Implicit–Explicit Multistep Methods for Nonlinear Convection–Diffusion Equations

Georgios Akrivis and Emmanuil H. Georgoulis

Abstract We construct and analyze implicit–explicit multistep schemes for nonlinear evolution convection–diffusion partial differential equations. We establish optimal order a priori error estimates. We are particularly interested in the dependence of the stability constants on the ratio between the convection and diffusion coefficients, the so-called Péclet number, and on the diffusion coefficient ε itself. In particular, we show that the second order implicit–explicit backward differentiation formula (BDF) admits stability constant independent of the Péclet number.

1 Introduction

This work is concerned with the design and analysis of implicit–explicit multistep methods for parabolic semilinear convection-diffusion partial differential equation (p.d.e.) problems. In order to focus on the time-stepping issues, only the time-discrete schemes are discussed.

Previous works on multistep methods for evolution p.d.e. problems include [1–8, 11, 12]. For the discretization of convection–diffusion equations by Runge–Kutta

G. Akrivis

Department of Computer Science & Engineering, University of Ioannina, Ioannina, Greece

Institute of Applied and Computational Mathematics, FORTH, Heraklion, Crete, Greece
e-mail: akrivis@cse.uoi.gr

E. H. Georgoulis (✉)
Department of Mathematics, University of Leicester, Leicester, UK

School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Zografou, Greece

Institute of Applied and Computational Mathematics, FORTH, Heraklion, Crete, Greece
e-mail: Emmanuil.Georgoulis@le.ac.uk

methods we refer to [9, 10]. The standard monograph for numerical methods for parabolic equations is [15].

Multistep methods can be computationally attractive, as they do not require the calculation of intermediate stages (in contrast to, e.g., Runge–Kutta time-stepping schemes) to achieve high order convergence rates in time. The use of carefully constructed implicit–explicit schemes can further reduce the computational cost by requiring the solution of one linear equation at each time step.

More specifically, we shall construct and analyze implicit–explicit multistep methods for the following initial and boundary value problem: seek a function $u : \bar{\Omega} \times [0, T] \to \mathbb{R}$ satisfying

$$
\begin{cases}
u_t - \varepsilon \Delta u + \nabla \cdot \big(ub(x, t)\big) + c(x, t)u = f(u, x, t) & \text{in } \Omega \times (0, T), \\
u = 0 & \text{on } \partial\Omega \times (0, T), \\
u(\cdot, 0) = u^0 & \text{in } \Omega.
\end{cases} \tag{1.1}
$$

Here $\Omega \subset \mathbb{R}^d$ is a bounded domain, $\bar{\Omega}$ and $\partial\Omega$ are the closure and the boundary of Ω, respectively, T and ε are given positive numbers, and

$$
\nabla \cdot (ub) := \sum_{i=1}^{d} (b_i u)_{x_i}.
$$

The convective coefficient b, the coefficient c, the initial value u^0 and the forcing term f are given functions. We assume the vector-valued function $b : \bar{\Omega} \times [0, T] \to \mathbb{R}^d$ is continuously differentiable, $c : \bar{\Omega} \times [0, T] \to \mathbb{R}$ is continuous, $u^0 \in L^2(\Omega)$ and $f : \mathbb{R} \times \bar{\Omega} \times [0, T] \to \mathbb{R}$ is globally Lipschitz continuous in its first argument, uniformly with respect to its second and third arguments,

$$
\exists L \geqslant 0 \, \forall x \in \bar{\Omega} \, \forall t \in [0, T] \, \forall y_1, y_2 \in \mathbb{R} \quad |f(y_1, x, t) - f(y_2, x, t)| \leqslant L|y_1 - y_2|, \tag{1.2}
$$

and such that $f(0, \cdot, t) \in L^2(\Omega)$, for all $t \in [0, T]$. It is then easily seen that $f(v, \cdot, t) \in L^2(\Omega)$, for any $v \in L^2(\Omega)$ and all $t \in [0, T]$; indeed, using the Lipschitz condition (1.2) and elementary inequalities, we have

$$
\begin{aligned}
\big|f\big(v(x), x, t\big)\big|^2 &\leqslant 2\big|f\big(v(x), x, t\big) - f(0, x, t)\big|^2 + 2|f(0, x, t)|^2 \\
&\leqslant 2L^2|v(x)|^2 + 2|f(0, x, t)|^2,
\end{aligned}
$$

and our claim is evident. Additional hypotheses on the data will be imposed below.

We assume that the initial and boundary value problem (1.1) admits a sufficiently smooth solution u.

Let now (α, β) be a strongly $A(0)$-stable q-step scheme and (α, γ) be an explicit q-step scheme, characterized by three polynomials α, β and γ,

$$\alpha(\zeta) = \sum_{i=0}^{q} \alpha_i \zeta^i, \quad \beta(\zeta) = \sum_{i=0}^{q} \beta_i \zeta^i, \quad \gamma(\zeta) = \sum_{i=0}^{q-1} \gamma_i \zeta^i.$$

For simplicity, we assume that the order of both q-step schemes, the implicit (α, β) and the explicit (α, γ), is p, i.e.,

$$\sum_{i=0}^{q} i^\ell \alpha_i = \ell \sum_{i=0}^{q} i^{\ell-1} \beta_i = \ell \sum_{i=0}^{q-1} i^{\ell-1} \gamma_i, \quad \ell = 0, 1, \ldots, p. \tag{1.3}$$

As an example of schemes satisfying our assumptions we mention the implicit–explicit BDF methods, described by the polynomials

$$\alpha(\zeta) = \sum_{j=1}^{q} \frac{1}{j} \zeta^{q-j} (\zeta - 1)^j, \quad \beta(\zeta) = \zeta^q, \quad \gamma(\zeta) = \zeta^q - (\zeta - 1)^q. \tag{1.4}$$

The corresponding implicit (α, β)-schemes are the well-known BDF methods, which are strongly $A(0)$-stable for $q = 1, \ldots, 6$; their order is $p = q$. For a given α, the scheme (α, γ) is the unique explicit q-step scheme of order $p = q$. Thus, the implicit–explicit BDF methods satisfy the order conditions (1.3) with $p = q$.

Let $N \in \mathbb{N}$, $k := T/N$ be the constant time step, and $t^n := nk, n = 0, \ldots, N$, be a uniform partition of the interval $[0, T]$. We assume that starting approximations $U^0, \ldots, U^{q-1}$ are given, as we shall be concerned with q-step schemes for (1.1).

We first write the differential equation in (1.1) in the form

$$u_t + Au + C(t)u = B(t, u(t)), \quad t \in (0, T), \tag{1.5}$$

with appropriate for our purposes operators A, B and C. The numerical scheme depends on the particular choice of the operators A, B and C in (1.5). We shall discuss specific choices later on. We discretize the operators A and C implicitly, by the implicit scheme (α, β), and the operator B explicitly, by the explicit scheme (α, γ). Thus, we define approximations U^m to the nodal values $u^m := u(\cdot, t^m)$ of the exact solution by

$$\sum_{i=0}^{q} \left[\alpha_i I + k\beta_i \left(A + C(t^{n+i}) \right) \right] U^{n+i} = k \sum_{i=0}^{q-1} \gamma_i B(t^{n+i}, U^{n+i}), \tag{1.6}$$

for $n = 0, \ldots, N - q$.

1.1 Consistency Error

We shall now discuss a suitable representation of the consistency error of the implicit–explicit scheme (1.6), which will later be used to derive optimal order consistency estimates; see also [1]. We assume that the order of both schemes (α, β) and (α, γ) is p; cf. (1.3).

The consistency error E^n of the scheme (1.6) for the solution u of (1.1), i.e., the amount by which the exact solution fails to satisfy (1.6), is given by

$$kE^n = \sum_{i=0}^{q} \left[\alpha_i I + k\beta_i \left(A + C(t^{n+i})\right)\right] u^{n+i} - k \sum_{i=0}^{q-1} \gamma_i B(t^{n+i}, u^{n+i}), \qquad (1.7)$$

$n = 0, \ldots, N - q$. First, letting

$$E_1^n := \sum_{i=0}^{q} \left[\alpha_i u^{n+i} - k\beta_i u_t(t^{n+i})\right], \quad E_2^n := k \sum_{i=0}^{q} (\beta_i - \gamma_i) B(t^{n+i}, u^{n+i}),$$

with $\gamma_q := 0$, and using the differential equation in (1.5), we infer that

$$kE^n = E_1^n + E_2^n. \qquad (1.8)$$

Furthermore, via a Taylor expansion about t^n, we see that, due to the order conditions of the implicit (α, β)-scheme (i.e., the first equality in (1.3)) and the second equality in (1.3), respectively, leading terms of order up to $p - 1$ cancel, and we obtain

$$\begin{cases} E_1^n = \dfrac{1}{p!} \sum_{i=1}^{q} \int_{t^n}^{t^{n+i}} (t^{n+i} - s)^{p-1} \left[\alpha_i (t^{n+i} - s) - pk\beta_i\right] \dfrac{\partial^{p+1} u}{\partial t^{p+1}}(s)\, ds, \\[2ex] E_2^n = \dfrac{k}{(p-1)!} \sum_{i=1}^{q} (\beta_i - \gamma_i) \int_{t^n}^{t^{n+i}} (t^{n+i} - s)^{p-1} \dfrac{d^p}{dt^p} B(s, u(s))\, ds. \end{cases} \qquad (1.9)$$

This representation of the consistency error will allow us to derive optimal order consistency estimates in suitable norms, under reasonable regularity assumptions.

The remainder of this work is structured as follows. In Sect. 2, we study linearly implicit numerical schemes, characterised by the same discrete operator for all time levels, and we derive optimal order error estimates. We show that, under such general setting, the constant in the stability estimate for this family of methods depends on the Péclet number. Moreover, the constant in the error estimate depends also on the diffusion parameter ε implicitly, through high order Sobolev norms of the exact solution u. In such general setting one does not expect better dependence on the singular perturbation parameter and the results should be treated as an indication

of the numerical challenges for this class of methods. At the other end of the spectrum, in Sect. 3, we focus on first and second order schemes of BDF type, requiring again one linear solve at every time level to advance in time. Crucially, however, we shall make use of possibly time-dependent *nonsymmetric*, in general, linear discrete operators, stemming from possible respective time-dependence of the convective and/or reaction coefficients b and c. This somewhat nonstandard choice will be justified below. Using the A-stability of these low order schemes, we improve crucially on the estimates of Sect. 2: the stability constant is now *independent* of the diffusion parameter ε, while the constant in the error estimate depends on it only implicitly, through appropriate norms of the exact solution u. To test the potential of the proposed method, in Sect. 4, we present a series of numerical experiments demonstrating the performance of the second order BDF method in the discretization of semilinear convection-diffusion equations, with nonlinearities admitting non-Lipschitz growth. Although the latter result is somewhat in the folklore of this class of methods, we were not able to locate a proof under the same assumptions.

2 Error Estimates with Constants Depending on the Péclet Number

Let $c^\star$ be a fixed positive number, let the operator C in (1.5) vanish, and choose the operators A and B as $Av := -\varepsilon\Delta v + c^\star v$ and $B(t, v) := f\big(v(\cdot), \cdot, t\big) - \nabla \cdot \big(vb(t)\big) + \big(c^\star - c(t)\big)v$. We can then, obviously, write the p.d.e. in (1.1) in the form (1.5). For simplicity, we suppressed the dependence on x; we shall follow this convention also below.

With this splitting, the scheme (1.6) takes the form

$$\sum_{i=0}^{q}(\alpha_i I + k\beta_i A)U^{n+i} = k \sum_{i=0}^{q-1}\gamma_i B(t^{n+i}, U^{n+i}), \tag{2.1}$$

$n = 0, \ldots, N - q$.

To advance with (2.1) in time, i.e., to compute the unknown U^{n+q}, we need to solve one linear equation, with the same operator for all time levels. As we shall see below, this operator is self-adjoint and positive definite; in particular, the approximate solutions are well defined.

We shall follow the analysis in [4, 7] and [1] to study the implicit–explicit numerical scheme (1.6).

2.1   Notation and Lipschitz Conditions

We let $H := L^2(\Omega)$, denote by $(\cdot, \cdot)$ and $\|\cdot\|$ its inner product and norm, respectively, and we recall the standard Sobolev spaces. Evidently, the operator $A : \mathscr{D}(A) = H^2(\Omega) \cap H_0^1(\Omega) \to H$ is linear, self-adjoint and positive definite; the domain $V := \mathscr{D}(A^{1/2})$ of $A^{1/2}$ is $V = H_0^1(\Omega)$. We denote by V' the dual of V, with respect to the pivot space H, i.e., $V' = H^{-1}(\Omega)$, and we introduce the norms $\|\cdot\|$ and $\|\cdot\|_\star$ in V and V', respectively, by

$$\|v\| := \|A^{1/2}v\| = \big(-\varepsilon(\Delta v, v) + c^\star(v, v)\big)^{1/2} = \big(\varepsilon\|\nabla v\|^2 + c^\star\|v\|^2\big)^{1/2}$$

and

$$\|v\|_\star := \|A^{-1/2}v\| = \big(v, (-\varepsilon\Delta + c^\star I)^{-1}v\big)^{1/2}.$$

In a standard fashion, A can be extended to an operator from V onto V'; the notation $(\cdot, \cdot)$ will additionally signify the duality pairing between V' and V. The operator $B(t, \cdot) : \mathscr{D}(A) \to H$ can also be viewed as an operator from V into V'.

For notational convenience, we split the operator B into two parts, $B = B_1 + B_2$, with

$$B_1(t)v := -\nabla \cdot \big(vb(t)\big) \quad \text{and} \quad B_2(t, v) := f\big(v(\cdot), \cdot, t\big) + \big(c^\star - c(t)\big)v.$$

Useful Estimates for $\|v\|_\star$ We have

$$\|v\|_\star = \sup_{u \in V \setminus \{0\}} \frac{(v, u)}{\big(\varepsilon\|\nabla u\|^2 + c^\star\|u\|^2\big)^{1/2}},$$

which leads to the estimate

$$\|v\|_\star \leqslant \min\left\{\frac{1}{\sqrt{c^\star}}\|v\|, \frac{1}{\sqrt{\varepsilon}}\|v\|_{H^{-1}}\right\}. \tag{2.2}$$

Indeed, first

$$\|v\|_\star \leqslant \sup_{u \in V \setminus \{0\}} \frac{(v, u)}{\sqrt{c^\star}\|u\|} \leqslant \sup_{u \in V \setminus \{0\}} \frac{\|v\|\,\|u\|}{\sqrt{c^\star}\|u\|},$$

whence

$$\|v\|_\star \leqslant \frac{1}{\sqrt{c^\star}}\|v\|; \tag{2.3}$$

furthermore,

$$\|v\|_\star \leqslant \sup_{u \in V \setminus \{0\}} \frac{(v, u)}{\sqrt{\varepsilon}\|\nabla u\|} \leqslant \sup_{u \in V \setminus \{0\}} \frac{\|v\|_{H^{-1}}\|\nabla u\|}{\sqrt{\varepsilon}\|\nabla u\|} \leqslant \frac{1}{\sqrt{\varepsilon}}\|v\|_{H^{-1}}.$$

Lipschitz Conditions First, for $v, \tilde{v} \in V$, we have

$$(B_1(t, v), \tilde{v}) = \int_\Omega v b(\cdot, t) \cdot \nabla \tilde{v} \, dx = \sum_{i=1}^d \int_\Omega b_i(\cdot, t) v(\tilde{v})_{x_i} \, dx,$$

whence, with

$$\hat{b} := \max_{1 \leqslant i \leqslant d} \max_{\substack{x \in \bar{\Omega} \\ 0 \leqslant t \leqslant T}} |b_i(x, t)|,$$

we obtain

$$\left|\left(B_1(t, v), \tilde{v}\right)\right| \leqslant \hat{b}\|v\| \sum_{i=1}^d \|(\tilde{v})_{x_i}\| \leqslant \hat{b}\sqrt{d}\|v\| \left(\sum_{i=1}^d \|(\tilde{v})_{x_i}\|^2\right)^{1/2} = \hat{b}\sqrt{d}\|v\| \|\nabla \tilde{v}\|.$$

Therefore, since $\sqrt{\varepsilon}\|\nabla \tilde{v}\| \leqslant \|\tilde{v}\|$,

$$\forall t \in [0, T] \, \forall v \in V \quad \|B_1(t, v)\|_\star \leqslant \mu_1\|v\| \quad \text{with} \quad \mu_1 := \frac{\hat{b}}{\sqrt{\varepsilon}}\sqrt{d}. \tag{2.4}$$

Furthermore, with

$$\hat{c} := \max_{\substack{x \in \bar{\Omega} \\ 0 \leqslant t \leqslant T}} |c^\star - c(x, t)|,$$

in view of (1.2), for $v, \tilde{v} \in V$, we have

$$\|B_2(t, v) - B_2(t, \tilde{v})\|^2 \leqslant \int_\Omega \left[(L + \hat{c})|v(x) - \tilde{v}(x)|\right]^2 dx = (L + \hat{c})^2\|v - \tilde{v}\|^2,$$

whence, according to (2.3),

$$\|B_2(t, v) - B_2(t, \tilde{v})\|_\star \leqslant \frac{1}{\sqrt{c^\star}}\|B_2(t, v) - B_2(t, \tilde{v})\| \leq \frac{L + \hat{c}}{\sqrt{c^\star}}\|v - \tilde{v}\|,$$

i.e.,

$$\forall t \in [0, T] \, \forall v, \tilde{v} \in V \quad \|B_2(t, v) - B_2(t, \tilde{v})\|_\star \leqslant \mu_2\|v - \tilde{v}\| \quad \text{with} \quad \mu_2 := \frac{L + \hat{c}}{\sqrt{c^\star}}. \tag{2.5}$$

From (2.4) and (2.5) we obtain the desired global Lipschitz condition

$$\forall t \in [0, T] \; \forall v, \tilde{v} \in V \quad \|\!|B(t, v) - B(t, \tilde{v})|\!\|_\star \leqslant \mu \|v - \tilde{v}\|, \tag{2.6}$$

with Lipschitz constant $\mu := \mu_1 + \mu_2$,

$$\mu = \frac{\hat{b}}{\sqrt{\varepsilon}}\sqrt{d} + \frac{L + \hat{c}}{\sqrt{c^\star}}. \tag{2.7}$$

Notice that the Lipschitz constant μ is bounded for uniformly bounded Péclet numbers $\hat{b}/\sqrt{\varepsilon}$.

2.2 Consistency

From the representations (1.8) and (1.9) of the consistency error, we immediately obtain, under obvious regularity requirements, the desired optimal order consistency estimate

$$\max_{0 \leqslant n \leqslant N-q} \|\!|E^n|\!\|_\star \leqslant ck^p; \tag{2.8}$$

the coefficient c depends on the diffusion parameter ε, through appropriate norms of the exact solution u.

2.3 Stability

Let $U^m, V^m \in V, m = 0, \ldots, N$, satisfy (2.1) and

$$\sum_{i=0}^{q}(\alpha_i I + k\beta_i A)V^{n+i} = k \sum_{i=0}^{q-1} \gamma_i B(t^{n+i}, V^{n+i}), \tag{2.9}$$

$n = 0, \ldots, N - q$, respectively. Then, with $\vartheta^m := U^m - V^m, m = 0, \ldots, N$, we have the stability estimate

$$\|\vartheta^n\|^2 + k \sum_{\ell=0}^{n} \|\!|\vartheta^\ell|\!\|^2 \leqslant c \sum_{j=0}^{q-1} \left(\|\vartheta^j\|^2 + k\|\!|\vartheta^j|\!\|^2\right), \quad n = q - 1, \ldots, N, \tag{2.10}$$

with a constant c independent of U^m, V^m and k; see [7, Theorem 2.1]. Actually, since the Lipschitz constant μ in the Lipschitz condition (2.6) depends on ε only through the Péclet number, the stability constant c in (2.10) depends on the diffusion

parameter ε also only through the Péclet number; indeed, it is bounded, for bounded Péclet numbers.

2.4 Error Estimates

According to [7, Theorem 2.1], we have the following error estimate:

Theorem 2.1 (Error Estimate) *Let* (α, β) *be a strongly* $A(0)$*-stable* q*-step scheme and* (α, γ) *be an explicit* q*-step scheme. Let the order of both schemes* (α, β) *and* (α, γ) *be* p*. Assume we are given starting approximations* $U^0, U^1, \ldots, U^{q-1} \in V$ *to* $u^0, \ldots, u^{q-1}$ *such that*

$$\max_{0 \leqslant j \leqslant q-1} \left(\|u^j - U^j\| + k^{1/2}\|u^j - U^j\| \right) \leqslant ck^p. \tag{2.11}$$

Let $U^n \in V, n = q, \ldots, N,$ *be recursively defined by* (2.1)*. Let* $e^n = u^n - U^n, n = 0, \ldots, N,$ *be the approximation error. Then, there exists a constant* c*, independent of* k *and* n*, depending exponentially on* $\hat{b}^2/\varepsilon$*, such that*

$$\|e^n\|^2 + k\sum_{\ell=0}^{n} \|e^\ell\|^2 \leqslant c\left\{ \sum_{j=0}^{q-1} \left(\|e^j\|^2 + k\|e^j\|^2 \right) + k\sum_{\ell=0}^{n-q} \|E^\ell\|_\star^2 \right\}, \tag{2.12}$$

$n = q - 1, \ldots, N,$ *whence, in view of* (2.11) *and* (2.8)*,*

$$\max_{0 \leqslant n \leqslant N} \|u(t^n) - U^n\| \leqslant ck^p. \tag{2.13}$$

In the estimate (2.12) E^ℓ *is the consistency error of the scheme; see* (1.7) *with* $C(t) = 0$. $\qquad\qquad\qquad\Box$

Remark 2.1 (BDF Schemes) We focus here on the case that the method (2.1) is the implicit–explicit q-step BDF scheme. Since in the Lipschitz condition (2.6) the norm $\|\!|\cdot|\!\|$ does not enter on the right-hand side, the stability constants λ in the notation of [7] vanishes; therefore, Remark 7.2 of [4] applies, and we can relax condition (2.11) on the starting approximations $U^0, \ldots, U^{q-1}$. More precisely, the statement of Theorem 2.1 is valid in this case under the assumption

$$\max_{0 \leqslant j \leqslant q-1} \|u^j - U^j\| \leqslant ck^p.$$

$\qquad\qquad\qquad\Box$

Remark 2.2 (A Wider Class of Linearly Implicit Methods) The splitting of the p.d.e. in (1.1) that we used in the scheme (2.1) satisfies also the assumptions imposed in [4]. Therefore, with this specific splitting, the initial and boundary value problem

(1.1) can be discretized by the wider class of linearly implicit methods discussed in [4]. Thus, the abstract results of [4] apply and lead to error estimates with error constants depending exponentially on $\hat{b}^2/\varepsilon$, as is the case in Theorem 2.1.   □

2.5 Alternative Forms of the Implicit–Explicit Schemes

Various possibilities of splitting the p.d.e. in (1.1) before discretizing by implicit–explicit multistep schemes are possible; here, we comment on two such alternatives.

2.5.1 First Choice

Consider now the operators $\tilde{A}$ and $\tilde{B}$ defined as $\tilde{A}v := -\varepsilon\Delta v$, and $\tilde{B}(t, v) := f(v(\cdot), \cdot, t) - \nabla \cdot (vb(t)) - c(t)v$. Notice that the only difference to the splitting we used in (2.1) is that here we set $c^\star = 0$. With this splitting, the scheme (1.6) takes the form

$$\sum_{i=0}^{q}(\alpha_i I + k\beta_i \tilde{A})U^{n+i} = k\sum_{i=0}^{q-1}\gamma_i \tilde{B}(t^{n+i}, U^{n+i}), \qquad (2.14)$$

$n = 0, \ldots, N - q$. In this case,

$$\|v\| := \|\tilde{A}^{1/2}v\| = \sqrt{\varepsilon}\|\nabla v\| \quad \text{and} \quad \|v\|_\star := \|\tilde{A}^{-1/2}v\| = \frac{1}{\sqrt{\varepsilon}}\|(-\Delta)^{-1/2}v\|.$$

and, with $\tilde{B}_2(t, v) := f(v(\cdot), \cdot, t)$, we have

$$\|\tilde{B}_2(t, v) - \tilde{B}_2(t, \tilde{v})\|^2 = \int_\Omega \left|f(v(x), x, t) - f(\tilde{v}(x), x, t)\right|^2 dx$$

$$\leqslant L^2 \int_\Omega |v(x) - \tilde{v}(x)|^2 dx = L^2\|v - \tilde{v}\|^2.$$

Furthermore, using the Poincaré–Friedrichs inequality $\|w\| \leqslant C_{PF}\|\nabla w\|$, for $w \in V$, we easily see that $\|v\|_\star \leqslant \frac{1}{\sqrt{\varepsilon}}C_{PF}\|v\|$, for $v \in H$. Thus, we infer that

$$\forall t \in [0, T] \, \forall v, \tilde{v} \in V \quad \|\tilde{B}_2(t, v) - \tilde{B}_2(t, \tilde{v})\|_\star \leqslant \tilde{\mu}_2\|v - \tilde{v}\| \qquad (2.15)$$

with $\tilde{\mu}_2 := \frac{1}{\sqrt{\varepsilon}}LC_{PF}$. Therefore, a straightforward application of the analysis of [7] leads to a stability estimate with constant depending exponentially on $1/\varepsilon$, rather than on the Péclet number. However, the scheme (2.14) can, obviously, be

equivalently written in the form

$$\sum_{i=0}^{q}(\alpha_i I + k\beta_i A)U^{n+i} - kc^\star \sum_{i=0}^{q}\beta_i U^{n+i} = k\sum_{i=0}^{q-1}\gamma_i \tilde{B}(t^{n+i}, U^{n+i}), \qquad (2.16)$$

$n = 0, \ldots, N - q$. Applying here the analysis of [1], rather than the one of [7], we can then easily see that the results of both Theorem 2.1 and Remark 2.1 remain valid also for the scheme (2.14).

2.5.2   Second Choice

We discuss here the discretization of the linear part of Eq. (1.1) implicitly, by the implicit scheme (α, β), and of the nonlinear part explicitly, by the explicit scheme (α, γ). With the operator A used above, $C(t)v := \nabla \cdot (vb(t)) + (c(t) - c^\star)v$, and $B(t, v) := f(v(\cdot), \cdot, t)$, the scheme can be written in the form (2.1). Applying the analysis of [1] (with the operator A rather than with $\tilde{A}$) we easily see that the results of Theorem 2.1 remain valid in this case as well. Furthermore, in the case of BDF schemes the requirement on the starting approximations can be relaxed as in Remark 2.1. Notice that $c^\star$ is used in the analysis of the schemes only; the schemes themselves are independent of $c^\star$.

Let us emphasize that, in contrast to the schemes (2.1) and (2.14), in this case the operator of the linear equations is in general time dependent, and so varies from one time level to the next, if b and/or c are/is time dependent. Also, the numerical approximations are well defined, provided the time step k is sufficiently small. More precisely, for a given $w \in V'$, it suffices to show that equation

$$\alpha_q v + k\beta_q\big[-\varepsilon\Delta v + \nabla \cdot (vb(t^{n+q})) + c(t^{n+q})v\big] = w, \qquad (2.17)$$

possesses a unique solution $v \in V$. A well-known property of $A(0)$-stable multistep schemes (α, β) is that the product $\alpha_q\beta_q$ is positive. Assume, without loss of generality, that α_q is positive. According to the Lax–Milgram Lemma, it obviously suffices to show that the bilinear form $a : V \times V \to \mathbb{R}$,

$$a(v, \tilde{v}) := \alpha_q(v, \tilde{v}) + k\beta_q\varepsilon(\nabla v, \nabla\tilde{v}) + k\beta_q\big(\nabla \cdot (vb(t^{n+q})), \tilde{v}\big) + k\beta_q\big(c(t^{n+q})v, \tilde{v}\big)$$

is coercive and continuous. First, for $v \in V$, using an elementary inequality, we have

$$a(v, v) = \alpha_q\|v\|^2 + k\beta_q\varepsilon\|\nabla v\|^2 - k\beta_q\big(vb(t^{n+q}), \nabla v\big) + k\beta_q\big(c(t^{n+q})v, v\big)$$

$$= \alpha_q\|v\|^2 + k\beta_q\varepsilon\|\nabla v\|^2 + \frac{1}{2}k\beta_q\big(v\nabla \cdot b(t^{n+q}), v\big)k\beta_q\big(c(t^{n+q})v, v\big)$$

$$\geqslant \Big(\alpha_q - \frac{1}{2}\beta_q k\|\nabla \cdot b(t^{n+q})\|_{L^\infty(\Omega)} - \beta_q k\|c(t^{n+q})\|_{L^\infty(\Omega)}\Big)\|v\|^2 + k\beta_q\varepsilon\|\nabla v\|^2,$$

and we infer that a is indeed coercive for sufficiently small k, independent of ε. Similarly, for $v, \tilde{v} \in V$, we have

$$|a(v, \tilde{v})| \leqslant \alpha_q \|v\| \, \|\tilde{v}\| + k\beta_q \varepsilon \|\nabla v\| \, \|\nabla \tilde{v}\| + k\beta_q b^\star \|v\| \, \|\nabla \tilde{v}\| + k\beta_q c^\star \|v\| \, \|\tilde{v}\|,$$

and see that a is also continuous.

It is evident from the above discussion that, Péclet-number independent stability analysis, in such general setting, is an essential challenge and may potentially not be true for schemes that are not A-stable. Indeed, as we shall see below, upon employing non-standard and specialised techniques, we are able to recover Péclet-number independent stability bounds for implicit–explicit Euler method (a known result, but nevertheless, included for completeness of the presentation) and for the classical implicit–explicit second order BDF method; the latter is an improvement upon the results presented in [12].

3   Low Order Schemes

Focusing, now, on low order time stepping schemes, our goal is to establish error estimates via the energy technique with stability constants independent of ε.

We consider, again, the initial and boundary value problem (1.1) and assume that

$$c(x, t) + \frac{1}{2}\nabla \cdot b(x, t) \geqslant \underline{b}^2 \quad \forall x \in \bar{\Omega} \ t \in [0, T], \tag{3.1}$$

for some positive $\underline{b}$; notice that this can always be achieved by either adding, if necessary, to both sides of the differential equation a term of the form au with a sufficiently large coefficient a, or by the change of variables $\tilde{u} := e^{-at}u$. Notice that this affects the constant in the error estimate. Indeed, in the first case the Lipschitz constant changes from L to $L + a$. In the second case $\tilde{f}(\tilde{u}, x, t) = e^{-at}f(e^{at}\tilde{u}, x, t)$ satisfies the Lipschitz condition with constant L, but we need to multiply the approximations $\tilde{U}^n$ of $\tilde{u}^n$ by e^{at^n} to obtain approximations U^n of u^n and thus $u^n - U^n = e^{at^n}(\tilde{u}^n - \tilde{U}^n)$.

For convenience we introduce the operator $A(t) : H^2(\Omega) \cap H_0^1(\Omega) \to L^2(\Omega) =: H, t \in [0, T]$, by

$$A(t)v := -\varepsilon \Delta v + \nabla \cdot \big(vb(x, t)\big) + c(x, t)v.$$

Notice, however, that $A(t)$ is not self-adjoint and possibly time-dependent. Obviously, $A(t)$ can be extended to an operator from $V := H_0^1(\Omega)$ to $V' = H^{-1}(\Omega)$.

Taking the duality paring with v and integrating by parts, we have

$$(A(t)v, v) = \varepsilon \|\nabla v\|^2 + (\nabla \cdot \big(vb(x, t)\big) + c(x, t)v, v)$$

$$= \varepsilon \|\nabla v\|^2 + ((c + \frac{1}{2}\nabla \cdot b)v, v).$$

Thus, in view of (3.1), $A(t)$ is coercive:

$$(A(t)v, v) \geqslant \varepsilon \|\nabla v\|^2 + \underline{b}^2 \|v\|^2 \quad \forall v \in V = H_0^1(\Omega). \tag{3.2}$$

3.1   Implicit–Explicit Euler Method

For completeness, we begin by presenting the lowest order case of implicit–explicit Euler method, keeping careful track of the stability constants.

We recursively define a sequence of approximations U^m to the nodal values $u(t^m)$ of the solution u of the initial and boundary value problem (1.1) by the implicit–explicit Euler method,

$$U^{n+1} + kA(t^{n+1})U^{n+1} = U^n + kf(U^n, \cdot, t^n), \quad n = 0, \dots, N-1, \tag{3.3}$$

with starting approximation $U^0 := u^0$. In view of the coercivity (3.2), it can be easily seen that the numerical approximations are well defined.

3.1.1   Consistency

The consistency error E^n of the implicit–explicit Euler scheme (3.3) for the solution u of (1.1),

$$kE^n = u^{n+1} + kA(t^{n+1})u^{n+1} - u^n - kf(u^n, \cdot, t^n), \quad n = 0, \dots, N-1, \tag{3.4}$$

can be written in the form $kE^n = E_1^n + E_2^n$ with

$$E_1^n = \int_{t^n}^{t^{n+1}} (t^n - s)u_{tt}(s)\,ds, \quad E_2^n = k \int_{t^n}^{t^{n+1}} \frac{d}{dt} f(u(s), \cdot, s)\,ds; \tag{3.5}$$

cf. (1.9). Therefore, under obvious regularity assumptions, we derive the desired optimal order consistency estimate

$$\max_{0 \leqslant n \leqslant N-1} \|E^n\| \leqslant ck \tag{3.6}$$

with a suitable positive constant c. (Of course, c depends implicitly on ε, since the solution u depends on ε.)

3.1.2 Stability

Let U^m, $V^m \in V$, $m = 0, \ldots, N$, satisfy (3.3) and

$$V^{n+1} + kA(t^{n+1})V^{n+1} = V^n + kf(V^n, \cdot, t^n), \quad n = 0, \ldots, N - 1, \tag{3.7}$$

respectively. Then, $\vartheta^m := U^m - V^m$, $m = 0, \ldots, N$, satisfy the relation

$$\vartheta^{n+1} + kA(t^{n+1})\vartheta^{n+1} = \vartheta^n + k\big[f(U^n, \cdot, t^n) - f(V^n, \cdot, t^n)\big], \tag{3.8}$$

$n = 0, \ldots, N - 1$. Taking in (3.8) the inner product with ϑ^{n+1} and utilizing (3.2), we obtain

$$\|\vartheta^{n+1}\|^2 + k\varepsilon\|\nabla\vartheta^{n+1}\|^2 + k\underline{b}^2\|\vartheta^{n+1}\|^2 \leqslant (\vartheta^n, \vartheta^{n+1}) + k(f(U^n, \cdot, t^n) - f(V^n, \cdot, t^n), \vartheta^{n+1}).$$

Therefore,

$$\begin{aligned}
\|\vartheta^{n+1}\|^2 + k\varepsilon\|\nabla\vartheta^{n+1}\|^2 + k\underline{b}^2\|\vartheta^{n+1}\|^2 &\leqslant \frac{1}{2}\|\vartheta^n\|^2 + \frac{1}{2}\|\vartheta^{n+1}\|^2 \\
&\quad + k\|f(U^n, \cdot, t^n) - f(V^n, \cdot, t^n)\| \, \|\vartheta^{n+1}\| \\
&\leqslant \frac{1}{2}\|\vartheta^n\|^2 + \frac{1}{2}\|\vartheta^{n+1}\|^2 + \frac{1}{2}Lk\|\vartheta^n\|^2 + \frac{1}{2}Lk\|\vartheta^{n+1}\|^2,
\end{aligned}$$

whence, by multiplying by 2,

$$\big[1 - (L - 2\underline{b}^2)k\big]\|\vartheta^{n+1}\|^2 \leq (1 + Lk)\|\vartheta^n\|^2. \tag{3.9}$$

Now, for sufficiently small k,

$$\|\vartheta^{n+1}\|^2 \leqslant (1 + c_\star k)\|\vartheta^n\|^2, \tag{3.10}$$

with a suitable constant $c_\star$. This is obviously valid in the case $2\underline{b}^2 \geqslant L$, with $c_\star = L$. In the case $L > 2\underline{b}^2$, (3.9) yields, for $k < 1/(L - 2\underline{b}^2)$,

$$\|\vartheta^{n+1}\|^2 \leq \frac{1 + Lk}{1 - (L - 2\underline{b}^2)k}\|\vartheta^n\|^2. \tag{3.11}$$

For any fixed $c_\star$ (strictly) larger than $2(L - \underline{b}^2)$, it is easily seen that

$$\frac{1 + Lk}{1 - (L - 2\underline{b}^2)k} \leqslant 1 + c_\star k, \tag{3.12}$$

provided that k is sufficiently small,

$$k \leqslant \frac{c_\star - 2(L - \underline{b}^2)}{c_\star(L - 2\underline{b}^2)},$$ (3.13)

and (3.10) follows from (3.11) and (3.12).

Now, from (3.10) we obtain

$$\|\vartheta^n\|^2 \leqslant (1 + c_\star k)^n \|\vartheta^0\|^2, \quad n = 0, \dots, N,$$

and thus

$$\|\vartheta^n\|^2 \leqslant e^{c_\star nk} \|\vartheta^0\|^2, \quad n = 0, \dots, N.$$

Hence, we arrive at the desired stability estimate

$$\max_{1 \leqslant n \leqslant N} \|\vartheta^n\| \leqslant e^{c_\star T/2} \|\vartheta^0\|.$$ (3.14)

Crucially, the above stability constant is *independent* of the diffusion coefficient ε.

3.1.3 Error Estimate

Let $e^n := u^n - U^n, n = 0, \dots, N$. Subtracting (3.3) from (3.4), we obtain the error equation

$$e^{n+1} + kA(t^{n+1})e^{n+1} = e^n + k\big[f(u^n, \cdot, t^n) - f(U^n, \cdot, t^n)\big] + kE^n,$$ (3.15)

$n = 0, \dots, N - 1$. Taking here the inner product with e^{n+1}, proceeding as in the stability proof, and utilizing the consistency estimate (3.6) as well as the fact that e^0 vanishes, we easily derive the desired error estimate

$$\max_{1 \leqslant n \leqslant N} \|e^n\| \leqslant ck.$$ (3.16)

The constant c on the right-hand side of (3.16) depends on ε only implicitly through Sobolev norms of the solution u (see (3.5) and (3.6)).

3.2 Implicit–Explicit Two-Step BDF Method

Here we present a robust error analysis for the IMEX using a two-step BDF method. Although we present the time-discrete analysis only, the result can be used to improve fully discrete a priori error bounds for fully discrete BDF-IMEX schemes

for various stable spatial discretizations, e.g., [12]; in particular, the exponential dependence of the a priori error bound constant on the Péclet number from [12] can be avoided.

With starting approximation $U^0 := u^0$, we first perform one step of the implicit–explicit Euler scheme to compute U^1, i.e., we let U^1 be given by

$$U^1 + kA(t^1)U^1 = U^0 + kf(U^0, \cdot, t^0), \tag{3.17}$$

and then let the approximations $U^2, \ldots, U^N$ be given by the implicit–explicit two-step BDF scheme,

$$\frac{3}{2}U^{n+2} - 2U^{n+1} + \frac{1}{2}U^n + kA(t^{n+2})U^{n+2} = 2kf(U^{n+1}, \cdot, t^{n+1}) - kf(U^n, \cdot, t^n), \tag{3.18}$$

$n = 0, \ldots, N - 2$. Again, in view of (3.2), it can be easily seen that the numerical approximations are well defined.

3.2.1 Consistency

The consistency error E^n of the implicit–explicit BDF scheme (3.18),

$$\begin{aligned} kE^n = {} & \frac{3}{2}u^{n+2} - 2u^{n+1} + \frac{1}{2}u^n + kA(t^{n+2})u^{n+2} \\ & - 2kf(u^{n+1}, \cdot, t^{n+1}) + kf(u^n, \cdot, t^n), \end{aligned} \tag{3.19}$$

$n = 0, \ldots, N - 2$, can be written in the form

$$kE^n = E_1^n + E_2^n$$

with

$$E_1^n = -\int_{t^n}^{t^{n+1}} (t^{n+1} - s)^2 u_{ttt}(s)\, ds + \frac{3}{4}\int_{t^n}^{t^{n+2}} (t^{n+2} - s)\left(t^{n+2} - s - \frac{4}{3}k\right)u_{ttt}(s)\, ds,$$

$$E_2^n = -2k\int_{t^n}^{t^{n+1}} (t^{n+1} - s)\frac{d^2}{dt^2}f(u(s), \cdot, s)\, ds + k\int_{t^n}^{t^{n+2}} (t^{n+2} - s)\frac{d^2}{dt^2}f(u(s), \cdot, s)\, ds;$$

cf. (1.9). Therefore, under the regularity assumptions

$$\|u_{ttt}(t)\|_\star \leqslant c_1 \quad \text{and} \quad \left\|\frac{d^2}{dt^2}f(u(t), \cdot, t)\right\| \leqslant c_2, \tag{3.20}$$

for all $t \in [0, T]$, we immediately conclude that

$$\max_{0 \leqslant n \leqslant N-2} \|E_1^n\| \leqslant 2c_1 k^2 \quad \text{and} \quad \max_{0 \leqslant n \leqslant N-2} \|E_2^n\| \leqslant 2c_2 k^2.$$

Thus, we obtain the desired estimate for the consistency error E^n,

$$\max_{0 \leqslant n \leqslant N-2} \|E^n\| \leqslant Ck^2. \tag{3.21}$$

Remark 3.1 (Regularity Requirement) Note that (3.20) can be replaced by slightly weaker $C^{2,1}$-requirements on u and f. Similar remark applies to (3.5) and (3.6).

3.2.2 Stability

Let $U^0, \ldots, U^N \in V$ satisfy (3.17) and (3.18), and $V^0, \ldots, V^N \in V$ satisfy

$$\frac{3}{2}V^{n+2} - 2V^{n+1} + \frac{1}{2}V^n + kA(t^{n+2})V^{n+2} = 2kf(V^{n+1}, \cdot, t^{n+1}) - kf(V^n, \cdot, t^n), \tag{3.22}$$

$n = 0, \ldots, N - 2$. Let

$$\vartheta^m := U^m - V^m \quad \text{and} \quad b^m := f(U^m, \cdot, t^m) - f(V^m, \cdot, t^m),$$

$m = 0, \ldots, N$. Subtracting (3.22) from (3.18), we obtain

$$\frac{3}{2}\vartheta^{n+2} - 2\vartheta^{n+1} + \frac{1}{2}\vartheta^n + kA(t^{n+2})\vartheta^{n+2} = 2kb^{n+1} - kb^n. \tag{3.23}$$

Now, we observe the identity

$$\left(\frac{3}{2}\vartheta^{n+2} - 2\vartheta^{n+1} + \frac{1}{2}\vartheta^n, \vartheta^{n+2}\right) = \frac{5}{4}\|\vartheta^{n+2}\|^2 - \|\vartheta^{n+1}\|^2 - \frac{1}{4}\|\vartheta^n\|^2$$
$$- \left((\vartheta^{n+2}, \vartheta^{n+1}) - (\vartheta^{n+1}, \vartheta^n)\right) + \frac{1}{4}\|\vartheta^{n+2} - 2\vartheta^{n+1} + \vartheta^n\|^2; \tag{3.24}$$

cf. [16]. We note that (3.24) stems from the G-stability of the BDF method (α, β) with the positive definite symmetric matrix G,

$$G = \frac{1}{4}\begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix};$$

see [14, Example 6.5].

Taking the inner product with ϑ^{n+2} in (3.23), and using (3.24) and (3.2), we have

$$\frac{5}{4}\|\vartheta^{n+2}\|^2 - \|\vartheta^{n+1}\|^2 - \frac{1}{4}\|\vartheta^n\|^2 - \big((\vartheta^{n+2}, \vartheta^{n+1}) - (\vartheta^{n+1}, \vartheta^n)\big)$$
$$+ k\varepsilon\|\nabla\vartheta^{n+2}\|^2 + k\underline{b}^2\|\vartheta^{n+2}\|^2 \leqslant 2k\|b^{n+1}\| \, \|\vartheta^{n+2}\| + k\|b^n\| \, \|\vartheta^{n+2}\|. \tag{3.25}$$

Now, in view of the Lipschitz condition (1.2), we have

$$\|b^m\| \leqslant L\|\vartheta^m\|; \tag{3.26}$$

therefore, (3.25) yields

$$\frac{5}{4}\|\vartheta^{n+2}\|^2 - \|\vartheta^{n+1}\|^2 - \frac{1}{4}\|\vartheta^n\|^2 - \big((\vartheta^{n+2}, \vartheta^{n+1}) - (\vartheta^{n+1}, \vartheta^n)\big)$$
$$+ k\varepsilon\|\nabla\vartheta^{n+2}\|^2 + k\underline{b}^2\|\vartheta^{n+2}\|^2 \leqslant 2Lk\|\vartheta^{n+1}\| \, \|\vartheta^{n+2}\| + Lk\|\vartheta^n\| \, \|\vartheta^{n+2}\|, \tag{3.27}$$

and, using a standard Poincaré-Friedrichs inequality $\|v\|^2 \leqslant c_{PF}\|\nabla v\|^2$, we infer

$$\frac{5}{4}\big(\|\vartheta^{n+2}\|^2 - \|\vartheta^{n+1}\|^2\big) + \frac{1}{4}\big(\|\vartheta^{n+1}\|^2 - \|\vartheta^n\|^2\big)$$
$$-\big((\vartheta^{n+2}, \vartheta^{n+1}) - (\vartheta^{n+1}, \vartheta^n)\big) \leqslant \big(\frac{3L}{2} - \delta\big)k\|\vartheta^{n+2}\|^2 + Lk\|\vartheta^{n+1}\|^2 + \frac{Lk}{2}\|\vartheta^n\|^2, \tag{3.28}$$

with $\delta := c_{PF}\varepsilon + \underline{b}^2$. Summing in (3.28) from $n = 0$ to $n = \ell$, we obtain

$$\frac{5}{4}\big(\|\vartheta^{\ell+2}\|^2 - \|\vartheta^1\|^2\big) + \frac{1}{4}\big(\|\vartheta^{\ell+1}\|^2 - \|\vartheta^0\|^2\big) - (\vartheta^{\ell+2}, \vartheta^{\ell+1})$$
$$\leqslant (3L - \delta)k \sum_{n=0}^{\ell+2}\|\vartheta^n\|^2 - (\vartheta^1, \vartheta^0),$$

whence, easily,

$$\frac{1}{4}\|\vartheta^{\ell+2}\|^2 \leqslant (3L - \delta)k \sum_{n=0}^{\ell+2}\|\vartheta^n\|^2 + \frac{1}{2}\big(\|\vartheta^1\|^2 + \|\vartheta^0\|^2\big).$$

Therefore, we have,

$$\|\vartheta^\ell\|^2 \leqslant 4(3L - \delta)k \sum_{n=0}^{\ell}\|\vartheta^n\|^2 + 2\big(\|\vartheta^1\|^2 + \|\vartheta^0\|^2\big), \tag{3.29}$$

$\ell = 2, \ldots, N$. Now, for sufficiently small k, whose size depends adversely *only* on L, application of the discrete Gronwall inequality leads to the desired local stability estimate

$$\|\vartheta^n\|^2 \leqslant c(\|\vartheta^1\|^2 + \|\vartheta^0\|^2), \quad n = 1, \ldots, N. \tag{3.30}$$

Note that, in particular, the dependence of the stability constant c on ε is desirable, in that it diminishes as $\varepsilon \to 0$ and can even be beneficial for large ε.

Now, let V^1 and V^0 be related by

$$V^1 + kA(t^1)V^1 = V^0 + kf(V^0, \cdot, t^0), \tag{3.31}$$

i.e., starting with initial value V^0 we obtain V^1 by performing one step with the implicit–explicit Euler scheme to the differential equation in (1.1); see (3.17) and (3.18). Obviously, in view of the stability property (3.14) of the implicit–explicit Euler method, $\|\vartheta^1\| \leqslant c\|\vartheta^0\|$, which combined with (3.30) leads to our final stability estimate

$$\max_{1 \leqslant n \leqslant N} \|\vartheta^n\| \leqslant c\|\vartheta^0\|, \tag{3.32}$$

with a constant $c > 0$ independent of ε.

3.2.3   Error Estimates

Let the implicit–explicit BDF2 approximations $U^0, \ldots, U^N$ be given by (3.17) and (3.18). Assume that the solution u of (1.1) is sufficiently smooth, such that (3.21) and (3.6) be valid. Then, combining stability and consistency in the standard way we establish the following optimal order error estimate

$$\max_{0 \leqslant n \leqslant N} \|u(t^n) - U^n\| \leqslant ck^2. \tag{3.33}$$

Again, here the constant c in (3.32) depends only implicitly on ε, through its dependence on Sobolev norms of the exact solution u, via the consistency estimates (3.21) and (3.6).

Remark 3.2 (Energy Technique for Higher Order BDF Methods) Proceeding as in Sect. 2.1, we can see that

$$|(A(t)v, u)| \leqslant (\sqrt{\varepsilon}\|\nabla v\| + \mu\|v\|)(\varepsilon\|\nabla u\|^2 + \underline{b}^2\|u\|^2)^{1/2} \quad \forall v, u \in V = H_0^1(\Omega), \tag{3.34}$$

with a constant μ depending on d, $\max_{x,t} |c(x, t)|$ and the Péclet number $\hat{b}/\sqrt{\varepsilon}$. In view of the coercivity condition (3.2) and (3.34), as well as of the fact that the time

interval $[0, T]$ is bounded, a slight modification of the stability analysis of [2] leads to optimal order error estimates in our case with constants depending on the Péclet number for BDF methods up to order five. Notice that (3.34) is a slight relaxation of the corresponding boundedness condition [2, (1.7)]. The energy stability analysis in [2] is based on the Nevanlinna–Odeh multiplier technique. Moreover, it is not clear if it is possible to improve upon these estimates to arrive to a Péclet-number independent stability analysis for $A(\alpha)$-stable BDF methods with $\alpha < \pi/2$. □

The proposed implicit-explicit methods are somewhat non-standard in that they require the solution of a *non-symmetric* linear system per time-step. Indeed, it is a usual practice to treat convection explicitly also in an effort to arrive at symmetric linear systems instead. Such methods, however, require careful tuning of the discretization parameters, as hyperbolic-type CFL restrictions are introduced by the explicit treatment of the dominant convection term. The latter is, crucially, *not* the case for the low order schemes studied in this work.

We shall now argue that the implicit treatment of the predominantly skew-symmetric convection term does *not* hinder the computational efficiency of the proposed methods in any essential fashion. This is because nonsymmetric linear systems arising from the discretization of the convection-diffusion spatial operator through some stable finite elements (e.g., streamline upwinded Petrov-Galerkin methods, discontinuous Galerkin approaches, etc.) admit a number of special properties that can be exploited in the design of scalable preconditioning strategies. For instance, for discontinuous Galerkin methods for steady convection-diffusion problems it has been shown in [13] that, preconditioning by the symmetric part of the convection-diffusion stiffness matrix for the interior penalty discontinuous Galerkin method within a preconditioned GMREs iteration, provides a *three-step recurrence* Krylov method that converges independently of the spatial mesh-size. This means that, up to the cost of inversion of a symmetric preconditioner, the complexity of the preconditioned GMREs is comparable to that of a standard Conjugate Gradient iteration that would normally be used for the respective symmetric linear system of the classical diffusion-only implicit IMEX scheme. At the same time the step of inverting the symmetric part of the stiffness matrix can be efficiently tackled by standard multilevel approaches, whose convergence is further aided by the presence of a strong reaction coefficient stemming from the time discretization.

4　Numerical Experiments

We present some numerical experiments investigating the convergence rates for the implicit-explicit second order BDF (BDF2) method, described in Sect. 3.2, as well as its robustness with respect to the Péclet number. To fully asses the applicability of the proposed method, our numerical investigations are not confined to globally Lipschitz nonlinearities.

Table 1 Example 1: error and convergence rates

$\varepsilon = 1$			$\varepsilon = 10^{-1}$		
k	$\|e\|_{L^\infty(L^2)}$	Rate	k	$\|e\|_{L^\infty(L^2)}$	Rate
1.0e−1	2.519e−4	–	1.0e−1	1.646e−4	–
5.0e−2	6.603e−5	1.93	5.0e−2	4.289e−5	1.94
1.0e−2	2.674e−6	1.99	1.0e−2	2.221e−6	1.84
5.0e−3	6.687e−7	1.99	5.0e−3	5.832e−7	1.93
1.0e−3	2.682e−8	1.99	1.0e−3	2.451e−8	1.97
$\varepsilon = 10^{-3}$			$\varepsilon = 10^{-5}$		
k	$\|e\|_{L^\infty(L^2)}$	Rate	k	$\|e\|_{L^\infty(L^2)}$	Rate
1.0e−1	3.075e−4	–	1.0e−1	3.790e−4	–
5.0e−2	8.958e−5	1.78	5.0e−2	1.102e−4	1.78
1.0e−2	3.874e−6	1.95	1.0e−2	4.786e−6	1.95
5.0e−3	9.731e−7	1.99	5.0e−3	1.203e−6	1.99
1.0e−3	3.897e−8	2.00	1.0e−3	4.572e−8	2.03

4.1 Example 1

We begin with considering the semilinear convection-diffusion problem for $\Omega = [0, 1]^2$, $T = 1$, $b = (1, 1)^T$, $c = 0$ and $f(u, x, t) = -u^2 + g(x, t)$, with g such that the exact solution of the problem is given by

$$u(x, t) := (1 - \exp(-t))x_1 x_2 (1 - x_1)(1 - x_2), \quad x := (x_1, x_2)^T.$$

The implicit-explicit BDF2 method is implemented for $\varepsilon = 1, 10^{-1}, 10^{-3}, 10^{-5}$, using the finite element library FEniCS, with spatial discretization via conforming finite elements on a 32×32 triangular mesh. The mesh is fine enough to ensure that the time-discretization error dominates the spatial error, which is, generally, non-zero as cubic conforming elements on triangular meshes are used for all computations but one, namely for $k = 10^{-3}$, $\varepsilon = 10^{-5}$, where quadric elements are used.

The errors and the convergence rates are given in Table 1, where k is the time-step size, $\|e\|_{L^\infty(L^2)} := \max_{0 \leqslant n \leqslant N} \|u(t^n) - U^n\|$, and 'rate' is the respective convergence rate between two consecutive time-step sizes. As predicted by the theory, second order convergence with respect to k is observed in all cases.

4.2 Example 2

Next, we consider a convective variant of the classical Fisher equation, namely problem (1.1) with $\Omega = [0, 1]^2$, $T = 1$, $b = (1, 1)^T$, $c = 0$ and $f(u, x, t) = 10u(1 - u)$. We apply the implicit-explicit BDF2 method for $\varepsilon = 10^{-1}, 10^{-2}$, with spatial discretization via conforming quadratic finite elements on a 64×64

Table 2 Example 2: error and convergence rates

$\varepsilon = 10^{-1}$			$\varepsilon = 10^{-2}$		
k	$\|e\|_{L^\infty(L^2)}$	Rate	k	$\|e\|_{L^\infty(L^2)}$	Rate
1.0e−1	4.982e−2	–	5.0e−2	6.743e−1	–
5.0e−2	1.647e−2	1.58	2.5e−2	2.653e−1	1.34
2.5e−2	4.348e−3	1.87	1.25e−2	2.382e−2	3.48
1.25e−2	8.677e−4	2.12	6.125e−3	6.793e−3	1.81

triangular mesh. The finite element space is accurate enough to ensure that the time-discretization error dominates the spatial error and that the boundary layers are, therefore, sufficiently resolved.

As no analytical solution is available, the time-discretization error is computed by comparing the numerical solution to a much finer reference numerical solution $\tilde{U}^n$, $n = 0, 1, \ldots, N$. The reference numerical solution is computed using the implicit-explicit Euler method from Sect. 3.1, with $k = 2.5 \cdot 10^{-4}$ and cubic conforming finite elements on the same meshes as the numerical solution. The errors and the convergence rates are given in Table 2, where k is the time-step size, $\|e\|_{L^\infty(L^2)} := \max_{0 \leqslant n \leqslant N} \|\tilde{U}^n - U^n\|$, and 'rate' is the respective convergence rate between two consecutive time-step sizes. Approximately second order convergence with respect to k is observed in this case also.

References

1. Akrivis, G.: Implicit–explicit multistep methods for nonlinear parabolic equations. Math. Comput. **82**, 45–68 (2013)
2. Akrivis, G.: Stability of implicit–explicit backward difference formulas for nonlinear parabolic equations. SIAM J. Numer. Anal. **53**, 464–484 (2015)
3. Akrivis, G.: Stability of implicit and implicit–explicit multistep methods for nonlinear parabolic equations. IMA J. Numer. Anal. **38**, 1768–1796 (2018)
4. Akrivis, G., Crouzeix, M.: Linearly implicit methods for nonlinear parabolic equations. Math. Comput. **73**, 613–635 (2004)
5. Akrivis, G., Lubich, C.: Fully implicit, linearly implicit and implicit–explicit backward difference formulae for quasi-linear parabolic equations. Numer. Math. **131**, 713–735 (2015)
6. Akrivis, G., Crouzeix, M., Makridakis, Ch.: Implicit–explicit multistep finite element methods for nonlinear parabolic problems. Math. Comput. **67**, 457–477 (1998)
7. Akrivis, G., Crouzeix, M., Makridakis, Ch.: Implicit–explicit multistep methods for quasilinear parabolic equations. Numer. Math. **82**, 521–541 (1999)
8. Banjai, L., Peterseim, D.: Parallel multistep methods for linear evolution problems. IMA J. Numer. Anal. **32**, 1217–1240 (2011)
9. Burman, E., Ern, A.: Implicit-explicit Runge-Kutta schemes and finite elements with symmetric stabilization for advection-diffusion equations. ESAIM Math. Model. Numer. Anal. **46**, 681–707 (2012)
10. Burman, E., Ern, A., Fernández, M.A.: Explicit Runge-Kutta schemes and finite elements with symmetric stabilization for first-order linear PDE systems. SIAM J. Numer. Anal. **48**, 2019–2042 (2010)

11. Crouzeix, M.: Une méthode multipas implicite–explicite pour l'approximation des équations d'évolution paraboliques. Numer. Math. **35**, 257–276 (1980)
12. Dolejší, V., Vlasák, M.: Analysis of a BDF–DGFE scheme for nonlinear convection–diffusion problems. Numer. Math. **110**, 405–447 (2008)
13. Georgoulis, E.H., Loghin, D.: Norm preconditioners for discontinuous Galerkin hp-finite element methods. SIAM J. Sci. Comput. **30**, 2447–2465 (2008)
14. Hairer, E., Wanner, G.: Solving Ordinary Differential Equations II: Stiff and Differential–Algebraic Problems. Springer Series in Computational Mathematics v. 14, 2nd revised edn. Springer, Berlin (2002)
15. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems, 2nd edn. Springer, Berlin (2006)
16. Zlámal, M.: Finite element methods for nonlinear parabolic equations. RAIRO **11**, 93–107 (1977)