use of a higher order discretization formula (a MIRK formula) within a deferred correction (DC) framework, and (iii) the product of the defect estimate and an estimate of the BVODE conditioning constant (CO). We compare their performance with respect to accuracy and efficiency. We then consider a further modified version of `BVP_SOLVER` that has an option for estimation and *control* of the GE. This new version of `BVP_SOLVER` provides options for GE control, defect control, and combinations thereof. We provide numerical results comparing these options.

It is possible that a numerical solution with an estimated maximum defect that satisfies a user tolerance can nonetheless have a large GE. In extreme cases, a defect control solver can return a numerical solution for a problem that has no solution. The paper [14] refers to such solutions as *pseudosolutions* and provides examples where the defect control codes `bvp4c` [10] and `MIRKDC` [6] return pseudosolutions under certain conditions. It should be emphasized that such a solution is in fact an acceptable numerical solution in the following sense: the solver has returned a numerical solution whose defect satisfies the user tolerance. In such cases, the numerical solution is the exact solution to a BVODE that is reasonably close to the original one. However, if the BVODE is ill-conditioned and the tolerance is coarse, the solution of the perturbed problem may not be close to the solution of the original problem. This suggests that it can be important for a defect control solver to also provide an assessment of the GE of the defect controlled numerical solution it computes. This report will also include experiments in which we investigate the role that a GE estimate can play within a defect control code when we consider a problem that has a pseudosolution.

The report is organized as follows. Section 2 reviews the algorithms used in `BVP_SOLVER`. Section 3 describes the three alternative GE estimation techniques and their efficient implementation within `BVP_SOLVER`. This section also briefly discusses a slight modification of the RE based approach implemented in the original version of `BVP_SOLVER`. Section 4 presents numerical experiments comparing the four GE estimators with respect to accuracy and efficiency. Section 5 considers a new version of `BVP_SOLVER` that provides options for defect and GE control as well as options for combinations of defect and GE control. This section also provides results associated with the treatment of a problem having a pseudosolution. Section 6 provides our conclusions.

2 Review of `BVP_SOLVER`

`BVP_SOLVER` is capable of solving a first order system of n ODEs of the form

$$\mathbf{y}'(x) = \left(\frac{\mathbf{\Lambda}}{x-a}\right)\mathbf{y}(x) + \mathbf{f}(x, \mathbf{y}(x), \mathbf{p}), \qquad a \le x \le b,$$

subject to separated nonlinear two-point boundary conditions (BCs)

$$\mathbf{g}_a(\mathbf{y}(a), \mathbf{p}) = \mathbf{0}, \qquad \mathbf{g}_b(\mathbf{y}(b), \mathbf{p}) = \mathbf{0}.$$

Here $\mathbf{y}$ and $\mathbf{f}$ are vectors of length n and $\mathbf{p}$ is an optional vector of length n_p of unknown parameters. The vector $[\mathbf{g}_a^T, \mathbf{g}_b^T]^T$ is of length $n + n_p$. The $n \times n$ constant matrix $\mathbf{\Lambda}$ is optional. In this report, we assume the simpler form (1).

In order to solve a BVODE, `BVP_SOLVER` generates a system of nonlinear equations for which the unknowns, $\mathbf{y}_i$, are approximations to the solution values, $\mathbf{y}(x_i)$, at the points of a mesh that partitions the problem domain: $a = x_0 < x_1 < \ldots < x_N = b$. Let $h_{i+1} = x_{i+1} - x_i$, $i = 0, 1, \ldots, N-1$. On the subinterval, $[x_i, x_{i+1}]$, these nonlinear equations have the form

$$\boldsymbol{\phi}_{i+1}(\mathbf{y}_i, \mathbf{y}_{i+1}) = \mathbf{y}_{i+1} - \mathbf{y}_i - h_{i+1} \sum_{j=1}^{s} b_j \mathbf{f}(x_i + c_j h_{i+1}, \mathbf{y}_{ij}) = \mathbf{0}, \qquad (2)$$

where

$$\mathbf{y}_{ij} = (1 - v_j)\mathbf{y}_i + v_j \mathbf{y}_{i+1} + h_{i+1} \sum_{k=1}^{j-1} a_{j,k} \mathbf{f}(x_i + c_k h_{i+1}, \mathbf{y}_{ik}), \qquad (3)$$

for $j = 1, 2, \ldots, s$, are the stages of the MIRK method. The coefficients, $v_j, b_j, a_{j,k}$, $j = 1, \ldots, s$, $k = 1, 2, \ldots, j-1$, define the MIRK method, and $c_j = v_j + \sum_{k=1}^{j-1} a_{j,k}$.

Equation (2) represents n nonlinear equations involving the unknowns $\mathbf{y}_i$ and $\mathbf{y}_{i+1}$. Taking these equations for all subintervals together with the BCs gives a system of $(N+1)n$ nonlinear equations whose solution gives a discrete approximate solution at the mesh points, $\mathbf{Y} \equiv [\mathbf{y}_0^T, \mathbf{y}_1^T, \ldots, \mathbf{y}_N^T]^T$. This nonlinear system has the form

$$\mathbf{\Phi}(\mathbf{Y}) \equiv \begin{pmatrix} \mathbf{g}_a(\mathbf{y}_0) \\ \boldsymbol{\phi}_1(\mathbf{y}_0, \mathbf{y}_1) \\ \vdots \\ \boldsymbol{\phi}_N(\mathbf{y}_{N-1}, \mathbf{y}_N) \\ \mathbf{g}_b(\mathbf{y}_N) \end{pmatrix} = \mathbf{0}. \tag{4}$$

System (4) is solved using a modified Newton iteration, which requires the evaluation and factorization of the Jacobian

$$\boldsymbol{J}_{\mathbf{\Phi}}(\mathbf{Y}) \equiv \frac{\partial \mathbf{\Phi}(\mathbf{Y})}{\partial \mathbf{Y}}. \tag{5}$$

`BVP_SOLVER` solves the nonlinear system (4) via a hybrid damped Newton/fixed Jacobian iteration. When there are convergence issues, the solver re-evaluates the Jacobian and uses a damping factor to control the contribution of the Newton correction to the next iterate. Otherwise, it holds the Jacobian constant and takes full Newton steps as long as convergence is sufficiently rapid. Once the Newton iteration converges, we obtain the discrete solution, $\{\mathbf{y}_i\}_{i=0}^N$, which serves as the basis for a (vector) piecewise polynomial, $\mathbf{S}(x)$, that is based on a continuous MIRK (CMIRK) formula [11]. On the subinterval, $[x_i, x_{i+1}]$, $\mathbf{S}(x)$ takes the form

$$\mathbf{S}(x_i + \theta h_{i+1}) = \mathbf{y}_i + h_{i+1} \sum_{j=1}^{s^*} b_j(\theta) \mathbf{f}(x_i + c_j h_{i+1}, \mathbf{y}_{ij}),$$

where $0 \leq \theta \leq 1$ and $s^* \geq s$. In the above equation, each $b_j(\theta)$ is a known polynomial in θ, defined by the CMIRK method. Because $s^* \geq s$, it follows that $\mathbf{S}(x)$ may need to use extra stages; each such stage has the same general form as in (3). The piecewise polynomial, $\mathbf{S}(x)$, is a C^1-continuous approximation to the exact solution to the BVODE, $\mathbf{y}(x)$.

On each subinterval, `BVP_SOLVER` computes a scaled defect, $\boldsymbol{\delta}(x)$, of the approximate solution at several points on each subinterval. The jth component of $\boldsymbol{\delta}(x)$ is

$$\delta_j(x) = \frac{|S_j'(x) - f_j(x, \mathbf{S(x)})|}{1 + |f_j(x, \mathbf{S(x)})|}. \tag{6}$$

The maximum of these scaled defect samples is taken to be an estimate of the maximum scaled defect (MSD) for the subinterval. If the estimated MSD is greater than the user-prescribed tolerance on any subinterval, the current solution approximation is rejected and estimates of the MSD on each subinterval are then used to guide a process that attempts to construct a new mesh such that (i) the MSD estimates are approximately equidistributed over the subintervals of the new mesh, and (ii) the MSD estimate on each subinterval of the new mesh is less than the user tolerance. Achieving such a mesh typically involves changing the total number of mesh points and redistributing them over the problem domain. Once a new mesh is obtained, a new continuous solution approximation is computed and the defect sampling process is repeated. If the estimated MSD for each subinterval is less than the user tolerance, the solution is accepted.

The current version of `BVP_SOLVER` simply samples the defect at two points on each subinterval; a more robust estimate of the MSD on each subinterval can be obtained at a modest increase in cost using an approach called *asymptotically correct defect control* [7]. The approach relies on the development of a new type of interpolant for the continuous solution approximation on each subinterval (building upon the CMIRK interpolants mentioned earlier) for which the maximum defect is (asymptotically) guaranteed to occur at a known, problem independent location within each subinterval.

Although `BVP_SOLVER` does not attempt to directly control the GE, as mentioned earlier, it does provide the option for the computation of an a posteriori estimate of the GE based on RE, which we now briefly describe (see also, e.g., Section 5.5.2 of [2]). Let π be the final mesh upon which the accepted, defect controlled numerical solution is obtained. Let $\mathbf{Y}_\pi$ be the numerical solution evaluated at the points of the mesh π. Let $\mathbf{Y}_\pi^{(i,j)}$ be the jth component of the numerical solution evaluated at the ith mesh point of π. Let π_2 be a new mesh is obtained by halving each subinterval of π. In the RE scheme, a second discrete solution is computed on this new mesh, using the same MIRK scheme that was used to obtain $\mathbf{Y}_\pi$. Define $\mathbf{Y}_{\pi_2}$ to be this second solution *evaluated only at the points of* π. Let $\mathbf{Y}_{\pi_2}^{(i,j)}$ be the jth component of this second numerical solution at the ith mesh point of π. The GE estimate by RE is then given by

$$\left(\frac{2^p}{2^p-1}\right)\max_{i,j}\frac{|\mathbf{Y}_\pi^{(i,j)}-\mathbf{Y}_{\pi_2}^{(i,j)}|}{1+|\mathbf{Y}_\pi^{(i,j)}|}, \tag{7}$$

where p is the order of the MIRK method used to compute these approximate solutions.

The computation of $\mathbf{Y}_{\pi_2}$ requires the setup and solution via Newton's method of a nonlinear system similar in form to (4) but with approximately twice as many nonlinear equations and unknowns. In `BVP_SOLVER` this nonlinear system is solved using the same modified Newton iteration as described for the computation of the primary solution, using a Newton tolerance that is half the size. Because at least one Jacobian matrix must be evaluated and factored, this scheme can be quite computationally expensive. The initial estimate of the solution provided to the Newton iteration for the determination of $\mathbf{Y}_{\pi_2}$ is obtained from the evaluation of the continuous numerical solution, $\mathbf{S}(x)$, at the points of the mesh π_2.

3 Alternative GE Estimators for BVP_SOLVER

In the following we will assume that the defect controlled numerical solution has been accepted by the solver and that we wish to compute an a posteriori estimate of the GE of the defect controlled solution. Subsection 3.1 describes a GE estimation technique based on the direct use of higher order MIRK formulas. Subsection 3.2 describes an approach for GE estimation based on the use of higher order MIRK formulas within a deferred correction framework. Subsection 3.3 examines defect control from a backward error analysis viewpoint and describes a GE bound based on the norm of the defect and an estimate of a conditioning constant for the BVODE. Subsection 3.4 discusses an improved implementation of the RE algorithm described in the previous section.

Except for the approach based on the conditioning constant, the other approaches mentioned above are all examples of well known techniques for the estimation of the GE. The paper [13] considers mesh adaptation based on a number of error estimation schemes and looks at relationships between them; the focus is on collocation methods for the discretization but the authors indicate that the conclusion of the paper may be relevant to other approaches.

3.1 Direct Use of a Higher Order MIRK Formula for GE Estimation (HO)

Assuming that the primary solution is obtained using a MIRK method of order p on some final mesh, we obtain a second numerical solution of order $p+2$ on the same mesh by constructing and solving a nonlinear system of the form (4) using a MIRK method of order $p+2$. We choose a method 2 orders higher (rather than only 1) because it is important to employ symmetric Runge–Kutta methods when solving a BVODE, and such symmetric methods have only even orders; see, e.g., [12]. This computation yields only a discrete solution approximation.

BVP_SOLVER can solve BVODEs using a second, fourth, or sixth order MIRK method; see [12] for the tableaux that define these formulas and their associated interpolants. Thus for primary solutions obtained using a second or fourth order MIRK formula, there is a natural MIRK formula available for the computation of the higher order numerical solution. For the case when the primary solution is obtained by using the sixth order MIRK formula, we have added an eighth order MIRK method [8] to BVP_SOLVER. Although the formulas included in BVP_SOLVER are optimal in a certain sense (see [12]), there is no reason to expect that the eighth order formula from [8] is also optimal. Although further work could be done to develop an optimal eighth order MIRK formula, because the formula is only used on the final mesh, it is not clear that the use of an optimized eighth order formula would lead to much improvement overall.

For completeness, we provide here the tableau for the ten stage, eighth order MIRK formula from [8] that we have implemented in BVP_SOLVER:

$$
\begin{array}{c|c|cccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{4} & \frac{5}{32} & \frac{9}{64} & -\frac{3}{64} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{3}{4} & \frac{27}{32} & \frac{3}{64} & \frac{-9}{64} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & -\frac{5}{24} & \frac{5}{24} & \frac{2}{3} & -\frac{2}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{8} & \beta & a_1 & a_2 & a_3 & a_4 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{7}{8} & 1-\beta & -a_2 & -a_1 & -a_4 & -a_3 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{7-\sqrt{21}}{14} & \Theta & a_5 & a_6 & 0 & 0 & 0 & a_7 & a_8 & 0 & 0 & 0 \\
\frac{7+\sqrt{21}}{14} & 1-\Theta & -a_6 & -a_5 & 0 & 0 & 0 & -a_8 & -a_7 & 0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & \frac{29}{896} & -\frac{29}{896} & 0 & 0 & 0 & \frac{-2}{21} & \frac{2}{21} & \frac{7\sqrt{21}}{128} & -\frac{7\sqrt{21}}{128} & 0 \\
\hline
 & & \frac{1}{20} & \frac{1}{20} & 0 & 0 & 0 & 0 & 0 & \frac{49}{180} & \frac{49}{180} & \frac{16}{45}
\end{array}
$$

where β is a free parameter, and

$$
\begin{aligned}
a_1 &= \frac{757}{9216} - \frac{\beta}{18}, & a_2 &= \frac{43}{9216} - \frac{\beta}{18}, & a_3 &= \frac{235}{4608} - \frac{4\beta}{9}, \\
a_4 &= -\frac{59}{4608} - \frac{4\beta}{9}, & a_5 &= \frac{3451 + 717\sqrt{21}}{139258}, & a_6 &= \frac{-3451 + 717\sqrt{21}}{139258}, \\
a_7 &= \frac{64}{1029} + \frac{1024\sqrt{21}}{69629}, & a_8 &= \frac{-64}{1029} + \frac{1024\sqrt{21}}{69629}, & \Theta &= \frac{1}{2} + \frac{2211\sqrt{21}}{19894}.
\end{aligned}
$$

This MIRK formula contains embedded formulas of orders 2, 4, and 6, but if the embedded sixth order formula is not required, stage five can be ignored. We do not make use of the embedded formulas, and thus we implement this formula as a nine stage method. We also choose the free parameter β of this eighth order formula to be 0, which is a reasonable approximation to the optimal value $\beta \approx 0.006970$ that we have computed that minimizes the 2-norm of the principal error function (see, e.g., [12]) for this eighth order method. (The corresponding values for the norms of the ninth and tenth order principal error coefficients are 6.9950×10^{-6} and 6.9751×10^{-6}, respectively.)

Let $\mathbf{Y}_p$ be the primary solution of order p, evaluated at the mesh points of the final mesh and let $\mathbf{Y}_{p+2}$ be the solution of order $p+2$, evaluated at the same set of points. Let $\mathbf{Y}_p^{(i,j)}$ be the jth component of $\mathbf{Y}_p$ at the ith mesh point and let $\mathbf{Y}_{p+2}^{(i,j)}$ be the corresponding component of $\mathbf{Y}_{p+2}$. Then the estimate of the GE for $\mathbf{Y}_p$ in this case is

$$\max_{i,j} \frac{|\mathbf{Y}_p^{(i,j)} - \mathbf{Y}_{p+2}^{(i,j)}|}{1 + |\mathbf{Y}_p^{(i,j)}|}. \tag{8}$$

When implementing this scheme, several observations were exploited in order to obtain substantial savings in computation time.

1. From our numerical experiments, we have observed that the primary solution, $\mathbf{Y}_p$, proves to be an effective initial guess for the solution of the system of nonlinear equations based on the higher order MIRK formula. Because $\mathbf{Y}_p$ is saved in the solution structure employed by `BVP_SOLVER`, it is available for use as the initial guess at no additional computational cost.

2. From our numerical experiments, we have observed that the Jacobian matrix from the primary solution computation is a good approximation to the Jacobian associated with the nonlinear system based on the higher order MIRK formula. This matrix is also available within one of the arrays used by `BVP_SOLVER` during the computation of the primary solution. We are therefore able to avoid the expensive evaluation and factorization of this matrix during the Newton iteration for the determination of $\mathbf{Y}_{p+2}$.

3. We initially employed the same form of Newton iteration for the determination of $\mathbf{Y}_{p+2}$ that is used in the computation of the primary solution; this meant that full and damped Newton steps were allowed, and the iteration terminated when an appropriately scaled norm of the Newton correction was less than the user tolerance. However, we have experimented with a simpler version of the iteration in which only one Newton correction is performed. Our experiments showed that the resultant estimates were essentially the same as those obtained by the full Newton iteration scheme. Accordingly, we perform only one Newton correction for the computation of $\mathbf{Y}_{p+2}$.

By making use of the (factored) Jacobian from the primary computation and by employing only one Newton iteration, the implementation of this GE estimate involves only one backward substitution, based on one evaluation of $\mathbf{\Phi}$ in (4).

3.2 GE Estimation based on Deferred Correction (DC)

When the MIRK method upon which (4) is based is of order p, we rewrite (4) as

$$\mathbf{\Phi}_p(\mathbf{Y}_p) = \mathbf{0},$$

where the pth order discrete solution, obtained by solving this system, is $\mathbf{Y}_p$. In [5], the authors describe a deferred correction method based on MIRK formulas. They demonstrate how solutions of orders four, six, and eight can be computed by solving the systems

$$\mathbf{\Phi}_4(\mathbf{Y}_4) = \mathbf{0}, \quad \mathbf{\Phi}_4(\mathbf{Y}_6) = -\mathbf{\Phi}_6(\mathbf{Y}_4), \quad \mathbf{\Phi}_4(\mathbf{Y}_8) = \mathbf{\Phi}_4(\mathbf{Y}_6) - \mathbf{\Phi}_8(\mathbf{Y}_6), \qquad (9)$$

for $\mathbf{Y}_4, \mathbf{Y}_6$, and $\mathbf{Y}_8$. In [5], the authors first compute a solution of order four and then use two steps of deferred correction to efficiently compute the solutions of orders six and eight. The fourth order solution $\mathbf{Y}_4$ is computed using the first equation from (9), and the sixth and eighth order solutions, $\mathbf{Y}_6$ and $\mathbf{Y}_8$, are obtained by solving the second and third equations from (9).

BVP_SOLVER uses MIRK formulas and thus we can use an approach similar to that of [5]; however, we use only one step of deferred correction to obtain a higher order solution. When BVP_SOLVER solves a BVODE it returns a primary solution of order two, four, or six, depending on what the user has selected. Assume that the primary solution is of order p; i.e., we have $\mathbf{Y}_p$ available from the primary defect controlled computation, and that the deferred correction equation that allows us to obtain a higher order solution, $\mathbf{Y}_{p+2}$, is

$$\mathbf{\Phi}_p(\mathbf{Y}_{p+2}) = -\mathbf{\Phi}_{p+2}(\mathbf{Y}_p).$$

That is, we need to solve the nonlinear system

$$\mathbf{\Phi}_p(\mathbf{z}) + \mathbf{\Phi}_{p+2}(\mathbf{Y}_p) = \mathbf{0}, \tag{10}$$

for $\mathbf{z} = \mathbf{Y}_{p+2}$. The primary expense is the construction and factorization of the Jacobian matrix of this nonlinear system. However, the system,

$$\mathbf{\Phi}_p(\mathbf{z}) = \mathbf{0}, \tag{11}$$

is the one that was just solved during the primary computation to get $\mathbf{Y}_p$. The corresponding Jacobian (evaluated at $\mathbf{Y}_p$ or an approximation to it) was computed and factored for use in that iteration; a significant advantage of employing (10) to determine $\mathbf{Y}_{p+2}$ is that it has the same Jacobian matrix as (11), and thus the Jacobian and its factorization are available at no cost. Furthermore, a natural initial guess for $\mathbf{Y}_{p+2}$ to start the Newton iteration for (10) is $\mathbf{Y}_p$. As in the approach described in Section 3.1, we also apply only one Newton step to obtain an estimate of $\mathbf{Y}_{p+2}$.

Once $\mathbf{Y}_{p+2}$ is available, the estimate of the norm of the GE for $\mathbf{Y}_p$ is computed as in (8). The computational costs incurred in this approach involve the computation of the correction term, $\mathbf{\Phi}_{p+2}(\mathbf{Y}_p)$, one evaluation of $\mathbf{\Phi}_p(\mathbf{z})$, and one backward substitution associated with applying Newton's method to (10).

3.3 A GE Bound based on Estimating the BVODE Conditioning Constant (CO)

The third GE estimation approach we consider is based on a backward error analysis for the numerical solution of a BVODE. Here we briefly review the main points; see, e.g., [14] for further details.

We consider a linear BVODE and assume that the exact solution, $\mathbf{y}(x)$, satisfies

$$\mathbf{y}'(x) = \mathbf{A}(x)\mathbf{y}(x) + \mathbf{q}(x), \quad \mathbf{B}_a\mathbf{y}(a) + \mathbf{B}_b\mathbf{y}(b) = \boldsymbol{\beta}. \tag{12}$$

In (12), $\mathbf{A}(x), \mathbf{B}_a, \mathbf{B}_b \in R^{n\times n}$ and $\mathbf{q}(x), \mathbf{y}(x), \boldsymbol{\beta} \in R^n$.

Then the approximate solution, $\mathbf{S}(x)$, exactly satisfies the perturbed BVODE and BCs

$$\mathbf{S}'(x) = \mathbf{A}(x)\mathbf{S}(x) + \mathbf{q}(x) + \boldsymbol{\delta}(x), \quad \mathbf{B}_a\mathbf{S}(a) + \mathbf{B}_b\mathbf{S}(b) = \boldsymbol{\beta} + \boldsymbol{\sigma},$$

where $\boldsymbol{\delta}(x) = \mathbf{S}'(x) - \mathbf{A}(x)\mathbf{S}(x) - \mathbf{q}(x)$ and $\boldsymbol{\sigma} = \mathbf{B}_a\mathbf{S}(a) + \mathbf{B}_b\mathbf{S}(b) - \boldsymbol{\beta}$ are the defects associated with the ODE and the BCs, respectively.

The main result is that

$$||\mathbf{y}(x) - \mathbf{S}(x)||_{\mathbf{W}_3} \leq \kappa \max\left(||\boldsymbol{\delta}(x)||_{\mathbf{W}_1}, ||\boldsymbol{\sigma}||_{\mathbf{W}_2}\right), \tag{13}$$

where κ is a conditioning constant for the BVODE and the weighted norms are defined as follows:

$$||\boldsymbol{\delta}(x)||_{\mathbf{W}_1} = \max_{a \leq x \leq b} ||\mathbf{W}_1^{-1}(x)\boldsymbol{\delta}(x)||_\infty, \quad ||\boldsymbol{\sigma}||_{\mathbf{W}_2} = ||\mathbf{W}_2^{-1}\boldsymbol{\sigma}||_\infty,$$

$$||\mathbf{S}(x) - \mathbf{y}(x)||_{\mathbf{W}_3} = \max_{a \leq x \leq b} ||\mathbf{W}_3^{-1}(x)\left(\mathbf{S}(x) - \mathbf{y}(x)\right)||_\infty.$$

Here $\mathbf{W}_1(x)$, $\mathbf{W}_2$, and $\mathbf{W}_3(x)$ are $n \times n$ diagonal matrices with positive entries. The matrix $\mathbf{W}_3(x)$ is associated with the scaling of the defect (6); the matrix $\mathbf{W}_1(x)$ is associated with the scaling for the GE (8). Because we do not scale the BCs, $\mathbf{W}_2$ is taken to be the identity matrix.

The conditioning constant, κ, depends on the fundamental solution of the corresponding homogeneous BVODE and the boundary condition matrices $\mathbf{B}_a$ and $\mathbf{B}_b$. (The conditioning constant is given by $\kappa = \max(\kappa_1, \kappa_2)$, where κ_1 is the conditioning constant for the BCs and κ_2 is the conditioning constant for the ODEs. However, in `BVP_SOLVER`, the BCs are solved much more accurately than the ODEs; so in fact only κ_2 is relevant in the present context.)

The paper [14] also explains how to compute an estimate of κ. Let

$$\overline{\mathbf{W}}_{12} = \text{diag}\{\mathbf{W}_1(x_1), \ldots, \mathbf{W}_1(x_N), \mathbf{W}_2\}, \quad \overline{\mathbf{W}}_3 = \text{diag}\{\mathbf{W}_3(x_0), \ldots, \mathbf{W}_3(x_N)\}.$$

Then in [14], it is shown that, for a sufficiently fine mesh,

$$\kappa \approx ||\overline{\mathbf{W}}_3^{-1}\mathbf{J}_{\boldsymbol{\Phi}}^{-1}\overline{\mathbf{W}}_{12}||_\infty,$$

where $\mathbf{J}_{\boldsymbol{\Phi}}$ is the Jacobian matrix (5). The paper [14] suggests the use of the Higham–Tisseur algorithm [9] for the efficient estimation of this norm.

Because the factored Newton matrix from the primary solution computation is available, once the primary solution is accepted, the Higham–Tisseur algorithm can be used to obtain the estimate for κ using only a few additional back solves involving the matrix $\overline{\mathbf{W}}_{12}^{-1}\mathbf{J}_{\boldsymbol{\Phi}}\overline{\mathbf{W}}_3$. (The right hand sides involved in these backward substitutions are generated by the software based on the Higham–Tisseur algorithm and represent no significant computational cost.)

We have modified `BVP_SOLVER` to provide an option to efficiently estimate κ after the primary solution is accepted. The product of κ and the estimate of the maximum norm of the defect is then used to obtain an upper bound on the GE as in (13). However, it is worth noting that, especially for a defect control solver, it may be useful to estimate and return κ itself because this quantity gives a measure of the sensitivity of the solution to small changes in the problem.

3.4 Modification of the Implementation of the RE based GE Estimate (RE)

Based on numerical experiments, we found that the treating the nonlinear system associated with the RE approach by performing only one full Newton step yielded error estimates that compared well with those obtained from a full tolerance controlled Newton iteration. We therefore employ a more efficient implementation of the RE based GE estimate that computes and factors a new Jacobian matrix and then performs only one full Newton step.

4 Comparison of the GE Estimates

With the implementations described in the previous section, `BVP_SOLVER` now has four possible methods for estimating the GE. In this section, we discuss accuracy and computational efficiency results for these four estimators. All GE estimates are for the scaled norms specified earlier. We consider three test problems and examine the performance of the GE estimators for the three MIRK formula order options (2, 4, and 6) and for the range of tolerance values $10^{-4}, 10^{-5}, \ldots, 10^{-8}$. All test problems were converted to first order systems as required by `BVP_SOLVER`.

Each problem is solved a number of times in succession in order to obtain cumulative timings that are large enough to be reliable, i.e., on the order of 10 seconds. Each problem also depends on a positive parameter ϵ, where the problem difficulty increases as ϵ decreases. Values of ϵ are chosen according to those suggested by their sources in the literature unless this led to excessively large solution times, in which case the value of ϵ was increased slightly. Consequently, the number of timing runs varies according to the problem solved and the order of the discretization. The minimum time from three timing runs is reported.

The computations were performed using an Intel Xeon w3520 quad core processor running at 2.667 GHz. The RAM consisted of 16GB of DDR3 memory running at 1.333 GHz. The operating system was 64-bit Ubuntu 10.04.2 LTS with kernel 2.6.32-32-generic and the `Fortran` compiler was gfortran with gcc 4.4.3-4ubuntu5.

4.1 Test Problems

1. The first problem is

$$\epsilon y'' + (y')^2 = 1, \tag{14a}$$

with BCs

$$y(0) = 1 + \epsilon \ln\cosh\left(\frac{-0.745}{\epsilon}\right), \quad y(1) = 1 + \epsilon \ln\cosh\left(\frac{0.255}{\epsilon}\right), \tag{14b}$$

and exact solution

$$y(x) = 1 + \epsilon \ln\cosh\left(\frac{x - 0.745}{\epsilon}\right).$$

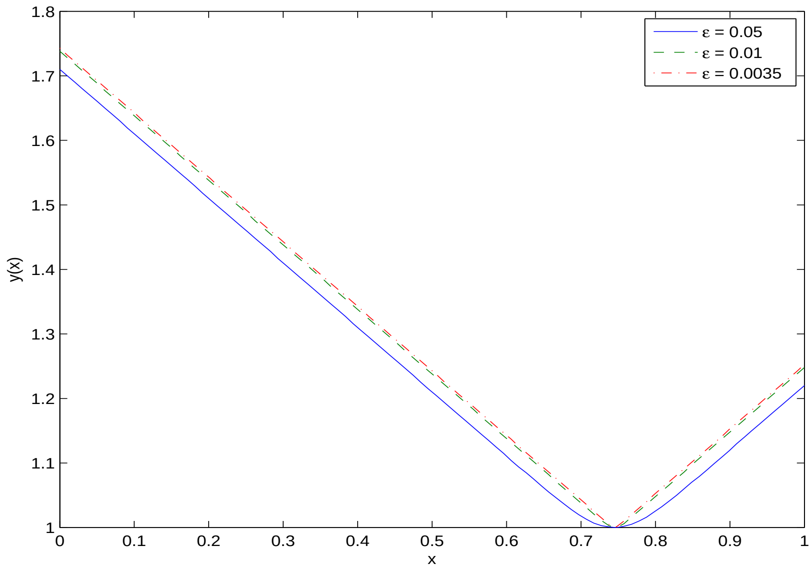

Figure 1: Solution $y(x)$ of problem (14) for $\epsilon = 0.05$, 0.01, and 0.0035.

This is Problem 20 from `www.ma.ic.ac.uk/~jcash/BVP_software`; see also [4]. For MIRK order 2, we use $\epsilon = 0.05$. For MIRK orders 4 and 6, we use $\epsilon = 0.0035$; in Section 5 we also use $\epsilon = 0.01$. The solutions $y(x)$ for these values of ϵ are displayed in Figure 1. We use an initial guess of $y(x) \equiv \frac{1}{2}$, $y'(x) \equiv 0$. Timing results are the average of 500 runs.

2. The second problem is

$$\epsilon y'' = y + y^2 - \exp\left(\frac{-2x}{\sqrt{x}}\right), \tag{15a}$$

with BCs,

$$y(0) = 1, \quad y(1) = \exp\left(\frac{-1}{\sqrt{\epsilon}}\right), \tag{15b}$$

and exact solution

$$y(x) = \exp\left(\frac{-x}{\sqrt{\epsilon}}\right).$$

This is Problem 21 from `www.ma.ic.ac.uk/~jcash/BVP_software`. For MIRK order 2, we use $\epsilon = 10^{-7}$. For MIRK order 4, we use $\epsilon = 5 \times 10^{-8}$. For MIRK order 6, we use $\epsilon = 10^{-8}$. The solutions $y(x)$ for these values of ϵ are displayed in Figure 2. (Note that, in order to make visible the differences between the solutions associated with different ϵ values, the

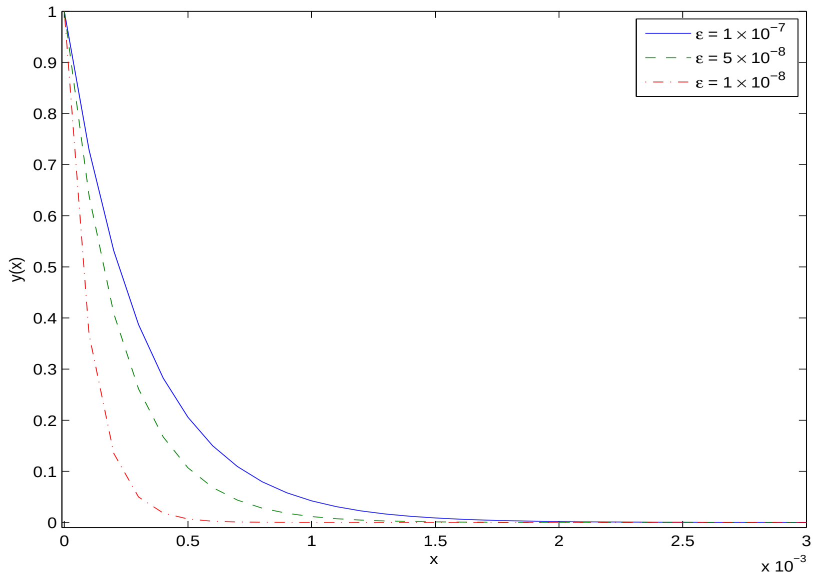

Figure 2: Zoom-in of solution $y(x)$ of problem (15) for $\epsilon = 1 \times 10^{-7}$, 5×10^{-8}, and 1×10^{-8}.

horizontal axis in Figure 2 includes only the region $[0, 0.003]$.) We use an initial guess of $y(x) \equiv \frac{1}{2}$, $y'(x) \equiv 0$. Timing results are the average of 100 runs.

3. The third problem is

$$\epsilon f'''' + ff''' + gg' = 0, \quad \epsilon g'' + fg' - f'g = 0, \tag{16a}$$

with BCs,

$$f(0) = f(1) = f'(0) = f'(1) = 0, \; g(0) = \Omega_0, \; g(1) = \Omega_1, \tag{16b}$$

where $\Omega_0 = -1$, $\Omega_1 = 1$, and $\epsilon = 9 \times 10^{-5}$; in Section 5 we also use $\epsilon = 5\times10^{-3}$. This is Example 1.20 of [2]. Because no exact solution for this problem is known, a reference solution was computed by `BVP_SOLVER` using the sixth order MIRK method with a tolerance of 10^{-11}. The solutions $f(x)$ and $g(x)$ for these values of ϵ are displayed in Figure 3. We use an initial guess of $f(x) = f'(x) = f''(x) = f'''(x) = 0$ and $g(x) = 2x - 1$, $g'(x) = 2$. Timing results are the average of 3000 runs.

4.2 Results for Second Order

There is generally good agreement between the true GE and the estimated GE from RE, the approach based on the use of a higher order method (HO), and

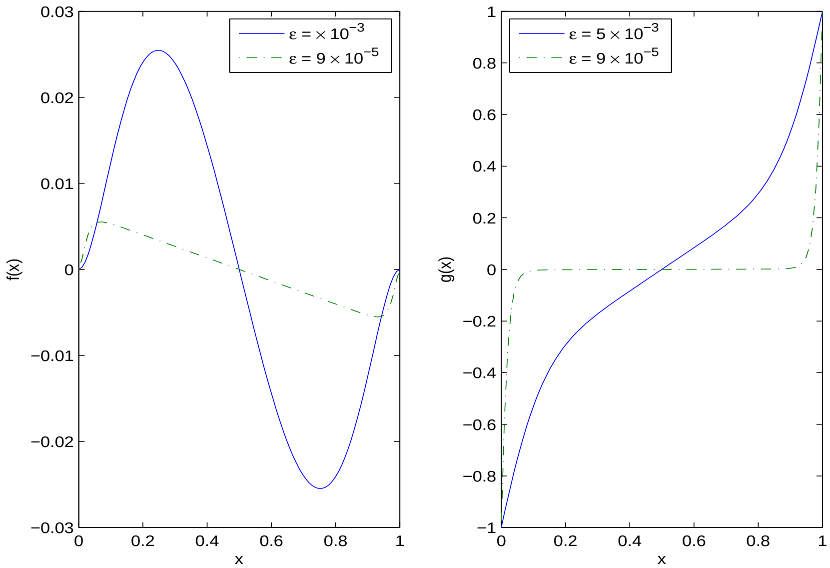

Figure 3: Solutions $f(x)$ and $g(x)$ of problem (16) for $\epsilon = 5\times10^{-3}$ and 9×10^{-5}.

the approach based on deferred correction (DC). All results from the use of the conditioning constant estimate (CO) give a substantial overestimate of the GE, however, typically by several orders of magnitude. Tables 1, 2, and 3 provide results for second order for the three test problems. "% Total" is the time for the computation of the error estimate expressed as a percentage of the time taken to obtain the original defect controlled solution. "τ" is the absolute difference between the actual error and the estimated error.

We next consider the relative efficiency of the estimators by considering plots of execution time of each estimator (relative to the time required to compute the primary solution) vs. the tolerance. Typically, the execution time for the RE estimator is a much higher percentage of the primary solution computation time than the other estimators; see Figure 4. For problem (14), the relative costs of the RE estimator are approximately between 23% and 40% for all tolerances. The relative costs for the HO scheme are about 4% to 5%; the relative costs for the DC scheme are somewhat larger, ranging from about 4% to 10%; the CO scheme costs range from about 4% to 8%. Similar results are obtained for test problem (15); see Figure 5. The relative costs of the RE estimator are approximately between 10% and 15%. The relative costs for the HO scheme are about 2% to 3%; the DC costs are about 4% to 5%; the CO costs are about 2%. For problem (16), we see slightly different results; see Figure 6. The relative cost of the RE estimator increases as the tolerance becomes sharper. The cost of the RE estimator is approximately 4% for Tol $= 10^{-4}$ and steadily increases to

	Tol	Time	% Total	Actual Error	Estimated Error	τ
RE	1E–4	1.90×10^{-3}	23.46 %	1.237×10^{-5}	1.239×10^{-5}	1.679×10^{-8}
	1E–5	5.70×10^{-3}	40.14 %	1.406×10^{-6}	1.406×10^{-6}	1.384×10^{-10}
	1E–6	1.41×10^{-2}	27.24 %	1.360×10^{-7}	1.362×10^{-7}	2.344×10^{-10}
	1E–7	4.35×10^{-2}	32.68 %	1.720×10^{-8}	1.724×10^{-8}	4.255×10^{-11}
	1E–8	1.63×10^{-1}	26.85 %	1.324×10^{-9}	1.324×10^{-9}	4.921×10^{-13}
CO	1E–4	5.40×10^{-4}	6.67 %	1.237×10^{-5}	2.073	2.073
	1E–5	1.20×10^{-3}	8.45 %	1.406×10^{-6}	6.557×10^{-1}	6.557×10^{-1}
	1E–6	2.26×10^{-3}	4.37 %	1.360×10^{-7}	1.501×10^{-1}	1.501×10^{-1}
	1E–7	7.68×10^{-3}	5.77 %	1.720×10^{-8}	5.650×10^{-2}	5.650×10^{-2}
	1E–8	2.92×10^{-2}	4.80 %	1.324×10^{-9}	1.616×10^{-2}	1.616×10^{-2}
HO	1E–4	3.00×10^{-4}	3.70 %	1.237×10^{-5}	1.264×10^{-5}	2.704×10^{-7}
	1E–5	6.00×10^{-4}	4.23 %	1.406×10^{-6}	1.406×10^{-6}	5.855×10^{-10}
	1E–6	3.34×10^{-3}	6.46 %	1.360×10^{-7}	1.364×10^{-7}	4.206×10^{-10}
	1E–7	8.48×10^{-3}	6.37 %	1.720×10^{-8}	1.728×10^{-8}	7.989×10^{-11}
	1E–8	3.18×10^{-2}	5.23 %	1.324×10^{-9}	1.326×10^{-9}	2.389×10^{-12}
DC	1E–4	4.60×10^{-4}	5.68 %	1.237×10^{-5}	1.264×10^{-5}	2.704×10^{-7}
	1E–5	6.20×10^{-4}	4.37 %	1.406×10^{-6}	1.406×10^{-6}	5.855×10^{-10}
	1E–6	5.12×10^{-3}	9.91 %	1.360×10^{-7}	1.364×10^{-7}	4.206×10^{-10}
	1E–7	1.35×10^{-2}	10.17 %	1.720×10^{-8}	1.728×10^{-8}	7.989×10^{-11}
	1E–8	5.16×10^{-2}	8.49 %	1.324×10^{-9}	1.326×10^{-9}	2.389×10^{-12}

Table 1: GE Estimates applied to (14), MIRK order 2. Number of points in the final mesh for each tolerance, respectively: 394, 1162, 3722, 10459, and 37853.

approximately 27% as when Tol $= 10^{-8}$. The relative costs for the HO scheme are about 1% to 4%; the costs for the DC scheme are about 1% to 5%; the costs for the CO estimator are higher and range from about 2% to about 15%.

4.3 Results for Fourth Order

In this case, we again find that there is excellent agreement between the true GE and the estimated GE from the RE, HO, and DC approaches. For all cases, the CO approach gives a significant overestimate of the GE. Tables 4, 5, and 6 provide results for fourth order for the three test problems. "% Total" is the time for the computation of the error estimate expressed as a percentage of the time taken to obtain the original defect controlled solution. "τ" is the absolute difference between the actual error and the estimated error.

We next consider the relative efficiency of the estimators by considering plots

	Tol	Time	% Total	Actual Error	Estimated Error	τ
RE	1E–4	4.22×10^{-2}	11.89 %	1.560×10^{-6}	1.560×10^{-6}	3.122×10^{-11}
	1E–5	4.23×10^{-2}	11.44 %	1.501×10^{-6}	1.501×10^{-6}	3.851×10^{-13}
	1E–6	4.25×10^{-2}	10.43 %	1.470×10^{-6}	1.470×10^{-6}	1.330×10^{-13}
	1E–7	1.45×10^{-1}	14.00 %	1.350×10^{-7}	1.350×10^{-7}	6.569×10^{-16}
	1E–8	4.28×10^{-1}	14.87 %	1.744×10^{-8}	1.744×10^{-8}	2.192×10^{-16}
CO	1E–4	6.80×10^{-3}	1.92 %	1.560×10^{-6}	1.279×10^{4}	1.279×10^{4}
	1E–5	5.80×10^{-3}	1.57 %	1.501×10^{-6}	5.819×10^{3}	5.819×10^{3}
	1E–6	6.60×10^{-3}	1.62 %	1.470×10^{-6}	5.581×10^{3}	5.581×10^{3}
	1E–7	2.22×10^{-2}	2.14 %	1.350×10^{-7}	2.950×10^{3}	2.950×10^{3}
	1E–8	6.47×10^{-2}	2.25 %	1.744×10^{-8}	9.741×10^{2}	9.741×10^{2}
HO	1E–4	8.70×10^{-3}	2.45 %	1.560×10^{-6}	1.560×10^{-6}	2.651×10^{-10}
	1E–5	8.40×10^{-3}	2.27 %	1.501×10^{-6}	1.501×10^{-6}	6.467×10^{-11}
	1E–6	8.50×10^{-3}	2.09 %	1.470×10^{-6}	1.470×10^{-6}	9.650×10^{-12}
	1E–7	2.93×10^{-2}	2.83 %	1.350×10^{-7}	1.350×10^{-7}	7.125×10^{-15}
	1E–8	8.55×10^{-2}	2.97 %	1.744×10^{-8}	1.744×10^{-8}	3.708×10^{-16}
DC	1E–4	1.41×10^{-2}	3.97 %	1.560×10^{-6}	1.560×10^{-6}	3.592×10^{-10}
	1E–5	1.45×10^{-2}	3.92 %	1.501×10^{-6}	1.501×10^{-6}	6.625×10^{-11}
	1E–6	1.43×10^{-2}	3.51 %	1.470×10^{-6}	1.470×10^{-6}	9.657×10^{-12}
	1E–7	4.74×10^{-2}	4.57 %	1.350×10^{-7}	1.350×10^{-7}	7.125×10^{-15}
	1E–8	1.35×10^{-1}	4.70 %	1.744×10^{-8}	1.744×10^{-8}	3.708×10^{-16}

Table 2: GE Estimates applied to (15), MIRK order 2. Number of points in the final mesh for each tolerance, respectively: 9217, 9217, 9217, 30166, and 83883.

of execution time of each estimator (relative to the time required to compute the primary solution) vs. the tolerance. Figure 7 shows results for test problem (14). The cost of the RE estimator is somewhat larger than those of all other error estimators. However, in all cases, the cost of the GE estimate is small compared to that of the primary computation. Figure 8 shows results for test problem (15). Relative costs for the RE scheme range from about 9% to 12%; the HO scheme costs are about 2%; the DC costs range from 3% to 4%; the CO costs are about 1%. For test problem (16), the cost of the RE estimator steadily increases as the tolerance becomes sharper (approximately between 8% and 21%). The HO scheme costs are about 1% to 4%; the DC costs are about 1% to 2%; the CO costs are about 2% to 6%. Figure 9 shows results for test problem (16).

	Tol	Time	% Total	Actual Error	Estimated Error	τ
RE	1E–4	2.87×10^{-2}	4.04 %	1.095×10^{-2}	1.101×10^{-2}	5.999×10^{-5}
	1E–5	9.07×10^{-2}	9.47 %	7.255×10^{-4}	7.240×10^{-4}	1.462×10^{-6}
	1E–6	2.74×10^{-1}	15.47 %	6.967×10^{-5}	6.963×10^{-5}	3.559×10^{-8}
	1E–7	8.35×10^{-1}	21.53 %	7.759×10^{-6}	7.758×10^{-6}	4.924×10^{-10}
	1E–8	2.48	26.92 %	8.083×10^{-7}	8.082×10^{-7}	4.168×10^{-11}
CO	1E–4	1.25×10^{-2}	1.76 %	1.095×10^{-2}	3.683×10^{8}	3.683×10^{8}
	1E–5	3.97×10^{-2}	4.14 %	7.255×10^{-4}	8.577×10^{7}	8.577×10^{7}
	1E–6	1.54×10^{-1}	8.72 %	6.967×10^{-5}	2.607×10^{7}	2.607×10^{7}
	1E–7	4.69×10^{-1}	12.09 %	7.759×10^{-6}	8.056×10^{6}	8.056×10^{6}
	1E–8	1.39	15.11 %	8.083×10^{-7}	2.718×10^{6}	2.718×10^{6}
HO	1E–4	5.33×10^{-3}	0.75 %	1.095×10^{-2}	1.148×10^{-2}	5.281×10^{-4}
	1E–5	1.29×10^{-2}	1.35 %	7.255×10^{-4}	7.306×10^{-4}	5.050×10^{-6}
	1E–6	4.24×10^{-2}	2.40 %	6.967×10^{-5}	6.975×10^{-5}	7.952×10^{-8}
	1E–7	1.26×10^{-1}	3.26 %	7.759×10^{-6}	7.760×10^{-6}	9.911×10^{-10}
	1E–8	3.78×10^{-1}	4.10 %	8.083×10^{-7}	8.083×10^{-7}	1.407×10^{-11}
DC	1E–4	4.89×10^{-3}	0.69 %	1.095×10^{-2}	1.148×10^{-2}	5.281×10^{-4}
	1E–5	1.88×10^{-2}	1.96 %	7.255×10^{-4}	7.306×10^{-4}	5.050×10^{-6}
	1E–6	5.47×10^{-2}	3.09 %	6.967×10^{-5}	6.975×10^{-5}	7.952×10^{-8}
	1E–7	1.66×10^{-1}	4.29 %	7.759×10^{-6}	7.760×10^{-6}	9.912×10^{-10}
	1E–8	4.97×10^{-1}	5.39 %	8.083×10^{-7}	8.083×10^{-7}	1.407×10^{-11}

Table 3: GE estimation applied to (16), MIRK order 2. Number of points in the final mesh for each tolerance, respectively: 2372, 7330, 21856, 65445, and 193040.

4.4 Results for Sixth Order

We again find that the accuracy of the error estimates for the RE, HO, and DC approaches is excellent. As in the previous cases, the CO estimates are several orders of magnitude too large. Tables 7, 8, and 9 provide results for sixth order for the three test problems. "% Total" is the time for the computation of the error estimate expressed as a percentage of the time taken to obtain the original defect controlled solution. "τ" is the absolute difference between the actual error and the estimated error.

We next consider the relative efficiency of the estimators by considering plots of execution time of each estimator (relative to the time required to compute the primary solution) vs. the tolerance. Figure 10 shows results for test problem (14). For problem (14), the costs of the RE estimator are small (approximately

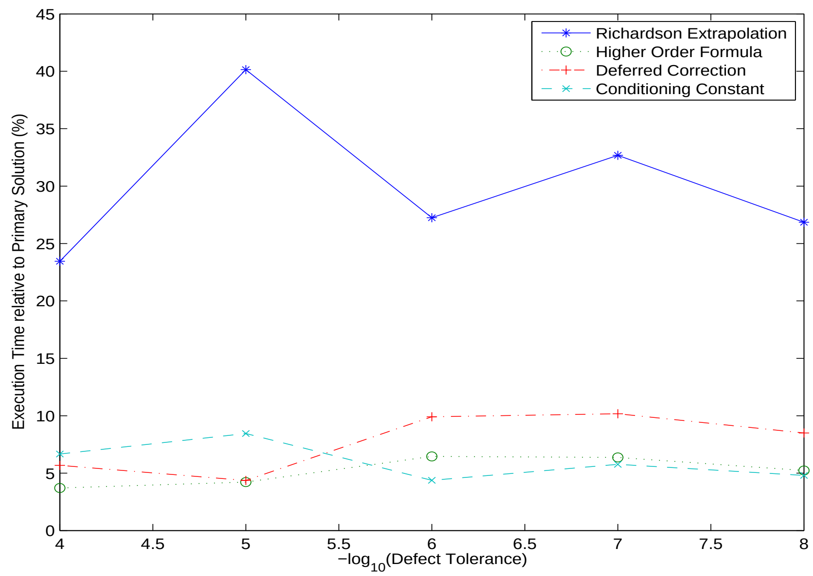

Figure 4: Relative execution time of GE estimators vs. $-\log_{10}$ of defect tolerance with second order MIRK formula for test problem (14).

between 1% and 3%). The relative costs of the other error estimators are less than approximately 1%. Figure 11 shows results for test problem (15). For problem (15), the costs of the RE estimator are larger (approximately between 8% and 17%). The HO costs are 2% to 3%; the DC costs are 3% to 6%; the CO costs are 1% to 2%. For problem (16), the relative cost of RE increases from approximately 8% to 26% as the tolerance grows sharper. The relative costs of the CO estimator are less than 5%. The relative costs the HO scheme are 1% to 2%; the DC costs are 1% to 4% approximately. See Figure 12.

5 BVP_SOLVER with Defect/Global Error Control

5.1 Comparison of Defect/GE Control Schemes

The version of BVP_SOLVER that we consider in this section provides an option for the computation of a *GE controlled* numerical solution. This version of the code performs the same basic computation to obtain a discrete numerical solution at the mesh points of a given mesh as does the original. Once this discrete numerical solution is obtained, the new version of the solver can then compute an estimate of the (scaled) GE of that solution using one of the GE estimation algorithms analyzed in the previous section. That is, this version of

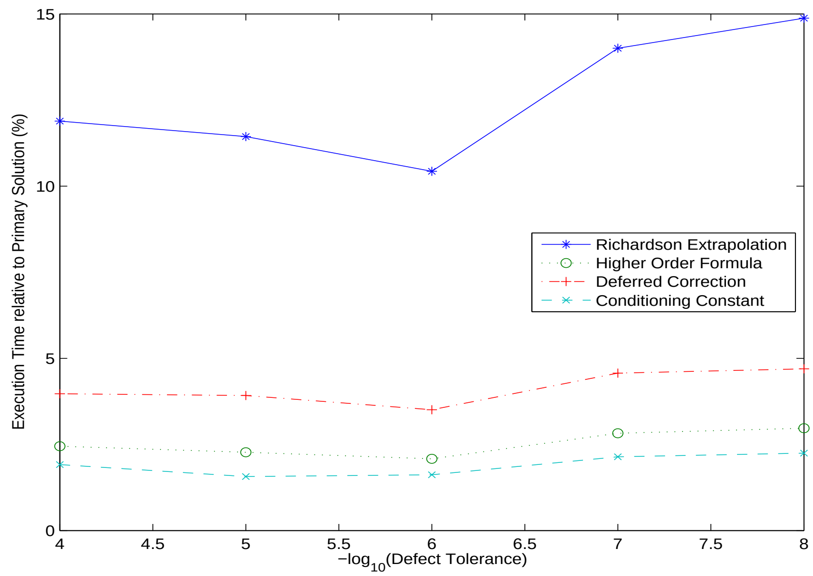

Figure 5: Relative execution time of GE estimators vs. $-\log_{10}$ of defect tolerance with second order MIRK formula for test problem (15).

BVP_SOLVER is able to compute an estimate of the GE for the discrete numerical solution obtained on each intermediate mesh rather than only at the end of the computation, as considered in the previous section. In this version of the code, if the estimate of the GE does not satisfy the tolerance, the GE estimates for each subinterval are passed to the mesh adaptation algorithm, where they are used to construct a new mesh.

The mesh adaptation algorithm is identical to what is used in the defect control case except for one parameter setting, which we now describe. Two important quantities that are computed in the BVP_SOLVER mesh adaptation routine are related to the maximum GE or defect over all subintervals and the average GE or defect over all subintervals. The ratio of the former quantity to the latter is computed and compared to a parameter called ρ. If the ratio is greater than or equal to ρ, a new mesh is constructed based on equidistribution of the GE or defect. Otherwise, a new mesh is constructed by halving each subinterval of the current mesh. In the original version of BVP_SOLVER, $\rho = 1$, and this forces an equidistribution process for the construction of *every* new mesh. When we tried to use $\rho = 1$ for the new GE controlled version of BVP_SOLVER, we found that it was impossible to obtain a solution for any of the test problems, even using meshes with millions of points. It was necessary to use a larger value of ρ (we chose $\rho = 2$, a common choice in the literature) in order

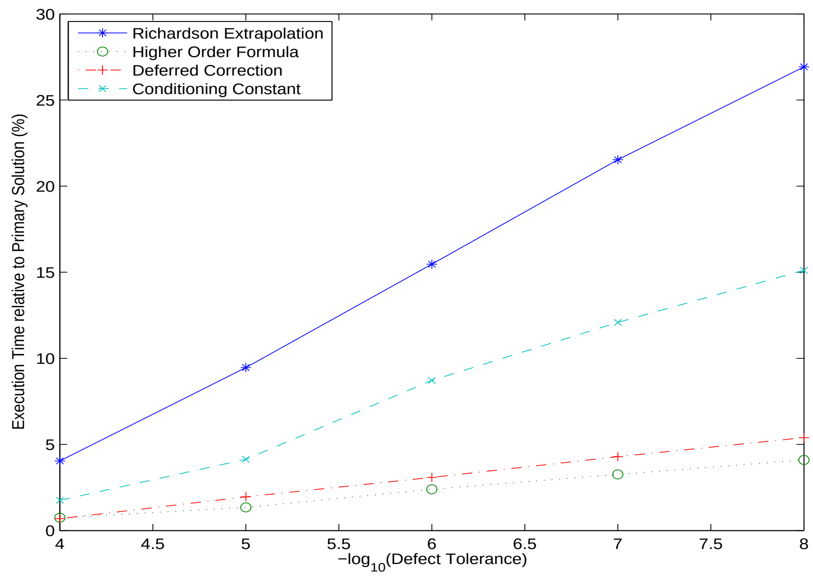

Figure 6: Relative execution time of GE estimators vs. $-\log_{10}$ of defect tolerance with second order MIRK formula for test problem (16).

to force an occasional global refinement (via mesh halving) of the mesh. This was necessary to reduce the size of non-local contributions to the GE on each subinterval. See [6] for further details regarding the mesh adaptation algorithm employed in BVP_SOLVER and [2] for further discussion of mesh adaptation based on error estimates.

In this section we present some results that represent a preliminary investigation of the use of the GE control mode in the new version of BVP_SOLVER. We have considered numerical experiments employing the test problems (14) with $\epsilon = 0.01$ (see Figure 1) and (16) with $\epsilon = 0.005$ (see Figure 3), for all three orders, and for a range of tolerances. In each table presented here, we report, for each tolerance, the required CPU time, the number of points used in the final mesh (N), the estimated and true maximum GE, and the estimated and true maximum defect. (The true maximum GE and defect were obtained by sampling them at 10 points per subinterval.) The DC algorithm of the previous section was used to estimate the GE. BVP_SOLVER is run in each of four control modes: defect control (DefC), GE control (GEC), and sequential and parallel combinations of defect and GE control, which we now describe.

- In sequential combination control (SCC) mode, a defect controlled solution is computed, and it and its corresponding mesh are passed as the initial data to a GE controlled computation. When the BVODE is poorly condi-

	Tol	Time	% Total	Actual Error	Estimated Error	τ
RE	1E–4	6.30×10^{-3}	2.55 %	6.491×10^{-6}	6.492×10^{-6}	5.702×10^{-10}
	1E–5	2.82×10^{-3}	1.15 %	8.836×10^{-10}	8.836×10^{-10}	1.448×10^{-14}
	1E–6	2.38×10^{-3}	0.97 %	8.656×10^{-10}	8.656×10^{-10}	1.391×10^{-14}
	1E–7	2.38×10^{-3}	0.96 %	8.654×10^{-10}	8.654×10^{-10}	1.365×10^{-14}
	1E–8	1.96×10^{-3}	0.79 %	8.653×10^{-10}	8.653×10^{-10}	1.393×10^{-14}
CO	1E–4	7.80×10^{-4}	0.32 %	6.491×10^{-6}	4.245×10^{1}	4.245×10^{1}
	1E–5	3.00×10^{-4}	0.12 %	8.836×10^{-10}	1.914×10^{-2}	1.914×10^{-2}
	1E–6	4.20×10^{-4}	0.17 %	8.656×10^{-10}	1.788×10^{-2}	1.788×10^{-2}
	1E–7	2.80×10^{-4}	0.11 %	8.654×10^{-10}	1.807×10^{-2}	1.807×10^{-2}
	1E–8	2.40×10^{-4}	0.10 %	8.653×10^{-10}	1.785×10^{-2}	1.785×10^{-2}
HO	1E–4	1.24×10^{-3}	0.50 %	6.491×10^{-6}	6.627×10^{-6}	1.354×10^{-7}
	1E–5	6.60×10^{-4}	0.27 %	8.836×10^{-10}	8.836×10^{-10}	2.544×10^{-14}
	1E–6	4.00×10^{-4}	0.16 %	8.656×10^{-10}	8.656×10^{-10}	2.406×10^{-14}
	1E–7	3.60×10^{-4}	0.14 %	8.654×10^{-10}	8.654×10^{-10}	2.356×10^{-14}
	1E–8	5.20×10^{-4}	0.21 %	8.653×10^{-10}	8.653×10^{-10}	2.406×10^{-14}
DC	1E–4	1.80×10^{-3}	0.73 %	6.491×10^{-6}	6.595×10^{-6}	1.038×10^{-7}
	1E–5	1.02×10^{-3}	0.42 %	8.836×10^{-10}	8.836×10^{-10}	2.284×10^{-14}
	1E–6	6.60×10^{-4}	0.27 %	8.656×10^{-10}	8.656×10^{-10}	2.145×10^{-14}
	1E–7	6.40×10^{-4}	0.26 %	8.654×10^{-10}	8.654×10^{-10}	2.088×10^{-14}
	1E–8	9.40×10^{-4}	0.38 %	8.653×10^{-10}	8.653×10^{-10}	2.145×10^{-14}

Table 4: GE Estimates applied to (14), MIRK order 4. Number of points in the final mesh for each tolerance, respectively: 1153, 577, 577, 577, and 577.

tioned, the conditioning constant is large and (from (13)) we can expect the GE to be larger than the defect. It will thus be easier to compute a defect controlled numerical solution than a GE controlled numerical solution, potentially making it more efficient in such cases to first compute a defect controlled solution rather than directly compute a GE controlled numerical solution. When the defect controlled solution is passed to `BVP_SOLVER` in GE control mode, the solver first estimates the GE of that solution. If this estimate does not satisfy the tolerance, then a new mesh is constructed based on the GE estimate and the GE control mode computation begins.

- In parallel combination control (PCC) mode, both the (scaled) defect and the (scaled) GE estimate are obtained for each numerical solution, and a linear combination of the two is used as input to the mesh refinement

	Tol	Time	% Total	Actual Error	Estimated Error	τ
RE	1E–4	2.84×10^{-2}	12.00 %	3.350×10^{-12}	3.355×10^{-12}	5.496×10^{-15}
	1E–5	2.88×10^{-2}	11.63 %	3.078×10^{-12}	3.078×10^{-12}	3.755×10^{-16}
	1E–6	2.90×10^{-2}	10.85 %	2.949×10^{-12}	2.948×10^{-12}	6.299×10^{-16}
	1E–7	2.88×10^{-2}	10.30 %	2.888×10^{-12}	2.887×10^{-12}	5.757×10^{-16}
	1E–8	2.80×10^{-2}	9.34 %	2.842×10^{-12}	2.841×10^{-12}	6.191×10^{-16}
CO	1E–4	2.10×10^{-3}	0.89 %	3.350×10^{-12}	1.956×10^{2}	1.956×10^{2}
	1E–5	2.90×10^{-3}	1.17 %	3.078×10^{-12}	7.233	7.233
	1E–6	2.10×10^{-3}	0.79 %	2.949×10^{-12}	2.605×10^{-1}	2.605×10^{-1}
	1E–7	3.60×10^{-3}	1.29 %	2.888×10^{-12}	5.180×10^{-2}	5.180×10^{-2}
	1E–8	2.00×10^{-3}	0.67 %	2.842×10^{-12}	1.043×10^{-1}	1.043×10^{-1}
HO	1E–4	5.60×10^{-3}	2.37 %	3.350×10^{-12}	3.349×10^{-12}	5.684×10^{-16}
	1E–5	5.90×10^{-3}	2.38 %	3.078×10^{-12}	3.078×10^{-12}	3.129×10^{-16}
	1E–6	6.20×10^{-3}	2.32 %	2.949×10^{-12}	2.948×10^{-12}	4.757×10^{-16}
	1E–7	6.30×10^{-3}	2.25 %	2.888×10^{-12}	2.888×10^{-12}	4.059×10^{-16}
	1E–8	6.10×10^{-3}	2.03 %	2.842×10^{-12}	2.841×10^{-12}	4.722×10^{-16}
DC	1E–4	1.03×10^{-2}	4.35 %	3.350×10^{-12}	3.441×10^{-12}	9.167×10^{-14}
	1E–5	1.00×10^{-2}	4.04 %	3.078×10^{-12}	3.079×10^{-12}	3.129×10^{-16}
	1E–6	1.00×10^{-2}	3.74 %	2.949×10^{-12}	2.948×10^{-12}	4.757×10^{-16}
	1E–7	1.01×10^{-2}	3.61 %	2.888×10^{-12}	2.888×10^{-12}	3.807×10^{-16}
	1E–8	1.02×10^{-2}	3.40 %	2.842×10^{-12}	2.841×10^{-12}	4.722×10^{-16}

Table 5: GE Estimates applied to (15), MIRK order 4. Number of points in the final mesh for each tolerance is 4609 in all cases.

process and the termination criterion. For this preliminary investigation, we consider only the simple sum of the scaled defect and the scaled GE estimate. In this case, *the resultant numerical solution has a defect estimate and a GE estimate that both satisfy the user tolerance, and the numerical solution can also be said to be the exact solution to a BVODE that is a perturbation (on the order of the tolerance) of the original BVODE.*

In Table 10 we report results for test problem (14) with $\epsilon = 0.01$ and using MIRK formula order 2. In Table 11 we report results for test problem (14) with $\epsilon = 0.01$ and using MIRK formula order 4. In Table 12 we report results for test problem (14) with $\epsilon = 0.01$ and using MIRK formula order 6. In all three cases, the cost of computing a defect controlled solution is greater than the cost of computing a GE controlled solution for coarser tolerances and only slightly smaller for sharper tolerances. The estimated GE for the defect controlled solution is less than the tolerance, and thus in SCC mode the code does no

	Tol	Time	% Total	Actual Error	Estimated Error	τ
RE	1E–4	2.00×10^{-3}	7.82 %	7.339×10^{-4}	8.108×10^{-4}	7.686×10^{-5}
	1E–5	3.52×10^{-3}	12.14 %	3.215×10^{-5}	3.218×10^{-5}	3.453×10^{-8}
	1E–6	5.37×10^{-3}	13.86 %	9.032×10^{-5}	8.975×10^{-5}	5.669×10^{-7}
	1E–7	1.09×10^{-2}	19.95 %	1.071×10^{-5}	1.070×10^{-5}	1.504×10^{-8}
	1E–8	1.94×10^{-2}	21.34 %	7.346×10^{-7}	7.341×10^{-7}	5.033×10^{-10}
CO	1E–4	4.30×10^{-4}	1.68 %	7.339×10^{-4}	3.524×10^{6}	3.524×10^{6}
	1E–5	1.07×10^{-3}	3.70 %	3.215×10^{-5}	7.554×10^{5}	7.554×10^{5}
	1E–6	2.13×10^{-3}	5.50 %	9.032×10^{-5}	2.572×10^{5}	2.572×10^{5}
	1E–7	1.74×10^{-3}	3.17 %	1.071×10^{-5}	6.239×10^{4}	6.239×10^{4}
	1E–8	3.72×10^{-3}	4.09 %	7.346×10^{-7}	4.743×10^{3}	4.743×10^{3}
HO	1E–4	1.70×10^{-4}	0.67 %	7.339×10^{-4}	7.143×10^{-4}	1.955×10^{-5}
	1E–5	4.43×10^{-4}	1.53 %	3.215×10^{-5}	3.269×10^{-5}	5.413×10^{-7}
	1E–6	3.17×10^{-4}	0.82 %	9.032×10^{-5}	9.159×10^{-5}	1.265×10^{-6}
	1E–7	1.22×10^{-3}	2.23 %	1.071×10^{-5}	1.075×10^{-5}	3.693×10^{-8}
	1E–8	3.56×10^{-3}	3.92 %	7.346×10^{-7}	7.354×10^{-7}	7.140×10^{-10}
DC	1E–4	2.93×10^{-4}	1.15 %	7.339×10^{-4}	7.143×10^{-4}	1.956×10^{-5}
	1E–5	5.47×10^{-4}	1.89 %	3.215×10^{-5}	3.269×10^{-5}	5.413×10^{-7}
	1E–6	8.76×10^{-4}	2.26 %	9.032×10^{-5}	9.159×10^{-5}	1.265×10^{-6}
	1E–7	1.02×10^{-3}	1.86 %	1.071×10^{-5}	1.075×10^{-5}	3.693×10^{-8}
	1E–8	9.19×10^{-4}	1.01 %	7.346×10^{-7}	7.354×10^{-7}	7.140×10^{-10}

Table 6: GE estimation applied to (16), MIRK order 4. Number of points in the final mesh for each tolerance, respectively: 125, 224, 345, 595, and 1105.

extra adaptation — it simply stops after determining that the estimated GE is less than the tolerance. We observe that the final mesh is the same as the final mesh used in the defect control case. The cost of running in SCC mode is greater than the cost of running in GE control mode for coarser tolerances and only slightly smaller for sharper tolerances. For this test problem, there is thus no significant advantage to using SCC mode. The costs for PCC mode are slightly higher than those for either defect control mode or GE control mode.

In Table 13 we report results for test problem (16) with $\epsilon = 0.005$ and MIRK formula order 2. In this case, the cost of computing a defect controlled solution is significantly less than the cost of computing a GE controlled solution. As well, the estimated GE for the defect controlled solutions is greater than the corresponding tolerance. In the SCC case, substantial additional computation is required to go from the defect controlled solution to the GE controlled solution, and the costs are greater than for direct GE control. On the other hand, except

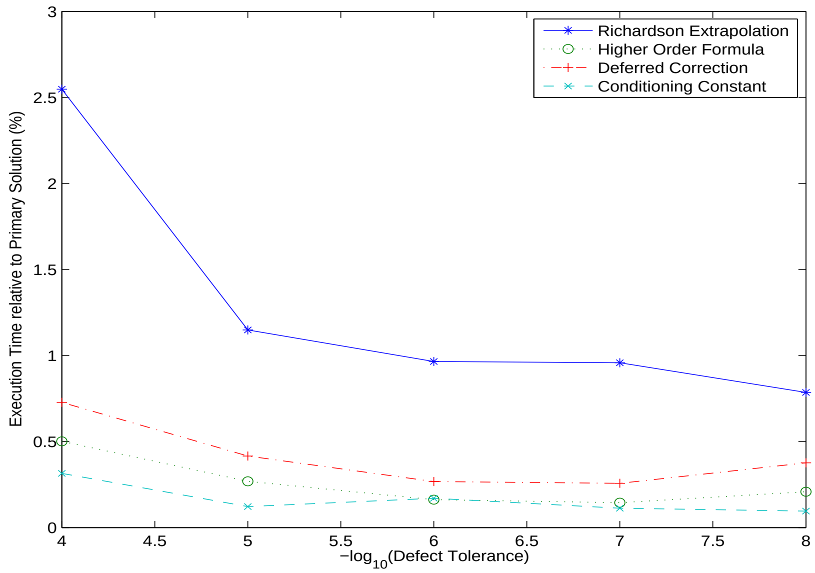

Figure 7: Relative execution time of global error estimators vs. $-\log_{10}$ of defect tolerance with fourth order MIRK formula for test problem (14).

for the sharpest tolerance, *the solution obtained through PCC control costs less than the computation in which only the GE estimate is controlled*, suggesting that the inclusion of information about the defect is of some benefit to GE control. In Table 14 we report results for test problem (16) with $\epsilon = 0.005$ and MIRK formula order 4. The cost of computing a defect controlled solution is less than or equal to the cost of computing a GE controlled solution. The estimated GE for the defect controlled solutions is less than the tolerance. *In SCC mode, except for the two coarsest tolerances, the costs for obtaining a GE controlled solution are less than for the direct GE control case.* The solution obtained through PCC control costs somewhat more than the other cases. In Table 15 we report results for test problem (16) with $\epsilon = 0.005$ and MIRK formula order 6. The cost of computing a defect controlled solution is somewhat less than the cost of computing a GE controlled solution. Except for the sharpest tolerance, the estimated GE for the defect controlled solutions is less than the corresponding tolerance. *In SCC mode, except for the coarsest tolerance, the costs for obtaining a GE controlled solution are less than for the direct GE control case.* The solution obtained through PCC control costs approximately the same as the computation in which only the GE estimate is controlled.

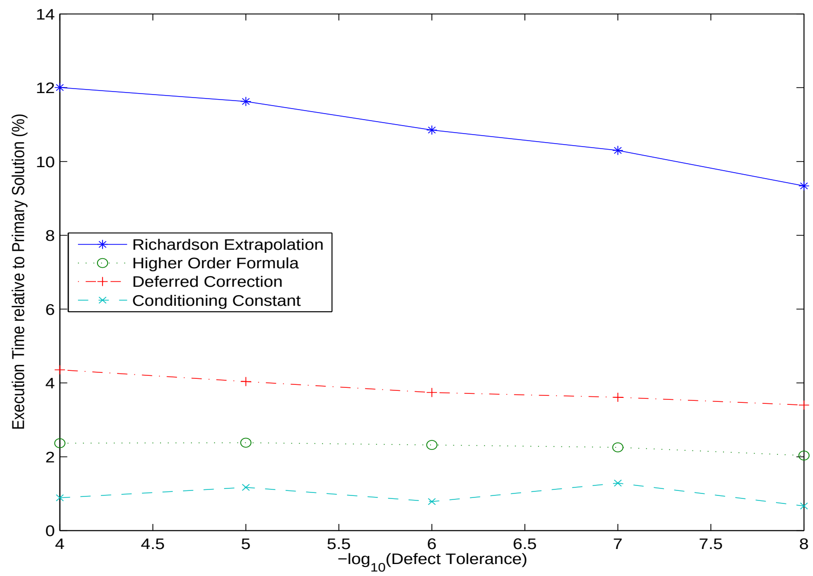

Figure 8: Relative execution time of global error estimators vs. $-\log_{10}$ of defect tolerance with fourth order MIRK formula for test problem (15).

5.2 Use of BVP_SOLVER in GE Control Mode on a Problem with a Pseudosolution

In this subsection we briefly consider the application of BVP_SOLVER in GE control mode to a problem that has a pseudosolution.

The BVODE [14]

$$y''(x) + |y(x)| = 0, \quad 0 < x < \pi, \quad y(0) = 0, \quad y(\pi) = y_\pi, \tag{17}$$

has a unique solution for $y_\pi < 0$, infinitely many solutions for $y_\pi = 0$, and no solution for $y_\pi > 0$. We first run BVP_SOLVER in its original defect control mode and are able to obtain two pseudosolutions when $y_\pi = 0.001$. We obtain one pseudosolution with the second order MIRK method and a second pseudosolution with the fourth order MIRK method. In both cases, a tolerance of 10^{-6} is used and an initial guess of $y(x) = 1.0$ and $y'(x) = 0.0$ for $0 \le x \le \pi$ is provided. For both orders, BVP_SOLVER indicates that it finds a solution with a defect norm well below the tolerance. However, if we employ the option within BVP_SOLVER to compute an a posteriori GE estimate (using RE) we find that for both MIRK orders the estimated GE is quite large, signaling the presence of a pseudosolution; see Table 16. When we attempt to use BVP_SOLVER in defect control mode using the sixth order MIRK method to solve (17) with $y_\pi = 0.001$,

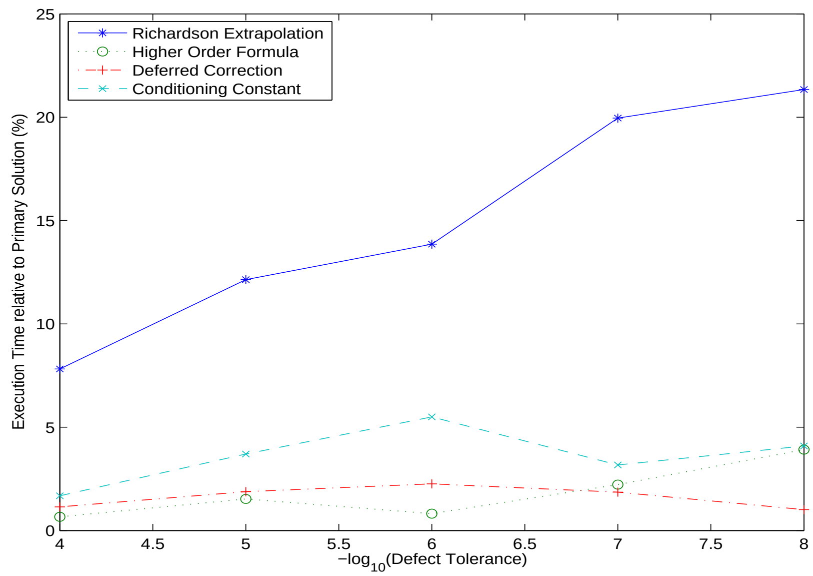

Figure 9: Relative execution time of global error estimators vs. $-\log_{10}$ of defect tolerance with fourth order MIRK formula for test problem (16).

the Newton iteration fails to converge and even a defect controlled numerical solution cannot be obtained.

We next tried using `BVP_SOLVER` *in GE control mode* to solve (17) with $y_\pi = 0.001$, using second and fourth order MIRK methods, with a tolerance of 10^{-6}, i.e., the cases that yield pseudosolutions in defect control mode. We found that `BVP_SOLVER` in GE control mode was unable to significantly reduce the GE even using a million mesh points and thus, appropriately, is not able to obtain a GE controlled numerical solution.

6 Conclusions

6.1 Conclusions: Alternative GE Estimators

In this report we have discussed the efficient implementation of three well known approaches to estimating the GE of the numerical solution of a BVODE within a defect control solver. We have also considered an approach for obtaining a bound on the GE that is based on an estimate of the defect and an estimate of a conditioning constant for the BVODE. The approaches based on the HO scheme and the DC scheme are generally less expensive than the approach based on the RE scheme while achieving a GE estimate with the same overall quality. The CO

	Tol	Time	% Total	Actual Error	Estimated Error	τ
RE	1E–4	9.04×10^{-3}	2.45 %	1.072×10^{-7}	1.087×10^{-7}	1.526×10^{-9}
	1E–5	9.34×10^{-3}	2.51 %	1.826×10^{-8}	1.831×10^{-8}	5.280×10^{-11}
	1E–6	8.76×10^{-3}	2.35 %	1.502×10^{-8}	1.500×10^{-8}	1.759×10^{-11}
	1E–7	4.26×10^{-3}	1.15 %	1.525×10^{-12}	1.530×10^{-12}	4.845×10^{-15}
	1E–8	4.78×10^{-3}	1.28 %	1.521×10^{-12}	1.526×10^{-12}	4.804×10^{-15}
CO	1E–4	6.60×10^{-4}	0.18 %	1.072×10^{-7}	1.926	1.926
	1E–5	7.60×10^{-4}	0.20 %	1.826×10^{-8}	1.056×10^{-1}	1.056×10^{-1}
	1E–6	8.00×10^{-4}	0.21 %	1.502×10^{-8}	1.041×10^{-1}	1.041×10^{-1}
	1E–7	3.60×10^{-4}	0.10 %	1.525×10^{-12}	2.724×10^{-5}	2.724×10^{-5}
	1E–8	4.40×10^{-4}	0.12 %	1.521×10^{-12}	2.783×10^{-5}	2.783×10^{-5}
HO	1E–4	2.20×10^{-3}	0.60 %	1.072×10^{-7}	1.050×10^{-7}	2.175×10^{-9}
	1E–5	1.90×10^{-3}	0.51 %	1.826×10^{-8}	1.780×10^{-8}	4.649×10^{-10}
	1E–6	3.10×10^{-3}	0.83 %	1.502×10^{-8}	1.485×10^{-8}	1.635×10^{-10}
	1E–7	1.08×10^{-3}	0.29 %	1.525×10^{-12}	1.530×10^{-12}	4.616×10^{-15}
	1E–8	9.80×10^{-4}	0.26 %	1.521×10^{-12}	1.526×10^{-12}	4.770×10^{-15}
DC	1E–4	2.16×10^{-3}	0.59 %	1.072×10^{-7}	2.070×10^{-7}	9.978×10^{-8}
	1E–5	2.86×10^{-3}	0.77 %	1.826×10^{-8}	1.973×10^{-8}	1.470×10^{-9}
	1E–6	2.30×10^{-3}	0.62 %	1.502×10^{-8}	1.386×10^{-8}	1.163×10^{-9}
	1E–7	1.40×10^{-3}	0.38 %	1.525×10^{-12}	1.525×10^{-12}	2.810×10^{-16}
	1E–8	9.80×10^{-4}	0.26 %	1.521×10^{-12}	1.521×10^{-12}	2.004×10^{-16}

Table 7: GE Estimates applied to (14), MIRK order 6. Number of points in the final mesh for each tolerance, respectively: 1153, 1153, 1153, 577, and 577.

approach generally has a low cost but does not have good accuracy. Nonetheless, a bound on the conditioning constant may be useful for the detection of ill-conditioning for a given BVODE [14].

We can draw the following conclusions from the results presented in this report:

(i) The a posteriori GE estimation employed by `BVP_SOLVER` should be based on the HO or DC estimate rather than the RE estimate.

(ii) The CO approach provides a less accurate estimate of the GE because the estimate of the conditioning constant does not provide a tight upper bound. When one employs `BVP_SOLVER` with the option to compute an estimate of the GE, our results suggest that one can then obtain a better estimate of the conditioning constant by using the GE estimate and the

	Tol	Time	% Total	Actual Error	Estimated Error	τ
RE	1E–4	1.46×10^{-1}	17.14 %	7.694×10^{-5}	7.697×10^{-5}	2.942×10^{-8}
	1E–5	1.45×10^{-1}	16.72 %	1.058×10^{-5}	1.058×10^{-5}	2.015×10^{-9}
	1E–6	7.27×10^{-2}	8.64 %	7.411×10^{-16}	3.750×10^{-16}	3.661×10^{-16}
	1E–7	7.68×10^{-2}	8.42 %	1.108×10^{-15}	2.125×10^{-16}	8.959×10^{-16}
	1E–8	7.62×10^{-2}	7.84 %	1.055×10^{-15}	2.246×10^{-16}	8.308×10^{-16}
CO	1E–4	1.87×10^{-2}	2.20 %	7.694×10^{-5}	3.807×10^{7}	3.807×10^{7}
	1E–5	1.74×10^{-2}	2.01 %	1.058×10^{-5}	5.708×10^{6}	5.708×10^{6}
	1E–6	4.90×10^{-3}	0.58 %	7.411×10^{-16}	1.200×10^{-2}	1.200×10^{-2}
	1E–7	4.50×10^{-3}	0.49 %	1.108×10^{-15}	1.618×10^{-4}	1.618×10^{-4}
	1E–8	5.50×10^{-3}	0.57 %	1.055×10^{-15}	1.309×10^{-4}	1.309×10^{-4}
HO	1E–4	3.22×10^{-2}	3.78 %	7.694×10^{-5}	7.755×10^{-5}	6.109×10^{-7}
	1E–5	3.10×10^{-2}	3.57 %	1.058×10^{-5}	1.063×10^{-5}	4.578×10^{-8}
	1E–6	1.64×10^{-2}	1.95 %	7.411×10^{-16}	4.217×10^{-16}	3.194×10^{-16}
	1E–7	1.73×10^{-2}	1.90 %	1.108×10^{-15}	2.149×10^{-16}	8.936×10^{-16}
	1E–8	1.77×10^{-2}	1.82 %	1.055×10^{-15}	2.173×10^{-16}	8.381×10^{-16}
DC	1E–4	4.99×10^{-2}	5.86 %	7.694×10^{-5}	7.755×10^{-5}	6.109×10^{-7}
	1E–5	4.90×10^{-2}	5.65 %	1.058×10^{-5}	1.063×10^{-5}	4.578×10^{-8}
	1E–6	2.42×10^{-2}	2.87 %	7.411×10^{-16}	4.409×10^{-16}	3.003×10^{-16}
	1E–7	2.60×10^{-2}	2.85 %	1.108×10^{-15}	2.218×10^{-16}	8.867×10^{-16}
	1E–8	2.62×10^{-2}	2.70 %	1.055×10^{-15}	2.218×10^{-16}	8.336×10^{-16}

Table 8: GE Estimates applied to (15), MIRK order 6. Number of points in the final mesh for each tolerance, respectively: 15213, 15424, 7752, 7787, and 7825.

defect estimate; i.e., rewriting (13), we get

$$\frac{||\mathbf{y}(x) - \mathbf{S}(x)||_{\mathbf{W}_3}}{\max(||\boldsymbol{\delta}(x)||_{\mathbf{W}_1}, ||\boldsymbol{\sigma}||_{\mathbf{W}_2})} \leq \kappa,$$

giving a lower bound on κ.

(iii) The results presented in this report may also be relevant for BVODE *GE control solvers*. In particular, it may be possible to improve the efficiency of the GE estimation approach employed by `COLSYS`[1] or `COLNEW`[3] because these solvers employ RE for one type of GE estimation. It may be worthwhile to investigate the use of the HO or DC approach, with appropriate modifications, within these solvers. It may be possible to obtain a higher order approximate solution using a higher order collocation method

	Tol	Time	% Total	Actual Error	Estimated Error	τ
RE	1E–4	8.17×10^{-4}	8.21 %	6.175×10^{-4}	6.226×10^{-4}	5.102×10^{-6}
	1E–5	1.01×10^{-3}	9.77 %	1.536×10^{-5}	1.530×10^{-5}	6.543×10^{-8}
	1E–6	2.98×10^{-3}	22.83 %	9.056×10^{-7}	9.061×10^{-7}	5.469×10^{-10}
	1E–7	3.20×10^{-3}	23.17 %	3.393×10^{-7}	3.393×10^{-7}	1.961×10^{-11}
	1E–8	6.01×10^{-3}	26.05 %	9.737×10^{-8}	9.734×10^{-8}	3.197×10^{-11}
CO	1E–4	1.03×10^{-4}	1.04 %	6.175×10^{-4}	1.073×10^{5}	1.073×10^{5}
	1E–5	2.50×10^{-4}	2.41 %	1.536×10^{-5}	9.898×10^{3}	9.898×10^{3}
	1E–6	4.00×10^{-4}	3.07 %	9.056×10^{-7}	1.459×10^{3}	1.459×10^{3}
	1E–7	5.50×10^{-4}	3.99 %	3.393×10^{-7}	5.966×10^{2}	5.966×10^{2}
	1E–8	9.80×10^{-4}	4.24 %	9.737×10^{-8}	3.206×10^{2}	3.206×10^{2}
HO	1E–4	7.67×10^{-5}	0.77 %	6.175×10^{-4}	6.305×10^{-4}	1.302×10^{-5}
	1E–5	1.00×10^{-4}	0.96 %	1.536×10^{-5}	1.525×10^{-5}	1.128×10^{-7}
	1E–6	2.23×10^{-4}	1.71 %	9.056×10^{-7}	9.076×10^{-7}	2.007×10^{-9}
	1E–7	3.70×10^{-4}	2.68 %	3.393×10^{-7}	3.395×10^{-7}	1.491×10^{-10}
	1E–8	4.93×10^{-4}	2.14 %	9.737×10^{-8}	9.731×10^{-8}	6.310×10^{-11}
DC	1E–4	6.00×10^{-5}	0.60 %	6.175×10^{-4}	6.305×10^{-4}	1.302×10^{-5}
	1E–5	1.37×10^{-4}	1.32 %	1.536×10^{-5}	1.525×10^{-5}	1.128×10^{-7}
	1E–6	4.03×10^{-4}	3.09 %	9.056×10^{-7}	9.076×10^{-7}	2.007×10^{-9}
	1E–7	4.99×10^{-4}	3.62 %	3.393×10^{-7}	3.395×10^{-7}	1.491×10^{-10}
	1E–8	8.03×10^{-4}	3.48 %	9.737×10^{-8}	9.731×10^{-8}	6.310×10^{-11}

Table 9: GE estimation applied to (16), MIRK order 6. Number of points in the final mesh for each tolerance, respectively: 50, 76, 104, 136, and 213.

applied on the final mesh from the computation of the primary collocation solution.

6.2 Conclusions: GE/Defect Control Modes

Because both the GE and the defect provide valid measures of solution quality, the results presented here suggest that it may be worthwhile to have a BVODE solver that can employ either GE control, defect control, or a hybrid GE/defect control strategy. In particular, some of the results indicate that a hybrid control scheme can yield a GE controlled numerical solution more efficiently than a scheme that controls only the GE.

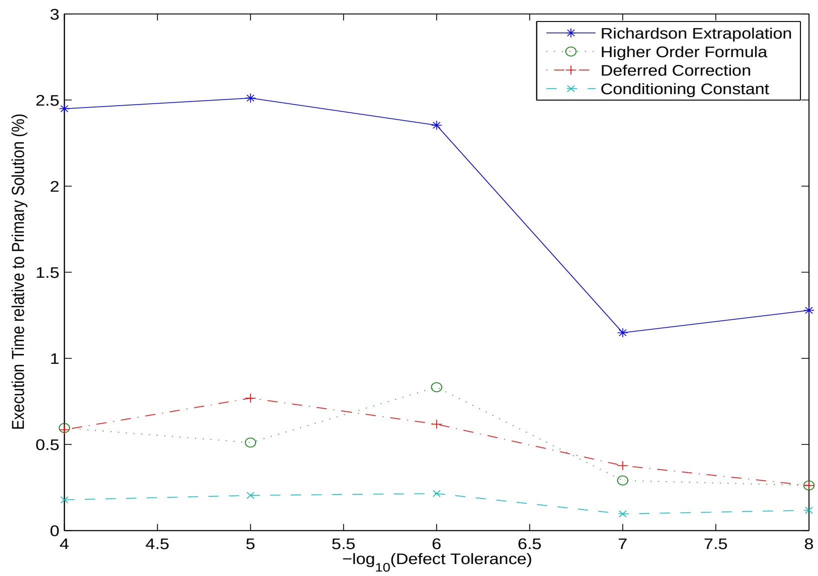

Figure 10: Relative execution time of global error estimators vs. $-\log_{10}$ of defect tolerance with sixth order MIRK formula for test problem (14).

References

[1] Ascher, U. M., Christiansen, J., and Russell, R. D. Collocation software for boundary value odes. *ACM Trans. Math. Softw. 7* (1981), 209–222.

[2] Ascher, U. M., Mattheij, R. M. M., and Russell, R. D. *Numerical solution of boundary value problems for ordinary differential equations*, vol. 13 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995. Corrected reprint of the 1988 original.

[3] Bader, G., and Ascher, U. A new basis implementation for a mixed order boundary value ODE solver. *SIAM J. Sci. Statist. Comput. 8*, 4 (1987), 483–500.

[4] Cash, J. R., and Mazzia, F. Hybrid mesh selection algorithms based on conditioning for two-point boundary value problems. *JNAIAM J. Numer. Anal. Ind. Appl. Math. 1*, 1 (2006), 81–90.

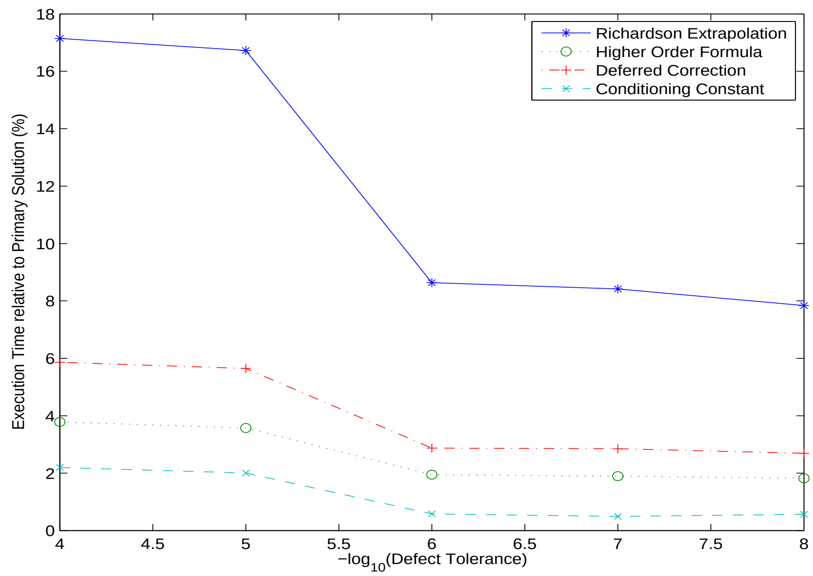

Figure 11: Relative execution time of global error estimators vs. $-\log_{10}$ of defect tolerance with sixth order MIRK formula for test problem (15).

[5] Cash, J. R., and Wright, M. H. A deferred correction method for nonlinear two-point boundary value problems: implementation and numerical evaluation. *SIAM J. Sci. Statist. Comput. 12*, 4 (1991), 971–989.

[6] Enright, W. H., and Muir, P. H. Runge-Kutta software with defect control for boundary value ODEs. *SIAM J. Sci. Comput. 17*, 2 (1996), 479–497.

[7] Enright, W. H., and Muir, P. H. New interpolants for asymptotically correct defect control of BVODEs. *Numer. Algorithms 53*, 2-3 (2010), 219–238.

[8] Gupta, S. An adaptive boundary value Runge-Kutta solver for first order boundary value problems. *SIAM J. Numer. Anal. 22*, 1 (1985), 114–126.

[9] Higham, N. J. FORTRAN codes for estimating the one-norm of a real or complex matrix, with applications to condition estimation. *ACM Trans. Math. Software 14*, 4 (1988), 381–396 (1989).

[10] Kierzenka, J., and Shampine, L. F. A BVP solver based on residual control and the MATLAB PSE. *ACM Trans. Math. Software 27*, 3 (2001), 299–316.

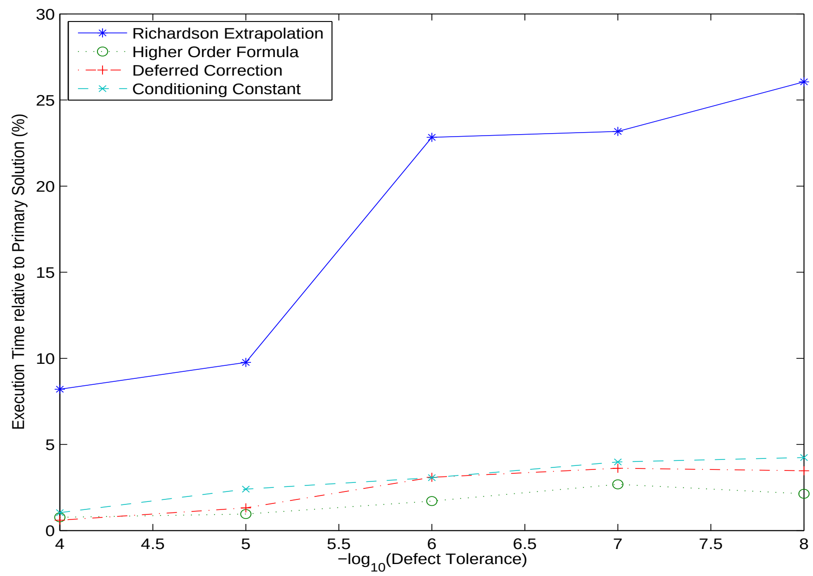

Figure 12: Relative execution time of global error estimators vs. $-\log_{10}$ of defect tolerance with sixth order MIRK formula for test problem (16).

[11] Muir, P., and Owren, B. Order barriers and characterizations for continuous mono-implicit Runge-Kutta schemes. *Math. Comp. 61*, 204 (13), 675–699.

[12] Muir, P. H. Optimal discrete and continuous mono-implicit Runge-Kutta schemes for BVODEs. *Adv. Comput. Math. 10*, 2 (1999), 135–167.

[13] Russell, R. D., and Christiansen, J. Adaptive mesh selection strategies for solving boundary value problems. *SIAM J. Numer. Anal. 15*, 1 (1978), 59–80.

[14] Shampine, L. F., and Muir, P. H. Estimating conditioning of BVPs for ODEs. *Math. Comput. Modelling 40*, 11-12 (2004), 1309–1321.

[15] Shampine, L. F., Muir, P. H., and Xu, H. A user-friendly Fortran BVP solver. *JNAIAM J. Numer. Anal. Ind. Appl. Math. 1*, 2 (2006), 201–217.

DefC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	1.128×10^{-2}	662	4.233×10^{-5}	5.597×10^{-5}	7.953×10^{-6}	7.872×10^{-6}
1×10^{-5}	3.550×10^{-2}	2193	5.279×10^{-6}	6.980×10^{-6}	7.163×10^{-7}	7.091×10^{-7}
1×10^{-6}	7.780×10^{-2}	6015	9.289×10^{-7}	1.228×10^{-6}	9.365×10^{-8}	9.380×10^{-8}
1×10^{-7}	1.440×10^{-1}	16067	9.852×10^{-8}	1.303×10^{-7}	1.334×10^{-8}	1.312×10^{-8}
1×10^{-8}	4.759×10^{-1}	51236	6.326×10^{-9}	8.366×10^{-9}	1.297×10^{-9}	1.287×10^{-9}
GEC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	1.030×10^{-2}	541	6.380×10^{-3}	8.408×10^{-3}	4.045×10^{-5}	2.774×10^{-5}
1×10^{-5}	3.004×10^{-2}	1859	1.035×10^{-3}	1.366×10^{-3}	3.408×10^{-6}	3.540×10^{-6}
1×10^{-6}	3.938×10^{-2}	2709	1.152×10^{-4}	1.523×10^{-4}	8.999×10^{-7}	8.565×10^{-7}
1×10^{-7}	5.236×10^{-2}	5069	1.309×10^{-5}	1.730×10^{-5}	9.750×10^{-8}	9.751×10^{-8}
1×10^{-8}	2.816×10^{-1}	31165	4.780×10^{-7}	6.320×10^{-7}	2.565×10^{-9}	2.565×10^{-9}
SCC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	1.188×10^{-2}	662	4.233×10^{-5}	5.597×10^{-5}	7.953×10^{-6}	7.872×10^{-6}
1×10^{-5}	3.724×10^{-2}	2193	5.279×10^{-6}	6.980×10^{-6}	7.163×10^{-7}	7.091×10^{-7}
1×10^{-6}	8.630×10^{-2}	6015	9.289×10^{-7}	1.228×10^{-6}	9.365×10^{-8}	9.380×10^{-8}
1×10^{-7}	1.662×10^{-1}	16067	9.852×10^{-8}	1.303×10^{-7}	1.334×10^{-8}	1.312×10^{-8}
1×10^{-8}	5.500×10^{-1}	51236	6.326×10^{-9}	8.366×10^{-9}	1.297×10^{-9}	1.287×10^{-9}
PCC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	1.498×10^{-2}	1291	8.098×10^{-5}	1.071×10^{-4}	1.705×10^{-6}	1.706×10^{-6}
1×10^{-5}	5.044×10^{-2}	4293	3.783×10^{-6}	5.002×10^{-6}	1.600×10^{-7}	1.600×10^{-7}
1×10^{-6}	1.212×10^{-1}	7470	7.699×10^{-7}	1.018×10^{-6}	4.345×10^{-8}	4.252×10^{-8}
1×10^{-7}	3.116×10^{-1}	33647	4.738×10^{-8}	6.265×10^{-8}	2.345×10^{-9}	2.386×10^{-9}
1×10^{-8}	1.359	147457	4.514×10^{-9}	5.969×10^{-9}	2.433×10^{-10}	2.393×10^{-10}

Table 10: Error control methods applied to (14) with $\epsilon = 0.01$, MIRK order 2.

DefC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	4.657×10^{-3}	62	4.434×10^{-5}	8.908×10^{-5}	4.381×10^{-6}	4.750×10^{-6}
10^{-5}	6.566×10^{-3}	106	2.385×10^{-6}	5.813×10^{-6}	3.594×10^{-7}	3.930×10^{-7}
10^{-6}	6.927×10^{-3}	191	6.012×10^{-7}	1.208×10^{-6}	2.703×10^{-8}	2.874×10^{-8}
10^{-7}	8.178×10^{-3}	281	9.429×10^{-8}	1.894×10^{-7}	5.358×10^{-9}	5.651×10^{-9}
10^{-8}	9.947×10^{-3}	485	9.264×10^{-9}	1.861×10^{-8}	5.649×10^{-10}	5.831×10^{-10}
GEC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	3.660×10^{-3}	47	8.676×10^{-4}	1.648×10^{-3}	4.309×10^{-5}	4.931×10^{-5}
10^{-5}	4.310×10^{-3}	83	9.105×10^{-5}	1.757×10^{-4}	4.237×10^{-6}	4.729×10^{-6}
10^{-6}	5.243×10^{-3}	145	1.266×10^{-5}	2.452×10^{-5}	4.513×10^{-7}	4.857×10^{-7}
10^{-7}	9.898×10^{-3}	303	2.675×10^{-5}	5.374×10^{-5}	5.942×10^{-8}	1.207×10^{-7}
10^{-8}	1.009×10^{-2}	529	2.037×10^{-6}	4.093×10^{-6}	4.668×10^{-9}	9.050×10^{-9}
SCC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	5.480×10^{-3}	62	4.434×10^{-5}	8.908×10^{-5}	4.381×10^{-6}	4.750×10^{-6}
10^{-5}	6.699×10^{-3}	106	2.385×10^{-6}	5.813×10^{-6}	3.594×10^{-7}	3.930×10^{-7}
10^{-6}	8.236×10^{-3}	191	6.012×10^{-7}	1.208×10^{-6}	2.703×10^{-8}	2.874×10^{-8}
10^{-7}	9.631×10^{-3}	281	9.429×10^{-8}	1.894×10^{-7}	5.358×10^{-9}	5.651×10^{-9}
10^{-8}	9.997×10^{-3}	485	9.264×10^{-9}	1.861×10^{-8}	5.649×10^{-10}	5.831×10^{-10}
PCC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	4.800×10^{-3}	110	2.544×10^{-5}	4.665×10^{-5}	1.916×10^{-6}	2.122×10^{-6}
10^{-5}	5.403×10^{-3}	145	8.356×10^{-6}	1.614×10^{-5}	6.310×10^{-7}	6.685×10^{-7}
10^{-6}	9.759×10^{-3}	235	4.847×10^{-7}	9.737×10^{-7}	5.700×10^{-9}	5.780×10^{-9}
10^{-7}	9.955×10^{-3}	403	4.278×10^{-8}	8.594×10^{-8}	6.297×10^{-10}	6.347×10^{-10}
10^{-8}	1.084×10^{-2}	1047	7.168×10^{-9}	1.440×10^{-8}	1.376×10^{-11}	3.628×10^{-11}

Table 11: Error control methods applied to (14) with $\epsilon = 0.01$, MIRK order 4.

DefC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	5.133×10^{-3}	32	2.392×10^{-5}	7.236×10^{-5}	1.194×10^{-6}	1.594×10^{-6}
1×10^{-5}	4.850×10^{-3}	40	7.391×10^{-6}	8.207×10^{-6}	4.671×10^{-7}	4.700×10^{-7}
1×10^{-6}	6.419×10^{-3}	68	2.470×10^{-7}	2.728×10^{-7}	1.529×10^{-8}	1.622×10^{-8}
1×10^{-7}	6.238×10^{-3}	81	1.991×10^{-8}	2.308×10^{-8}	1.441×10^{-9}	1.569×10^{-9}
1×10^{-8}	7.042×10^{-3}	116	7.742×10^{-9}	7.910×10^{-9}	1.464×10^{-10}	1.516×10^{-10}
GEC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	4.797×10^{-3}	19	2.197×10^{-2}	2.217×10^{-2}	2.788×10^{-5}	5.660×10^{-5}
1×10^{-5}	4.917×10^{-3}	25	6.023×10^{-4}	6.078×10^{-4}	5.129×10^{-6}	5.191×10^{-6}
1×10^{-6}	5.193×10^{-3}	40	2.713×10^{-5}	2.820×10^{-5}	2.644×10^{-7}	2.898×10^{-7}
1×10^{-7}	5.396×10^{-3}	51	5.954×10^{-6}	6.225×10^{-6}	5.824×10^{-8}	6.416×10^{-8}
1×10^{-8}	4.446×10^{-3}	75	5.409×10^{-7}	5.727×10^{-7}	5.477×10^{-9}	6.394×10^{-9}
SCC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	5.340×10^{-3}	32	2.392×10^{-5}	7.236×10^{-5}	1.194×10^{-6}	1.594×10^{-6}
1×10^{-5}	4.967×10^{-3}	40	7.391×10^{-6}	8.207×10^{-6}	4.671×10^{-7}	4.700×10^{-7}
1×10^{-6}	6.675×10^{-3}	68	2.470×10^{-7}	2.728×10^{-7}	1.529×10^{-8}	1.622×10^{-8}
1×10^{-7}	6.531×10^{-3}	81	1.991×10^{-8}	2.308×10^{-8}	1.441×10^{-9}	1.569×10^{-9}
1×10^{-8}	7.521×10^{-3}	116	7.742×10^{-9}	7.910×10^{-9}	1.464×10^{-10}	1.516×10^{-10}
PCC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	5.607×10^{-3}	42	9.438×10^{-6}	1.278×10^{-5}	2.686×10^{-7}	3.093×10^{-7}
1×10^{-5}	6.073×10^{-3}	58	8.441×10^{-7}	8.911×10^{-7}	3.786×10^{-8}	3.824×10^{-8}
1×10^{-6}	6.752×10^{-3}	82	1.048×10^{-7}	1.245×10^{-7}	4.577×10^{-9}	4.806×10^{-9}
1×10^{-7}	7.674×10^{-3}	116	1.430×10^{-8}	1.563×10^{-8}	5.476×10^{-10}	5.465×10^{-10}
1×10^{-8}	8.543×10^{-3}	145	3.796×10^{-9}	3.893×10^{-9}	1.396×10^{-10}	1.393×10^{-10}

Table 12: Error control methods applied to (14) with $\epsilon = 0.01$, MIRK order 6.

DefC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	1.828×10^{-2}	935	6.430×10^{-5}	8.232×10^{-5}	4.303×10^{-4}	4.303×10^{-4}
10^{-5}	4.608×10^{-2}	2621	8.181×10^{-6}	1.069×10^{-5}	5.409×10^{-5}	5.460×10^{-5}
10^{-6}	1.664×10^{-1}	8491	6.788×10^{-7}	8.945×10^{-7}	5.177×10^{-6}	5.188×10^{-6}
10^{-7}	5.794×10^{-1}	27546	6.140×10^{-8}	8.110×10^{-8}	4.922×10^{-7}	4.924×10^{-7}
10^{-8}	1.218	71641	9.405×10^{-9}	1.243×10^{-8}	7.278×10^{-8}	7.280×10^{-8}
GEC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	2.086×10^{-1}	9217	1.551×10^{-3}	1.786×10^{-3}	5.354×10^{-5}	5.354×10^{-5}
10^{-5}	8.790×10^{-1}	36865	1.117×10^{-4}	1.417×10^{-4}	3.333×10^{-6}	3.333×10^{-6}
10^{-6}	1.772	73729	2.871×10^{-5}	3.714×10^{-5}	8.333×10^{-7}	8.333×10^{-7}
10^{-7}	7.218	294913	1.832×10^{-6}	2.410×10^{-6}	5.208×10^{-8}	5.208×10^{-8}
10^{-8}	2.912×10^{1}	1179649	1.151×10^{-7}	1.520×10^{-7}	3.255×10^{-9}	3.257×10^{-9}
SCC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	3.590×10^{-1}	14697	3.426×10^{-5}	4.423×10^{-5}	2.917×10^{-5}	2.917×10^{-5}
10^{-5}	1.137	46233	2.812×10^{-6}	3.697×10^{-6}	3.074×10^{-6}	3.074×10^{-6}
10^{-6}	3.606	145681	2.777×10^{-7}	3.671×10^{-7}	3.108×10^{-7}	3.108×10^{-7}
10^{-7}	1.145×10^{1}	460041	2.777×10^{-8}	3.671×10^{-8}	3.116×10^{-8}	3.116×10^{-8}
10^{-8}	3.591×10^{1}	1454553	2.774×10^{-9}	3.668×10^{-9}	3.119×10^{-9}	3.122×10^{-9}
PCC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	1.017×10^{-1}	4355	4.846×10^{-5}	6.230×10^{-5}	2.990×10^{-5}	2.993×10^{-5}
10^{-5}	5.790×10^{-1}	20275	2.045×10^{-6}	2.688×10^{-6}	1.940×10^{-6}	1.940×10^{-6}
10^{-6}	1.372	61907	6.660×10^{-7}	8.776×10^{-7}	3.249×10^{-7}	3.249×10^{-7}
10^{-7}	6.938	294913	5.342×10^{-8}	7.058×10^{-8}	1.943×10^{-8}	1.943×10^{-8}
10^{-8}	3.236×10^{1}	1007643	1.622×10^{-9}	2.144×10^{-9}	1.547×10^{-9}	1.549×10^{-9}

Table 13: Error control methods applied to (16) with $\epsilon = 0.005$, MIRK order 2.

DefC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	2.433×10^{-4}	39	9.072×10^{-5}	2.688×10^{-4}	3.812×10^{-5}	3.892×10^{-5}
1×10^{-5}	6.200×10^{-4}	69	5.324×10^{-6}	2.540×10^{-5}	3.384×10^{-6}	3.388×10^{-6}
1×10^{-6}	1.520×10^{-3}	119	6.478×10^{-7}	2.700×10^{-6}	3.710×10^{-7}	3.714×10^{-7}
1×10^{-7}	3.585×10^{-3}	202	9.185×10^{-8}	2.885×10^{-7}	4.376×10^{-8}	4.380×10^{-8}
1×10^{-8}	9.689×10^{-3}	374	4.934×10^{-9}	2.204×10^{-8}	3.717×10^{-9}	3.719×10^{-9}
GEC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	2.433×10^{-4}	37	2.965×10^{-4}	6.397×10^{-4}	4.552×10^{-5}	4.547×10^{-5}
1×10^{-5}	6.667×10^{-4}	73	3.892×10^{-5}	7.114×10^{-5}	2.797×10^{-6}	2.799×10^{-6}
1×10^{-6}	2.383×10^{-3}	145	4.284×10^{-6}	6.605×10^{-6}	1.734×10^{-7}	1.737×10^{-7}
1×10^{-7}	8.320×10^{-3}	289	4.031×10^{-7}	5.299×10^{-7}	1.082×10^{-8}	1.083×10^{-8}
1×10^{-8}	1.340×10^{-2}	577	3.301×10^{-8}	3.867×10^{-8}	6.760×10^{-10}	6.767×10^{-10}
SCC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	3.133×10^{-4}	39	9.072×10^{-5}	2.688×10^{-4}	3.812×10^{-5}	3.892×10^{-5}
1×10^{-5}	7.667×10^{-4}	69	5.324×10^{-6}	2.540×10^{-5}	3.384×10^{-6}	3.388×10^{-6}
1×10^{-6}	1.743×10^{-3}	119	6.478×10^{-7}	2.700×10^{-6}	3.710×10^{-7}	3.714×10^{-7}
1×10^{-7}	4.331×10^{-3}	202	9.185×10^{-8}	2.885×10^{-7}	4.376×10^{-8}	4.380×10^{-8}
1×10^{-8}	1.042×10^{-2}	374	4.934×10^{-9}	2.204×10^{-8}	3.717×10^{-9}	3.719×10^{-9}
PCC						
Tol	CPU Times (s)	N	Est. Defect	True Defect	Est. GE	True GE
1×10^{-4}	7.200×10^{-4}	73	3.892×10^{-5}	7.114×10^{-5}	2.797×10^{-6}	2.799×10^{-6}
1×10^{-5}	2.530×10^{-3}	145	4.284×10^{-6}	6.605×10^{-6}	1.734×10^{-7}	1.737×10^{-7}
1×10^{-6}	8.403×10^{-3}	289	4.031×10^{-7}	5.299×10^{-7}	1.082×10^{-8}	1.083×10^{-8}
1×10^{-7}	1.457×10^{-2}	577	3.301×10^{-8}	3.867×10^{-8}	6.760×10^{-10}	6.767×10^{-10}
1×10^{-8}	3.280×10^{-2}	1153	2.429×10^{-9}	2.774×10^{-9}	4.225×10^{-11}	4.265×10^{-11}

Table 14: Error control methods applied to (16) with $\epsilon = 0.005$, MIRK order 4.

DefC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	5.000×10^{-5}	16	3.768×10^{-5}	9.267×10^{-5}	8.900×10^{-5}	1.483×10^{-4}
10^{-5}	9.000×10^{-5}	22	5.583×10^{-6}	1.813×10^{-5}	8.158×10^{-6}	2.181×10^{-5}
10^{-6}	2.967×10^{-4}	35	2.651×10^{-7}	9.181×10^{-7}	6.303×10^{-7}	9.015×10^{-7}
10^{-7}	5.933×10^{-4}	49	3.741×10^{-8}	1.283×10^{-7}	9.248×10^{-8}	1.307×10^{-7}
10^{-8}	7.933×10^{-4}	68	7.000×10^{-9}	2.160×10^{-8}	1.306×10^{-8}	1.813×10^{-8}
GEC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	8.000×10^{-5}	19	1.475×10^{-5}	5.395×10^{-5}	2.293×10^{-5}	8.404×10^{-5}
10^{-5}	3.233×10^{-4}	37	4.392×10^{-7}	1.734×10^{-6}	1.129×10^{-6}	1.289×10^{-6}
10^{-6}	1.330×10^{-3}	73	1.432×10^{-8}	4.625×10^{-8}	1.755×10^{-8}	2.005×10^{-8}
10^{-7}	1.373×10^{-3}	73	1.432×10^{-8}	4.625×10^{-8}	1.755×10^{-8}	2.005×10^{-8}
10^{-8}	5.024×10^{-3}	145	4.173×10^{-10}	1.049×10^{-9}	2.738×10^{-10}	3.169×10^{-10}
SCC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	8.667×10^{-5}	16	3.768×10^{-5}	9.267×10^{-5}	8.900×10^{-5}	1.483×10^{-4}
10^{-5}	1.067×10^{-4}	22	5.583×10^{-6}	1.813×10^{-5}	8.158×10^{-6}	2.181×10^{-5}
10^{-6}	3.933×10^{-4}	35	2.651×10^{-7}	9.181×10^{-7}	6.303×10^{-7}	9.015×10^{-7}
10^{-7}	6.300×10^{-4}	49	3.741×10^{-8}	1.283×10^{-7}	9.248×10^{-8}	1.307×10^{-7}
10^{-8}	3.942×10^{-3}	135	2.003×10^{-10}	4.755×10^{-10}	2.417×10^{-10}	2.833×10^{-10}
PCC						
Tol	CPU Time (s)	N	Est. Defect	True Defect	Est. GE	True GE
10^{-4}	9.000×10^{-5}	19	1.475×10^{-5}	5.395×10^{-5}	2.293×10^{-5}	8.404×10^{-5}
10^{-5}	3.700×10^{-4}	37	4.392×10^{-7}	1.734×10^{-6}	1.129×10^{-6}	1.289×10^{-6}
10^{-6}	1.287×10^{-3}	73	1.432×10^{-8}	4.625×10^{-8}	1.755×10^{-8}	2.005×10^{-8}
10^{-7}	1.423×10^{-3}	73	1.432×10^{-8}	4.625×10^{-8}	1.755×10^{-8}	2.005×10^{-8}
10^{-8}	5.268×10^{-3}	145	4.173×10^{-10}	1.049×10^{-9}	2.738×10^{-10}	3.169×10^{-10}

Table 15: Error control methods applied to (16) with $\epsilon = 0.005$, MIRK order 6.

Order	Defect	GE
2	6.23×10^{-7}	5.17
4	7.21×10^{-7}	164.55

Table 16: Result of solving (17) with `BVP_SOLVER` using defect control with an a posteriori GE estimate. Order is the order of the MIRK formula, Defect is the estimated defect max norm, and GE is the estimated GE max norm. We see that the defect is less than the requested tolerance of 10^{-6} but the GE is quite large, signaling the presence of a pseudosolution.