Convergence Analysis of Schwarz Waveform Relaxation for Nonlocal Diffusion Problems

Ke Li,1,2 Dali Guo,1 and Yunxiang Zhao1

1School of Sciences, Southwest Petroleum University, Chengdu, Sichuan 610500, China
2School of Mathematics and Statistics, Sichuan University of Science and Engineering, Zigong, Sichuan 643000, China

Correspondence should be addressed to Dali Guo; guodali@swpu.edu.cn

Received 26 February 2021; Revised 17 May 2021; Accepted 1 June 2021; Published 22 July 2021

Diffusion equations with Riemann–Liouville fractional derivatives are Volterra integro-partial differential equations with weakly singular kernels and present fundamental challenges for numerical computation. In this paper, we make a convergence analysis of the Schwarz waveform relaxation (SWR) algorithms with Robin transmission conditions (TCs) for these problems. We focus on deriving good choice of the parameter involved in the Robin TCs, at the continuous and fully discretized level. Particularly, at the space-time continuous level, we show that the derived Robin parameter is much better than the one predicted by the well-understood equioscillation principle. At the fully discretized level, the problem of determining a good Robin parameter is studied in the convolution quadrature framework, which permits us to precisely capture the effects of different temporal discretization methods on the convergence rate of the SWR algorithms. The results obtained in this paper will be preliminary preparations for our further study of the SWR algorithms for integro-partial differential equations.

1. Introduction

Fractional kinetic equations have been proved particularly useful to model the anomalous slow diffusion (subdiffusion) [1]. Subdiffusive motion is characterized by an asymptotically long-time behavior of the mean square displacement, such as

$$\langle x^2(t) \rangle \sim \frac{2}{\Gamma(1+\gamma)} t^\gamma, \quad t \to +\infty,$$  \hspace{1cm} (1)

where $\gamma$ is the anomalous diffusion exponent. The process is usually referred to as subdiffusive when $\gamma \in (0, 1)$. For the ordinary (or Brownian) diffusion (this corresponds to $\gamma = 1$), from a continuous point of view the diffusion process is described by the well-known diffusion equation $u_t(x, t) = u_{xx}(x, t)$, where $u(x, t)$ represents the probability density of finding a “particle” at space $x$ and at time $t$. It turns out that the probability density function $u(x, t)$ that describes anomalous (sub) diffusive particles follows the fractional diffusion equation [1–4]:

$$\frac{\partial u(x, t)}{\partial t} = (D_t^{1-\gamma}) \frac{\partial^2 u(x, t)}{\partial x^2} + f, \hspace{1cm} (2)$$

where $D_t^{1-\gamma}$ is the Riemann–Liouville fractional derivative defined by (6) below.

Subdiffusive equation (2) is the fundamental model for other anomalous diffusion observed in many fields, such as the transport of fluid in porous media [5], diffusion on liquid surfaces [6], and turbulent flow [7]. Other applications of fractional differential equations can be found in [8–17]. Very recently, the so-called fractional cable equation was introduced for modeling the anomalous diffusion in spiny neuronal dendrites [18–20]:

$$C_m \frac{\partial u(x, t)}{\partial t} = \mathcal{D}_x^\alpha \frac{d}{4R_A} \frac{\partial^2 u(x, t)}{\partial x^2} - \mathcal{D}_t^\beta \left[ \frac{u(x, t) + E}{R_m} \right], \hspace{1cm} (3)$$

where $\alpha, \beta \in (0, 1)$. This equation models spatial-temporal dynamics of the membrane potential along the axial direction of an approximately cylindrical segment of a nerve.
cell, which takes into account anomalous transport in an external field. In (3), \( d \) denotes the diameter of the cable, \( E \) is the equilibrium potential of the membrane, \( R_m \) is a membrane constant and \( C_m \) is a constant independent of the membrane area, and \( R_A \) is an axial material constant, associated to the given cable. By taking \( \tau_m = R_mC_m, \nu = (dR_m/4R_A) \) and then letting \( f = -\mathcal{D}_t^\nu (E/R_m), \bar{T} = (t/\tau_m), \bar{x} = \sqrt{x/\tau_m} \), cable equation (3) can be transformed to \( \bar{\partial}_t \mathcal{D}_x^\nu \bar{u} - \mu \mathcal{D}_x^\alpha \bar{u} + f \) with \( \mu = \tau_m^{-1} \). Dropping the tilde superscript, we finally get
\[
\partial_t u = \mathcal{D}_t^\nu \partial_x^\nu u - \mu \mathcal{D}_x^\alpha u + f. \tag{4}
\]

Unlike the integer order PDEs, a simple and explicit analytical solution to a fractional PDE is usually unavailable and therefore numerical methods play an important role in investigating the properties of these equations. Moreover, compared to the integer order problems, numerical computation for the fractional PDEs is much more time consuming, because we need to treat convolutions with singular kernels at every time grid. Here, our interest is to apply the Schwarz waveform relaxation (SWR) algorithms to solve the fractional diffusion equations. These algorithms belong to the widely used domain decomposition methods, but with completely new implementation strategy for time-dependent problems: the classical strategy employs the alternating or parallel Schwarz methods to the elliptic PDEs which result from semidiscretizing the time-dependent PDEs in time by some implicit time-integrator [21,22], but in the SWR framework one directly decomposes the space domain into many subdomains and then solve each subproblem simultaneously or alternately. This provides flexibility for treating different subproblems numerically differently with an adapted procedure for each subdomain (see [23,24] for the original idea of the SWR algorithms). There are also other efficient algorithms for handling the fractional PDEs, such as the interesting parallel-in-time algorithms studied in [25–29].

For integer order differential equations, the SWR algorithms have been investigated widely and deeply by Gander and his colleagues [30–35]. However, much less results are known about the algorithms applied to fractional order PDEs; actually, according to our best knowledge, this is the first time that the SWR algorithms are analyzed for fractional problems. In this paper, we analyze the algorithms with Robin transmission conditions (TCs) for the aforementioned fractional cable equation, which, as we will show in Section 2, represents a large class of fractional problems that received considerable attention in the literature. The Robin TCs contain a free parameter, which has a significant influence on the convergence rate of the SWR algorithms; hence, finding a good choice of the Robin parameter is one of the top-priority matters. For the integer order PDEs, optimizing the Robin parameter is deeply studied in [30,35–37] and the main finding is that a good parameter can be computed by solving one or two simple nonlinear equations. This is the so-called equioscillation principle, and nowadays, it is well-understood and widely used in the study of the SWR algorithms.

However, as we will show in Section 3, for the time-fractional diffusion problems, the Robin parameter predicted by the equioscillation principle is unsatisfactory and there exists a much better choice for specified problem/discretization parameters. To make the SWR algorithms practical in real computation, we also perform a convergence analysis at the fully discretized level in Section 4, in the framework of convolution quadratures [38–40]. This framework allows for a convenient treatment of the convergence analysis based on discrete Laplace transform. We validate the theoretical conclusions by numerical results in Section 5 and we finish this paper with some concluding remarks in Section 6.

2. The Model Problem and the SWR Algorithms

Motivated by the fractional cable equation, which includes fundamental equation (2), our model studied throughout this paper is the following initial value problem:
\[
\begin{aligned}
\partial_t u - \mathcal{D}_x^\nu \partial_x^\nu u + \mu \mathcal{D}_x^\alpha u &= f(x,t), \quad (x,t) \in \mathbb{R} \times (0,T), \\
\quad u(x,0) &= u_0, \quad x \in \mathbb{R},
\end{aligned}
\]
where \( \mu \geq 0 \) and \( \alpha, \beta \in (0,1) \). This model gets considerable attention in recent years and we can refer to [19,20] for the construction of solutions based on Green’s functions. For any \( \gamma \in (0,1) \), we denote by \( \mathcal{D}_t^\gamma \) the Riemann–Liouville fractional derivative:
\[
\mathcal{D}_t^\gamma v(t) = \frac{1}{\Gamma(1-\gamma)} \frac{d}{dt} \int_0^t \frac{1}{(t-\tau)^\gamma} v(\tau) d\tau, \tag{6}
\]
where \( \Gamma \) is the gamma function \( \Gamma(\tau) = \int_0^\infty e^{-t} t^{-\tau-1} dt \). From [41] (Chapter 2), for any \( v \in C^1(0,T) \) it holds that \( \lim_{\gamma \rightarrow 1^{-1}} \mathcal{D}_t^\gamma v = \nu(t) \) (\( \forall t \in (0,T) \)). Hence, the Riemann–Liouville fractional derivative builds a “bridge” from \( v(t) \) to \( \nu(t) \).

In the field of fractional PDEs, there exists another definition of the fractional derivative, the Caputo derivative:
\[
\mathcal{C}_t^\gamma \nu(t) = \frac{1}{\Gamma(1-\gamma)} \int_0^t \frac{d}{dt} \nu(\tau) (t-\tau)^{-\gamma} d\tau. \tag{7}
\]

The two definitions are linked by the following relation [41], (pp. 62–77):
\[
\mathcal{D}_t^\gamma v(t) = \frac{1}{\Gamma(1-\gamma)} \int_0^t \frac{\nu'(\tau)}{(t-\tau)^\gamma} d\tau + \frac{\nu(0) t^{-\gamma}}{\Gamma(1-\gamma)} \quad \forall \gamma \in (0,1),
\]

\[
\text{Caputo derivative, } \mathcal{C}_t^\gamma \nu(t) \tag{8}
\]

Thanks to this connection, we shall only consider the Riemann–Liouville derivatives, since the results can be directly generalized to problems with Caputo derivatives.

There exists another class of fractional PDEs, which are also widely studied in the literature [41,42], the fractional heat equations:
\[
\mathcal{D}_t^\alpha u = \mathcal{D}_x^\nu \partial_x^\nu u + f. \tag{9a}
\]
From [41] (pp. 65–75, Chapter 2), we know that (9a) can be rewritten as
\[ \partial_t u = \mathcal{D}^\alpha u + f, \quad f \text{ depends on } f \text{ and } u_0. \]  

(9b)

Hence, the analysis for the model problem (5) can be also applied to (9a).

We now introduce the Schwarz waveform relaxation algorithms for (5). We divide the spatial domain into two subdomains, \( \Omega = \Omega_1 \cup \Omega_2, \) \( \Omega_1 = (-\infty, L) \) and \( \Omega_2 = [0, +\infty). \) Here, \( L > 0 \) denotes the size of overlap. Then, the SWR algorithms for (5) are
\[
\begin{aligned}
\partial_t u_i^k - \mathcal{D}^\alpha \partial_x u_i^k + \mu \mathcal{W}^\beta u_i^k = f(x, t), & \quad x \in \Omega_i, \\
(\partial_x + (1-j)^i!) u_i^k((2-j)L, t) = (\partial_x + (1-j)^i!) u_{i-1}^k((2-j)L, t), & \quad \text{for } i = 1, 2 \\
u_0^k(x, 0) = u_0, & \quad \text{for } j = 1, \ldots, p,
\end{aligned}
\]

where \( p \in \mathbb{R} \) is a free parameter, \( k \geq 1 \) is iteration index, and \( j = 1, 2. \) It is clear that the proposed SWR algorithm does not directly depend on the orders of the differential equations. So, the algorithm is applicable to variable-order time-fractional PDEs. However, the convergence analysis will be different.

3. Convergence Analysis at the Continuous Level

We are now in a position to analyze SWR algorithms (10). In this section, we perform an analysis at the continuous level, and in the next section, we analyze the algorithm at the fully discretized level. For any real-valued function \( v(t): \mathbb{R}^+ \rightarrow \mathbb{R}, \) we denote by \( \mathcal{L}(v) \) the Laplace transform of \( v, \) defined by
\[
\mathcal{L}(v)(s) = \int_0^{\infty} e^{-st} v(t) dt, \quad s = \sigma + iw, \ \sigma > 0.
\]

(11)

Let \( \epsilon_j^k(x, t) = u_j^k(x, t) - u(x, t)|_{x \in \Omega_j}. \) Then, it is easy to know that \( \epsilon_j^k \) satisfies
\[
\begin{aligned}
\partial_t \epsilon_j^k - \mathcal{D}^\alpha \partial_x \epsilon_j^k + \mu \mathcal{W}^\beta \epsilon_j^k = 0, \\
(\partial_x + (1-j)^i!) \epsilon_j^k((2-j)L, t) = (\partial_x + (1-j)^i!) \epsilon_{j-1}^k((2-j)L, t), \\
\epsilon_j^k(x, 0) = 0.
\end{aligned}
\]

(12)

Applying Laplace transform to the error equation (12), we get
\[
\begin{aligned}
\mathcal{L}(\partial_t \epsilon_j^k) - \lambda(\lambda - 1) \mathcal{L}(\epsilon_j^k) + \mu \mathcal{L}(\mathcal{W}^\beta \epsilon_j^k) = 0, \\
(\partial_x + (1-j)^i!) \mathcal{L}(\epsilon_j^k((2-j)L, t)) = (\partial_x + (1-j)^i!) \mathcal{L}(\epsilon_{j-1}^k((2-j)L, t)).
\end{aligned}
\]

(13)

Proof. Let \( s = re^{i\varphi} \) with \( r > 0 \) and \( \varphi \in (-\pi/2, \pi/2). \) Then, we have
\[
\mathfrak{R}(s^{1-\beta} + \mu s^{\alpha-\beta}) = r^{1-\beta} \cos((1-\beta)\varphi) + \mu r^{\alpha-\beta} \cos((\alpha - \beta)\varphi).
\]

(18)

This implies that, for \( \mu \geq 0 \) and \( \alpha, \beta \in (0, 1) \) the function \( \lambda(s) \) defined in (13) is analytic in the right half of the complex plane, since the quantity under the square root avoids the negative real axis. For \( p > 0, \) the denominator \( \lambda(s) + p \) does not vanish in the right half of the complex plane; hence, \( (p - \lambda(s))/p + \lambda(s) \) is an analytic function. Since the product of two analytic functions is still analytic, \( \tilde{\rho}_c \) is analytic. \( \square \)

Note that, for any \( s = re^{i\varphi} \) with \( \varphi \in (-\pi/2, \pi/2), \) the limit of \( \tilde{\rho}_c \) as \( r \to +\infty \) is zero. Hence, by using the maximum principle for complex analytic functions, we know that the maximum of \( \tilde{\rho}_c(p, L, s) \) is obtained on the imaginary axis, i.e., \( s = iw. \) Moreover, in a numerical computation, the angular frequency \( \omega \) cannot be arbitrarily large or small; there is a range of frequencies, which can be represented on a time grid with size \( \Delta t \) and a length of time window \( (0, T), \) as \( |\omega| \in [(\pi/T, \pi/(\Delta t)]. \) Hence, it is reasonable to define \( \tilde{\rho}_c(p, L) \), which we call convergence factor of the SWR algorithm at the space-time continuous level, as

\[ \text{Lemma 1. Assume } \alpha, \beta \in (0, 1), \mu \geq 0 \text{ and } p > 0. \text{ Then, the argument } \tilde{\rho}_c \text{ defined by (16) is analytic in the right half of the complex plane.} \]
\[
\rho_c (p, L) = \max_{|\omega| \in [(\pi T), (\pi / \Delta t)]} \left| \tilde{\rho}_c (p, L, i \omega) \right|^2.
\]

The subscript “c” appearing in \( \tilde{\rho}_c \) and \( \rho_c \) denotes “continuous.” Clearly, an ideal \( p \) is the minimizer of the convergence factor \( \rho_c (p, L) \), i.e.,

\[
\frac{i (\pm \omega) + \mu (i (\pm \omega))^\beta}{(i (\pm \omega))^\beta} = \frac{(\pm i)^{1-\beta} \omega + \mu (\pm i)^{\alpha-\beta} \omega^\alpha}{\omega^\beta} = \omega \left[ \cos (\pm 1 - \beta / 2\pi) + i \sin (\pm 1 - \beta / 2\pi) \right] + \mu \omega^\alpha \left[ \cos (\pm \alpha - \beta / 2\pi) + i \sin (\pm \alpha - \beta / 2\pi) \right] \omega^\beta
\]

\[
= \omega^{1-\beta} \cos \left( \frac{1-\beta}{2} \pi \right) + \mu \omega^{\alpha-\beta} \cos \left( \frac{\alpha-\beta}{2} \pi \right) + i \left[ \omega^{1-\beta} \sin \left( \frac{1-\beta}{2} \pi \right) + \mu \omega^{\alpha-\beta} \sin \left( \frac{\alpha-\beta}{2} \pi \right) \right].
\]

Thus, for any \( \omega > 0 \), we can rewrite \( \lambda (\pm \omega) = \lambda_R (\omega) \pm i \lambda_I (\omega) \), where

\[
\lambda_R (\omega) = \sqrt{\frac{\lambda_R^2 (\omega) + \lambda_I^2 (\omega)}{2}},
\]

\[
\lambda_I (\omega) = \sqrt{\frac{\lambda_R^2 (\omega) + \lambda_I^2 (\omega) - \lambda_R (\omega)}{2}},
\]

\[
\chi_R (\omega) = \omega^{1-\beta} \cos \left( \frac{1-\beta}{2} \pi \right) + \mu \omega^{\alpha-\beta} \cos \left( \frac{\alpha-\beta}{2} \pi \right),
\]

\[
\chi_I (\omega) = \omega^{1-\beta} \sin \left( \frac{1-\beta}{2} \pi \right) + \mu \omega^{\alpha-\beta} \sin \left( \frac{\alpha-\beta}{2} \pi \right).
\]

This gives

\[
[\tilde{\rho}_c (p, L, \pm i \omega)]^2 = \left( \frac{\lambda_R (\omega) - p}{\lambda_R (\omega) + p} \right)^2 + \frac{\lambda_I^2 (\omega) - 2 \lambda_R (\omega)}{\lambda_R^2 (\omega) + \lambda_I^2 (\omega)} \chi_R (\omega), \quad \forall \omega \in \left[ \frac{\pi}{T}, \frac{\pi}{\Delta t} \right].
\]

Clearly, for any \( p < 0 \), it holds \( |\tilde{\rho}_c| > 1 \). Thus, min-max problem (20) is equivalent to

\[
\min_{p \geq 0} \max_{\omega \in [(\pi T), (\pi / \Delta t)]} \tilde{\mathcal{R}} (\omega, q),
\]

where

\[
q = 2 L p, \quad \tilde{\mathcal{R}} (\omega, q) = \frac{(\lambda_R (\omega) - q)^2 + \lambda_I^2 (\omega)}{(\lambda_R (\omega) + q)^2 + \lambda_I^2 (\omega)} \chi_R (\omega).
\]

\[
= 2 L \lambda_{R,I} (\omega).
\]

3.1. Towards a Good Choice of \( q (= 2L p) \): The Equioscillation Principle. A well-understood formula for choosing the free parameter \( p \) is based on the so-called equioscillation principle, established recently by Gander and Halpern [30,35]:

\[
\tilde{\mathcal{R}} (\omega_{\text{min}}, q) = \tilde{\mathcal{R}} (\pi (q), q),
\]

where \( \omega_{\text{min}} = (\pi / T) \) and \( \pi (q) \) denotes the local maximizer of \( \tilde{\mathcal{R}} \). The solution of (25), namely, \( q_{\text{eqiu}} \), can be obtained by solving the following two nonlinear equations:

\[
\left \{ \begin{array}{l}
\tilde{\mathcal{R}} (\omega_{\text{min}}, q) = \tilde{\mathcal{R}} (\pi (q), q), \\
\partial_q \tilde{\mathcal{R}} (\omega, q) = 0,
\end{array} \right.
\]

(26)

together with a constraint \( \tilde{\mathcal{R}}^2 (\omega, q) < 0 \) (equations (26) may have several solutions and only the one satisfying this constraint corresponds to the solution of (25)). Solving such a nonlinear problem is quite easy from a numerical point of view. For integer order problems, such as the advection-reaction-diffusion equations, formula (25) gives the best solution of (25), namely, \( q_{\text{eqiu}} \), can be obtained by solving the following two nonlinear equations:

\[
\left \{ \begin{array}{l}
\tilde{\mathcal{R}} (\omega_{\text{min}}, q) = \tilde{\mathcal{R}} (\pi (q), q), \\
\partial_q \tilde{\mathcal{R}} (\omega, q) = 0,
\end{array} \right.
\]

For time-fractional PDEs, problems exist for the equioscillation principle in twofold:

1. It is unclear for the moment whether equation (25) is well defined or not. Is there a unique positive solution? Is the solution located in the relevant interval \([(\pi T), (\pi / \Delta t)]\)? For integer order time-dependent PDEs, the answers for these questions are positive, at least in the asymptotic sense, i.e., \( L = O (\Delta x) \), \( \Delta t = O (\Delta x^2) \), and \( \Delta x \) small; see [30,35]. However, in the fractional situation, we know nothing about (25).

2. More importantly, the parameter predicted by the equioscillation principle is unsatisfactory. For example, with problem parameters listed in the title of
Figure 1: (a) An illustration of equioscillation, i.e., the parameter \( q = q_{\text{equi}} \) is obtained by solving (26). The plot of \( \mathcal{R}(\omega, q) \) in the relevant frequency domain \( \omega \in [(\pi/T), (\pi/\Delta t)] \) for two choices of \( q \), (b) \( q = q_{\text{equi}} \) and (c) \( q = q_{\text{opt}} \) given by Lemma 3.

Since the applicability of the equioscillation principle to time-fractional problems is unclear, we introduce in Lemma 4 a new formula for computing \( q \). In Figure 1(b), with \( \Delta t = 0.001, T = 50, L = 0.02, \mu = 1.3, \) and two groups of \( \alpha \) and \( \beta \), we plot the function \( \mathcal{R}(\omega, q) \) for \( \omega \in [(\pi/T), (\pi/\Delta t)] \), where we can see that compared to \( q_{\text{equi}} \) the new parameter \( q_{\text{opt}} \) leads to much smaller convergence factor and that with the new parameter the function \( \mathcal{R}(\omega, q) \) does not equioscillate.

3.2. Towards a Good Choice of \( q (= 2L \rho) \): A New Formula. Because of the high complexity of the arguments \( \lambda_{1,R}(\omega) \), it is difficult to present a closed formula for the solution of minmax problem (24a) and (24b). In what follows, we focus on deriving a reliable approximate solution. We need two auxiliary lemmas.

**Lemma 2.** Let \( \alpha, \beta \in (0, 1), \mu \geq 0 \) and

\[
S(x) = \frac{\sqrt{x^2 + 2\mu x \cos((1 - \alpha)/2)\pi + \mu^2} - [x \cos((1 - \beta)/2)\pi + \mu \cos((\alpha - \beta)/2)\pi]}{\sqrt{x^2 + 2\mu x \cos((1 - \alpha)/2)\pi} + \mu^2 + [x \cos((1 - \beta)/2)\pi + \mu \cos((\alpha - \beta)/2)\pi]}
\] (27)
Then, it holds that \( \max_{x \in [x_0,x_1]} S(x) = \max\{S(x_0), S(x_1)\} \), where \( 0 < x_0 < x_1 \).

**Proof.** Clearly, we only need to consider the case \( \mu > 0 \), since for \( \mu = 0 \) the function \( S(x) \) is a constant and therefore \( \max_{x \in [x_0,x_1]} S(x) = \max\{S(x_0), S(x_1)\} \).

Let \( R_0(x) = x^2 + 2\mu x \cos((1 - \alpha/2)\pi) + \mu^2 \) and \( R_1(x) = x \cos((1 - \beta/2)\pi) + \mu \cos((\alpha - \beta/2)\pi) \). We then have \( S(x) = (\sqrt{R_0(x)} - R_1(x))/\sqrt{R_0(x)} + R_1(x) \) and \( S'(x) = (R_1(x)R_0''(x) - 2R_0(x)R_1''(x))x^2 \). Suppose \( S'(x) \) has a positive root located at \( x = x^* \), i.e., \( R_1(x^*)R_0''(x^*) = 2R_0(x^*)R_1''(x^*) \). Then, we have

\[
S''(x) = \frac{R_1(x)R_0''(x) - 2R_0(x)R_1''(x)}{\sqrt{R_0(x)}[\sqrt{R_0(x)} + R_1(x)]^2}.
\]

For any \( x \), it holds that

\[
R_0''(x) = 2x + 2\mu \cos\left(\frac{1 - \alpha}{2}\pi\right),
\]

\[
R_1''(x) = \cos\left(\frac{1 - \beta}{2}\pi\right),
\]

\[
R_0''(x) = 2,
\]

\[
R_1''(x) = 0,
\]

and this, together with \( \mu > 0 \) and \( \alpha, \beta \in (0, 1) \), gives

\[
S''(x^*) = 2\mu \cos((1 - \beta/2)\pi) - \cos((1 - \beta/2)\pi)\cos((1 - \alpha/2)\pi) \frac{\sqrt{R_0(x^*)} + R_1(x^*)}{\left[\sqrt{R_0(x^*)} + R_1(x^*)\right]^2} > 0,
\]

\[
\text{Hence, if } x^* \text{ corresponds to a local extremum of } S(x), \text{ it must be a minimizer.}
\]

**Lemma 3.** Let \( \mu \geq 0 \), \( \alpha, \beta \in (0, 1) \), and \( \phi = \max\{S((\pi/T)^{-1} - \omega), S((\pi/\Delta t)^{-1} - \omega)\} \) with \( S \) being the function defined in Lemma 2. Then, for all \( \omega \in [(\pi/T), (\pi/\Delta t)] \), we have \( \lambda_1^2(\omega) \leq \phi \lambda_R^2(\omega) \).

**Proof.** From (22), \( \chi(\omega) = (\omega - \alpha) \sin((1 - \beta/2)\pi) + \mu \sin((\alpha - \beta/2)\pi) \omega^{-2} \cos((1 - \beta/2)\pi) + \mu \cos((\alpha - \beta/2)\pi) \). Since \( \alpha, \beta \in (0, 1) \),

\[
\cos\left(\frac{1 - \beta}{2}\pi\right) \sin\left(\frac{\alpha - \beta}{2}\pi\right) \leq \sin\left(\frac{1 - \beta}{2}\pi\right) \cos\left(\frac{\alpha - \beta}{2}\pi\right).
\]

\[
\text{Hence,}
\]

\[
\phi(\omega) = (\alpha - \beta) \omega^{-2} \cos((1 - \beta/2)\pi) + \mu \cos((\alpha - \beta/2)\pi)
\]

is a decreasing function of \( \omega^{-1} \); this gives

\[
\frac{\chi(\omega)}{\lambda_R(\omega)} \leq \frac{\chi(\omega)}{\lambda_R(\omega)} \leq \frac{\chi(\omega)}{\lambda_R(\omega)} \forall \omega \in \left[\frac{\pi}{\pi}, \frac{\pi}{\Delta t}\right].
\]

By noticing \( \lambda_1^2(\omega) + \chi^2(\omega) = \omega^2(1 - \beta) + \mu^2 \omega^2(\alpha - \beta) + 2\mu \alpha^1 \omega^{-2b} \cos((1 - \alpha/2)\pi) \), we have

\[
\frac{\lambda_1^2(\omega)}{\lambda_R^2(\omega)} \leq \frac{\chi^2(\omega) + \chi_x^2(\omega) - \chi_R(\omega)}{\chi^2(\omega) + \chi_x^2(\omega) + \chi_R(\omega)} = S(\omega^{-1}).
\]

Hence, from Lemma 2, we get

\[
(\lambda_1^2(\omega) - \lambda_R^2(\omega)) \leq \max\{S((\pi/T)^{-1} - \omega), S((\pi/\Delta t)^{-1} - \omega)\} = \phi
\]

and this completes the proof.

Now, for min-max problem (24a) and (24b), by using Lemma 3, we have

\[
\left(\frac{\lambda_R(\omega) - p}{\lambda_R(\omega) + p}\right)^2 + \phi \lambda_R^2(\omega) e^{-2L\lambda_R(\omega)} \leq \left(\frac{\lambda_R(\omega) - p}{\lambda_R(\omega) + p}\right)^2 + \phi \lambda_R^2(\omega) e^{-2L\lambda_R(\omega)}
\]

for any \( p > 0 \), this leads to the following inequality:

\[
\max_{\omega \in [(\pi/\piT), (\pi/\Delta t)]} \left(\frac{\lambda_R(\omega) - p}{\lambda_R(\omega) + p}\right)^2 + \phi \lambda_R^2(\omega) e^{-2L\lambda_R(\omega)} \leq \left(\frac{\lambda_R(\omega) - p}{\lambda_R(\omega) + p}\right)^2 + \phi \lambda_R^2(\omega) e^{-2L\lambda_R(\omega)}
\]

where

\[
y_0 = 2L \min_{\omega \in [(\pi/\piT), (\pi/\Delta t)]} \lambda_R(\omega),
\]

\[
y_1 = 2L \max_{\omega \in [(\pi/\piT), (\pi/\Delta t)]} \lambda_R(\omega).
\]

We present the solution of (36a) and (36b) in a very general setting: we only assume \( \phi = (0, 1) \) and \( y_1 > y_0 > 0 \). The concrete proof is presented in Appendix.
Lemma 4. Parameter optimization at the continuous level: assume $\phi \in (0, 1)$ and $y_1 > y_0 > 0$. Let $q_0 = y_0 \sqrt{1 + \phi}$, $q_1 = y_1 \sqrt{1 + \phi}$, $\mathcal{R}(q) = \max_{y \in [y_0, y_1]} \mathcal{R}(y, q)$, and $q_{opt}$ be the solution of min-max problem (36a) and (36b). Then,

(a) If $y_0 \geq \sqrt{1 + \phi} (\sqrt{1 + 1/\phi} - 1)$, we have

$$q_{opt} = q_1, \quad \mathcal{R}_{\max}(q_{opt}) = \mathcal{R}^+(q_{opt}), \quad (37a)$$

where $\mathcal{R}^+(q) = \max\{\mathcal{R}(y_0, q), \mathcal{R}(y_1, q)\}$ and $q_1$ is defined by

$$q_1 = \begin{cases} q_0, & \text{if } \mathcal{R}(y_0, q_0) \geq \mathcal{R}(y_1, q_0), \\ q_1, & \text{if } \mathcal{R}(y_1, q_1) \geq \mathcal{R}(y_0, q_1), \\ \text{otherwise}. & \end{cases} \quad (37b)$$

The argument $\tilde{q}_1 \in (q_0, q_1)$ is the unique root of $\mathcal{R}(y_0, q) = \mathcal{R}(y_1, q)$.

(b) If $y_0 < \sqrt{1 + \phi} (\sqrt{1 + 1/\phi} - 1) < y_1$, we have

$$q_{opt} = \begin{cases} q_0, & \text{if } q_d > q_1, \mathcal{R}(q_1) < y_1 \text{ and } \mathcal{R}(q_1) > \mathcal{R}^+(q_1), \\ q_1, & \text{otherwise}, \end{cases} \quad (38a)$$

and

$$\mathcal{R}_{\max}(q_{opt}) = \begin{cases} \mathcal{R}(q_{opt}), & \text{if } q_d > q_1, \mathcal{R}(q_1) < y_1, \text{ and } \mathcal{R}(q_1) > \mathcal{R}^+(q_1), \\ \mathcal{R}^+(q_{opt}), & \text{otherwise}, \end{cases} \quad (38b)$$

where $q_d = (1 + \phi) (\sqrt{1 + 1/\phi} - 1)$, $\mathcal{R}(q) = \mathcal{R}(\mathcal{R}_+(q), q)$, $d(q) = -\phi q^2 - 2(1 + \phi)q + 1 + \phi)^2$, and $q_0 \in (q^1, q_d)$ is the unique solution of

$$\mathcal{R}(q) = \mathcal{R}(y_0, q). \quad (39b)$$

(c) If $y_0 \leq \sqrt{1 + \phi} (\sqrt{1 + 1/\phi} - 1)$, then

$$q_{opt} = \begin{cases} q_0, & \text{if } \mathcal{R}_+(q_0) \geq y_1, \text{ or } \mathcal{R}(q_1) < y_1 \text{ and } \mathcal{R}(q_1) \leq \mathcal{R}^+(q_1), \\ q_1, & \text{otherwise}, \end{cases} \quad (39a)$$

and

$$\mathcal{R}_{\max}(q_{opt}) = \begin{cases} \mathcal{R}^+(q_{opt}), & \text{if } \mathcal{R}_+(q_0) \geq y_1, \text{ or } \mathcal{R}(q_1) < y_1, \text{ and } \mathcal{R}(q_1) \leq \mathcal{R}^+(q_1), \\ \mathcal{R}(q_{opt}), & \text{otherwise}, \end{cases} \quad (39b)$$

where $q_0^* \in (q^1, q_1)$ is the unique solution of the equation (38b).

Theorem 1. Let $S$ be the function defined in Lemma 2 and $\phi = \max\{S(\pi/T)^{1-\alpha}), S(\pi/\Delta T)^{1-\alpha})\}$. Then, for specified $\alpha, \beta \in (0, 1)$, $\mu \geq 0$, and $T, \Delta T > 0$, the parameter $p$ involved in the Robin TCs can be chosen as $p^* = q_{opt}/(2L)$, where $q_{opt}$ given by Lemma 4, is the solution of min-max problem (36a) and (36b) and $L \geq 0$ denotes the size of overlap. The quantities $y_0$ and $y_1$ in (36b) can be explicitly written down, as

$$y_0 = \begin{cases} 2L \lambda_1 (\frac{\pi}{T}), & \text{if } \alpha \geq \beta, \\ 2L \min\left\{ \lambda_1 (\frac{\pi}{T}), \lambda_2 (\frac{\pi}{\Delta T}) \right\}, & \text{if } \alpha < \beta, \end{cases} \quad (40)$$

$$y_1 = \begin{cases} 2L \lambda_1 (\frac{\pi}{\Delta T}), & \text{if } \alpha \geq \beta, \\ 2L \max\left\{ \lambda_1 (\frac{\pi}{T}), \lambda_2 (\frac{\pi}{\Delta T}) \right\}, & \text{if } \alpha < \beta. \end{cases}$$

With $p^*$, the convergence factor $p_\epsilon$ defined by (19) can be bounded by
\[ \rho_c(p^*, L) \leq R_{\text{max}}(q_{\text{opt}}), \]  

(41)

where \( R_{\text{max}}(q_{\text{opt}}) \) is given by (37a), or (38a), or (39b).

**Proof.** Bound (41) can be derived from (34) and (36b):

\[ (34), (36b) \Rightarrow \rho_c(p, L) \leq R_{\text{max}}(q) = \max_{\forall \delta \in [y_0, y_1]} R(y, q), \]

\[ \forall q (= 2L) > 0. \]

Hence, it remains to prove that the arguments \( y_0 \) and \( y_1 \) defined by (36b) can be explicitly expressed as (40). To this end, we define

\[ S_0(x) = \sqrt{x^2 + 2\mu x \cos \left( \frac{1 - \alpha}{2} \pi \right) + \mu^2 + x \cos \left( \frac{1 - \beta}{2} \pi \right)}, \]

(42)

Clearly, for any \( \alpha, \beta \in (0, 1) \) and \( \mu \geq 0 \), the function \( S_0(x) \) is increasing for \( x > 0 \). \( \square \)

**Case 1.** \( \alpha \geq \beta \). In this case, by noticing

\[ 2\lambda^2_R(\omega) = \omega^{a-b} S_0(\omega^{-a}), \]

(43)

we know that the function \( \lambda_R(\omega) \) is increasing for \( \omega \in [(\pi/T), (\pi/\Delta t)] \). Hence, we have

\[ y_0 = 2L \min_{\omega \in [(\pi/T), (\pi/\Delta t)]} \lambda_R(\omega) = L \sqrt{\left( \frac{\pi}{T} \right)^{\alpha - \beta} S_0 \left( \frac{\pi}{T} \right)^{1 - a}}. \]

\[ y_1 = 2L \max_{\omega \in [(\pi/T), (\pi/\Delta t)]} \lambda_R(\omega) = L \sqrt{\left( \frac{\pi}{\Delta t} \right)^{\alpha - \beta} S_0 \left( \frac{\pi}{\Delta t} \right)^{1 - a}}. \]

(44)

**Case 2.** \( \alpha < \beta \). In this case, from (43), we get

\[ 2\lambda^2_R(\omega) = S(x) := S_0(x)x^{- (b - a) - a}, \]

where \( x = \omega^{1-a} \). We have

\[ S^R(x) = \left[ x S_0(x) - (\beta - \alpha/1 - a) S_0(x) \right] x^{- (1 + \beta - 2\alpha + 1 - a)} \]

and for \( \mu = 0 \), it holds that \( S^R(x) = (1 - \beta/1 - a) S_0(x) x^{- (1 + \beta - 2\alpha + 1 - a)} > 0 \) if \( x > 0 \) and \( \alpha, \beta \in (0, 1) \); therefore, \( y_0 \) and \( y_1 \) satisfy (40).

We now consider the case \( \mu > 0 \). Suppose the function \( \tilde{S}(x) \) has a local extremum located at \( x = x_0 \). Then, we have \( x, S^R_0(x_0) = (\beta - \alpha/1 - a) S_0(x_0) \). Hence, it holds that \( S^R_0(x_0) = \left[ x_0 S_0(x_0) + (1 - \beta/1 - a) S_0(x_0) \right] x_0^{- (1 + \beta - 2\alpha + 1 - a)} \). We claim \( S^R_0(x_0) > 0 \). Indeed, from (42), it is easy to get \( S^R_0(x_0) > 0 \) and this, together with \( \alpha, \beta \in (0, 1) \), implies \( (1 - \beta/1 - a) S^R_0(x_0) > 0 \). For \( S^R_0(x_0) \), a tedious but routine calculation yields

\[ S^R_0(x_0) = \left[ x_0^2 + 2\mu x_0 \cos \left( (1 + \alpha/2) \pi \right) + \mu^2 - (x_0 + \mu \cos \left( (1 + \beta/2) \pi \right))^2 \right] \]

(45)

\[ \left( x_0^2 + 2\mu x_0 \cos \left( (1 + \beta/2) \pi \right) + \mu^2 \right) \left( x_0^2 + 2\mu x_0 \cos \left( (1 - \beta/2) \pi \right) + \mu^2 \right) > 0. \]

This implies that \( \tilde{S}(x) \) has no local maximum(s) and thus \( y_{0,1} \) also satisfy (40).

We next comment in two remarks the parameter optimization studied here for the time-fractional problems. First, in Remark 1, we show that the analysis of finding the solution of min-max problem (36a) and (36b) is different from the analysis existing in the literature. Then, in Remark 2, we show that the parameter given by Lemma 4 is close to the best one and that we can get through numerical optimizations.

**Remark 1.** About min-max problem (36a) and (36b): for the integer order reaction-diffusion equation \( \partial_t u = \partial_{xx} u - \mu u + f \) (or more generally, the advection-reaction-diffusion equation \( \partial_t u = \partial_{xx} u + b \partial_x u - \mu u + f \)), the related min-max problem is

\[ \min_{\varphi \geq 0} \max_{q \in [y_0, y_1]} \mathcal{R}_{\text{int-order}}(y, q), \]

with

\[ \mathcal{R}_{\text{int-order}}(y, q) = \frac{(y - q)^2 + y^2 - y_0^2}{(y + q)^2 + y^2 - y_0^2}, \]

(46)

where \( y_1 > y_0 > 0 \) are constants related to the problem/ discretization parameters (see [35] for more details). An important property of this min-max problem is that, as we already mentioned above, the solution is determined by the equioscillation principle [30,35]. Although the basic mathematical tools used for analyzing (46) and (36a) and (36b) are the same (such as monotonicity, continuity of real-valued functions, and roots of quadratic/cubic polynomials), the concrete procedures are different. This difference can be seen from two points:
(1) The two functions, $R_{\text{int-order}}$ in (40) and $R$ in (36a) and (36b), are different.

(2) Because of this difference, the solution $q_{\text{opt}}$ of min-max problem (36a) and (36b) is not always determined by the equioscillation principle.

For min-max problem (36a) and (36b), there are two reasons that the equioscillation principle does not hold. First, the local maximizer $\gamma_r(q_{\text{opt}})$ does not exist in the interval $[y_0, y_1]$ (see Figure 2(a)). Second, more typically, even though the local maximizer $\gamma_r(q_{\text{opt}})$ exists, $R(y_0, q_{\text{opt}}) \neq R(\gamma_r(q_{\text{opt}}), q_{\text{opt}})$ (see Figure 2(b)). What we observed in these two subfigures does not happen occasionally, and it seems an inherent feature for the time-fractional problems. Precisely, as will be shown in our forthcoming paper, in the asymptotic sense, i.e., $L = \mathcal{O}(\Delta x)$, $\Delta t = \mathcal{O}(\Delta x^\alpha)$ and $\Delta x$ small, the equioscillation property always does not hold for min-max problem (36a) and (36b) when $r < (1/1 - \beta)$ and $\alpha, \beta \in (0, 1)$.

**Remark 2.** About the parameter $q = q_{\text{opt}}$: the solution $q_{\text{opt}}$ of min-max problem (36a) and (36b) depends on $y_{0,1}$ and $\phi$.

\[ q_{\min} = 2L \min_{\omega \in ([\pi T], (\pi / \Delta t))] \sqrt{\lambda_2^2(\omega) + \lambda_1^2(\omega)}, \quad q_{\max} = 2L \max_{\omega \in ([\pi T], (\pi / \Delta t))] \sqrt{\lambda_2^2(\omega) + \lambda_1^2(\omega)}. \tag{48} \]

(2) Compute and then compare the values of $R_{\max}(q) = \max_{\omega \in ([\pi T], (\pi / \Delta t))] \|R(w, q)\|$, i.e., the convergence factor of the SWR algorithm with $q = q_{\text{opt}}$ (the approximate solution) and $q = q_{\text{best}}$ obtained through numerical optimizations.

Let $\Delta t = 0.02$, $T = 40$, $L = 0.04$, and $\mu = 8$. Then, in Figure 3, we plot $R_{\max}(q_{\text{opt}})$ (Figure 3(a)) and $R_{\max}(q_{\text{best}})$ (Figure 3(b)) for $\alpha, \beta \in [0, 1]$, where a small step-size, $\Delta q = 10^{-3}$, is used in the above numerical optimization procedure to get $q_{\text{best}}$. In Figure 3(c), we plot which are related to $\alpha, \beta, \mu, T$ (problem parameters) and $\Delta t, L$ (discretization parameters). Hence, we denote $q_{\text{opt}}$ by

\[ q_{\text{opt}} = \hat{\gamma}_{\text{new}}(\alpha, \beta, \mu, \Delta t, L, T), \tag{47} \]

which we call “parameter formula” and its concrete expression is given by Lemma 5 as we have seen. Note that $q_{\text{opt}}$ is not the solution of original min-max problem (24a) and (24b); it is only an approximate solution. Hence, it is important to present evidence to show that $q_{\text{opt}}$ is really a good approximate solution. Our evidence is based on numerical optimization:

(1) Find the best parameter $q_{\text{best}}$ through repetitive numerical optimizations for original min-max problem (24a) and (24b):

(i) For $q_j = q_{\min} + j \Delta q$, compute $R_j := \max_{\omega \in ([\pi T], (\pi / \Delta t))] \|R(w, q_j)\|$ by the global optimization toolbox in Matlab, where $j = 0, 1, \ldots, J = (q_{\max} - q_{\min}/\Delta q)$ (since sign $(\partial_q \tilde{R}(w, q)) = \text{sign}(q - 2L/\lambda_2^2(\omega) + \lambda_1^2(\omega))$, it suffices to consider $q \in [q_{\min}, q_{\max}]$)

\[ \tilde{R}_{\max}(q_{\text{opt}}) - \tilde{R}_{\max}(q_{\text{best}}), \text{i.e., the error between the convergence factors using the two parameters. This figure clearly shows that the quantity $q_{\text{opt}}$, given by Lemma 4, is really a good approximate solution of original min-max problem (24a) and (24b).}

### 4. Convergence Analysis at the Discrete Level

To make the SWR algorithms practical for real computation, we need an analysis at the fully discretized level. To this end, we first make a transform for (5), which avoids derivatives. Integrating (10) from 0 to $t$, we get

\[ u_j^k - \mathcal{J}_t^{1-\beta} \partial_x u_j^k + \mu \mathcal{J}_t^{1-a} u_j^k = \tilde{f}(x, t), \quad x \in \Omega_j, \]

\[ (\partial_x + (-1)^{j+1}) u_j^k ((2j)L, t) = (\partial_x + (-1)^{j+1}) u_{j-1}^k ((2j-2)L, t), \quad j = 0, \ldots, \frac{L}{\Delta x}, \tag{49} \]

where $\tilde{f} = \int_0^t f(x, \tau) d\tau + u_0(x)$. For any $\gamma \in [0, 1]$, $\mathcal{J}_t^{\gamma}$ denotes fractional integral.
\[ R(y, q_{opt}) = R(y_0, q_{opt}) - R(\bar{y} + (q_{opt}), y_{opt}) \]

Figure 2: Two typical behaviors of \( R(y, q_{opt}) \), where the equioscillation principle does not hold. Here, \( \alpha, \beta, \) and \( \mu \) are the problem parameters, \( L \) is the overlap size, and \( T \) is the length of time-interval.

Figure 3: Comparisons of \( \tilde{R}_{\text{max}}(q) \), i.e., the convergence factor of the SWR algorithm with Robin TCs, under two choices of the parameter \( q \). (a) \( q = q_{opt} \) and (b) \( q = q_{best} \). (c) We show the error \( \tilde{R}_{\text{max}}(q_{opt}) - \tilde{R}_{\text{max}}(q_{best}) \).

\[
\mathcal{J}_t^\gamma v(t) = \begin{cases} 
v(t), & \text{if } \gamma = 0, \\
\frac{1}{\Gamma(\gamma)} \int_0^t 1 \ (t - \tau)^{-\gamma} v(\tau) d\tau, & \text{if } \gamma \in (0, 1].
\end{cases}
\]

(50)

\[
U_{j,m}^k(t) - \mathcal{J}_t^{1-\beta} \delta_t^{x} U_{j,m}^k(t) + \mu \mathcal{J}_t^{1-\alpha} U_{j,m}^k(t) = f(x_m, t), \quad x_m \in \Omega, \ j = 1, 2,
\]

(51a)

where \( U_{j,m}^k(t) = u_j^k(x_m, t) \), \( \delta_t^{x} U_{j,m}^k(t) = (U_{j,m+1}^k(t) - 2U_{j,m}^k(t) + U_{j,m-1}^k(t))/\Delta x^2 \), \( x_m = m\Delta x \), and \( m \in \mathbb{Z} \). To simplify the convergence analysis, we assume \( L = \Delta x \), i.e., the size of overlap equals a single space mesh size. Then, the artificial boundary interfaces \( x = L \) (resp. \( x = 0 \)) corresponds to \( x_1 \) (resp. \( x_0 \)) and the Robin TCs are discretized as

\[
\frac{U_{1,1}^k(t) - U_{1,0}^k(t)}{\Delta x} + pU_{1,1}^k(t) = \frac{U_{2,1}^{k-1}(t) - U_{2,0}^{k-1}(t)}{\Delta x} + pU_{2,1}^{k-1}(t),
\]

(51b)
Here, we discretized the spatial derivatives in the Robin TCs as
\[
\partial_x u^k_i(x_1, t) \approx \frac{U^k_{i,1}(t) - U^k_{i,0}(t)}{\Delta x},
\]
(51c)
\[
\partial_x u^k_i(x_0, t) \approx \frac{U^k_{i,2}(t) - U^k_{i,0}(t)}{\Delta x}.
\]
(51c)

Note that, (51c) is a first-order discretization, and therefore, a question arises naturally: does this simple discretization affect the accuracy of the converged solutions? Following the proof given in [43], Lemma 1, this worry can be removed.

In (51a), we need to deal with convolutions, as
\[
\int_0^{t} \mathcal{K}_y(t - \tau) V(\tau) d\tau, \quad \text{with kernel function } \mathcal{K}_y(t) = \frac{t^{r-1}}{\Gamma(r)},
\]
(52)
where \( y = 1 - \alpha \) (or \( 1 - \beta \) and \( V \) is not known beforehand but are computed as the computing proceeds from time step to time step. A naive implementation would require \( \Theta(N^2) \) operations and \( \Theta(N^t) \) memory for per expansion coefficients over \( N \) time steps. In [44], fast convolution algorithms are constructed, which implements, over \( N \) time steps, with \( \Theta(N, \log(N)) \) operations and \( \Theta(\log(N)) \) active memory.

The starting point of the fast convolution algorithms is to replace the kernel function \( \mathcal{K}_y(t - \tau) \) in (52) by its inverse Laplace transform, along some contour \( Y \):
\[
\int_0^{t} \left( \frac{1}{2\pi i} \int Y e^{rac{s}{\tau - \tau}} \mathcal{K}_y(s) ds \right) V(\tau) d\tau,
\]
(53)
where \( \mathcal{K}_y(s) = \mathcal{L}(\mathcal{K}_y) \), \( s^{-1} \) denotes the Laplace forward transform of the kernel function \( \mathcal{K}_y \). The contour \( Y \) is a curve in the sector oriented with an increasing imaginary part and going to infinity with an acute angle to the negative real half-axis, so that \( e^s \) decays rapidly for growing \( s \) along \( Y \). In (53), reversing the order of integration gives
\[
\frac{1}{2\pi i} \int Y \mathcal{K}_y(s) \left( \int_0^{\tau} e^{(r - r)} V(\tau) d\tau \right) ds = \frac{1}{2\pi i} \int Y \mathcal{K}_y(s) y(0, t; s) ds.
\]
(54)
The function \( y(0, t; s) \) is the solution of the following linear scale ODE at \( \tau = t \):
\[
\begin{cases}
  y' \tau = s \tau + V \tau, & \tau \in (0, t), \\
  y(0) = 0.
\end{cases}
\]
(55)

Multiplying (56) by \( \zeta^n \) and summing over \( n \) from 0 to \( +\infty \), we obtain
\[
\left( \sum_{j=0}^{K} a_j \zeta^{-j} \right) y(\zeta) = \left( \sum_{j=0}^{K} b_j \zeta^{-j} \right) [\Delta t y(\zeta) + \Delta t V(\zeta)],
\]
(57)
with the generating (formal) power series \( y(\zeta) = \sum_{n=0}^{+\infty} \eta_n z^n \) and \( V(\zeta) = \sum_{n=0}^{+\infty} \nu_n z^n \). Solving \( y(\zeta) \) from (57), we get
\[
\left( \frac{\Lambda(\zeta)}{\Delta t} - s \right) y(\zeta) = V(\zeta), \quad \text{with } \Lambda(\zeta) = a_0 + \cdots + a_{K-1} \zeta + a_K
\]
(58)
\[
b_0 \zeta^K + \cdots + b_{K-1} \zeta^{K-1} + b_K
\]
where \( \Lambda(\zeta) \) is called the quotient of the generating polynomial of a linear multistep method. Clearly, the solution \( \eta_n \) of (56) is the \( n \)-th coefficient of the power series \( \Lambda(\zeta)/(\Delta t - s)^{-1} V(\zeta) \), that is, \( y_0(0, t_0; s) = n \)-th coefficient of \( \Lambda(\zeta)/(\Delta t - s)^{-1} V(\zeta) \). By substituting this into the right-hand side of (54), we get
\[
\frac{1}{2\pi i} \int Y \mathcal{K}_y(s) \left( \int_0^{\tau} e^{(r - r)} V(\tau) d\tau \right) ds
\]
\[
\approx n - \text{th coefficient of } \frac{1}{2\pi i} \int Y \mathcal{K}_y(s) \left( \frac{\Lambda(\zeta)}{\Delta t} - s \right)^{-1} ds V(\zeta)
\]
\[
\approx n - \text{th coefficient of } \frac{1}{2\pi i} \int Y \mathcal{K}_y(s) \left( \frac{\Lambda(\zeta)}{\Delta t} - s \right)^{-1} ds V(\zeta)
\]
\[
\approx n - \text{th coefficient of } \frac{1}{2\pi i} \int Y \mathcal{K}_y(s) \left( \frac{\Lambda(\zeta)}{\Delta t} - s \right)^{-1} ds V(\zeta)
\]
(59)

Now, we expand \( \mathcal{K}_y(\Lambda(\zeta)/\Delta t) \) in the formal power series, as
\[
\mathcal{K}_y(\Lambda(\zeta)/\Delta t) = \sum_{j=0}^{+\infty} c_j(y, \Delta t) \zeta^j.
\]
(60)

Then, the \( n \)-th coefficient of \( \mathcal{K}_y(\Lambda(\zeta)/\Delta t) V(\zeta) \) is given by
\[
\sum_{j=0}^{n} c_{n-j}(y, \Delta t) V_j.
\]
(61)

Suppose the time-integrator applied to (55) is an A-stable method with order \( p \). Then, from [40], Theorem 2.1, we know that (61) is an approximation to (52) and
\[
\int_0^{t} \mathcal{K}_y(t - \tau) V(\tau) d\tau = \sum_{j=0}^{n} c_{n-j}(y, \Delta t) V_j + \Theta(h^p), \quad \forall n \geq 1.
\]
(62)

The last step is to determine the weights \( \{c_j(y, \Delta t)\}_{j=0}^{+\infty} \). For general case, we have to rely on numerical calculation to get \( \{c_j(y, \Delta t)\}_{j=0}^{+\infty} \) (see [44] for more details). For some special cases, such as the linear \( \Theta \)-method \((\Lambda(\zeta) = (1 - \zeta)(\theta + (1 - \theta)(\zeta))) \) and the two-step backward difference formula (BDF2, \( \Lambda(\zeta) = \sum_{j=1}^{2} (1 - \zeta)/j \)), these weights can be written down, as
\[ c_j(\gamma, \Delta t) = \begin{cases} (\theta \Delta t)^{i - 1} \frac{1 - \theta}{\theta} \sum_{l=0}^{j-i-1} \left( \begin{array}{c} \theta \Delta t \\ j-i-l \end{array} \right) \left( \begin{array}{c} \gamma \\ l \end{array} \right), \theta \in [0.5, 1), (\Delta t)^{i-1} \left( \begin{array}{c} -\gamma \\ l \end{array} \right), \theta = 1, \end{cases} \]

BDF2: \[ c_j(\gamma, \Delta t) = \frac{2\Delta t}{3} \left( 1 - \right)^{i-1} \sum_{l=0}^{j-i-1} \left( \begin{array}{c} -\gamma \\ l \end{array} \right). \]

With the quadrature weights \( c_n(\gamma, \Delta t) \) in our hands, we denote the discrete convolution (61) by \( \mathcal{H}_{\gamma, \Delta t} V(n) \), that is,

\[ \left( \mathcal{H}_{\gamma, \Delta t} V \right)(n) = \sum_{l=0}^{n} c_{n-l}(\gamma, \Delta t)V_l. \]

Then, the scheme of the fully discrete SWR algorithms on the \( j \)-th subdomain is

\[ U_{jm}^k(n) - \left( \mathcal{H}_{1-\beta, \delta} \mathcal{H}_{\gamma, \Delta t} \right)(n) + \mu \left( \mathcal{H}_{1-\alpha, \delta} \right) U_{jm}^k(n) = \tilde{f}(x_m, t_n), \quad x_m \in \Omega_j. \]  

4.2. Convergence Analysis. Since the quadrature weights \( c_n(\gamma, \Delta t) \) are the coefficients of the formal power series of \( \mathcal{H}_{\gamma, \Delta t} \), the discrete Laplace transform of the sequence \( \{(\mathcal{H}_{\gamma, \Delta t} V(n) \} \) defined by (64) is

\[ \mathcal{E}_{jm}^k(z) = \frac{1}{\Lambda(z)} \mathcal{E}_{jm}^k(z) + \mu \mathcal{E}_{jm}^k(z) = 0, \quad x_m \in \Omega_j, \]

where

\[ \tilde{\Delta} = \frac{\Lambda(z)}{\Delta T}, a(\tilde{\Delta}) = 1 + \Delta x \rho. \]

Define

\[ \eta = 1 + \Delta x \rho. \]

Then, interface conditions (51b) are transformed to

\[ \eta \mathcal{E}_{1,1}^k(z) - \mathcal{E}_{1,0}^k(z) = \eta \mathcal{E}_{2,1}^k(z) - \mathcal{E}_{2,0}^k(z), \]

\[ \eta \mathcal{E}_{2,0}^k(z) - \mathcal{E}_{2,1}^k(z) = \eta \mathcal{E}_{2,0}^k(z) - \mathcal{E}_{2,1}^k(z). \]

From (67), the arguments \( E_{jm}^k \) can be expressed as

\[ \mathcal{E}_{jm}^k(z) = \frac{1}{\Lambda(z)} \mathcal{E}_{jm}^k(z) + \mu \mathcal{E}_{jm}^k(z) = 0, \quad x_m \in \Omega_j, \]

i.e.,

\[ \mathcal{E}_{jm}^k(z) - \mathcal{E}_{jm}^k(z) = 0, \]

where

\[ \tilde{\Delta} = \frac{\Lambda(z)}{\Delta T}, a(\tilde{\Delta}) = 1 + \Delta x \rho. \]

Lemma 5. Let \( \sigma, \Delta t, T > 0, \epsilon = e^{-\Delta t}, \) and \( c_{\text{max}} = \cos(\pi/T). \) Then, for \( z = e^{-\Delta t} \) with \( s = \sigma + i\omega \) and \( |\omega| \in [(\pi/T), (\pi/\Delta t)], \) we have

1. If \( \Lambda(z) = (1-z)/(\theta + (1-\theta)z) \) (linear \( \theta \)-method):

   (i) \( \Re(\Lambda(z)) \in [(1-c_{\text{max}})(2\theta-1)/(1-\theta)^2 + \theta^2 + 2\theta(1-\theta)c_{\text{max}}, 2(1-\theta-1)] \) and \( |\Im(\Lambda(z))| \leq 1/(1-2\theta)^2 \), if \( \theta \in ((1/2), 1]; \)

   (ii) \( \Re(\Lambda(z)) \in [2(1-\epsilon/1+\epsilon), 2(1+\epsilon/1-\epsilon)] \) and \( |\Im(\Lambda(z))| \leq 4(1+\epsilon^2/(1-\epsilon)^2) \sin(\arccos(2\epsilon/1+\epsilon^2)), \) if \( \theta = (1/2); \)

2. If \( \Lambda(z) = \sum_{i=1}^{\infty} (1-z)/j \) (BDF2 method):

   \( \Re(\Lambda(z)) \in [(1-c_{\text{max}}^2)^{3/2}, 3/2], |\Im(\Lambda(z))| \leq 2.202. \)
Proof. Let \( c = \cos(\omega \Delta t) \) and \( \bar{c} = \sin(\omega \Delta t) \). We have, \( t \in (0, 1) \), \( c \in [-1, c_{\max}] \), and \( \bar{c} \in [-1, 1] \). The proof for the two time-integrators is given as follows:

\[
\Lambda(z) = \frac{1 - \epsilon (c + \bar{c})}{\theta + \epsilon (1 - \theta)(c + \bar{c})} = \Lambda_R(e, c) - i\Lambda_I(e, c, \bar{c}) = \frac{N(e, c)}{D(e, c)} - i \frac{\epsilon \bar{c}}{D(e, c)}
\]

where \( N(e, c) = -e^2 (1 - \theta) + e (1 - 2\theta)c + \theta \) and \( D(e, c) = e^2 (1 - \theta)^2 + 2\epsilon\theta(1 - \theta)c + \theta^2 \). For \( \theta = 0.5 \), the bounds for \( \Re(\Lambda) \) and \( \Im(\Lambda) \) can be obtained easily. For \( \theta \in (0.5, 1] \),

\[
\begin{align*}
\text{sign}(\partial_z \Lambda_R) &= \text{sign}\left[-e^2 (1 - \theta)^2 c - 2\epsilon (1 - \theta)(1 - \theta^2)c\right], \\
\text{sign}(\partial_z \Lambda_I) &= \text{sign}(\bar{c}) - \left(\frac{\theta}{1 - \theta} - \epsilon\right) = \text{sign}(\bar{c}), \\
\text{sign}(\partial_c \Lambda_R) &= \text{sign}\left(\frac{e - \theta}{1 - \theta}\right), \text{sign}(\partial_c \Lambda_I) = \text{sign}(\bar{c}).
\end{align*}
\]

(73)

Clearly, it holds that \( \partial_z \Lambda_R < 0 \) and this implies

\[
\begin{align*}
\max_{\epsilon \in [0, 1]} \max_{\epsilon \in [-1, c_{\max}]} \Lambda_R(e, c) &= N(1, -1) = \frac{2}{(2\theta - 1)}, \\
\min_{\epsilon \in [0, 1]} \min_{\epsilon \in [-1, c_{\max}]} \Lambda_R(e, c) &= N(1, c_{\max}) > \frac{N(1, 1)}{D(1, 1)} = 0.
\end{align*}
\]

This gives the bounds of \( \Re(\Lambda) \) and \( \Im(\Lambda) \) for \( \theta \in (0.5, 1] \).

(2) For the BDF2 method: \( \Lambda(z) = \sum_{j=1}^{2} (1 - z)^j / j \). In this case, we have

\[
|\Im(\Lambda(z))| = \sqrt{1 - \epsilon^2 (2 - \epsilon^2)c} \leq \sqrt{1 - \epsilon^2 (2 - c)}, \quad \text{for} \ \epsilon \in (0, 1),
\]

(77)

\( \xi_+ (3) \) are also analytic. Since \( \Re(1 + \alpha^2 (3)) > 1 \), we have \( \Re(\xi_+ (3)) > 1 \Rightarrow |\xi_+ (3)| > 1 \); this, together with \( \xi_+ (5) \xi_+ (3) \equiv 1 \), gives \( |\xi_+ (5)| < 1 \).

Lemma 6. Let \( \alpha, \beta \in (0, 1) \), \( \mu > 0 \), and \( H := \{ s \in C^+: \Im(s) \in [(\pi/T), (\pi/\Delta t)] \} \). Then, for the linear \( \theta \)-method with \( \theta \in [(1/2), 1) \) and the BDF2 method, the arguments \( \xi_+ (3) \) with \( 3 = e^{-\alpha \Delta t} \) are analytic for \( s \in H \) and satisfy \( |\xi_+ (3)| > 1 \) and \( |\xi_+ (5)| < 1 \).

Proof. For \( s \in H \), from Lemma 5, we have \( 3 = (\Lambda(e^{-\alpha \Delta t})/\Delta t) \in C^+ \). Then, for \( \mu > 0 \), it holds that \( \Re(\alpha(3)) > 1 \) for \( 3 \in C^+ \), where \( \alpha(3) \) is defined by (68). This implies \( \Re(1 + \alpha^2 (3)) > 1 \), and thus, the composite functions

(1) For the linear \( \theta \)-method: \( \Lambda(z) = (1 - z)/(\theta + (1 - \theta)z) \). In this case, we have

\[
\Lambda(z) = \frac{1 - \epsilon (c + \bar{c})}{\theta + \epsilon (1 - \theta)(c + \bar{c})} = \Lambda_R(e, c) - i\Lambda_I(e, c, \bar{c}) = \frac{N(e, c)}{D(e, c)} - i \frac{\epsilon \bar{c}}{D(e, c)}
\]

Now, by using the boundedness conditions at infinity for each subsystem, the constants \( A_k \) and \( B_k \) in (71) can be determined as \( B_k = 0 \) and \( A_k = 0 \) for the first and second subsystems, respectively. We then have \( E_{\theta, m}^k = A_k \xi_m \) and \( E_{\theta, m}^k = B_k \xi_m \). By using the transmission conditions in (70), we get the recurrence relation \( A_k (\eta - \xi_1) = B_{k-1} (\eta - \xi_1) \) and \( B_k (\eta - \xi_1) = A_{k-1} (\eta - \xi_1) \), i.e.,
\[ A_k = \tilde{p}_d (s) A_{k-2}, \quad B_k = \tilde{p}_d (s) B_{k-2}, \quad \text{with} \quad \tilde{p}_d (s) = \left( \frac{\eta - \xi_s (s)}{\eta \xi_s (s) - 1} \right)^2, \]

where the subscript "\( \tilde{\cdot} \)" denotes "discrete" and

\[ s = \Lambda \left( e^{-\Delta t \Delta} \right), \quad s \in \mathbb{H} = \{ s \in \mathbb{C}^+ : [\mathcal{F} (s)] \in [(\pi/T), (\pi/\Delta t)] \}. \]

(78b)

Mathematically, we expect \(|\rho_d| \ll 1\), which leads to the following min-max problem:

\[ \min_{\eta \in \mathbb{R}} \max_{s = (\Lambda (e^{-\Delta t \Delta}) \Delta) t \in \mathbb{H}} \left| \frac{\eta - \xi_s (s)}{\eta \xi_s (s) - 1} \right|^2. \]

(79)

Let \( \xi_s (s) = \xi_R (s) + i \xi_I (s) \). Then, from Lemma 6, we know \( \xi_R (s) > 1 \). For \( \eta \leq 1 \), we have \((\xi_s (s))^2 < 1\) and \((\xi_s (s))^2 > 1\). Hence, \( \xi_s \) approaches to 1 as \( \eta \rightarrow (0, \pm \pi) \) in the complex plane. This implies that, it is sufficient to consider \( \eta > 1 \) in (79).

For \( \eta \rightarrow 0.5 \), the quantity \( \bar{s} = (\Lambda (e^{-\Delta t \Delta}) \Delta) t \) approaches to \( s \rightarrow (0, \pm \pi) \). This implies that \( \max_{s \in \mathbb{H}} |\rho_d (s)| \approx 1 \), independent of \( \eta \).

For \( \bar{s} \in (0.5, 1) \), we know from Lemma 5 that the origin is kept away from \( \bar{s} \), and therefore, it is possible to expect

\[ D(\bar{s}_{\min}, \bar{s}_{\max}, \bar{s}_{\max}) = \{ s \in \mathbb{C}^+ : \bar{s}_{\min} \leq \Re (s) \leq \bar{s}_{\max}, 0 \leq \Im (s) \leq \bar{s}_{\max} \}, \]

(80)

which lies in the right-hand side of the vertical line \( s = 1 \) (i.e., \( \bar{s}_{\min} > 1 \)) and covers the region mapped by \( \bar{s} \); see Figure 4 for an illustration. Then, regarding \( \xi_s \) as a single complex variable, solve (79) over this rectangular region.

For the linear \( \theta \)-method with \( \theta \in (0.5, 1) \) and the BDF2 method, if \( \alpha > \beta \), by using Lemma 5 and Lemma 7 given below, the boundaries of the rectangular region \( D \), i.e., \( \bar{s}_{\min}, \bar{s}_{\max}, \bar{s}_{\max} \), can be expressed explicitly, as

\[ \bar{s}_{\min} = \xi_R (\bar{s}_{\min}, \bar{s}_{\max}), \quad \bar{s}_{\max} = \xi_I (\bar{s}_{\min}, \bar{s}_{\max}), \]

(81a)

where \( \xi_R, \xi_I \) are defined by (87a). The quantities \( \bar{s}_{\min}, \bar{s}_{\max}, \bar{s}_{\max} \), respectively, are given by

\[ \bar{s}_{\min} = \frac{1 + \Delta x^2}{2} \left[ \sigma_0^{1-\beta} + \mu \sigma_0^{a-\beta} \right], \]

\[ \bar{s}_{\max} = \frac{1 + \Delta x^2}{2} \left[ \sigma_{\max}^{1-\beta} \sin((1 - \beta)\psi_1) + \mu (\sigma_{\max}^{a-\beta}) \cos((\alpha - \beta)\psi_1) \right], \]

(81b)

with \( \psi_0 = \arccos \left( \frac{\sigma_0}{\sqrt{\sigma_0^2 + \omega_{\max}^2}} \right), \psi_1 = \arccos \left( \frac{\sigma_1}{\sqrt{\sigma_1^2 + \omega_{\max}^2}} \right) \).
Lemma 7. \((a \geq b)\). Assume \(0 < \beta \leq a < 1\) and \(\mu \geq 0\). Let \(D_0 = \{z \in \mathbb{C} : 0 < \sigma_0 \leq \Re (s) \leq \sigma_1, 0 \leq \Im (s) \leq \omega_{\text{max}}\}\) and \(D = \{s \in \mathbb{C} : \sigma_{\text{min}} \leq \Re (s) \leq \sigma_{\text{max}}, 0 \leq \Im (s) \leq \bar{\omega}_{\text{max}}\}\), where \(\sigma_{\text{min}}, \sigma_{\text{max}}\) and \(\omega_{\text{max}}\) are given by (81a) and (81b). Then, we have \(\xi_i(s) \in \hat{D}\) for \(s \in D_0\).

\[
\begin{pmatrix}
\sigma_0 &=& \frac{(1-c_{\text{max}})(2\theta-1)}{\Delta t \left( (1-\theta)^2 + \theta^2 + 2(1-\theta)c_{\text{max}} \right)} \\
\sigma_1 &=& \frac{2}{(2\theta-1)\Delta t} \omega_{\text{max}} = \frac{1}{(2\theta-1)^2 \Delta t}
\end{pmatrix}
\]

where \(\theta \in ((1/2), 1]\) and \(c_{\text{max}} = \cos(\pi/T)\). In general, we have to rely on numerical optimizations to obtain these three quantities, i.e., \(\sigma_0 = \min_{e^{[\lambda \left( 1-e^{-\omega_d/\Delta t} \right), e^{\Delta t}}\Re (\xi_i(s))}, \sigma_1 = \max_{e^{[\lambda \left( 1-e^{-\omega_d/\Delta t} \right), e^{\Delta t}}\Re (\xi_i(s))}, \sigma_{\text{max}} = \max_{e^{[\lambda \left( 1-e^{-\omega_d/\Delta t} \right), e^{\Delta t}}\Re (\xi_i(s))}\).

Proof. With the arguments \(\sigma_0, \sigma_1, \) and \(\bar{\omega}_{\text{max}}\) given by (81b), we define \(\hat{D} = \{s \in \mathbb{C} : 0 < \sigma_0 \leq \Re (s) \leq \sigma_1, 0 \leq \Im (s) \leq \bar{\omega}_{\text{max}}\}\). Then, the proof is divided into two steps: \(a(s) \in \hat{D}\) if \(s \in D_0\) and \(s + \sqrt{s^2 - 1} \in \hat{D}\) if \(s \in \hat{D}\). The function \(a(s)\) is defined by (68).

Step 1. \(a(s) \in \hat{D}\) for \(s \in D_0\). Let \(s = \sigma + i\omega \in D_0\). Then, we can rewrite \(a(s)\) as

\[
\begin{align*}
\left( \frac{\Delta x^2}{2} \left( e^{G_{a}(1-\beta,\sigma)} + \mu e^{G_{a}(a-\beta,\sigma)} \right) \right) &+ i \left( \frac{(\Delta x^2)}{2} \left( e^{G_{b}(1-\beta,\sigma)} + \mu e^{G_{b}(a-\beta,\sigma)} \right) \right),
\end{align*}
\]

where \(r = \sqrt{\sigma^2 + \omega^2}\). For any \(y \in (0, 1)\), the functions \(G_{R,f}\) are defined by

\[
\begin{align*}
G_R(y, r, \sigma) &= y \log r + \log \cos(y \psi(r, \sigma)), \\
G_f(y, r, \sigma) &= y \log r + \log \sin(y \psi(r, \sigma)),
\end{align*}
\]

where \(\psi(r, \sigma) = \arccos(\sigma/r)\). Since \(\sigma > 0\), we have \(\psi(r) \in (0, \pi/2)\). For fixed \(\sigma \in [\sigma_0, \sigma_1]\), because \(\psi(r, \sigma)\) is an increasing function of \(r\), \(G_f\) is also an increasing function of \(r\). Hence, for \(y \in [0, 1]\), \(\max_{\sigma \in [\sigma_0, \sigma_1]} G_f(y, \sqrt{\sigma^2 + \omega_{\text{max}}^2}, \sigma)\) is defined as \(\overline{\omega}_{\text{max}}\), shown in (81b).

It remains to consider the lower and upper bounds of \(\Re (a(s))\). For \(r = \sqrt{\sigma^2 + \omega^2}\) and \(\sigma \in [\sigma_0, \sigma_1]\), we have \(\partial_r G_R(y, r, \sigma) = \max_{\omega_{\text{max}}/r} (\omega_{\text{max}}/r)\). Hence, for \(y \in (0, 1)\), it holds that \(\partial_r G_R(y, r, \sigma) \geq 0\) and this gives

\[
\begin{align*}
\text{sign} [\partial_r (y \log r + \log \cos(y \psi(r)))] &= \text{sign} (y \log r + \log \cos(y \psi(r))) = \text{sign} [(\omega_{\text{max}}/r) \log \cos(y \psi(r)) - \omega_{\text{max}}] = \text{sign} [\partial_r G_{R,f}(y, \sigma)],
\end{align*}
\]

Hence, \(G_R\) attains its global maximum at \(\sigma = \sigma_1\); this gives \(\overline{\sigma}_1\), the upper bound of \(\Re (a(s))\).
\[ \xi_R (\sigma, \omega) = \sigma + \sqrt{\frac{\left( \omega^2 - 1 \right)^2 + 4\sigma^2 \omega^2 + (\sigma^2 - \omega^2 - 1)}{2}}, \]  
(87a)

\[ \xi_I (\sigma, \omega) = \omega + \sqrt{\frac{\left( \omega^2 - 1 \right)^2 + 4\sigma^2 \omega^2 - (\sigma^2 - \omega^2 - 1)}{2}}. \]  
(87b)

Clearly, it holds \( \partial_{\omega} \xi_I (\sigma, \omega) > 0 \) and \( \partial_{\sigma} \xi_R (\sigma, \omega) > 0 \).

Moreover, some simple algebra shows

\[
\text{sign}(\partial_{\omega} \xi_R (\sigma, \omega)) = \text{sign}\left( \frac{\omega^2 + 1 - \sqrt{(\omega^2 - 1)^2 + 4\sigma^2 \omega^2}}{\omega^2 - \sigma^2} \right) \geq 0,
\]

(87b)

Hence, \( \max_{ \omega \in [0, \omega_{\text{max}}] } \xi_I (\sigma, \omega) = \xi_I (\sigma, \omega_{\text{max}}) \) and from the second inequality given in (87b), we know that \( \xi_I (\sigma, \omega_{\text{max}}) \) attains its maximum at \( \sigma = \sigma_0 \); this gives the upper bound of \( \mathfrak{S} (\xi_+ (s)) \), i.e., \( \xi_I (\sigma_0, \omega_{\text{max}}) \).

Similarly, we have

\[
\min_{ \sigma \in [\sigma_{\text{min}}, \sigma_1] } \min_{ \omega \in [0, \omega_{\text{max}}] } \xi_R (\sigma, \omega) = \xi_R (\bar{\sigma}, 0) = \bar{\sigma} + \sqrt{\bar{\sigma}^2 - 1},
\]

(88)

and based on this relation, it is reasonable to consider the following equation:

\[
\min_{ \eta > 1 } \max_{ x \in \mathbb{D} (\mathfrak{S}_{\text{min}}, \mathfrak{S}_{\text{max}}) } \left| \frac{\eta - \mathfrak{S}_{\text{min}} (\bar{\sigma})}{\eta s - 1} \right|^2 \leq \max_{ x \in \mathbb{D} (\mathfrak{S}_{\text{max}}, \mathfrak{S}_{\text{max}}) } \left| \frac{\eta - s}{\eta s - 1} \right|^2.
\]

(89)

To analyze the solution of this min-max problem, we define

\[
\kappa(x, y) = \frac{Z + \sqrt{Z^2 - 4}}{2}, \quad \text{with} \quad Z = \frac{1 + x^2 + y^2}{\bar{x}},
\]

(91a)

\[
\eta_+ = \kappa(\sigma_{\text{max}}, 0), \quad \eta_+ = \kappa(\sigma_{\text{min}}, \omega_{\text{max}}), \quad \mathfrak{T}_+ (\sigma, \omega, \eta) = \frac{(\sigma - \eta)^2 + \omega^2}{(\sigma \eta - 1)^2 + \eta^2 \omega^2}.
\]

Then, for \( \eta_{\text{min}} < \eta_+ < \eta_{\text{max}} \), we defined two functions \( \mathcal{T}_- (\eta) \) and \( \mathcal{T}_+ (\eta) \) as

\[
\kappa(x, y) = \frac{Z + \sqrt{Z^2 - 4}}{2}, \quad \text{with} \quad Z = \frac{1 + x^2 + y^2}{\bar{x}},
\]

(91a)

\[
\eta_+ = \kappa(\sigma_{\text{max}}, 0), \quad \eta_+ = \kappa(\sigma_{\text{min}}, \omega_{\text{max}}), \quad \mathfrak{T}_+ (\sigma, \omega, \eta) = \frac{(\sigma - \eta)^2 + \omega^2}{(\sigma \eta - 1)^2 + \eta^2 \omega^2}.
\]
Theorem 2. Parameter optimization at the discrete level: let \( L = \Delta x \), \( \alpha, \beta \in (0, 1) \) and \( \mu \geq 0 \). Then, the parameter \( p \) involved in the discrete SWR algorithms can be chosen as \( p^*_d = (\eta^*_\text{opt} - 1)/\Delta x \), where \( \eta^*_\text{opt} \) is the solution of min-max problem (90), is given by

\[
\eta^*_\text{opt} = \begin{cases} 
\eta_{\text{min}}, & \text{if } \mathcal{F}_+ (\eta_{\text{min}}) \geq \mathcal{F}_- (\eta_{\text{max}}), \\
\eta_{\text{max}}, & \text{if } \mathcal{F}_- (\eta_{\text{max}}) \geq \mathcal{F}_+ (\eta_{\text{max}}), \\
\eta', & \text{otherwise}, 
\end{cases}
\]

where \( \eta_{\text{min}} = \kappa (\sigma_{\text{min}}, 0) \), \( \eta_{\text{max}} = \kappa (\sigma_{\text{max}}, \omega_{\text{max}}) \), and \( \eta' \in (\eta_{\text{min}}, \eta_{\text{max}}) \) are the unique root of \( \mathcal{F}_+ (\eta) = \mathcal{F}_- (\eta) \). With the parameter \( p_\beta \), the convergence factor \( p_d = \max_{\sigma \in D} |\rho_{\beta}(\sigma)| \) at the discrete level satisfies

\[
|\rho_d| \leq \max_{\sigma \in D} |\rho_{\beta}(\sigma)| \leq \max_{\sigma \in D} |\rho_{\beta}(\sigma)|, \quad \text{where } \rho_{\beta}(\sigma) \text{ is given by (78a)}. 
\]

Proof. Let \( \eta = \sigma + i\omega \in D(\sigma_{\text{min}}, \sigma_{\text{max}}, \omega_{\text{max}}) \). Then, it holds that \( |\eta - \eta_0| = |\eta - \eta_0 - 1| \geq \mathcal{F}(\sigma, \omega, \eta) \). For \( \eta > 1 \), a routine calculation yields

\[
\text{sign}(\partial_\eta \mathcal{F}(\sigma, \omega, \eta)) = \text{sign}(\eta - \kappa (\sigma, \omega)).
\]

Hence, the solution of (90) should satisfy \( \eta^*_\text{opt} \in [\eta_{\text{min}}, \eta_{\text{max}}] \), since otherwise, if \( \eta^*_\text{opt} < \eta_{\text{min}} \) (resp. \( \eta^*_\text{opt} > \eta_{\text{max}} \)), we can further minimize \( \mathcal{F} \) by increasing (resp. decreasing) \( \eta^*_\text{opt} \).

We have \( \text{sign}(\partial_\sigma \mathcal{F}(\sigma, \omega, \eta)) = \text{sign}(2\sigma - 1 - \eta^2) \), and therefore, for fixed \( \kappa \), it holds that \( \max_{\eta \in [0, \omega_{\text{max}}]} \mathcal{F}(\sigma, \omega) = \max_{\eta \in [0, \omega_{\text{max}}]} \mathcal{F}(\sigma, \omega, \eta) \). Hence, \( \max_{\eta \in D(\sigma_{\text{min}}, \sigma_{\text{max}}, \omega_{\text{max}})} \mathcal{F}(\sigma, \omega, \eta) = \max_{\eta \in D(\sigma_{\text{min}}, \sigma_{\text{max}}, \omega_{\text{max}})} \mathcal{F}(\sigma, \omega, \eta) \). Another routine calculation yields

\[
\partial_\sigma \mathcal{F}(\sigma, \omega, \eta) = \frac{2\eta^2 - 2\sigma - 1 + \eta^2 + \eta^2\omega^2}{\eta(\sigma - 1)^2 + \sigma^2}.
\]

(94)

Then, for fixed \( \eta > 1 \) and \( \omega \geq 0 \), if we vary \( \sigma \) from 1 to \( +\infty \), we first meet \( \partial_\sigma \mathcal{F}(\sigma, \omega, \eta) < 0 \), and then \( \partial_\sigma \mathcal{F}(\sigma, \omega, \eta) > 0 \). Hence, for fixed \( \omega \), the function \( \mathcal{F}(\sigma, \omega, \eta) \) does not have local maximum(s) for \( \sigma \in [\sigma_{\text{min}}, \sigma_{\text{max}}] \). This implies that for given \( \eta > 1 \) the global maximum of \( \mathcal{F} \) on the region \( D \) is obtained at the four corners, i.e.,

\[
\max_{\sigma \in D} \mathcal{F}(\sigma, \omega, \eta) = \max_{\sigma \in D} \mathcal{F}(\sigma_{\text{min}}, \sigma_{\text{max}}, \omega_{\text{max}}) = \max_{\sigma \in D} \mathcal{F}(\sigma_{\text{max}}, \sigma_{\text{max}}, \omega_{\text{max}}).
\]

(95)

With the functions \( \mathcal{F}_+ (\eta) \) defined by (91b) and (91c), we have

\[
\min_{\eta \in \mathbb{R}^+} \min_{[\eta_{\text{min}}, \eta_{\text{max}}]} \max_{\sigma \in D(\sigma_{\text{min}}, \sigma_{\text{max}}, \omega_{\text{max}})} \mathcal{F}(\sigma, \omega, \eta) = \min_{\eta \in \mathbb{R}^+} \min_{[\eta_{\text{min}}, \eta_{\text{max}}]} \max_{\sigma \in \mathbb{R}^+} \mathcal{F}_+ (\eta, \mathcal{F}_- (\eta)).
\]

(96)
It remains to analyze the monotonic properties of $\mathcal{T}_+ (\eta)$ and $\mathcal{T}_- (\eta)$. From (93), we know that $\mathcal{T}_+ (\sigma_{\min}, 0, \eta)$ and $\mathcal{T}_- (\sigma_{\max}, 0, \eta)$ are, respectively, increasing and decreasing for $\eta \in [\eta_{\min}, \eta_{\max}]$. Similarly, $\mathcal{T}_+ (\sigma_{\max}, 0, \eta)$ is decreasing for $\eta \in [\eta_{\min}, \eta_{\max}]$; $\mathcal{T}_- (\sigma_{\max}, 0, \eta)$ is decreasing for $\eta \in [\eta_{\min}, \eta_{\max}]$ and increasing for $\eta \in [\eta_{\min}, \eta_{\max}]$. Hence, the functions $\mathcal{T}_+ (\eta)$ and $\mathcal{T}_- (\eta)$, defined by (91b) and (91c), are, respectively, increasing and decreasing for $\eta \in [\eta_{\min}, \eta_{\max}]$. This, together with some simple algebra, gives the solution of (96), i.e., $\eta_{\text{opt}}$, as shown in (92).

Remark 3. The optimization procedure at the discrete level consists of two steps:

1. For specified quotient $\Lambda (\zeta)$ and problem/discretization parameters, compute the boundaries of the rectangular region $\Omega$, i.e., $\sigma_{\min}, \sigma_{\max}$, and $\omega_{\max}$
2. Compute $\eta_{\text{opt}}$ according to formula (92) and then let $P_{\text{opt}} = (\eta_{\text{opt}} - 1)/\Delta x$

The quotient $\Lambda (\zeta)$ associated with a specified time-integrator determines the boundaries of $\Omega$, which further determines the parameter $\eta_{\text{opt}}$. To make a convergent SWR algorithm, $\Lambda (e^{-\Delta t})$ should lay in $C^+$ and should be away from the origin, $\forall s \in \{ s \in C^+ : |\Im (s)| \in [\pi/T, (\pi/\Delta t)] \}$.

Remark 4. Large overlap size: if we use an overlap size $L = M \Delta x$ with $M \geq 2$, the following min-max problem occurs:

$$\min_{\eta = 1} \max_{\sigma \in \Omega} \mathcal{T}(s, \sigma)$$

with $\mathcal{T}(s, \sigma) = \frac{\eta - s}{e^{\sigma \Delta t} (|s - \eta|)}$.

(97)

For $M \geq 2$ and $\eta > 1$, the function $\mathcal{T}(s, \sigma)$ has local maximums in the interior of $\Omega$, and therefore, the analysis is much more complicated compared to the case $M = 1$. The proof of Theorem 2 cannot be straightforwardly generalized to (97), and we will investigate this general min-max problem in other places.

4.3. An Application in the Field of Fractional Circuits. In the last part of this section, we show that the analysis performed at the discrete level can be directly applied to fractional diffusive (RC) circuits. In recent years, there is an increasing interest in the fractional order circuits, thanks to the fact that they can more precisely describe the dynamic behavior of the supercapacitors (also called ultracapacitors or electrochemical double-layer capacitors); see the recent papers [46,47] for more details.

For the model circuit shown in Figure 5, consisting of infinite number of series connected cell-circuits, the state (or modified-nodal-analysis) circuit equation is

$$\frac{d^b x(t)}{dt^b} = A x(t) + f(t),$$

$$A = \begin{pmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ & & \ddots \end{pmatrix},$$

(98)

where $x(t) = (\cdots, x_{-1}(t), x_0(t), x_1(t), \cdots)^T$ denotes the vector of the nodal voltage values, $c = (1/\sqrt{RC})$ with $C$ the value of the fractional capacitor and $R$ the value of the resistance, as shown in Figure 5(a). The fractional derivative in (98) is defined in the Caputo sense (see (8)), and the argument $b \in (0, 1)$ is determined by the branches of the circuit shown on the right. Mathematically, (98) can be equivalently rewritten as $x(t) = F_b^\Lambda (\mathcal{A}(x)) + f(t)$, with $F_t^\Lambda$ the fractional integral defined by (50) and some other function $f$. Clearly, this is a special case of (4), which can be seen by letting $1 - \beta = b$ and $\mu = 0$ and then by replacing $U^b_{\text{int}} (t)$ by $x_m(t)$ and $\delta_x^b$ by $A$ there. Therefore, the semi-discrete SWR algorithm (51a) and (51b) corresponds to the waveform relaxation method for (98).

For an analysis of the continuous WR algorithm without time discretization, i.e., (51a) and (51b), we will also arrive at (78a), but with

$$\tilde{s} = s \in \mathbb{H} := \{ s \in C^+ : |\Im (s)| \in \left[ \frac{\pi}{T}, \frac{\pi}{\Delta t} \right] \},$$

(99)

instead of $\tilde{s} = \Lambda (e^{-\Delta t})/\Delta t$ as shown in (78b). Hence, the rectangular region $\mathcal{D}$ covering $\xi$, (3) is a semi-infinite horizontal stripe, i.e., $\sigma_{\max} = +\infty$ in (80). In this limit case, Theorem 2 is still applicable, but we need to replace $\eta_{\text{max}}$ by $+\infty$ there and redefine $\mathcal{T}_+ (\eta)$ as

$$\mathcal{T}_+ (\eta) = \frac{\max \{ \mathcal{T}(\sigma_{\min}, 0, \eta), \mathcal{T}(\sigma_{\max}, 0, \eta) \} \mathcal{T}(\sigma_{\min}, \sigma_{\max}, \eta) \} \eta > \eta_1,$$

$$\mathcal{T}_- (\eta) = \frac{\max \{ \mathcal{T}(\sigma_{\min}, 0, \eta), \mathcal{T}(\sigma_{\min}, \sigma_{\max}, \eta) \} \eta > \eta_1,$$

(100)

where $\eta_1$ and $\eta_{\text{min}}$ are the same quantities as we have stated in Theorem 2.

5. Numerical Results

In this section, we test and compare the performance of the parameters analyzed at the continuous and discrete level. Here, we restrict ourselves to the case $L = \Delta x$. Since the parameter $p$ appearing in the continuous Robin TCs is transformed to $\eta = 1 + \Delta x \rho$ in the discrete circumstance, we shall only mention $\eta$ throughout this section.
At the moment, for given problem/discretization parameters, we have two choices of $\eta$: $\eta = 1 + (\Delta x \eta_{\text{opt}} / (2L)) = 1 + (\eta_{\text{opt}} / 2)$ from the continuous analysis (i.e., Lemma 4) and $\eta = \eta_{\text{opt}}$ from the discrete analysis (i.e., Theorem 2). Both are approximate values of the best one that we can make through numerical optimizations (as described in Remark 2). To compare the performance of these two parameters, in Figure 6 we show the convergence factor $\rho_\ell = \max_{0 \leq j \leq L} |\rho_j|$ based on three choices of $\eta$, where $\rho_\ell$ and $\|\|$ are given in (78a) and (78b). Here, the problem/discretization parameters are chosen as $\mu = 4.8$, $T = 18$, $\Delta t = 0.025$, and $\Delta x = 0.04$. To get $\eta_{\text{opt}}$ from Theorem 2, the boundaries of the rectangular region $D$, i.e., $\bar{\sigma}_{\text{min}}, \bar{\sigma}_{\text{max}}$, and $\bar{\sigma}_{\text{max}}$, are obtained through numerical optimizations (for the case $\alpha \geq \beta$, if we use the explicit expressions (81a)–(81c) to get these three quantities, the corresponding convergence factor looks similar as what we observed in Figure 6).

From Figure 6, we see that for the backward Euler (or BDF2) method, the convergence factors under the three choices of $\eta$ are almost equal. For a specified choice of $\eta$, the convergence factors associated with the two time-integrators are almost equal. These can be regarded as theoretical predictions, and we now provide numerical results for validation. We consider the following equation with zero initial and boundary values:

\[ u(x,t) - J^{1-\beta}_t \frac{\partial^2 u}{\partial x^2} + \mu J^{1-\alpha}_t u(x,t) = \sin(x\sqrt{t}), \quad (x,t) \in (0,4) \times (0,18), \tag{101} \]

which is equivalent to model problem (5) after an integral transform. We choose $\alpha = 0.84$, $\beta = 0.5$, and $\mu = 4.8$ for the problem parameters, and $\Delta t = 0.025$ and $\Delta x = 0.04$ for the discretization parameters. For each experiment, the initial iterate is chosen randomly.

In Figure 7(a), we plot the convergence rates of the SWR algorithms, using $\eta = 1.12031$ from the continuous analysis. (For the parameters obtained from the discrete analysis, this error plot looks similar.) The three lines shown in this figure correspond to three time-integrators, trapezoidal (dot line),
backward Euler (dot-dash line), and BDF2 (solid line). We see that (1) the convergence rates associated with backward Euler and BDF2 are very close and (2) both are significantly better than that of the trapezoidal method. The first observation is consistent with Figure 6, and the second one is consistent with our analysis at the discrete level, since for the trapezoidal method, we always have $\max_{\eta \in [1]} |\tilde{E}_\eta| = 1$, independent of $\eta$. Then, in Figure 7, we verify whether the theoretically analyzed parameter is close to the best numerical one or not. Precisely, we show the error obtained by running the SWR algorithm after 4 iterations using various choices of $\eta$.

FIGURE 7: (a) Measured convergence rates of the SWR algorithms for three time-integrators, using the parameter $\eta = 1.12031$ obtained from the continuous analysis. (b) For BDF2 and (c) backward Euler, the error after 4 iterations of the SWR algorithms using various choices of $\eta$ for the Robin TCs; the parameters analyzed at the continuous and discrete level are denoted by a circle and a star.

At the end of this section, we show in Figure 8 the effects of the discretization parameters and the length of time-interval on the convergence rate of the SWR algorithms.

\begin{itemize}
  \item Trapezoidal
  \item BDF2
  \item Backward euler
\end{itemize}

\begin{itemize}
  \item $\eta = 1.12031$ (from continuous analysis)
  \item $\eta = 1.12031$ (from discrete analysis)
\end{itemize}
using the backward Euler method as the time-integrator in the case of two and eight subdomains. For each \( \Delta t \) and \( T \), we use two choices for the parameter \( \eta \) involved in the Robin TCs, the one \( \eta = 1 + \Delta xq_{\text{opt}} \) from the continuous analysis and the one \( \eta = \eta_{\text{opt}} \) from the numerical optimization. (For the parameter \( \eta = \eta_{\text{opt}} \) from the discrete analysis, the plots look similar.) For each \( \Delta t \) and \( T \), the iteration number is measured when the error between the iterate and the reference solution is less than \( 10^{-8} \). The reference solution is obtained by applying the same discretization to (101) on the whole space-time domain. We see that the SWR method has very similar asymptotic convergence rate, depending on the discretization parameters and the length of time-interval, in the case of two and many subdomains. We can also see that the convergence rates of the SWR method using the theoretically analyzed parameter and the one from the numerical optimization are very similar in the two subdomain case, while for the eight-subdomain case, these two Robin parameters lead to slightly different outcomes. This is natural since the Robin parameter \( \eta = 1 + \Delta xq_{\text{opt}} \) is analyzed in the two-subdomain case, and in the many subdomain case, an optimal choice of \( \eta \) should depend on the number of subdomains (see, e.g., [48]).

6. Conclusions

We studied the Schwarz waveform relaxation (SWR) algorithms with Robin transmission conditions (TCs) for a class of representative time-fractional PDEs. The problem of determining the parameter for the Robin TCs was investigated at the continuous and discrete level. Particularly, at the continuous level, the parameter obtained by solving a transformed min-max problem (i.e., (36a) and (36b)) is found much better than the one predicted by the equioscillation principle [30,35]. The solution of the transformed min-max problem is not always determined by the equioscillation principle.

At the discrete level, we used the centered finite difference scheme to discretize the spatial variable and then we used the convolution quadratures [38–40], based on linear multistep methods, to get time discretizations. This framework permits us to treat the time discretizations as a whole; that is, the discrete Laplace transform of the sequences generated by the quadratures has a uniform expression, which is mainly determined by the quotient associated with a specified time-integrator. Three time-integrators were carefully studied (i.e., backward Euler, trapezoidal, and BDF2), and it was shown that the trapezoidal method leads to a convergence factor equaling 1 (independent of the Robin parameter) and that for the backward Euler and BDF2 methods, the convergence factors are (almost) equal and both are satisfactory.

Based on the work in this paper, our ongoing research is twofold. First, we attempt to generalize the two-subdomain analysis to the multisubdomain case; this is a necessary step to make the SWR algorithm practical for real parallel computation. Second, we try to solve the general min-max problem (97) (i.e., large overlap size \( L = M \Delta x \) with \( M \geq 2 \)) and perform an asymptotic analysis for the convergence factor with respect to the discretization parameters \( \Delta t \) (and \( \Delta x \)) and \( T \), the length of the time-interval.
Appendix.

A. The Proof of Lemma 4

Proof. For the function \( R(y, q) \) defined by (36b), we have
\[
\partial_q R(y, q) = 4y \frac{q^2 - (1 + \phi)y^2}{\left[(y + q)^2 + \phi y^2\right]^2} - y.
\]
(A.1)

Hence, for \( y \in [y_0, y_1] \), the solution \( q_{opt} \) of (36a) satisfies \( \sqrt{1 + \phi} y_0 \leq q_{opt} \leq \sqrt{1 + \phi} y_1 \), since otherwise, if \( q_{opt} < \sqrt{1 + \phi} y_0 \) (resp. \( q_{opt} > \sqrt{1 + \phi} y_1 \)), we can further uniformly reduce \( R \) by increasing (resp. decreasing) the value of \( q_{opt} \). Hence, it suffices to study
\[
\min_{\mathcal{D}} \max \{ R(y, q) \}
\]
(A.2)

where \( q_0 = y_0 \sqrt{1 + \phi} \) and \( q_1 = y_1 \sqrt{1 + \phi} \).

A routine calculation shows that, for a given \( q \in [q_0, q_1] \), the function \( R \) has at most two local extremums, located at
\[
\mathcal{T}_+(q) = \frac{1}{1 + \phi} \sqrt{(1 - \phi)q^2 + 2(1 + \phi)q ± 2q\sqrt{d(q)}},
\]
(A.3)

where \( d(q) = -\phi q^2 - 2\phi(1 + \phi)q + (1 + \phi)^2 \) is the discriminant. Since \( R(y, q) \to 0 \) as \( y \to +\infty \), the larger one \( \mathcal{T}_+(q) \) must be a local maximizer. This local maximizer does not always exist, and by letting \( d(q) > 0 \), we know that it exists for \( q \in (0, q_d) \), where \( q_d = (1 + \phi)\left(\sqrt{1 + (1/\phi)} - 1\right) \). By using \( R(+\infty, q) = 0 \) and \( R(0, q) = 1 \), we know that for \( q = q_d \), i.e., \( d(q) = 0 \), the function \( R \) is strictly decreasing for \( y > 0 \). \( \square \)

Case A. If \( y_0 \geq \sqrt{1 + \phi} \left(\sqrt{1 + 1/\phi} - 1\right) \). In this case, we have \( d(q) \leq 0 \) for \( q \in [q_0, q_1] \). Hence, \( R \) is increasing (resp. decreasing) function of \( q \) when \( q \in [q_0, q_1] \). Then, it can be shown that for \( q \in [q_0, q_1] \), the function \( R \) attains its global minimum at \( q = q_1 \), where \( q_1 \) is given by (37b); this implies \( q_{opt} = q_1 \). We have
\[
R_1(q) = \max \{ R(y_0, q), R(y_1, q) \} = \begin{cases} R(y_1, q), & q \in \left[q_0, q_1\right), \\ R(y_0, q), & q \in (q_1, q_1]. \end{cases}
\]
(A.4)

where for any real numbers \( I_1 \) and \( I_2 \), the set \( I_1 \cup I_2 \) (or \( (I_1, I_2) \)) is empty if \( I_1 \geq I_2 \).

Case B. \( y_0 < \sqrt{1 + \phi} \left(\sqrt{1 + 1/\phi} - 1\right) < y_1 \). In this case, it is clear that \( q_1 > q_d > q_0 \). We claim \( \mathcal{T}_+(q) < y_0 \) for \( q \in [q_d, q_d] \). Note that, for any \( q \in [q_0, q_1] \), we have
\[
\mathcal{T}_+(q) > y_0 \iff (1 - \phi)q^2 + 2(1 + \phi)q + 2q\sqrt{d(q)} > (1 + \phi)^2 y_0^2.
\]
(A.5)

Hence, it is sufficient to prove
\[
(1 - \phi)q_d^2 + 2(1 + \phi)q_d + 2q_d\sqrt{d(q_d)} > (1 + \phi)^2 y_0^2.
\]
(A.6)

it is sufficient to prove \( \sqrt{1 + \phi} \left(\sqrt{1 + 1/\phi} - 1\right) < \sqrt{1 + \phi} / \phi \). By noticing
\[
\forall \phi \in (0, 1), \quad \sqrt{1 + \phi} < \sqrt{1 + \phi} \left(\sqrt{1 + (1/\phi)} - 1\right) < \sqrt{1 + \phi} / \phi.
\]
(A.7)

we finally get \( \mathcal{T}_+(q) > y_0 \) for \( q \in [q_0, q_d] \). Now, by using \( d(q_d) = 0 \) and the continuity of \( \max_{y \in [y_0, y_1]} R(y, q) \) with respect to \( q \), it is easy to get
\[
R(q_d) \leq R(1, q_d).
\]
(A.8)

We next claim \( dR(q)/dq < 0 \) for \( q \in [q_0, q_d] \). From (A.1), we have
\[
\frac{dR(q)}{dq} = \frac{\partial R(\mathcal{T}_+(q), q)}{\partial q} + \frac{\partial R(\mathcal{T}_+(q), q)}{\partial y} \times \frac{d\mathcal{T}_+(q)}{dq}.
\]
(A.9)

where we used the fact that \( \mathcal{T}_+ \) is the local maximizer of \( R \), that is \( (\partial R(\mathcal{T}_+, q)/\partial y) = 0 \). Then, it holds that
\[
q - \sqrt{1 + \phi} \mathcal{T}_+(q) < 0 \iff \phi q - 1 - \phi < \sqrt{d(q)} \iff \phi q - 1 - \phi < 0 = \phi q_qd - 1 - \phi < 0,
\]
(A.10)
where the last inequality holds for any \( \phi \in (0, 1) \), because
\[
\phi d - 1 - \phi < 0 \iff \phi \left( \frac{1}{1 + \phi} - 1 \right) - 1 < 0 \iff \phi < \sqrt{1 + \phi}.
\]

(A.11)

Therefore, \( d \mathcal{R}(q)/dq < 0 \) and this implies that the function \( \mathcal{R}(q) \) is decreasing for \( q \in [q_d, q_1] \). Based on the above analysis, we next consider the following two cases:

1. \( q_1 < q_d, \ y(q_1) < y_1, \text{ and } \mathcal{R}(q_1) > \mathcal{R}^+(q_1) \). In this case, it must hold \( q_1 < q_\ast \), since otherwise we have \( q_d > q_1 \) and this leads to contradiction. Hence, it holds that \( \mathcal{R}^+(q) = \mathcal{R}(y_0, q) \) for \( q \in [q_0, q_1] \). Then, by using \( \mathcal{R}(q_1) > y_1, \mathcal{R}(q_1) > \mathcal{R}^+(q_1) \), and (A.8), we know that the equation \( \mathcal{R}(y, q_1) = \mathcal{R}(q) \) has a unique root \( q_\ast \in (q_0, q_1) \) and that \( \mathcal{R}\max(q) (= \max_{y \in [y_0, y_1]} \mathcal{R}(y, q)) \) attains its global minimum at \( q_\ast \), because \( \mathcal{R}(q) \) and \( \mathcal{R}(y_0, q) \) are, respectively, decreasing and increasing functions of \( q \).

2. If any one of the conditions in 1 is not satisfied, we claim
\[
\max_{y \in [y_0, y_1]} \mathcal{R}(y, q_1) = \mathcal{R}^+(q_1).
\]

(A.12)

Indeed, this can be proved by using (A.8) and the fact that “both \( \mathcal{R}(q) \) and \( \mathcal{R}(y_1, q) \) are decreasing functions, but \( \mathcal{R}(y_0, q) \) is an increasing function.” For example, if \( q_1 = q_0, \) by using (A.8), we know that (A.12) holds. For the case \( q_1 > q_0 \) or \( \mathcal{R}(q_1) \geq y_1, \) (A.12) can be proved similarly. Now, let \( Q = \{ q: \mathcal{R}(q) \} \) and \( Q_0 = [q_0, q_1] \). Then, we have

\[
\min_{q \in Q_0} \max_{y \in [y_0, y_1]} \mathcal{R}(y, q) = \min_{q \in Q_0} \max_{y \in [y_0, y_1]} \mathcal{R}(y, q) = \min_{q \in Q_0} \mathcal{R}^+(q) = \mathcal{R}^+(q_1).
\]

(A.13)

**Conflicts of Interest**

It is declared that there are no conflicts of interest among the authors regarding this manuscript.

**Acknowledgments**

This paper was supported financially by the National Science and Technology Major Project of China (no. 2016ZX05065 and no. 2016ZX05042-003), by the Sichuan Science and Technology Program (2019Y0541), and by the Opening Project of Sichuan Province University Key Laboratory of Bridge Non-Destruction Detecting and Engineering Computing (2019QZJ03).

**References**

[1] R. Metzler and J. Klafter, “The random walk’s guide to anomalous diffusion: a fractional dynamics approach,” *Physics Reports*, vol. 339, no. 1, pp. 1–77, 2000.

[2] V. Balakrishnan, “Anomalous diffusion in one dimension,” *Physica A: Statistical Mechanics and Its Applications*, vol. 132, no. 2-3, pp. 569–580, 1985.

[3] W. Wyss, “The fractional diffusion equation,” *Journal of Mathematical Physics*, vol. 27, no. 11, pp. 2782–2785, 1986.

[4] W. R. Schneider and W. Wyss, “Fractional diffusion and wave equations,” *Journal of Mathematical Physics*, vol. 30, no. 1, pp. 134–144, 1989.

[5] J.-P. Bouchaud and A. Georges, “Anomalous diffusion in disordered media: statistical mechanisms, models and

**Data Availability**

The data used to support the findings of this study are included within the manuscript.
physical applications,” *Physics Reports*, vol. 195, no. 4-5, pp. 127–293, 1990.

[6] E. K. Lenzi, R. S. Mendes, K. S. Fa, L. C. Malacarne, and L. R. da Silva, “Anomalous diffusion: fractional Fokker-Planck equation and its solutions,” *Journal of Mathematical Physics*, vol. 44, no. 5, pp. 2179–2185, 2003.

[7] E. R. Weeks, *Experimental studies of anomalous diffusion, blocking phenomena, and two-dimensional turbulence*, Ph.D. thesis, University of Texas, Austin, TX, USA, 1997.

[8] J. Daniel Djida, J. J. Nieto, and I. Area, “Parabolic problem with fractional time derivative with nonlocal and nonsingular Mittag-Leffler kernel,” *Discrete & Continuous Dynamical Systems*, vol. 13, no. 3, pp. 609–627, 2020.

[9] Z.-J. Fu, S. Reutskiy, H.-G. Sun, J. Ma, and M. A. Khan, “A robust kernel-based solver for variable-order time fractional PDEs under 2D/3D irregular domains,” *Applied Mathematics Letters*, vol. 94, pp. 105–111, 2019.

[10] Z. J. Fu, W. Chen, and H. T. Yang, “Boundary particle method for Laplace transformed time fractional diffusion equations,” *Journal of Computational Physics*, vol. 235, pp. 52–66, 2013.

[11] W. Hu, C. Ruan, H. Nian, and D. Sun, “Zero-sequence current suppression strategy with common-mode voltage control for open-end winding PMSM drives with common DC bus,” *IEEE Transactions on Industrial Electronics*, vol. 68, no. 6, pp. 4691–4702, 2021.

[12] N. Huy Tuan, D. O’Regan, and T. Bao Ngoc, “Continuity with respect to fractional order of the time fractional diffusion-wave equation,” *Evolution Equations & Control Theory*, vol. 9, no. 3, pp. 773–793, 2020.

[13] M. Usman, M. Hamid, Z. H. Khan, and R. U. Haq, “Neuronal dynamics and electrophysiology fractional model: a modified wavelet approach,” *Physica A: Statistical Mechanics and Its Applications*, vol. 570, p. 125805, 2021.

[14] R. Wang, Q. Sun, W. Hu, Y. Li, D. Ma, and P. Wang, “SoC-based droop coefficients stability region analysis of the battery for stand-alone supply systems with constant power loads,” *IEEE Transactions on Power Electronics*, vol. 36, no. 7, pp. 7866–7879, in press, 2021.

[15] R. Wang, Q. Sun, D. Ma, and Z. Liu, “The small-signal stability analysis of the droop-controlled converter in electromagnetic timescale,” *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1459–1469, 2019.

[16] R. Wang, Q. Sun, D. Ma, and X. Hu, “Line impedance co-operative stability region identification method for grid-tied inverters under weak grids,” *IEEE Transactions on Smart Grid*, vol. 11, pp. 2856–2866, 2020.

[17] D. Zifabec, B. Lwy, X. B. Qiang, and B. Csl, “A boundary collocation method for anomalous heat conduction analysis in functionally graded materials,” *Computers & Mathematics with Applications*, vol. 88, pp. 91–109, 2020.

[18] B. I. Henry, T. A. M. Langlands, and S. L. Wearne, “Fractional cable models for spiny neuronal dendrites,” *Physical Review Letters*, vol. 100, no. 12, p. 128103, 2008.

[19] T. A. M. Langlands, B. I. Henry, and S. L. Wearne, “Fractional cable equation models for anomalous electrodiffusion in nerve cells: finite domain solutions,” *SIAM Journal on Applied Mathematics*, vol. 71, no. 4, pp. 1168–1203, 2011.

[20] T. A. M. Langlands, B. I. Henry, and S. L. Wearne, “Fractional cable equation models for anomalous electrodiffusion in nerve cells: infinite domain solutions,” *Journal of Mathematical Biology*, vol. 59, no. 6, pp. 761–808, 2009.

[21] X.-C. Cai, “Additive Schwarz algorithms for parabolic convection-diffusion equations,” *Numerische Mathematik*, vol. 60, no. 1, pp. 41–61, 1991.

[22] X.-C. Cai, “Multiplicative Schwarz methods for parabolic problems,” *SIAM Journal on Scientific Computing*, vol. 15, no. 3, pp. 587–603, 1994.

[23] M. J. Gander and A. M. Stuart, “Space-time continuous analysis of waveform relaxation for the heat equation,” *SIAM Journal on Scientific Computing*, vol. 19, no. 6, pp. 2014–2031, 1998.

[24] E. Gilardi and H. B. Keller, “Space-time domain decomposition for parabolic problems,” *Numerische Mathematik*, vol. 93, no. 2, pp. 279–313, 2002.

[25] X.-M. Gu and S.-L. Wu, “A parallel-in-time iterative algorithm for Volterra partial integro-differential problems with weakly singular kernel,” *Journal of Computational Physics*, vol. 417, p. 109576, 2020.

[26] S.-L. Wu and T. Zhou, “Parallel implementation for the two-stage SDIRK methods via diagonalization,” *Journal of Computational Physics*, vol. 428, p. 110076, 2021.

[27] S.-L. Wu, T. Zhou, and X. Chen, “A gauss-seidel type method for dynamic nonlinear complementarity problems,” *SIAM Journal on Control and Optimization*, vol. 58, no. 6, pp. 3389–3412, 2020.

[28] S.-L. Wu, “Toward parallel coarse grid correction for the parareal algorithm,” *SIAM Journal on Scientific Computing*, vol. 40, no. 3, pp. A1446–A1472, 2018.

[29] S.-L. Wu and T. Zhou, “Acceleration of the two-level MGRIT algorithm via the diagonalization technique,” *SIAM Journal on Scientific Computing*, vol. 41, no. 5, pp. A3421–A3448, 2019.

[30] D. Bennequin, M. J. Gander, and L. Halpern, “A homographic best approximation problem with application to optimized Schwarz waveform relaxation,” *Mathematics of Computation*, vol. 78, no. 265, p. 185, 2009.

[31] D. Bennequin, M. J. Gander, L. Gouarin, and L. Halpern, “Optimized Schwarz waveform relaxation for advection reaction diffusion equations in two dimensions,” *Numerische Mathematik*, vol. 134, no. 3, pp. 513–567, 2016.

[32] M. E. Bouajaji, V. Dolean, M. J. Gander, and S. Lanteri, “Optimized Schwarz methods for the time-harmonic Maxwell equations with damping,” *SIAM Journal on Scientific Computing*, vol. 34, no. 4, pp. A2048–A2071, 2012.

[33] V. Dolean, M. J. Gander, S. Lanteri, J.-F. Lee, and Z. Peng, “Effective transmission conditions for domain decomposition methods applied to the time-harmonic curl-curl Maxwell’s equations,” *Journal of Computational Physics*, vol. 280, pp. 232–247, 2015.

[34] V. Dolean, M. J. Gander, and L. Gerardo-Giorda, “Optimized Schwarz methods for Maxwell’s equations,” *SIAM Journal on Scientific Computing*, vol. 31, no. 3, pp. 2193–2213, 2009.

[35] M. J. Gander and L. Halpern, “Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems,” *SIAM Journal on Numerical Analysis*, vol. 45, no. 2, pp. 666–697, 2007.

[36] S.-L. Wu and Y. Xu, “Convergence analysis of Schwarz waveform relaxation with convolution transmission conditions,” *SIAM Journal on Scientific Computing*, vol. 39, no. 3, pp. A890–A921, 2017.

[37] S.-L. Wu, “Optimized overlapping Schwarz waveform relaxation for a class of time-fractional diffusion problems,” *Journal of Scientific Computing*, vol. 72, no. 2, pp. 842–862, 2017.

[38] C. Lubich, “Convolution quadrature and discretized operational calculus. I,” *Numerische Mathematik*, vol. 52, no. 2, pp. 129–145, 1988.
[39] C. Lubich, “Convolution quadrature and discretized operational calculus. II,” Numerische Mathematik, vol. 52, no. 4, pp. 413–425, 1988.
[40] C. Lubich, “Convolution quadrature revisited,” BIT Numerical Mathematics, vol. 44, no. 3, pp. 503–514, 2004.
[41] I. Podlubny, Fractional Differential Equations, Academic Press, New York, NY, USA, 1999.
[42] X. Li and C. Xu, “A space-time spectral method for the time fractional diffusion equation,” SIAM Journal on Numerical Analysis, vol. 47, no. 3, pp. 2108–2131, 2009.
[43] S.-L. Wu and M. D. Al-Khaleel, “Semi-discrete Schwarz waveform relaxation algorithms for reaction diffusion equations,” BIT Numerical Mathematics, vol. 54, no. 3, pp. 831–866, 2014.
[44] A. Schädle, M. López-Fernández, and C. Lubich, “Fast and oblivious convolution quadrature,” SIAM Journal on Scientific Computing, vol. 28, no. 2, pp. 421–438, 2006.
[45] S. Vandewalle and M. J. Gander, “Optimized overlapping Schwarz methods for parabolic PDEs with time-delay,” Lecture Notes in Computer Science, vol. 40, pp. 291–298, 2005.
[46] W. Mitkowski and P. Skruch, “Fractional-order models of the supercapacitors in the form of RC ladder networks,” Bulletin of the Polish Academy of Sciences: Technical Sciences, vol. 52, pp. 109–114, 2013.
[47] D. Sierociuk, I. Podlubny, and I. Petras, "Experimental evidence of variable-order behavior of ladders and nested ladders," IEEE Transactions on Control Systems Technology, vol. 21, no. 2, pp. 459–466, 2013.
[48] O. Dubois, M. J. Gander, S. Loisel, A. St-Cyr, and D. B. Szyld, "The optimized Schwarz method with a coarse grid correction," SIAM Journal on Scientific Computing, vol. 34, no. 1, pp. A421–A458, 2012.