SIAM J. NUMER. ANAL.
Vol. 58, No. 1, pp. 189–210

LONG-TIME ACCURATE SYMMETRIZED IMPLICIT-EXPLICIT BDF METHODS FOR A CLASS OF PARABOLIC EQUATIONS WITH NON-SELF-ADJOINT OPERATORS*

BUYANG LI†, KAI WANG†, AND ZHI ZHOU†

Abstract. An implicit-explicit multistep method based on the backward difference formulae (BDF) is proposed for time discretization of parabolic equations with a non-self-adjoint operator. Implicit and explicit schemes are used for the self-adjoint and anti-self-adjoint parts of the operator, respectively. For a k-step method, some correction terms are added to the starting $k-1$ steps to maintain kth-order convergence without imposing further compatibility conditions at the initial time. Long-time kth-order convergence for the numerical method is proved under the assumptions that the operator is coercive and that the non-self-adjoint part is low order. Such an operator often appears in practical computation (such as the Stokes–Darcy system) but may violate the standard sectorial angle condition used in the literature for analysis of BDF. In particular, the proposed method and analysis in this paper extend the long-time energy error analysis of the Stokes–Darcy system in Chen et al. [*SIAM J. Numer. Anal.,* 51 (2013), pp. 2563–2584; *Numer. Math.,* 134 (2016), pp. 857–879] to general symmetrized and decoupled BDF methods up to order 6 by using the generating function technique.

Key words. backward difference formula, implicit-explicit, non-self-adjiont operator, initial correction, long-time stability, Stokes–Darcy system, parabolic equation, error estimate, sectorial angle

AMS subject classifications. 65J08, 65L06, 65M12

DOI. 10.1137/18M1227536

1. Introduction.

Let $V \subset H = H' \subset V'$ be a Gelfand triple of complex Hilbert spaces, where the superscript $'$ denotes the dual. Namely, the embedding $V \hookrightarrow H$ is continuous and dense, and

$$\langle u, v\rangle = (u, v) \quad \forall\, u \in H \hookrightarrow V',\ \forall\, v \in V \hookrightarrow H,$$

where $\langle \cdot,\cdot\rangle$ is the duality pairing between V' and V (sesquilinear in the first argument and linear in the second argument), and $(\cdot,\cdot)$ is the inner product on H.

We consider an abstract parabolic initial value problem: find

$$u \in L^2((0,T);V) \cap H^1((0,T);V') \hookrightarrow C([0,T];H)$$

such that

$$\begin{cases} \partial_t u(t) - Au(t) = f(t), & 0 < t < T, \\ u(0) = u_0 \in H, \end{cases} \tag{1.1}$$

where $A : V \to V'$ is a bounded linear operator (possibly non-selfa-djoint) with the following properties:

*Received by the editors November 19, 2018; accepted for publication (in revised form) October 21, 2019; published electronically January 8, 2020.
https://doi.org/10.1137/18M1227536
Funding: The work of the first and second authors was partially supported by a Hong Kong RGC grant 15300817, and the work of the third author was partially supported by a Hong Kong RGC grant (25300818).
†Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong (buyang.li@polyu.edu.hk, kai-r.wang@connect.polyu.hk, zhi.zhou@polyu.edu.hk).

$$\beta^{-1}\|u\|_V^2 \le -\mathrm{Re}\langle Au, u\rangle \le \beta\|u\|_V^2 \quad \forall\, u \in V, \tag{P1}$$

$$|((A - A^*)u, u)| \le \varepsilon\|u\|_V^2 + C_\varepsilon\|u\|_H^2 \quad \forall\, u \in V,\ \forall\, \varepsilon \in (0,1), \tag{P2}$$

where β is a constant, and C_ε is a constant depending on ε. Properties (P1)–(P2) imply that the operator A generates a bounded analytic semigroup $E(t)_{t>0}$ on H such that $u(t) = E(t)u_0$ is the solution of (1.1) in the case $f = 0$. This further implies the existence of an angle $\alpha \in (0, \frac{\pi}{2})$ such that (cf. [7, Theorem 3.7.11] and [36])

$$\|(z - A)^{-1}\|_{H\to H} \le C|z|^{-1}, \quad \forall\, z \in \Sigma_{\pi-\alpha} = \{z \in \mathbb{C} : |\arg(z)| \le \pi - \alpha\}. \tag{1.2}$$

The infimum of the angle α satisfying (1.2) will be called the sectorial angle of A. For an elliptic partial differential operator A, we have (cf., e.g., [8, Theorem 1.4])

$$\alpha = \pi - \inf_{\lambda\in\mathcal{N}(A)} |\arg(\lambda)|,$$

where $\mathcal{N}(A)$ denotes the numerical range of A, i.e.,

$$\mathcal{N}(A) = \left\{(Av, v) : v \in \mathrm{D}(A) \ \text{ and } \ \|v\|_{L^2(\Omega)} = 1\right\}.$$

In this paper, we propose and analyze a k-step symmetrized implicit-explicit backward difference formula (BDF) method for (1.1), with $k = 1, \ldots, 6$. For $n \ge 1$, the method seeks $u_n \in V$ such that

$$\begin{aligned}\frac{1}{\tau}\sum_{j=0}^{\min(n,k)} \delta_j(u_{n-j} - u_0) - Du_n = f(t_n) &+ \sum_{j=0}^{\min(n,k)-1} \gamma_j L u_{n-1-j} + c_n^{(k)} L u_0 \\ &+ a_n^{(k)}(Du_0 + f(0)) + \sum_{\ell=1}^{k-2} b_{\ell,n}^{(k)} \tau^\ell \partial_t^\ell f(0),\end{aligned} \tag{1.3}$$

where $D = \frac{1}{2}(A + A^*)$ and $L = \frac{1}{2}(A - A^*)$ denote the self-adjoint and anti-self-adjoint parts of the operator A, respectively, τ denotes the step size of time discretization and $t_n = n\tau$, and the constants δ_j and γ_j are coefficients of the polynomials $\delta(\zeta) = \sum_{\ell=1}^k \frac{1}{\ell}(1-\zeta)^\ell$ and $\gamma(\zeta) = \frac{1}{\zeta}\left[1 - (1-\zeta)^k\right]$, i.e.,

$$\delta(\zeta) = \sum_{i=0}^k \delta_i \zeta^i, \qquad \gamma(\zeta) = \sum_{i=0}^{k-1} \gamma_i \zeta^i.$$

For $n \ge k$, we choose $a_n^{(k)}$, $b_{\ell,n}^{(k)}$, and $c_n^{(k)}$ to be zero. Since $\frac{1}{\tau}\sum_{j=0}^k \delta_j(u_{n-j} - u_0) = \frac{1}{\tau}\sum_{j=0}^k \delta_j u_{n-j}$, it follows that $\frac{1}{\tau}\sum_{j=0}^k \delta_j(u_{n-j} - u_0)$ is the standard k-step approximation of $\partial_t u(t_n)$ by the BDF method, and $\sum_{j=0}^{k-1} \gamma_j L u_{n-1-j}$ approximates $Lu(t_n)$ by the k-step extrapolation. For $1 \le n \le k-1$, these constants will be determined in section 3 to maintain the long-time kth-order accuracy of the numerical scheme for general initial data $u_0 \in V$. The numerical scheme is called symmetrized because one only needs to solve an equation with a self-adjoint operator $\tau^{-1}\delta_0 - D$ at each time step.

The implicit k-step BDF method is known to be A(α_k)-stable with angles $\alpha_1 = \alpha_2 = 0.5\pi, \alpha_3 = 0.478\pi, \alpha_4 = 0.408\pi, \alpha_5 = 0.288\pi$, and $\alpha_6 = 0.099\pi$; see [20, section V.2]. The error estimate for implicit A(α)-stable multistep methods for nonlinear parabolic equations was given by Lubich [27] under the sectorial angle condition $\alpha < \alpha_k$. The convergence of implicit A(α)-stable multistep methods, including A(α_k)-stable BDF methods, for linear parabolic equations was proved by Savaré [34] for operators having properties (P1)–(P2) under the condition

$$|\arg\langle Au, u\rangle| \le \alpha - \varepsilon_0 \quad \forall\, u \in V \quad \text{for some constant } \varepsilon_0.$$

For the implicit k-step BDF method, the condition above implies that the sectorial angle of A defined in (1.2) satisfies $\alpha < \alpha_k$. Stability analysis of implicit multistep methods in a Banach space setting was done by Palencia [32] under the same sectorial angle condition. If an operator does not satisfy (1.2) but satisfies the estimate in (1.2) in a shifted sector $c + \Sigma_{\pi-\alpha}$ with $\alpha < \alpha_k$, then the error estimate still holds but the constant in the error estimate generally grows exponentially as $T \to \infty$.

Implicit-explicit BDF methods were analyzed for linear parabolic equations by Crouzeix [17]. For nonlinear equations, implicit-explicit BDF methods were studied by Akrivis [2, 3] and Akrivis, Crouzeix, and [4] based on spectral and Fourier techniques and studied by Akrivis [1], Akrivis and Katsoprinakis [5], and Akrivis and Lubich [6] based on energy techniques. In these articles, either the assumptions imply (1.2) for some $\alpha < \alpha_k$ or the constant in the error estimate grows exponentially as $T \to \infty$. More recently, Chen et al. [15, 16] have constructed long-time accurate second-order and third-order implicit-explicit methods for a specific Stokes–Darcy system with a nonsymmetric operator, where the analyses are based on energy techniques without using the sectorial angle condition. In particular, the operator in the Stokes–Darcy model has the properties (P1)–(P2) but does not satisfy the sectorial angle condition $\alpha < \alpha_k$ for $k \geq 3$.

In practical computation, many examples of the abstract problem (1.1) have properties (P1)–(P2), but the sectorial angle α may not be smaller than α_k for a given k-step BDF method. Two examples are given in section 2, including the convection-diffusion equation and the coupled Stokes–Darcy system. For the convection-diffusion equation, the proposed method yields a symmetrized semi-implicit scheme. For the coupled Stokes–Darcy system, the proposed method is not only symmetrized but also decoupled (decoupling the Stokes and Darcy equations), a desired property for solving the Stokes–Darcy system.

For these examples, properties (P1)–(P2) guarantee that the argument of the points with large modulus in the numerical range of A is close to π. Therefore, the resolvent estimate in (1.2) may be satisfied for z in a shifted sector $c_k + \Sigma_{\pi-\alpha_k/2}$, but analysis using such a shifted sector will lead to an exponentially growing constant with respect to T in the error estimate. In this paper we will adopt a different approach (without using such a shifted sector) based on the Laplace transform techniques and the observation that the graph of the generating function $\delta(\zeta)$, with $\zeta \in \mathbb{C}$ and $|\zeta| = 1$, is tangent to the vertical axis at the origin. This property guarantees that the contour of $\delta(\zeta)/\tau$, when τ is smaller than some constant, will be on the right side of the points with small modulus in the numerical range of A even if the argument of these points exceeds α_k. For the implicit-explicit method proposed in this paper, this approach requires proving the boundedness of the resolvent operators

$$\left(\tau^{-1}\delta(\zeta) - D - \zeta\gamma(\zeta)L\right)^{-1} L : V \to V, \tag{1.4}$$

$$\left(\tau^{-1}\delta(\zeta) - D - \zeta\gamma(\zeta)L\right)^{-1} D : V \to V \tag{1.5}$$

for ζ in some contour on the complex plane.

The objective of this paper is to determine the coefficients $a_n^{(k)}$, $b_{\ell,n}^{(k)}$, and $c_n^{(k)}$ in (1.3) for the starting $k-1$ steps and to long-time analysis for the proposed symmetrized implicit-explicit BDF method (1.3) without assuming the sectorial angle condition $\alpha < \alpha_k$. In particular, we prove the following theorem.

THEOREM 1.1. *Under the assumptions* (P1)–(P2)*, there exists a positive constant* $\tau_0 > 0$ *such that for* $0 < \tau < \tau_0$*,* $u_0 \in V$*, and* $f \in W^{k+1,1}(\mathbb{R}_+; V') \hookrightarrow C^k([0, t_n]; V')$*, the numerical solution* u_n *given by* (1.3) *satisfies the following error estimate:*

$$\|u_n - u(t_n)\|_V \le c\,\tau^k\left(t_n^{-k}\|u_0\|_V + \sum_{\ell=0}^{k} t_n^{\ell-k}\|\partial_t^\ell f(0)\|_{V'} + \int_0^{t_n}\|\partial_t^{k+1} f(t)\|_{V'}dt\right),$$

where the constant c is independent of τ and t_n.

The modification of the starting $k-1$ steps is motivated by the correction scheme in [28] and [24]. The former was constructed for improving the accuracy in evaluating convolution integrals, and the latter was introduced for an implicit scheme for time-fractional evolution equations. The connection between the two approaches can be found in [24, Appendix A]. From the analysis in the following sections we can see that the coefficients $a_n^{(k)}$ and $b_{\ell,n}^{(k)}$ coincide with the coefficients constructed in [24]. However, for the symmetrized implicit-explicit scheme considered in this paper, an additional term $c_n^{(k)} L u_0$ in (1.3) needs to be introduced to balance the extrapolation error of the anti-self-adjoint part.

The rest of this paper is organized as follows. In section 2, we present two examples of the abstract problem (1.1) that have properties (P1)–(P2) but may violate the sectorial angle condition, i.e., (1.2) may not be satisfied for $\alpha \le \alpha_k$ for a given $k \ge 3$. In section 3, we determine the coefficients $a_n^{(k)}$, $b_{\ell,n}^{(k)}$, and $c_n^{(k)}$ in (1.3) for long-time optimal-order accuracy of the proposed method. In section 4, we present estimates for the operators (1.4)–(1.5) and use the results to prove Theorem 1.1.

Throughout this paper, we denote by c a generic positive constant that is independent of τ and t_n.

2. Examples of the abstract problem (1.1). In this section, we introduce two examples of the abstract problem (1.1) having properties (P1)–(P2) but that may violate the sectorial angle condition $\alpha < \alpha_k$.

Example 2.1 (convection-diffusion equation). Consider the convection-diffusion equation

$$\begin{cases} \partial_t u - \kappa\Delta u + \boldsymbol{v}\cdot\nabla u = f & \text{in } \Omega\times(0,T], \\ u = 0 & \text{on } \partial\Omega\times(0,T], \\ u(0) = u_0 & \text{in } \Omega, \end{cases} \tag{2.1}$$

in a bounded domain $\Omega \subset \mathbb{R}^d$ with a divergence-free velocity field $\boldsymbol{v} \in L^\infty(\Omega)^d$ (incompressible fluid flow), with $d \ge 1$. Problem (2.1) is an example of the abstract problem (1.1) with $V = H_0^1(\Omega)$, $H = L^2(\Omega)$, and the operator $A : V \to V'$ defined (via duality) by

$$\langle Au, w\rangle = -(\kappa\nabla u, \nabla w) - (\boldsymbol{v}\cdot\nabla u, w) \quad \forall\, u, w \in V.$$

Since $\boldsymbol{v}$ is divergence free, we have $D = \frac{1}{2}(A + A^*) = \kappa\Delta$ and $L = \frac{1}{2}(A - A^*) = -\boldsymbol{v}\cdot\nabla$.

It is straightforward to verify properties (P1)–(P2), i.e.,

$$-\mathrm{Re}\langle Au, u\rangle = \kappa\|\nabla u\|_{L^2(\Omega)}^2 \ge C^{-1}\kappa\|u\|_{H_0^1(\Omega)}^2,$$

$$\begin{aligned} |((A - A^*)u, u)| &= 2|(\boldsymbol{v}\cdot\nabla u, u)| \\ &\le 2\|\boldsymbol{v}\|_{L^\infty(\Omega)}\|\nabla u\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)} \\ &\le \varepsilon\|u\|_{H_0^1(\Omega)}^2 + \varepsilon^{-1}\|\boldsymbol{v}\|_{L^\infty(\Omega)}^2\|u\|_{L^2(\Omega)}^2. \end{aligned}$$

However, the sectorial angle α of the operator A depends on the smallness of κ. If κ is sufficiently small compared with $\|\boldsymbol{v}\|_{L^\infty(\Omega)}$, then the sectorial angle α may be close to $\frac{\pi}{2}$ (therefore α may not be smaller than α_k for a given $k \ge 3$).

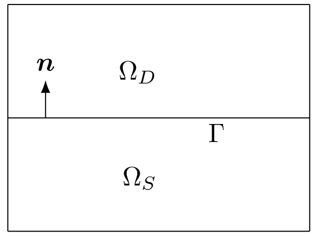

FIG. 1. *Computational domain of the Stokes–Darcy system.*

Example 2.2 (the Stokes–Darcy system). The Stokes–Darcy system describes the coupling of flow in a porous media region $\Omega_D \subset \mathbb{R}^d$ and a free-flow region $\Omega_S \subset \mathbb{R}^d$, with $d \in \{2,3\}$, separated by an interface Γ (as shown in Figure 1), which has many applications in groundwater system [22, 23, 26], industrial filtrations [21, 30], petroleum extraction [9, 14], and so on.

This model consists of a parabolic equation

$$
\begin{cases}
\partial_t\phi - \nabla\cdot(\kappa\nabla\phi) = f_D & \text{in } \ \Omega_D\times(0,T],\\
\phi = 0 & \text{on } \ \partial\Omega_D\backslash\overline{\Gamma}\times(0,T],\\
-\kappa\nabla\phi\cdot\boldsymbol{n} = \boldsymbol{u}\cdot\boldsymbol{n} & \text{on } \ \Gamma\times(0,T],\\
\phi(0) = \phi_0 & \text{in } \ \Omega_D,
\end{cases}
\tag{2.2}
$$

which describes the Darcy flow in the porous media region Ω_D through the unknown hydraulic head ϕ, and a time-dependent Stokes equation

$$
\begin{cases}
\partial_t\boldsymbol{u} - \nabla\cdot\mathbb{T}(\boldsymbol{u},p) = \boldsymbol{f}_S & \text{in } \ \Omega_S\times(0,T],\\
\nabla\cdot\boldsymbol{u} = 0 & \text{in } \ \Omega_S\times(0,T],\\
\boldsymbol{u} = 0 & \text{on } \ \partial\Omega_S\backslash\overline{\Gamma}\times(0,T],\\
-\mathbb{T}(\boldsymbol{u},p)\boldsymbol{n} = g\phi\boldsymbol{n} + \mu(\boldsymbol{u}-(\boldsymbol{u}\cdot\boldsymbol{n})\boldsymbol{n}) & \text{on } \ \Gamma\times(0,T],\\
\boldsymbol{u}(0) = \boldsymbol{u}_0 & \text{in } \ \Omega_S,
\end{cases}
\tag{2.3}
$$

which describes free flow in the region Ω_S through the fluid velocity $\boldsymbol{u}$, where $\mathbb{T}(\boldsymbol{u},p) = 2\nu\mathbb{D}(\boldsymbol{u}) - p\mathbb{I}$ denotes the stress tensor, in which $\mathbb{D}(\boldsymbol{u}) = \frac{1}{2}(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^\top)$ is the deformation tensor and $\mathbb{I}$ is the $d\times d$ identity matrix. The physical parameters κ, g, μ, and ν in this model are positive constants, and f_D and $\boldsymbol{f}_S$ are given source terms.

Homogeneous Dirichlet boundary conditions are imposed on $\partial\Omega_D\backslash\Gamma$ and $\partial\Omega_S\backslash\Gamma$. The interface conditions on Γ in (2.2) and (2.3) represent conservation of mass and balance of force, respectively, where $\boldsymbol{n}$ denotes the unit normal vector on $\partial\Omega_S$ as shown in Figure 1.

Let $H = L^2(\Omega_D)\times L^2(\Omega_S)^d$ and $V = H^1_\Gamma(\Omega_D)\times\mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0)$, where

$$
\begin{aligned}
&H^1_\Gamma(\Omega_D) = \{\varphi\in H^1(\Omega_D) : \varphi = 0 \text{ on } \partial\Omega_D\backslash\Gamma\},\\
&\mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0) = \{\boldsymbol{v}\in H^1(\Omega_S)^d : \nabla\cdot\boldsymbol{v} = 0 \text{ in } \Omega_S \text{ and } \boldsymbol{v} = 0 \text{ on } \partial\Omega_S\backslash\Gamma\}.
\end{aligned}
$$

The weak formulation of (2.2)–(2.3) reads, Find $(\phi,\boldsymbol{u})\in L^2((0,T);V)\cap H^1((0,T);V') \hookrightarrow C([0,T];H)$ satisfying the following equations for all test functions $(\varphi,\boldsymbol{v})\in L^2((0,T);V)$:

$$(\partial_t\phi,\varphi)_D + (\kappa\nabla\phi,\nabla\varphi)_D - (\boldsymbol{u}\cdot\boldsymbol{n},\varphi)_\Gamma = (f_D,\varphi), \tag{2.4}$$

$$(\partial_t\boldsymbol{u},\boldsymbol{v})_S + (2\nu\mathbb{D}(\boldsymbol{u}),\mathbb{D}(\boldsymbol{v}))_S + (g\phi,\boldsymbol{v}\cdot\boldsymbol{n})_\Gamma + \mu(\boldsymbol{u}-(\boldsymbol{u}\cdot\boldsymbol{n})\boldsymbol{n},\boldsymbol{v}-(\boldsymbol{v}\cdot\boldsymbol{n})\boldsymbol{n})_\Gamma = (\boldsymbol{f}_S,\boldsymbol{v}), \tag{2.5}$$

where $(\cdot,\cdot)_D$ is the pairing between $H^1_\Gamma(\Omega_D)'$ and $H^1_\Gamma(\Omega_D)$, $(\cdot,\cdot)_S$ is the pairing between $\mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0)'$ and $\mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0)$, and $(\cdot,\cdot)_\Gamma$ is the inner product on $L^2(\Gamma)$.

Let the operators $A_1 : H^1_\Gamma(\Omega_D)\to H^1_\Gamma(\Omega_D)'$, $A_2 : \mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0)\to\mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0)'$, $B : \mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0)\to H^1_\Gamma(\Omega_D)'$, and $B^* : H^1_\Gamma(\Omega_D)\to\mathbf{H}^1_\Gamma(\Omega_1;\mathrm{div}_0)'$ be defined via duality by

$$\begin{aligned}
&(A_1\phi,\varphi)_D = (\kappa\overline{\nabla\phi},\nabla\varphi)_D && \forall\,\phi,\varphi\in H^1_\Gamma(\Omega_D),\\
&(A_2\boldsymbol{u},\boldsymbol{v})_S = (2\nu\overline{\mathbb{D}(\boldsymbol{u})},\mathbb{D}(\boldsymbol{v}))_S+\mu(\overline{\boldsymbol{u}-(\boldsymbol{u}\cdot\boldsymbol{n})\boldsymbol{n}},\boldsymbol{v}-(\boldsymbol{v}\cdot\boldsymbol{n})\boldsymbol{n})_\Gamma && \forall\,\boldsymbol{u},\boldsymbol{v}\in\mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0),\\
&(B\boldsymbol{u},\varphi)_D = (\overline{\boldsymbol{u}\cdot\boldsymbol{n}},\varphi)_\Gamma \quad \forall\,\boldsymbol{u}\in\mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0)\hookrightarrow L^2(\Gamma), && \forall\,\varphi\in H^1_\Gamma(\Omega_D)\hookrightarrow L^2(\Gamma),\\
&(B^*\phi,\boldsymbol{v})_D = (\phi,\overline{\boldsymbol{v}\cdot\boldsymbol{n}})_\Gamma \quad \forall\,\phi\in H^1_\Gamma(\Omega_D)\hookrightarrow L^2(\Gamma), && \forall\,\boldsymbol{v}\in\mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0)\hookrightarrow L^2(\Gamma).
\end{aligned}$$

Then the weak formulation (2.4)–(2.5) can be written as

$$\partial_t\phi + A_1\phi - B\boldsymbol{u} = f_D, \tag{2.6}$$

$$\partial_t\boldsymbol{u} + A_2\boldsymbol{u} + gB^*\phi = \boldsymbol{f}_S. \tag{2.7}$$

By denoting

$$u = \begin{pmatrix} \phi \\ g^{-\frac12}\boldsymbol{u}\end{pmatrix},\quad f = \begin{pmatrix} f_D \\ g^{-\frac12}\boldsymbol{f}_S\end{pmatrix},\quad \text{and}\quad A = -\begin{pmatrix} A_1 & -g^{\frac12}B \\ g^{\frac12}B^* & A_2\end{pmatrix}, \tag{2.8}$$

(2.6)–(2.7) are equivalent to the abstract problem (1.1). Furthermore, properties (P1)–(P2) hold, i.e.,

$$\begin{aligned}
-\mathrm{Re}\langle Au,u\rangle &= \mathrm{Re}(A_1\phi,\phi)_D + \mathrm{Re}(A_2\boldsymbol{u},\boldsymbol{u})_S\\
&= \kappa\|\nabla\phi\|^2_{L^2(\Omega_D)} + 2\nu\|\mathbb{D}(\boldsymbol{u})\|^2_{L^2(\Omega_S)^3} + \mu\|\boldsymbol{u}-(\boldsymbol{u}\cdot\boldsymbol{n})\boldsymbol{n}\|^2_{L^2(\Gamma)}\\
&\ge \alpha(\|\phi\|^2_{H^1_\Gamma(\Omega_D)} + \|\boldsymbol{u}\|^2_{\mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0)})\\
&= \alpha\|u\|^2_V
\end{aligned}$$

and

$$\begin{aligned}
|((A-A^*)u,u)| &= |-2(B\boldsymbol{u},\phi)_D + 2(B^*\phi,\boldsymbol{u})_D|\\
&= 4\,|\mathrm{Im}(B\boldsymbol{u},\phi)_D|\\
&= 4\,|\mathrm{Im}(\boldsymbol{u}\cdot\boldsymbol{n},\phi)_\Gamma|\\
&\le C\|\boldsymbol{u}\|_{L^2(\Gamma)^3}\|\phi\|_{L^2(\Gamma)}\\
&\le C\|\boldsymbol{u}\|^2_{L^2(\Gamma)^3} + C\|\phi\|^2_{L^2(\Gamma)}\\
&\le \varepsilon(\|\boldsymbol{u}\|^2_{H^1(\Omega_S)^3} + \|\phi\|^2_{H^1(\Omega_D)}) + C_\varepsilon(\|\boldsymbol{u}\|^2_{L^2(\Omega_S)^3} + \|\phi\|^2_{L^2(\Omega_D)})\\
&\le \varepsilon\|u\|^2_V + C_\varepsilon\|u\|^2_H.
\end{aligned}$$

Therefore, the Stokes–Darcy system (2.2)–(2.3) is an example of the abstract problem (1.1) with properties (P1)–(P2). The sectorial angle condition $\alpha<\alpha_k$ may be violated for $k\ge 3$ if the coefficients κ and ν are sufficiently small.

Many time-stepping and decoupling methods have been developed and analyzed for the Stokes–Darcy model [11, 12, 15, 16, 23, 29, 33]. In particular, the long-time accuracy of numerical solutions was investigated in [15, 16] based on long-time regularity assumptions on the solution.

For the Stokes–Darcy system, the proposed method (1.3) can be equivalently written as

$$\begin{aligned}\frac{1}{\tau}\sum_{j=0}^{\min(n,k)}\delta_j(\phi_{n-j}-\phi_0)+A_1\phi_n &= f_D(t_n)+\sum_{j=0}^{\min(n,k)-1}\gamma_j B\boldsymbol{u}_{n-1-j}+c_n^{(k)}B\boldsymbol{u}_0\\ &\quad + a_n^{(k)}(-A_1\phi_0+f_D(0))+\sum_{\ell=1}^{k-2}b_{\ell,n}^{(k)}\tau^\ell\partial_t^\ell f_D(0),\end{aligned}\tag{2.9}$$

$$\begin{aligned}\frac{1}{\tau}\sum_{j=0}^{\min(n,k)}\delta_j(\boldsymbol{u}_{n-j}-\boldsymbol{u}_0)+A_2\boldsymbol{u}_n &= \boldsymbol{f}_S(t_n)-\sum_{j=0}^{\min(n,k)-1}\gamma_j gB^*\phi_{n-1-j}-c_n^{(k)}gB^*\phi_0\\ &\quad + a_n^{(k)}(-A_2\boldsymbol{u}_0+\boldsymbol{f}_S(0))+\sum_{\ell=1}^{k-2}b_{\ell,n}^{(k)}\tau^\ell\partial_t^\ell \boldsymbol{f}_S(0).\end{aligned}\tag{2.10}$$

This scheme decouples ϕ and $\boldsymbol{u}$. At each time step, one only needs to solve two linear systems with symmetric coefficient matrices. The error estimate presented in this paper (for the abstract problem) implies the kth-order convergence of the symmetrized and decoupled method (2.9)–(2.10) for general initial data $\phi_0\in H^1_\Gamma(\Omega_D)$, $\boldsymbol{u}_0\in\mathbf{H}^1_\Gamma(\Omega_S;\mathrm{div}_0)$ and general smooth source terms f_D and $\boldsymbol{f}_S$, without assumptions on the regularity of the solution.

Remark 2.1. For the examples presented in section 2 (and similar examples that can arise from a PDE problem), the self-adjoint part D and anti-self-adjoint part L can be generated using bilinear forms in the weak formulation of the PDE problem (upon finite element discretization in space, for example). This only requires the same computational cost as generating the full matrix A from the bilinear form. In this case, one only needs to solve a system of linear equations with matrix $\tau^{-1}I-D$. This is simpler than solving a system of linear equations with matrix $\tau^{-1}I-A$ (using the fully implicit method).

In general, if the matrix A is given (instead of to be generated from a bilinear form), then we do not need to compute two matrices $D=(A+A^*)/2$ and $L=(A-A^*)/2$ explicitly. Instead, we only need to solve systems of linear equations with matrix D using iterative solvers, which require matrix-vector multiplications of the forms $\frac12 Av+\frac12 A^*v$ and $\frac12 Av-\frac12 A^*v$, while matrix-vector multiplications of type A^*v have the same computational cost as Av. But solving systems of linear equations with matrix D has better convergence (for the iterative solver) than solving systems of linear equations with matrix A. Therefore, the computational cost of using the proposed method is practically much less than that of using the fully implicit scheme.

3. Choice of the coefficients $a_n^{(k)}$, $b_{\ell,n}^{(k)}$, and $c_n^{(k)}$. The method is to compare the expressions of the numerical solution and the exact solution. To this end, we first derive the expression of the numerical solution by using techniques of generating functions and inverse Laplace transform.

By using the Taylor expansion

$$f(t_n)=f(0)+\sum_{\ell=1}^{k-2}\frac{t_n^\ell}{\ell!}\partial_t^\ell f(0)+R_k(t_n),\tag{3.1}$$

where $R_k(t_n) = \frac{t_n^{k-1}}{(k-1)!}\partial_t^{k-1} f(0) + \frac{t_n^k}{k!}\partial_t^k f(0) + \frac{t_n^k}{k!} * \partial_t^{k+1} f(t_n)$ denotes the remainder, and $\frac{t_n^k}{k!} * \partial_t^{k+1} f(t_n) = \int_0^{t_n} \frac{(t_n-t)^k}{k!}\partial_t^{k+1} f(t)dt$. We introduce $v_n = u_n - u_0$ and rewrite (1.3) as

$$\begin{aligned}\frac{1}{\tau}\sum_{j=0}^{\min(n,k)} \delta_j v_{n-j} - Dv_n = \sum_{j=0}^{\min(n,k)-1} \gamma_j L v_{n-1-j} + (d_n^{(k)} + c_n^{(k)})Lu_0 + R_k(t_n)\\ + (1 + a_n^{(k)})(Du_0 + f(0)) + \sum_{\ell=1}^{k-2}\left(\frac{t^\ell}{\ell!} + b_{\ell,n}^{(k)}\tau^\ell\right)\partial_t^\ell f(0),\end{aligned} \tag{3.2}$$

where $d_n^{(k)} = \sum_{j=0}^{\min(n,k)-1}\gamma_j$. For any given bounded sequence $(w_n)_{n=1}^\infty$ we define its generating function

$$\widetilde{w}(\zeta) = \sum_{n=1}^\infty w_n \zeta^n \quad \text{for } \zeta \in \mathbb{C},\ |\zeta| < 1. \tag{3.3}$$

Then, multiplying (3.2) by ζ^n and summing up the result for $n = 1, 2, \ldots$, we obtain

$$\begin{aligned}&(\tau^{-1}\delta(\zeta) - D - \zeta\gamma(\zeta)L)\widetilde{v}(\zeta)\\ &= \left(\sum_{n=1}^\infty \zeta^n + \sum_{n=1}^{k-1} a_n^{(k)}\zeta^n\right)(Du_0 + f(0)) + \left(\sum_{n=1}^\infty \zeta^n + \sum_{n=1}^{k-1}(d_n^{(k)} + c_n^{(k)} - 1)\zeta^n\right)Lu_0\\ &\quad + \sum_{\ell=1}^{k-2}\left(\sum_{n=1}^\infty \frac{t_n^\ell}{\ell!}\zeta^n + \sum_{n=1}^{k-1} b_{\ell,n}^{(k)}\tau^\ell\zeta^n\right)\partial_t^\ell f(0) + \widetilde{R}_k(\zeta)\\ &= \delta(\zeta)^{-1}\mu(\zeta)(Du_0 + f(0)) + \delta(\zeta)^{-1}\chi(\zeta)Lu_0 + \sum_{\ell=1}^{k-2}\eta_\ell(\zeta)\tau^\ell\partial_t^\ell f(0) + \widetilde{R}_k(\zeta),\end{aligned}$$

where

$$\mu(\zeta) = \delta(\zeta)\left(\frac{\zeta}{1-\zeta} + \sum_{n=1}^{k-1} a_n^{(k)}\zeta^n\right), \tag{3.4}$$

$$\chi(\zeta) = \delta(\zeta)\left(\frac{\zeta}{1-\zeta} + \sum_{n=1}^{k-1}(d_n^{(k)} + c_n^{(k)} - 1)\zeta^n\right), \tag{3.5}$$

$$\eta_\ell(\zeta) = \frac{1}{\ell!}\left(\zeta\frac{\mathrm{d}}{\mathrm{d}\zeta}\right)^\ell \frac{1}{1-\zeta} + \sum_{n=1}^{k-1} b_{\ell,n}^{(k)}\zeta^n, \tag{3.6}$$

$$\widetilde{R}_k(\zeta) = \sum_{n=1}^\infty R_k(t_n)\zeta^n. \tag{3.7}$$

Let

$$K(z,\xi) = z^{-1}(zI - D - \xi L)^{-1}. \tag{3.8}$$

Then, we have

$$\begin{aligned}\widetilde{v}(\zeta) &= K\left(\tau^{-1}\delta(\zeta), \zeta\gamma(\zeta)\right)\left(\tau^{-1}\mu(\zeta)(Du_0 + f(0)) + \tau^{-1}\chi(\zeta)Lu_0\right)\\ &\quad + K\left(\tau^{-1}\delta(\zeta), \zeta\gamma(\zeta)\right)\left(\tau^{-1}\delta(\zeta)\sum_{\ell=1}^{k-2}\eta_\ell(\zeta)\tau^\ell\partial_t^\ell f(0) + \tau^{-1}\delta(\zeta)\widetilde{R}_k(\zeta)\right).\end{aligned} \tag{3.9}$$

By using Cauchy's integral formula, we obtain for arbitrary $0 < \rho < 1$

$$v_n = \frac{1}{2\pi \mathrm{i}} \int_{|\zeta|=\rho} \zeta^{-n-1} \widetilde{v}(\zeta) d\zeta. \tag{3.10}$$

Then, by a change of variable $\zeta = e^{-z\tau}$ we obtain

$$v_n = \frac{\tau}{2\pi \mathrm{i}} \int_{\Gamma^\tau} e^{zt_n} \widetilde{v}(e^{-z\tau})\, dz, \tag{3.11}$$

with Γ^τ being a vertical segment, i.e.,

$$\Gamma^\tau := \{z = -\ln(\rho)/\tau + \mathrm{i}y : y \in \mathbb{R} \text{ and } |y| \le \pi/\tau\}. \tag{3.12}$$

Next, we deform the contour Γ^τ in the integral (3.11) to a contour

$$\Gamma_\theta^\tau = \{z \in \mathbb{C} : |\arg(z)| = \theta,\ |\mathrm{Im}(z)| \le \pi/\tau\} \quad \text{for some } \theta > \frac{\pi}{2}, \tag{3.13}$$

which is on the left half of the complex plane. Since the $\mathrm{A}(\alpha_k)$-stability of the k-step BDF method says that

$$\delta(e^{-z\tau}) \in \Sigma_{\pi-\alpha_k} = \{z \in \mathbb{C} : |\arg(z)| < \pi - \alpha_k\}$$

for

$$z \in \Sigma_{\pi/2}^\tau = \{z \in \mathbb{C} : |\arg(z)| < \pi/2,\ |\mathrm{Im}(z)| \le \pi/\tau\},$$

it follows that, by slightly increasing the angle of z, there exists an angle $\theta \in (\frac{\pi}{2}, \pi)$ (sufficiently close to $\frac{\pi}{2}$) such that

$$\delta(e^{-z\tau}) \in \Sigma_{\pi-\alpha_k/2} = \{z \in \mathbb{C} : |\arg(z)| < \pi - \alpha_k/2\} \quad \text{for } z \in \Sigma_\theta^\tau. \tag{3.14}$$

A proof of this can be found in [24, Lemma B.1]. Therefore, the function $\widetilde{v}(e^{-z\tau})$ defined by (3.9) is analytic for $z \in \Sigma_\theta^\tau$, which is enclosed by Γ_θ^τ and the two horizontal lines $\Gamma_\pm^\tau := \{x \pm \mathrm{i}\pi/\tau : x \ge \cot(\theta)\pi/\tau\}$.

Due to the periodicity of the function $e^{zt_n}\widetilde{v}(e^{-z\tau})$ in the imaginary direction of z, the values of $e^{zt_n}\widetilde{v}(e^{-z\tau})$ on the two horizontal lines $\Gamma_\pm^\tau$ coincide. Therefore, Cauchy's theorem allows deforming the contour Γ^τ to Γ_θ^τ in the integral (3.11) (integrations on $\Gamma_\pm^\tau$ cancel each other), which yields

(3.15)

$$\begin{aligned}
v_n &= \frac{\tau}{2\pi \mathrm{i}} \int_{\Gamma_\theta^\tau} e^{zt_n} \widetilde{v}(e^{-z\tau}) dz \\
&= \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta^\tau} e^{zt_n} K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))\mu(e^{-z\tau})(Du_0 + f(0)) dz \\
&\quad + \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta^\tau} e^{zt_n} K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))\chi(e^{-z\tau}) L u_0 dz \\
&\quad + \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta^\tau} e^{zt_n} \tau^{-1}\delta(e^{-z\tau})) K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau})) \sum_{\ell=1}^{k-2} \eta_\ell(e^{-z\tau})\tau^{\ell+1}\partial_t^\ell f(0) dz \\
&\quad + \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta^\tau} e^{zt_n} \tau^{-1}\delta(e^{-z\tau})) K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))\, \tau \widetilde{R}_k(e^{-z\tau}) dz \\
&= \sum_{j=1}^{4} I_j^{\tau,n}.
\end{aligned}$$

Similarly, the function $v(t) := u(t) - u_0$ satisfies $v(0) = 0$ and

$$\partial_t v - Av = Au_0 + f(0) + \sum_{\ell=1}^{k-2} \frac{t^\ell}{\ell!}\partial_t^\ell f(0) + R_k. \tag{3.16}$$

By taking Laplace transform in time and using the identity $\widehat{\partial_t v}(z) = z\widehat{v}(z)$, we obtain

$$z\widehat{v}(z) - A\widehat{v}(z) = z^{-1}(Au_0 + f(0)) + \sum_{\ell=1}^{k-2} \frac{1}{z^{\ell+1}}\partial_t^\ell f(0) + \widehat{R}_k(z).$$

Then, the inverse Laplace transform of the above equation yields

$$\begin{aligned} v(t) =& \frac{1}{2\pi \mathrm{i}}\int_{\Gamma_\theta} e^{zt}K(z,1)\big(Du_0 + f(0)\big)dz \\ &+ \frac{1}{2\pi \mathrm{i}}\int_{\Gamma_\theta} e^{zt}K(z,1)Lu_0 dz \\ &+ \frac{1}{2\pi \mathrm{i}}\int_{\Gamma_\theta} e^{zt}zK(z,1)\sum_{\ell=1}^{k-2}\frac{1}{z^{\ell+1}}\partial_t^\ell f(0)dz \\ &+ \frac{1}{2\pi \mathrm{i}}\int_{\Gamma_\theta} e^{zt}zK(z,1)\widehat{R}_k(z)dz \\ =& \sum_{j=1}^{4} I_j(t), \end{aligned} \tag{3.17}$$

where

$$\Gamma_\theta = \{z \in \mathbb{C} : |\arg(z)| = \theta\}. \tag{3.18}$$

It is known that the k-step BDF method automatically satisfies (see [36, equation (10.6)])

$$|\delta_\tau(e^{-z\tau}) - z| \le c|z|^{k+1}\tau^k, \qquad |e^{-z\tau}\gamma(e^{-z\tau}) - 1| \le c|z|^k\tau^k. \tag{3.19}$$

By comparing the kernel functions in (3.15) and (3.17), in order to have kth-order accuracy, we need to choose the coefficients $a_n^{(k)}$, $b_{\ell,n}^{(k)}$, and $c_n^{(k)}$ to satisfy the following kth-order conditions (see (3.4)–(3.6)):

$$|\mu(e^{-z\tau}) - 1| \le c|z|^k\tau^k, \tag{3.20}$$

$$|\chi(e^{-z\tau}) - 1| \le c|z|^k\tau^k, \tag{3.21}$$

$$\left|\eta_\ell(e^{-z\tau})\tau^{\ell+1} - \frac{1}{z^{\ell+1}}\right| \le c|z|^{k-\ell-1}\tau^k. \tag{3.22}$$

If these conditions are satisfied for $z \in \Gamma_\theta^\tau$, then kth-order convergence can be proved (see section 4). Since $|1 - e^{-z\tau}| \le c|z|\tau$ for $z \in \Gamma_\theta^\tau$, by using a change of variables $\zeta = e^{-z\tau}$ we only need to impose the following conditions:

$$|\mu(\zeta) - 1| \le c|1-\zeta|^k, \tag{3.23}$$

$$|\chi(\zeta) - 1| \le c|1-\zeta|^k, \tag{3.24}$$

$$\left|\frac{\gamma_\ell(\zeta)}{\ell!} + \sum_{j=1}^{k-1} b_{\ell,j}^{(k)}\zeta^j - \frac{1}{\delta(\zeta)^{\ell+1}}\right| \le c|1-\zeta|^{k-\ell-1}, \quad \ell = 1, \ldots, k-2. \tag{3.25}$$

The first and third conditions determine the coefficients $a_n^{(k)}$ and $b_{\ell,n}^{(k)}$ similarly as in [24], as shown in Tables 1–2. The second condition above determines

$$c_n^{(k)} = 1 + a_n^{(k)} - d_n^{(k)}, \tag{3.26}$$

as shown in Table 3. The new coefficients $c_n^{(k)}$ are caused by the extrapolation term $\sum_{j=0}^{\min(n,k)-1} \gamma_j L u_{n-1-j}$, which does not enter the fully implicit scheme studied in [24].

TABLE 1
The coefficients $a_n^{(k)}$.

Order of BDF	$a_1^{(k)}$	$a_2^{(k)}$	$a_3^{(k)}$	$a_4^{(k)}$	$a_5^{(k)}$
$k = 2$	$\frac{1}{2}$				
$k = 3$	$\frac{11}{12}$	$-\frac{5}{12}$			
$k = 4$	$\frac{31}{24}$	$-\frac{7}{6}$	$\frac{3}{8}$		
$k = 5$	$\frac{1181}{720}$	$-\frac{177}{80}$	$\frac{341}{240}$	$-\frac{251}{720}$	
$k = 6$	$\frac{2837}{1440}$	$-\frac{2543}{720}$	$\frac{17}{5}$	$-\frac{1201}{720}$	$\frac{95}{288}$

TABLE 2
The coefficients $b_{\ell,n}^{(k)}$.

Order of BDF		$b_{\ell,1}^{(k)}$	$b_{\ell,2}^{(k)}$	$b_{\ell,3}^{(k)}$	$b_{\ell,4}^{(k)}$	$b_{\ell,5}^{(k)}$
$k = 3$	$\ell = 1$	$\frac{1}{12}$	0			
$k = 4$	$\ell = 1$	$\frac{1}{6}$	$-\frac{1}{12}$	0		
	$\ell = 2$	0	0	0		
$k = 5$	$\ell = 1$	$\frac{59}{240}$	$-\frac{29}{120}$	$\frac{19}{240}$	0	
	$\ell = 2$	$\frac{1}{240}$	$-\frac{1}{240}$	0	0	
	$\ell = 3$	$\frac{1}{720}$	0	0	0	
$k = 6$	$\ell = 1$	$\frac{77}{240}$	$-\frac{7}{15}$	$\frac{73}{240}$	$-\frac{3}{40}$	0
	$\ell = 2$	$\frac{1}{96}$	$-\frac{1}{60}$	$\frac{1}{160}$	0	0
	$\ell = 3$	$-\frac{1}{360}$	$\frac{1}{720}$	0	0	0
	$\ell = 4$	0	0	0	0	0

TABLE 3
The coefficients $c_n^{(k)}$.

Order of BDF	$c_1^{(k)}$	$c_2^{(k)}$	$c_3^{(k)}$	$c_4^{(k)}$	$c_5^{(k)}$
$k = 2$	$-\frac{1}{2}$				
$k = 3$	$-\frac{13}{12}$	$\frac{7}{12}$			
$k = 4$	$-\frac{41}{24}$	$\frac{11}{6}$	$-\frac{5}{8}$		
$k = 5$	$-\frac{1699}{720}$	$\frac{303}{80}$	$-\frac{619}{240}$	$\frac{469}{720}$	
$k = 6$	$-\frac{4363}{1440}$	$\frac{4657}{720}$	$-\frac{33}{5}$	$\frac{2399}{720}$	$-\frac{193}{288}$

4. Long-time error analysis. In this section, we prove Theorem 1.1, which is concerned with long-time error of the numerical solution.

From the solution representation formula (3.15) we see that estimates of the following discrete resolvents are needed:

$$K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau})),$$
$$K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))D,$$
$$K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))L,$$

where $K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))$ is defined in (3.8). These discrete resolvents are estimated in section 4.1. The differences between these discrete resolvents and the continuous resolvents are estimated in section 4.2. By using these results, we present error estimates in section 4.3 to prove Theorem 1.1.

4.1. Resolvent estimates.

LEMMA 4.1. *There exist* $\tau_0 > 0$, $\theta \in (\frac{\pi}{2}, \pi)$ *and a positive constant* c*, depending on the constants in* (P1)–(P2)*, such that for* $0 < \tau < \tau_0$ *the operator*

$$\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L \quad \textit{with } z \in \Gamma_\theta^\tau$$

is invertible and satisfies the following estimate:

$$\|(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}\|_{H\to H} \le c|z|^{-1}. \tag{4.1}$$

Proof. There exist $\theta_0 \in (\frac{\pi}{2}, \pi)$ and positive constants c_1, c_2 such that for $\theta \in (\frac{\pi}{2}, \theta_0]$ the following estimates hold (cf. [24, Lemma B.1]):

$$c_1|z| \le |\tau^{-1}\delta(e^{-\tau z})| \le c_2|z| \qquad \forall\, z \in \Gamma_\theta^\tau, \tag{4.2}$$
$$|\tau^{-1}\delta(e^{-\tau z}) - z|/|z| \le c\tau^k|z|^k \qquad \forall\, z \in \Gamma_\theta^\tau, \tag{4.3}$$
$$\tau^{-1}\delta(e^{-\tau z}) \in \Sigma_{\pi-\alpha_k/2} \qquad \forall\, z \in \Gamma_\theta^\tau. \tag{4.4}$$

If $\varphi \in H$ satisfies $(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)\varphi = f \in H$, then

$$\begin{aligned}\mathrm{Re}(f,\varphi) &= \mathrm{Re}((\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)\varphi, \varphi)\\ &= \mathrm{Re}(\tau^{-1}\delta(e^{-\tau z}))\|\varphi\|_H^2 - \mathrm{Re}(A\varphi,\varphi) + \mathrm{Re}((1 - e^{-\tau z}\gamma(e^{-\tau z}))L\varphi, \varphi)\\ &\ge \mathrm{Re}(\tau^{-1}\delta(e^{-\tau z}))\|\varphi\|_H^2 + \beta^{-1}\|\varphi\|_V^2 - |1 - e^{-\tau z}\gamma(e^{-\tau z})|(\varepsilon\|\varphi\|_V^2 + C_\varepsilon\|\varphi\|_H^2), \end{aligned} \tag{4.5}$$

where we have used (P1)–(P2) in the last inequality. Besides, we also have

$$\begin{aligned}|\mathrm{Im}(f,\varphi)| &= |\mathrm{Im}((\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)\varphi, \varphi)|\\ &\ge |\mathrm{Im}\,\tau^{-1}\delta(e^{-\tau z})|\|\varphi\|_H^2 - |e^{-\tau z}\gamma(e^{-\tau z})||(L\varphi,\varphi)|\\ &\ge |\mathrm{Im}\,\tau^{-1}\delta(e^{-\tau z})|\|\varphi\|_H^2 - c(\varepsilon\|\varphi\|_V^2 + C_\varepsilon\|\varphi\|_H^2), \end{aligned} \tag{4.6}$$

where we have used the fact that $|e^{-\tau z}| \le c$ for $z \in \Gamma_\theta^\tau$.

If $\mathrm{Re}(\tau^{-1}\delta(e^{-\tau z})) \ge 0$, then

$$\begin{aligned}&\frac{1}{2}|\tan(\pi - \alpha_k/2)|\mathrm{Re}(\tau^{-1}\delta(e^{-\tau z})) + |\mathrm{Im}(\tau^{-1}\delta(e^{-\tau z}))|\\ &\ge \min\left(\frac{1}{2}|\tan(\pi - \alpha_k/2)|, 1\right)(|\mathrm{Re}(\tau^{-1}\delta(e^{-\tau z}))| + |\mathrm{Im}(\tau^{-1}\delta(e^{-\tau z}))|)\\ &\ge \min\left(\frac{1}{2}|\tan(\pi - \alpha_k/2)|, 1\right)|\tau^{-1}\delta(e^{-\tau z})|. \end{aligned}$$

Otherwise, $\mathrm{Re}(\tau^{-1}\delta(e^{-\tau z})) \le 0$ and

$$\mathrm{Re}(\tau^{-1}\delta(e^{-\tau z})) \ge \cos(\pi - \alpha_k/2)|\tau^{-1}\delta(e^{-\tau z})|,$$
$$|\mathrm{Im}(\tau^{-1}\delta(e^{-\tau z}))| \ge \sin(\pi - \alpha_k/2)|\tau^{-1}\delta(e^{-\tau z})|.$$

In this case,

$$\frac{1}{2}|\tan(\pi - \alpha_k/2)|\mathrm{Re}(\tau^{-1}\delta(e^{-\tau z})) + |\mathrm{Im}(\tau^{-1}\delta(e^{-\tau z}))| \ge \frac{1}{2}\sin(\pi - \alpha_k/2)|\tau^{-1}\delta(e^{-\tau z})|.$$

In view of (4.2), there exists a constant c_3 such that

$$\frac{1}{2}|\tan(\pi - \alpha_k/2)|\mathrm{Re}(\tau^{-1}\delta(e^{-\tau z})) + |\mathrm{Im}(\tau^{-1}\delta(e^{-\tau z}))| \ge c_3|z|.$$

Combining (4.5) and (4.6), we obtain

$$c_3|z|\|\varphi\|_H^2 + \beta^{-1}\|\varphi\|_V^2 \le c_4|(f,\varphi)| + c_4(\varepsilon\|\varphi\|_V^2 + C_\varepsilon\|\varphi\|_H^2). \tag{4.7}$$

By choosing ε sufficiently small, the term $c_4\varepsilon\|\varphi\|_V^2$ can be absorbed by the left side. We obtain

$$c_3|z|\|\varphi\|_H^2 + \frac{\beta^{-1}}{2}\|\varphi\|_V^2 \le c_4|(f,\varphi)| + c_5\|\varphi\|_H^2.$$

If $|z| \ge 2c_3^{-1}c_5$, then the last term on the right side can be absorbed by the left side, and therefore

$$|z|\|\varphi\|_H^2 \le 2c_3^{-1}c_4|(f,\varphi)| \le 8c_3^{-2}c_4^2|z|^{-1}\|f\|_H^2 + \frac{1}{2}|z|\|\varphi\|_H^2,$$

which implies

$$\|\varphi\|_H \le 4c_3^{-1}c_4|z|^{-1}\|f\|_H \quad \text{if } |z| \ge 2c_3^{-1}c_5. \tag{4.8}$$

If $|z| \le 2c_3^{-1}c_5$, then (4.3) implies

$$|\arg(\tau^{-1}\delta(e^{-\tau z})) - \arg(z)| \le c_6\tau^k,$$

which together with (4.5) implies

$$\begin{aligned}\mathrm{Re}(f,\varphi) &\ge \cos(\theta + c_6\tau^k)|z|\|\varphi\|_H^2 + \alpha\|\varphi\|_V^2 - \tau^k|z|^k(\varepsilon_1\|\varphi\|_V^2 + C_{\varepsilon_1}\|\varphi\|_H^2)\\ &\ge \cos(\theta + c_6\tau^k)4c_3^{-1}c_5\|\varphi\|_H^2 + \alpha\|\varphi\|_V^2 - \tau^k 4c_3^{-1}c_5(\varepsilon_1\|\varphi\|_V^2 + C_{\varepsilon_1}\|\varphi\|_H^2).\end{aligned}$$

If τ is sufficiently small and θ is sufficiently close to $\frac{\pi}{2}$, then the first and third terms on the right side above are smaller than $\frac{\alpha}{2}\|\varphi\|_V^2$, and the inequality reduces to

$$\frac{\alpha}{2}\|\varphi\|_V^2 \le \mathrm{Re}(f,\varphi), \tag{4.9}$$

which implies

$$\|\varphi\|_V \le c\|f\|_H \le c2c_3^{-1}c_5|z|^{-1}\|f\|_H \quad \text{if } |z| \le 2c_3^{-1}c_5. \tag{4.10}$$

Combining the two estimates (4.8) and (4.10), we obtain the desired estimate (4.1).

The above analysis shows that if φ is a solution of $(\tau^{-1}\delta(e^{-\tau z})-D-e^{-\tau z}\gamma(e^{-\tau z})L)\varphi = f \in V \hookrightarrow H$, then (4.1) holds. This implies that the operator

$$\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L : V \to V'$$

is one-to-one. By using the same analysis as above, one can prove that the dual operator $\tau^{-1}\delta(e^{-\tau\bar{z}}) - D - e^{-\tau\bar{z}}\gamma(e^{-\tau\bar{z}})L^*$ is also one-to-one. Since both the operator and its dual operator are one-to-one, it follows that the operator is invertible (both one-to-one and onto). ∎

LEMMA 4.2. *Let τ_0 and θ be as required in Lemma 4.1. For $0 < \tau < \tau_0$ the following estimates hold:*

$$\|(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}L\|_{V\to V} \le c \qquad \forall\, z \in \Gamma_\theta^\tau, \tag{4.11}$$

$$\|(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}D\|_{V\to V} \le c \qquad \forall\, z \in \Gamma_\theta^\tau, \tag{4.12}$$

$$\|(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}\|_{V\to V} \le c|z|^{-1} \qquad \forall\, z \in \Gamma_\theta^\tau. \tag{4.13}$$

Proof. Let $\varphi \in V$ be the solution of

$$(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)\varphi = Lf \in V'.$$

Then replacing f by Lf in (4.7) yields

$$c_3|z|\|\varphi\|_H^2 + \beta^{-1}\|\varphi\|_V^2 \le c_4|(Lf,\varphi)| + c_4(\varepsilon\|\varphi\|_V^2 + C_\varepsilon\|\varphi\|_H^2).$$

By choosing $\varepsilon = \beta^{-1}/2c_4$, we obtain that

$$c_3|z|\|\varphi\|_H^2 + \frac{\beta^{-1}}{2}\|\varphi\|_V^2 \le c_4|(Lf,\varphi)| + c_5\|\varphi\|_H^2.$$

If $|z| \ge 2c_3^{-1}c_5$, then

$$c_5\|\varphi\|_H^2 + \frac{\beta^{-1}}{2}\|\varphi\|_V^2 \le c_4|(Lf,\varphi)| \le c_6\|f\|_V\|\varphi\|_V \le \frac{c_6^2}{\alpha}\|f\|_V^2 + \frac{\alpha}{4}\|\varphi\|_V^2,$$

which implies $\|\varphi\|_V \le c\|f\|_V$.

If $|z| \le 2c_3^{-1}c_5$, then replacing f by Lf in (4.9) yields

$$\frac{\alpha}{2}\|\varphi\|_V^2 \le \mathrm{Re}(Lf,\varphi) \le c\|f\|_V\|\varphi\|_V,$$

which again implies $\|\varphi\|_V \le c\|f\|_V$.

In either case, we have proved $\|\varphi\|_V \le c\|f\|_V$, which implies (4.11).

Replacing Lf by Df in the analysis above yields (4.12). Moreover, (4.11) and (4.12) together imply

$$\|(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}(D + e^{-\tau z}\gamma(e^{-\tau z})L)\|_{V\to V} \le c.$$

Since

$$\begin{aligned}&(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}\tau^{-1}\delta(e^{-\tau z})\\ &= I + (\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}(D + e^{-\tau z}\gamma(e^{-\tau z})L),\end{aligned}$$

with I denoting the identity operator on V, it follows that

$$\|(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}\tau^{-1}\delta(e^{-\tau z})\|_{V\to V} \le c.$$

Since $|\tau^{-1}\delta(e^{-\tau z})| \ge c|z|$, as shown in (4.2), the estimate above implies (4.13). ∎

By taking limit $\tau \to 0$ in Lemma 4.2, we also obtain the following estimates for the continuous operators.

LEMMA 4.3. *The operator A satisfies the following estimates for all $z \in \Gamma_\theta$:*

$$\|(z-A)^{-1}\|_{V\to V} \le c|z|^{-1},$$
$$\|(z-A)^{-1}L\|_{V\to V} \le c,$$
$$\|(z-A)^{-1}D\|_{V\to V} \le c.$$

4.2. Error between resolvents. By using Lemmas 4.2–4.3 we can prove the following estimates for the differences between the discrete and continuous resolvent operators.

LEMMA 4.4. *Let τ_0 and θ be as required in Lemma 4.1. For $0 < \tau < \tau_0$ the following estimates hold:*

$$\|(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}L - (z-A)^{-1}L\|_{V\to V} \le c\tau^k|z|^k, \qquad z \in \Gamma_\theta^\tau,$$
$$\|(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}D - (z-A)^{-1}D\|_{V\to V} \le c\tau^k|z|^k, \qquad z \in \Gamma_\theta^\tau.$$

Proof. Let u and v be solutions of

$$(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)u = Lf,$$
$$(z-A)v = Lf,$$

respectively. Then the difference between the two equations gives

$$(z-A)(v-u) = (\tau^{-1}\delta(e^{-\tau z}) - z)u + (1 - e^{-\tau z}\gamma(e^{-\tau z}))Lu,$$

and therefore

$$v - u = (\tau^{-1}\delta(e^{-\tau z}) - z)(z-A)^{-1}u + (1 - e^{-\tau z}\gamma(e^{-\tau z}))(z-A)^{-1}Lu.$$

By using (4.3) and the identity $1 - e^{-\tau z}\gamma(e^{-\tau z}) = (1-e^{-\tau z})^k$, we have

$$\begin{aligned}
\|v-u\|_V &\le |\tau^{-1}\delta(e^{-\tau z}) - z|\|(z-A)^{-1}u\|_V + |1 - e^{-\tau z}\gamma(e^{-\tau z})|\|(z-A)^{-1}Lu\|_V \\
&\le c\tau^k|z|^{k+1}\|(z-A)^{-1}u\|_V + c\tau^k|z|^k\|(z-A)^{-1}Lu\|_V \\
&\le c\tau^k|z|^k\|u\|_V \quad \text{(Lemma 4.3 is used here)} \\
&= c\tau^k|z|^k\|(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1}Lf\|_V \\
&\le c\tau^k|z|^k\|f\|_V \quad \text{(Lemma 4.2 is used here).}
\end{aligned}$$

This proves the first inequality of Lemma 4.4. The second inequality can be proved similarly. ∎

Since $K(z,1) = z^{-1}(z-A)^{-1}$ and

$$K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau})) = (\tau^{-1}\delta(e^{-\tau z}))^{-1}(\tau^{-1}\delta(e^{-\tau z}) - D - e^{-\tau z}\gamma(e^{-\tau z})L)^{-1},$$

the estimates (3.20)–(3.22) and Lemmas 4.2–4.4 immediately imply the following results (the proof is omitted).

LEMMA 4.5. *Let τ_0 and θ be as required in Lemma 4.1. For $0 < \tau < \tau_0$ the following estimates hold for $z \in \Gamma_\theta^\tau$:*

$$\left\|\left[K(z,1) - K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))\mu(e^{-z\tau})\right]D\right\|_{V\to V} \le c\tau^k|z|^{k-1},$$

$$\left\|\left[z^{-\ell-1}zK(z,1) - \tau^{\ell+1}\eta_\ell(e^{-z\tau})\tau^{-1}\delta(e^{-z\tau})K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))\right]D\right\|_{V\to V}$$
$$\le c\tau^k|z|^{k-\ell-1}.$$

4.3. Error estimate. In this subsection, we present an error estimate for the proposed method (1.3) by using Lemmas 4.2–4.5. To this end, we recall the solution representations (3.15) and (3.17) and split the error into four parts, i.e.,

$$\|u(t_n) - u_n\|_V = \|v(t_n) - v_n\|_V \le \sum_{j=1}^{4} \|I_j(t_n) - I_j^{\tau,n}\|_V.$$

For each $\|I_j(t_n) - I_j^{\tau,n}\|_V$, $j = 1,2,3,4$, we establish an error bound in the lemma below.

LEMMA 4.6. *Let τ_0 be as required in Lemma* 4.1, *and assume that* $0 < \tau < \tau_0$. *For $u_0 \in V$ and $f(0) \in V'$, the following estimate holds:*

$$\|I_1(t_n) - I_1^{\tau,n}\|_V + \|I_2(t_n) - I_2^{\tau,n}\|_V \le c\tau^k t_n^{-k} \left(\|u_0\|_V + \|f(0)\|_{V'}\right).$$

Proof. We split the error into two parts, i.e.,

$$\begin{aligned}
&I_1(t_n) - I_1^{\tau,n} \\
&= \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta^\tau} e^{zt_n} \Big(K(z,1) - K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))\mu(e^{-z\tau}) \Big) D(u_0 + D^{-1}f(0))\, dz \\
&\quad + \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta \backslash \Gamma_\theta^\tau} e^{zt_n} K(z,1) D\big(u_0 + D^{-1}f(0)\big) dz \\
&=: \mathrm{I} + \mathrm{II}.
\end{aligned}$$

By using Lemma 4.5, we bound the first term I by

$$\begin{aligned}
\|\mathrm{I}\|_V &\le c\tau^k \|u_0 + D^{-1}f(0)\|_V \int_{\Gamma_\theta^\tau} e^{\mathrm{Re}(z)t_n} |z|^{k-1} |dz| \\
&\le c\tau^k \|u_0 + D^{-1}f(0)\|_V \int_0^{\pi/(\tau\sin\theta)} e^{-rt_n|\cos\theta|} r^{k-1} dr \\
&\le c\tau^k t_n^{-k} \|u_0 + D^{-1}f(0)\|_V \\
&\le c\tau^k t_n^{-k} (\|u_0\|_V + \|f(0)\|_{V'}).
\end{aligned}$$

For the second term II, we appeal to Lemma 4.3, which implies

$$\|K(z,1)D\|_{V\to V} \le c|z|^{-1},$$

and therefore

$$\begin{aligned}
\|\mathrm{II}\|_V &\le c\|u_0 + D^{-1}f(0)\|_V \int_{\Gamma_\theta^\tau} e^{\mathrm{Re}(z)t_n} |z|^{-1} |dz| \\
&\le c\|u_0 + D^{-1}f(0)\|_V \int_{\pi/(\tau\sin\theta)}^{\infty} e^{-rt_n|\cos\theta|} r^{-1}\, dr \\
&\le c\tau^k \|u_0 + D^{-1}f(0)\|_V \int_0^{\infty} e^{-rt_n|\cos\theta|} r^{k-1}\, dr \quad (1 \le \tau^k|z|^k \text{ is used here}) \\
&\le c\tau^k t_n^{-k} \|u_0 + D^{-1}f(0)\|_V \\
&\le c\tau^k t_n^{-k} (\|u_0\|_V + \|f(0)\|_{V'}).
\end{aligned}$$

This proves the desired estimate for $\|I_1(t_n) - I_1^{\tau,n}\|_V$.

Since χ and μ satisfy the same estimates, i.e., (3.20)–(3.21), the above proof for $\|I_1(t_n) - I_1^{\tau,n}\|_V$ also works for $\|I_2(t_n) - I_2^{\tau,n}\|_V$. ∎

LEMMA 4.7. *Let τ_0 be as required in Lemma 4.1, and assume that $0 < \tau < \tau_0$. If $u_0 \in V$ and $\partial_t^\ell f(0) \in V'$, $\ell = 1, 2, \dots, k-2$, then*

$$\|I_3(t_n) - I_3^{\tau,n}\|_V \le c\tau^k \sum_{\ell=1}^{k-2} t_n^{\ell-k} \|\partial_t^\ell f(0)\|_{V'}.$$

Proof. We split the error into two parts, i.e.,

$$I_3(t_n) - I_3^{\tau,n} = \sum_{\ell=1}^{k-2} e_{1,\ell}^n + \sum_{\ell=1}^{k-2} e_{2,\ell}^n,$$

with

$$\begin{aligned} e_{1,\ell}^n &= \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta^\tau} e^{zt_n} B_\ell(z) D^{-1} \partial_t^\ell f(0)\, dz, \\ e_{2,\ell}^n &= \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_\theta \backslash \Gamma_\theta^\tau} e^{zt_n} z^{-\ell-1} zK(z,1) D D^{-1} \partial_t^\ell f(0) dz, \end{aligned}$$

where

$$B_\ell(z) = z^{-\ell-1} zK(z,1) - \tau^{\ell+1} \eta_\ell(e^{-z\tau}) \tau^{-1} \delta(e^{-z\tau}) K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau})).$$

Lemma 4.5 implies $\|B_\ell(z)\|_{V\to V} \le c\tau^k |z|^{k-\ell-1}$, and therefore

$$\begin{aligned} \|e_{1,\ell}^n\|_V &\le c\tau^k \|D^{-1}\partial_t^\ell f(0)\|_V \int_{\Gamma_\theta^\tau} e^{\mathrm{Re}(z)t_n} |z|^{k-\ell-1} |dz| \\ &\le c\tau^k \|D^{-1}\partial_t^\ell f(0)\|_V \int_0^{\pi/(\tau \sin\theta)} e^{-rt_n|\cos\theta|} r^{k-\ell-1} dr \\ &\le c\tau^k t_n^{\ell-k} \|D^{-1}\partial_t^\ell f(0)\|_V. \end{aligned}$$

Lemma 4.3 implies $\|zK(z,1)D\|_{V\to V} \le c$, and therefore

$$\begin{aligned} \|e_{2,\ell}^n\|_V &\le c\|D^{-1}\partial_t^\ell f(0)\|_V \int_{\Gamma_\theta\backslash\Gamma_\theta^\tau} e^{\mathrm{Re}(z)t_n} |z|^{-\ell-1}\, |dz| \\ &\le c\|D^{-1}\partial_t^\ell f(0)\|_V \int_{\pi/(\tau\sin\theta)}^\infty e^{-rt_n|\cos\theta|} r^{-\ell-1}\, dr \\ &\le c\tau^k \|D^{-1}\partial_t^\ell f(0)\|_V \int_0^\infty e^{-rt_n|\cos\theta|} r^{k-\ell-1}\, dr \quad (1 \le \tau^k|z|^k \text{ is used here}) \\ &\le c\tau^k t_n^{\ell-k} \|D^{-1}\partial_t^\ell f(0)\|_V. \end{aligned}$$

This proves the desired estimate. ∎

LEMMA 4.8. *Let τ_0 be as required in Lemma 4.1, and assume that $0 < \tau < \tau_0$. If $u_0 = 0$ and $f \in W^{k+1,1}(\mathbb{R}_+; V')$ satisfies $\partial_t^\ell f(0) = 0$ for $\ell = 0, 1, \dots, k-2$, then*

$$\|I_4(t_n) - I_4^{\tau,n}\|_V \le c\tau^k \left(t_n^{-1} \|\partial_t^{k-1} f(0)\|_{V'} + \|\partial_t^k f(0)\|_{V'} + \int_0^{t_n} \|\partial_t^{k+1} f(t)\|_{V'}\, dt \right).$$

Proof. We recall the definition of $R_k(t_n)$ in (3.1) and use the splitting

$$R_k(t) = \frac{t^{k-1}}{(k-1)!} \partial_t^{k-1} f(0) + \frac{t^k}{k!} \partial_t^k f(0) + \frac{t^k}{k!} * \partial_t^{k+1} f(t) =: \sum_{i=1}^3 R_k^{(i)}(t).$$

The generating function of the sequence $(R_k(t_n))_{n=1}^\infty$ is

$$
\begin{aligned}
\widetilde{R}_k(\zeta) &= \sum_{n=1}^{\infty} R_k(t_n)\zeta^n \\
&= \sum_{n=1}^{\infty} \frac{t_n^{k-1}}{(k-1)!}\partial_t^{k-1} f(0)\zeta^n + \sum_{n=1}^{\infty} \frac{t^k}{k!}\partial_t^k f(0)\zeta^n + \sum_{n=1}^{\infty} R_k^{(3)}(t_n)\zeta^n \\
&= \eta_{k-1}(\zeta)\tau^{k-1}\partial_t^{k-1} f(0) + \eta_k(\zeta)\tau^k\partial_t^k f(0) + \sum_{n=1}^{\infty} R_k^{(3)}(t_n)\zeta^n \\
&=: \widetilde{R}_k^{(1)}(\zeta) + \widetilde{R}_k^{(2)}(\zeta) + \widetilde{R}_k^{(3)}(\zeta),
\end{aligned}
$$

where $\eta_{k-1}(\zeta)$ and $\eta_k(\zeta)$ are given by (3.6) with $b_{k-1,n}^{(k)} = b_{k,n}^{(k)} = 0$. Then we have

$$
I_3^{\tau,n} - I_3(t_n) = \sum_{i=1}^{3} e_i^n
$$

with

$$
\begin{aligned}
e_i^n &= \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_\theta^\tau} e^{zt_n}\tau^{-1}\delta(e^{-z\tau})K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))\eta_{k-2+i}(\zeta)\tau^{k-1+i}\partial_t^{k-2+i} f(0)dz \\
&\quad - \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_\theta} e^{zt_n} zK(z,1)\frac{t^{k-2+i}}{(k-2+i)!}\partial_t^{k-1} f(0)dz \qquad \text{for } i = 1, 2, \\
e_3^n &= \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_\theta^\tau} e^{zt_n}\tau^{-1}\delta(e^{-z\tau})K(\tau^{-1}\delta(e^{-z\tau}), e^{-z\tau}\gamma(e^{-z\tau}))\tau\widetilde{R}_k^{(3)}(e^{-z\tau})dz \\
&\quad - \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_\theta} e^{zt_n} zK(z,1)\widehat{R}_k^{(3)}(z)dz.
\end{aligned}
$$

Therefore, repeating the proof of Lemma 4.7 yields the following estimates:

$$
\begin{aligned}
\|e_1^n\|_V &\le c\tau^k t_n^{-1}\|\partial_t^{k-1} f(0)\|_{V'}, \\
\|e_2^n\|_V &\le c\tau^k \|\partial_t^k f(0)\|_{V'}.
\end{aligned}
$$

Note that e_3^n can be viewed as the error of numerical solution to an equation with sufficiently smooth right side f, with sufficiently many compatibility conditions $\partial_t^m f(0) = 0$ for $m = 0, 1, \ldots, k$. For such a smooth right side f, by using the resolvent estimates established in sections 4.1–4.2, the estimation of e_3^n can be done similarly (hence omitted) as that given in [25, Lemma 3.7], i.e.,

$$
\|e_3^n\|_V \le c\tau^k \int_0^{t_n} \|\partial_s^{k+1} f(s)\|_{V'}ds.
$$

The estimates above imply Lemma 4.8. ∎

Combining the estimates in Lemmas 4.6–4.8, we obtain the desired error bound in Theorem 1.1.

5. Numerical examples. In this section, we present two numerical examples to support the theoretical analysis.

Example 5.1 (diffusion equation with an imaginary reaction coefficient). We begin with the following one-dimensional diffusion equation which involves an imaginary reaction coefficient, i.e.,

$$
\begin{cases}
\partial_t u - u_{xx} - 20\mathrm{i}u = f & \text{in} \quad \Omega \times (0,T], \\
u = 0 & \text{on} \quad \partial\Omega \times (0,T],
\end{cases}
$$

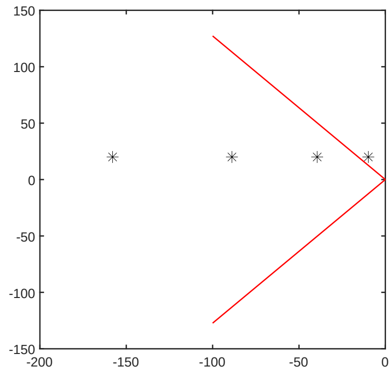

FIG. 2. *Distribution of first several (approximate) eigenvalues of operator A.*

TABLE 4
Example 5.1: errors and convergence rates of u_n for k-step BDF with $k = 1, 2, \ldots, 5$ at $t_n = 1$.

$k \backslash \tau$	1/60	1/120	1/240	1/480	Convergence rate
1	5.70E-03	9.51E-04	3.39E-04	1.55E-04	≈ 1.13 (1.00)
2	8.13E-04	2.62E-04	6.81E-05	1.70E-05	≈ 2.01 (2.00)
3	1.30E-03	1.06E-04	1.01E-05	1.24E-06	≈ 3.03 (3.00)
4	6.83E-04	7.33E-05	5.10E-06	3.17E-07	≈ 4.01 (4.00)
5	3.17E-02	2.90E-05	6.63E-07	2.02E-08	≈ 5.04 (5.00)

with $\Omega = (0, 1)$, where the initial data and source term are given by

$$\begin{aligned} u_0(x) &= \sin(\pi x), \\ f(x,t) &= \sin(x)(1+t)^{-2}\cos(50t). \end{aligned}$$

We divide the interval Ω into 1000 equally spaced subintervals with a mesh size $h = 1/1000$ and apply the P1 Galerkin FEM in space. With the fine meshes in space, the first several eigenvalues of A are approximated accurately. The distribution of the smallest in the sense of modulus four eigenvalues is shown in Figure 2, where the two straight lines are the sectorial angle condition for the five-step BDF. We observe that the first eigenvalue λ_1 lies in the right side of the red line, which indicates that $\alpha = \pi - \arg(\lambda_1) > \alpha_5$, violating the sectorial angle condition $|\alpha| \le \alpha_5$.

The errors of the numerical solutions at time level $t_n = 1$ are presented in Table 4 for $k = 1, 2, \ldots, 5$. Since the closed form of the solution is unavailable, we present the approximate errors

$$e_n^{(\tau)} = \|u_n^{(\tau)} - u_n^{(\tau/2)}\|_V,$$

instead of the exact errors, where $u_n^{(\tau)}$ denotes the numerical solution at time t_n calculated by using step size τ. The convergence rates of the numerical solutions are calculated by using ratios between the errors with consecutive step sizes, i.e.,

$$\text{convergence rate } = \log\left(\frac{\|u_n^{(\tau)} - u_n^{(\tau/2)}\|_V}{\|u_n^{(\tau/2)} - u_n^{(\tau/4)}\|_V}\right) \Big/ \log(2).$$

TABLE 5
Example 5.1: errors and convergence rates of u_n for k-step BDF with $k = 1, 2, \ldots, 5$ at different time levels.

t_n	k \ τ	1/60	1/120	1/240	1/480	Convergence rate
5	1	1.12E-04	5.25E-05	2.56E-05	1.27E-05	≈ 1.01 (1.00)
	2	6.80E-05	1.61E-05	5.07E-06	1.41E-06	≈ 1.84 (2.00)
	3	1.21E-04	1.75E-05	2.02E-06	2.31E-07	≈ 3.12 (3.00)
	4	1.19E-04	4.36E-06	3.27E-07	2.42E-08	≈ 3.75 (4.00)
	5	1.27E-04	3.93E-06	1.23E-07	3.54E-09	≈ 5.12 (5.00)
10	1	1.60E-05	9.88E-06	5.53E-06	2.93E-06	≈ 0.92 (1.00)
	2	3.23E-05	8.10E-06	1.85E-06	4.32E-07	≈ 2.10 (2.00)
	3	2.70E-05	2.83E-06	4.78E-07	6.73E-08	≈ 2.82 (3.00)
	4	3.26E-05	2.55E-06	1.40E-07	7.69E-09	≈ 4.19 (4.00)
	5	3.20E-05	6.35E-07	2.60E-08	9.62E-10	≈ 4.75 (5.00)
50	1	1.61E-06	8.02E-07	4.05E-07	2.04E-07	≈ 0.99 (1.00)
	2	7.50E-07	3.09E-07	9.24E-08	2.43E-08	≈ 1.93 (2.00)
	3	1.95E-06	1.99E-07	1.89E-08	2.02E-09	≈ 3.23 (3.00)
	4	1.09E-06	8.00E-08	6.66E-09	4.52E-10	≈ 3.88 (4.00)
	5	1.31E-06	5.23E-08	1.30E-09	3.40E-11	≈ 5.25 (5.00)

From Table 4 we see that the numerical schemes are unconditionally stable with kth-order convergence, which are consistent with the theoretical results proved in Theorem 1.1. In the case $k = 5$, the sectorial angle condition is violated, while the numerical solutions are still stable and convergent when the step size τ is smaller than $\tau_0 = 1/50$. This step size condition is due to the violation of the angle condition $\pi - \arg(\lambda_1) > \alpha_5$. The threshold τ_0 of the step size depends only on the operator A, independent of spatial mesh size h. Moreover, we present long-time errors in Table 5 and observe that the errors keep small even for large time, e.g., $t_n = 5$, 10, and 50. This is also consistent with our analysis on the long-time accuracy of the numerical method.

Example 5.2 (the Stokes–Darcy system). We solve the Stokes–Darcy system (2.2)–(2.3) in the domains $\Omega_D = (-1,1) \times (0,1)$ and $\Omega_S = (-1,1) \times (-1,0)$ (as shown in Figure 1), by using the proposed method (2.9)–(2.10). The domains Ω_D and Ω_S are discretized into regular right triangles with spatial step sizes $h_x = h_y = 1/64$. For simplicity, we choose the parameters $\kappa = \beta = g = \nu = 1$ in (2.2)–(2.3). The initial values and the source terms are given by

$$\begin{aligned}
\phi_0(x,y) &= 0,\\
\boldsymbol{u}_0(x,y) &= \Big(\cos(\pi x)\cos(\pi y + \pi/2)\,,\,0\Big),\\
f_D(x,y,t) &= \sin(x)\cos(t) + \cos(y)\sin(t),\\
\boldsymbol{f}_S(x,y,t) &= \Big(\sin(x)\cos(t) + \cos(y)\sin(t), \cos(x)\cos(t) + \sin(y)\sin(t)\Big).
\end{aligned}$$

The errors of numerical solutions at $t_n = 1$ are presented in Tables 6 and 7 for different step sizes. The numerical results in Tables 6 and 7 show that the proposed k-step method has kth-order convergence, which is consistent with the theoretical result proved in Theorem 1.1.

TABLE 6
Example 5.2: errors and convergence rates of ϕ_n, using k-step BDF, at $t_n = 1$.

$k \backslash \tau$	1/32	1/64	1/128	1/256	Convergence rate
1	3.31E-02	1.72E-02	8.79E-03	4.44E-03	≈ 0.99 (1.00)
2	1.65E-03	4.32E-04	1.10E-04	2.76E-05	≈ 1.99 (2.00)
3	1.52E-04	1.84E-05	2.33E-06	2.93E-07	≈ 3.00 (3.00)
4	7.92E-03	1.05E-06	4.51E-08	2.72E-09	≈ 4.05 (4.00)

TABLE 7
Example 5.2: errors and convergence rates of $\boldsymbol{u}_n$, using k-step BDF, at $t_n = 1$.

$k \backslash \tau$	1/32	1/64	1/128	1/256	Convergence rate
1	1.96E-02	9.89E-03	4.97E-03	2.49E-03	≈ 1.00 (1.00)
2	2.45E-03	5.93E-04	1.46E-04	3.61E-05	≈ 2.01 (2.00)
3	2.10E-04	2.68E-05	3.41E-06	4.29E-07	≈ 2.99 (3.00)
4	6.18E-03	1.24E-06	4.97E-08	4.03E-09	≈ 4.03 (4.00)

REFERENCES

[1] G. AKRIVIS, *Stability of implicit-explicit backward difference formulas for nonlinear parabolic equations*, SIAM J. Numer. Anal., 53 (2015), pp. 464–484.
[2] G. AKRIVIS, *Stability properties of implicit-explicit multistep methods for a class of nonlinear parabolic equations*, Math. Comp., 85 (2016), pp. 2217–2229.
[3] G. AKRIVIS, *Stability of implicit-explicit multistep methods for nonlinear parabolic equations*, IMA J. Numer. Anal., 38 (2018), pp. 1768–1796.
[4] G. AKRIVIS, M. CROUZEIX, AND C. MAKRIDAKIS, *Implicit-explicit multistep methods for quasilinear parabolic equations*, Numer. Math., 82 (1999), pp. 521–541.
[5] G. AKRIVIS AND E. KATSOPRINAKIS, *Backward difference formulae: new multipliers and stability properties for parabolic equations*, Math. Comp., 85 (2016), pp. 2195–2216.
[6] G. AKRIVIS AND C. LUBICH, *Fully implicit, linearly implicit and implicit-explicit backward difference formulae for quasi-linear parabolic equations*, Numer. Math., 131 (2015), pp. 713–735.
[7] W. ARENDT, C. J. K. BATTY, M. HIEBER, AND F. NEUBRANDER, *Vector-Valued Laplace Transforms and Cauchy Problems*, 2nd ed., Birkhäuser, Springer, Basel, 2011.
[8] W. AUZINGER, *Sectorial operators and normalized numerical range*, Appl. Numer. Math., 45 (2003), pp. 367–388.
[9] T. ARBOGAST AND M. GOMEZ, *A discretization and multigrid solver for a Stokes–Darcy system of three dimensional vuggy porous media*, Comput. Geosci., 13 (2009), pp. 331–348.
[10] T. ARBOGAST AND H. L. LEHR, *Homogenization of a Stokes–Darcy system modeling vuggy porous media*, Comput. Geosci., 10 (2006), pp. 291–302.
[11] M. CAI, M. MU, AND J. XU, *Numerical solution to a mixed Navier–Stokes/Darcy model by the two-grid approach*, SIAM J. Numer. Anal., 47 (2009), pp. 3325–3338.
[12] Y. CAO, M. GUNZBURGER, X. HE, AND X. WANG, *Parallel, non-iterative, multi-physics domain decomposition methods for time-dependent Stokes–Darcy systems*, Math. Comp., 83 (2014), pp. 1617–1644.
[13] Y. CAO, M. GUNZBURGER, X. HU, F. HUA, X. WANG, AND W. ZHAO, *Finite element approximation for Stokes–Darcy flow with Beavers-Joseph interface conditions*, SIAM. J. Numer. Anal., 47 (2010), pp. 4239–4256.
[14] J. CHEN, S. SUN, AND X. WANG, *A numerical method for a model of two-phase flow in a coupled free flow and porous media system*, J. Comput. Phys., 268 (2014), pp. 1–16.
[15] W. CHEN, M. GUNZBURGER, D. SUN, AND X. WANG, *Efficient and long-time accurate second-order methods for the Stokes–Darcy system*, SIAM J. Numer. Anal., 51 (2013), pp. 2563–2584.
[16] W. CHEN, M. GUNZBURGER, D. SUN, AND X. WANG, *An efficient and long-time accurate third-order algorithm for the Stokes–Darcy system*, Numer. Math., 134 (2016), pp. 857–879.
[17] M. CROUZEIX, *Une méthode multipas implicite–explicite pour l'approximation des équations d'évolution paraboliques*, Numer. Math., 35 (1980), pp. 257–276.

[18] V. J. Ervin, E. W. Jenkins, and S. Sun, *Coupled generalized nonlinear Stokes flow with flow through a porous medium*, SIAM J. Numer. Anal., 47 (2009), pp. 929–952.

[19] G. P. Galdi, *An Introduction to the Mathematical Theory of the Navier-Stokes Equations*, Springer Monogr. Math., 39, Springer-Verlag, New York, 2011.

[20] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations* II: *Stiff and Differential–Algebraic Problems*, 2nd ed., Springer Ser. Comput. Math. 14, Springer-Verlag, Berlin, 2002.

[21] N. Hanspal, A. Waghode, V. Nassehi, and R. Wakeman, *Numerical analysis of coupled Stokes/Darcy flow in industrial filtrations*, Transp. Porous Media, 64 (2006), pp. 73–101.

[22] R. Hoppe, P. Porta, and Y. Vassilevski, *Computational issues related to iterative coupling of subsurface and channel flows*, Calcolo, 44 (2007), pp. 1–20.

[23] J. Hou, X. He, C. Guo, M. Wei, and B. Bai, *A dual-porosity-Stokes model and finite element method for coupling dual-porosity flow and free flow*, SIAM J. Sci. Comput., 38 (2016), pp. B710–B739.

[24] B. Jin, B. Li, and Z. Zhou, *Correction of high-order BDF convolution quadrature for fractional evolution equations*, SIAM J. Sci. Comput., 39 (2017), pp. A3129–A3152.

[25] B. Jin, B. Li, and Z. Zhou, *An analysis of the Crank-Nicolson method for subdiffusion*, IMA J. Numer. Anal., 38 (2018), pp. 518–541.

[26] W. Layton, H. Tran, and C. Trenchea, *Analysis of long time stability and errors of two partitioned methods for uncoupling evolutionary groundwater-surface water flows*, SIAM J. Numer. Anal., 51 (2013), pp. 248–272.

[27] C. Lubich, *On the convergence of multistep methods for nonlinear stiff differential equations*, Numer. Math., 58 (1991), pp. 839–853; 61 (1992), pp. 277–279.

[28] C. Lubich, *Convolution quadrature revisited*, BIT, 44 (2004), pp. 503–514.

[29] A. Márquez, S. Meddahi, and F. J. Sayas, *A decoupled preconditioning technique for a mixed Stokes–Darcy model*, J. Sci. Comput., 57 (2013), pp. 174–192.

[30] V. Nassehi, *Modelling of combined Navier-Stokes and Darcy flows in crossflow membrane filtration*, Chem. Eng. Sci., 53 (1998), pp. 1253–1265.

[31] O. Nevanlinna and F. Odeh, *Multiplier techniques for linear multistep methods*, Numer. Funct. Anal. Optim., 3 (1981), pp. 377–423.

[32] C. Palencia, *Stability of rational multistep approximations of holomorphic semigroups*, Math. Comp., 64 (1995), pp. 591–599.

[33] I. Rybak, and J. Magiera, *A multiple-time-step technique for coupled free flow and porous medium systems*, J. Comput. Phys., 272 (2014), pp. 327–342.

[34] G. Savaré, *$A(\Theta)$–stable approximations of abstract Cauchy problems*, Numer. Math., 65 (1993), pp. 319–335.

[35] L. Shan and H. Zheng, *Partitioned time stepping method for fully evolutionary Stokes–Darcy flow with Beavers-Joseph interface conditions*, SIAM J. Numer. Anal., 51 (2013), pp. 813–839.

[36] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer-Verleg, Berlin, 2006.