SIAM J. NUMER. ANAL.
Vol. 62, No. 5, pp. 2196–2221

SOME GRÖNWALL INEQUALITIES FOR A CLASS OF DISCRETIZATIONS OF TIME FRACTIONAL EQUATIONS ON NONUNIFORM MESHES*

YUANYUAN FENG†, LEI LI‡, JIAN-GUO LIU§, AND TAO TANG¶

Abstract. We consider the completely positive discretizations of fractional ordinary differential equations (FODEs) on nonuniform meshes. Making use of the resolvents for nonuniform meshes, we first establish comparison principles for the discretizations. Then we prove some discrete Grönwall inequalities using the comparison principles and careful analysis of the solutions to the time continuous FODEs. Our results do not have restriction on the step size ratio. The Grönwall inequalities for dissipative equations can be used to obtain the uniform-in-time error control and decay estimates of the numerical solutions. The Grönwall inequalities are then applied to subdiffusion problems and the time fractional Allen–Cahn equations for illustration.

Key words. resolvent kernel, complete positivity, comparison principle, nonuniform mesh, dissipative equation

MSC codes. 34A08, 45E10, 65L05, 65R20

DOI. 10.1137/24M1631614

1. Introduction.

The time fractional differential equations with Caputo derivatives [1, 5, 12] have been widely used to model the power law memory effects of energy dissipation for some anelastic materials, and soon became a useful modeling tool in engineering and physical sciences to construct physical models for nonlocal interactions in time (see [35]). The models with Caputo derivatives may also result from the complexity reduction and the generalized Langevin dynamics [13, 19]. The Caputo derivatives are more suitable for studying the initial value problems as it removes the singularity in the Riemann–Liouville derivatives [16]. In analyzing these models, especially the a priori energy estimates for time fractional PDEs, it is crucial

*Received by the editors January 16, 2024; accepted for publication (in revised form) June 21, 2024; published electronically September 18, 2024.
https://doi.org/10.1137/24M1631614

Funding: This work was partially supported by the National Key R&D Program of China, projects 2021YFA1002800 and 2020YFA0712000. The work of the first author was partially sponsored by NSFC 12301283, Shanghai Sailing program 23YF1410300, and Science and Technology Commission of Shanghai Municipality (22DZ2229014). The work of the second author was partially supported by NSFC 12371400 and 12031013, the Strategic Priority Research Program of Chinese Academy of Sciences, grant XDA25010403. The work of the third author was partially supported by NSF DMS-2106988. The work of the fourth author was partially supported by the Guangdong Provincial Key Laboratory of Interdisciplinary Research and Application for Data Science under UIC2022B1212010006, and by Science Challenge Project (grant TZ2018001) and NSFC 11731006 and K20911001.

†School of Mathematical Sciences, Key Laboratory of MEA (Ministry of Education) and Shanghai Key Laboratory of PMMP, East China Normal University, Shanghai, 200241, China (yyfeng@math.ecnu.edu.cn).

‡Corresponding author. School of Mathematical Sciences, Institute of Natural Sciences, MOE-LSC, Shanghai Jiao Tong University, Shanghai, 200240, China (leili2010@sjtu.edu.cn).

§Department of Mathematics, Department of Physics, Duke University, Durham, NC 27708 USA (jliu@math.duke.edu).

¶School of Electrical and Computer Engineering, Guangzhou Nanfang College, Guangzhou 510970, China and Guangdong Provincial Key Laboratory of Interdisciplinary Research and Application for Data Science, BNU-HKBU United International College, Zhuhai, 519087, China (ttang@uic.edu.cn).

to have the comparison principles and Grönwall inequalities for the time fractional ordinary differential equations (FODEs) that take values in $\mathbb{R}$.

The FODE taking values in $\mathbb{R}$ with Caputo derivative of order $\alpha \in (0,1)$ can be written as

$$D_c^\alpha u = f(t,u), \quad u(0) = u_0, \tag{1.1}$$

where $f : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ is assumed to be locally Lipschitz and $u : [0,T) \to \mathbb{R}$ for some $T > 0$ is the unknown function. Here, the Caputo derivative is defined by

$$D_c^\alpha u = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{u'(s)}{(t-s)^\alpha} ds, \tag{1.2}$$

where $\Gamma(\cdot)$ is the gamma function. Often the comparison principles may involve functions u that are continuous but not absolutely continuous so the generalized definitions of Caputo derivatives in [16, 17] might be considered in these cases. One may refer to section 2.1 for more details. The FODE (1.1) is equivalent to the integral equation (see [5] and also [16, 17] for generalized versions)

$$u(t) = u_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s, u(s)) ds. \tag{1.3}$$

In other words, the FODE is equivalent to a Volterra equation with the Abel integral kernel

$$g_\alpha(t) := \frac{1}{\Gamma(\alpha)} t_+^{\alpha-1}, \tag{1.4}$$

where $t_+ = t1_{t\ge 0}$ and $1_{t\ge 0}$ is the standard Heaviside step function. The kernel g_α is known to be completely monotone [34, 31], and thus log-convex and completely positive [3]. A function $a : (0,\infty) \to \mathbb{R}$ is said to be completely monotone if $(-1)^n a^{(n)}(t) \ge 0$ for $n = 0, 1, \ldots$. A completely monotone function that is not identically zero is strictly positive (i.e., bigger than zero everywhere on $(0,\infty)$) by the Bernstein theorem [34, 31]. A locally integrable function $a : [0,\infty) \to \mathbb{R}$ is said to be completely positive if the resolvent kernels given in Definition 2.4 below are nonnegative (see the original paper [3] for more details). It has been pointed out in [3] that completely monotone functions are completely positive (as locally integrable functions, the definition at $t = 0$ does not matter). Using the integral formulation, it is clear that as $\alpha \equiv 1$, (1.1) reduces to the usual ODE.

The first type of comparison principle is for two solution curves of (1.1) or (1.3). If the initial value of one solution is smaller, the solution is always smaller. Such results are well-established and one may refer to [9, 6] for examples. A more useful type of comparison principle is for inequalities, which can give estimates to some energy functionals and norms for the solutions. Such results using the differential inequalities are actually also well-known [30, 9] but the proofs there are not easy to generalize to numerical schemes. For uniform meshes, the comparison principles for the differential form have been established in [18] and [20]. The proof there, however, heavily relies on the discretization and is intrinsically different from the proof for the time continuous version in [30, 9]. It is thus desired that the proof can be motivated from the analysis for the time continuous equations. Besides, due to the weak singularity in the memory kernels, the FODE models often exhibit multiscale behaviors and the solution has singulairty at $t = 0$ [4, 32, 33]. The adaptive time-stepping is often adopted to address this issue [28, 23, 32, 15, 26]. Hence, it is desired to establish the comparison principles for the variable step discretizations.

Establishing explicit bounds for the discrete time fractional inequalities, or the discrete Grönwall inequalities, is of great significance to get a priori bounds and to prove stability of numerical schemes. The Grönwall inequalites with linear function $f(u) = \beta u + c$ may be established using the comparison principles because the solutions to the FODEs with linear $f(\cdot)$ are explicitly known. There are several results about the discrete Grönwall inequalities in literature [8, 21, 23]. The Grönwall inequality in [8] is restricted to uniform meshes. The ones in [21] are based on the special form of L1 scheme so they only apply to L1 discretizations. The ones in [23] can apply to a broader class but there is step ratio restriction and are not directly applicable to estimate the decaying rate of solutions. Often the error control for dissipative systems like the subdiffusion problems relies on the maximum principles [32] so it cannot get the decay bounds for the numerical solutions.

In this work, we aim to establish the comparison principles and Grönwall inequalities for completely positive discretizations on nonuniform meshes (detailed in section 2.3). Our new approach is to discover a different proof for the comparison principles for time continuous FODEs based on the resolvents [3, 29], using similar techniques as in [9, 7]. With the so-called pseudo-convolution (see the details in section 2.2), one can define the resolvent kernels for nonuniform meshes as well. Then, we can establish the comparison principles for the variable-step discretizations using the resolvent kernels by generalizing the argument for the continuous case. Based on the comparison principles, we establish Grönwall inequalities for the discretizations making use of some key properties of the solutions to continuous equations. There is no monotonicity assumption on the function f. The main results can be summarized as below.

THEOREM 1.1 (informal version of Theorems 4.3 and 4.8). *Consider the discretization of the differential form* (2.17) *that is completely positive and assume that* c_{n-j}^n *is comparable to the average of* $g_{1-\alpha}(t_n - \cdot)$ *on the* j*th interval in some sense (see the corresponding sections for the details). If a nonnegative sequence* v_n *satisfies that* $\mathcal{D}_\tau^\alpha v_n \le -\lambda v_n + c$ *for* $\lambda > 0$, *then for some constants* $\nu > 0$ *and* $\tilde{\sigma} > 0$, *it holds that*
(a) $v_n \le v_0 E_\alpha(-\nu^{-1}\lambda t_n^\alpha) + (c/\lambda)(1 - E_\alpha(-\nu^{-1}\lambda t_n^\alpha))$ *if* $v_0 \le c/\lambda$;
(b) $v_n \le (v_0 - c/\lambda) E_\alpha(-\tilde{\sigma}\lambda t_n^\alpha) + c/\lambda$ *if* $v_0 > c/\lambda$.

For the two cases, the final statements are similar. We divide them into $v_0 \le c/\lambda$ and $v_0 > c/\lambda$ because the proofs rely on different properties of the solutions to the continuous equations. The result in (a) is useful for the uniform bound estimate of the numerical solutions while the result in (b) is useful to get the decay rate of the solutions. One may obtain the decay estimates of norms of the numerical solutions and the uniform-in-time error estimates for dissipative systems based on these results. Another result is the following.

THEOREM 1.2 (informal version of Theorem 4.13). *Consider the discretization of the differential form* (2.17) *that is completely positive. Assume that* c_{n-j}^n *is comparable to the average of* $g_{1-\alpha}(t_n - \cdot)$ *on the* j*th interval and the stepsize* τ_n *satisfies that* $\lambda \tau_n^\alpha \le \delta$ *for some* $\delta > 0$ *(see the corresponding section for the details). If* $\mathcal{D}_\tau^\alpha v_n \le \lambda v_n + c$ *for* $\lambda > 0$, *then it holds for some* $\mu > 0$ *that* $v_n \le (v_0 + c/\lambda) E_\alpha(\mu^{-1}\lambda t_n^\alpha) - c/\lambda$.

In this result, we have removed the usual constraint on the stepsize ratio in literature. The result is based on the comparison principles and careful analysis of the asymptotic behaviors of the solutions to the continuous equation.

The rest of this paper is organized as follows. In section 2, some preliminary concepts and results are reviewed, including the definition of generalized Caputo derivatives, the behaviors of the Mittag–Leffler functions, and the discretization we conisder in this work on nonuniform meshes. In section 3, we present a new proof of comparision principles using the resolvent kernels for the time continuous equations and then generalize it to completely positive discretizations on nonuniform meshes. Our main results for Grönwall inequalities are then established in section 4 and some applications to dissipative systems are presented in section 5.

2. Preliminaries and setup.

In this section, we review some basic concepts and results for later sections.

2.1. The generalized Caputo derivatives and the resolvent kernels.

In this subsection, we summarize the generalized Caputo derivative introduced in [16, 17]. This generalized definition is theoretically convenient, since it allows us to use the underlying group structure. Moreover, the generalized definition allows us to consider the fractional *inequalities* for functions that are merely continuous.

The standard one-sided convolution for two functions u and v defined on $[0,\infty)$ is given by

$$(u * v)(t) = \int_{[0,t]} u(s)v(t-s)\,ds, \tag{2.1}$$

which can be generalized to distributions whose supports are on $[0,\infty)$ (see [16, sections 2.1 and 2.2]). Recall the Abel integral kernels for $\alpha > 0$ in (1.4). Let $g_0 = \delta$ and

$$g_\beta(t) = \frac{1}{\Gamma(1+\beta)} D\left(t_+^\beta\right), \beta \in (-1, 0). \tag{2.2}$$

Here D means the distributional derivative on $\mathbb{R}$. Then, for any $\beta_1 > -1$ and $\beta_2 > -1$,

$$g_{\beta_1} * g_{\beta_2} = g_{\beta_1+\beta_2}. \tag{2.3}$$

We remark that g_β can indeed be defined for $\beta \in \mathbb{R}$ (see [16]) so that $\{g_\beta : \beta \in \mathbb{R}\}$ forms a convolutional group. We introduce the generalized definition.

DEFINITION 2.1 (see [16, 17]). *Let $0 < \alpha < 1$ and $T > 0$. For $u \in L^1_{\mathrm{loc}}[0,T)$ and a given $u_0 \in \mathbb{R}$, the αth order generalized Caputo derivative of u associated with initial value u_0 is a distribution with support in $[0,T)$ given by*

$$D_c^\alpha u = g_{-\alpha} * \Big((u-u_0)1_{t\geq 0}\Big). \tag{2.4}$$

It has been verified in [16] that if the function u is absolutely continuous, the generalized definition reduces to the classical definition (1.2). A function $u \in L^1_{\mathrm{loc}}[0,T)$ is a weak solution to (1.1) on $[0,T)$ with initial value u_0 if the equality holds in distribution. A weak solution u is a strong solution if $\lim_{t\to 0+} \frac{1}{t}\int_0^t |u(s)-u_0|ds = 0$ and both sides of (1.1) are locally integrable on $[0,T)$. Using the group property (2.3), one may obtain directly the following.

PROPOSITION 2.2 (see [16, Proposition 4.2]). *Suppose $f \in L^\infty_{\mathrm{loc}}([0,\infty)\times\mathbb{R};\mathbb{R})$. Fix $T > 0$. Then, $u(t) \in L^1_{\mathrm{loc}}[0,T)$ with initial value u_0 is a strong solution of* (1.1) *on $(0,T)$ if and only if it solves the following integral equation:*

$$u(t) = u_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1} f(s,u(s))ds \quad \forall t \in (0,T). \tag{2.5}$$

Using this integral formulation, the following has been shown in [16].

LEMMA 2.3. *Suppose $f:[0,\infty)\times(u_*,u^*)\to\mathbb{R}$ is continuous and locally Lipschitz in u where $u_*\in[-\infty,\infty)$ and $u^*\in(-\infty,\infty]$. For any $u_0\in(u_*,u^*)$, there is a maximum time $T_b>0$ and a unique weak solution on $[0,T_b)$ satisfying $u(0+)=u_0$. This weak solution is a strong solution, and if $T_b<\infty$, then either $\limsup_{t\to T_b^-}u(t)=u^*$ or $\liminf_{t\to T_b^-}u(t)=u_*$.*

Since the weak solution becomes the strong solution and, in fact, belongs to $C^1(0,T_b)\cap C[0,T_b)$, the Caputo derivative then reduces to the classical one. Consider the following linear FODE:

$$D_c^\alpha v=\beta v+c. \tag{2.6}$$

The solution exists globally (i.e., $T_b=\infty$) and is given by

$$v=\left(v_0+\frac{c}{\beta}\right)E_\alpha(\beta t^\alpha)-\frac{c}{\beta}. \tag{2.7}$$

Here, $E_\alpha(z):=E_{\alpha,1}(z)$ and the Mittag–Leffler function $E_{\alpha,\beta}$ is an entire function given by (see, for example, [27, 10])

$$E_{\alpha,\beta}(z)=\sum_{k=0}^{\infty}\frac{z^k}{\Gamma(\alpha k+\beta)},\alpha>0,z\in\mathbb{C}. \tag{2.8}$$

It is then clear that $E'_\alpha(z)=\frac{E_{\alpha,0}(z)}{\alpha z}=\alpha^{-1}E_{\alpha,\alpha}(z)$. The function $E_{\alpha,\beta}(z)$ has the following integral representation (see [10]) for $\alpha\in(0,2)$:

$$E_{\alpha,\beta}(z)=\begin{cases}\dfrac{1}{\alpha}z^{(1-\beta)/\alpha}e^{z^{1/\alpha}}+\dfrac{1}{2\pi i\alpha}\displaystyle\int_{\gamma(\epsilon;\delta)}\frac{e^{\zeta^{1/\alpha}}\zeta^{(1-\beta)/\alpha}}{\zeta-z}d\zeta, & z\in R_+,\\ \dfrac{1}{2\pi i\alpha}\displaystyle\int_{\gamma(\epsilon;\delta)}\frac{e^{\zeta^{1/\alpha}}\zeta^{(1-\beta)/\alpha}}{\zeta-z}d\zeta & z\in R_-,\end{cases} \tag{2.9}$$

where $\gamma(\epsilon;\delta)$ is the curve consisting of $\{re^{-i\delta}:r\geq\epsilon\}$, $\{\epsilon e^{i\theta}:-\delta\leq\theta\leq\delta\}$, and $\{re^{i\delta}:r\geq\epsilon\}$, going from $\infty e^{-i\delta}$ to $\infty e^{i\delta}$. The parameter δ satisfies $\frac{\pi\alpha}{2}<\delta\leq\min(\alpha\pi,\pi)$. The region R_+ is the one on the right of $\gamma(\epsilon;\delta)$ while R_- is on the left. Note that the two expressions appear different, but they are actually connected continuously across the curve $\gamma(\epsilon;\delta)$.

Next, we recall the resolvent kernels (see [2, 3, 29]). In [8], the resolvent kernels have been used to establish the monotonicity of the solutions to autonomous fractional ODEs, and then generalized to discrete schemes on nonuniform meshes in [7].

DEFINITION 2.4 (see [3, 29]). *Let $\lambda>0$. The resolvent kernels r_λ and s_λ for a are defined, respectively, by*

$$r_\lambda+\lambda r_\lambda*a=\lambda a,\quad s_\lambda+\lambda s_\lambda*a=1_{t\geq0}. \tag{2.10}$$

It has been shown in [7, Proposition 2.1] that both r_λ and s_λ are completely monotone for all $\lambda>0$ if a is completely monotone. For the kernel $a=g_\alpha$, r_λ and s_λ are thus completely monotone and strictly positive. Hence, the Abel kernel g_α is completely positive for $\alpha\in(0,1)$. In fact, it has been mentioned in the proof of [8, Lemma 3.4] that the resolvent r_λ for $g_\alpha=\frac{1}{\Gamma(\alpha)}t_+^{\alpha-1}$ is

$$r_\lambda(t)=-\frac{d}{dt}E_\alpha(-\lambda t^\alpha)>0, \tag{2.11}$$

and thus $s_\lambda(t)=1-E_\alpha(-\lambda t^\alpha)>0$.

2.2. Discretization on nonuniform meshes. There are two ways to discretize the fractional ODEs. One is to discretize the differential form (1.1) even though the derivative may be understood in the generalized one (2.4), and the other way is to discretize the integral form (1.3).

Let the computational time interval be $[0,T]$, and let $0=t_0<t_1<t_2<\cdots<t_N=T$ be the grid points. We allow $T=\infty$ in some applications and in this case, there are infinitely many grid points t_n. Define $\mathcal{N}:=\{1,\ldots,N\}$ or $\mathcal{N}=\mathbb{N}_+$ when $T=\infty$. For $n\in\mathcal{N}$, define

$$\tau_n := t_n - t_{n-1}. \tag{2.12}$$

Let u_n be the numerical solution at t_n.

The integral form (1.3) may be discretized for $\theta\in[0,1]$ by

$$u_n = u_0 + \sum_{j=1}^{n} \bar{a}_{n-j}^n f_j^\theta \tau_j = u_0 + \sum_{j=1}^{n} a_{n-j}^n f_j^\theta =: u_0 + \mathcal{I}_\tau^\alpha f_n^\theta. \tag{2.13}$$

Here, $\{\bar{a}_{n-j}^n\}$ is an approximation of the average of $g_\alpha(t_n-s)=\frac{1}{\Gamma(\alpha)}(t_n-s)^{\alpha-1}$ on $[t_{j-1},t_j]$ while a_{n-j}^n is like the integral of $g_\alpha(t_n-s)$ on this interval. The notation f_j^θ means an approximation of f at $(1-\theta)t_j+\theta t_{j-1}$. We consider two examples here. The first is

$$f_j^\theta = (1-\theta) f(t_j,u_j) + \theta f(t_{j+1},u_{j+1}) \tag{2.14}$$

while the second is

$$f_j^\theta = f(t_j^\theta, u_j^\theta), \quad t_j^\theta := (1-\theta)t_j + \theta t_{j-1}, \quad u_j^\theta := (1-\theta)u_j + \theta u_{j-1}. \tag{2.15}$$

These two approximations will have no big difference in our analysis later.

Next, we consider discretization of the differential form (1.1). One can first approximate the derivative and define

$$\nabla_\tau u_j := u_j - u_{j-1}. \tag{2.16}$$

One may then introduce the following approximation inspired by (1.2):

$$\mathcal{D}_\tau^\alpha u_n := \sum_{j=1}^{n} c_{n-j}^n \nabla_\tau u_j = f_n^\theta, \quad n\in\mathcal{N}. \tag{2.17}$$

Here, $\nabla_\tau u_j \approx u'(t_j)\tau_j$ so c_{n-j}^n is an approximation of the average of $g_{1-\alpha}(t_n-s)$ on $[t_{j-1},t_j)$.

We remark that (2.13) and (2.17) are related by the so-called pseudo-convolution in [7] (also used in [24] for defining the discrete orthogonal convolution (DOC) kernel). We arrange the kernel (a_{n-j}^n) into a lower triangular array A of the following form with size $|\mathcal{N}|\times|\mathcal{N}|$

$$A = \begin{bmatrix} a_0^1 & & & & \\ a_1^2 & a_0^2 & & & \\ \cdots & \vdots & \vdots & & \\ a_{n-1}^n & \cdots & a_1^n & a_0^n & \\ \cdots & \vdots & \vdots & & \vdots \end{bmatrix}. \tag{2.18}$$

The pseudo-convolution $\bar{*}$ between two array kernels is given by the usual matrix product between two arrays of the form (2.18) [7]. In particular, $C = A\,\bar{*}\,B$ is given by

$$c_k^n = \sum_{j=0}^{k} a_{k-j}^n b_j^{n+j-k} \quad \text{or} \quad c_{n-k}^n = \sum_{j=k}^{n} a_{n-j}^n b_{j-k}^j. \tag{2.19}$$

There are some special kernels that play important roles. The first is I with $I_{n-j}^n = \delta_{nj}$. This is the kernel that is 1 on the diagonal. The other one is L with $L_{n-j}^n = 1$ for all $j \le n$. This kernel corresponds to the Heaviside function $1_{t\ge 0}$ for the continuous case. The inverse of L, $L^{(-1)}$, satifies that $(L^{(-1)})_{n-j}^n = 1$ for $j = n$, $(L^{(-1)})_{n-j}^n = -1$ for $j = n-1$ and 0 otherwise. This corresponds to the finite difference operator in (2.16). It can be verified that I is the identify for the pseudo-convolution, and the following properties hold.

(a) The pseudo-convolution is associative.

(b) If B is an inverse of A, namely $A\,\bar{*}\,B = I$, then $B\,\bar{*}\,A = I$.

The pseudo-convolution provides us a convenient way to investigate the properties of the kernels. For example, the monotonicity preserving properties of some discretizations have been established using such tools in [6].

For a given A, the kernel C_R with $A\,\bar{*}\,C_R = L$ is called the right complementary kernel. The kernel C_L with $C_L\,\bar{*}\,A = L$ is called the left complementary kernel. The kernel C_R is, in fact, the so-called right convolutional complementary (RCC) kernel (see [22]) and C_L is the discrete complementary convolution (DCC) kernel (see [23]). If A is a kernel that is invertible, then direct verification tells us that $C_R = A^{(-1)}\bar{*}L$ and $C_L = L\bar{*}A^{(-1)}$. Moreover, a_j^n is nonincreasing in n if and only if the inverse of C_R has nonpositive off-diagonals; a_j^n is nonincreasing in j if and only if the inverse of C_L has nonpositive off-diagonals.

The pseudo-convolution is also defined between a kernel and a vector $x = (x_j)_{j\in\mathcal{N}}$:

$$y = A\bar{*}x \quad \Longleftrightarrow \quad y_n = \sum_{j=1}^{n} a_{n-j}^n x_j \quad \forall n \in \mathcal{N}, \tag{2.20}$$

then it holds that $A\bar{*}(B\bar{*}x) = (A\bar{*}B)\bar{*}x$.

One has the following simple conclusion.

Lemma 2.5. *Consider the kernel $C := (c_{n-j}^n)$ for $\mathcal{D}_\tau^\alpha$ in* (2.17) *and the kernel $A := (a_{n-j}^n)$ for $\mathcal{I}_\tau^\alpha$ in the integral scheme* (2.13)*. If C is the right complementary kernel of A, then the two schemes are equivalent.*

Proof. It is clear that $\mathcal{D}_\tau^\alpha u_n = C\bar{*}\nabla_\tau u_n = C\bar{*}L^{(-1)}\bar{*}(u-u_0)_n$, while $\mathcal{I}_\tau^\alpha f(t_n^\theta, u_n^\theta) = A\bar{*}f_n^\theta$. Hence, if C is the right complementary kernel of A, $C\bar{*}L^{(-1)}$ is the inverse of A and the claim follows. ∎

Below, we will mainly focus on schemes for the differential form. We then have

$$B := A^{-1} = C\bar{*}L^{-1} \Leftrightarrow b_0^n = c_0^n, \quad b_{n-j}^n = c_{n-j}^n - c_{n-j-1}^n, \quad j \le n-1. \tag{2.21}$$

Regarding the solvability, the following is straightforward and we omit the proof.

Lemma 2.6. *Suppose $f(t,\cdot)$ is uniformly Lipschitz with constant M. If $\theta M a_0^n < 1$ or, equivalently, $\theta M < c_0^n$ for all $n \in \mathcal{N}$, the numerical solution to* (2.13) *or, equivalently, to* (2.17) *is uniquely solvable.*

Remark 2.7. As commented in [6], if c_0^n is like the average of $g_{1-\alpha}(t_n - s)$, then $a_0^n = 1/c_0^n$ is like the integral of g_α on $[t_{n-1}, t_n)$ so the discussion above is quite natural.

2.3. Completely positive discretizations. In this subsection, we introduce the class of variable-step discretizations we consider in this work.

One may define the resolvent R_λ using the pseudo-convolution similar to the continuous case as in Definition 2.4,

$$R_\lambda + \lambda R_\lambda \bar{*} A = \lambda A \quad \Longleftrightarrow \quad A - R_\lambda \bar{*} A = \frac{1}{\lambda} R_\lambda. \tag{2.22}$$

The resolvent kernels have the following simple facts as proved in [7].

Lemma 2.8. *Suppose the diagonal elements of A are positive. The resolvent kernel R_λ always exists for $\lambda > 0$ and $R_\lambda \bar{*} A = A \bar{*} R_\lambda$. Moreover, as $\lambda \to \infty$,*

$$R_\lambda = I - \lambda^{-1} A^{(-1)} + O(\lambda^{-2}), \tag{2.23}$$

where the $O(\lambda^{-2})$ is elementwise.

As a discrete analogue of the complete positivity for the continuous kernels studied in [3], namely those with nonnegative resolvents, we say the kernel A (or the discretization) is *completely positive* if for all $\lambda > 0$, R_λ has nonnegative entries, $0 < (R_\lambda)_0^n < 1$ and $\sum_{j=1}^n (R_\lambda)_{n-j}^n \le 1$.

The following basic condition is present for many usual discretizations [23].

Condition 2.9 (equivalent condition for complete positivity). *The array kernel A is invertible and the inverse $B = A^{(-1)} = (b_{n-j}^n)$ satisfies for all $n \in \mathcal{N}$ that*

$$\begin{aligned} &b_0^n > 0, \quad b_{n-j}^n \le 0 \quad \forall j < n, \\ &\sum_{j=1}^n b_{n-j}^n \ge 0. \end{aligned} \tag{2.24}$$

The following has been proved in [7, Theorem 5.1].

Proposition 2.10. *The array kernel A is completely positive if and only if it satisfies Condition* 2.9*.*

The class of discretizations we study, therefore, would be those satisfying Condition 2.9, or equivalently, those are completely positive. For such discretizations, one has the following observation which might be used in applications. The proof uses the signs in $B = A^{(-1)}$ and is similar to those in [18, Theorem 3] and [20, Proposition 2.2]. We thus omit the proof.

Lemma 2.11. *Suppose H is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and $\varphi : H \to \mathbb{R}$ is a convex function. If the kernel $B = C \bar{*} L^{(-1)}$ for $\mathcal{D}_\tau^\alpha$ satisfies* (2.24)*, then*

$$\mathcal{D}_\tau^\alpha \varphi(u_n) \le \langle \varphi'(u_n), \mathcal{D}_\tau^\alpha u_n \rangle. \tag{2.25}$$

In particular, if $u_n \ne 0$, then one has

$$\mathcal{D}_\tau^\alpha \|u_n\| \le \left\langle \frac{u_n}{\|u_n\|}, \mathcal{D}_\tau^\alpha u_n \right\rangle. \tag{2.26}$$

The following gives a simple condition for complete positivity.

Lemma 2.12. *If $a_{j-1}^{n-1} \ge a_j^n$ for all $n \in \mathcal{N}$, $n \ge 2$, and $j \le n-1$, and the inverse B satisfies $b_0^n > 0$, $b_{n-j}^n \le 0$ $\forall j < n$, then A is completely positive.*

In fact, if $b_0^n > 0$ and $b_{n-j}^n \le 0$ for all $j < n$, A must have nonnegative entries by [7, Lemma 4.3]. Then, together with the fact that $a_{j-1}^{n-1} \ge a_j^n$, one can easily verify that $C = A^{(-1)} \bar{*} L$ is nonnegative.

Below we look at examples of discretizations on nonuniform meshes. We say a discretization is completely positive if the corresponding kernel A is completely positive (note that the schemes for differential form and the integral form are equivalent by Lemma 2.5). The L1 scheme [25, 32] is the most popular and simplest discretization, obtained by piecewise linear interpolation of the derivative u in the differential form, which can be written as

$$D^\alpha u(t_n) \approx \mathcal{D}_\tau^\alpha u_n := C \bar{*} \nabla_\tau u_n = C \bar{*} L^{(-1)} \bar{*} (u - u_0)_n, \tag{2.27}$$

where

$$c_{n-j}^n = \frac{1}{\tau_j \Gamma(1-\alpha)} \int_{t_{j-1}}^{t_j} (t_n - s)^{-\alpha}\, ds. \tag{2.28}$$

It can be verified directly that $B^{(-1)} := C \bar{*} L^{(-1)}$ satisfies (2.24) so that A is completely positive. Hence, the $L1$ scheme is completely positive (see also [6, section 6]).

One can also approximate the integral formulation. In [6, section 6], the integral scheme

$$u_n = u_0 + \sum_{j=1}^n a_{n-j}^n f(t_j, u_j), \quad a_{n-j}^n = \frac{1}{\Gamma(\alpha)} \int_{t_{j-1}}^{t_j} (t_n - s)^{-\alpha}\, ds \tag{2.29}$$

is proposed as approximation and has been proved to be completely positive.

For uniform schemes, all the CM-preserving schemes discussed in [20] are completely positive. This includes the standard Grünwald–Letnikov (GL) scheme, the convolutional quadrature (CQ) with θ-method etc. The GL and CQ methods are not easy to generalize to nonuniform meshes. The requirement on completely positive schemes is quite reasonable. Some related discussions can be found in [23, 22]. The schemes considered in [23] are a subclass of completely positive schemes. Moreover, the second order Crank–Nicolson scheme with the $L1^+$ discretization can be decomposed into a completely positive scheme plus a local difference operator and the comparison principle can be established similarly, as we shall see in section 3.2.

Remark 2.13. The completely positive schemes may have some order barrier, which means that the high order schemes cannot be completely positive. Often, in these high order schemes, only the first few b_j^n terms (like $j = 1, 2$) cause issues. It is possible to decompose the kernel for a high order scheme into a completely positive kernel and the kernel for some local difference operator. We leave this for future study.

3. The comparison principles.

In this section, we establish some comparison principles. First, we give an alternative proof for continuous problems based on the resolvents. Then, we generalize the proof to discrete schemes.

3.1. Comparison principles for time continuous case.

THEOREM 3.1. *Suppose f is continuous and locally Lipschitz in u. Let y and z be two continuous functions satisfying $D_c^\alpha y \ge f(t, y)$, $D_c^\alpha z \le f(t, z)$ in the distributional sense, with $y(0) \ge z(0)$, then $y(t) \ge z(t)$ on the common interval of existence. If $y(0) > z(0)$, then $y(t) > z(t)$ on the interval considered.*

This result is, in fact, known. One can refer to [30, Theorem 2.3] or [9, Theorem 2.2] for the generalized version with less regularity. However, the proof there relies on the continuity of the continuous functions, which is not applicable for discrete sequences. We provide another proof using the resolvents, which can then be generalized the discrete schemes.

By the definition (2.4), $D_c^\alpha u = g_{-\alpha} * ((u-u_0)1_{t\ge 0})$. Using the definition of the resolvent, we find that $r_\lambda * g_{-\alpha} = \lambda(\delta - r_\lambda)$. Hence, one has

$$\lambda^{-1} r_\lambda * D_c^\alpha u = (u-u_0)1_{t\ge 0} - r_\lambda * (u-u_0). \tag{3.1}$$

If u is absolutely continuous, one can see this more clearly. In fact, $g_{-\alpha} * (1_{t\ge 0}(u - u_0)) = g_{1-\alpha} * (u')$. Here, $g_{1-\alpha}$ is the complementary kernel of g_α in the sense that $g_{1-\alpha} * g_\alpha = 1_{t\ge 0}$. It can then be derived from the definition of the resolvent that

$$r_\lambda * g_{1-\alpha} = \lambda s_\lambda = \lambda(1_{t\ge 0} - r_\lambda * 1_{t\ge 0}). \tag{3.2}$$

The main observation is that r_λ is nonnegative so that the inequality can be preserved when convolving both sides with $\lambda^{-1} r_\lambda$.

Proof of Theorem 3.1. Define $u := y - z$. Taking the difference between the two relations, one has (in the distributional sense) that

$$D_c^\alpha u \ge h(t)u(t), \quad h(t) = \int_0^1 \partial_u f(\eta y + (1-\eta)z)\,d\eta.$$

Convolving both sides with $\lambda^{-1} r_\lambda$, which is nonnegative, one then has

$$(u-u_0)1_{t\ge 0} - r_\lambda * (u-u_0) \ge \lambda^{-1} r_\lambda * (hu).$$

This implies for $t \ge 0$ that

$$\begin{aligned} u(t) &\ge (u_0 - r_\lambda * u_0) + r_\lambda * (u + \lambda^{-1} h u) \\ &= u_0 s_\lambda + \int_0^t r_\lambda(t-s)(1 + h(s)/\lambda)u(s)\,ds. \end{aligned} \tag{3.3}$$

If $u_0 > 0$, one can see that $u(t) > 0$ for t small enough. Then, for all $t > 0$ considered, we can take λ large enough such that $1 + h(s)/\lambda > 0$ on $[0,t]$, and (3.3) implies that $u(t) > 0$.

Consider $u_0 = 0$. If $u(t) < 0$ somewhere, we set $t_1 \ge 0$ to be the first time when $u(t) < 0$ on $(t_1, t_1 + \epsilon)$ for some $\epsilon > 0$. Then, $u(t_1) = 0$. It is clear that $u(t) = 0$ on $[0, t_1]$, otherwise (3.3) would yield a contradiction by setting $t = t_1$ and λ large enough. Take λ large enough such that $1 + h(s)/\lambda > 0$ on $(t_1, t_1+\epsilon)$ and set

$$A := \sup_{s\in(t_1, t_1+\epsilon)} 1 + h(s)/\lambda.$$

Take $\delta \le \epsilon$ such that $\int_0^\delta r_\lambda(s)\,ds \le 1/(2A)$. Moreover, let

$$t_2 = \mathrm{argmin}_{s\in[t_1, t_1+\delta]} u(s) \in (t_1, t_1+\delta].$$

Then, $u(t_2) < 0$ and

$$u(t_2) \ge \int_{t_1}^{t_2} r_\lambda(t_2 - s)(1 + h(s)/\lambda)u(s)\,ds \ge Au(t_2)\int_{t_1}^{t_2} r_\lambda(t_2 - s)\,ds \ge u(t_2)/2.$$

This is then a contradiction. Hence, the claim is proved. ∎

If the inequality is given in integral form, the technique here does not apply, since $\delta - r_\lambda$ is not nonnegative. Currently, the comparison principle is known to hold for integral forms only when the function $f(t,\cdot)$ is nondecreasing as in [16]. Whether it can be generalized to general f is an interesting question.

3.2. Comparison principles for numerical schemes. The argument in [30, Theorem 2.3] or [9, Theorem 2.2] cannot be generalized to discrete schemes easily as it relies on the continuity. Motivated by the proof above based on resolvents and the pseudo-convolution, we are then able to establish a series of comparison principles for the discrete schemes.

The following is for the implicit scheme ($\theta = 1$).

THEOREM 3.2. *Suppose that the discretization is completely positive and C is the corresponding right complementary kernel. If $y_0 \ge z_0$ and for all $n \in \mathcal{N}$,*

$$\mathcal{D}_\tau^\alpha y_n \ge f(t, y_n), \quad \mathcal{D}_\tau^\alpha z_n \le f(t, z_n), \tag{3.4}$$

then the following comparison principles hold:

(1) *If $f(t, \cdot)$ is nonincreasing for all t, then $y_n \ge z_n$. If C has positive entries, then $y_0 > z_0$ implies that $y_n > z_n$.*

(2) *Suppose that $f(t, \cdot)$ is uniformly Lipschitz such that $|\partial_u f(t,u)| \le M$ for all t, u, and the discretization satisfies the stability condition $M/c_0^n < 1$, then one has $y_n \ge z_n$. If C has positive entries, then $y_0 > z_0$ implies that $y_n > z_n$.*

Proof. Define $u_n = y_n - z_n$. Then, it holds that

$$\mathcal{D}_\tau^\alpha u_n \ge h_n u_n, \tag{3.5}$$

where $h_n = \int_0^1 \partial_u f(\eta y_n + (1-\eta) z_n)\, d\eta$. Recall

$$\mathcal{D}_\tau^\alpha u_n = C \bar{*} (\nabla_\tau u_n) = A^{(-1)} \bar{*} u_n.$$

Taking pseudo-convolution with $\lambda^{-1} R_\lambda$ on both sides of (3.5) and noting that R_λ has nonnegative entries, one has

$$u_n \ge u_0 \left(1 - \sum_{j=1}^{n} (R_\lambda)_{n-j}^n\right) + R_\lambda \bar{*} [(1 + h_n/\lambda) u_n].$$

Hence,

$$(1 - (R_\lambda)_0^n (1 + h_n/\lambda)) u_n \ge u_0 \left(1 - \sum_{j=1}^{n} (R_\lambda)_{n-j}^n\right) + \sum_{j=1}^{n-1} (R_\lambda)_{n-j}^n (1 + h_j/\lambda) u_j.$$

Assume that $u_j \ge 0$ for $j \le n-1$ has been established. For fixed n, one can choose λ-large enough such that the coefficients on the righthand side are all nonnegative, and also such that $1 - \sum_{j=1}^n (R_\lambda)_{n-j}^n > 0$ if C has positive entries by Lemma 2.8.

If f is nonincreasing, $1 - (R_\lambda)_0^n - h_n/\lambda > 0$ always holds, then simple induction yields that $u_n \ge 0$ if $u_0 \ge 0$.

If f is assumed to be Lipschitz, then by Lemma 2.8, $(R_\lambda)_0^n = 1 - \lambda^{-1} b_0^n + O(\lambda^{-2})$ and $b_0^n = c_0^n$. Consequently,

$$1 - (R_\lambda)_0^n (1 + h_n/\lambda) = \lambda^{-1}(c_0^n - h_n) + O(\lambda^{-2}).$$

Hence, if $c_0^n > M$, one can choose λ large enough such that the coefficient is positive. Then, $u_n > 0$. ∎

Taking $f(t,u) \equiv 0$, then $\mathcal{D}_\tau^\alpha y_n \ge 0$ implies that $y_n \ge 0$, which gives the following.

COROLLARY 3.3. *For any completely positive discretization, if* $\mathcal{D}_\tau^\alpha x_n \ge \mathcal{D}_\tau^\alpha y_n$ *for all* $n \in \mathcal{N}$ *and* $x_0 \ge y_0$*, then* $x_n \ge y_n$.

Below, we consider weighted implicit schemes for $\theta \in [0,1)$.

THEOREM 3.4. *Assume the discretization is completely positive and* C *is the corresponding right complementary kernel. Suppose* y *and* z *are two sequences with* $y_0 \ge z_0$ *and satisfy the following for* $\theta \in [0,1)$ *and all* $n \in \mathcal{N}$:

$$\mathcal{D}_\tau^\alpha y_n \ge f_n^\theta[y], \quad \mathcal{D}_\tau^\alpha z_n \le f_n^\theta[z], \tag{3.6}$$

where $f_n^\theta[y]$ *and* $f_n^\theta[z]$ *are either given by (2.14) or (2.15).*

(1) *Suppose* f *is nondecreasing in* u*. If* $c_0^n > \theta M$*, then* $y_n \ge z_n$.
(2) *If* $f(t,\cdot)$ *is uniformly Lipschitz with constant* M*,* $c_0^n > \theta M$*, and* $c_0^n - c_1^n \ge (1-\theta)M$*, then* $y_n \ge z_n$*. Moreover, if* C *has positive entries, then* $y_0 > z_0$ *implies that* $y_n > z_n$.

Proof. Define $u_n = y_n - z_n$. It holds that

$$\mathcal{D}_\tau^\alpha u_n = C_R \bar{*} L^{(-1)} \bar{*} (u-u_0)_n \ge \theta h_{n,1} u_n + (1-\theta) h_{n,2} u_{n-1}.$$

In the case of (2.14), $h_{n+1,2} = h_{n,1} = \int_0^1 \partial_u f(t_n, \eta y_n + (1-\eta) z_n)\, d\eta$. In the case of (2.15), $h_{n,1} = h_{n,2} = \int_0^1 \partial_u f(t_j^\theta, \eta y_n^\theta + (1-\eta) z_n^\theta)\, d\eta$.

Taking the pseudo-convolution with R_λ on the left for both sides, one has by the nonnegativity of the elements of $\lambda^{-1} R_\lambda$ that

$$u_n - u_0 - R_\lambda \bar{*} (u-u_0) \ge \lambda^{-1} R_\lambda \bar{*} (\theta h_{n,1} u_n + (1-\theta) h_{n,2} u_{n-1}),$$

which implies that

$$\begin{aligned}(1-(R_\lambda)_0^n(1+\theta h_{n,1}/\lambda))u_n \ge& \left(1 - \sum_{j=1}^n (R_\lambda)_{n-j}^n + \lambda^{-1}(1-\theta)(R_\lambda)_{n-1}^n h_{1,2}\right) u_0 \\ &+ \sum_{j=1}^{n-1} \big((R_\lambda)_{n-j}^n (1+\theta h_{j,1}/\lambda) \\ &+ \lambda^{-1}(1-\theta)(R_\lambda)_{n-j-1}^n h_{j+1,2}\big)\, u_j.\end{aligned} \tag{3.7}$$

The observation is that u_n does not depend on λ so for each n one can inductively take λ large enough to show $u_n \ge 0$.

If $f(t,\cdot)$ is nondecreasing, then $h_{n,\ell} \ge 0$, $\ell = 1,2$ and the coefficients on the right-hand side of (3.7) are all nonnegative. By Lemma 2.8 and (2.21),

$$(R_\lambda)_0^n = 1 - \lambda^{-1} c_0^n + O(\lambda^{-2}).$$

Since $|h_{n,\ell}| \le M$, if $c_0^n > \theta M$, then $1 - (R_\lambda)_0^n(1+\theta h_{n,1}/\lambda) > 0$ for λ large enough. This then implies that $u_n \ge 0$ for all $n \in \mathcal{N}$ if $u_0 \ge 0$.

Now, we focus on the second case. Due to the same reason as above, if $c_0^n > \theta M$, then the coefficient of u_n is positive when λ is large enough. Again, by Lemma 2.8 and (2.21), one has

$$\sum_{j=1}^n (R_\lambda)_{n-j}^n = 1 - \lambda^{-1} \sum_{j=1}^n b_{n-j}^n + O(\lambda^{-1}) = 1 - \lambda^{-1} c_{n-1}^n + O(\lambda^{-2}).$$

Moreover, $(R_\lambda)_{n-1}^n = -\lambda^{-1} b_{n-1}^n + O(\lambda^{-2})$. Hence, the leading term in the coefficient of u_0 is $\lambda^{-1} c_{n-1}^n$ with $c_{n-1}^n \ge 0$. For the last term in (3.7), as $\lambda \to \infty$, the coefficient for $j = n-1$ satisfies

$$(R_\lambda)_1^n (1 + \theta h_{n-1,1}/\lambda) + \lambda^{-1}(1-\theta)(R_\lambda)_0^n h_{n,2} = \lambda^{-1}(-b_1^n + (1-\theta) h_{n,2}) + O(\lambda^{-2}).$$

If $c_0^n - c_1^n \ge (1-\theta)M$, then noting $-b_1^n = c_0^n - c_1^n$ and $|h_{n-1}| \le M$, one finds that the leading term in the coefficient for $j = n-1$ is $\lambda^{-1}(-b_1^n + (1-\theta)h_{n,2})$ with $-b_1^n + (1-\theta)h_{n,2} \ge 0$. For $j < n-1$, the coefficient is like $(R_\lambda)_{n-j}^n + O(\lambda^{-2}) = \lambda^{-1}(-b_{n-j}^n) + O(\lambda^{-2})$. Since $-b_{n-j}^n = c_{n-j-1}^n - c_{n-j}^n \ge 0$, then one can multiply λ on both sides to take the limit $\lambda \to \infty$. This then yields that $u_n \ge 0$ by simple induction.

If C has positive entries, $c_{n-1}^n > 0$. The coefficient of u_0 is like $\lambda^{-1} c_{n-1}^n$ when λ is large enough. The coefficients in other terms at this order are all nonnegative. This will then naturally yield $u_n > 0$ by choosing λ sufficiently large. ∎

For a typical scheme, since c_j^n is like the average of $g_{1-\alpha}(t_n - s)$ on (t_{n-j-1}, t_{n-j}),

$$c_0^n \sim \frac{1}{\Gamma(2-\alpha)} \tau_n^{-\alpha}, \quad c_1^n \sim \frac{1}{\Gamma(2-\alpha)} \frac{(\tau_n + \tau_{n-1})^{1-\alpha} - \tau_n^{1-\alpha}}{\tau_{n-1}}.$$

Hence, when the step sizes are small, the conditions listed above are expected to hold.

Below, we perform a discussion for the Crank–Nicolson scheme. Consider the so-called $L1^+$ scheme for the derivative at $t_{n-1/2}$:

$$(\mathcal{D}_\tau^\alpha u)^{n-1/2} = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \int_0^t g_{1-\alpha}(t-s)(\Pi_1 u)'(s)\, ds, \tag{3.8}$$

where $\Pi_1 u(s)$ is the piecewise linear approximation of u such that $\Pi_1 u(t_j) = u_j$. Define

$$\chi_{n-j}^n = \frac{1}{\tau_n \tau_k} \int_{t_{n-1}}^{t_n} \int_{t_{k-1}}^{\min(t,t_k)} g_{1-\alpha}(t-s)\, ds dt, \quad 1 \le j \le n. \tag{3.9}$$

Then, the Crank–Nicolson scheme is given by

$$(\mathcal{D}_\tau^\alpha u)^{n-1/2} = \sum_{j=1}^n \chi_{n-j}^n \nabla_\tau u_j = f_n^{1/2}[u]. \tag{3.10}$$

It has been shown in [22] that χ_j^n is not monotone in j because χ_0^n could be smaller than χ_1^n. Only when $\alpha > \alpha_c$ for some critical value $\alpha_c \in (0,1)$, one has $\chi_0^n > \chi_1^n$. If one defines

$$c_0^n = 2\chi_0^n, \quad c_j^n = \chi_j^n, j \ge 1, \tag{3.11}$$

then $C = (c_{n-j}^n)$ is monotone along both the columns and rows with the log-convex condition $c_{j-1}^{n-1} c_{j+1}^n \ge c_j^n c_j^{n-1}$. The kernel associated with C is completely positive (see [6]). The Crank–Nicolson scheme can then be written as

$$C \bar{*} \nabla_\tau u_n = C \bar{*} L^{(-1)} \bar{*} (u - u_0)_n = \chi_0^n (u_n - u_{n-1}) + f_n^{1/2}[u]. \tag{3.12}$$

PROPOSITION 3.5. *If $\alpha > \alpha_c$, the Crank–Nicolson scheme satisfies the following comparison principle. Suppose $f(t,\cdot)$ is uniformly Lipschitz with constant M, and two sequences $\{y_n\}$, $\{z_n\}$ satisfy $(\mathcal{D}_\tau^\alpha y)^{n-1/2} \ge f_n^{1/2}[y]$, $(\mathcal{D}_\tau^\alpha z)^{n-1/2} \le f_n^{1/2}[z]$, with $y_0 \ge z_0$. If the discretization satisfies that $\chi_0^n > M/2$ and $\chi_0^n - \chi_1^n \ge M/2$, then $y_n \ge z_n$. If the function is nondecreasing, one only needs $\chi_0^n > M/2$ for the comparison principle to hold.*

The sequence $u_n = y_n - z_n$ satisfies

$$C \,\bar{*}\, L^{(-1)} \,\bar{*}\, (u_n - u_0) \ge \chi_0^n (u_n - u_{n-1}) + \frac{1}{2}(h_{n,1} u_n + h_{n,2} u_{n-1}), \tag{3.13}$$

where $h_{n,\ell}$ ($\ell = 1, 2$) are the same as in the proof of Theorem 3.4. The argument above can then be carried here with minor modification. We thus skip the proof here.

We remark that some versions of the comparison principles above can also be established using the technique as in the proof of [18, Theorem 3], using the signs of $A^{(-1)}$. However, such a proof heavily relies on the properties in the discretized scheme and has no analogue for the time continuous version.

4. Grönwall inequalities for completely positive schemes. In this section, we will establish some Grönwall inequalities. The versions for $f(u) = -\lambda u + c$, $\lambda > 0$ could be used for uniform error control and decay estimates for dissipative systems. We will only consider implicit schemes. The Grönwall inequality for the weighted implicit or explicit schemes can be obtained similarly.

We start with some basic facts under the following assumption. We recall that $T \in (0, \infty]$ (we allow $T = \infty$) and $\mathcal{N}$ is the set of index n such that $t_n \in (0, T] \cap (0, \infty)$.

Assumption 4.1. There exists $\nu > 0$ such that for all $n \in \mathcal{N}$ one has

$$c_{n-j}^n \ge \nu \frac{1}{\tau_j} \int_{t_{j-1}}^{t_j} g_{1-\alpha}(t_n - s)\, ds \quad \forall 1 \le j \le n. \tag{4.1}$$

This assumption says that c_{n-j}^n is bounded by from below by a fraction of the average of $g_{1-\alpha}(t_n - \cdot)$ on $(t_{j-1}, t_j]$, which is clearly natural for the approximation of the continuous derivative in (1.2).

LEMMA 4.2. *Under Assumption* 4.1*, one has for concave function $v(\cdot)$ with $v'(\cdot) \ge 0$ that*

$$\mathcal{D}_\tau^\alpha v(t_n) \ge \nu \frac{1}{\Gamma(1-\alpha)} \int_0^{t_n} (t_n - s)^{-\alpha} v'(s)\, ds = \nu D^\alpha v(t_n) \quad \forall n \in \mathcal{N}. \tag{4.2}$$

Moreover, for the L1 scheme, $\nu = 1$.

Proof. By Assumption 4.1, consider each term in $\mathcal{D}_\tau^\alpha v(t_n)$, one has

$$c_{n-j}^n (v(t_j) - v(t_{j-1})) \ge \nu \frac{1}{\tau_j} \int_{t_{j-1}}^{t_j} g_{1-\alpha}(t_n - s)\, ds \int_{t_{j-1}}^{t_j} v'(s)\, ds.$$

Here, $v'(s)$ is nonincreasing. By Chebyshev's sorting inequality [11, item 236], one has

$$\frac{1}{\tau_j} \int_{t_{j-1}}^{t_j} g_{1-\alpha}(t_n - s)\, ds \int_{t_{j-1}}^{t_j} v'(s)\, ds \ge \int_{t_{j-1}}^{t_j} g_{1-\alpha}(t_n - s) v'(s)\, ds.$$

In fact, a direct computation can also verify this:

$$\int_{t_{j-1}}^{t_j}\int_{t_{j-1}}^{t_j} g_{1-\alpha}(t_n-s)v'(z)\,dsdz - \int_{t_{j-1}}^{t_j}\int_{t_{j-1}}^{t_j} g_{1-\alpha}(t_n-s)v'(s)\,dsdz$$
$$= \int_{t_{j-1}}^{t_j}\int_{t_{j-1}}^{s} [g_{1-\alpha}(t_n-s)-g_{1-\alpha}(t_n-z)](v'(z)-v'(s))\,dsdz \ge 0.$$

The equality above is by changing the order of integration for the second term. Summing over j and using (4.1), one then obtains the desired result. ∎

As a special case, one has

$$\mathcal{D}_\tau^\alpha\left(\frac{1}{\Gamma(1+\alpha)}t_n^\alpha\right) \ge \nu, \quad n\in\mathcal{N}. \tag{4.3}$$

4.1. A Grönwall inequality for uniform bound. Next, we will consider $T=\infty$ and suppose Assumption 4.1 holds. In this case, $|\mathcal{N}|=\infty$ and we consider all $n\ge 1$. We aim to establish a Grönwall inequality that is useful for the uniform bound if the system is dissipative. This could be useful for a uniform-in-time error estimate.

THEOREM 4.3. *Suppose the discretization is completely positive and Assumption* 4.1 *holds for some* $\nu>0$ *and all* $n\ge 1$. *Consider a nonnegative sequence* $\{v_n\}_{n\in\{0\}\cup\mathcal{N}}$ *satisfying for some* $\lambda\ge 0$ *that*

$$\mathcal{D}_\tau^\alpha v_n \le -\lambda v_n + c, \quad n\in\mathcal{N}. \tag{4.4}$$

If $\lambda>0$ *and* $v_0\le c/\lambda$, *then one has*

$$v_n \le (v_0-c/\lambda)E_\alpha(-\nu^{-1}\lambda t_n^\alpha)+c/\lambda. \tag{4.5}$$

If $\lambda=0$ *and* $v_0\ge 0$, *then one has*

$$v_n \le v_0+\nu^{-1}c\frac{1}{\Gamma(1+\alpha)}t_n^\alpha. \tag{4.6}$$

Proof. Consider only $\lambda>0$ ($\lambda=0$ is similar). Consider the auxiliary problem

$$D^\alpha y^\nu = \nu^{-1}(-\lambda y^\nu + c), y^\nu(0)=v_0.$$

If $v_0\le c/\lambda$, the solution is concave using the explicit formula (2.7) and the fact that $t\mapsto E_\alpha(-\lambda t^\alpha)$ is completely monotone. Then,

$$\mathcal{D}_\tau^\alpha y^\nu \ge \nu(\nu^{-1}(-\lambda y^\nu + c)) = -\lambda y^\nu + c.$$

Since $v\mapsto -\lambda v+c$ is nonincreasing, the comparison principle in Theorem 3.2 holds for any completely positive discretization. Then, it holds that $v_n\le y^\nu(t_n)$, which is the desired result. ∎

This results above hold with no restriction on the step size ratio and the largest step size. They imply that the solution v_n of the implicit scheme $\mathcal{D}_\tau^\alpha v_n = -\lambda v_n + c$ is bounded above by the exact solution when $\mathcal{D}_\tau^\alpha$ is defined using the L1 approximation. In fact, if $\alpha=1$ (the ODE case), it is known that $v_n=(v_0-c/\lambda)(1+\lambda\tau)^{-n}+c/\lambda\le (v_0-c/\lambda)e^{-\lambda n\tau}+c/\lambda = y(t_n)$.

Often for a dissipative system, one may obtain that the error satisfies

$$\mathcal{D}_\tau^\alpha \|e_n\|^2 \le -\lambda\|e_n\|^2 + C\tau^\beta, \tag{4.7}$$

for some $\beta>0$. Provided that $\|e_0\|=0$, the error is controlled uniformly in time as

$$\|e_n\|^2 \le \frac{C\tau^\beta}{\lambda}(1-E_\alpha(-\nu^{-1}\lambda t_n^\alpha)) \le \frac{C}{\lambda}\tau^\beta. \tag{4.8}$$

4.2. A Grönwall inequality for decay estimates. In this subsection, we consider a Grönwall inequality that can control a sequence above by a decreasing sequence. In particular, consider the discrete inequality

$$\mathcal{D}_\tau^\alpha u_n \le f(u_n), \tag{4.9}$$

where $f(u_0) < 0$. We aim to find an upper boud for u_n. We first explain our strategy. Suppose that for a class of nonnegative functions $w(t)$, there is a constant ρ such that

$$c_{n-j}^n \int_{t_{j-1}}^{t_j} w(s)\,ds \le \rho \int_{t_{j-1}}^{t_j} g_{1-\alpha}(t_n - s) w(s)\,ds \quad \forall j \le n. \tag{4.10}$$

Consider the function $u^\rho(\cdot)$ solving

$$D^\alpha u^\rho = \rho^{-1} f(u^\rho), \quad u^\rho(0) = u_0. \tag{4.11}$$

Then $t \mapsto u^\rho(t)$ is nonincreasing and $w(t) = -\frac{d}{dt} u^\rho(t) \ge 0$ for $t > 0$. If $w(\cdot)$ is in the class such that (4.10) holds, one then has

$$\mathcal{D}_\tau^\alpha (u^\rho(t_n)) \ge \rho D^\alpha u^\rho = f(u^\rho). \tag{4.12}$$

The comparison principle in Theorem 3.2 implies that $u_n \le u^\rho(t_n)$.

To establish (4.10), we introduce another assumption.

Assumption 4.4. There exists $\rho_1 > 0$ such that for all $n \in \mathcal{N}$,

$$c_{n-j}^n \le \rho_1 \frac{1}{\tau_j} \int_{t_{j-1}}^{t_j} g_{1-\alpha}(t_n - s)\,ds \quad \forall 1 \le j \le n. \tag{4.13}$$

This assumption is a mirror version of Assumption 4.1, which says that the c_{n-j}^n is bounded above by a multiple of the average of $g_{1-\alpha}(t_n - \cdot)$ on $(t_{j-1}, t_j]$. We emphasize that these two assumptions put no restriction on the step size ratio. In fact, for L1 schemes, c_{n-j}^n equals exactly the average so these two assumptions hold. Clearly, for any nonuniform grid, one can define the L1 scheme.

Since it holds for $s' \in (t_{j-1}, t_j)$ that

$$\frac{1}{\tau_j} \int_{t_{j-1}}^{t_j} (t_n - s)^{-\alpha}\,ds \le \frac{1}{t_n - t_{j-1}} \int_{t_{j-1}}^{t_n} (t_n - s)^{-\alpha}\,ds \le \frac{1}{1-\alpha}(t_n - s')^{-\alpha}, \tag{4.14}$$

one has the following.

LEMMA 4.5. *Suppose Assumption* 4.4 *holds. Then* (4.10) *holds with* $\rho = \frac{\rho_1}{1-\alpha}$.

The above constant ρ is not satisfactory. In fact, if α is close to 1, it blows up, which should not happen in practice. Below, we will consider the special case

$$f(u) = -\lambda u + c, \quad \lambda > 0 \tag{4.15}$$

and make use of the information of the solution given by (2.7), with $\beta = -\lambda$. Clearly, it suffices to consider $y(t) = E_\alpha(-\lambda t^\alpha)$. Define

$$w(t) := -y'(t). \tag{4.16}$$

Since $y(\cdot)$ is completely monotone, w is also completely monotone. Our observation is that there could be another constant σ such that

$$\frac{1}{\tau_j} \int_{t_{j-1}}^{t_j} w(s)\,ds \le \sigma w(t_j), \tag{4.17}$$

which is good for $\alpha \in [1/2, 1)$. Then, one can take $\rho = \rho_1 \min(1/(1-\alpha), \sigma)$ for $\alpha \in [1/2, 1)$.

LEMMA 4.6. *The function $w(t)$ in (4.16) is a completely monotone function so that $t \mapsto \frac{w(t+\tau)}{w(t)}$ is nondecreasing for any $\tau > 0$. Consequently, for $\alpha \in [1/2, 1)$ and τ_n with $\lambda\tau_n^\alpha \le 1$, one has for two universal constants c_1, c_2 such that*

$$\frac{1}{\tau_j}\int_{t_{j-1}}^{t_j} w(s)\,ds \le \frac{c_2}{c_1}\max(2^{1-\alpha}, \alpha^{-1})w(t_j). \tag{4.18}$$

One can thus take $\sigma = 2c_2/c_1 > 1$ for (4.17).

Proof. Since $y(\cdot)$ is completely monotone, $w(t) = -y'(t)$ is completely monotone by definition. By the discussion in [2, 29], $t \mapsto w(t+\tau)/w(t)$ is nondecreasing. Let n_0 be the smallest one such that $\lambda t_{n_0}^\alpha \ge 1$. For all $t \ge t_{n_0}$ and τ with $\lambda\tau^\alpha \le 1$, one has $\tau \le t_{n_0}$. Recalling that $w(t) = \alpha\lambda t^{\alpha-1}(\alpha^{-1}E_{\alpha,\alpha}(-\lambda t^\alpha))$, one thus has

$$\frac{w(t)}{w(t+\tau)} \le \frac{w(t_{n_0})}{w(t_{n_0}+\tau)} = \left(\frac{t_{n_0}}{t_{n_0}+\tau}\right)^{\alpha-1} \frac{\alpha^{-1}E_{\alpha,\alpha}(-\lambda t_{n_0}^\alpha)}{\alpha^{-1}E_{\alpha,\alpha}(-\lambda(t_{n_0}+\tau)^\alpha)}.$$

The right-hand side can be bounded easily. In fact, the function (see [27])

$$z \mapsto \alpha^{-1}E_{\alpha,\alpha}(-z) = \sum_{k=0}^{\infty}\frac{(k+1)(-1)^k z^k}{\Gamma((k+1)\alpha+1)} = E_\alpha'(-z) = \int_0^\infty M_\alpha(r) r e^{-rz}\,dr$$

is positive, continuous, and nonincreasing in $z > 0$ ($M_\alpha(\cdot)$ is nonnegative). Moreover, it is continuous in α on $[1/2, 1]$. Hence, there exist $0 < c_1 < c_2$ such that for all $\alpha \in [1/2, 1)$ and all $z \le 3$, $c_1 \le \alpha^{-1}E_{\alpha,\alpha}(z) \le c_2$. Clearly, $\lambda(t_{n_0-1} + \tau_{n_0} + \tau)^\alpha \le 3^\alpha \lambda \max(t_{n_0}, \tau_{n_0}, \tau)^\alpha \le 3^\alpha$. Then, $\lambda(t_{n_0} + \tau)^\alpha \in (0, 3]$ and $w(t)/w(t+\tau) \le w(t_{n_0})/w(t_{n_0}+\tau) \le 2^{1-\alpha}c_2/c_1$. Hence, for $j \ge n_0$,

$$\frac{1}{\tau_j}\int_{t_{j-1}}^{t_j} w(s)\,ds \le \frac{2^{1-\alpha}c_2}{c_1} w(t_j).$$

Next, for $j \le n_0 - 1$, one has $c_1\lambda\alpha t^{\alpha-1} \le w(t) \le c_2\lambda\alpha t^{\alpha-1}$. It follows that

$$\frac{1}{\tau_j}\int_{t_{j-1}}^{t_j} w(s)\,ds \le \frac{1}{t_j}\int_0^{t_j} w(s)\,ds \le \frac{c_2\lambda t_j^{\alpha-1}}{\alpha} \le \frac{c_2}{c_1\alpha}w(t_j).$$

Combining these two cases, the conclusion then follows. ∎

Remark 4.7. The upper bound in $\lambda\tau_n^\alpha \le 1$ is not crucial. In fact, for fixed $\gamma > 0$, if we restrict $\lambda\tau_n^\alpha \le \gamma$, then we can find corresponding constant σ.

One can then establish the following Grönwall inequalities.

THEOREM 4.8. *Consider a variable-step completely positive discretization and a nonnegative sequence $\{v_n\}_{n\in\{0\}\cup\mathcal{N}}$. Let $\lambda > 0$ and $v_0 > c/\lambda$.*

(1) *Suppose Assumption 4.1 holds for some $\nu > 0$. If $\mathcal{D}_\tau^\alpha v_n \ge -\lambda v_n + c$ for all $n \in \mathcal{N}$, then it holds that*

$$v_n \ge \left(v_0 - \frac{c}{\lambda}\right) E_\alpha(-\nu^{-1}\lambda t_n^\alpha) + \frac{c}{\lambda}. \tag{4.19}$$

(2) *Suppose Assumption 4.4 holds for some $\rho_1 > 0$. If $\mathcal{D}_\tau^\alpha v_n \le -\lambda v_n + c$ for all $n \in \mathcal{N}$, then it holds that*

$$v_n \le \left(v_0 - \frac{c}{\lambda}\right) E_\alpha\left(-\frac{\lambda(1-\alpha)}{\rho_1} t_n^\alpha\right) + \frac{c}{\lambda}. \tag{4.20}$$

If, moreover, it holds that $\lambda\tau_n^\alpha \le 1$ for all $n \in \mathcal{N}$, then for the universal constant σ introduced in Lemma 4.6, it holds that

$$v_n \le \left(v_0 - \frac{c}{\lambda}\right) E_\alpha\left(-\frac{\lambda}{\sigma\rho_1}t_n^\alpha\right) + \frac{c}{\lambda}. \tag{4.21}$$

Proof. (1) Consider the auxiliary equation

$$D_c^\alpha y^\nu = \nu^{-1}(-\lambda y^\nu + c), y^\nu(0) = y_0.$$

Now that $v_0 > c/\lambda$, y^ν is a nonincreasing, convex function. Then, one may apply Lemma 4.2 to $-y^\nu$ to obtain

$$\mathcal{D}_\tau^\alpha y^\nu \le \nu D^\alpha y^\nu(t_n) = -\lambda y^\nu + c.$$

The comparison principle in Theorem 3.2 then implies $v_n \ge y^\nu(t_n)$, giving the desired result.

(2) Consider the auxiliary problem

$$D_c^\alpha y^\rho = \rho^{-1}(-\lambda y^\rho + c), y^\rho(0) = y_0.$$

By Lemma 4.5, one can take $\rho = \rho_1/(1-\alpha)$ and then

$$\mathcal{D}_\tau^\alpha y^\rho \ge -\lambda y^\rho + c.$$

Then, by Theorem 3.2, one concludes that $y(t_n) \le y^\rho(t_n)$.

If the discretization satisfies $\lambda\tau_n^\alpha \le 1$, motivated by Lemmas 4.5 and 4.6, one can take for all $\alpha \in (0,1)$ that

$$\rho = \rho_1 \frac{2c_2}{c_1} = \rho_1\sigma.$$

Note that for $\alpha \le 1/2$, $\rho_1/(1-\alpha) < \rho_1\sigma$ so taking $\rho = \rho_1\sigma$ also works for $\alpha \le 1/2$. Now, since $\rho > 1$, $\rho^{-1}\lambda\tau_n^\alpha \le 1$ holds, Lemma 4.6 applying to $y^\rho(\cdot)$ then gives

$$\mathcal{D}_\tau^\alpha y^\rho(t_n) \ge -\lambda y^\rho + c.$$

The comparison principle in Theorem 3.2 then yields the last result. ∎

COROLLARY 4.9. *Consider a variable-step completely positive discretization. If $\mathcal{D}_\tau^\alpha v_n \le -\lambda v_n$ for a nonnegative sequence $\{v_n\}_{n\in\{0\}\cup\mathcal{N}}$ and some $\lambda > 0$, v_n has an upper bound $v_n \le v_0 E_\alpha(-\frac{\lambda}{\rho}t_n^\alpha)$, where $\rho = \rho_1/(1-\alpha)$. If $\lambda\tau_n^\alpha \le 1$ for all $n \in \mathcal{N}$, one can take $\rho = \rho_1\sigma$.*

Remark 4.10. For $L1$ scheme, the solution to the implicit scheme satisfies

$$\left(v_0 - \frac{c}{\lambda}\right) E_\alpha(-\lambda t_n^\alpha) + \frac{c}{\lambda} \le v_n \le \left(v_0 - \frac{c}{\lambda}\right) E_\alpha\left(-\frac{\lambda}{\rho}t_n^\alpha\right) + \frac{c}{\lambda}, \tag{4.22}$$

where $\rho_1 = 1$ so that $\rho = \min(1/(1-\alpha), \sigma)$. The lower bound holds for any discretization, while the upper bound holds for $\lambda\tau_n^\alpha \le 1$. This means that the solution of the implicit scheme is above the exact solution.

4.3. A Grönwall inequality for growing linear functions. We aim to establish a Grönwall inequality for $f(u) = \lambda u + c$. Compared to [23], we aim to remove the requirement on the step size ratio. The key is again to show for some $\mu \in (0, 1]$ that $\mathcal{D}_\tau^\alpha y \geq \mu D^\alpha y$ where y is the solution to the auxiliary equation. We assume Assumption 4.1 and need

$$c_{n-j}^n (y(t_j) - y(t_j - 1)) \geq \frac{\mu}{\Gamma(1-\alpha)} \int_{t_{j-1}}^{t_j} (t_n - s)^{-\alpha} y'(s)\, ds. \tag{4.23}$$

There are two special cases to guarantee this:

- If $y''(s) \leq 0$ for $s \in [t_{j-1}, t_j]$, then under Assumption 4.1, one can take $\mu = \nu$.
- If $c_{n-j}^n \geq \mu_1 (t_n - s)^{-\alpha}$ or $\tau_j^{-1}(y(t_j) - y(t_{j-1})) \geq \mu_1 y'(s)$ for $s \in [t_{j-1}, t_j]$, one can then take $\mu = \nu\mu_1$.

The function $y(s)$ is not concave unless $\lambda = 0$. It is concave only near $t = 0$. In fact, using $y'(s) = \lambda s^{\alpha-1} E_{\alpha,\alpha}(\lambda s^\alpha)$ and the power series of Mittag–Leffler function in (2.8), one has the following claim.

LEMMA 4.11. *Let $w(t) = y'(t)$ where $y(t) = E_\alpha(\lambda t^\alpha)$, $\lambda > 0$ and $\alpha \in (0, 1)$. Then, there exists $t_* > 0$ such that w is decreasing on $(0, t_*)$ while increasing on (t_*, ∞).*

Proof. Since $w(t) = \lambda t^{\alpha-1} E_{\alpha,\alpha}(\lambda t^\alpha)$, we set $z = \lambda t^\alpha$. Then,

$$w = \lambda^{1/\alpha} z^{(\alpha-1)/\alpha} E_{\alpha,\alpha}(z) =: F(z).$$

Recalling the power series of $E_{\alpha,\alpha}(z)$ in (2.8), one has

$$\begin{aligned} \lambda^{-1/\alpha} F'(z) &= (1 - \alpha^{-1}) z^{-\alpha^{-1}} E_{\alpha,\alpha}(z) + z^{1-\alpha^{-1}} E'_{\alpha,\alpha}(z) \\ &= z^{-\alpha^{-1}} \sum_{k=0}^{\infty} \frac{(1 - \alpha^{-1} + k) z^k}{\Gamma(\alpha(k+1))}. \end{aligned}$$

Let k_0 be the integer such that $1 - \alpha^{-1} + k_0 \leq 0$ while $1 - \alpha^{-1} + k_0 + 1 > 0$. Such k_0 exists and $k_0 \geq 0$. Define

$$A(z) = \sum_{k=0}^{k_0} \frac{(\alpha^{-1} - k - 1) z^k}{\Gamma(\alpha(k+1))}, \quad B(z) = \sum_{k=k_0+1}^{\infty} \frac{(1 - \alpha^{-1} + k) z^k}{\Gamma(\alpha(k+1))}.$$

Let z_* be the first point such that $A(z_*) = B(z_*)$. It is clear that z_* exists and $z_* > 0$. For $z < z_*$, $B(z) - A(z) < 0$. For any $z > z_*$, let $\xi := z/z_* > 1$. Then, it is clear that $A(z) \leq \xi^{k_0} A(z_*)$ while $B(z) > \xi^{k_0} B(z_*)$. Then, $B(z) - A(z) > \xi^{k_0}(B(z_*) - A(z_*)) = 0$. This means that there is only one point z_* such that $F'(z)$ is zero and thus only one t_* such that $w'(t_*) = 0$. ∎

Then, for $t_n < t_*$, one can use the concavity as in Lemma 4.2. For large t_n, we turn to the second case above. Clearly, $c_{n-j}^n \geq \mu_1 (t_n - s)^{-\alpha}$ for $s \in [t_{j-1}, t_j]$ cannot hold if s is near t_n. Hence, we seek a lower bound for $y'(t)/y'(t + \tau)$. By the integral representation (2.9), one finds that

$$E_{\alpha,\beta}(\lambda t^\alpha) = \frac{1}{\alpha} \lambda^{(1-\beta)/\alpha} t^{(1-\beta)} e^{\lambda^{1/\alpha} t} + \frac{1}{2\pi i \alpha} \int_{\gamma(\epsilon;\delta)} \frac{e^{\zeta^{1/\alpha}} \zeta^{(1-\beta)/\alpha}}{\zeta - \lambda t^\alpha} d\zeta. \tag{4.24}$$

The first term grows exponentially while the second term goes to zero algebraically as $t \to \infty$. This observation gives a way to control $y'(t)/y'(t + \tau)$ for t large.

LEMMA 4.12. *Let $u(t)=E_\alpha(\lambda t^\alpha)$ and $w(t):=u'(t)=\lambda t^{\alpha-1}E_{\alpha,\alpha}(\lambda t^\alpha)$. For any $0\le t<t+\tau\le t'$ with $\lambda\tau^\alpha\le 1$ where t' is a given upper bound, there is a universal constant $\mu_1\in(0,1)$ (independent of all parameters) such that*

$$\frac{1}{\tau}\int_t^{t+\tau}(t'-s)^{-\alpha}\,ds\int_t^{t+\tau}w(s)\,ds\ge \mu_1\int_t^{t+\tau}(t'-s)^{-\alpha}w(s)\,ds. \tag{4.25}$$

Proof. Note that w decreases first and then increases. Let t_* be the transition point. We aim to find $\mu_1\in(0,1)$ such that when $t_*\le s<s+\tau'$ with $\tau'\le\tau$ such that

$$\frac{w(s)}{w(s+\tau')}\ge\mu_1. \tag{4.26}$$

We first prove the claim by assuming (4.26). In fact,

- If $t+\tau\le t_*$, then w is nonincreasing on this interval and thus,
$$\frac{1}{\tau}\int_t^{t+\tau}(t'-s)^{-\alpha}\,ds\int_t^{t+\tau}w(s)\,ds\ge \int_t^{t+\tau}(t'-s)^{-\alpha}w(s)\,ds.$$
- If $t\ge t_*$, then by (4.26), one has
$$\frac{1}{\tau}\int_t^{t+\tau}(t'-s)^{-\alpha}\,ds\int_t^{t+\tau}w(s)\,ds\ge \mu_1\int_t^{t+\tau}(t'-s)^{-\alpha}w(t+\tau)\,ds.$$
- For $t_*\in(t,t+\tau)$, define
$$\tilde{w}(s)=\begin{cases}w(s), & s\le t_*,\\ w(t_*), & s>t_*\end{cases}$$
Then, $\tilde{w}$ is nonincreasing so one has by Chebyshev's sorting inequality that
$$\begin{aligned}\frac{1}{\tau}\int_t^{t+\tau}(t'-s)^{-\alpha}\,ds\int_t^{t+\tau}w(s)\,ds&\ge \frac{1}{\tau}\int_t^{t+\tau}(t'-s)^{-\alpha}\,ds\int_t^{t+\tau}\tilde{w}(s)\,ds\\ &\ge\int_t^{t+\tau}(t'-s)^{-\alpha}\tilde{w}(s)\,ds.\end{aligned}$$
For $s\ge t_*$, $\tilde{w}(s)=w(t_*)\ge\mu_1 w(s)$. Hence, the claim still holds for this case.

Next, we establish (4.26). Letting $z=\lambda t^\alpha$ and taking $\delta=\alpha\pi$ in (4.24), for $\alpha\in(0,1)$ and $\beta=\alpha<1+\alpha$, one can take $\epsilon\to0$ to have

$$\begin{aligned}E_{\alpha,\alpha}(z)&=\frac{1}{\alpha}z^{(1-\alpha)/\alpha}e^{z^{1/\alpha}}+\frac{1}{\pi\alpha}\int_0^\infty r^{(1-\alpha)/\alpha}e^{-r^{1/\alpha}}\frac{r\sin(\pi(1-\alpha))}{r^2-2rz\cos(\pi\alpha)+z^2}dr\\ &=:I_1+I_2.\end{aligned}$$

Using $\sin(\pi\alpha/2)\ge\alpha$ and the simple bound $e^{-r^{1/\alpha}}r^{1/\alpha}\le e^{-1}$, one has

$$I_2\le\frac{\sin(\pi\alpha)}{ze\pi\alpha}\int_{-1}^\infty\frac{1}{y^2+4(y+1)\alpha^2}dy\le\frac{\sin(\pi\alpha)}{ze\pi\alpha}\Big(\frac{\pi}{4\alpha}+\frac{\pi}{2\sqrt{2}\alpha}+1\Big)\le\frac{\sin(\pi\alpha)}{\pi\alpha}\frac{\pi}{z\alpha e}.$$

Here, the integral is compared to that of $1/(y^2+4\alpha^2)$, $1/(y^2+2\alpha^2)$, and $1/y^2$, respectively, for $y\ge0$, $-1/2\le y<0$, and $y<-1/2$. Hence, one has

$$\frac{I_2(z)}{I_1(z)}\le\frac{\pi}{z^{1/\alpha}e^{1+z^{1/\alpha}}}. \tag{4.27}$$

Below, we consider two cases.

Case 1. $\alpha \geq 1/2$.

For $z \geq 1$, one has $I_2(z)/I_1(z) \leq \pi/e^2$. Decompose

$$w(t) = \lambda t^{\alpha-1} I_1(\lambda t^\alpha) + \lambda t^{\alpha-1} I_2(\lambda t^\alpha) =: w_1(t) + w_2(t).$$

Then, $w_1(t) = \alpha^{-1}\lambda^{1/\alpha} e^{\lambda^{1/\alpha} t}$, and $w_2/w_1 \leq \pi/e^2$, which implies that

$$\frac{w(t)}{w(t+\tau)} \geq \frac{w_1(t)}{w_1(t+\tau) + w_2(t+\tau)} \geq \frac{1}{1+\frac{\pi}{e^2}} \frac{w_1(t)}{w_1(t+\tau)} \geq \frac{1}{1+\frac{\pi}{e^2}} e^{-1}.$$

For $z \leq 1$, one has

$$\alpha^{-1} E_{\alpha,\alpha}(z) = \sum_{k=0}^\infty \frac{z^k}{\alpha\Gamma((k+1)\alpha)} = \sum_{k=0}^\infty \frac{(k+1)z^k}{\Gamma((k+1)\alpha+1)}.$$

Since $4/5 < \Gamma(1+\alpha) < 1$ and the $\alpha \mapsto \Gamma((k+1)\alpha+1)$ is increasing since $(k+1)\alpha+1 \geq 1.5$, one finds that $\alpha^{-1}E_{\alpha,\alpha}(z)$ is uniformly bounded as

$$1 \leq \alpha^{-1} E_{\alpha,\alpha}(z) \leq 2E_{1/2,1/2}(1) =: c'.$$

Then, one finds that $\alpha\lambda t^{\alpha-1} \leq w(t) \leq \alpha\lambda t^{\alpha-1} c'$.

Combining these two results, if t_* is the transition point for w, then for all $t_* \leq s \leq s+\tau'$ with $\lambda(\tau')^\alpha \leq 1$:

$$\frac{w(s)}{w(s+\tau')} \geq \frac{1}{c'} \frac{1}{1+\frac{\pi}{e^2}} e^{-1} =: \mu_{1,1}.$$

Case 2. $\alpha \leq 1/2$.

We take $M > 1$ to be determined later. For $z \geq M^{-\alpha}$, one has by (4.27) that

$$\frac{I_2(z)}{I_1(z)} \leq \frac{M\pi}{e^{1+1/M}}.$$

Using similar argument as above, if $z = \lambda t^\alpha \geq M^{-\alpha}$, one has

$$\frac{w(t)}{w(t+\tau)} \geq \frac{1}{1+\frac{M\pi}{e^{1+1/M}}} e^{-1} =: \mu_{1,2}.$$

We show that w is monotone for $z = \lambda t^\alpha < M^{-\alpha}$ if M is large so that $\lambda t_*^\alpha \geq M^{-\alpha}$. Then, taking $\mu_1 = \min(\mu_{1,1}, \mu_{1,2})$ clearly makes (4.26) hold.

Recall

$$w(t) = F(\lambda t^\alpha), \quad F(z) = \lambda^{1/\alpha} z^{(\alpha-1)/\alpha} E_{\alpha,\alpha}(z).$$

Since $dz/dt > 0$, one need only consider

$$\frac{d}{dz}\log F(z) = \frac{\alpha-1}{\alpha z} + \frac{F_1'(z)}{F_1(z)}, \quad F_1(z) = \alpha^{-1}E_{\alpha,\alpha}(z).$$

Noting that $\inf_{a>0}\Gamma(a) > 0.8$, one has

$$F_1'(z) = \sum_{k=1}^\infty \frac{(k+1)kz^{k-1}}{\Gamma((k+1)\alpha+1)} \leq \frac{5}{4}\sum_{k=1}^\infty (k+1)kz^{k-1} = \frac{5}{4}\frac{1}{(1-z)^3}.$$

Moreover, since $\Gamma(a)<1$ for $a\in(1,2)$, one has

$$F_1(z)=\sum_{k=0}^{\infty}\frac{(k+1)z^k}{\Gamma((k+1)\alpha+1)}\ge\sum_{k=0}^{[1/\alpha]}(k+1)z^k=\frac{d}{dz}\frac{1-z^{[1/\alpha]+2}}{1-z}\ge\frac{d}{dz}\frac{1-z^{1/\alpha}}{1-z}.$$

Here, we used the fact that $\frac{d}{dz}\frac{z^m}{1-z}=\frac{z^{m-1}[m(1-z)+z]}{(1-z)^2}$ is decreasing in m for $m>1$. In fact, the derivative of $m\mapsto\log(z^{m-1}[m(1-z)+z])$ is negative. These imply that

$$\frac{F_1'(z)}{F_1(z)}\le\frac{5}{2}\frac{1}{(1-z)(1-(1-z+\alpha z)z^{1/\alpha}/(\alpha z))}.$$

Consider the expression

$$A:=(1-z+\alpha z)z^{1/\alpha}/(\alpha z)=(1-z)\frac{z^{1/\alpha}}{\alpha z}+z^{1/\alpha}<(1-z)\frac{z^{1/\alpha}}{\alpha z}+M^{-1}.$$

Optimizing over $\alpha\in(0,1/2]$, one has

$$\frac{z^{1/\alpha}}{\alpha}\le\begin{cases}2z^2, & z\le 1/\sqrt{e},\\ \dfrac{1}{-\ln(z)e}, & z>1/\sqrt{e}.\end{cases}$$

Then, one finds that $A<1/2+1/M$ (note that $e^{-1}(z^{-1}-1)/\ln(1+(z^{-1}-1))$ is monotone in the case $z>1/\sqrt{e}$). Hence, for $z<M^{-\alpha}$,

$$\frac{F_1'(z)}{F_1(z)}\le\frac{5}{1-2M^{-1}}\frac{1}{1-z}.$$

Moreover, $1-z\ge 1-M^{-\alpha}>\alpha(\ln M)M^{-\alpha}>\alpha(\ln M)z$. Then, for $z<M^{-\alpha}$, one has

$$\frac{d}{dz}\log F(z)<\frac{\alpha-1+5/(\ln M(1-2M^{-1}))}{\alpha z},\quad \alpha\le 1/2.$$

Hence, for M large enough, this is negative. This means that $\lambda t_*^{\alpha}\ge M^{-\alpha}$. ∎

THEOREM 4.13. *Consider a variable-step completely positive discretization. Suppose Assumption* 4.1 *holds for some* $\nu>0$. *Let* $\mu:=\nu\mu_1$ *where* μ_1 *is the constant in Lemma* 4.12. *Assume* $\lambda\tau_n^{\alpha}<\min(\mu,\nu/\Gamma(2-\alpha))$ *for all* $n\in\mathcal{N}$. *Suppose a nonnegative sequence* $\{v_n\}_{n\in\{0\}\cup\mathcal{N}}$ *satisfies for some* $\lambda>0$ *that*

$$\mathcal{D}_\tau^{\alpha}v_n\le\lambda v_n+c,\quad n\in\mathcal{N}. \tag{4.28}$$

Then one has for $n\in\mathcal{N}$ *that*

$$v_n\le\left(v_0+\frac{c}{\lambda}\right)E_\alpha(\mu^{-1}\lambda t_n^{\alpha})-\frac{c}{\lambda}. \tag{4.29}$$

Proof. Consider the equation

$$D^{\alpha}y^{\mu}=\mu^{-1}(\lambda y^{\mu}+c),\quad y^{\mu}(0)=y_0.$$

Note that $\mu^{-1}\lambda\tau_n^{\alpha}<1$. By Lemma 4.12, one has

$$\mathcal{D}_\tau^{\alpha}y^{\mu}(t_n)\ge\lambda y^{\mu}+c.$$

If $\lambda\tau_n^{\alpha}<\nu/\Gamma(2-\alpha)$, one then has $c_0^n>M$. By the comparison principle in Theorem 3.2 and formula (2.7), one has $v_n\le y^{\mu}(t_n)$, which is the desired result. ∎

5. Applications to dissipative systems. In this section, we consider two examples for the dissipative systems to illustrate how the Grönwall inequalities above can be applied. The first is the standard subdiffusion equation while the second is the time fractional Allen–Cahn equation.

5.1. Example 1: Subdiffusion equation. Consider the following subdiffusion equation on a bounded domain $\Omega \subset \mathbb{R}^d$ with $d \geq 1$:

$$\begin{aligned} &D_c^\alpha u = \Delta u + f(x), \quad x \in \Omega \\ &u|_{\partial\Omega} = 0, \quad u(x,0) = u_0(x). \end{aligned} \tag{5.1}$$

Consider the approximation where the time derivative is discretized by the L1 scheme on nonuniform mesh (i.e., (2.27)–(2.28)) and the Laplace operator Δ is discretized by the centered difference method Δ_h with spatial step h. Then, one has

$$\mathcal{D}_\tau^\alpha u_n = \Delta_h u_n + f(x). \tag{5.2}$$

The truncation error for spatial derivative is clearly $O(h^2)$. Regarding the temporal truncation error, if the solution is assumed to be smooth, the truncation error for $L1$ scheme is $\tau^{2-\alpha}$ [25], where τ is the maximum time step. However, taking into account the singularity near $t = 0$, the truncation error reduces to τ. Using graded mesh, $t_n = T(n/N)^r$ can improve the accuracy [32]. Here, we consider the truncation error on a general nonuniform mesh

$$R_n := \mathcal{D}_\tau^\alpha u(\cdot, t_n) - D_c^\alpha u(\cdot, t_n) = \mathcal{D}_\tau^\alpha u(\cdot, t_n) - (\Delta_h u(\cdot, t_n) + f(\cdot)). \tag{5.3}$$

Let Ω_h be the set of spatial grid points and consider

$$\langle u, v\rangle_{\Omega_h} := \sum_{x \in \Omega_h} u(x)v(x)h^d, \quad \|u\|_{\ell^2}^2 := \langle u, u\rangle_{\Omega_h}. \tag{5.4}$$

Then, the truncation error is bounded by

$$\|R_n\|_{\ell^2} \leq C(\tau + h^2), \tag{5.5}$$

where $\tau = \max_i \tau_i$, C is independent of n.

For the discrete Laplacian Δ_h, there exists a constant $\kappa > 0$ such that for all discrete functions v being zero on $\partial\Omega_h$, one has

$$-\langle v, \Delta_h v\rangle_{\Omega_h} \geq \kappa \|v\|_{\ell^2}^2. \tag{5.6}$$

We have the following conclusions.

PROPOSITION 5.1. *Let u_∞ be the steady solution of* (5.1), *and let u_∞^h be the steady solution of the numerical scheme. If $\kappa\tau_n^\alpha \leq 1$, then for the universal constant σ introduced in Lemma* 4.6, *one has*

$$\|u_n - u_\infty^h\|_{\ell^2} \leq \|u_0 - u_\infty^h\|_{\ell^2} E_\alpha(-\sigma^{-1}\kappa t_n^\alpha). \tag{5.7}$$

Moreover, the error satisfies

$$\sup_n \|u_n - u(t_n)\|_{\ell^2} \leq \frac{C}{\kappa}(\tau + h^2)(1 - E_\alpha(-\kappa t_n^\alpha)). \tag{5.8}$$

Consequently, $\|u_\infty - u_\infty^h\|_{\ell^2} \leq C(\tau + h^2)$.

Proof. Since u_∞^h is a steady solution to the numerical scheme, one has

$$\mathcal{D}_\tau^\alpha(u_n - u_\infty^h) = \Delta_h(u_n - u_\infty^h).$$

Then, by Lemma 2.11,

$$\mathcal{D}_\tau^\alpha \|u_n - u_\infty^h\|_{\ell^2} \le \left\langle \frac{u_n - u_\infty^h}{\|u_n - u_\infty^h\|_{\ell^2}}, \mathcal{D}_\tau^\alpha(u_n - u_\infty^h) \right\rangle \le -\kappa \|u_n - u_\infty^h\|_{\ell^2}.$$

Theorem 4.8 then gives the desired result (Assumption 4.4 holds with $\rho_1 = 1$ for L1 discretization).

By the definition of truncation error,

$$\mathcal{D}_\tau^\alpha u(\cdot, t_n) = \Delta_h u(\cdot, t_n) + f(\cdot) + R_n.$$

If one defines the error $e_n := u_n - u(\cdot, t_n)$, one has

$$\mathcal{D}_\tau^\alpha e_n = \Delta_h e_n - R_n.$$

Pairing with $\frac{e_n}{\|e_n\|}$, one has

$$\mathcal{D}_\tau^\alpha \|e_n\|_{\ell^2} \le -\kappa \|e_n\|_{\ell^2} + C(\tau + h^2).$$

Applying Theorem 4.3 (with $v_0 = 0$ and $\nu = 1$) gives the desired control to the error. ☐

5.2. Example 2: Time fractional Allen–Cahn equation. We consider the following one-dimensional (1D) time fractional Allen–Cahn equation as an example [14]:

$$\begin{aligned} &D_c^\alpha u = \kappa^2 \partial_{xx} u + (u - u^3), \quad x \in \mathbb{T}, \\ &u|_{t=0} = u_0. \end{aligned} \tag{5.9}$$

Here, $\mathbb{T}$ is the 1D torus with length 2π (i.e., $[-\pi, \pi)$ with periodic boundary condition). The equation is associated with a free energy

$$E(u) = \int_{\mathbb{T}} \frac{\kappa^2}{2} |\partial_x u|^2 + \frac{1}{4}(u^2 - 1)^2 \, dx, \tag{5.10}$$

and the equation is actually the time fractional gradient flow of this free energy in $L^2(\mathbb{T})$,

$$D_c^\alpha u = -\frac{\delta E}{\delta u}. \tag{5.11}$$

See [18] for some discussion on how one uses the discretization to analyze the behaviors of time fractional gradient flows.

Below, we consider $\kappa > 1$ and the discretization with $L1$ scheme:

$$\mathcal{D}_\tau^\alpha u_n = \kappa^2 D^2 u_n + (u_n - u_n^3). \tag{5.12}$$

Here, D^2 means that the spatial derivative is discretized by the Fourier spectral method with uniform spatial stepsize h. For any $p > 0$, there exists $C > 0$ such that the truncation error satisfies

$$r_n := \|\mathcal{D}_\tau^\alpha u(\cdot, t_n) - (\kappa^2 \partial_{xx} u_n + (u_n - u_n^3))\|_{\ell^2} \le C(\tau + h^p). \tag{5.13}$$

The truncation error for the time discretization has been discussed above in the first example. The spatial truncation error is standard for spectral method. For spectral discretization, one has for a sequence v with zero Fourier mode $\hat{v}_0 = 0$ that

$$-\langle v, D^2 v \rangle \ge \|v\|_{\ell^2}^2. \tag{5.14}$$

Proposition 5.2. *Suppose that u_0 is an odd function on $[-\pi,\pi]$ and $\kappa^2>1$. Assume also that $(\kappa^2-1)\tau_n^\alpha \le 1$, then the L^2 norm of the numerical solution satisfies*

$$\|u_n\|_{\ell^2} \le CE_\alpha(-\sigma^{-1}(\kappa^2-1)t_n^\alpha) \sim Ct_n^{-\alpha}, \quad n\to\infty. \tag{5.15}$$

The error satisfies

$$\|u(t_n)-u_n\|_{\ell^2} \le \frac{C(\tau+h^p)}{\kappa^2-1}\left(1-E_\alpha(-(\kappa^2-1)t_n^\alpha)\right). \tag{5.16}$$

If u_0 is an odd function, then the zero Fourier mode of u_n preserves to be zero. Then,

$$\begin{aligned}\left\langle \frac{u_n}{\|u_n\|_{\ell^2}}, \kappa^2 D^2 u_n + (u_n - u_n^3)\right\rangle &\le -\kappa^2\|u_n\|_{\ell^2} + \|u_n\|_{\ell^2} - \frac{1}{\|u_n\|_{\ell^2}}\|u_n\|_{\ell^4}^4 \\ &\le -(\kappa^2-1)\|u_n\|_{\ell^2}.\end{aligned} \tag{5.17}$$

The detailed proof would be the same as that for Proposition 5.1, so we omit it.

REFERENCES

[1] M. Caputo, *Linear models of dissipation whose Q is almost frequency independent*–II, Geophys. J. Int., 13 (1967), pp. 529–539.
[2] P. Clément and J. A. Nohel, *Abstract linear and nonlinear Volterra equations preserving positivity*, SIAM J. Math. Anal., 10 (1979), pp. 365–388, https://doi.org/10.1137/0510035.
[3] P. Clément and J. A. Nohel, *Asymptotic behavior of solutions of nonlinear Volterra equations with completely positive kernels*, SIAM J. Math. Anal., 12 (1981), pp. 514–535.
[4] E. Cuesta, C. Lubich, and C. Palencia, *Convolution quadrature time discretization of fractional diffusion-wave equations*, Math. Comp., 75 (2006), pp. 673–696.
[5] K. Diethelm, *The Analysis of Fractional Differential Equations: An Application-oriented Exposition Using Differential Operators of Caputo Type*, Springer, 2010.
[6] Y. Feng and L. Li, *A class of monotonicity-preserving variable-step discretizations for Volterra integral equations*, BIT, 64 (2024), 24.
[7] Y. Feng and L. Li, *On the completely positive kernels for nonuniform meshes*, Quart. Appl. Math., 82 (2024), pp. 801–817.
[8] Y. Feng, L. Li, J.-G. Liu, and X. Xu, *Continuous and discrete one dimensional autonomous fractional ODEs*, Discrete Contin. Dyn. Syst. Ser. B, 23 (2018), pp. 3109–3135.
[9] Y. Feng, L. Li, J.-G. Liu, and X. Xu, *A note on one-dimensional time fractional ODEs*, Appl. Math. Lett., 83 (2018), pp. 87–94.
[10] R. Gorenflo, J. Loutchko, and Y. Luchko, *Computation of the Mittag-Leffler function $E_{\alpha,\beta}(z)$ and its derivative*, Fract. Calc. Appl. Anal., 5 (2002), pp. 491–518.
[11] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, 1952.
[12] B. Jin, *Fractional Differential Equations: An Approach Via Fractional Derivatives*, Appl. Math. Sci. 206., Springer, Cham, 2021.
[13] S. C. Kou and X. S. Xie, *Generalized Langevin equation with fractional Gaussian noise: Subdiffusion within a single protein molecule*, Phys. Rev. Lett., 93 (2004), 180603.
[14] D. Li, C. Quan, T. Tang, and W. Yang, *Sharp convergence to steady states of Allen-Cahn*, Commun. Math. Anal. Appl., 1 (2022), pp. 355–394.
[15] D. Li, C. Wu, and Z. Zhang, *Linearized Galerkin FEMs for nonlinear time fractional parabolic problems with non-smooth solutions in time direction*, J. Sci. Comput., 80 (2019), pp. 403–419.
[16] L. Li and J.-G. Liu, *A generalized definition of Caputo derivatives and its application to fractional ODEs*, SIAM J. Math. Anal., 50 (2018), pp. 2867–2900, https://doi.org/10.1137/17M1160318.
[17] L. Li and J.-G. Liu, *Some compactness criteria for weak solutions of time fractional PDEs*, SIAM J. Math. Anal., 50 (2018), pp. 3963–3995, https://doi.org/10.1137/17M1145549.
[18] L. Li and J.-G. Liu, *A discretization of Caputo derivatives with application to time fractional SDEs and gradient flows*, SIAM J. Numer. Anal., 57 (2019), pp. 2095–2120, https://doi.org/10.1137/19M123854X.

[19] L. Li, J.-G. Liu, and J. Lu, *Fractional stochastic differential equations satisfying fluctuation-dissipation theorem*, J. Stat. Phys., 169 (2017), pp. 316–339.

[20] L. Li and D. Wang, *Complete monotonicity-preserving numerical methods for time fractional ODEs*, Commun. Math. Sci., 19 (2021), pp. 1301–1336.

[21] H.-L. Liao, D. Li, and J. Zhang, *Sharp error estimate of the nonuniform $L1$ formula for linear reaction-subdiffusion equations*, SIAM J. Numer. Anal., 56 (2018), pp. 1112–1133, https://doi.org/10.1137/17M1131829.

[22] H.-L. Liao, N. Liu, and P. Lyu, *Discrete gradient structure of a second-order variable-step method for nonlinear integro-differential models*, SIAM J. Numer. Anal., 61 (2023), pp. 2157–2181, https://doi.org/10.1137/22M1520050.

[23] H.-L. Liao, W. McLean, and J. Zhang, *A discrete Gronwall inequality with applications to numerical schemes for subdiffusion problems*, SIAM J. Numer. Anal., 57 (2019), pp. 218–237, https://doi.org/10.1137/16M1175742.

[24] H.-L. Liao, T. Tang, and T. Zhou, *Positive definiteness of real quadratic forms resulting from the variable-step $L1$-type approximation of convolution operators*, Sci. China Math., Sci. China Math., 67 (2024), pp. 237–252.

[25] Y. Lin and C. Xu, *Finite difference/spectral approximations for the time-fractional diffusion equation*, J. Comput. Phys., 225 (2007), pp. 1533–1552.

[26] P. Lyu and S. Vong, *A symmetric fractional-order reduction method for direct nonuniform approximations of semilinear diffusion-wave equations*, J. Sci. Comput., 93 (2022), 34.

[27] F. Mainardi, P. Paradisi, and R. Gorenflo, *Probability Distributions Generated by Fractional Diffusion Equations*, preprint, https://arxiv.org/abs/0704.0320, 2007.

[28] W. McLean, V. Thomée, and L. B. Wahlbin, *Discretization with variable time steps of an evolution equation with a positive-type memory term*, J. Comput. Appl. Math., 69 (1996), pp. 49–69.

[29] R. Miller, *On Volterra integral equations with nonnegative integrable resolvents*, J. Math. Anal. Appl., 22 (1968), pp. 319–340.

[30] J. D. Ramírez and A. S. Vatsala, *Generalized monotone iterative technique for Caputo fractional differential equation with periodic boundary condition via initial value problem*, Int. J. Differ. Equ., 2012 (2012), 842813.

[31] R. L. Schilling, R. Song, and Z. Vondracek, *Bernstein Functions: Theory and Applications*, De Gruyter Stud. Math. 37, Walter de Gruyter, Berlin, 2012.

[32] M. Stynes, E. O'Riordan, and J. L. Gracia, *Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation*, SIAM J. Numer. Anal., 55 (2017), pp. 1057–1079, https://doi.org/10.1137/16M1082329.

[33] T. Tang, H. Yu, and T. Zhou, *On energy dissipation theory and numerical stability for time-fractional phase-field equations*, SIAM J. Sci. Comput., 41 (2019), pp. A3757–A3778, https://doi.org/10.1137/18M1203560.

[34] D. Widder, *The Laplace Transform*, Princeton University Press, Princeton, NJ, 1941.

[35] G. M. Zaslavsky, *Chaos, fractional kinetics, and anomalous transport*, Phys. Rep., 371 (2002), pp. 461–580.