Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 205 (2007) 777–790
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Blended implicit methods for solving ODE and DAE problems, and their extension for second-order problems☆

Luigi Brugnano*, Cecilia Magherini

Dipartimento di Matematica "U. Dini", Viale Morgagni 67/A, 50134 Firenze, Italy

Received 22 July 2005

Abstract

The use of implicit numerical methods is mandatory when solving general stiff ODE/DAE problems. Their use, in turn, requires the solution of a corresponding discrete problem, which is one of the main concerns in the actual implementation of the methods. In this respect, blended implicit methods provide a general framework for the efficient solution of the discrete problems generated by block implicit methods. In this paper, we review the main facts concerning blended implicit methods for the numerical solution of ODE and DAE problems.

In addition to this, we study the extension of blended implicit methods for solving second-order problems, which results in a straightforward generalization of the basic theory for such methods.

Finally, a few numerical tests obtained with the computational code `BiMD`, implementing a variable order-variable stepsize blended implicit method, are also reported, in order to confirm the effectiveness of the approach.

MSC: 65L05; 65L06; 65L80; 65H10

Keywords: Ordinary differential equations; Initial value problems; Stiff problems; Second-order problems; Differential algebraic equations; Numerical methods; Iterative solution of algebraic systems; Blended implicit methods

1. Introduction

When solving general stiff ODE and DAE problems, the use of implicit methods is mandatory, because of the weak stability properties of explicit methods. On the other hand, implicit methods require the solution, at each integration step, of a corresponding discrete problem, whose dimension is a multiple of that of the continuous one. Often, this is the main drawback when implementing such methods in a computational code, since the cost for solving the discrete problem may be very high. In particular, we consider here block implicit methods [25], namely methods that, when applied to the IVP

$$y' = f(t, y), \quad t \in [t_0, T], \quad y(t_0) = y_0 \in \mathbb{R}^m, \tag{1}$$

☆ Research supported by I.N.d.A.M. and Italian M.I.U.R.
* Corresponding author. Fax: +39 55 4222605.
E-mail addresses: brugnano@math.unifi.it (L. Brugnano), magherini@math.unifi.it (C. Magherini).

0377-0427/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2006.02.057

generate, at the nth integration step, a discrete problem in the form

$$A \otimes I_m \mathbf{y}_n - h_n B \otimes I_m \mathbf{f}_n = \boldsymbol{\eta}_n, \tag{2}$$

where the matrices $A, B \in \mathbb{R}^{r\times r}$ (which we assume non-singular hereafter) define the method, h_n is the current stepsize, the block vectors

$$\begin{aligned} &\mathbf{y}_n = (y_{n1}, \ldots, y_{nr})^{\mathrm{T}}, \quad \mathbf{f}_n = (f_{n1}, \ldots, f_{nr})^{\mathrm{T}}, \\ &f_{nj} = f(t_{nj}, y_{nj}), \quad t_{nj} = t_n + c_j h_n, \quad j = 1, \ldots, r, \end{aligned} \tag{3}$$

contain the discrete solution and, finally, $\boldsymbol{\eta}_n$ only depends on already known quantities. This is a rather general framework, including the majority of implicit Runge–Kutta methods, for which $A = I$, the identity matrix, a number of general linear methods (see, for example [11,18,19]), and, more recently, block boundary value methods [7]. Many techniques have been derived, across the years, for efficiently solving the discrete problem (2) (see, for example, [1,10,14,20,21]) and/or to define methods for which it has a suitably simple structure, thus gaining efficiency (see, for example, [9,12,24]). In this context, blended implicit methods [3,4,8] provide a quite general framework, allowing a re-formulation of the discrete problem generated by a given block implicit method, for which an efficient non-linear splitting is straightforwardly defined. Blended implicit methods have been implemented in the computational code `BiM` [5] for the numerical solution of stiff ODEs. Recently, this approach has been extended for handling linearly implicit DAE problems [6], and a corresponding code, `BiMD`, is now available at the same web page [27] of `BiM`. Moreover, in this paper we extend blended implicit methods for solving special second-order problems.

With this premise, the paper is then organized as follows: in Sections 2 and 3 we recall the main facts about blended implicit methods for ODEs and their extension for DAE problems, respectively; in Section 4 we study the extension of blended implicit methods for second-order problems and, finally, in Section 5 we present some numerical tests along with a few concluding remarks.

2. Blended implicit methods for ODE problems

Let now summarize the main facts about *blended implicit methods*, which have been introduced and analyzed in [3–5,8]. In order to conveniently introduce such methods, and to carry out a corresponding linear analysis of convergence (according to [20,21]), let us consider the application of the block method (2) to the usual test equation,

$$y' = \mu y, \quad y(t_0) = y_0, \quad Re(\mu) \leqslant 0.$$

Moreover, for sake of simplicity, we consider the very first application of the method, since the same arguments apply to each step of integration. This allows us to avoid, hereafter, the subscript n and the discrete problem results in the following system of linear equations,

$$(A - qB)\mathbf{y} = \boldsymbol{\eta}, \quad q = h\mu. \tag{4}$$

We observe that the non-singularity of both matrices A and B implies that the discrete problem (4) can be cast, by setting $\gamma > 0$ a free parameter and

$$C = A^{-1}B, \tag{5}$$

into either one of the following two *equivalent* equations:

$$(I - qC)\mathbf{y} = A^{-1}\boldsymbol{\eta} \equiv \boldsymbol{\eta}_1, \quad \gamma(C^{-1} - qI)\mathbf{y} = \gamma B^{-1}\boldsymbol{\eta} \equiv \boldsymbol{\eta}_2. \tag{6}$$

Let us now introduce the *weighting function*

$$\theta(q) = (I - q\gamma I)^{-1}. \tag{7}$$

We observe that $\theta(q)$ is analytical for $q \in \mathbb{C}^-$ and, moreover,

$$\theta(0) = I, \quad \theta(q) \to O, \quad \text{as } q \to \infty, \tag{8}$$

with O denoting the zero matrix. The *blended implicit method* based on the block method (4) is then defined through the following discrete problem,

$$M(q)\mathbf{y} \equiv (A(q) - qB(q))\mathbf{y} = \theta(q)\boldsymbol{\eta}_1 + (I - \theta(q))\boldsymbol{\eta}_2 \equiv \hat{\boldsymbol{\eta}}(q), \tag{9}$$

where

$$A(q) = \theta(q)I + (I - \theta(q))\gamma C^{-1}, \quad B(q) = \theta(q)C + (I - \theta(q))\gamma I. \tag{10}$$

The problem (9)–(10) is evidently obtained as the *blending* of the two equivalent equations (6) with weights $\theta(q)$ and $I - \theta(q)$, respectively: it is, therefore, still equivalent to (4). As a consequence, order and stability properties are fully inherited from the block method (2). However, the problem (9)–(10) has a relevant advantage with respect to (4). In fact, by considering that, because of (7) and (8),

$$M(q) = I + O(q), \quad \text{when } q \approx 0, \quad M(q) = q(\gamma I + O(q^{-1})), \quad \text{as } q \to \infty,$$

then, the splitting matrix

$$N(q) = I - q\gamma I \equiv \theta(q)^{-1} \tag{11}$$

is naturally induced. Indeed, it coincides with $M(q)$ at $q = 0$ and at ∞. The corresponding *blended iteration* associated to the blended method (9)–(10) is

$$N(q)\mathbf{y}^{(i)} = (N(q) - M(q))\mathbf{y}^{(i-1)} + \hat{\boldsymbol{\eta}}(q), \quad i = 1, 2, \ldots, \tag{12}$$

which is convergent iff the spectral radius, say $\rho(q)$, of the iteration matrix,

$$Z(q) = I - N(q)^{-1}M(q), \tag{13}$$

is less than 1. From the previous arguments, we observe that, whatever the choice of the parameter $\gamma > 0$, one has that

$$\rho(0) = 0 \quad \text{and} \quad \rho(q) \to 0, \quad \text{as } q \to \infty.$$

According to [20,21], and by using the maximum modulus principle, the iteration (12) is said to be L-convergent if the *maximum amplification factor*,

$$\rho_\gamma^* = \max_{x>0} \rho(\mathrm{i}x), \tag{14}$$

is smaller than 1, where, as usual, i denotes the imaginary unit. It is worth to emphasize that an L-convergent iteration is highly desirable, if the underlying block method is L-stable. In (14), the subscript γ does denote the dependence of this factor from the choice of the positive parameter γ. A little algebra shows that [4], for the iteration matrix (13),

$$\lambda \in \sigma(C) \iff \frac{q(\lambda - \gamma)^2}{\lambda(1 - q\gamma)^2} \in \sigma(Z(q)). \tag{15}$$

Moreover, for the methods implemented in the codes `BiM` [5] and `BiMD`, and, more in general, for all block methods such that the characteristic polynomial of the matrix (5) coincides (up to a variable scaling) with the reciprocal polynomial at the denominator of the (s, r) Padé approximation to the exponential, $s = r - 2, r - 1, r$, the following result holds true for all practical values of r [4].

Theorem 1. *Let $\lambda_1 = \phi_1 \mathrm{e}^{\mathrm{i}\xi_1}$, $\phi_1, \xi_1 \geqslant 0$, be the eigenvalue of C having minimum modulus (and non-negative imaginary part). Then*

$$\rho(q) = \left| \frac{q(\lambda_1 - \gamma)^2}{\lambda_1(1 - q\gamma)^2} \right| \quad \text{and} \quad \rho_\gamma^* \equiv \rho(\mathrm{i}\gamma^{-1}) = \frac{|\lambda_1 - \gamma|^2}{2\gamma\phi_1}. \tag{16}$$

The latter quantity is minimized by choosing $\gamma = \phi_1$, for which

$$\rho_{\phi_1}^* \equiv \rho^* = 1 - \cos\xi_1.$$

Table 1
Data for the methods implemented in the codes BiM and BiMD

r	Padé	Order	γ	ρ^*
3	(2, 3)	4	.7387	.3398
4	(2, 4)	6	.8482	.5291
6	(4, 6)	8	.7285	.6299
8	(6, 8)	10	.6745	.6885
10	(8, 10)	12	.6433	.7276

Table 2
Data for the Radau IIA and Lobatto IIIA/Gauss methods

r	Padé	Order	γ	ρ^*
2	(1, 2)	3	.4082	.1835
3	(2, 3)	5	.2462	.3398
4	(3, 4)	7	.1738	.4416
5	(4, 5)	9	.1334	.5123
6	(5, 6)	11	.1079	.5644
7	(6, 7)	13	$.9032 \cdot 10^{-1}$	.6045
2	(2, 2)	4	.2887	.1340
3	(3, 3)	6	.1967	.2765
4	(4, 4)	8	.1475	.3793
5	(5, 5)	10	.1173	.4544
6	(6, 6)	12	$.9710 \cdot 10^{-1}$	.5114
7	(7, 7)	14	$.8265 \cdot 10^{-1}$	.5561

With this choice, for all practical values of r one obtains L-convergent iterations, which are associated to corresponding L-stable methods, when $s < r$, or A-stable methods, when $s = r$. In Table 1, we list the data corresponding to the methods implemented in the current versions of the codes BiM (Rel. 2.0, April 2005) and BiMD (Rel. 1.0, October 2005) [27], for which (see (3)) $c_j = j,\ \ j = 1, \ldots, r$. By the way, we mention that Theorem 1 also applies, for example, to Radau IIA, Lobatto IIIA, and Gauss methods: in Table 2, we list the information concerning such methods. We end this section by observing that, when the blended method is applied to the more general problem (1), then the iteration (9)–(12) formally becomes

$$\begin{aligned}\boldsymbol{\delta}^{(i)} &= N^{-1}(\theta((I - \gamma C^{-1}) \otimes I_m \mathbf{y}^{(i-1)} - h(C - \gamma I) \otimes I_m \mathbf{f}^{(i-1)}) \\ &\quad + \gamma(C^{-1} \otimes I_m \mathbf{y}^{(i-1)} - hI \otimes I_m \mathbf{f}^{(i-1)}) - \hat{\boldsymbol{\eta}}),\end{aligned}$$

$$\mathbf{y}^{(i)} = \mathbf{y}^{(i-1)} - \boldsymbol{\delta}^{(i)}, \quad i = 1, 2, \ldots, \tag{17}$$

where, by setting J the Jacobian of $f(t, y)$ evaluated at (t_0, y_0),

$$N^{-1} \equiv \theta = I \otimes \Omega^{-1}, \quad \Omega = I_m - h\gamma J. \tag{18}$$

Consequently, if ω iterations are required to obtain convergence, the overall leading cost, to carry out the iteration (17)–(18), amounts to:

- 1 Jacobian evaluation,
- 1 factorization of the $m \times m$ matrix Ω,

- $r\omega$ function evaluations, and
- $2r\omega$ system solvings with the factors of Ω.

For more details, we refer to [5]. An additional, remarkable, feature of the iteration (17)–(18) is the block diagonal structure of the splitting matrix N. For later reference, we observe that, in the case of the linear system of ODEs, $y' = Jy + g(t)$, with constant $J \in \mathbb{R}^{m\times m}$, the iteration (17)–(18) induces the iteration matrix

$$Z = C^{-1}(C - \gamma I)^2 \otimes hJ(I_m - h\gamma J)^{-2}, \tag{19}$$

for which the previous linear analysis of convergence, summarized by the result of Theorem 1, applies.

3. Blended implicit methods for DAE problems

We now summarize the extension of blended implicit methods for solving linearly implicit DAE problems, according to [6]. Therefore, we now consider problems in the form

$$Ky' = f(t, y), \quad t \in [t_0, T], \quad y(t_0) = y_0 \in \mathbb{R}^m, \tag{20}$$

where the (constant) *mass matrix* K may be singular. Since we are interested in a linear analysis of convergence for the corresponding blended iteration, we confine ourselves to the case of linear equations in the form

$$Ky' = Jy + g(t), \tag{21}$$

where K and J are constant matrices and $g(t)$ is a vector valued function. The following notions (see, for example, [2,17,19]) are briefly recalled, for completeness and later reference. First of all, the *matrix pencil*

$$\mu K - J \tag{22}$$

is associated to Eq. (21). The pencil is said to be *regular* if its determinant is not identically 0, as a function of μ. Moreover, Eq. (21) is solvable iff the pencil (22) is regular, which we assume hereafter. In such a case, the pencil (22) can be cast into its *Kronecker canonical form*,

$$PKQ = \begin{pmatrix} I_d & \\ & H \end{pmatrix}, \quad PJQ = \begin{pmatrix} G & \\ & I_a \end{pmatrix}, \tag{23}$$

where P, $Q \in \mathbb{R}^{m\times m}$ are non-singular, I_d and I_a are the identity matrices of dimension d and a, respectively ($d+a=m$), $G \in \mathbb{R}^{d\times d}$, and $H \in \mathbb{R}^{a\times a}$,

$$H = \begin{pmatrix} H_1 & & \\ & \ddots & \\ & & H_k \end{pmatrix}, \quad H_i = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{R}^{\nu_i\times\nu_i}, \quad i = 1, \ldots, k,$$

with $\sum_{i=1}^{k} \nu_i = a$. Consequently,

$$H^{\nu} = O, \quad \nu = \max_{i=1,\ldots,k} \nu_i, \tag{24}$$

where ν is the *index* (*of nilpotency*) of the DAE (21). It is known that the higher the index, the more difficult the problem. By means of the Kronecker normal form (23), the problem decouples as follows:

$$x' = Gx + \phi(t), \quad Hz' = z + \psi(t), \tag{25}$$

where

$$Q^{-1}y \equiv \begin{pmatrix} x \\ z \end{pmatrix}, \quad Pg(t) \equiv \begin{pmatrix} \phi(t) \\ \psi(t) \end{pmatrix}.$$

With these premises and notations, let us generalize the blended iteration (17)–(18) for solving the DAE (21). The only differences with what has been seen in Section 2 amount to the weighting function and the splitting matrix,

$$\theta = I \otimes K\Omega^{-1}, \quad N = I \otimes \Omega, \quad \Omega = K - h\gamma J, \tag{26}$$

in place of (18). We observe that we get exactly the formulae previously seen in Section 2, in the case $K = I_m$, i.e., when (21) reduces to an ODE. By taking into account (21) and (26), it is not difficult to see that the blended iteration (17) now, formally, becomes

$$\begin{aligned} \boldsymbol{\delta}^{(i)} &= N^{-1}(\theta((I - \gamma C^{-1}) \otimes K - h(C - \gamma I) \otimes J)\mathbf{y}^{(i-1)} \\ &\quad + \gamma(C^{-1} \otimes K - hI \otimes J)\mathbf{y}^{(i-1)} - \hat{\boldsymbol{\eta}}), \\ \mathbf{y}^{(i)} &= \mathbf{y}^{(i-1)} - \boldsymbol{\delta}^{(i)}, \quad i = 1, 2, \dots . \end{aligned} \tag{27}$$

The corresponding iteration matrix is then given by

$$\begin{aligned} Z = I \otimes I_m - N^{-1}(\theta((I - \gamma C^{-1}) \otimes K - h(C - \gamma I) \otimes J) \\ + \gamma(C^{-1} \otimes K - hI \otimes J)). \end{aligned} \tag{28}$$

The following result then holds true [6].

Lemma 2. *The iteration matrix* (28) *is similar to the block diagonal matrix* (*see* (23))

$$\begin{pmatrix} Z_d & \\ & Z_a \end{pmatrix}, \quad \text{where} \quad \begin{cases} Z_d = C^{-1}(C - \gamma I)^2 \otimes hG(I_d - h\gamma G)^{-2}, \\ Z_a = C^{-1}(C - \gamma I)^2 \otimes hH(H - h\gamma I_a)^{-2}. \end{cases} \tag{29}$$

Remark 3. We observe that the two blocks Z_d and Z_a in (29) correspond, respectively, to the ODE and to the normalized DAE in (25).

It is evident that the matrix Z_d in (29) formally coincides with (19), via the substitutions $G \longleftrightarrow J$ and $I_a \longleftrightarrow I_m$. Consequently, the result of Theorem 1 applies to it. On the other hand, for Z_a the following result follows at once from (24).

Theorem 4. *The matrix Z_a is nilpotent of index v.*

Remark 5. As a consequence, one obtains that the convergence properties of the blended iteration (27) are essentially unaffected by the algebraic part of the problem. In fact, they depend only on the matrix Z_d, corresponding to the differential equation in (25), provided that at least v iterations are carried out, if v is the index of the DAE.

4. Extension for second-order problems

We now study the extension of blended implicit methods for solving second-order problems of special type,

$$y'' = f(t, y), \quad t \in [t_0, T], \quad y(t_0) = y_0, \quad y'(t_0) = y'_0 \in \mathbb{R}^m, \tag{30}$$

where stiffness is present, i.e., with $y(t)$ combining components with dominant short frequencies and components with large frequencies and small amplitudes. In this case, it is customary to resort to the following linear test equation [23],

$$y'' = -\mu^2 y, \quad \mu \in \mathbb{R},$$

which we use for the linear analysis of convergence. In particular, methods which are popular for solving this problem (for example, in the class of implicit Runge–Kutta–Nyström methods [16,18]), lead to a discrete problem in the form

$$(I - q^2 C^2)\mathbf{y} = \boldsymbol{\eta}_1, \quad q = \mathrm{i}h\mu \equiv \mathrm{i}x, \quad x \in \mathbb{R}, \tag{31}$$

where the (non-singular) matrix $C \in \mathbb{R}^{r\times r}$ corresponds to a suitable block method (2)–(5). For example, when speaking about Runge–Kutta methods, Gauss formulae with $s = r$ stages and Lobatto IIIA formulae with $s = r + 1$ stages. Then, we can define the following *equivalent* formulation of (31),

$$\gamma^2(C^{-2} - q^2 I)\mathbf{y} = \gamma^2 C^{-2}\boldsymbol{\eta}_1 \equiv \boldsymbol{\eta}_2,$$

with $\gamma > 0$ a free parameter, and the weighting function

$$\theta(q) = (I - q^2\gamma^2 I)^{-1},$$

such that the resulting blended method corresponding to (31) is formally still given by (9), with

$$A(q) = \theta(q)I + (I - \theta(q))\gamma^2 C^{-2}, \quad B(q) = \theta(q)C^2 + (I - \theta(q))\gamma^2 I, \tag{32}$$

instead of (10). Consequently, it induces the blended iteration (12) with the splitting matrix

$$N(q) = I - q^2\gamma^2 I \equiv \theta(q)^{-1} \tag{33}$$

in place of (11). The corresponding iteration matrix is formally still given by (13) and, again, $N(q)$ coincides with $M(q)$ at $q = 0$ and at ∞. The iteration is convergent iff the spectral radius of $Z(q)$, which we again denote by $\rho(q)$, is less than 1, and the corresponding maximum amplification factor is formally still given by (14) (see (31)). In analogy with the related notion of P-stability [23], we give the following definition.

Definition 6. We say that the iteration (12) and (32)–(33) is P-convergent if the corresponding maximum amplification factor (14) is smaller than 1.

Also in this case, we are interested in choosing the parameter γ in order for ρ^*_γ to be minimized. For this purpose, we observe that the two iterations, respectively, given by (9)–(12), and (9), (12), (32)–(33), are formally the same via the substitutions:

$$\gamma \longleftrightarrow \gamma^2, \quad q \longleftrightarrow q^2, \quad C \longleftrightarrow C^2.$$

Consequently, we obtain at once from (15) that the eigenvalues of the iteration matrix $Z(q)$ now satisfy

$$\lambda \in \sigma(C) \iff \frac{q^2(\lambda^2 - \gamma^2)^2}{\lambda^2(1 - q^2\gamma^2)^2} \in \sigma(Z(q)).$$

The following result is then readily established.

Lemma 7. *For a fixed* $\gamma > 0$, *one has that* (*see* (31))

$$\rho^*_\gamma \equiv \rho(\mathrm{i}\gamma^{-1}) = \frac{1}{4\gamma^2}\max_{\lambda\in\sigma(C)}\frac{|\lambda^2 - \gamma^2|^2}{|\lambda|^2}. \tag{34}$$

Since C is a real matrix and $\gamma^2 > 0$, we can consider, in the study of the last member of Eq. (34), only the eigenvalues with non-negative imaginary part. Let us assume that they are given by

$$\lambda_j = \phi_j \mathrm{e}^{\mathrm{i}\xi_j}, \quad j = 1, \ldots, \ell \equiv \lceil r/2 \rceil, \tag{35}$$

and moreover they satisfy

$$0 < \phi_1 \leqslant \cdots \leqslant \phi_\ell, \quad \frac{\pi}{2} > \xi_1 > \cdots > \xi_\ell \geqslant 0. \tag{36}$$

The following result then holds true.

Theorem 8. *If conditions* (35)–(36) *are fulfilled and*, *moreover*,

$$\phi_1^2 \geqslant \frac{\phi_j^2}{1 + 4(\cos^2\xi_j - \cos^2\xi_1)}, \quad j = 2, \ldots, \ell, \tag{37}$$

Table 3
Parameters for Gauss and Lobatto IIIA methods

r	Order	γ	$\hat{\rho}^*$	ρ^*	$\hat{\rho}_0$	$\tilde{\rho}_0$	$\hat{\rho}_\infty$	$\tilde{\rho}_\infty$
2	4	.2887	.2500	.1340	.0833	.0774	$1.2 \cdot 10^{1}$	$9.3 \cdot 10^{-1}$
3	6	.1967	.4765	.2765	.0738	.1088	$4.9 \cdot 10^{1}$	$2.8 \cdot 10^{0}$
4	8	.1475	.6147	.3793	.0535	.1119	$1.1 \cdot 10^{2}$	$5.1 \cdot 10^{0}$
5	10	.1173	.7024	.4544	.0387	.1066	$2.0 \cdot 10^{2}$	$7.7 \cdot 10^{0}$
6	12	$.9710 \cdot 10^{-1}$	.7613	.5114	.0287	.0993	$3.2 \cdot 10^{2}$	$1.1 \cdot 10^{1}$
7	14	$.8265 \cdot 10^{-1}$	.8029	.5561	.0219	.0919	$4.7 \cdot 10^{2}$	$1.3 \cdot 10^{1}$
8	16	$.7185 \cdot 10^{-1}$	.8336	.5921	.0172	.0851	$6.5 \cdot 10^{2}$	$1.6 \cdot 10^{1}$
9	18	$.6348 \cdot 10^{-1}$	.8569	.6218	.0138	.0789	$8.5 \cdot 10^{2}$	$2.0 \cdot 10^{1}$
10	20	$.5681 \cdot 10^{-1}$	.8752	.6467	.0113	.0735	$1.1 \cdot 10^{3}$	$2.3 \cdot 10^{1}$

then (*see* (34)),

$$\rho(q) = \left| \frac{q(\lambda_1^2 - \gamma^2)}{\lambda_1(1 - q^2\gamma^2)} \right|^2 \quad \textit{and} \quad \rho_\gamma^* = \frac{|\lambda_1^2 - \gamma^2|^2}{4\gamma^2\phi_1^2}. \tag{38}$$

The latter quantity is minimized by choosing $\gamma = \phi_1$. *With this choice, one obtains* $\rho_{\phi_1}^* \equiv \hat{\rho}^* = (1 - \cos^2 \xi_1)$. *Therefore, the corresponding iteration* (9), (12), *and* (32)–(33) *turns out to be P-convergent*.

Proof. In order for the second equation in (38) to be satisfied, one must have

$$|\lambda_1^2 - \gamma^2|^2\phi_j^2 \geqslant |\lambda_j^2 - \gamma^2|^2\phi_1^2, \quad j = 2, \ldots, \ell.$$

By considering (35), one then obtains

$$\gamma^4(\phi_j^2 - \phi_1^2) + 4\gamma^2\phi_1^2\phi_j^2(\cos^2 \xi_j - \cos^2 \xi_1) \geqslant \phi_1^2\phi_j^2(\phi_j^2 - \phi_1^2).$$

By taking into account (36), a sufficient condition for this inequality to be satisfied is given by

$$\gamma^2 \geqslant \frac{\phi_j^2 - \phi_1^2}{4(\cos^2 \xi_j - \cos^2 \xi_1)}, \quad j = 2, \ldots, \ell. \tag{39}$$

Let us assume, for a moment, that (39) holds true. Then, from (38) one obtains, after a few calculations,

$$\rho_\gamma^* = \frac{1}{4}\left(\frac{\phi_1^2}{\gamma^2} + \frac{\gamma^2}{\phi_1^2} - 2\cos 2\xi_1 \right) \equiv g(\gamma^2).$$

This is a strictly convex function, since $g''(\gamma^2) > 0$. Its global minimum is obtained by solving $g'(\gamma^2) = 0$, thus leading to $\gamma = \phi_1$. As a consequence, (39) turns out to be equivalent to (37), which holds true by hypothesis. Moreover, we get: $\rho_{\phi_1}^* = g(\phi_1^2) = (1 - \cos 2\xi_1)/2 = 1 - \cos^2 \xi_1 \equiv \hat{\rho}^*$. Finally, having fixed γ, and by considering that we have just proved that the maximum in (34) occurs at $\lambda = \lambda_1$, the first equality in (38) immediately follows. □

Remark 9. We observe that, from a geometric point of view, the conditions (36)–(37) require the eigenvalues of the matrix C to be contained in a suitable annulus of the complex plane, whose internal radius equals ϕ_1.

It turns out that the conditions (36)–(37) are satisfied by both Runge–Kutta Gauss and Lobatto IIIA methods, for which we list in Table 3 the corresponding parameters γ and $\hat{\rho}^*$, as defined in Theorem 8. Such values provide results, for Gauss methods, which appear to be comparable, for example, with those recently obtained in [15] (though, after a more complicated analysis). We can then conclude that the iteration (9), (12), and (32)–(33), based on Gauss or Lobatto IIIA methods, provide *P*-convergent implementations of high order, *P*-stable methods for problem (30) [16]. Moreover,

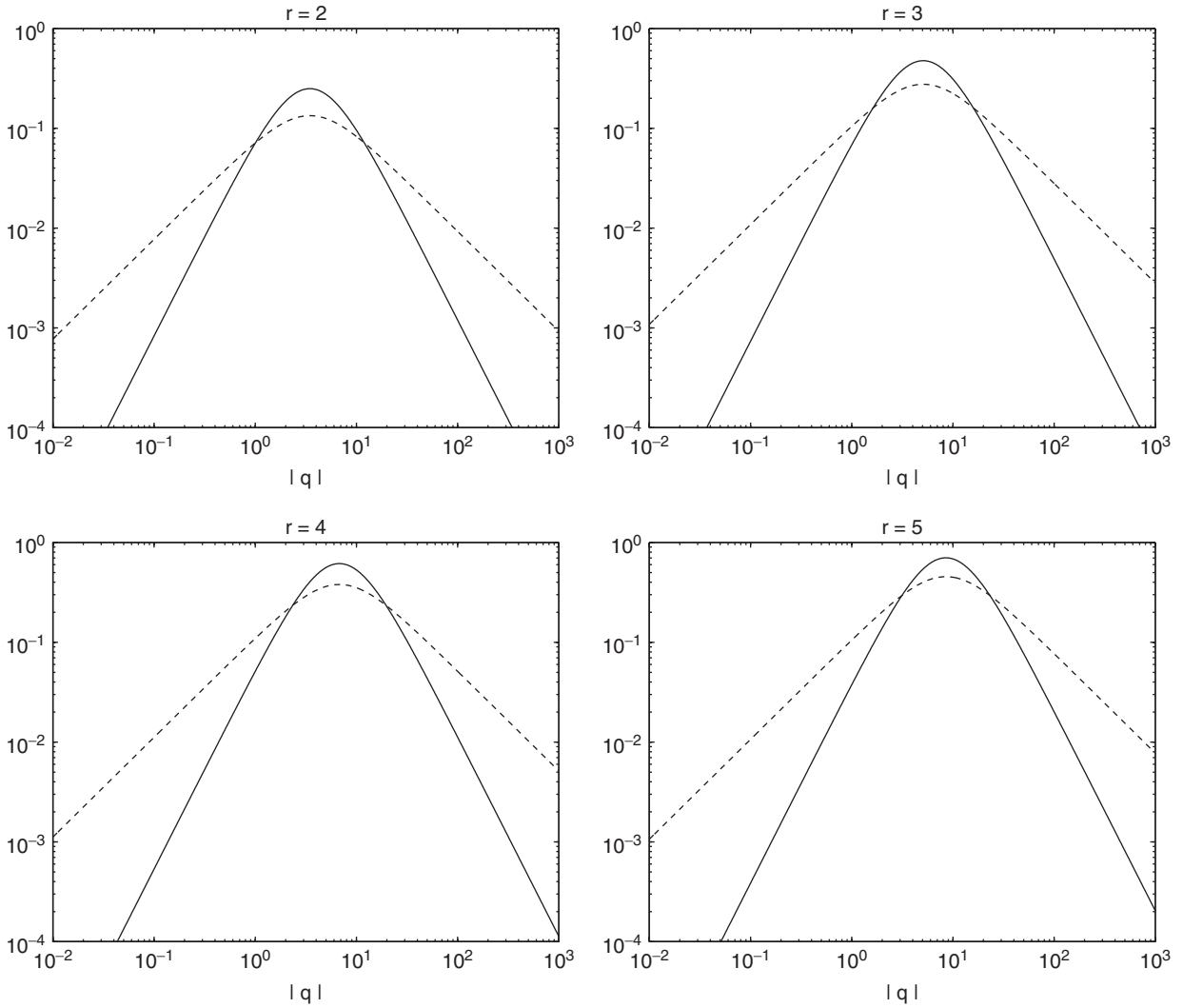

Fig. 1. Spectral radii of the iteration matrices, $r = 2, 3, 4, 5$, corresponding to the second-order (solid line) and first-order (dashed line) formulations.

the fact that the iteration matrix is zero at ∞, makes such iteration very well suited for stiff problems. Finally, in the case of problem (30), the corresponding blended iteration is formally still given by (17)–(18), via the substitutions

$$\gamma \longleftrightarrow \gamma^2, \quad h \longleftrightarrow h^2, \quad C \longleftrightarrow C^2.$$

4.1. Some remarks

It is noticeable that the optimal choice for the parameter γ, provided by the result of Theorem 8 for second-order problems, nicely coincides with that for first-order problems stated in Theorem 1. However, we observe that the maximum amplification factor $\hat{\rho}^*$, as defined in Theorem 8 (see the fourth column in Table 3) is larger than the corresponding value of the amplification factor ρ^*, as defined in Theorem 1. The latter would result from the application of the blended method based on Gauss or Lobatto IIIA formulae to the corresponding first-order system (see the fifth column in Table 3). Nevertheless, though $\hat{\rho}^* > \rho^*$, we observe that the overall behavior of the spectral radius of the iteration matrix, for the second-order formulation, is generally more favorable, as a function of $q = \mathrm{i}x$, $x > 0$, than that of the corresponding first-order formulation (see, for example, Fig. 1, for $r = 2, 3, 4, 5$). In order to better measure

Table 4
Parameters for the iteration in [15] for Gauss methods

r	Order	$\hat{\rho}^*$	$\hat{\rho}_0$	$\hat{\rho}_\infty$	Nilpotency index
2	4	.2500	.0833	$1.2 \cdot 10^1$	2
3	6	.4753	.0733	$4.9 \cdot 10^1$	2
4	8	.6437	.0619	$1.2 \cdot 10^2$	2

this aspect, we observe that (see (31) and (38)), for a given $q = \mathrm{i}h\mu$, the spectral radius of the iteration matrix for the second-order formulation is given by

$$\left|\frac{q^2(\lambda_1^2-\gamma^2)^2}{\lambda_1^2(1-q^2\gamma^2)^2}\right| \approx \begin{cases} \hat{\rho}_0|q|^2, & |q| \approx 0, \\ \hat{\rho}_\infty|q|^{-2}, & |q| \gg 0, \end{cases} \tag{40}$$

where

$$\hat{\rho}_0 = \left|\frac{(\lambda_1^2-\gamma^2)}{\lambda_1}\right|^2 \equiv (2\gamma)^2\hat{\rho}^*, \quad \hat{\rho}_\infty \equiv \frac{\hat{\rho}_0}{\gamma^4}.$$

On the other hand, for the first-order formulation one obtains that (see (16)) the spectral radius of the corresponding iteration matrix is given by

$$\left|\frac{q(\lambda_1-\gamma)^2}{\lambda_1(1-q\gamma)^2}\right| \approx \begin{cases} \tilde{\rho}_0|q|, & |q| \approx 0, \\ \tilde{\rho}_\infty|q|^{-1}, & |q| \gg 0, \end{cases} \tag{41}$$

where

$$\tilde{\rho}_0 = \left|\frac{(\lambda_1-\gamma)^2}{\lambda_1}\right| \equiv 2\gamma\rho^*, \quad \tilde{\rho}_\infty \equiv \frac{\tilde{\rho}_0}{\gamma^2}.$$

In the last four columns in Table 3, the two couples of parameters are listed, showing that, for each value of r, the parameters $\hat{\rho}_0$ and $\tilde{\rho}_0$ are almost of the same size for both formulations, whereas the ratio $\hat{\rho}_\infty/\tilde{\rho}_\infty \sim 2/\gamma$ (which is, therefore, bounded). As a consequence, by considering that the power (positive or negative) of $|q|$ is *doubled* for the second-order formulation (compare (40) and (41)), one obtains that this formulation is much more appealing, when $|q|$ is either very small or very large (i.e., in the case of stiff problems). Last, but not least, we observe that one blended iteration for the second-order problem turns out to be cheaper than that for the corresponding first order one. In light of the above facts, we can then conclude that the former iteration is generally preferable than the latter one.

For completeness, in Table 4 we also list corresponding parameters for the iteration recently defined in [15] for second-order problems solved by means of Gauss methods. In addition to the parameters $\hat{\rho}^*$, $\hat{\rho}_0$, and $\hat{\rho}_\infty$ above described, i.e., by setting as usual $\rho(q)$ the spectral radius of the corresponding iteration matrix,

$$\hat{\rho}^* = \max_{x \geqslant 0} \rho(\mathrm{i}x) \quad \text{and} \quad \rho(q) \approx \begin{cases} \hat{\rho}_0|q|^2, & |q| \approx 0, \\ \hat{\rho}_\infty|q|^{-2}, & |q| \gg 0, \end{cases}$$

we also list the index of nilpotency of the iteration matrix at ∞. By comparing these entries with the corresponding ones listed in Table 3 for the extension of blended implicit methods for second-order problems, we can conclude that the two iterations appear to have very similar convergence properties: actually, the plot of the spectral radii of the corresponding iteration matrices, as functions of q, are almost identical. Nevertheless, we want to emphasize the conceptual simplicity of the analysis concerning the use of blended implicit methods, for which the convergence properties of the corresponding blended iteration do only depend on the choice of one parameter (i.e., γ). Moreover, the optimal choice of such parameter is easily obtained for most of the methods of practical interest.

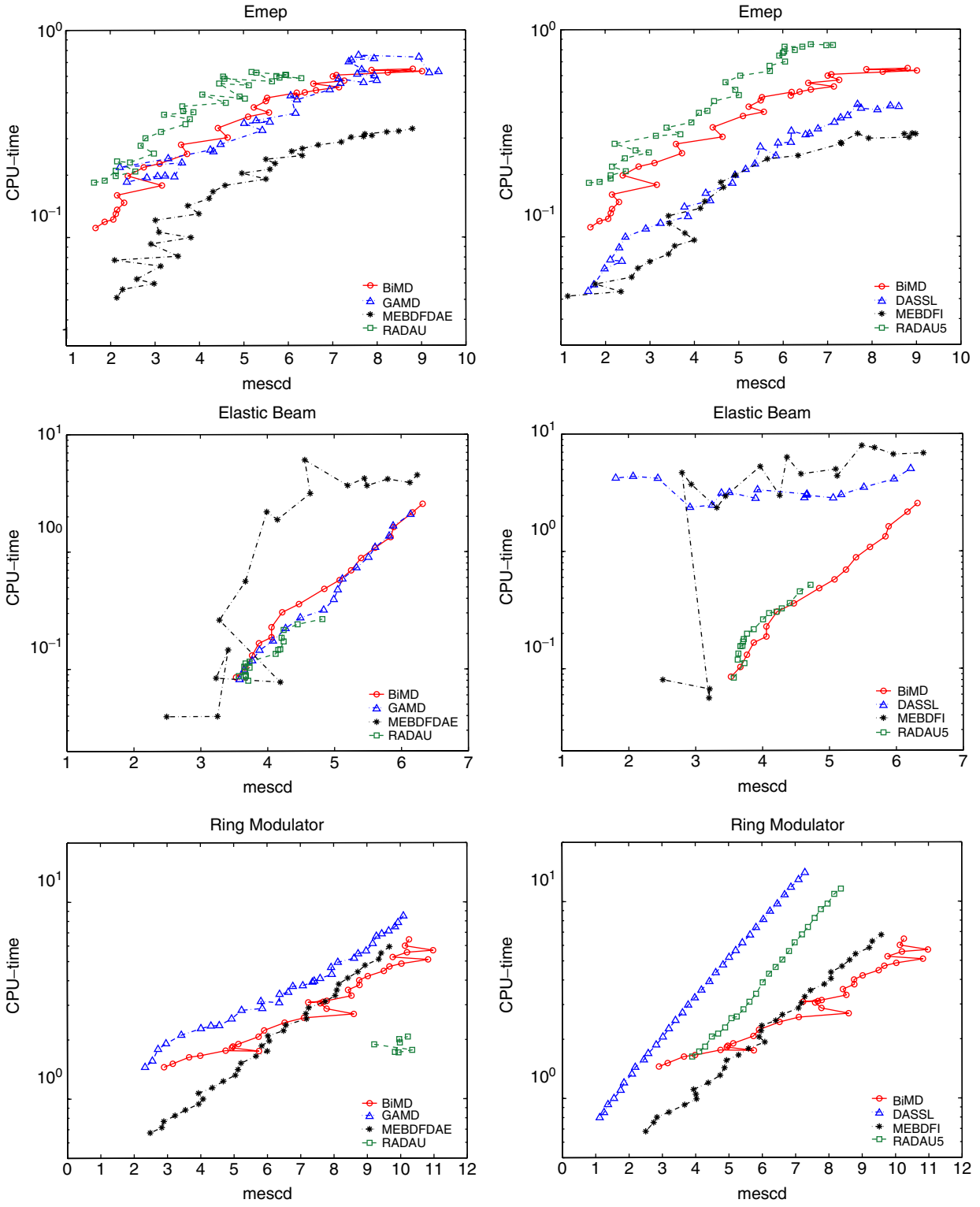

Fig. 2. Numerical results for ODEs.

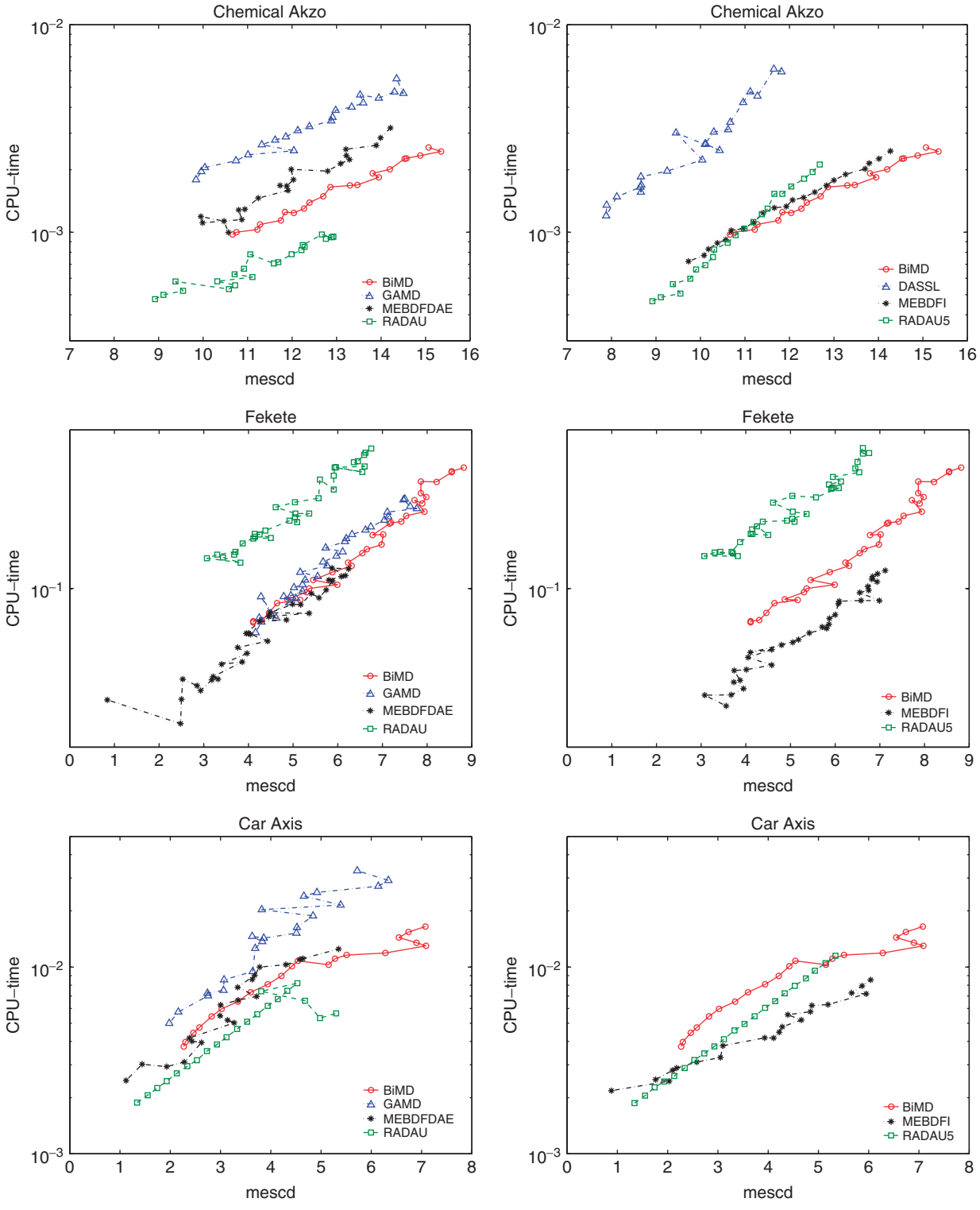

Fig. 3. Numerical results for DAEs.

5. Numerical tests and conclusions

As previously mentioned, blended implicit methods were first implemented in the Fortran 77 code BiM [5], for the numerical solution of stiff ODEs. Subsequently, the extension of such code, called BiMD, for the numerical solution of stiff ODEs and linearly implicit DAEs in the form (20) up to index three, has been released, and it is available at the same web page [27] of BiM. We shall here use the code BiMD to perform some numerical tests, which compare this code with the codes available in the current release of the *Test Set for IVP Solvers* [26] (Rel. 2.2, August 2003). In particular, the codes compared with BiMD are DASSL [2], GAMD [22], MEBDFDAE and MEBDFI [13], RADAU5 and RADAU [19]. All the codes are able to solve ODEs and DAEs up to index three, with the only exception of DASSL, which is designed to work with ODEs and index 1 DAEs. Also the test problems considered here are taken from the *Test Set*: additional numerical results, with respect to those reported here, can be found at the web page [27], where the code BiMD is available. For each problem, the obtained results are displayed by means of corresponding *work-precision diagrams*, i.e., by using the same standard used in the *Test Set* (see also [19]).

In Fig. 2, we plot the results obtained for the following first-order ODE problems: *Emep*, *Elastic Beam*, *Ring Modulator*. In Fig. 3, we plot the results for the following linearly implicit DAE problems: *Chemical Akzo Nobel Problem* (index 1), *Fekete* (index 2), *Car Axis* (index 3). For all tests, the accuracy of the numerical solution (measured in terms of *mixed error significant computed digits* (*mescd*) [26]) is plotted against the execution time. The used input parameters are essentially those used in [26]. The tests have been done on a IBM SP Power 5 computer, by using the `xl` Fortran compiler with the optimization options `-O5 -qstrict`. We mention that the codes RADAU5 and RADAU fail to solve the *Ring Modulator* problem, when using the coarsest accuracy tolerances. Moreover, DASSL, MEBDFDAE, and MEBDFI are less efficient, when solving the *Elastic Beam* problem, because of the lack of A-stability of their respective higher order formulae (see also [19]).

In this paper, we reviewed the main facts concerning the definition of blended implicit methods for the numerical solution of ODE–IVPs and linearly implicit DAEs. The corresponding linear convergence analysis has been recalled, showing the conceptual simplicity of this approach. Moreover, its usefulness is also confirmed by some numerical tests based on the codes and problems included in the *Test Set for IVP Solvers* [26]. In the tests, the code BiMD, implementing a variable order–variable stepsize blended implicit method, compares well with the best codes currently available, both in terms of efficiency and reliability. Finally, blended implicit methods have been here extended for the numerical solution of the special second-order problem (30), resulting in straightforward extensions of the basic theory and results. In the future, we plan to study further extensions of this approach such as, for example, for the numerical solution of evolutionary PDE problems.

References

[1] P. Amodio, L. Brugnano, A note on the efficient implementation of implicit methods for ODEs, J. Comput. Appl. Math. 87 (1997) 1–9.
[2] K.E. Brenan, S.L. Campbell, L.R. Petzold, Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations, Classics in Applied Mathematics, vol. 14, SIAM, Philadelphia, 1996.
[3] L. Brugnano, Blended block BVMs (B_3VMs): a family of economical implicit methods for ODEs, J. Comput. Appl. Math. 116 (2000) 41–62.
[4] L. Brugnano, C. Magherini, Blended implementation of block implicit methods for ODEs, Appl. Numer. Math. 42 (2002) 29–45.
[5] L. Brugnano, C. Magherini, The BiM code for the numerical solution of ODEs, J. Comput. Appl. Math. 164–165 (2004) 145–158.
[6] L. Brugnano, C. Magherini, F. Mugnai, Blended implicit methods for the numerical solution of DAE problems, J. Comput. Appl. Math. 189 (2006) 34–50.
[7] L. Brugnano, D. Trigiante, Solving Differential Problems by Multistep Initial and Boundary Value Methods, Gordon and Breach, Amsterdam, 1998.
[8] L. Brugnano, D. Trigiante, Block implicit methods for ODEs, in: D. Trigiante (Ed.), Recent Trends in Numerical Analysis, Nova Science Publ. Inc., New York, 2001, pp. 81–105.
[9] K. Burrage, A special family of Runge–Kutta methods for solving stiff differential equations, BIT 18 (1988) 22–41.
[10] J.C. Butcher, On the implementation of implicit Runge–Kutta methods, BIT 6 (1976) 237–240.
[11] J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations: Runge–Kutta and General Linear Methods, Wiley, Chichester, 1987.
[12] J.R. Cash, The integration of stiff initial value problems in ODEs using modified extended backward differentiation formulae, Comput. Math. Appl. 9 (1983) 645–657.
[13] J.R. Cash, S. Considine, An MEBDF code for stiff initial value problems, ACM Trans. Math. Software 18 (2) (1992) 142–158.
[14] G.J. Cooper, J.C. Butcher, An iteration scheme for implicit Runge–Kutta methods, IMA J. Numer. Anal. 3 (1983) 127–140.
[15] S. Gonzales-Pinto, S. Pérez-Rodríguez, R. Rojas-Bello, Efficient iterations for Gauss methods on second order problems, J. Comput. Appl. Math. 189 (2006) 80–97.
[16] E. Hairer, Unconditionally stable methods for second order differential equations, Numer. Math. 32 (1979) 373–379.

[17] E. Hairer, C. Lubich, M. Roche, The Numerical Solution of Differential-Algebraic Systems by Runge–Kutta Methods, Lecture Notes in Mathematics, vol. 1409, Springer, Berlin, 1989.
[18] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations I, second ed., Springer Series in Computational Mathematics, vol. 8, Springer, Berlin, 1993.
[19] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems, second ed., Springer Series in Computational Mathematics, vol. 14, Springer, Berlin, 1996.
[20] P.J. van der Houwen, J.J.B. de Swart, Triangularly implicit iteration methods for ODE-IVP solvers, SIAM J. Sci. Comput. 18 (1997) 41–55.
[21] P.J. van der Houwen, J.J.B. de Swart, Parallel linear system solvers for Runge–Kutta methods, Adv. Comput. Math. 7 (1–2) (1997) 157–181.
[22] F. Iavernaro, F. Mazzia, Solving ordinary differential equations by generalized Adams methods: properties and implementation techniques, Appl. Numer. Math. 28 (1998) 107–126.
[23] J.D. Lambert, I.A. Watson, Symmetric multistep methods for periodic initial value problems, J. Inst. Math. Appl. 18 (1976) 189–202.
[24] S.P. Nørsett, Runge–Kutta methods with a multiple real eigenvalue only, BIT 16 (1976) 388–393.
[25] H.A. Watts, L.F. Shampine, *A*-stable block one-step methods, BIT 12 (1972) 252–266.
[26] Test Set for IVP Solvers: ⟨http://pitagora.dm.uniba.it/∼testset/⟩.
[27] Code BiM/BiMD home page: ⟨http://www.math.unifi.it/∼brugnano/BiM/⟩.