Bergische Universität Wuppertal

Fakultät für Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 22/23

Zijun Zheng, Gang Pang, Matthias Ehrhardt, and Baiyili Liu

An efficient second-order method for the linearized Benjamin-Bona-Mahony equation with artificial boundary conditions

December 9, 2022

http://www.imacm.uni-wuppertal.de

An efficient second-order method for the linearized Benjamin-Bona-Mahony equation with artificial boundary conditions

Zijun Zheng[b], Gang Pang[a,*], Matthias Ehrhardt[c], Baiyili Liu[d]

[a] *School of Mathematical Science, Beihang University, Beijing 102206, China*
[b] *Chongqing University of technology, Chongqing 400054, China*
[c] *Chair of Applied and Computational Mathematics, School of Mathematics and Natural Sciences, University of Wuppertal, D-42119, Wuppertal, Germany*
[d] *School of Physics and Electronic Engineering, Centre for Computational Sciences, Sichuan Normal University, Chengdu 610066, China*

Abstract

In this paper, we present a fully discrete finite difference scheme with a fast convolution of artificial boundary conditions for solving the Cauchy problem of the one-dimensional linearized Benjamin-Bona-Mahony equation.

The Padé expansion of the square root function in the complex plane is used to implement the fast convolution thereby significantly reducing the computational costs incurred by the time convolution.

By introducing a constant damping term into the governing equations, the convergence analysis is performed when the damping term satisfies certain conditions. The theoretical analysis is supported by numerical examples that demonstrate the performance of the proposed fast numerical method.

Keywords: Benjamin-Bona-Mahony equation, artificial boundary condition, fast convolution quadrature, Padé approximation, convergence analysis.
2000 MSC: 65M06, 65M12, 65M85, 76M20

1. Introduction

The *Benjamin-Bona-Mahony* (BBM) equation is a classical nonlinear dispersive equation governing the unidirectional propagation of weakly nonlinear long waves in the presence of dispersion. The theoretical and numeri-

*Corresponding author, gangpang@buaa.edu.cn

cal aspects of the BBM equation have been studied in some important works [1, 2, 3]. In this paper, we mainly focus on the following linearized BBM equation, cf. [3], namely,

$$\partial_t u + c\partial_x u = \kappa \partial_{xxt} u, \quad x \in \mathbb{R}, t > 0 \tag{1}$$

The original system (1) is derived and defined on the whole space. However, the domain of numerical investigation is restricted to a bounded domain and one must prescribe appropriate boundary conditions. To obtain a reliable computational method, a standard approach is to truncate the infinite domain around the region of interest and introduce an *artificial boundary condition* (ABC) at the fictitious boundary. There are quite extensive studies [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22] on ABCs for some types of wave equations such as the classical Schrödinger equation. In general, the exact ABCs for the wave equations are nonlocal in time and contain some temporal convolutions in the formulations. Noble proposed the exact ABC for the linearized BBM equation in [3]. The ABC for the linearized BBM equation is also studied in [4], where the boundary condition is treated by numerical convolutions.

However, the nonlocal convolutions in the exact ABC cause a huge computational cost for the truncated problem in bounded domains (this occurs in the long term evolution: at the n-th time step $\mathcal{O}(n)$ operations are required to compute the convolution integral). Therefore, fast algorithms are introduced to reduce the computational load. Some types of fast approximation for kernel symbols are studied in [23, 24, 25, 26, 27, 28]. In most fast algorithms, summation of exponentials is used to approximate the convolution kernel, which can reduce the computational load from $\mathcal{O}(n)$ to $\mathcal{O}(1)$. To reduce the computational cost incurred by the exact convolution in time, we have proposed in this paper a convergent numerical method for solving the Cauchy problem of the one-dimensional linearized BBM equation (1), which integrates a fast evaluation of the exact ABC.

To this end, we first reformulate the BBM equation (1) into an equivalent form by introducing a constant damping term, and then construct a Crank-Nicolson scheme to discretize the equivalent problem in time. Specifically, for the reformulated Crank-Nicolson scheme, we derive a semi-discrete ABC for the temporally discretized problem by applying the $\mathcal{Z}$-transform, and then we propose a second-order finite difference scheme for further spatial discretization. A fast algorithm is introduced to approximate the discrete convolution kernel involved in the exact semi-discrete ABC by using the Padé rational expansion of the square root function [27] where the symbol

for the exact semi-discrete ABC can be approximated by a rational function that allows a fast convolution calculation, as the multi-pole method does. The damping term is chosen to preserve the convergence of the resulting fully discrete numerical method [31]. Finally, a stability analysis for the proposed numerical method is presented.

The remainder of this paper is organized as follows. In Section 2 the problem formulation is introduced, a damping factor in time is introduced to reformulate the problem, and the exact ABC for the semi-discrete reformulated scheme is obtained. In Section 3 a semi-discrete exact ABC for the time discretized Crank-Nicolson scheme for the reformulated problem is derived by applying the $\mathcal{Z}$-transform. Then the time and space discretizations for the reformulated system on a truncated computational domain are proposed. In Section 4 we introduce a Padé approximation for the exact discrete kernel and the resulting fast algorithm for practical computations. In Section 5 we give the properties of the approximation for the exact discrete kernel and determine the order of the Padé expansion. In Section 6 we give the convergence analysis of the fast numerical solutions. In Section 7, numerical examples are given to illustrate the effectiveness of the proposed numerical method. In Section 8 we draw a conclusion.

2. Exact ABCs for the one-dimensional linearized BBM equation

We consider the initial value problem (IVP) for the linearized Benjamin-Bona-Mahony (BBM) equation defined on the whole real line,

$$\begin{aligned} &\partial_t u(x,t) + c\partial_x u(x,t) = \kappa \partial_{xxt} u(x,t), && \forall\, x \in \mathbb{R},\ \forall\, t > 0, \\ &u(x,0) = u(x), && \forall\, x \in \mathbb{R}, \\ &\lim_{|x|\to+\infty} u(x,t) = 0, && \forall\, t > 0. \end{aligned} \tag{2}$$

Let us introduce the new function

$$v(x,t) = e^{-\sigma t} u(x,t),$$

where $\sigma > 0$ denotes an auxiliary parameter for controlling the stability of a fast algorithm to be introduced later in this paper. It is straightforward to verify that the function $v(x,t)$ solves the following IVP:

$$\begin{aligned} \partial_t v(x,t) + \sigma v(x,t) + c\partial_x v(x,t) &= \kappa \partial_{xx}\big(\partial_t v(x,t) + \sigma v(x,t)\big),\ \forall\, x \in \mathbb{R},\ \forall\, t > 0, \\ v(x,0) &= u(x), \quad \forall\, x \in \mathbb{R}, \\ \lim_{|x|\to+\infty} v(x,t) &= 0, \quad \forall\, t > 0. \end{aligned} \tag{3}$$

To obtain exact ABCs for the problem (3), we first consider the following exterior problem on the semi-infinite interval $[x_+, +\infty)$:

$$\partial_t v(x,t) + \sigma v(x,t) + c\partial_x v(x,t) = \kappa \partial_{xx}\Big(\partial_t v(x,t) + \sigma v(x,t)\Big), \quad \forall\, x \in [x_+, +\infty),\ \forall\, t > 0, \tag{4a}$$

$$v(x,0) = 0, \quad \forall\, x \in [x_+, +\infty), \tag{4b}$$

$$\lim_{x\to+\infty} v_1(x,t) = 0, \quad \forall\, t > 0. \tag{4c}$$

The Laplace transform of (4) in time yields

$$(s+\sigma)\widehat{v}(x,s) + c\partial_x \widehat{v}(x,s) = \kappa(s+\sigma)\partial_{xx}\widehat{v}(x,s), \quad \forall\, x \in [x_+, \infty),\ \forall\, s \in \mathbb{C}_+, \tag{5}$$

$$\lim_{x\to\infty} \widehat{v}(x,s) = 0, \quad \forall\, s \in \mathbb{C}_+, \tag{6}$$

where $\mathbb{C}_+$ stands for the right half part of the complex plane. Thus, the general solution of the equation (5) reads

$$\widehat{v}(x,s) = c_1(s)\exp\Big(x\xi_-(s)\Big) + c_2(s)\exp\Big(x\xi_+(s)\Big),$$

with

$$\xi_\pm(s) = \frac{c}{2\kappa(s+\sigma)}\Big(1 \pm \sqrt{1 + \frac{4\kappa(s+\sigma)^2}{c^2}}\Big)$$

where $\sqrt{\cdot}$ denotes the branch of the square root with non-negative real part. Clearly, the infinity boundary condition (6) implies $c_2(s) = 0$. Consequently, by differentiating the last equation we obtain

$$\begin{aligned}\partial_x \widehat{v}(x,s) &= \xi_-(s)\widehat{v}(x,s)\\ &= \frac{c}{2\kappa(s+\sigma)}\Big(1 - \sqrt{1 + \frac{4\kappa(s+\sigma)^2}{c^2}}\Big)\widehat{v}(x,s),\ \forall\, x \in [x_+, +\infty),\ \forall\, s \in \mathbb{C}_+,\end{aligned}$$

whose inverse Laplace transform yields an ABC at x_+:

$$\partial_x v(x_+,t) = \frac{c}{2\kappa(\partial_t+\sigma)}\Big(1 - \sqrt{1 + \frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)v(x_+,t), \quad \forall\, t > 0. \tag{7}$$

In the above equation (7), The factor

$$\frac{c}{2\kappa(\partial_t+\sigma)}\sqrt{1 + \frac{4\kappa(\partial_t+\sigma)^2}{c^2}}$$

stands for the multiplier operator (in time) associated with the symbol

$$\frac{c}{2\kappa(s+\sigma)}\sqrt{1+\frac{4\kappa(s+\sigma)^2}{c^2}},$$

namely,

$$\frac{c}{2\kappa(\partial_t+\sigma)}\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\,v_2(x_+,t)$$
$$:=\mathcal{L}^{-1}[\frac{c}{2\kappa(s+\sigma)}\sqrt{1+\frac{4\kappa(s+\sigma)^2}{c^2}}\,\widehat{v}_2(x_+,s)](t),\quad \forall\, t>0,$$

with $\mathcal{L}^{-1}$ denoting the inverse Laplace transform with respect to the s-variable. A similar boundary condition can be derived at x_-:

$$\partial_x v(x_-,t)=\frac{c}{2\kappa(\partial_t+\sigma)}\Big(1+\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)v(x_-,t),\quad \forall\, t>0. \tag{8}$$

In view of (7) and (8), the solution of (3) is the same as the solution of the following problem posed in a bounded domain:

$$\partial_t v(x,t)+\sigma v(x,t)+c\partial_x v(x,t)=\kappa\partial_{xx}\Big(\partial_t v(x,t)+\sigma v(x,t)\Big),$$
$$\forall x\in(x_-,x_+),\ \forall\, t>0, \tag{9}$$

$$\partial_\nu v(x_\pm,t)=\pm\frac{c}{2\kappa(\partial_t+\sigma)}\Big(1\mp\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)v(x_\pm,t),\quad \forall\, t>0. \tag{10}$$

$$v(x,0)=u(x),\quad \forall\, x\in[x_-,x_+], \tag{11}$$

where ∂_ν denotes the outward normal derivative at the boundary points $x_\pm$.

3. Discretization of the one-dimensional linearized BBM equation with exact semi-discrete ABC

In this section, we discretize the one-dimensional linearized BBM equation in time using the Crank-Nicolson scheme and derive an exact semi-discrete ABC. Then we propose a second order finite difference scheme for further spatial discretization. To do this, we first introduce the necessary notations related to the $\mathcal{Z}$-transform.

3.1. $\mathcal{Z}$-transform of a sequence of functions

Given a Hilbert space $\mathcal{H}$ with inner product $(\cdot,\cdot)_{\mathcal{H}}$ and induced norm $\|\cdot\|_{\mathcal{H}}$, we introduce the semi-infinite sequence spaces:

$$\ell^2(\mathcal{H}) = \Big\{u = \{u^n\}_{n=0}^{\infty} : \|u\|_{\ell^2(\mathcal{H})} \equiv \Big(\sum_{n=0}^{\infty} |u^n|^2\Big)^{\frac{1}{2}} < \infty\Big\},$$
$$\ell_0^2(\mathcal{H}) = \Big\{u = \{u^n\}_{n=0}^{\infty} \in \ell^2(\mathcal{H}) : u^0 = 0\Big\}$$

equipped with the inner product

$$(u,v)_{\ell^2(\mathcal{H})} \equiv \sum_{n=0}^{\infty} \overline{u^n} v^n, \quad \forall u, v \in \ell^2(\mathcal{H}).$$

Next, we define the $\mathcal{Z}$-transform as

$$\widetilde{u}(z) = \sum_{n=0}^{\infty} u^n z^n \quad \text{for} \quad u = \{u^n\}_{n=0}^{\infty} \in \ell^2(\mathcal{H}). \tag{12}$$

It is well-known that the following Parseval's identity holds:

$$(u,v)_{\ell^2(\mathcal{H})} = \int_{\partial\mathbb{D}} \overline{\widetilde{u}(z)}\widetilde{v}(z)\ \mu(dz), \quad \forall u, v \in \ell^2(\mathcal{H}). \tag{13}$$

For a sequence $u = \{u^n\}_{n=0}^{\infty} \in \ell^2(\mathcal{H})$, the *shift operator* S is defined by $Su = \{u^{n+1}\}_{n=0}^{\infty}$. The *average operator* E and the *forward difference quotient operator* D_τ are defined by

$$E = \frac{S+I}{2} \quad \text{and} \quad D_\tau = \frac{S-I}{\tau},$$

respectively. It is convention to define that

$$Su^n = (Su)^n, \quad Eu^n = (Eu)^n, \quad D_\tau u^n = (D_\tau u)^n.$$

3.2. Exact boundary conditions for the semi-discretized linearized BBM equation

Let $\tau > 0$ denote the time step where $N\tau = T$ with T being the total computation time. Let us set $t_n = n\tau$. We discretize (3) as follows

$$(D_\tau + \sigma E)v^n(x) + c\partial_x E v^n(x) = \kappa\partial_{xx}(D_\tau + \sigma E)v^n(x), \quad \forall x \in \mathbb{R},\ \forall\, n \geq 0, \tag{14}$$
$$v^0(x) = u(x), \quad \forall\, x \in \mathbb{R}, \tag{15}$$
$$\lim_{|x|\to+\infty} v^n(x) = 0, \quad \lim_{|x|\to+\infty} v_2^n(x) = 0, \quad \forall\, n \geq 1, \tag{16}$$

where $v(x)^n \approx v(x, t_n)$.

Suppose the initial data $u_1(x)$ and $u_2(x)$ are compact on the finite interval $[x_-, x_+]$. On the interval $[x_+, +\infty)$ the semi-discrete problem (14) reduces to

$$(D_\tau + \sigma E)v^n(x) + c\partial_x E v^n(x) = \kappa \partial_{xx}(D_\tau + \sigma E)v^n(x), \quad \forall x \in [x_+, +\infty),\ \forall n \geq 0, \tag{17}$$

$$v^0(x) = 0, \quad \forall x \in [x_+, +\infty), \tag{18}$$

$$\lim_{x \to +\infty} v^n(x) = 0, \quad \forall n \geq 1. \tag{19}$$

Let $\widetilde{u}(x, z)$ denote the $\mathcal{Z}$-transform of the sequence $\{u^n(x)\}_{n=0}^{\infty}$. Applying the $\mathcal{Z}$-transform to (17), we obtain

$$\frac{1}{\kappa}\widetilde{v}(x, z) + \frac{c\tau(1+z)}{\kappa\Big(2 - 2z + \sigma\tau(1+z)\Big)}\widetilde{v}(x, z) - \partial_{xx}\widetilde{v}(x, z) = 0,\ \forall x \in [x_+, +\infty),$$

$$\lim_{x \to +\infty} \widetilde{v}(x, z) = 0,$$

whose solution can be generally expressed as

$$\widetilde{v}(x, z)s = c_1^+ \exp\big(x\eta_-(z)\big) + c_2^+ \exp\big(x\eta_+(z)\big),$$

with

$$\eta_\pm(z) = \frac{c\tau(1+z)}{2\kappa(2 - 2z + \sigma\tau(1+z))}\Big(1 \pm \sqrt{1 + \frac{4\kappa(2 - 2z + \sigma\tau(z+1))^2}{c^2\tau^2(1+z)^2}}\Big).$$

The decay condition $\lim_{x \to +\infty} \widetilde{v}(x, z) = 0$ implies $c_1^+ = 0$. This leads to the following identity (by differentiating $\widetilde{v}(x, z)$ with respect to x):

$$\partial_x \widetilde{v}(x_+, z) = \eta_-(z)\widetilde{v}(x_+, z) = \widetilde{\mathcal{B}}_+(z)\widetilde{v}(x_+, z), \quad \forall z \in \mathbb{D}. \tag{20}$$

We note that the function $\widetilde{\mathcal{B}}_+(z)$ is analytic in the unit disk $\mathbb{D}$, i.e. it has a power series expansion

$$\widetilde{\mathcal{B}}_+(z) = \sum_{j=0}^{\infty} (\mathcal{B}_+)^j z^j, \quad \forall\, z \in \mathbb{D}. \tag{21}$$

Substituting (21) and $\widetilde{v}(x, z) = \sum_{n=0}^{\infty} v^n(x) z^n$ into (20) yields an exact ABC for (14) at the right artificial boundary point $x = x_+$:

$$(\mathcal{B}_+ * v)^n(x_+) = \partial_\nu v^n(x_+), \quad \forall\, n \geq 0,$$

where $\mathcal{B}_+*$ denotes the convolution quadrature operator corresponding to the symbol $\widetilde{\mathcal{B}}_+(z)$

$$(\mathcal{B}_+ * v_2)^n = \sum_{j=0}^{n} (\mathcal{B}_+)^j v_2^{n-j}. \tag{22}$$

For simplicity, for a function $v(x,t)$ we use the notation

$$\mathcal{B}_+ * v(x,t_n) = \sum_{j=0}^{n} \mathcal{B}_+^j \, v(x,t_{n-j}).$$

Analogously, by analyzing the problem (14) on $(-\infty, x_-]$, we derive an exact ABC at the left artificial boundary point $x = x_-$:

$$\partial_x \widetilde{v}(x_-,z) = \eta_+(z)\widetilde{v}(x_-,z), \quad \forall z \in \mathbb{D}.$$

thus

$$\partial_x \widetilde{v}(x_-,z) = \frac{c\tau(1+z)}{2\kappa(2-2z+\sigma\tau(1+z))} \Big(1+\sqrt{1+\frac{4\kappa(2-2z+\sigma\tau(z+1))^2}{c^2\tau^2(1+z)^2}}\Big)\widetilde{v}(x_-,z), \quad \forall z \in \mathbb{D}, \tag{23}$$

which leads to

$$\partial_\nu v^n(x_-) = (\mathcal{B}_- * v)^n(x_-), \quad \forall\, n \geq 1,$$

with

$$\widetilde{\mathcal{B}}_-(z) = -\frac{c\tau(1+z)}{2\kappa(2-2z+\sigma\tau(1+z))}\Big(1+\sqrt{1+\frac{4\kappa(2-2z+\sigma\tau(z+1))^2}{c^2\tau^2(1+z)^2}}\Big).$$

Hence, the semi-discrete problem (14) defined on the the whole real line, can be reduced to the following semi-discrete problem on a bounded interval:

$$(D_\tau + \sigma E)v^n(x) + c\partial_x E v^n(x) = \kappa\partial_{xx}(D_\tau+\sigma E)v^n(x), \quad \forall x \in (x_-,x_+),\ \forall\, n \geq 0, \tag{24}$$

$$\partial_\nu v^n(x_\pm) = (\mathcal{B}_\pm * v)^n(x_\pm), \quad \forall\, n \geq 0, \tag{25}$$

$$v(x) = u(x), \quad \forall\, x \in [x_-,x_+]. \tag{26}$$

Comparing (24) with (9), we see that the equation is discretized by a Crank-Nicolson scheme subject to a perturbation of order $\mathcal{O}(\tau^2)$, with a convolution quadrature approximation of the fractional order time derivative

at the boundary points $x_\pm$. Since the time discretization (14) is of second order in the whole space, it follows that the induced convolution quadrature at the boundary points $x_\pm$ in (24) is also of second order:

$$|(\mathcal{B}_\pm * v)^n(x_\pm) \mp \frac{c}{2\kappa(\partial_t+\sigma)}\Big(1 \mp \sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)v(x_\pm,t_n)| \le C\tau^2.$$

A proof of the above estimate is given in Section 6.1.

3.3. Spatial discretization

Let M be a positive integer, $h = (x_+ - x_-)/M$ be the mesh size, and $\tau > 0$ be the time step. We define the mesh points

$$\begin{aligned} &x_k = x_- + (k-1/2)h, && k = 0,1,\dots,M+1,\\ &t_n = n\tau, && n = 0,1,\dots,N, \end{aligned}$$

with x_0 and x_{M+1} being two ghost points.

In the time-stepping scheme (24) v_k^n denotes the numerical solution of $v^n(x_k)$ with $0 \le k \le M+1$. Let $v^n = (v_0^n,\dots,v_{M+1}^n)$. Given a vector $\omega = (\omega_0,\dots,\omega_{M+1}) \in \mathbb{R}^{M+2}$, we introduce the discrete gradient $\nabla_h w$ by

$$\nabla_h \omega = \Big(\frac{\omega_1-\omega_0}{h}, \frac{\omega_2-\omega_1}{h}, \dots, \frac{\omega_{M+1}-\omega_M}{h}\Big)$$

The linear operator $\mathcal{P}$ maps the $(M+2)$-dimensional vector $\omega = (\omega_0,\dots,\omega_{M+1})$ to the M-dimensional vector $(\omega_1,\dots,\omega_M)$. Next, we define the Neumann and Dirichlet data associated with the $(M+2)$-dimensional vector ω as

$$\partial_\nu^- \omega = \frac{\omega_0-\omega_1}{h},\ \partial_\nu^+\omega = \frac{\omega_{M+1}-\omega_M}{h},\ \gamma^- w = \frac{\omega_0+\omega_1}{2},\ \gamma^+ w = \frac{\omega_{M+1}+\omega_M}{2}.$$

We introduce an inner product for M-dimensional vectors

$$\phi_1 = \big((\phi_1)_1,\dots,(\phi_1)_M\big) \quad \text{and} \quad \phi_2 = \big((\phi_2)_1,\dots,(\phi_2)_M\big)$$

as

$$(\phi_1,\phi_2)_h = h\sum_{k=1}^{M} \overline{(\phi_1)_k}(\phi_2)_k,$$

an inner product for $(M+2)$-dimensional vectors

$$\omega_1 = \big((\omega_1)_0,\dots,(\omega_1)_{M+1}\big) \quad \text{and} \quad \omega_2 = \big((\omega_2)_0,\dots,(\omega_2)_{M+1}\big)$$

as

$$\langle \omega_1, \omega_2 \rangle_h = \frac{h}{2}\overline{(\omega_1)_0}(\omega_2)_0 + h\sum_{k=1}^{M}\overline{(\omega_1)_k}(\omega_2)_k + \frac{h}{2}\overline{(\omega_1)_{M+1}}(\omega_2)_{M+1}.$$

Correspondingly, the induced norms will be denoted by

$$\|\phi\|_h = \sqrt{(\phi,\phi)_h}, \qquad |\omega|_h = \sqrt{\langle \omega, \omega \rangle_h}.$$

We introduce a second-order spatial discretization Δ_h, which maps the $(M+2)$-dimensional ω to the M-dimensional vector space:

$$\Delta_h w = \Big(\frac{w_0 - 2w_1 + w_2}{h^2}, \ldots, \frac{w_{M-1} - 2w_M + w_{M+1}}{h^2}\Big).$$

Thus we have

$$(\mathcal{P}\omega_2, \Delta_h\omega_1)_h = -\langle \nabla_h\omega_2, \nabla_h\omega_1 \rangle_h + \overline{\gamma^+\omega_2}\cdot\partial^+\omega_1 + \overline{\gamma^-\omega_2}\cdot\partial^-\omega_1. \tag{27}$$

We also introduce a second-order centered spatial discretization ∇_h^m, which maps the $(M+2)$-dimensional ω to the M-dimensional vector space:

$$\nabla_h^m w = \Big(\frac{w_2 - w_0}{2h}, \ldots, \frac{w_{M+1} - w_{M-1}}{2h}\Big).$$

Thus we have

$$\mathrm{Re}(\mathcal{P}\omega, \nabla_h^m\omega)_h = \mathrm{Re}(\overline{w}_M w_{M+1}) - \mathrm{Re}(\overline{w}_1 w_0). \tag{28}$$

The vector $\nabla_h^m\mathcal{H}v^n$ can be defined by

$$\nabla_h^m\mathcal{H}v^n = \Big(\frac{\mathcal{H}v_2^n - \mathcal{H}v_0^n}{2h}, \ldots, \frac{\mathcal{H}v_{M+1}^n - \mathcal{H}v_{M-1}^n}{2h}\Big),$$

where $\mathcal{H}$ is any operator that only works on the time direction. We can also define the vector $\Delta_h\mathcal{H}v_2^n$ by

$$\Delta_h\mathcal{H}v^n = \Big(\frac{\mathcal{H}v_0^n - 2\mathcal{H}v_1^n + \mathcal{H}v_2^n}{h^2}, \ldots, \frac{\mathcal{H}v_{M-1}^n - 2\mathcal{H}v_M^n + \mathcal{H}v_{M+1}^n}{h^2}\Big).$$

It is easy to see

$$\nabla_h\mathcal{H}v^n = \mathcal{H}\nabla_h v^n, \quad \Delta_h\mathcal{H}v^n = \mathcal{H}\Delta_h v^n.$$

Replacing in the time-stepping scheme (24) the function $v^n(x)$ by the vector $v^n = (v_0^n, \ldots, v_{M+1}^n)$ and replacing the continuous operator ∂_{xx} by its discrete analog Δ_h, we obtain the fully discrete finite difference scheme

$$\begin{aligned}
&(D_\tau + \sigma E)\mathcal{P}v^n + \nabla_h^m E v^n = \kappa\Delta_h(D_\tau + \sigma E)v^n, && \forall\, n \geq 0,\\
&(\mathcal{B}_\pm * \gamma^\pm v)^n - \partial_\nu^\pm v^n = 0, && \forall\, n \geq 0,\\
&v^0 = (u(x_0), \ldots, u(x_{M+1})).
\end{aligned} \tag{29}$$

4. Fast approximation of the discrete convolution $(\boldsymbol{\mathcal{B}}_{\pm} * \boldsymbol{\gamma}^{\pm}\boldsymbol{v_2})^n$

In this section, a fast algorithm for approximating the boundary convolution $(\mathcal{B}_{\pm} * \gamma^{\pm} v_2)^n$ in (29) is presented. The stability of the proposed fast algorithm is shown in the next section.

4.1. Rational approximation of the convolution quadrature

In [27] it was shown that for non-negative integer $m > 0$ the Padé approximation for the function $\sqrt{1+s}$ can be expressed as

$$\sqrt{1+s} \approx 1 + \sum_{j=1}^{m} \frac{\alpha_j s}{1+\beta_j s},$$

where

$$\alpha_j = \frac{2}{2m+1} \sin^2 \frac{j\pi}{2m+1}, \quad \beta_j = \cos^2 \frac{j\pi}{2m+1}, \quad j = 1, \ldots, m.$$

Based on this Padé approximation, a rational approximation for the square root function $\sqrt{s}$ on the closed right half complex plane can be written as:

$$\sqrt{s} = \sqrt{1+s-1} \approx 1 + \sum_{j=1}^{m} \frac{\alpha_j (s-1)}{1+\beta_j (s-1)} \equiv R_m(s), \quad \mathrm{Re}(s) \geq 0.$$

Thus,

$$\begin{aligned} R_m(s) &= \lambda - \sum_{j=1}^{m} \frac{1}{g_j s + h_j}, \quad \lambda = 1 + \sum_{j=1}^{m} \alpha_j \beta_j^{-1}, \\ h_j &= \alpha_j^{-1} \beta_j (1-\beta_j), \quad g_j = \alpha_j^{-1} \beta_j^2, \quad j = 1, \ldots, m. \end{aligned} \tag{30}$$

We define $\kappa_1 = 4\kappa/c^2$ for convenience. For all $\tau > 0$ and $\sigma > 0$, by defining $s(z)$ such that,

$$s(z) = \frac{c\tau(1+z)}{2\kappa\big(2 - 2z + \sigma\tau(1+z)\big)}, \tag{31}$$

and

$$\begin{aligned} \mathcal{S}(z) = \Big(\frac{c\tau(1+z)}{2\kappa\big(2-2z+\sigma\tau(1+z)\big)}\Big)^2 \Big(1 + \kappa_1 \frac{\big(2-2z+\sigma\tau(z+1)\big)^2}{\tau^2(1+z)^2}\Big) \\ = s(z)^2 + \frac{1}{\kappa}, \end{aligned} \tag{32}$$

let us introduce the rational approximation $\widetilde{\mathcal{B}}_+^{(m)}(z)$ of the symbol $\widetilde{\mathcal{B}}_+(z)$:

$$\widetilde{\mathcal{B}}_+^{(m)}(z) := s(z) - R_m\Big(\mathcal{S}(z)\Big), \quad \forall\, m \geq 0. \tag{33}$$

We denote by $\mathcal{B}_+^{(m)}*$ the convolution operator defined analogously to (22) by replacing the convolution coefficients $(B_\pm)^j$ in (22) by the series expansion coefficients of the function $\widetilde{\mathcal{B}}^{(m)}(z)$. After replacing the convolution operator $\mathcal{B}_\pm*$ in (29) by its rational approximation $\mathcal{B}_\pm^{(m)}*$, we obtain the following fully discrete scheme:

$$(D_\tau + \sigma E)\mathcal{P}v^n + c\nabla_h^m E v^n = \kappa\Delta_h(D_\tau + \sigma E)v^n, \qquad \forall\, n \geq 0, \tag{34}$$

$$(\mathcal{B}_\pm^{(m)} * \gamma^\pm v)^n - \partial_\nu^\pm v^n = 0, \qquad \forall\, n \geq 0, \tag{35}$$

$$v^0 = (u(x_0), \ldots, u(x_{M+1})). \tag{36}$$

Eq. (35) can be solved by the fast algorithm described in the next Section 4.2.

4.2. Fast algorithm

By applying (30) to (33), we derive

$$\begin{aligned}
&\widetilde{\mathcal{B}}_\pm^{(m)}(z) \mp s(z) = -R_m\Big(\mathcal{S}(z)\Big) = -\frac{1}{\sqrt{\kappa}} R_m\Big(\kappa\Big(s(z)\Big)^2 + 1\Big) \\
&= -\frac{1}{\sqrt{\kappa}}\Big[\lambda - \sum_{j=1}^m \frac{1}{g_j \frac{c^2\tau^2(1+z)^2}{4\kappa\Big(2-2z+\sigma\tau(1+z)\Big)^2} + h_j}\Big] \\
&= -\frac{1}{\sqrt{\kappa}}\Big[\lambda - \sum_{j=1}^m \frac{4\kappa\Big(2+\sigma\tau+(\sigma\tau-2)z\Big)^2}{4\kappa h_j\Big(2+\sigma\tau+(\sigma\tau-2)z\Big)^2 + g_j c^2\tau^2(1+z)^2}\Big] \\
&= -\frac{1}{\sqrt{\kappa}}\Big[\lambda - \sum_{j=1}^m \Big(\lambda_j + \frac{e_j z + f_j}{(a_j z + b_j)^2 - (c_j z + d_j)^2}\Big)\Big] \\
&= -\frac{1}{\sqrt{\kappa}}\Big[\lambda - \sum_{j=1}^m \lambda_j - \sum_{j=1}^m \Big(\frac{A_j}{(a_j + c_j)z + b_j + d_j} + \frac{B_j}{(a_j - c_j)z + b_j - d_j}\Big)\Big],
\end{aligned} \tag{37}$$

where we have set

$$\begin{aligned}
\lambda_j &= \frac{4\kappa(\tau\sigma-2)^2}{4\kappa h_j(\tau\sigma-2)^2+g_jc^2\tau^2},\\
e_j &= 8\kappa(1-h_j\lambda_j)(\sigma^2\tau^2-4)-2c^2\tau^2\lambda_jg_j,\\
f_j &= 4\kappa(1-h_j\lambda_j)(2+\sigma\tau)^2-c^2\tau^2\lambda_jg_j,
\end{aligned}$$

$$\begin{aligned}
a_j &= 2\sqrt{\kappa h_j}(\sigma\tau-2), \qquad b_j = 2\sqrt{\kappa h_j}(\sigma\tau+2),\\
c_j &= i\sqrt{g_j}c\tau, \qquad d_j = i\sqrt{g_j}c\tau,\\
A_j &= \frac{-e_jc_j+b_jf_j-e_ja_j+f_jd_j}{a_j^2+b_j^2-c_j^2-d_j^2},\\
B_j &= \frac{e_jc_j+b_jf_j-e_ja_j-f_jd_j}{a_j^2+b_j^2-c_j^2-d_j^2}.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\widetilde{\mathcal{B}}_{\pm}^{(m)}(z) &= -\frac{1}{\sqrt{\kappa}}\Big[\lambda-\sum_{j=1}^m\lambda_j-\sum_{j=1}^m\Big(\frac{A_j}{(a_j+c_j)z+b_j+d_j}\\
&\quad+\frac{B_j}{(a_j-c_j)z+b_j-d_j}\Big)\Big]\pm s(z)\\
&= -\frac{1}{\sqrt{\kappa}}(\lambda-\sum_{j=1}^m\lambda_j)+\frac{1}{\sqrt{\kappa}}\sum_{j=1}^m\frac{A_j}{b_j+d_j}\sum_{n=0}^\infty\Big(-\frac{a_j+c_j}{b_j+d_j}\Big)^nz^n\\
&\quad+\frac{1}{\sqrt{\kappa}}\sum_{j=1}^m\frac{B_j}{b_j-d_j}\sum_{n=0}^\infty\Big(-\frac{a_j-c_j}{b_j-d_j}\Big)^nz^n\\
&\quad\pm\Big(\frac{c\tau}{2\kappa(\sigma\tau-2)}+\frac{2c\tau}{\kappa(2-\sigma\tau)}\frac{1}{(2+\sigma\tau)+(\sigma\tau-2)z}\Big)\\
&= -\frac{1}{\sqrt{\kappa}}(\lambda-\sum_{j=1}^m\lambda_j)+\frac{1}{\sqrt{\kappa}}\sum_{j=1}^m\frac{A_j}{b_j+d_j}\sum_{n=0}^\infty\Big(-\frac{a_j+c_j}{b_j+d_j}\Big)^nz^n\\
&\quad+\frac{1}{\sqrt{\kappa}}\sum_{j=1}^m\frac{B_j}{b_j-d_j}\sum_{n=0}^\infty\Big(-\frac{a_j-c_j}{b_j-d_j}\Big)^nz^n\\
&\quad\pm\frac{c\tau}{2\kappa(\sigma\tau-2)}\pm\frac{2c\tau}{\kappa(4-\sigma^2\tau^2)}\sum_{j=1}^m\Big(\frac{2-\sigma\tau}{2+\sigma\tau}\Big)^nz^n.
\end{aligned}$$

Thus, $\widetilde{\mathcal{B}}_{\pm}^{(m)}(z)$ can be rewritten as

$$\widetilde{\mathcal{B}}_{\pm}^{(m)}(z) = \sum_{k=1}^{3m+1} \sum_{n=0}^{\infty} C_k^{\pm} (\gamma_k^{\pm})^n z^n,$$

which implies that

$$(\mathcal{B}_{\pm}^{(m)})^j = \sum_{k=1}^{2m+1} C_k^{\pm} (\gamma_k^{\pm})^j.$$

Therefore $(\mathcal{B}_{\pm}^{(m)} * \gamma^{\pm} v_2)^n = \sum_{j=0}^{n} (\mathcal{B}_{\pm}^{(m)})^j (\gamma^{\pm} v_2)^{n-j}$ can be implemented by fast convolution in (35).

5. Properties of the rational approximation $\widetilde{\mathcal{B}}_{\pm}^{(m)}(z)$

In this section, we prove the following Proposition 1 for the properties of $(\mathcal{B}_{\pm}^{(m)})^j$ used later for the error estimation:

Proposition 1 *Under the condition $\sigma \geq 1/\sqrt{\kappa_1}$ and for a time step τ small enough, if m is large enough such that*

$$2m+1 \geq \frac{\ln \epsilon}{\ln(1-\delta)}, \quad \text{for some } \epsilon \in \left(0, \frac{\mu\sigma\kappa}{4c\sqrt{\kappa_1\sigma^2+1}}\tau^3\right], \tag{38}$$

the following inequalities hold:

$$\max_{z\in\partial\mathbb{D}} |\widetilde{\mathcal{B}}_{\pm}^{(m)}(z) - \widetilde{\mathcal{B}}_{\pm}(z)| \leq \mu \frac{\tau^3}{2}, \tag{39}$$

$$\mathrm{Re} \sum_{k=0}^{n} \overline{(\tau D_\tau + \sigma E) u^k} (D_\tau + \sigma E)(\mathcal{B}_{\pm}^{(m)} * u)^k \leq 0, \tag{40}$$

for all complex sequences $u = \{u^n\}_{n=0}^{\infty}$ with $u^0 = 0$. Here

$$\mu(\kappa,\sigma) = \min\Big[\frac{2\sigma}{c(1+\sqrt{1+\kappa_1\sigma^2})} \cos\Big(\theta_m(\kappa,\sigma)\Big), \\ \frac{2}{c}\Big(\frac{4\sigma^2}{\kappa_1(1+4\kappa_1\sigma^2)}\Big)^{\frac{1}{4}} \cos\Big(\theta_m(\kappa,\sigma)\Big)\Big],$$

and

$$\delta(\kappa,\sigma) = \min\Big(\frac{\sqrt{2}\mathcal{S}_1}{(\mathcal{S}_1)^2+\sqrt{2}\mathcal{S}_1+1}, \frac{\sqrt{2}\mathcal{S}_2}{(\mathcal{S}_2)^2+\sqrt{2}\mathcal{S}_2+1}\Big),$$

where

$$\mathcal{S}_1 = \frac{c\sqrt{\kappa_1\sigma^2+1}}{2\kappa\sigma}, \quad \mathcal{S}_2 = \frac{c}{2\kappa}\Big(\frac{4\kappa_1^3\sigma^2}{1+4\kappa_1\sigma^2}\Big)^{\frac{1}{4}},$$

and $\theta_m(\kappa,\sigma)$ *satisfies* $0 \le \theta_m(\kappa,\sigma) < \pi/2$.

Now, let

$$r(s) := \frac{\sqrt{s}-1}{\sqrt{s}+1}, \tag{41}$$

one can prove that the symbol $\widetilde{\mathcal{B}}_{\pm}(z)$ satisfies the following inequalities:

Lemma 1 *Under the setting of Proposition 1,*

$$\frac{2\sigma}{c(1+\sqrt{1+\kappa_1\sigma^2})} \le |\widetilde{\mathcal{B}}_+(z)| \le \frac{2}{c}\Big(\frac{1+4\kappa_1\sigma^2}{4\kappa_1^3\sigma^2}\Big)^{\frac{1}{4}}, \qquad \textit{for } z\in\partial\mathbb{D}. \tag{42}$$

$$\frac{2}{c}\Big(\frac{4\sigma^2}{\kappa_1(1+4\kappa_1\sigma^2)}\Big)^{\frac{1}{4}} \le |\widetilde{\mathcal{B}}_-(z)| \le c\frac{1+\sqrt{1+\kappa_1\sigma^2}}{2\kappa\sigma}, \qquad \textit{for } z\in\partial\mathbb{D}. \tag{43}$$

Furthermore, under the conditions $\sigma \ge 1/\sqrt{\kappa_1}$, *we have*

$$\max_{z\in\partial\mathbb{D}} |r(\mathcal{S}(z))| \le 1-\delta(\kappa,\sigma), \tag{44}$$

$$\max_{z\in\partial\mathbb{D}} \mathrm{Re}\,[\widetilde{\mathcal{B}}_{\pm}(z)] \le -\mu(\kappa,\sigma), \tag{45}$$

$$\begin{aligned} -\arctan(\frac{1}{\sqrt{\tau}}) &\le \arg_{z\in\partial\mathbb{D}}\Big[-\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_{\pm}(z)\Big] \\ &\le \arctan(\frac{1}{\sqrt{\tau}}), \end{aligned} \tag{46}$$

where $\mathcal{S}(z)$ *is defined in* (32).

Proof 1 *Let* $\rho = i^{-1}\dfrac{2(1-z)}{\tau(1+z)}$ *for* $z\in\partial\mathbb{D}$. *Then we have*

$$\begin{aligned} \max_{z\in\partial\mathbb{D}}|\widetilde{\mathcal{B}}_+(z)| &= \frac{2}{c}\max_{\rho\in\mathbb{R}}\Big|\frac{i\rho+\sigma}{1+\sqrt{1+\kappa_1(i\rho+\sigma)^2}}\Big| \le \frac{2}{c}\max_{\rho\in\mathbb{R}}\Big|\frac{i\rho+\sigma}{\sqrt{1+\kappa_1(i\rho+\sigma)^2}}\Big| \\ &= \frac{2}{c}\left(\frac{(\rho^2+\sigma^2)^2}{\kappa_1^2(\rho^2+\sigma^2)^2+1+2\kappa_1\sigma^2-2\rho^2\kappa_1}\right)^{\frac{1}{4}} \\ &= \frac{2}{c}\left(\frac{1}{\kappa_1^2+(1+2\kappa_1\sigma^2-2\rho^2\kappa_1)/(\rho^2+\sigma^2)^2}\right)^{\frac{1}{4}} \\ &= \frac{2}{c}\left(\frac{1}{\kappa_1^2+\frac{1+4\kappa_1\sigma^2}{(\rho^2+\sigma^2)^2}-\frac{2\kappa_1}{\rho^2+\sigma^2}}\right)^{\frac{1}{4}}. \end{aligned}$$

Thus,

$$|\widetilde{\mathcal{B}}_+(z)| \le \frac{2}{c}\Big(\frac{1+4\kappa_1\sigma^2}{4\kappa_1^3\sigma^2}\Big)^{\frac{1}{4}}, \qquad \textit{for } z\in\partial\mathbb{D}.$$

This leads to

$$|\widetilde{\mathcal{B}}_-(z)| = \Big|-\frac{1}{\kappa\widetilde{\mathcal{B}}_+(z)}\Big| \ge \frac{2}{c}\Big(\frac{4\sigma^2}{\kappa_1(1+4\kappa_1\sigma^2)}\Big)^{\frac{1}{4}}, \qquad \textit{for } z\in\partial\mathbb{D}.$$

In the same way, we have

$$\begin{aligned}
\max_{z\in\partial\mathbb{D}}|\widetilde{\mathcal{B}}_-(z)| &= \max_{z\in\partial\mathbb{D}}\Big|\frac{1}{\kappa\widetilde{\mathcal{B}}_+(z)}\Big| = \frac{c}{2\kappa}\max_{\rho\in\mathbb{R}}\Big|\frac{1+\sqrt{1+\kappa_1(i\rho+\sigma)^2}}{i\rho+\sigma}\Big| \\
&\le \frac{c}{2\kappa}\max_{\rho\in\mathbb{R}}\Big|\frac{1}{i\rho+\sigma}\Big| + \frac{c}{2\kappa}\max_{\rho\in\mathbb{R}}\Big|\frac{\sqrt{1+\kappa_1(i\rho+\sigma)^2}}{i\rho+\sigma}\Big| \\
&= \frac{c}{2\kappa\sigma} + \frac{c}{2\kappa}\max_{\rho\in\mathbb{R}}\left(\frac{\kappa_1^2(\rho^2+\sigma^2)^2+1+2\kappa_1\sigma^2-2\rho^2\kappa_1}{(\rho^2+\sigma^2)^2}\right)^{\frac{1}{4}} \\
&= \frac{c}{2\kappa\sigma} + \frac{c}{2\kappa}\max_{\rho\in\mathbb{R}}\Big(\kappa_1^2+(1+2\kappa_1\sigma^2-2\rho^2\kappa_1)/(\rho^2+\sigma^2)^2\Big)^{\frac{1}{4}} \\
&= c\frac{1+\sqrt{1+\kappa_1\sigma^2}}{2\kappa\sigma},
\end{aligned}$$

which leads to

$$|\widetilde{\mathcal{B}}_+(z)| = \Big|-\frac{1}{\kappa\widetilde{\mathcal{B}}_-(z)}\Big| \ge \frac{2\sigma}{c(1+\sqrt{1+\kappa_1\sigma^2})}, \qquad \textit{if} \quad z\in\partial\mathbb{D},$$

which finishes the proof of (42) *and* (43).

Recalling $\widetilde{\mathcal{B}}_-(z) = -c\dfrac{1+\sqrt{1+\kappa_1(i\rho+\sigma)^2}}{2\kappa_1(i\rho+\sigma)}$, *we take in the sequel* $\rho>0$ *for our discussion. It is clear that* $0\le\arg\Big[(i\rho+\sigma)^2\Big]\le\pi$. *Thus, we have*

$$\arg\Big[1+\sqrt{1+\kappa_1(i\rho+\sigma)^2}\Big] \le \arg\Big[\sqrt{1+\kappa_1(i\rho+\sigma)^2}\Big],$$

which leads to

$$\begin{aligned}
\arg\Big[\frac{i\rho+\sigma}{1+\sqrt{1+\kappa_1(i\rho+\sigma)^2}}\Big] &\ge \arg\Big[\frac{i\rho+\sigma}{\sqrt{1+\kappa_1(i\rho+\sigma)^2}}\Big] \\
&\ge \frac{1}{2}\arg\left(\frac{2\sigma\rho+i(\rho^2-\sigma^2)}{2\sigma\rho+i(\rho^2-\sigma^2-1/\kappa_1)}\right) \ge 0.
\end{aligned}$$

There are two cases for $\rho>0$.

- $\rho \le \sigma$.

$$\arg\Big[\frac{i\rho+\sigma}{1+\sqrt{1+\kappa_1(i\rho+\sigma)^2}}\Big] \le \arg\Big[i\rho+\sigma\Big] \le \arg\Big[i\sigma+\sigma\Big] = \pi/4.$$

- $\rho \ge \sigma$. *Taking* $A(\rho) = |1+\kappa_1(i\rho+\sigma)^2|^{\frac{1}{2}}$,

$$\begin{aligned}
&\arg\Big[\frac{i\rho+\sigma}{1+\sqrt{1+\kappa_1(i\rho+\sigma)^2}}\Big] \le \arg\Big[\frac{i\rho+\sigma}{1+e^{i\pi/4}A(\rho)}\Big] \\
&= \arg\Big[\sigma+\frac{\sqrt{2}}{2}A(\rho)\cdot\sigma+\frac{\sqrt{2}}{2}A(\rho)\cdot\rho+i\Big(\rho+\frac{\sqrt{2}}{2}A(\rho)\cdot\rho-\frac{\sqrt{2}}{2}A(\rho)\cdot\sigma\Big)\Big] \\
&= \arctan\Big[\frac{\rho+\frac{\sqrt{2}}{2}A(\rho)\cdot\rho-\frac{\sqrt{2}}{2}A(\rho)\cdot\sigma}{\sigma+\frac{\sqrt{2}}{2}A(\rho)\cdot\sigma+\frac{\sqrt{2}}{2}A(\rho)\cdot\rho}\Big].
\end{aligned}$$

It is clear that

$$\frac{\rho+\frac{\sqrt{2}}{2}\rho\cdot A(\rho)-\frac{\sqrt{2}}{2}\sigma\cdot A(\rho)}{\sigma+\frac{\sqrt{2}}{2}\sigma\cdot A(\rho)+\frac{\sqrt{2}}{2}\rho\cdot A(\rho)} > 0$$

and

$$\lim_{\rho\to\infty}\frac{\rho+\frac{\sqrt{2}}{2}\rho\cdot A(\rho)-\frac{\sqrt{2}}{2}\sigma\cdot A(\rho)}{\sigma+\frac{\sqrt{2}}{2}\sigma\cdot A(\rho)+\frac{\sqrt{2}}{2}\rho\cdot A(\rho)} = 1.$$

Thus, for $\rho \ge \sigma$, $\dfrac{\rho+\frac{\sqrt{2}}{2}\rho\cdot A(\rho)-\frac{\sqrt{2}}{2}\sigma\cdot A(\rho)}{\sigma+\frac{\sqrt{2}}{2}\sigma\cdot A(\rho)+\frac{\sqrt{2}}{2}\rho\cdot A(\rho)}$ *can obtain maximum* $A_m(\sigma,\kappa)$.

Combing the two cases for $\rho > 0$, *one obtains*

$$\begin{aligned}
&\arg_{\rho>0}\Big[\frac{i\rho+\sigma}{1+\sqrt{1+\kappa_1(i\rho+\sigma)^2}}\Big] \\
&\qquad \le \theta_m(\kappa,\sigma) = \max\Big(\pi/4, \arctan(A_m(\sigma,\kappa))\Big) < \pi/2. \quad (47)
\end{aligned}$$

In the same way for the case of $\rho < 0$ *we have*

$$0 \ge \arg_{\rho\le 0}\Big[\frac{i\rho+\sigma}{1+\sqrt{1+\kappa(i\rho+\sigma)^2}}\Big] \ge -\theta_m, \quad (48)$$

which leads to

$$-\theta_m(\kappa,\sigma) \le \arg_{z\in\partial\mathbb{D}}[-\widetilde{\mathcal{T}}_-(z)] \le \theta_m(\kappa,\sigma). \tag{49}$$

By the same manner, for $\mathcal{B}_+(z)$ *one obtains*

$$-\theta_m(\kappa,\sigma) \le \arg_{z\in\partial\mathbb{D}}[-\widetilde{\mathcal{B}}_+(z)] \le \theta_m(\kappa,\sigma). \tag{50}$$

In addition, recalling (42) *and* (43)*, above two estimates yield*

$$\mathrm{Re}\,[\widetilde{\mathcal{B}}_+(z)] \le -\frac{2\sigma}{c(1+\sqrt{1+\kappa_1\sigma^2})}\cos\theta_m(\kappa,\sigma).$$

and

$$\mathrm{Re}\,[\widetilde{\mathcal{B}}_-(z)] \le -\frac{2}{c}\Big(\frac{4\sigma^2}{\kappa_1(1+4\kappa_1\sigma^2)}\Big)^{\frac{1}{4}}\cos\theta_m(\kappa,\sigma).$$

This proves (45).

Recalling $\mathcal{S}(z) = \Big(\frac{c\tau(1+z)}{2\kappa(2-2z+\sigma\tau(1+z))}\Big)^2\Big(1+\kappa_1\frac{(2-2z+\sigma\tau(z+1))^2}{\tau^2(1+z)^2}\Big)$, *for* $z\in\partial\mathbb{D}$,

$$\mathcal{S}(z) = \Big(\frac{c}{2\kappa}\Big)^2\frac{1+\kappa_1(i\rho+\sigma)^2}{(i\rho+\sigma)^2},$$

with $\rho = i^{-1}\dfrac{2(1-z)}{\tau(1+z)} \in \mathbb{R}$. *It is easy to verify*

$$|\sqrt{\mathcal{S}(z)}| = \frac{c}{2\kappa}\Big(\kappa_1^2+(1+2\kappa_1\sigma^2-2\rho^2\kappa_1)/(\rho^2+\sigma^2)^2\Big)^{1/4} \le \mathcal{S}_1 = \frac{c\sqrt{\kappa_1\sigma^2+1}}{2\kappa\sigma}, \tag{51}$$

and

$$|\sqrt{\mathcal{S}(z)}| \ge \mathcal{S}_2 = \frac{c}{2\kappa}\Big(\frac{4\kappa_1^3\sigma^2}{1+4\kappa_1\sigma^2}\Big)^{\frac{1}{4}}. \tag{52}$$

It is also easy to for $\rho \ge 0$,

$$\theta = \arg\Big(\mathcal{S}(z)\Big) = \arg\Big(\frac{2\sigma\rho+i(\rho^2-\sigma^2-1/\kappa_1)}{2\sigma\rho+i(\rho^2-\sigma^2)}\Big),$$

from which we have

$$\begin{aligned}\theta &= -\arctan\frac{2\sigma\rho/\kappa_1}{\rho^4+(2\sigma^2-1/\kappa_1)\rho^2+\sigma^4+\sigma^2/\kappa_1}\\ &= -\arctan\frac{2\sigma\rho/\kappa_1}{\Theta(\rho)},\end{aligned} \tag{53}$$

with $\Theta(\rho)=\rho^4+(2\sigma^2-1/\kappa_1)\rho^2+\sigma^4+\sigma^2/\kappa_1$. *It is straightforward to derive that* $\Theta(\rho)\geq 0$, *for* $\sigma\geq\dfrac{1}{\sqrt{\kappa_1}}$ $(\sigma^2\geq\dfrac{1}{\kappa_1})$. *Therefore, from* (53) *we derive that* $-\pi/2\leq\theta\leq 0$ *for* $\rho\in[0,+\infty)$, *which means* $-\pi/4\leq\theta/2\leq 0$. *In the same manner, one derives* $0\leq\theta\leq\pi/2$ *for* $\rho\in(-\infty,0]$, *which means* $0\leq\theta/2\leq\pi/4$. *Therefore by*

$$\sqrt{\mathcal{S}(z)}=\sqrt{|\mathcal{S}(z)|}\bigg(\cos\Big(\frac{\theta}{2}\Big)+i\sin\Big(\frac{\theta}{2}\Big)\bigg),$$

and $\cos\big(\dfrac{\theta}{2}\big)\geq\dfrac{\sqrt{2}}{2}$, *for* $z\in\partial\mathbb{D}$ *we have*

$$\begin{aligned}|r(\mathcal{S}(z))|=\left|\frac{\sqrt{\mathcal{S}(z)}-1}{\sqrt{\mathcal{S}(z)}+1}\right|&=\sqrt{1-\frac{4\sqrt{|\mathcal{S}(z)|}\cos\big(\theta/2\big)}{|\mathcal{S}(z)|+2\sqrt{|\mathcal{S}(z)|}\cos\big(\theta/2\big)+1}}\\&\leq 1-\frac{\sqrt{2}\sqrt{|\mathcal{S}(z)|}}{|\mathcal{S}(z)|+\sqrt{2}\sqrt{|\mathcal{S}(z)|}+1},\end{aligned}$$

where the last inequality is due to Taylor's expansion $(1-x)^{\frac{1}{2}}=1-\frac{1}{2}x-\frac{1}{8}x^2+\cdots\leq 1-\frac{1}{2}x$. *Recalling* (51) *and* (52), *by considering*

$$\frac{\sqrt{2}r}{r^2+\sqrt{2}r+1}=\frac{\sqrt{2}}{r+\sqrt{2}+1/r},$$

we see that the minimum value of $\dfrac{\sqrt{2}\sqrt{|\mathcal{S}(z)|}}{|\mathcal{S}(z)|+\sqrt{2}\sqrt{|\mathcal{S}(z)|}+1}$ *is attained at* $|\mathcal{S}(z)|=(\mathcal{S}_1)^2$ *or* $|\mathcal{S}(z)|=(\mathcal{S}_2)^2$. *Thus, we obtain*

$$\begin{aligned}\frac{\sqrt{2}\sqrt{|\mathcal{S}(z)|}}{|\mathcal{S}(z)|+\sqrt{2}\sqrt{|\mathcal{S}(z)|}+1}&\geq\min\Big(\frac{\sqrt{2}\mathcal{S}_1}{(\mathcal{S}_1)^2+\sqrt{2}\mathcal{S}_1+1},\frac{\sqrt{2}\mathcal{S}_2}{(\mathcal{S}_2)^2+\sqrt{2}\mathcal{S}_2+1}\Big)\\&=\delta(\kappa,\sigma)<1,\end{aligned}$$

which completes the proof of (44).

We define $\mathrm{T}_{-}^{1}(\rho)$ *and* $\mathrm{T}_{-}^{2}(\rho)$ *in the following way,*

$$
\begin{aligned}
&\Big[-\frac{2\kappa}{c}\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{T}}_{-}(z)\Big]\\
&=\Big[\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\frac{\sqrt{1+\kappa_1\Big(\frac{2(z^{-1}-1)}{\tau(z^{-1}+1)}+\sigma\Big)^2}}{\frac{2(z^{-1}-1)}{\tau(z^{-1}+1)}+\sigma}\Big]+\frac{1}{\frac{2(z^{-1}-1)}{(z^{-1}+1)}+\sigma}\\
&=\Big[\frac{\rho i+\sigma}{\tau\rho i+\sigma}\frac{\sqrt{1+\kappa_1\Big(i\rho+\sigma\Big)^2}}{i\rho+\sigma}\Big]+\frac{1}{i\tau\rho+\sigma}\\
&=\mathrm{B}_{-}^{1}(\rho)+\mathrm{B}_{-}^{2}(\rho).
\end{aligned}
$$

Recalling (53)*, we can define* $F_1(\rho)$ *and* $F_2(\rho)$ *in the following way,*

$$
\begin{aligned}
\arg\left[\mathrm{B}_{-}^{1}(\rho)\right]&=\frac{1}{2}\Big[-\arctan\frac{2\sigma\rho/\kappa_1}{\Theta(\rho)}+2\arg\frac{\rho i+\sigma}{\tau\rho i+\sigma}\Big]\\
&=\arctan\frac{-2\sigma\rho/\kappa_1}{\sqrt{(\Theta)^2+(2\sigma\rho/\kappa_1)^2}+\Theta}+\arctan\frac{\sigma(1-\tau)\rho}{\sigma^2+\tau\rho^2},\qquad(54)\\
&=\arctan F_1(\rho)+\arctan F_2(\rho)=\arctan\frac{F_1(\rho)+F_2(\rho)}{1-F_1(\rho)F_2(\rho)}.
\end{aligned}
$$

It is easy to verify that,

$$
\begin{aligned}
\Big|F_1(\rho)F_2(\rho)\Big|&\le\frac{2\sigma^2\rho^2/\kappa_1}{2\Theta\sigma^2}=\frac{\rho^2/\kappa_1}{\Theta}\le\frac{\rho^2/\kappa_1}{\rho^4+(2\sigma^2-1/\kappa_1)\rho^2+\sigma^4+\sigma^2/\kappa_1}\\
&\le\frac{1/\kappa_1}{(2\sigma^2-1/\kappa_1)+2\sqrt{\sigma^4+\sigma^2/\kappa_1}}\le\frac{1}{\sqrt{3}}.\qquad(55)
\end{aligned}
$$

Then for $\rho>0$*,* (54) *and* (55) *indicates that,*

$$
\begin{aligned}
\arg[\mathrm{B}_{-}^{1}(\rho)]&=\arctan_{\rho\in\mathbb{R}^+}\frac{F_1(\rho)+F_2(\rho)}{1-F_1(\rho)F_2(\rho)}\le\arctan\Big[\frac{0+\frac{\sigma(1-\tau)\rho}{\sigma^2+\tau\rho^2}}{1-\frac{1}{\sqrt{3}}}\Big]\\
&\le\arctan\Big[\frac{\frac{\sigma(1-\tau)\rho}{\sigma^2+\tau\rho^2}}{1-\frac{1}{2}}\Big]\le\theta_\rho^{+}=\arctan\Big[\frac{2\sigma(1-\tau)\rho}{\sigma^2+\tau\rho^2}\Big]>0,
\end{aligned}
\qquad(56)
$$

and

$$\begin{aligned}\arg\left[\mathrm{B}^1_-(\rho)\right] &= \arctan_{\rho\in\mathbb{R}^+}\ \frac{F_1(\rho)+F_2(\rho)}{1-F_1(\rho)F_2(\rho)} \geq \arctan\Big[\frac{F_1(\rho)+0}{1-\frac{1}{\sqrt{3}}}\Big]\\ &\geq \arctan\Big[\frac{F_1(\rho)}{1-\frac{1}{2}}\Big] \geq \theta^-_\rho = -\arctan\Big[\frac{4\sigma\rho/\kappa_1}{\sqrt{(\Theta)^2+(2\sigma\rho/\kappa_1)^2}+\Theta}\Big]\\ &\geq -\arctan\Big[\frac{4\sigma\rho/\kappa_1}{2\sigma\rho/\kappa_1}\Big] = -\arctan(2). \end{aligned} \tag{57}$$

In addition,

$$\frac{d}{d\rho}\theta^+_\rho = \frac{d}{d\rho}\Big[\arctan\frac{2\sigma(1-\tau)\rho}{\sigma^2+\tau\rho^2}\Big] = \frac{\sigma(1-\tau)(\sigma^2-\tau\rho^2)}{(\sigma^2+\tau\rho^2)^2+\sigma^2(1-\tau)^2\rho^2},$$

where θ^+_ρ *increases with respect to* $|\rho|$ *in* $[0,\frac{\sigma}{\sqrt{\tau}}]$ *and decreases in* $[\frac{\sigma}{\sqrt{\tau}},\infty)$, *such that* $\theta^+_\rho \leq \theta^+_{\frac{\sigma}{\sqrt{\tau}}} = \arctan(\frac{1-\tau}{\sqrt{\tau}}) \leq \arctan\frac{1}{\sqrt{\tau}}$. *Thus, by* (56) *one obtains*

$$\begin{aligned}&\arg_{\rho>0}\Big[-\frac{2\kappa}{c}\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_-(z)\Big]\\ &= \arg_{\rho>0}\Big[\mathrm{B}^1_-(\rho)+\mathrm{B}^2_-(\rho)\Big] \leq \arg_{\rho>0}\Big[|\mathrm{B}^1_-(\rho)|e^{i\arg[\mathrm{B}^1_-(\rho)]}+\frac{1}{i\tau\rho+\sigma}\Big]\\ &= \arg_{\rho>0}\Big[|\mathrm{B}^1_-(\rho)|\cos\Big(\arg\left[\mathrm{B}^1_-(\rho)\right]\Big)+\frac{\sigma}{\sqrt{\tau^2\rho^2+\sigma^2}}\\ &\quad + i\Big(|\mathrm{B}^1_-(\rho)|\sin\Big(\arg\left[\mathrm{B}^1_-(\rho)\right]\Big)-\frac{\tau\rho}{\sqrt{\tau^2\rho^2+\sigma^2}}\Big)\Big]\\ &\leq \arctan_{\rho>0}\Big[\frac{|\mathrm{B}^1_-(\rho)|\sin\Big(\arg[\mathrm{B}^1_-(\rho)]\Big)}{|\mathrm{B}^1_-(\rho)|\cos\Big(\arg[\mathrm{B}^1_-(\rho)]\Big)}\Big] = \arg[\mathrm{B}^1_-(\rho)]\\ &\leq \theta^+_\rho = \arctan_{\rho>0}\Big[\frac{2\sigma(1-\tau)\rho}{\sigma^2+\tau\rho^2}\Big] \leq \arctan(\frac{1}{\sqrt{\tau}}).\end{aligned} \tag{58}$$

In the same way, by (57) *one gets*

$$\begin{aligned}&\arg_{\rho>0}\Big[-\frac{2\kappa}{c}\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_-(z)\Big] = \arg_{\rho>0}\Big[\mathrm{B}^1_-(\rho)+\mathrm{B}^2_-(\rho)\Big]\\ &\geq \arg_{\rho>0}\Big[|\mathrm{B}^1_-(\rho)|e^{i\theta^-_\rho}+\frac{1}{i\tau\rho+\sigma}\Big]\\ &= \arg_{\rho>0}\Big[|\mathrm{B}^1_-(\rho)|\cos\theta^-_\rho+\frac{\sigma}{\tau^2\rho^2+\sigma^2}+i(|\mathrm{B}^1_-(\rho)|\sin\theta^-_\rho-\frac{\tau\rho}{\tau^2\rho^2+\sigma^2})\Big],\end{aligned}$$

which leads to

$$
\begin{aligned}
&\arg\Big[\mathrm{B}_-^1(\rho)+\mathrm{B}_-^2(\rho)\Big] \geq -\arctan\frac{|\mathrm{B}_-^1(\rho)||\sin\theta_\rho^-|+\frac{\tau\rho}{\tau^2\rho^2+\sigma^2}}{|\mathrm{B}_-^1(\rho)|\cos\theta_\rho^-+\frac{\sigma}{\tau^2\rho^2+\sigma^2}}\\
&=-\arctan\Big[\frac{|\sin\theta_\rho^-|}{\cos\theta_\rho^-}+\frac{\frac{\tau\rho-|\sin\theta_\rho^-|\sigma/\cos\theta_\rho^-}{\tau^2\rho^2+\sigma^2}}{|\mathrm{B}_-^1(\rho)|\cos\theta_\rho^-+\frac{\sigma}{\tau^2\rho^2+\sigma^2}}\Big]\\
&=-\arctan\Big[\frac{|\sin\theta_\rho^-|}{\cos\theta_\rho^-}\\
&\quad+\frac{\tau\rho-|\sin\theta_\rho^-|\sigma/\cos\theta_\rho^-}{\sqrt[4]{\kappa_1^2\rho^4+2\kappa_1(\kappa_1\sigma^2-1)\rho^2+\kappa_1^2\sigma^4+2\kappa_1\sigma^2+1}\cdot\cos\theta_\rho^-+\frac{\sigma}{\sqrt{\tau^2\rho^2+\sigma^2}}}\Big]\\
&\geq-\arctan\Big[\frac{|\sin\theta_\rho^-|}{\cos\theta_\rho^-}\\
&\quad+\frac{\tau\rho}{\sqrt[4]{\kappa_1^2\rho^4+2\kappa_1(\kappa_1\sigma^2-1)\rho^2+\kappa_1^2\sigma^4+2\kappa_1\sigma^2+1}\cdot\cos\theta_\rho^-+\frac{\sigma}{\sqrt{\tau^2\rho^2+\sigma^2}}}\Big]\\
&=-\arctan\Big[\frac{|\sin\theta_\rho^-|}{\cos\theta_\rho^-}\\
&\quad+\frac{\tau}{\sqrt[4]{\kappa_1^2+2\kappa_1(\kappa_1\sigma^2-1)/\rho^2+(\kappa_1^2\sigma^4+2\kappa_1\sigma^2+1)/\rho^4}\cdot\cos\theta_\rho^-}\Big]\\
&\geq-\arctan\Big[\frac{|\sin\theta_\rho^-|}{\cos\theta_\rho^-}+\frac{\tau(\kappa_1\sigma^2+1)}{\sqrt{4\kappa_1^3\sigma^2}\cdot\cos\theta_\rho^-}\Big]=\arctan(-C_1),
\end{aligned}
\tag{59}
$$

where C_1 is a constant depended on κ and σ.

Therefore, by (58) *and* (59) *one has*

$$
\arctan(-C_1)\leq\arg_{\rho>0}\Big[\mathrm{B}_-^1(\rho)+\mathrm{B}_-^2(\rho)\Big]\leq\arctan(\frac{1}{\sqrt{\tau}}),
$$

and in the same manner, we can estimate the case for $\rho\leq 0$:

$$
-\arctan(\frac{1}{\sqrt{\tau}})\leq\arg_{\rho<0}\Big[\mathrm{B}_-^1(\rho)+\mathrm{B}_-^2(\rho)\Big]\leq\arctan(C_1).
$$

Combing the above two estimates, one derives

$$
\begin{aligned}
-\arctan(\frac{1}{\sqrt{\tau}})\leq\arg_{z\in\partial\mathbb{D}}\Big[-\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_-(z)\Big]\\
\leq\arctan(\frac{1}{\sqrt{\tau}}),
\end{aligned}
$$

which gives a half part of (46).

To prove another half part of (46), *we have*

$$\begin{aligned}
&\Big[-\frac{2\kappa}{c}\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_+(z)\Big]\\
&=\Big[\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\frac{\sqrt{1+\kappa_1\Big(\frac{2(z^{-1}-1)}{\tau(z^{-1}+1)}+\sigma\Big)^2}}{\frac{2(z^{-1}-1)}{\tau(z^{-1}+1)}+\sigma}\Big]\\
&\quad-\frac{1}{\frac{2(z^{-1}-1)}{(z^{-1}+1)}+\sigma}\\
&=\Big[\frac{\rho i+\sigma}{\tau\rho i+\sigma}\frac{\sqrt{1+\kappa_1\Big(i\rho+\sigma\Big)^2}}{i\rho+\sigma}\Big]-\frac{1}{i\tau\rho+\sigma}=\Big[\frac{\sqrt{1+\kappa_1\Big(i\rho+\sigma\Big)^2}-1}{i\tau\rho+\sigma}\Big].
\end{aligned}$$

We assume

$$\sqrt{1+\kappa_1\Big(i\rho+\sigma\Big)^2}=a+bi,$$

such that

$$a^2=\frac{1+\kappa_1\sigma^2-\kappa_1\rho^2+\sqrt{(1+\kappa_1\sigma^2-\kappa_1\rho^2)^2+4\kappa_1^2\sigma^2\rho^2}}{2},$$

and

$$b=\frac{\kappa_1\sigma\rho}{a}.$$

We can prove that $a>1$. *In fact, if* $a\leq 1$, *one has*

$$a^2=\frac{1+\kappa_1\sigma^2-\kappa_1\rho^2+\sqrt{(1+\kappa_1\sigma^2-\kappa_1\rho^2)^2+4\kappa_1^2\sigma^2\rho^2}}{2}\leq 1.$$

The above equality is equivalent to

$$\frac{\sqrt{(1+\kappa_1\sigma^2-\kappa_1\rho^2)^2+4\kappa_1^2\sigma^2\rho^2}}{2}\leq\frac{1-\kappa_1\sigma^2+\kappa_1\rho^2}{2},$$

from which one obtains

$$\kappa_1\sigma^2\rho^2\leq\rho^2-\sigma^2.$$

On the other hand, by $\sigma\geq 1/\sqrt{\kappa_1}$ *one derives*

$$\kappa_1\sigma^2\rho^2\geq\rho^2,$$

which is impossible.

Then it is easy to see $\dfrac{b}{a-1} = \dfrac{\kappa_1\sigma\rho}{a(a-1)}$ *has a maximum* $\vartheta_m(\kappa,\sigma)$ *for* $\rho\in[0,\infty)$. *Thus,*

$$\begin{aligned}
&\arg_{\rho>0}\Big[-\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_+(z)\Big]\\
&\qquad=\arg_{\rho>0}\Big[\frac{\sqrt{1+\kappa_1\Big(i\rho+\sigma\Big)^2}-1}{i\tau\rho+\sigma}\Big]\\
&\qquad=\arg\Big[\frac{a-1+bi}{i\tau\rho+\sigma}\Big]\le\arg\Big[a-1+bi\Big]\\
&\qquad\le\arctan(\frac{b}{a-1})\le\arctan(\vartheta_m).
\end{aligned}\tag{60}$$

On the other hand, one has

$$\begin{aligned}
&\arg\Big[-\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_+(z)\Big]\\
&\quad=\arg\Big[\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\frac{\frac{2(z^{-1}-1)}{\tau(z^{-1}+1)}+\sigma}{1+\sqrt{1+\kappa\Big(\frac{2(z^{-1}-1)}{\tau(z^{-1}+1)}+\sigma\Big)^2}}\Big]\\
&\quad=\arg\Big[\frac{i\rho+\sigma}{1+\sqrt{1+\kappa(i\rho+\sigma)^2}}\Big]+\arg\Big[\frac{\rho i+\sigma}{\rho\tau i+\sigma}\Big].
\end{aligned}$$

and

$$\begin{aligned}
\arg\Big[\frac{i\rho+\sigma}{1+\sqrt{1+\kappa(i\rho+\sigma)^2}}\Big]&\ge\arg\Big[\frac{i\rho+\sigma}{\sqrt{1+\kappa(i\rho+\sigma)^2}}\Big]\\
&\ge\arg\Big(\frac{2\sigma\rho+i(\rho^2-\sigma^2)}{2\sigma\rho+i(\rho^2-\sigma^2-1/\kappa)}\Big)\ge 0,
\end{aligned}$$

which leads to

$$\begin{aligned}
&\arg_{\rho>0}\Big[-\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_+(z)\Big]\\
&\qquad\ge\arg_{\rho>0}\Big[\frac{\rho i+\sigma}{\rho\tau i+\sigma}\Big]=\arg_{\rho>0}\Big[\sigma^2+\tau\rho^2+i\sigma(1-\tau)\rho\Big]\ge 0.
\end{aligned}\tag{61}$$

Combing (60) *and* (61)*, for small* τ *one gets*

$$0 \le \arg_{\rho>0}\Big[-\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_+(z)\Big] \le \arctan(\vartheta_m) \le \arctan(\frac{1}{\sqrt{\tau}}).$$

In the same manner we have

$$-\arctan(\frac{1}{\sqrt{\tau}}) \le \arg_{\rho<0}\Big[-\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_+(z)\Big] \le 0,$$

Finally we have

$$-\arctan(\frac{1}{\sqrt{\tau}}) \le \arg_{z\in\partial\mathbb{D}}\Big[-\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_+(z)\Big] \le \arctan(\frac{1}{\sqrt{\tau}}). \quad \square$$

For $R_m(s)$ defined in (30), the following error result was proved in [27].

Lemma 2 *For the error of the rational approximation for the square root*

$$e_m(s) := \sqrt{s} - R_m(s), \quad m = 0, 1, 2, \ldots$$

the following identity holds:

$$e_m(s) = 2\sqrt{s}\,\frac{r^{2m+1}(s)}{1+r^{2m+1}(s)}, \quad \textit{if} \;\; \mathrm{Re}(s) \ge 0 \;\; \textit{and} \;\; s \ne 0, \tag{62}$$

where $r(s)$ *is defined in* (41)*.*

Lemma 3 *Under the conditions* $\sigma \ge \frac{1}{\sqrt{\kappa_1}}$ *and the setting of Proposition 1, for small* τ *we have*

$$\max_{z\in\partial\mathbb{D}} \mathrm{Re}\ \widetilde{\mathcal{B}}_{\pm}^{(m)}(z) \le 0, \tag{63}$$

$$\max_{z\in\partial\mathbb{D}} \mathrm{Re}\ \Big[\frac{(z^{-1}-1)/\tau+\sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\widetilde{\mathcal{B}}_{\pm}^{(m)}(z)\Big] \le 0. \tag{64}$$

Proof 2 *From* (44) *of Lemma 1 we have following inequality*

$$\max_{z\in\partial\mathbb{D}} |r(\mathcal{S}(z))| \le 1 - \delta(\kappa, \sigma).$$

If $\sigma \geq \frac{1}{\sqrt{\kappa_1}}$ *and* m *satisfies* (38)*, then* $|r(\mathcal{S}(z))|^{2m+1} \leq [1-\delta]^{2m+1} \leq 1/2$. *As a result of* (62)*,*

$$\max_{z\in\partial\mathbb{D}} \left| \frac{\sqrt{\mathcal{S}(z)} - \sqrt{\mathcal{S}}^{(m)}(z)}{\sqrt{\mathcal{S}(z)}} \right| = \max_{z\in\partial\mathbb{D}} \left| \frac{2r^{2m+1}(\mathcal{S}(z))}{1+r^{2m+1}(\mathcal{S}(z))} \right| \leq \max_{z\in\partial\mathbb{D}} \frac{2\left|r(\mathcal{S}(z))\right|^{2m+1}}{1-\left|r(\mathcal{S}(z))\right|^{2m+1}}$$
$$\leq 4 \max_{z\in\partial\mathbb{D}} |r(\mathcal{S}(z))|^{2m+1},$$

then (42) *and* (45) *implies*

$$\begin{aligned}
\max_{z\in\partial\mathbb{D}} \mathrm{Re}\,[\widetilde{\mathcal{B}}_{\pm}^{(m)}(z)] &= \max_{z\in\partial\mathbb{D}} \Big[\mathrm{Re}\,\widetilde{\mathcal{B}}_{\pm}(z) - \mathrm{Re}\,\big(\widetilde{\mathcal{B}}_{\pm}(z) - \widetilde{\mathcal{B}}_{\pm}^{(m)}(z)\big)\Big] \\
&= \max_{z\in\partial\mathbb{D}} \Big[\mathrm{Re}\,\widetilde{\mathcal{B}}_{\pm}(z) \pm \mathrm{Re}\,\Big(\sqrt{\mathcal{S}(z)} - R_m(\mathcal{S}(z))\Big)\Big] \\
&\leq -\mu + \frac{\mu\sigma\kappa\tau^3}{c\sqrt{\kappa_1\sigma^2+1}} \max_{z\in\partial\mathbb{D}} \sqrt{|\mathcal{S}(z)|} \leq -\mu + \frac{\mu\sigma\kappa}{c\sqrt{\kappa_1\sigma^2+1}}\mathcal{S}_1 \\
&\leq -\mu + \mu/2 \leq 0,
\end{aligned}$$

which proves (63).

In addition, by (42) *and* (43)*, for* τ *small enough we have*

$$\begin{aligned}
&\arg_{z\in\partial\mathbb{D}} \Big[-\frac{(z^{-1}-1)/\tau + \sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2} \widetilde{\mathcal{B}}_{\pm}^{(m)}(z) \Big] \\
&= \arg_{z\in\partial\mathbb{D}} \Big[-\frac{(z^{-1}-1)/\tau + \sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2} \widetilde{\mathcal{B}}_{\pm}(z) \Big(1 + \frac{\widetilde{\mathcal{B}}_{\pm}^{(m)} - \widetilde{\mathcal{B}}_{\pm}}{\widetilde{\mathcal{B}}_{\pm}}\Big) \Big] \\
&= \arg_{z\in\partial\mathbb{D}} \Big[-\frac{(z^{-1}-1)/\tau + \sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2} \widetilde{\mathcal{B}}_{\pm}(z) \Big] \\
&\quad + \arg_{z\in\partial\mathbb{D}} \Big(1 + \frac{\widetilde{\mathcal{B}}_{\pm}^{(m)} - \widetilde{\mathcal{B}}_{\pm}}{\widetilde{\mathcal{B}}_{\pm}}\Big) \\
&= \arg_{z\in\partial\mathbb{D}} \Big[-\frac{(z^{-1}-1)/\tau + \sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2} \widetilde{\mathcal{B}}_{\pm}(z) \Big] \\
&\quad + \arg_{z\in\partial\mathbb{D}} \Big(1 \pm \frac{\sqrt{\mathcal{S}(z)} - R^{(m)}\Big(\mathcal{S}(z)\Big)}{\widetilde{\mathcal{B}}_{\pm}}\Big) \\
&\leq \arctan(\frac{1}{\sqrt{\tau}}) + \arg_{z\in\partial\mathbb{D}} \Big[1 + i\max\Big(\frac{c\mu(1+\sqrt{1+\kappa_1\sigma^2})}{4\sigma}, \\
&\qquad\qquad \frac{\mu c}{4}(\frac{\kappa_1(1+4\kappa_1\sigma^2)}{4\sigma^2})^{\frac{1}{4}}\Big)\tau^3\Big] \\
&\leq \arctan(\frac{1}{\sqrt{\tau}}) + \arctan(\frac{\sqrt{\tau}}{2}) \leq \arctan(\frac{4}{\sqrt{\tau}}).
\end{aligned}$$

Thus, for τ small enough,

$$-\arctan(\frac{4}{\sqrt{\tau}}) \le \arg_{z\in\partial\mathbb{D}} \Big(-\frac{(z^{-1}-1)/\tau + \sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\Big[\widetilde{\mathcal{B}}_{\pm}^{(m)}(z)\Big]\Big) \le \arctan(\frac{4}{\sqrt{\tau}}),$$

which completes the proof of (64). □

We give the proof of Proposition 1 with the consequences of (63), (64).

Proof 3 *Firstly, for small τ, if $\sigma \ge \frac{1}{\sqrt{\kappa_1}}$ and m satisfies* (38)*, then $[1-\delta]^{2m+1} \le \epsilon \le 1/2$. As a result, Lemma 2 implies*

$$\begin{aligned}&\max_{z\in\partial\mathbb{D}} \Big|\widetilde{\mathcal{B}}_{\pm}(z) - \widetilde{\mathcal{B}}_{\pm}^{(m)}(z)\Big| = \max_{z\in\partial\mathbb{D}} \Big|\sqrt{s}(z) - \sqrt{s}^{(m)}(z)\Big| \\ &\le \max_{z\in\partial\mathbb{D}} \frac{2\sqrt{|s(z)|}\,|r(s(z))|^{2m+1}}{1-|r(s(z))|^{2m+1}} \le \frac{\mu\tau^3}{2},\end{aligned}$$

which proves (39).

We assume $(D_\tau + \sigma E)u^k = 0$ for $k \ge n$. Thus, we have

$$u^{k+1} = \frac{1-\sigma\tau/2}{1+\sigma\tau/2}\,u^k \quad for \quad k \ge n,$$

which generates the sequence $\{u^k\}$ such that $(D_\tau + \sigma E)u^k = 0$ for $k \ge n$. From (63) *we have*

$$\begin{aligned}&\mathrm{Re}\sum_{k=0}^{n} \overline{(D_\tau+\sigma E)u^k}((D_\tau+\sigma E)\mathcal{T}_{\pm}^{(m)} * u)^k \\ &= \mathrm{Re}((D_\tau+\sigma E)u, (D_\tau+\sigma E)\mathcal{T}_{\pm}^{(m)} * u)_{\ell^2(\mathbb{C})} \\ &= \mathrm{Re}\int_{\partial\mathbb{D}} |\widetilde{u}(z)|^2 \overline{[z^{-1}-1+\sigma\tau(z^{-1}+1)/2]}\widetilde{\mathcal{B}}_{\pm}^{(m)}(z)[z^{-1}-1 \\ &\quad + \sigma\tau(z^{-1}+1)/2]\,\nu(dz)/\tau^2 \\ &= \mathrm{Re}\int_{\partial\mathbb{D}} |z|^{-2}|\widetilde{u}(z)|^2 \overline{[2-2z+\sigma\tau(1+z)]}\widetilde{\mathcal{B}}_{\pm}^{(m)}(z)[2-2z+\sigma\tau(1+z)]\,\nu(dz)/(4\tau^2) \\ &= \mathrm{Re}\int_{\partial\mathbb{D}} |z|^{-2}|\widetilde{u}(z)|^2 \widetilde{\mathcal{B}}_{\pm}^{(m)}(z)|[2-2z+\sigma\tau(1+z)]|^2\,\nu(dz)/(4\tau^2) \le 0.\end{aligned}$$

Analogously, we assume $(\tau D_\tau + \sigma E)u^k = 0$ for $k \ge n$. Thus, we have

$$u^{k+1} = \frac{1-\sigma/2}{1+\sigma/2}\,u^k \quad for \quad k \ge n,$$

which generates the sequence $\{u^k\}$ *such that* $Eu^k = 0$ *for* $k \geq n$. *We have*

$$\begin{aligned}
&\mathrm{Re}\sum_{k=0}^{n} \overline{(\tau D_\tau + \sigma E)u^k}((D_\tau + \sigma E)\mathcal{B}_\pm^{(m)} * u)^k \\
&= \mathrm{Re}((\tau D_\tau + \sigma E)u, (D_\tau + \sigma E)\mathcal{B}_\pm^{(m)} * u)_{\ell^2(\mathbb{C})} \\
&= \mathrm{Re}\int_{\partial\mathbb{D}} |\widetilde{u}(z)|^2 \widetilde{\mathcal{B}}_\pm^{(m)}(z)\overline{(z^{-1}-1)+\sigma(z^{-1}+1)/2}[(z^{-1}-1)/\tau \\
&\quad + \sigma(z^{-1}+1)/2]\,\nu(dz) \\
&= \mathrm{Re}\int_{\partial\mathbb{D}} |(z^{-1}-1)+\sigma(z^{-1}+1)/2|^2\Big(\frac{(z^{-1}-1)/\tau + \sigma(z^{-1}+1)/2}{(z^{-1}-1)+\sigma(z^{-1}+1)/2}\Big) \\
&\quad \widetilde{\mathcal{B}}_\pm^{(m)}(z)|\widetilde{u}(z)|^2\,\nu(dz) \leq 0.
\end{aligned}$$

This finishes the proof. □

6. Error estimate

Let $\varepsilon^n = \big(v_0^n - v(x_0,t_n), \ldots, v_{M+1}^n - v(x_{M+1},t_n)\big)$. We first give the main result on the error estimate:

Theorem 1 *Suppose that the solution* $u_1(x,t)$ *and* $u_2(x,t)$ *of* (2) *is sufficiently smooth, or equivalently that the solution* $v_1(x,t)$ *and* $v_2(x,t)$ *of* (3) *is sufficiently smooth. For* $\sigma \geq \frac{1}{\sqrt{\kappa_1}}$ *and a time step* τ *small enough, if* m *satisfies* (38) *with* μ *and* δ *given in Proposition 1, then we have the estimate:*

$$\max_{1\leq n\leq [T/\tau]} \big(\|\mathcal{P}\varepsilon^n\|_h^2 + |\nabla_h \varepsilon^n|_h^2\big) \leq \mathcal{O}(\tau^2 + h^2), \tag{65}$$

for a given computational time T.

It is easy to verify that the error vector ε^n defined in Theorem 1 satisfies the following equation:

$$(D_\tau + \sigma E)\mathcal{P}\varepsilon^n + c\nabla_h^m E\varepsilon^n = \kappa\Delta_h(D_\tau + \sigma E)\varepsilon^n + f^n, \quad \forall\, n \geq 0, \tag{66}$$

$$(\mathcal{B}_\pm^m * \gamma^\pm\varepsilon)^n - \partial_\nu^\pm \varepsilon^n = g_\pm^n, \quad \forall\, n \geq 0, \tag{67}$$

$$\varepsilon^0 = (0,\ldots,0), \tag{68}$$

where $f^n = \big(f_1^n, \ldots, f_M^n\big)$ and $g_\pm^n$ are given interior truncation errors and

boundary truncation errors of the time and space discretizations, i.e.

$$\begin{aligned} f_j^n =& -\left[(D_\tau+\sigma E)v(x_j,t_n)-(\partial_t v(x_j,t_{n+\frac{1}{2}})+\sigma v(x_j,t_{n+\frac{1}{2}}))\right]\\ &-\big[E(v(x_{j+1},t_n)-v(x_{j-1},t_n))/2h-\partial_x v(x_j,t_{n+\frac{1}{2}})\big]\\ &+\kappa\big[(D_\tau+\sigma E)(v(x_{j-1},t_n)-2v(x_j,t_n)+v(x_{j+1},t_n))/h^2\\ &\quad-\partial_x^2(\partial_t+\sigma)v(x_j,t_{n+\frac{1}{2}})\big], \qquad 1\le j\le M, \end{aligned} \tag{69}$$

$$\begin{aligned} g_\pm^n =& -(\mathcal{B}_\pm^{(m)}-\mathcal{B}_\pm)*\gamma^\pm v(t_n)\\ &-\big[\mathcal{B}_\pm*\gamma^\pm v(t_n)\mp\frac{c}{2\kappa(\partial_t+\sigma)}\Big(1\mp\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)\gamma^\pm v(t_n)\big]\\ &-\big[\pm\frac{c}{2\kappa(\partial_t+\sigma)}\Big(1\mp\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)\gamma^\pm v(t_n)\\ &\quad\mp\frac{c}{2\kappa(\partial_t+\sigma)}\Big(1\mp\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)v(x_\pm,t_n)\big]\\ &-\big[-\partial_\nu^\pm v(t_n)+\partial_\nu v(x_\pm,t_n)\big], \end{aligned} \tag{70}$$

with $v(t_n)=(v(x_1,t_n),\cdot,v(x_M,t_n))$.

The proof of Theorem 1 is presented in the following two subsections.

6.1. Estimate for the truncation errors

In this section we give the estimate for the truncation errors of the boundary scheme and the interior scheme.

Proposition 2

$$\|f^n\|_h+|g_\pm^n|+|D_\tau g_\pm^n|\le C(\tau^2+h^2). \tag{71}$$

Proof 4 Estimate of $|g_\pm^n|$. *Here we will prove*

$$g_\pm^n=\mathcal{O}(\tau^2+h^2). \tag{72}$$

We divide the proof into two steps.

Step 1*: First we prove*

$$\Big|\mathcal{B}_+*v(x_\pm,t)-\frac{c}{2\kappa(\partial_t+\sigma)}\Big(1-\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)v(x_+,t)\Big|\le\mathcal{O}(\tau^2). \tag{73}$$

Let us recall that

$$\widetilde{\mathcal{B}}_+(e^{-i\tau\xi}) = G\Big(\frac{2(1-e^{-i\tau\xi})}{\tau(1+e^{-i\tau\xi})}\Big),$$

with $G(s) = \dfrac{c}{2\kappa(s+\sigma)}\Big(1-\sqrt{1+\dfrac{4\kappa(s+\sigma)^2}{c^2}}\Big)$. *Therefore, we have*

$$\begin{aligned}\Big|\widetilde{\mathcal{B}}_+(e^{-i\tau\xi}) - G(i\xi)\Big| &= \Big|G(i\xi) - G(i\frac{2\tan(\tau\xi/2)}{\tau})\Big| = \Big|\int_{i\xi}^{i\frac{\tan(\tau\xi/2)}{\tau/2}} \frac{d}{ds}G(s)\,ds\Big| \\ &= \Big|\int_{i\xi}^{i\frac{\tan(\tau\xi/2)}{\tau/2}} \frac{2}{c}\frac{1}{\sqrt{1+\kappa_1(s+\sigma)^2}(1+\sqrt{1+\kappa_1(s+\sigma)^2})}ds\Big| \\ &\le C\Big|i\frac{\tan(\tau\xi/2)}{\tau/2} - i\xi\Big| \le C\Big|\int_0^\xi \Big(\frac{1}{1+\frac{\tau^2\xi_1^2}{4}} - 1\Big)d\xi_1\Big| \\ &= C\Big|\int_0^\xi \frac{\frac{\tau^2\xi_1^2}{4}}{1+\frac{\tau^2\xi_1^2}{4}}d\xi_1\Big| \le C\frac{\tau^2}{4}\int_0^{|\xi|}\xi_1^2 d\xi_1 \le C\tau^2|\xi|^3.\end{aligned}$$

Thus, one has

$$\Big|\widetilde{\mathcal{B}}_+(e^{-i\tau\xi}) - \frac{c}{2\kappa(i\xi+\sigma)}\Big(1-\sqrt{1+\frac{4\kappa(i\xi+\sigma)^2}{c^2}}\Big)\Big| \le C\tau^2|\xi|^3. \quad (74)$$

From (4a) *and* (4b) *we have*

$$\partial_t v(x,0) = \kappa\partial_{xx}\Big(\partial_t v(x,0)\Big), \quad \forall\, x\in[x_+,+\infty), \quad (75)$$

this implies that $\partial_t v(x,0) = Ce^{-\frac{x}{\sqrt{\kappa}}}$ *for* $x\in[x_+,+\infty)$. *Then we have*

$$\partial_t\partial_x v(x_+,0) = -\frac{C}{\sqrt{\kappa}}e^{-\frac{x_+}{\sqrt{\kappa}}}. \quad (76)$$

On the other hand, from (7) *one has*

$$\begin{aligned}\partial_x v(x_+,t) &= \frac{c}{2\kappa(\partial_t+\sigma)}\Big(1-\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)v_2(x_+,t) \\ &= \big(f * v(x_+,\cdot)\big)(t) = \int_0^t f(t-s)v(x_+,s)\,ds,\end{aligned}$$

with

$$f(t) = \mathcal{L}^{-1}\Big[\frac{c}{2\kappa(s+\sigma)}\Big(1-\sqrt{1+\frac{4\kappa(s+\sigma)^2}{c^2}}\Big)\Big].$$

This implies that

$$\partial_x \partial_t v(x_+, 0) = f(0)v(x_+, 0) + \big(\partial_t f * v(x_+, \cdot)\big)(0) = 0,$$

this means that the C in (76) is 0. Thus, $\partial_t v(x_+, 0) = 0$. Repeating this procedure, we can easily find that $v(x_+, t)$ and its time derivatives are zero at $t = 0$. Consequently, we obtain a sufficiently smooth function $v(x_+, t)$ defined for $t \in \mathbb{R}$ by extending $v(x_+, t)$ so that it is zero on $t \in (-\infty, 0]$. We define

$$\mathcal{B}_+ * v(x_+, t) := \sum_{j=0}^{\infty} \mathcal{B}_+^j v(x_+, t - j\tau), \quad \forall\, t \in \mathbb{R}, \tag{77}$$

which is consistent with the definition (22) at $t = t_n$. The Fourier transform in time of the last equation is

$$\begin{aligned}
&\mathcal{F}_t[\mathcal{B}_+ * v(x_+, t)](\xi) = \int_{\mathbb{R}} \mathcal{B}_+ * v(x_+, t) e^{-it\xi} dt \\
&= \sum_{j=0}^{\infty} \int_{\mathbb{R}} \mathcal{B}_+^j v(x_+, t - j\tau) e^{-it\xi}\, dt = \widetilde{\mathcal{B}}_+(e^{-i\tau\xi}) \mathcal{F}_t v(x_+, \xi) \\
&= \frac{c}{2\kappa(i\xi + \sigma)} \Big(1 - \sqrt{1 + \frac{4\kappa(i\xi + \sigma)^2}{c^2}}\Big) \mathcal{F}_t v(x_+, \xi) \\
&\quad + \big(\widetilde{\mathcal{B}}_+(e^{-i\tau\xi}) - \frac{c}{2\kappa(i\xi + \sigma)} \Big(1 - \sqrt{1 + \frac{4\kappa(i\xi + \sigma)^2}{c^2}}\Big)\big) \mathcal{F}_t v(x_+, \xi) \\
&= \mathcal{F}_t[\frac{c}{2\kappa(\partial_t + \sigma)} \Big(1 - \sqrt{1 + \frac{4\kappa(\partial_t + \sigma)^2}{c^2}}\Big) v(x_+, t)](\xi) \\
&\quad + \big(\widetilde{\mathcal{B}}_+(e^{-i\tau\xi}) - \frac{c}{2\kappa(i\xi + \sigma)} \Big(1 - \sqrt{1 + \frac{4\kappa(i\xi + \sigma)^2}{c^2}}\Big)\big) \mathcal{F}_t v(x_+, \xi)
\end{aligned}$$

which implies that

$$\begin{aligned}
&\Big|\mathcal{B}_+ * v(x_+,t) - \frac{c}{2\kappa(\partial_t+\sigma)}\Big(1-\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)v(x_+,t)\Big| \\
&= \Big|\mathcal{F}_\xi^{-1}\big[(\widetilde{\mathcal{B}}_+(e^{-i\tau\xi}) - \frac{c}{2\kappa(i\xi+\sigma)}\Big(1-\sqrt{1+\frac{4\kappa(i\xi+\sigma)^2}{c^2}}\Big))\mathcal{F}_t v(x_+,\xi)\big](t)\Big| \\
&\le \int_{\mathbb{R}} \Big|\widetilde{\mathcal{B}}_+(e^{-i\tau\xi}) - \frac{c}{2\kappa(i\xi+\sigma)}\Big(1-\sqrt{1+\frac{4\kappa(i\xi+\sigma)^2}{c^2}}\Big)\Big||\mathcal{F}_t v(x_+,\xi)|\,d\xi \\
&\le C\tau^2\int_{\mathbb{R}} |\xi|^3|\mathcal{F}_t v(x_+,\xi)|\,d\xi \\
&\le C\tau^2\int_{\mathbb{R}} \frac{1}{1+|\xi|}(1+|\xi|^4)|\mathcal{F}_t v(x_+,\xi)|\,d\xi \\
&\le C\tau^2\Big(\int_{\mathbb{R}}\frac{1}{(1+|\xi|)^2}d\xi\Big)^{\frac12}\Big(\int_{\mathbb{R}}(1+|\xi|^4)^2|\mathcal{F}_t v(x_+,\xi)|^2 d\xi\Big)^{\frac12} \\
&= C\tau^2\Big(\int_0^\infty (|v(x_+,t)|^2+|\partial_t^4 v(x_+,t)|^2)dt\Big)^{\frac12}
\end{aligned}$$

according to (74). *The estimate at* x_- *is the same and we obtain* (73).

Step 2*: Inequality* (39) *of Proposition 1 implies* $|\widetilde{\mathcal{B}}_\pm^{(m)}(z) - \widetilde{\mathcal{B}}_\pm(z)| \le C\tau^3$ *for* $|z|=1$. *Then*

$$(\mathcal{B}_\pm^{(m)})^j = \int_{\partial\mathbb{D}} \widetilde{\mathcal{B}}_\pm^{(m)}(z)z^{-j}\mu(dz) \quad \textit{and} \quad (\mathcal{B}_\pm)^j = \int_{\partial\mathbb{D}} \widetilde{\mathcal{B}}_\pm(z)z^{-j}\mu(dz)$$

imply that

$$|(\mathcal{B}_\pm^{(m)})^j - (\mathcal{B}_\pm)^j| \le \int_{\partial\mathbb{D}} |\widetilde{\mathcal{B}}_\pm^{(m)}(z) - \widetilde{\mathcal{B}}_\pm(z)|\mu(dz) \le C\tau^3.$$

Thus it holds that

$$\begin{aligned}
&\Big|\sum_{j=0}^n (\mathcal{B}_\pm^{(m)})^j v(t_{n-j}) - \sum_{j=0}^n (\mathcal{B}_\pm)^j v(t_{n-j})\Big| \\
&\qquad\qquad \le \sum_{j=0}^n |(\mathcal{B}_\pm^{(m)})^j - (\mathcal{B}_\pm)^j||v(t_{n-j})| \le \sum_{j=0}^n C\tau^3 \le C\tau^2,
\end{aligned}$$

which implies

$$(\mathcal{B}_\pm^{(m)} - \mathcal{B}_\pm) * \gamma^\pm v(t_n) = \mathcal{O}(\tau^2). \tag{78}$$

Besides, (73) *implies that*

$$\mathcal{B}_{\pm} * \gamma^{\pm} v(t_n) - \Big(\pm \frac{c}{2\kappa(\partial_t+\sigma)} \Big(1 \mp \sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big) \gamma^{\pm} v(t_n)\Big) = \mathcal{O}(\tau^2). \quad (79)$$

Since $\gamma^+ v = \dfrac{v_{M+1}+v_M}{2}$ *and* $x_+ = \dfrac{x_{M+1}+x_M}{2} = x_{M+\frac{1}{2}}$, *it follows*

$$\begin{aligned} &\Big| \frac{c}{2\kappa(\partial_t+\sigma)} \Big(1-\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big) \gamma^+ v(t_n) - \frac{c}{2\kappa(\partial_t+\sigma)} \Big(1-\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big) v(x_+, t_n) \Big| \\ =& \Big| \frac{\frac{c}{2\kappa(\partial_t+\sigma)} \Big(1-\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big) v(x_{M+1}, t_n) + \frac{c}{2\kappa(\partial_t+\sigma)} \Big(1-\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big) v(x_M, t_n)}{2} \\ &- \frac{c}{2\kappa(\partial_t+\sigma)} \Big(1-\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big) v(x_{M+\frac{1}{2}}, t_n) \Big| = \mathcal{O}(h^2), \end{aligned} \quad (80)$$

and

$$\begin{aligned} &\partial_\nu^+ v(t_n) - \partial_\nu v(x_+, t_n) = \frac{v(x_{M+1}, t_n) - v(x_M, t_n)}{h} - \partial_x v(x_{M+\frac{1}{2}}, t_n) = \mathcal{O}(h^2), \\ &\partial_\nu^- v(t_n) - \partial_\nu v(x_-, t_n) = -\frac{v(x_1, t_n) - v(x_0, t_n)}{h} + \partial_x v(x_{\frac{1}{2}}, t_n) = \mathcal{O}(h^2). \end{aligned} \quad (81)$$

Substituting (78)-(81) *into* (70) *yields* (72).

Estimate of $\|f^n\|_h$. *Here we will prove*

$$\|f^n\|_h \le \mathcal{O}(\tau^2 + h^2). \quad (82)$$

Recalling (69),*we estimate the three terms in the expression of* f_j^n *separately. Firstly, we have*

$$\begin{aligned} &(D_\tau + \sigma E) v(x_j, t_n) - (\partial_t v(x_j, t_{n+\frac{1}{2}}) + \sigma v(x_j, t_{n+\frac{1}{2}})) \\ &\quad = \Big(\frac{v(x_j, t_{n+1}) - v(x_j, t_n)}{\tau} - \partial_t v(x_j, t_{n+\frac{1}{2}}) \Big) \\ &\qquad + \sigma \Big(\frac{v(x_j, t_n) + v(x_j, t_{n+1})}{2} - v(x_j, t_{n+\frac{1}{2}}) \Big) \\ &\quad = \mathcal{O}(\tau^2). \end{aligned} \quad (83)$$

Secondly, it holds that

$$\big[E(v(x_{j+1},t_n)-v(x_{j-1},t_n))/(2h)-\partial_x v(x_j,t_{n+\frac{1}{2}})\big]=\mathcal{O}(\tau^2+h^2).$$

In the same manner, recalling (69),

$$(D_\tau+\sigma E)\frac{v(x_{j-1},t_n)-2v(x_j,t_n)+v(x_{j+1},t_n)}{h^2}-(\partial_t+\sigma)\partial_x^2 v_2(x_j,t_{n+\frac{1}{2}})$$
$$=\mathcal{O}(\tau^2+h^2),$$

Thus,

$$f_j^n=\mathcal{O}(\tau^2+h^2),\qquad 1\le j\le M. \tag{84}$$

Finally, from (83)–(84) *we obtain*

$$\|f^n\|_h=\mathcal{O}(\tau^2+h^2).$$

Estimate of $|D_\tau g_\pm^n|$. *Since*

$$\begin{aligned}
D_\tau g_\pm^n &= -(\mathcal{B}^{(m)}-\mathcal{B})*\gamma^\pm D_\tau v(t_n)\\
&-\big[\mathcal{B}*\gamma^\pm D_\tau v(t_n)\mp\frac{c}{2\kappa(\partial_t+\sigma)}\Big(1\mp\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)\gamma^\pm D_\tau v(t_n)\big]\\
&-\big[\pm\frac{c}{2\kappa(\partial_t+\sigma)}\Big(1\mp\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)\gamma^\pm D_\tau v(t_n)\\
&\mp\frac{c}{2\kappa(\partial_t+\sigma)}\Big(1\mp\sqrt{1+\frac{4\kappa(\partial_t+\sigma)^2}{c^2}}\Big)D_\tau v(x_\pm,t_n)\big]\\
&-\big[-\partial_\nu^\pm D_\tau v(t_n)+\partial_\nu D_\tau v(x_\pm,t_n)\big],
\end{aligned}\tag{85}$$

it follows that (85) *can be estimated similarly as* (70) *(replacing $v(x,t_n)$ by $D_\tau v(x,t_n)$), then we derive that*

$$D_\tau g_\pm^n=\mathcal{O}(\tau^2+h^2). \tag{86}$$

Combing (72), (82), *and* (86) *we prove* (71). □

6.2. Error estimate

Let us give the error estimate for (65). Firstly we prove the stability for the scheme (34)-(36) by the following estimate.

Lemma 4 *If $\sigma \geq 1/\sqrt{\kappa_1}$, and the order m of the Padé approximation satisfies* (38) *with μ and δ given in Proposition 1, the solution of* (34)–(36) *satisfies the following stability estimate:*

$$\max_{1\leq n\leq [T/\tau]} \big(\|\mathcal{P}v^n\|_h^2 + |\nabla_h v^n|_h^2\big) \leq C_T\big(\|\mathcal{P}v^0\|_h^2 + |\nabla_h v^0|_h^2\big), \tag{87}$$

where C_T is a constant depending on T.

Proof 5 *Due to*

$$D_\tau v_m^n \cdot E v_m^n = \frac{v_m^{n+1} - v_m^n}{\tau} \cdot \frac{v_m^{n+1} + v_m^n}{2} = \frac{|v_m^{n+1}|^2 - |v_m^n|^2}{2\tau},$$

Thus,

$$\mathrm{Re}\left(D_\tau \mathcal{P}v^n, E\mathcal{P}v^n\right)_h = D_\tau(\|\mathcal{P}v^n\|_h^2)/2.$$

Performing the inner product of (66) *with $(\tau D_\tau + \sigma E)\mathcal{P}v^n$ and taking the real part,*

$$\begin{aligned}\mathrm{Re}\left((\tau D_\tau + \sigma E)\mathcal{P}v^n, (D_\tau + \sigma E)\mathcal{P}v^n + \nabla_h^m E v^n\right)_h \\ = \mathrm{Re}\left((\tau D_\tau + \sigma E)\mathcal{P}v^n, \kappa\Delta_h^n (D_\tau + \sigma E)v\right)_h.\end{aligned} \tag{88}$$

The left of (88) *can be written as*

$$\begin{aligned}\tau\|D_\tau \mathcal{P}v^n\|_h^2 + \frac{\sigma(1+\tau)}{2} D_\tau(\|\mathcal{P}v^n\|_h^2) + \sigma^2\|E\mathcal{P}v^n\|_h^2 \\ + \mathrm{Re}\left((\tau D_\tau + \sigma E)\mathcal{P}v^n, \nabla_h^m E v^n\right)_h.\end{aligned} \tag{89}$$

By applying the discrete Green's formula (27), *the boundary conditions* (35) *and* (40), *the right of the above equality can be written as*

$$\begin{aligned}&-\kappa\langle\nabla_h(\tau D_\tau + \sigma E)v^n, \nabla_h(D_\tau + \sigma E)v^n\rangle_h \\ &\qquad + \kappa\,\mathrm{Re}\left(\overline{\gamma^\pm(\tau D_\tau + \sigma E)v^n}\partial^\pm(D_\tau + \sigma E)v^n\right) \\ &= -\kappa\langle\nabla_h(\tau D_\tau + \sigma E)v^n, \nabla_h(D_\tau + \sigma E)v^n\rangle_h \\ &\qquad + \kappa\,\mathrm{Re}\left(\overline{(\tau D_\tau + \sigma E)\gamma^\pm v^n}(D_\tau + \sigma E)\partial^\pm v^n\right) \\ &= -\kappa\langle\nabla_h(\tau D_\tau + \sigma E)v^n, \nabla_h(D_\tau + \sigma E)v^n\rangle_h \\ &\qquad + \kappa\,\mathrm{Re}\left(\overline{(\tau D_\tau + \sigma E)\gamma^\pm v^n}(D_\tau + \sigma E)(\mathcal{B}_\pm^{(m)} * \gamma^\pm v)^n\right) \\ &\leq -\kappa\langle\nabla_h(\tau D_\tau + \sigma E)v^n, \nabla_h(D_\tau + \sigma E)v^n\rangle_h.\end{aligned} \tag{90}$$

Combining (89) *and* (90) *we have*

$$\begin{aligned}
&\tau\|D_\tau\mathcal{P}v^n\|_h^2+\frac{\sigma(1+\tau)}{2}D_\tau(\|\mathcal{P}v^n\|_h^2)+\sigma^2\|E\mathcal{P}v^n\|_h^2\\
&\quad+\kappa\langle\nabla_h(\tau D_\tau+\sigma E)v^n,\nabla_h(D_\tau+\sigma E)v^n\rangle_h\\
&=\tau\|D_\tau\mathcal{P}v^n\|_h^2+\frac{\sigma(1+\tau)}{2}D_\tau(\|\mathcal{P}v^n\|_h^2)+\sigma^2\|E\mathcal{P}v^n\|_h^2\\
&\quad+\kappa\tau|D_\tau\nabla_h v^n|_h^2+\frac{\kappa\sigma(1+\tau)}{2}D_\tau(|\nabla_h v^n|_h^2)+\kappa\sigma^2|E\nabla_h v^n|_h^2\\
&\le-\mathrm{Re}\left((\tau D_\tau+\sigma E)\mathcal{P}v^n,\nabla_h^m Ev^n\right)_h,\\
&=\mathrm{Re}\left(\overline{(\tau D_\tau+\sigma E)v_1^n}Ev_0^n\right)/2-\mathrm{Re}\left(\overline{(\tau D_\tau+\sigma E)v_M^n}Ev_{M+1}^n\right)/2,
\end{aligned}$$

from which we derive that

$$\begin{aligned}
&D_\tau(\|\mathcal{P}v^n\|_h^2)+D_\tau(|\nabla_h v^n|_h^2)\\
&\le\mathcal{O}(1)\left(E(|\gamma^{\pm}v^n|^2)+E(|v_0^n|^2)+E(|v_1^n|^2)+E(|v_M^n|^2)+E(|v_{M+1}^n|^2)\right).
\end{aligned}$$

Summing up the above inequality from the 0*-th step to the* $n-1$*-th step yields*

$$\begin{aligned}
\|\mathcal{P}v^n\|_h^2+|\nabla_h v^n|_h^2&\le\|\mathcal{P}v^0\|_h^2+|\nabla_h v^0|_h^2\\
&\quad+\mathcal{O}(\tau)\sum_{k=0}^{n}\left(|\gamma^{\pm}v^n|^2+|v_0^n|^2+|v_1^n|^2+|v_M^n|^2+|v_{M+1}^n|^2\right).\quad(91)
\end{aligned}$$

Recalling the discrete Sobolev imbedding theorem

$$|\gamma^{\pm}v^n|^2\le C\|\mathcal{P}v^n\|_h^2+C|\nabla_h v^n|_h^2,$$

from (91) *we derive*

$$\|\mathcal{P}v^n\|_h^2+|\nabla_h v^n|_h^2\le\|\mathcal{P}v^0\|_h^2+|\nabla_h v^0|_h^2+\mathcal{O}(\tau)\sum_{k=1}^{n}\left(\|\mathcal{P}v^k\|_h^2+|\nabla_h v^k|_h^2\right).$$

Applying the discrete Gronwall's inequality to the above estimate, we derive (65)*. The proof of Lemma 4 is complete.* □

Lemma 4 assures that the numerical solution is bounded for a finite computational time.

Lemma 5 *If $\sigma \geq 1/\sqrt{\kappa_1}$, and the order m of the Padé approximation satisfies* (38) *with μ and δ given in Proposition 1, the solution of* (34)–(36) *satisfies the following estimate:*

$$\max_{1\leq n\leq [T/\tau]} \left(\|\mathcal{P}\varepsilon^n\|_h^2 + |\nabla_h \varepsilon^n|_h^2\right) \leq C_T\Big[\max_{0\leq k\leq n-1}(\|f^k\|_h^2 + \|D_\tau g_\pm^k\|_h^2) + \max_{0\leq k\leq n}|g_\pm^k|^2\Big], \quad (92)$$

where C_T is a constant depending on T.

Proof 6 *Due to*

$$D_\tau(\varepsilon)_m^n \cdot E(\varepsilon)_m^n = \frac{(\varepsilon)_m^{n+1} - (\varepsilon)_m^n}{\tau} \cdot \frac{(\varepsilon)_m^{n+1} + (\varepsilon)_m^n}{2} = \frac{|(\varepsilon)_m^{n+1}|^2 - |(\varepsilon)_m^n|^2}{2\tau},$$

Thus,

$$\mathrm{Re}\left(D_\tau \mathcal{P}\varepsilon^n, E\mathcal{P}\varepsilon^n\right)_h = D_\tau(\|\mathcal{P}\varepsilon^n\|_h^2)/2. \quad (93)$$

Using the inner product of (66) *with $(\tau D_\tau + \sigma E)\mathcal{P}\varepsilon^n$ and taking the real part, yields*

$$\mathrm{Re}\left((\tau D_\tau + \sigma E)\mathcal{P}\varepsilon^n, (D_\tau + \sigma E)\mathcal{P}\varepsilon^n + \nabla_h^m E\varepsilon^n\right)_h = \mathrm{Re}\left((\tau D_\tau + \sigma E)\mathcal{P}\varepsilon^n, \kappa\Delta_h^n(D_\tau + \sigma E)\varepsilon + f^n\right)_h. \quad (94)$$

By (93)*, the left side of* (94) *can be written as*

$$\tau\|D_\tau \mathcal{P}\varepsilon^n\|_h^2 + \frac{\sigma(1+\tau)}{2} D_\tau(\|\mathcal{P}\varepsilon^n\|_h^2) + \sigma^2\|E\mathcal{P}\varepsilon^n\|_h^2 + \mathrm{Re}\left((\tau D_\tau + \sigma E)\mathcal{P}\varepsilon^n, \nabla_h^m E\varepsilon^n\right)_h. \quad (95)$$

By applying the discrete Green's formula (27)*, the boundary conditions* (35)

and (40)*, the right of* (94) *can be written as*

$$
\begin{aligned}
&-\kappa\langle\nabla_h(\tau D_\tau+\sigma E)\varepsilon^n,\nabla_h(D_\tau+\sigma E)\varepsilon^n\rangle_h\\
&\quad+\kappa\,\mathrm{Re}\left(\overline{\gamma^{\pm}(\tau D_\tau+\sigma E)\varepsilon^n}\partial^{\pm}(D_\tau+\sigma E)\varepsilon^n\right)+\mathrm{Re}\left((\tau D_\tau+\sigma E)\mathcal{P}\varepsilon^n,f^n\right)_h\\
&=-\kappa\langle\nabla_h(\tau D_\tau+\sigma E)\varepsilon^n,\nabla_h(D_\tau+\sigma E)\varepsilon^n\rangle_h\\
&\quad+\kappa\,\mathrm{Re}\left(\overline{(\tau D_\tau+\sigma E)\gamma^{\pm}\varepsilon^n}(D_\tau+\sigma E)\partial^{\pm}\varepsilon^n\right)+\mathrm{Re}\left((\tau D_\tau+\sigma E)\mathcal{P}\varepsilon^n,f^n\right)_h\\
&=-\kappa\langle\nabla_h(\tau D_\tau+\sigma E)\varepsilon^n,\nabla_h(D_\tau+\sigma E)\varepsilon^n\rangle_h\\
&\quad+\kappa\,\mathrm{Re}\left(\overline{(\tau D_\tau+\sigma E)\gamma^{\pm}\varepsilon^n}(D_\tau+\sigma E)(\mathcal{B}_{\pm}^{(m)}*\gamma^{\pm}\varepsilon)^n\right)\\
&\quad-\kappa\,\mathrm{Re}\left(\overline{(\tau D_\tau+\sigma E)\gamma^{\pm}\varepsilon^n}(D_\tau+\sigma E)g_{\pm}^n\right)+\mathrm{Re}\left((\tau D_\tau+\sigma E)\mathcal{P}\varepsilon^n,f^n\right)_h\\
&\le-\kappa\langle\nabla_h(\tau D_\tau+\sigma E)\varepsilon^n,\nabla_h(D_\tau+\sigma E)\varepsilon^n\rangle_h\\
&\quad-\kappa\,\mathrm{Re}\left(\overline{(\tau D_\tau+\sigma E)\gamma^{\pm}\varepsilon^n}(D_\tau+\sigma E)g_{\pm}^n\right)+\mathrm{Re}\left((\tau D_\tau+\sigma E)\mathcal{P}\varepsilon^n,f^n\right)_h.
\end{aligned}
\tag{96}
$$

Combining (95) *and* (96) *we have*

$$
\begin{aligned}
&\tau\|D_\tau\mathcal{P}\varepsilon^n\|_h^2+\frac{\sigma(1+\tau)}{2}D_\tau(\|\mathcal{P}\varepsilon^n\|_h^2)+\sigma^2\|E\mathcal{P}\varepsilon^n\|_h^2\\
&+\kappa\langle\nabla_h(\tau D_\tau+\sigma E)\varepsilon^n,\nabla_h(D_\tau+\sigma E)\varepsilon^n\rangle_h\\
&=\tau\|D_\tau\mathcal{P}\varepsilon^n\|_h^2+\frac{\sigma(1+\tau)}{2}D_\tau(\|\mathcal{P}\varepsilon^n\|_h^2)+\sigma^2\|E\mathcal{P}\varepsilon^n\|_h^2+\\
&+\kappa\tau|D_\tau\nabla_h\varepsilon^n|_h^2+\frac{\kappa\sigma(1+\tau)}{2}D_\tau(|\nabla_h\varepsilon^n|_h^2)+\kappa\sigma^2|E\nabla_h\varepsilon^n|_h^2\\
&\le-\mathrm{Re}\left((\tau D_\tau+\sigma E)\mathcal{P}\varepsilon^n,\nabla_h^m E\varepsilon^n\right)_h+\mathrm{Re}\left((\tau D_\tau+\sigma E)\mathcal{P}\varepsilon^n,f^n\right)_h,\\
&-\kappa\,\mathrm{Re}\left(\overline{(\tau D_\tau+\sigma E)\gamma^{\pm}\varepsilon^n}(D_\tau+\sigma E)g_{\pm}^n\right),\\
&=\mathrm{Re}\left((\tau D_\tau+\sigma E)\mathcal{P}\varepsilon^n,f^n\right)_h-\kappa\,\mathrm{Re}\left(\overline{(\tau D_\tau+\sigma E)\gamma^{\pm}\varepsilon^n}(D_\tau+\sigma E)g_{\pm}^n\right)\\
&+\mathrm{Re}\left(\overline{(\tau D_\tau+\sigma E)\varepsilon_1^n}E\varepsilon_0^n\right)/2-\mathrm{Re}\left(\overline{(\tau D_\tau+\sigma E)\varepsilon_M^n}E\varepsilon_{M+1}^n\right)/2,
\end{aligned}
$$

from which we derive that

$$
\begin{aligned}
&D_\tau(\|\mathcal{P}\varepsilon^n\|_h^2)+D_\tau(|\nabla_h\varepsilon^n|_h^2)\\
&\le\mathcal{O}(1)\left(E(|\gamma^{\pm}\varepsilon^n|^2)+E(|\varepsilon_0^n|^2)+E(|\varepsilon_1^n|^2)+E(|\varepsilon_M^n|^2)+E(|\varepsilon_{M+1}^n|^2)\right)\\
&\quad+\mathcal{O}(1)\Big(\|f^n\|_h^2+|D_\tau g_{\pm}^n|^2+E(|g_{\pm}^n|^2)\Big).
\end{aligned}
$$

Summing up the above inequality from the 0*-th step to the* $n-1$*-th step, one obtains*

$$\begin{aligned}
&\|\mathcal{P}\varepsilon^n\|_h^2 + |\nabla_h\varepsilon^n|_h^2 \le \|\mathcal{P}\varepsilon^0\|_h^2 + |\nabla_h\varepsilon^0|_h^2 \\
&+ \quad \mathcal{O}(\tau)\sum_{k=0}^{n}\Big(|\gamma^{\pm}\varepsilon^n|^2 + |\varepsilon_0^n|^2 + |\varepsilon_1^n|^2 + |\varepsilon_M^n|^2 + |\varepsilon_{M+1}^n|^2\Big) \\
&+ \mathcal{O}(\tau)\sum_{k=0}^{n}|g_{\pm}^k|^2 + \mathcal{O}(\tau)\sum_{k=0}^{n-1}\Big(\|f^k\|_h^2 + |D_\tau g_{\pm}^k|^2\Big). \qquad (97)
\end{aligned}$$

By the discrete Sobolev imbedding theorem

$$|\gamma^{\pm}\varepsilon^n|^2 \le C\|\mathcal{P}\varepsilon^n\|_h^2 + C|\nabla_h\varepsilon^n|_h^2,$$

from which and (97) *we obtain*

$$\begin{aligned}
\|\mathcal{P}\varepsilon^n\|_h^2 + |\nabla_h\varepsilon^n|_h^2 \le \|\mathcal{P}\varepsilon^0\|_h^2 + |\nabla_h\varepsilon^0|_h^2 + \mathcal{O}(\tau)\sum_{k=1}^{n}\Big(\|\mathcal{P}\varepsilon^k\|_h^2 + |\nabla_h\varepsilon^k|_h^2\Big) \\
+ \mathcal{O}(\tau)\sum_{k=1}^{n}|g_{\pm}^k|^2 + \mathcal{O}(\tau)\sum_{k=0}^{n-1}\Big(\|f_2^k\|_h^2 + |D_\tau g_{\pm}^k|^2\Big). \qquad (98)
\end{aligned}$$

Applying the discrete Gronwall's inequality to the above estimate, we derive (65)*. The proof of Lemma 5 is complete.* □

Now, by Lemma 2 and Lemma 5, it is easy to prove Theorem 1.

7. Numerical results

We now perform numerical tests to confirm the theoretical results presented in the previous sections. In the calculations we determine the number of Padé expansion terms with the help of the following criterion:

$$m = \frac{\ln\epsilon}{2\ln(1-\delta)}, \qquad \epsilon = \frac{\mu\sigma\kappa}{4c\sqrt{\kappa_1\sigma^2+1}}\tau^3,$$

with μ and δ given in Proposition 1. When m is fixed, the computational cost of the fast convolution is of order $\mathcal{O}\big(mN\big) = \mathcal{O}\big(N\ln(N)\big)$ for $N\tau \le T$.

7.1. Example 1: Gaussian Initial Condition

We first consider an initial Gaussian distribution for the free surface elevation

$$u_0(x) = \exp\big(-400(x-1/2)^2\big).$$

The computational domain is $(t, x) \in [0, 1] \times [0, 1]$, and the parameters are taken as $c = 2$, $\kappa = 0.1$. We let $\sigma = 10$, which meets the condition $\sigma > 1/\sqrt{\kappa_1}$ in Theorem 1. The mesh sizes are $h = 10^{-5}$, $\tau = 10^{-4}$ respectively.

The displacement solution is depicted in the left panel of Figure 1. The influence of the wave has reached the boundary, where we do not see any significant artificial reflections, nor any numerical instability. A reference solution u_{ref} is computed with the same grid sizes but a much larger spatial domain ($x \in [-40, 40]$) to avoid any boundary reflections. The result is plotted in the right panel of Figure 1. It can be seen the computed solution matches u_{ref} well.

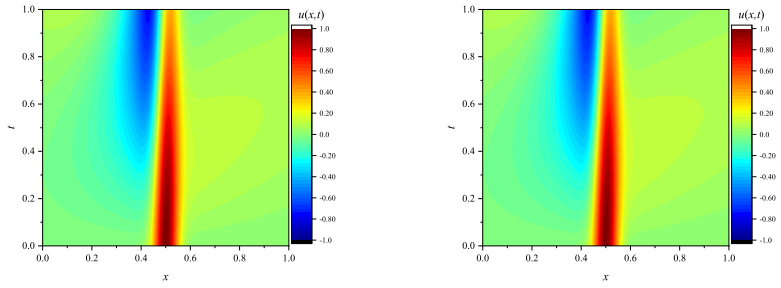

Figure 1: Solution to Example 1: (left) by the proposed method; (right) by using a large spatial domain

To study the performance of the proposed method, we fix the spatial mesh size $h = 10^{-5}$, but employ different time step sizes τ, ranging from 10^{-4} to 10^{-2}. The error of our solution is assessed by the relative difference between the obtained displacement vector at the final time step ($t = 1.0$) and the reference solution

$$\text{err} = \frac{\|u(x,1) - u_{\text{ref}}(x,1)\|_\infty}{\|u_{\text{ref}}(x,1)\|_\infty}.$$

The errors are then plotted vs. the time step size in the left panel of Figure 2. From the log-log plot, it can be seen that most data points are almost on a straight line with a slope of 2. It indicates a second-order convergence with respect to time step size. This agrees with our previous discussions.

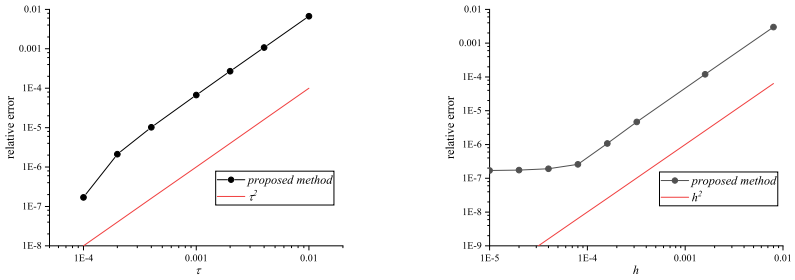

Figure 2: Performance of the proposed method in Example 1: (left) relative error versus time step size τ ($h = 10^{-5}$); (right) relative error versus spatial mesh size h ($\tau = 10^{-4}$)

Similarly, we fix $\tau = 10^{-4}$ but change the spatial mesh size h. The displacements at the nodes that overlay with the original mesh nodes are extracted and then compared with the reference solution. The relative error is shown in the right panel of Figure 2. We see a second-order convergence with respect to the spatial mesh size, which also agrees with our conclusion.

7.2. Example 2: Small Wave Packet

In this example, we further consider the following initial datum,

$$u_0(x) = \exp\big(-400(x - 1/2)^2\big)\sin(20\pi x).$$

To evaluate the long term performance of the proposed method, the equation parameters are taken as $c = 0.2$, $\kappa = 0.0001$ and the computer domain $(t, x) \in [0, 10] \times [0, 1]$. According to Theorem 1, we let $\sigma = 10$. The step sizes are taken as $h = 5 \times 10^{-5}, \tau = 2 \times 10^{-4}$.

The result is shown in the left panel of Figure 3. Since about $t = 3$, the wave packet has reached the right boundary, but no significant artificial reflection is observed throughout the computing time domain. The reference solution, which is computed in a much larger spatial domain $(t, x) \in [0, 10] \times [-40, 40]$ with the same step sizes ($h = 5 \times 10^{-5}, \tau = 2 \times 10^{-4}$), is given in the right panel of Figure 3 for comparison. As can be seen the two solutions match well.

The relative errors to the reference solution when employing a different h or τ are presented in Figure 4. Again we observe second-order convergence with respect to both spatial and time step sizes.

We now consider the computational cost. The CPU time is examined by increasing the total number of time steps N from $N = 1000$ up to $N = 50000$

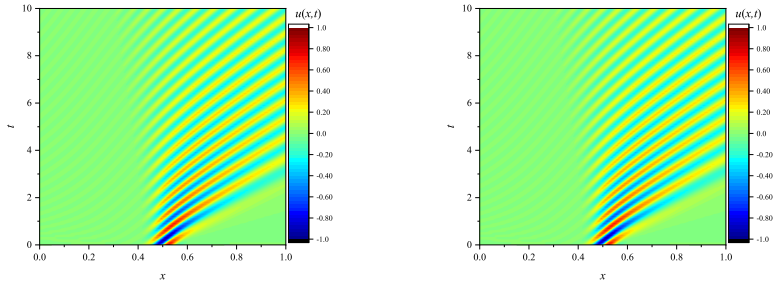

Figure 3: Solution to Example 2: (left) by proposed method; (right) by using a large spatial domain

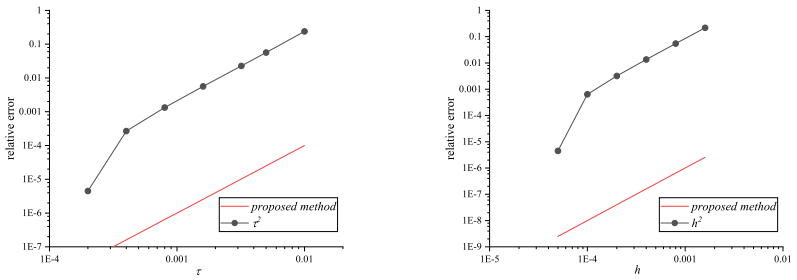

Figure 4: Performance of the proposed method in Example 2: (left) relative error versus time step size τ ($h = 5 \times 10^{-5}$); (right) relative error versus spatial mesh size h ($\tau = 2 \times 10^{-4}$)

with fixed m and $h = 5 \times 10^{-5}$. Figure 5 shows the CPU times for the fast convolution. One can observe the slope of 1.

7.3. Example 3: Using a Small σ

In previous section we proved the convergence of the proposed method under the condition $\sigma > 1/\sqrt{\kappa_1}$. However, in this example we show it is not a necessary condition. We employ the same initial condition and computing domain as in Example 1, but a different parameter set $c = 2$, $\kappa = 0.0001$. Then the previous condition leads to $\sigma \geq 100$, which is not favorable because a large σ will pollute the numbers when doing the transform $u = v \exp(\sigma t)$. In fact, by using a much smaller value, i.e., σ=0.01, we will see the proposed method still applies, even though right now we cannot prove its convergence theoretically in this condition.

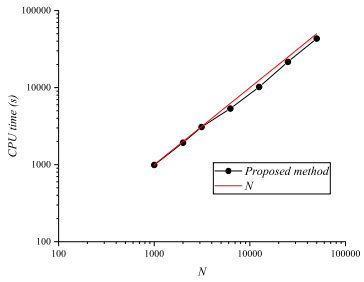

Figure 5: (Example 2) $Ln - Ln$ plot for the CPU time by fixing m with different N.

When mesh sizes of $h = 5 \times 10^{-5}, \tau = 5 \times 10^{-5}$ are employed, the displacement solutions by the proposed method and the reference method are shown in Figure 6. The two agree with each other well. In both solutions the main wave has reached the boundary at about $t = 0.25$, but no artificial reflection or numerical instability is noticed.

The error evolution with respect to different h and τ are shown in Figure 7. Second-order convergence can be observed from the log-log plots. The result suggests the proposed method works in a much wider condition than the assumption of Theorem 1, although the strict proof is currently beyond the scope of with work.

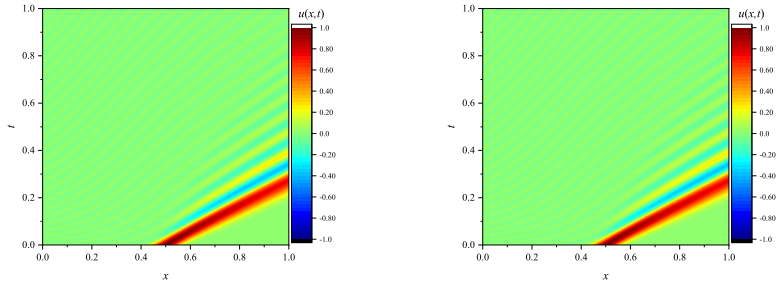

Figure 6: Solution to Example 3: (left) by proposed method; (right) by using a large spatial domain

8. Conclusion

A convergent fast numerical method for solving the Cauchy problem of the one-dimensional linearized Benjamin-Bona-Mahony (BBM) equation

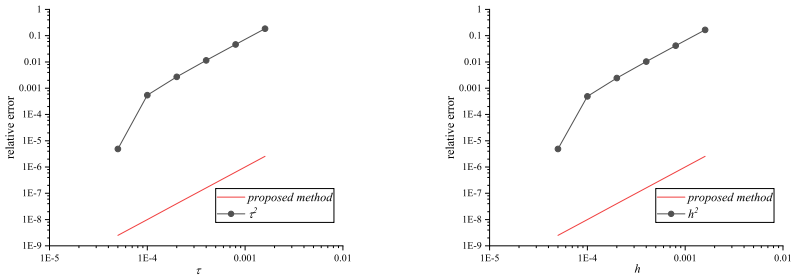

Figure 7: Performance of the proposed method in Example 3: (left) relative error versus time step size τ ($h = 5 \times 10^{-5}$); (right) relative error versus spatial mesh size h ($\tau = 5 \times 10^{-5}$)

is proposed to reduce the computational cost incurred by the exact convolution. To this end, the BBM equation in an unbounded domain was reformulated into an initial boundary value problem in a bounded domain of computational interest. A fully discrete Crank-Nicolson finite difference method has been proposed to solve the reformulated initial boundary value problem with an exact semi-discrete artificial boundary condition (ABC). A fast convolution algorithm is introduced to handle the convolutions for the exact semi-discrete ABC using the Padé rational expansion. A criterion for determining the damping term was proposed to guarantee convergence. In this case, it was theoretically proved that the corresponding numerical scheme can achieve second order accuracy. Numerical tests confirmed the effectiveness of the proposed numerical method.

The problem that remains to be solved is that the damping term $e^{-\sigma t}$ should satisfy the stability condition $\sigma \geq 1/\sqrt{\kappa_1}$. For small dispersion parameters κ, a too fast decaying damping term $e^{-\sigma t}$ leads to numerical errors. We will deal with this problem in a forthcoming paper.

Finally, we also have to deal with nonlinear systems, which is still an open problem. We would like to address these problems in the near future.

Acknowledgments

This research is partially supported by NSFC under grant Nos. 11502028 and Nos. 12102282.

References

[1] T. Achouri, N. Khiari, and K. Omrani, *On the convergence of difference schemes for the Benjamin Bona Mahony (BBM) equation*, Appl. Math. Comput. **182** (2006) 999–1005.

[2] A. Alazman, J. Albert, J. Bona, M. Chen, and J. Wu, *Comparisons between the BBM equation and a Boussinesq system*, Adv. Differ. Eq. **11** (2006) 121–166.

[3] C. Besse, B. Gireau, and P. Noble, *Artificial boundary conditions for the linearized Benjamin-Bona-Mahony equation*, Numer. Math. **139** (2018) 281–314.

[4] M. Kazakova and P. Nobel, *Discrete transparent boundary conditions for the linearized Green-Naghdi system of equations*, SIAM J. Numer. Anal. **58(1)** (2020) 657–683.

[5] C. Zheng, Q. Du, X. Ma, and J. Zhang, *Stability and error analysis for a second-order fast approximation of the local and nonlocal diffusion equations on the real Line*, SIAM J. Numer. Anal. **58(3)** (2020) 1893–1917.

[6] T. Fevens and H. Jiang, *Absorbing Boundary Conditions for the Schrödinger Equation*, SIAM J. Sci. Comput. **21** (1999) 255–282.

[7] X. Antoine, C. Besse and P. Klein, *Absorbing boundary conditions for the one-dimensional Schrödinger equation with an exterior repulsive potential*, J. Comput. Phys. **228** (2009) 312–335.

[8] X. Antoine and C. Besse, *Unconditionally stable discretization schemes of non-reflecting boundary conditions for the one-dimensional Schrödinger equation*, J. Comput. Phys. **188** (2003) 157–175.

[9] X. Wu and Z. Sun, *Convergence of difference scheme for heat equation in unbounded domains using artificial boundary conditions*, Appl. Num. Math. **50** (2004) 261–277.

[10] V. Baskakov and A. Popov, *Implementation of transparent boundaries for numerical solution of the Schrödinger equation*, Wave Motion **14** (1991) 123–128.

[11] H. Han and Z. Huang, *Exact and approximating boundary conditions for the parabolic problems on unbounded domains*, Comput. Math. Appl. **44** (2002) 655–666.

[12] H. Han and Z. Huang, *Exact artificial boundary conditions for the Schrödinger equation in* $\mathbb{R}^2$, Commun. Math. Sci. **2** (2004) 79–94.

[13] A. Arnold, M. Ehrhardt, M. Schulte, and I. Sofronov, *Discrete transparent boundary conditions for the Schrödinger equation on circular domains*, Commun. Math. Sci. **10** (2012) 889–916.

[14] H. Li, X. Wu, and J. Zhang, *Local artificial boundary conditions for Schrödinger and heat equations by using high-order azimuth derivatives on circular artificial boundary*, Comput. Phys. Commun. **185** (2014) 1606–1615.

[15] X. Antoine, C. Besse, and V. Mouysset, *Numerical schemes for the simulation of the two-dimensional Schrödinger equation using non-reflecting boundary conditions*, Math. Comput. **73** (2004) 1779–1799.

[16] X. Antoine, C. Besse, and P. Klein, *Absorbing Boundary Conditions for the Two-Dimensional Schrödinger Equation with an Exterior Potential. Part II: Discretization and Numerical Results*, Numer. Math. **125** (2013) 191–223.

[17] X. Antoine, C. Besse, and P. Klein, *Absorbing Boundary Conditions for General Nonlinear Schrödinger Equations*, SIAM J. Sci. Comput. **33** (2011) 1008–1033.

[18] G. Pang, Y. Yang, X. Antoine, and S. Tang, *Stability and convergence analysis of artificial boundary conditions for the Schrödinger equation on a rectangular domain*, Math. Comput. **90(332)** (2021), 2731–2756.

[19] X. Antoine, A. Arnold, C. Besse, M. Ehrhardt, and A. Schädle, *A Review of Transparent and Artificial Boundary Conditions Techniques for Linear and Nonlinear Schrödinger Equations*, Commun. Comput. Phys. **4** (2008) 729–796.

[20] G. Pang and S. Tang, *Approximate linear relations for Bessel functions*, Commun. Math. Sci. **15** (2017) 1967–1986.

[21] G. Pang, L. Bian, and S. Tang, *ALmost EXact boundary condition for one-dimensional Schrödinger equation*, Phys. Rev. E **86** (2012) 066709.

[22] S. Ji, Y. Yang, G. Pang, and X. Antoine, *Accurate artificial boundary conditions for the semi-discretized linear Schrödinger and heat equations on rectangular domains*, Comput. Phys. Commun. **222** (2018) 84–93.

[23] A. Arnold, M. Ehrhardt, and I. Sofronov, *Discrete transparent boundary conditions for the Schrödinger equation: fast calculation, approximation, and stability*, Commun. Math. Sci. **1** (2003) 501–556.

[24] S. Jiang and L. Greengard, *Fast evaluation of nonreflecting boundary conditions for the Schrödinger equation in one dimension*, Comput. Math. Appl. **47** (2004) 955–966.

[25] C. Zheng, *Approximation, stability and fast evaluation of exact artificial boundary condition for one-dimensional heat equation*, J. Comput. Math. **25** (2007) 730–745.

[26] B. Alpert, L. Greengard, and T. Hagstrom, *Rapid evaluation of nonreflecting boundary kernels for time-domain wave propagation*, SIAM J. Numer. Anal. **37** (2000) 1138–1164.

[27] Y. Lu, *A Padé approximation method for square roots of symmetric positive definite matrices*, SIAM J. Numer. Anal. **19** (1998) 833–845.

[28] C. Lubich and A. Schädle, *Fast convolution for nonreflecting boundary conditions*, SIAM J. Sci. Compt. **24** (2002) 161–182.

[29] Y. Feng and X. Wang, *Matching boundary conditions for Euler-Bernoulli beam*, Shock Vibr. (2021) 6685852.

[30] S. Tang and E. Karpov, *Artificial boundary conditions for Euler-Bernoulli beam equation*, Acta. Mech. Sinica-PRC **30** (2014) 687–692.

[31] B. Li, J. Zhang, and C. Zheng, *An efficient second-order finite difference method for the one-dimensional Schrödinger equation with absorbing boundary conditions*, SIAM J. Numer. Anal. **56** (2018) 766–791.