

ON THE CONSTRUCTION AND PROPERTIES OF M -STEP METHODS FOR FDES*

LIDIA ACETO[†], CECILIA MAGHERINI[‡], AND PAOLO NOVATI[§]

Abstract. In this paper we consider the numerical solution of Fractional Differential Equations by means of m -step recursions. The construction of such formulas can be obtained in many ways. Here we study a technique based on the rational approximation of the generating functions of Fractional Backward Differentiation Formulas (FBDFs). Accurate approximations lead to the definition of methods which simulate the underlying FBDF, with important computational advantages. Numerical experiments are presented.

Key words. fractional derivatives, matrix functions, contour integral approximation, Gauss-Jacobi rule

AMS subject classifications. 65L06,65F60,65D32,26A33,34A08

1. Introduction. This paper deals with the solution of Fractional Differential Equations (FDEs) of the type

$$(1.1) \quad {}_{t_0}D_t^\alpha y(t) = g(t, y(t)), \quad t_0 < t \leq T, \quad 0 < \alpha < 1,$$

where ${}_{t_0}D_t^\alpha$ denotes the Caputo's fractional derivative operator (see e.g. [21] for an overview) defined as

$$(1.2) \quad {}_{t_0}D_t^\alpha y(t) = \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^t \frac{y'(u)}{(t-u)^\alpha} du,$$

in which Γ denotes the Gamma function. As well known, the use of the Caputo's definition for the fractional derivative allows to treat the initial conditions at t_0 for FDEs in the same manner as for integer order differential equations. Setting $y(t_0) = y_0$ the solution of (1.1) exists and is unique under the hypothesis that g is continuous and fulfils a Lipschitz condition with respect to the second variable (see e.g. [4] for a proof).

As for the integer order case $\alpha = 1$, a classical approach for solving (1.1) is based on the discretization of the fractional derivative (1.2), which generalizes the well known Grünwald-Letnikov discretization (see [21, §2.2]), leading to the so-called Fractional Backward Differentiation Formulas (FBDFs) introduced in [16]. Taking a uniform mesh $t_0, t_1, \dots, t_N = T$ of the time domain with stepsize $h = (T - t_0)/N$, FBDFs are based on the full-term recursion

$$(1.3) \quad \sum_{j=0}^n \omega_{n-j}^{(p)} y_j = h^\alpha g(t_n, y_n), \quad p \leq n \leq N,$$

where $y_j \approx y(t_j)$ and $\omega_{n-j}^{(p)}$ are the Taylor coefficients of the generating function

$$(1.4) \quad \omega_p^{(\alpha)}(\zeta) = (a_0 + a_1 \zeta + \dots + a_p \zeta^p)^\alpha$$

$$(1.5) \quad = \sum_{i=0}^{\infty} \omega_i^{(p)} \zeta^i, \quad \text{for } 1 \leq p \leq 6,$$

*This work was supported by the GNCS-INdAM 2014 project "Metodi numerici per modelli di propagazione di onde elettromagnetiche in tessuti biologici".

[†]Dipartimento di Matematica, Università di Pisa, Italy (lidia.aceto@unipi.it)

[‡]Dipartimento di Matematica, Università di Pisa, Italy (cecilia.magherini@unipi.it)

[§]Dipartimento di Matematica, Università di Padova, Italy (novati@math.unipd.it)

being $\{a_0, a_1, \dots, a_p\}$ the coefficients of the underlying BDF. In [16] it is also shown that the order p of the BDF is preserved.

We remember that for this kind of equations, there is generally an intrinsic lack of regularity of the solution in a neighborhood of the starting point, that is, depending on the function g , one may have $y(t) \sim (t - t_0)^\alpha$ as $t \rightarrow t_0$. For this reason, in order to preserve the theoretical order p of the numerical method, formula (1.3) is generally corrected as

$$(1.6) \quad \sum_{j=0}^M w_{n,j} y_j + \sum_{j=0}^n \omega_{n-j}^{(p)} y_j = h^\alpha g(t_n, y_n),$$

where the sum $\sum_{j=0}^M w_{n,j} y_j$ is the so-called starting term, in which M and the weights $w_{n,j}$ depend on α and p (see [2, Chapter 6] for a discussion).

Denoting by Π_m the set of polynomials of degree not exceeding m , our basic idea is to design methods based on rational approximations of (1.4), i.e.,

$$(1.7) \quad R_m(\zeta) \approx \omega_p^{(\alpha)}(\zeta), \quad R_m(\zeta) = \frac{p_m(\zeta)}{q_m(\zeta)}, \quad p_m, q_m \in \Pi_m.$$

Writing $p_m(\zeta) = \sum_{j=0}^m \alpha_j \zeta^j$ and $q_m(\zeta) = \sum_{j=0}^m \beta_j \zeta^j$, the above approximation naturally leads to implicit m -step recursions of the type

$$(1.8) \quad \sum_{j=n-m}^n \alpha_{n-j} y_j = h^\alpha \sum_{j=n-m}^n \beta_{n-j} g(t_j, y_j), \quad n \geq \ell \geq m.$$

While the theoretical order of the FBDF is lost, we shall see that (1.8) is able to approximate the solution, with an accuracy closely related to the one of (1.6). Starting from the initial data $y(t_0) = y_0$, the first $\ell - 1$ approximations $y_1, \dots, y_{\ell-1}$ can be generated by the underlying FBDFs or even considering lower degree rational approximations (details are given in Section 4).

A formula of type (1.8) generalizes in some sense the methods based on the Short Memory Principle in which the truncated Taylor expansion of (1.4) is considered (see [21, §8.3] for some examples). Computationally, the advantages are noticeable, especially in terms of memory saving whenever (1.1) arises from the semi-discretization of fractional partial differential equations. We remark moreover that since the initial approximations are used only at the beginning of the process, there is no need to use a starting term to preserve the theoretical order as for standard full-recursion multistep formulas. Moreover, as remarked in [5], in particular when $\alpha \neq 1/2$, the use of a starting formula as in (1.6), that theoretically should ensure the order of the FBDF, in practice may introduce substantial errors, causing unreliable numerical solutions. For high-order formulas, this is due to the severe ill-conditioning of the Vandermonde type systems one has to solve at each integration step to generate the weights $w_{n,j}$ of the starting term. We also remark that in a typical application α , y_0 and possibly also the function g may be only known up to a certain accuracy (see [4] for a discussion), so that one may only be interested in having a rather good approximation of the true solution.

For the construction of formulas of type (1.8), in this paper we present a technique based on the rational approximation of the fractional derivative operator (cf. [19]). After considering a BDF discretization of order p of the first derivative operator, which can be represented by a $N \times N$ triangular banded Toeplitz matrix A_p , we approximate Caputo's fractional differential operator ${}_t D_t^\alpha$ by calculating A_p^α . This computation is performed by means of the contour integral approximation, which

leads to a rational approximation of A_p^α whose coefficients can be used to define (1.8). This technique is based on the fact that the first column of A_p^α contains the first N coefficients of the Taylor expansion of $\omega_p^{(\alpha)}(\zeta)$, so exploiting the equivalence between the approximation of A_p^α and $\omega_p^{(\alpha)}(\zeta)$. After suitable manipulations the contour integral is approximated by the Gauss-Jacobi rule that leads to the integration of analytic functions independently of α . This technique is completely new in this field. The paper also contains the error analysis of this approximation that yields indications about the proper definition of the parameters involved in this approach. Some of the theoretical properties of the arising formula (1.8), such as the zero-stability, the consistency, and the absolute (linear) stability are also considered in the paper.

For completeness we remember that in the more general framework of the evaluation of fractional integrals of the type

$$(1.9) \quad I^\alpha[f](t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t - \tau)^{\alpha-1} f(\tau) d\tau,$$

efficient algorithms that allow important memory saving have been recently proposed. We remark that the computation of (1.9) includes the solution of FDEs of type (1.1) since

$$y(t) = y_0 + I^\alpha[f](t)$$

for $f(t) = g(t, y(t))$ and $t_0 = 0$. These algorithms are all essentially based on the approximations of the convolution kernel $(t - \tau)^{\alpha-1}$ and therefore are different with respect to the one of the present work. Among the others we recall here the one presented in [14] where the convolution kernel is expressed in terms of a Laplace transform which is then approximated by suitable quadrature rules. Similarly to (1.8), the arising algorithms are able to work with a storage requirement of type $O(m)$ where m is the number of quadrature points. A similar approach was previously considered in [23], where a contour integral representation of the convolution kernel were considered. The arising ‘‘oblivious’’ algorithm uses $O(\log n)$ active memory to evaluate (1.9) at $t = t_n$. This approach was then extended to work with variable stepsize in [15]. Exploiting the decaying of the convolution kernel, in [6, 3] the authors proposed the use of nested meshes, that again allows the use of only $O(\log n)$ active memory. With the same memory requirements, a degenerate kernel approximation has been recently proposed in [18].

The outline of the paper is the following. As in [7], the contour integral is evaluated by means of the Gauss-Jacobi rule in Section 2. An error analysis of this approach is outlined in Section 3, together with some numerical experiments that show its accuracy. In Section 4 we investigate some basic features of this approach for the solution of fractional differential equations. In particular, we present some results concerning the consistency and the absolute stability. Finally, in Section 5 we consider the results of the method when applied to the discretization of two well-known models of fractional diffusion.

2. The approximation of the fractional derivative operator. Denoting by a_0, a_1, \dots, a_p the $p + 1$ coefficients of a Backward Differentiation Formula (BDF) of order p , with $1 \leq p \leq 6$, which discretizes the derivative operator (see [10, Chapter III.1] for a background), we consider lower triangular banded Toeplitz matrices of the

type

$$(2.1) \quad A_p = \begin{pmatrix} a_0 & 0 & & 0 \\ \vdots & a_0 & 0 & \\ a_p & \vdots & \ddots & 0 \\ 0 & \ddots & & \ddots & 0 \\ & 0 & a_p & \cdots & a_0 \end{pmatrix} \in \mathbb{R}^{(N+1) \times (N+1)}.$$

In this setting, $A_p^\alpha e_1$, $e_1 = (1, 0, \dots, 0)^T$, contains the whole set of coefficients of the corresponding FBDF for approximating the solution of (1.1) in t_0, t_1, \dots, t_N , that is

$$(2.2) \quad e_{j+1}^T A_p^\alpha e_1 = \omega_j^{(p)}, \quad 0 \leq j \leq N,$$

(cf. (1.5)). The constraint $p \leq 6$ is due to the fact that BDFs are not zero-stable for $p > 6$.

From the theory of matrix functions (see [12] for a survey), we know that the fractional power of matrix can be written as a contour integral

$$A^\alpha = \frac{A}{2\pi i} \int_{\Gamma} z^{\alpha-1} (zI - A)^{-1} dz,$$

where Γ is a suitable closed contour enclosing the spectrum of A , $\sigma(A)$, in its interior. The following known result (see, e.g., [1]) expresses A^α in terms of a real integral.

PROPOSITION 2.1. *Let $A \in \mathbb{R}^{N \times N}$ be such that $\sigma(A) \subset \mathbb{C} \setminus (-\infty, 0]$. For $0 < \alpha < 1$ the following representation holds*

$$(2.3) \quad A^\alpha = \frac{A \sin(\alpha\pi)}{\alpha\pi} \int_0^\infty (\rho^{1/\alpha} I + A)^{-1} d\rho.$$

Of course the above result holds also in our case since $\sigma(A_p) = \{a_0\}$ and $a_0 > 0$ for each $1 \leq p \leq 6$. At this point, for each suitable change of variable for ρ , a k -point quadrature rule for the computation of the integral in (2.3) yields a rational approximation of the type

$$(2.4) \quad A_p^\alpha \approx A_p \tilde{R}_k(A_p) := A_p \sum_{j=1}^k \gamma_j (\eta_j I + A_p)^{-1},$$

where the coefficients γ_j and η_j depend on the substitution and the quadrature formula. This technique has been used in [19], where the author applies the Gauss-Legendre rule to (2.3) after the substitution

$$\rho = a_0^\alpha (\cos \theta)^{-\alpha/(1-\alpha)} \sin \theta,$$

which generalizes the one presented in [11] for the case $\alpha = 1/2$.

In order to confine the dependence of α to a weight function, we consider the change of variable

$$(2.5) \quad \rho^{1/\alpha} = \tau \frac{1-t}{1+t}, \quad \tau > 0,$$

yielding

$$\begin{aligned} & \frac{1}{\alpha} \int_0^\infty (\rho^{1/\alpha} I + A_p)^{-1} d\rho \\ &= 2 \int_{-1}^1 \left(\tau \frac{1-t}{1+t} \right)^{\alpha-1} \left(\tau \frac{1-t}{1+t} I + A_p \right)^{-1} \frac{\tau}{(1+t)^2} dt \\ &= 2\tau^\alpha \int_{-1}^1 (1-t)^{\alpha-1} (1+t)^{-\alpha} (\tau(1-t)I + (1+t)A_p)^{-1} dt, \end{aligned}$$

and hence

$$(2.6) \quad A_p^\alpha = \frac{2 \sin(\alpha\pi) \tau^\alpha}{\pi} A_p \int_{-1}^1 (1-t)^{\alpha-1} (1+t)^{-\alpha} (\tau(1-t)I + (1+t)A_p)^{-1} dt.$$

The above formula naturally leads to the use of a k -point Gauss-Jacobi rule for the approximation of $A_p^\alpha e_1$ and hence to a rational approximation (2.4). In more details, if we denote with ϑ_j and w_j the j -th node and weight of the k -point Gauss-Jacobi quadrature rule with weight function $(1-t)^{\alpha-1}(1+t)^{-\alpha}$, then the resulting coefficients γ_j and η_j in (2.4) are given, respectively, by

$$(2.7) \quad \gamma_j = \frac{2 \sin(\alpha\pi) \tau^\alpha}{\pi} \frac{w_j}{1+\vartheta_j}, \quad \eta_j = \frac{\tau(1-\vartheta_j)}{1+\vartheta_j}.$$

The following result can be proved by direct computation.

PROPOSITION 2.2. *Let $A_p \in \mathbb{R}^{N \times N}$ be a matrix of type (2.1), and let $\bar{A}_p = \frac{1}{a_0} A_p$. Then the components of $(\xi I + \bar{A}_p)^{-1} e_1$, $\xi \neq -1$, are given by*

$$\begin{aligned} v_1^{(p)}(\xi) &= \frac{1}{\xi+1}, \\ v_j^{(p)}(\xi) &= \frac{c_{2,j}^{(p)}}{(\xi+1)^2} + \dots + \frac{c_{j,j}^{(p)}}{(\xi+1)^j}, \quad 2 \leq j \leq N, \end{aligned}$$

where the coefficients $c_{i,j}^{(p)}$ depend on the order p . For $p=1$ we simply have $\{a_0, a_1\} = \{1, -1\}$, and hence

$$v_j^{(1)}(\xi) = \frac{1}{(\xi+1)^j}, \quad 1 \leq j \leq N.$$

The above proposition shows that the components of

$$(\tau(1-t)I + (1+t)A_p)^{-1} e_1$$

are analytic functions in a suitable open set containing $[-1, 1]$ in its interior, since they are sum of functions of the type

$$(2.8) \quad \frac{(1+t)^{l-1}}{(\tau(1-t) + a_0(1+t))^l}, \quad l \geq 1,$$

whose pole lies outside $[-1, 1]$ for $\tau > 0$ (recall that $a_0 > 0$ for $1 \leq p \leq 6$). In this sense, the lack of regularity of the integrand in (2.6) due to the presence of α ,

Algorithm 1

-
- 1: Select p, k, τ and $\ell \geq m = kp$
 - 2: Use (2.7) to compute the coefficients γ_j 's and η_j 's
 - 3: Determine the coefficients of \tilde{p}_{k-1} and \tilde{q}_k in (2.9) by using an algorithm to transform partial fractions to polynomial quotient
 - 4: Compute the coefficients α_j 's and β_j 's by using (2.10)-(2.11)
 - 5: Use a suitable starting procedure to obtain an approximation of $y_1, \dots, y_{\ell-1}$
 - 6: Use the m -step recursion in (1.8) to advance the integration in time
-

is completely absorbed by the Jacobi weight function so that the Gauss-Jacobi rule yields a very efficient tool for the computation of A_p^α .

Increasing k the approximation (2.4) can be used to approximate the whole set of coefficients of the FBDFs. We remark that the computation of the vectors $(\eta_j I + A_p)^{-1} e_1$ does not constitute a problem because of the structure of A_p (see (2.1)). We also point out that since our aim is to construct reliable formulas of type (1.8) we actually do not need to evaluate (2.4). Indeed we just need to know the scalars γ_j and η_j , and then, using an algorithm to transform partial fractions to polynomial quotient, we obtain the approximation

$$(2.9) \quad z^\alpha \approx z \tilde{R}_k(z) = z \frac{\tilde{p}_{k-1}(z)}{\tilde{q}_k(z)}, \quad z = a_0 + a_1 \zeta + \dots + a_p \zeta^p,$$

where $\tilde{p}_{k-1} \in \Pi_{k-1}$ and $\tilde{q}_k \in \Pi_k$. This finally leads to the approximation (1.7) with

$$(2.10) \quad p_m(\zeta) = (a_0 + a_1 \zeta + \dots + a_p \zeta^p) \tilde{p}_{k-1}(a_0 + a_1 \zeta + \dots + a_p \zeta^p) \equiv \sum_{j=0}^m \alpha_j \zeta^j,$$

$$(2.11) \quad q_m(\zeta) = \tilde{q}_k(a_0 + a_1 \zeta + \dots + a_p \zeta^p) \equiv \sum_{j=0}^m \beta_j \zeta^j,$$

in which $m = kp$. We remark that whenever the procedure for the definition of the coefficients γ_j and η_j has been set for a given α , one can compute the corresponding coefficients in the m -step formula (1.8) once and for all. The main steps of the proposed procedure are summarized in Algorithm 1.

3. Theoretical error analysis. Denoting by $J_k(A_p)$ the result of the Gauss-Jacobi rule for the approximation of

$$J(A_p) = \int_{-1}^1 (1-t)^{\alpha-1} (1+t)^{-\alpha} (\tau(1-t)I + (1+t)A_p)^{-1} dt,$$

by (2.6) the corresponding approximation to A_p^α is given by

$$(3.1) \quad A_p^\alpha \approx A_p \tilde{R}_k(A_p), \quad \tilde{R}_k(A_p) = \frac{2 \sin(\alpha\pi) \tau^\alpha}{\pi} J_k(A_p).$$

In this section we analyze the error term componentwise, that is, (see (2.2)),

$$(3.2) \quad \begin{aligned} E_j &:= \omega_j^{(p)} - e_{j+1}^T A_p \tilde{R}_k(A_p) e_1 = e_{j+1}^T A_p \left(A_p^{\alpha-1} - \tilde{R}_k(A_p) \right) e_1 \\ &= \frac{2 \sin(\alpha\pi) \tau^\alpha}{\pi} e_{j+1}^T A_p (J(A_p) - J_k(A_p)) e_1, \quad 0 \leq j \leq N, \end{aligned}$$

which is the error in the computation of the j -th coefficient of the Taylor expansion of $\omega_p^{(\alpha)}(\zeta)$. Numerically one observes that the quality of the approximation tends to deteriorate when the dimension of the problem N grows. In this sense we are particularly interested in observing the dependence of the error term on j and k for $j \gg k$ and to find a strategy to define the parameter τ of the substitution (2.5) in this situation. As we shall see in the remainder of the paper, this parameter plays a crucial role for the quality of the approximation.

We restrict our analysis to the case of $p = 1$ for which $a_0 = -a_1 = 1$. In this situation, defining the vector

$$(3.3) \quad r := (J(A_1) - J_k(A_1)) e_1,$$

we have that

$$E_j = \frac{2 \sin(\alpha\pi) \tau^\alpha}{\pi} e_{j+1}^T A_1 r,$$

and therefore

$$(3.4) \quad |E_j| \leq \frac{2 \sin(\alpha\pi) \tau^\alpha}{\pi} (|r_j| + |r_{j-1}|).$$

The analysis thus reduces to the study of the components of the vector (3.3). By Proposition 2.2, see also (2.8), the j -th component of the vector

$$(\tau(1-t)I + (1+t)A_1)^{-1} e_1$$

is given by

$$(3.5) \quad f_j(t) = \frac{(1+t)^{j-1}}{(\tau(1-t) + 1+t)^j},$$

so that r_j is the error term of the k -point Gauss-Jacobi formula applied to the computation of

$$(3.6) \quad \int_{-1}^1 (1-t)^{\alpha-1} (1+t)^{-\alpha} f_j(t) dt.$$

We start with the following known result, [13].

THEOREM 3.1. *The error term of the k -point Gauss-Jacobi formula applied to the computation of*

$$\int_{-1}^1 (1-t)^{\alpha_1} (1+t)^{\alpha_2} g(t) dt, \quad \alpha_1, \alpha_2 > -1, \quad g \in C^{2k}([-1, 1]),$$

is given by

$$C_{k, \alpha_1, \alpha_2} g^{(2k)}(\xi), \quad -1 < \xi < 1,$$

where

$$(3.7) \quad C_{k, \alpha_1, \alpha_2} := \frac{\Gamma(k + \alpha_1 + 1) \Gamma(k + \alpha_2 + 1) \Gamma(k + \alpha_1 + \alpha_2 + 1) k!}{(2k + \alpha_1 + \alpha_2 + 1) [\Gamma(2k + \alpha_1 + \alpha_2 + 1)]^2 (2k)!} 2^{2k + \alpha_1 + \alpha_2 + 1}.$$

LEMMA 3.2. *Let $0 < \alpha < 1$ and $C_{k,\alpha} := C_{k,\alpha-1,-\alpha}$. Then*

$$C_{k,\alpha} \sim \frac{\pi 2^{1-2k}}{(2k)!}.$$

Proof. By (3.7) we easily obtain

$$C_{k,\alpha} = \frac{\Gamma(k+\alpha)\Gamma(k-\alpha+1)[\Gamma(k)]^2}{[\Gamma(2k)]^2(2k)!} 2^{2k-1}.$$

Using the Legendre formula

$$\Gamma\left(k + \frac{1}{2}\right)\Gamma(k) = \sqrt{\pi}\Gamma(2k)2^{1-2k},$$

we have that

$$C_{k,1/2} = \frac{\pi 2^{1-2k}}{(2k)!}.$$

Moreover, since for $a, b \in (0, 1)$

$$k^{b-a} \frac{\Gamma(k+a)}{\Gamma(k+b)} = 1 + O\left(\frac{1}{k}\right),$$

we have that $\Gamma(k+\alpha)\Gamma(k-\alpha+1) = [\Gamma(k+\frac{1}{2})]^2(1+O(k^{-1}))$ and, consequently, $C_{k,\alpha} \rightarrow C_{k,1/2}$ as $k \rightarrow \infty$. \square

REMARK 3.3. *If we set $\tau = 1$ in (2.5) we obtain $r_j = 0$ for each $j = 1, 2, \dots, 2k$ since $f_j \in \Pi_{j-1}$, see (3.5). From (2.9), with $p = 1$, and (3.1) one therefore gets*

$$(1-\zeta)^{\alpha-1} - \tilde{R}_k(1-\zeta) = (1-\zeta)^{\alpha-1} - \frac{\tilde{p}_{k-1}(1-\zeta)}{\tilde{q}_k(1-\zeta)} = O(\zeta^{2k}),$$

so that $\tilde{R}_k(1-\zeta)$ is the $(k-1, k)$ Padé approximant of $(1-\zeta)^{\alpha-1}$ with expansion point $\zeta = 0$. More generally, if $\tau \in (0, 1]$ then the resulting rational approximation coincides with the same Padé approximant with expansion point $\zeta = 1 - \tau$ (cf. [7, Lemma 4.4]).

Numerically, it is quite clear that the best results are obtained for τ strictly less than 1, so that in what follows we always assume to work with $\tau \in (0, 1)$. Indeed, in this situation we are able to approximate the Taylor coefficients with a more uniform distribution of the error with respect to j . By Theorem 3.1, we need to bound $\left|f_j^{(2k)}(t)\right|$ in the interval $[-1, 1]$ in order to bound the error term E_j . We start with the following result.

PROPOSITION 3.4. *Let $0 < \tau < 1$ and*

$$(3.8) \quad a := \frac{1+\tau}{1-\tau}.$$

For each j and k

$$\max_{[-1,1]} \left|f_j^{(2k)}(t)\right| \leq (2k)! \frac{\sqrt{a}}{(\sqrt{a}-1)^{2k+2}} \left(\frac{a+1}{2\sqrt{a}}\right)^j.$$

Proof. The function f_j can be written as

$$f_j(t) = \frac{(1+t)^{j-1}}{(a+t)^j} \left(\frac{a+1}{2} \right)^j, \quad a > 1.$$

Using the Cauchy integral formula we have

$$(3.9) \quad f_j^{(2k)}(t) = \frac{(2k)!}{2\pi i} \int_{\Gamma} \frac{f_j(w)}{(w-t)^{2k+1}} dw,$$

where Γ is a contour surrounding t but not the pole $-a < -1$. We take Γ as the circle centered at the origin and radius ρ such that $1 < \rho < a$, that is, we use the substitution $w = \rho e^{i\theta}$. We obtain

$$(3.10) \quad \max_{[-1,1]} |f_j^{(2k)}(t)| \leq (2k)! \frac{\rho}{(\rho-1)^{2k+1}} \max_{[0,2\pi]} |f_j(\rho e^{i\theta})|.$$

Taking $\rho = \sqrt{a}$ we have

$$\begin{aligned} |f_j(\rho e^{i\theta})| &= \left| \frac{1}{1+\rho e^{i\theta}} \right| \left| \frac{1+\rho e^{i\theta}}{a+\rho e^{i\theta}} \right|^j \left(\frac{a+1}{2} \right)^j \\ &= \left| \frac{1}{1+\sqrt{a}e^{i\theta}} \right| \left| \frac{1+\sqrt{a}e^{i\theta}}{a+\sqrt{a}e^{i\theta}} \right|^j \left(\frac{a+1}{2} \right)^j \\ &= \left| \frac{1}{1+\sqrt{a}e^{i\theta}} \right| \left(\frac{a+1}{2\sqrt{a}} \right)^j \\ &\leq \frac{1}{(\sqrt{a}-1)} \left(\frac{a+1}{2\sqrt{a}} \right)^j. \end{aligned}$$

By (3.10) we immediately achieve the result. \square

The above result is rather accurate only for small values of j . Since $f_j(t)$ is growing in the interval $[-1, 1]$, below we consider contours Γ in (3.9) which are dependent on t , in order to balance this effect.

PROPOSITION 3.5. *Let $0 < \tau < 1$ and a as in (3.8). For $j \geq 2k + 2$*

$$(3.11) \quad |f_j^{(2k)}(t)| \leq (2k)! \frac{\sqrt{a+1} + \sqrt{2}}{a-1} \frac{(1+t)^{\frac{j-1}{2}-k}}{(a+t)^{\frac{j}{2}+k}} \left(\frac{a+1}{2} \right)^j, \quad t \in [-1, 1].$$

Moreover

$$(3.12) \quad \max_{[-1,1]} |f_j^{(2k)}(t)| \leq (2k)! \left(\sqrt{a+1} + \sqrt{2} \right) \frac{\left(\frac{j-2k-1}{4k+1} \right)^{\frac{j-1}{2}-k}}{\left(\frac{j+2k}{4k+1} \right)^{\frac{j}{2}+k}} \left(\frac{a+1}{2} \right)^j (a-1)^{-2k-\frac{3}{2}}$$

for j such that

$$(3.13) \quad a \leq \frac{j+6k+1}{j-2k-1},$$

and

$$(3.14) \quad \max_{[-1,1]} |f_j^{(2k)}(t)| \leq (2k)! \frac{\sqrt{a+1} + \sqrt{2}}{a-1} \frac{(a+1)^{\frac{j}{2}-k}}{2^{\frac{j+1}{2}+k}}$$

otherwise.

Proof. In (3.9) we take Γ as the circle centered at t and of radius ρ such that $1+t < \rho < t+a$, that is, we use the substitution $w = t + \rho e^{i\theta}$. We obtain

$$(3.15) \quad |f_j^{(2k)}(t)| \leq (2k)! \frac{1}{\rho^{2k}} \max_{[0,2\pi]} |f_j(t + \rho e^{i\theta})|.$$

For $t > -1$ we define $\rho = \sqrt{(t+a)(1+t)}$, so that

$$\begin{aligned} |f_j(t + \rho e^{i\theta})| &= \left| \frac{1}{1+t+\rho e^{i\theta}} \right| \left| \frac{1+t+\rho e^{i\theta}}{a+t+\rho e^{i\theta}} \right|^j \left(\frac{a+1}{2} \right)^j \\ &= \left| \frac{1}{1+t+\sqrt{(t+a)(1+t)}e^{i\theta}} \right| \times \\ &\quad \left| \frac{1+t+\sqrt{(t+a)(1+t)}e^{i\theta}}{a+t+\sqrt{(t+a)(1+t)}e^{i\theta}} \right|^j \left(\frac{a+1}{2} \right)^j \\ &= \left| \frac{1}{1+t+\sqrt{(t+a)(1+t)}e^{i\theta}} \right| \left(\frac{1+t}{a+t} \right)^{\frac{j}{2}} \left(\frac{a+1}{2} \right)^j, \end{aligned}$$

and hence

$$\begin{aligned} \max_{[0,2\pi]} |f_j(t + \rho e^{i\theta})| &= \frac{1}{\sqrt{(t+a)(1+t)} - (1+t)} \left(\frac{1+t}{a+t} \right)^{\frac{j}{2}} \left(\frac{a+1}{2} \right)^j \\ &= \frac{\sqrt{(t+a)} + \sqrt{(1+t)}}{\sqrt{(1+t)(a-1)}} \left(\frac{1+t}{a+t} \right)^{\frac{j}{2}} \left(\frac{a+1}{2} \right)^j. \end{aligned}$$

By (3.15) we obtain the bound (3.11) for each j and k , for $t > -1$. By continuity, (3.11) holds for $t \in [-1, 1]$ if $j \geq 2k+2$.

Now, we observe that the maximum with respect to t of the function

$$\frac{(1+t)^{\frac{j-1}{2}-k}}{(a+t)^{\frac{j}{2}+k}}$$

is attained at

$$t^* = \frac{a(j-2k-1) - (j+2k)}{4k+1} \geq -1.$$

Moreover $t^* \leq 1$ for a verifying (3.13). Substituting t^* in (3.11) leads to (3.12). If $t^* > 1$ then the maximum of (3.11) is reached at $t = 1$ and hence we obtain (3.14). \square

In order to derive (3.12) and (3.14) we have assumed τ to be a priori fixed. Now, using these bounds, we look for the value of τ which minimize $\|f_j^{(2k)}\|$ for a given j . For $j \geq 2k+3$, the minimization of the term

$$(a+1)^j (a-1)^{-2k-\frac{3}{2}}$$

in (3.12) leads to

$$a^{(1)} = \frac{2j + 4k + 3}{2j - 4k - 3},$$

which satisfies (3.13). Consequently, by (3.8) we obtain

$$(3.16) \quad \tau^{(1)} = \frac{4k + 3}{2j}.$$

On the other side, for the same $j \geq 2k + 3$, the minimization of the term

$$\frac{\sqrt{a+1} + \sqrt{2}}{a-1} (a+1)^{\frac{j}{2}-k}$$

in (3.14) leads to a value

$$a^* \leq \frac{j - 2k + 2}{j - 2k - 2},$$

which also satisfies (3.13) and hence does not fulfill the requirement of (3.14). For $a > a^*$ the bound (3.14) is growing with a and consequently its minimum is attained just for

$$a^{(2)} = \frac{j + 6k + 1}{j - 2k - 1}.$$

Using this value in (3.14) leads to a bound that is coarser than the one obtained by replacing $a^{(1)}$ in (3.12). For this reason, with respect to our estimates, τ given by (3.16) represents the optimal value for the computation of (3.6) and consequently of $\omega_j^{(p)}$.

The following theorem summarizes the results obtained. The proof follows straightfully from (3.4), Theorem 3.1, Lemma 3.2, and Propositions 3.4, 3.5.

THEOREM 3.6. *Let $0 < \tau < 1$ and*

$$a := \frac{1 + \tau}{1 - \tau}.$$

Then

$$|E_j| \leq 2^{3-2k} \sin(\alpha\pi) \tau^\alpha \Psi(a, j, k).$$

where

$$\Psi(a, j, k) := \begin{cases} \frac{\sqrt{a}}{(\sqrt{a}-1)^{2k+2}} \left(\frac{a+1}{2\sqrt{a}}\right)^j, & j \leq 2k+1, \\ (\sqrt{a+1} + \sqrt{2}) \frac{\left(\frac{j-2k-1}{4k+1}\right)^{\frac{j-1}{2}-k}}{\left(\frac{j+2k}{4k+1}\right)^{\frac{j}{2}+k}} \left(\frac{a+1}{2}\right)^j (a-1)^{-2k-\frac{3}{2}}, & 2k+2 \leq j \leq \frac{6k+1+a(2k+1)}{a-1}, \\ \frac{\sqrt{a+1} + \sqrt{2}}{a-1} \frac{(a+1)^{\frac{j}{2}-k}}{2^{\frac{j+1}{2}+k}}, & j \geq \max\left(2k+2, \frac{6k+1+a(2k+1)}{a-1}\right). \end{cases}$$

For $\tau = \tau^{(1)}$ the corresponding expression of $\Psi(a, j, k)$ is minimized for $j \geq 2k + 3$.

3.1. Numerical experiments. As already mentioned, the aim of the whole analysis was to have indications about the choice of the parameter τ with respect to the degree k of the formula and the dimension of the problem N . Unfortunately, the definition of τ as in (3.16) depends on j , while we need a value which is as good as possible for each $1 \leq j \leq N$. In this sense, the idea, confirmed by the forthcoming experiments, is to use $\tau^{(1)}$ with $j = N/2$, that is, focusing the attention on the middle of the interval $[0, N]$. This leads to a choice of τ around the value $4k/N$. We remark that the previous analysis was restricted to the case of $p = 1$, because of the difficulties in dealing with the functions f_j for $p > 1$ (cf. Proposition 2.2).

Numerically, we can proceed as follows. If we define

$$\tau^* = \arg \min_{\tau} \mathcal{E}(\tau), \quad \mathcal{E}(\tau) := \left\| A_p^{\alpha-1} - \tilde{R}_k(A_p) \right\|_{\infty},$$

then, in principle, $\tau^* = \tau^*(\alpha, k, N, p)$. However, the numerical experiments done by using the `Matlab` optimization routine `fminsearch` indicate that the dependence on α is negligible with respect to the others. In particular, there is numerical evidence that $\mathcal{E}(\tau^*) \approx \mathcal{E}(\hat{\tau})$ where

$$(3.17) \quad \hat{\tau} = \frac{(7+p)}{2N}k.$$

In Figures 1-2 we report the values of $\mathcal{E}(\tau^*)$ and $\mathcal{E}(\hat{\tau})$ for $p = 1, 3$, respectively. We recall that the corresponding sets of coefficients $\{a_0, a_1, \dots, a_p\}$ in (2.1) are given by

$$\begin{aligned} p = 1 : & \quad \{1, -1\}, \\ p = 3 : & \quad \{11/6, -3, 3/2, -1/3\}. \end{aligned}$$

As one can see, all the curves are approximatively overlapped. A ‘‘quasi’’ optimal approximation of $A_p^{\alpha-1}$ can be therefore obtained by using the very simple formula in (3.17) for choosing τ . Moreover, it is important to remark that such approximations are surely satisfactory even with $k \ll N$.

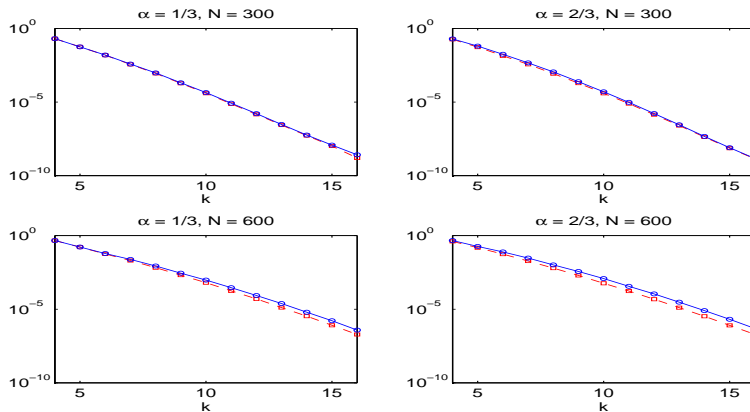


FIG. 1. Error behavior of the Gauss-Jacobi rule for the approximation of $A_1^{\alpha-1}$ for $\tau = \tau^*$ (dashed line) and $\tau = \hat{\tau} = 4k/N$ (solid line).

The previous results are all related to the overall error in the approximation of A_p^α . Considering that our final goal is the use of such approximation for the solution

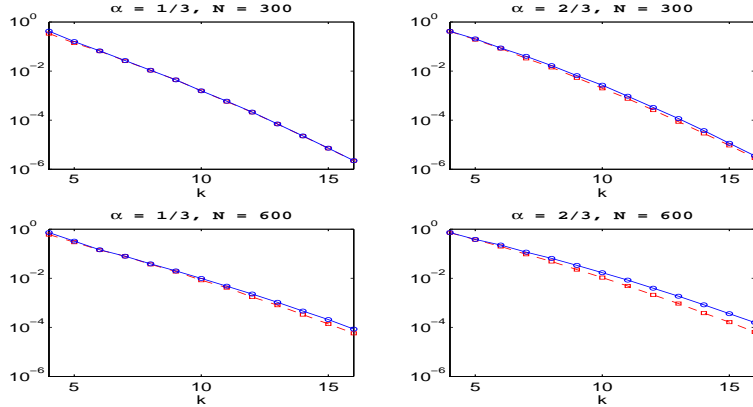


FIG. 2. Error behavior of the Gauss-Jacobi rule for the approximation of $A_3^{\alpha-1}$ for $\tau = \tau^*$ (dashed line) and $\tau = \hat{\tau} = 5k/N$ (solid line).

of FDEs, it is important to inspect also the componentwise error. As an example, in Figure 3, we report such errors, i.e., the values of E_j defined in (3.2), in the case of $N = 400$, $p = 1$, $\tau = \hat{\tau}$ for $\alpha = 0.3, 0.5, 0.7$, and different values of k . We also consider the componentwise errors of the polynomial approximation of the generating function obtained by truncating its Taylor series, with memory length equal to 16. Obviously this is equivalent to approximate with 0 the coefficients $\omega_i^{(p)}$ of (1.5), for $i > 16$, so that the error is just $|\omega_i^{(p)}|$. The competitiveness of the rational approach is undeniable.

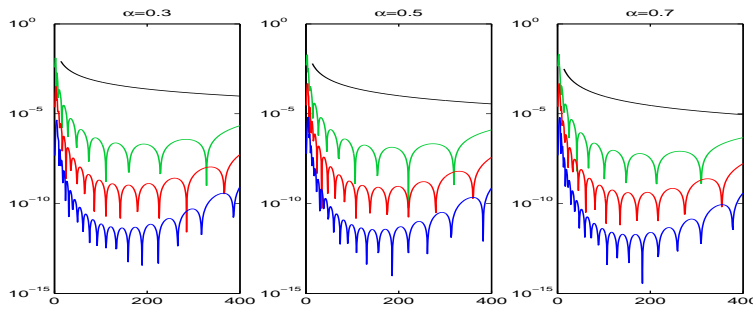


FIG. 3. Componentwise error of the Gauss-Jacobi rational approximation with memory length $k = 6, 9, 12$, and the polynomial approximation with $k = 12$.

4. The solution of FDEs. In this section we discuss the use of the described approximation of A_1^α for getting a k -step method that simulates the FBDF of order 1. The discrete problem provided by the latter method applied for solving (1.1) can be written in matrix form as follows

$$(4.1) \quad (A_1^\alpha \otimes I_s)(Y - \mathbf{1} \otimes y_0) = h^\alpha G(Y),$$

where s is the dimension of the FDE, I_s is the identity matrix of order s , $y_0 \in \mathbb{R}^s$ represents the initial value, $h = (T - t_0)/N$ is the stepsize, $\mathbf{1} = (1, 1, \dots, 1)^T \in \mathbb{R}^N$,

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} \approx \begin{pmatrix} y(t_1) \\ y(t_2) \\ \vdots \\ y(t_N) \end{pmatrix}, \quad G(Y) = \begin{pmatrix} g(t_1, y_1) \\ g(t_2, y_2) \\ \vdots \\ g(t_N, y_N) \end{pmatrix} \equiv \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_N \end{pmatrix}.$$

As described in Section 2, the use of a k -point Gauss-Jacobi rule for approximating (2.6) leads to

$$A_1^\alpha \approx \begin{pmatrix} \beta_0 & 0 & & 0 \\ \vdots & \beta_0 & & 0 \\ \beta_k & \vdots & \ddots & 0 \\ 0 & \ddots & & \ddots & 0 \\ 0 & 0 & \beta_k & \cdots & \beta_0 \end{pmatrix}^{-1} \begin{pmatrix} \alpha_0 & 0 & & 0 \\ \vdots & \alpha_0 & & 0 \\ \alpha_k & \vdots & \ddots & 0 \\ 0 & \ddots & & \ddots & 0 \\ 0 & \alpha_k & \cdots & \alpha_0 \end{pmatrix} \equiv B^{-1}C.$$

Here, the coefficients $\{\alpha_j\}_{j=0}^k$ and $\{\beta_j\}_{j=0}^k$ are related to the rational approximation through the formulas (2.9)–(2.11), with $m = k$ since $p = 1$,

$$(4.2) \quad p_k(\zeta) = (1 - \zeta)\tilde{p}_{k-1}(1 - \zeta) = \sum_{j=0}^k \alpha_j \zeta^j, \quad q_k(\zeta) = \tilde{q}_k(1 - \zeta) = \sum_{j=0}^k \beta_j \zeta^j.$$

If we replace A_1^α by $B^{-1}C$ in (4.1) and we multiply both side of the resulting equation from the left by $B \otimes I_s$, we obtain

$$(4.3) \quad (C \otimes I_s)Y - C\mathbf{1} \otimes y_0 = h^\alpha (B \otimes I_s)G(Y),$$

where now Y represents the numerical solution provided by the k -step method. In fact, considering that $(C\mathbf{1})_n = 0$ for each $n = k + 1, \dots, N$, since

$$(4.4) \quad p_k(1) = \sum_{j=0}^k \alpha_j = 0,$$

the discrete problem (4.3) simplifies to

$$(4.5) \quad \sum_{j=0}^{n-1} \alpha_j (y_{n-j} - y_0) = h^\alpha \sum_{j=0}^{n-1} \beta_j g_{n-j}, \quad n = 1, \dots, k,$$

$$(4.6) \quad \sum_{j=0}^k \alpha_j y_{n-j} = h^\alpha \sum_{j=0}^k \beta_j g_{n-j}, \quad n = k + 1, \dots, N.$$

Indeed, the equations in (4.5) allow to get an approximation of the solution over the first k meshpoints which are then used as starting values for the k -step recursion in (4.6). With reference to Algorithm 1, we have therefore decided to set $\ell = k + 1 = m + 1$ and the equations in (4.5) represent the implementation of step 5 we have used.

REMARK 4.1. From (4.4)–(4.6) follows that the method reproduces exactly constant solutions, i.e. it is exact if $g(t, y(t)) \equiv 0$.

As it happens in the case of ODEs, a localization of the zeros of the characteristic polynomials of the k -step method in (4.2) is required in order to study its stability properties. Clearly, such polynomials depend on the parameter τ , i.e. $p_k(\zeta) \equiv p_k(\zeta; \tau)$ and $q_k(\zeta) \equiv q_k(\zeta; \tau)$ since this dependence occurs in $\tilde{p}_{k-1}, \tilde{q}_k$. The method is therefore based on the following rational approximation

$$(4.7) \quad (1 - \zeta)^{\alpha-1} \approx \frac{\tilde{p}_{k-1}(1 - \zeta; \tau)}{\tilde{q}_k(1 - \zeta; \tau)} \equiv \tilde{R}_k(1 - \zeta; \tau).$$

THEOREM 4.2. *For each $\tau \in (0, 1]$, the adjoint of the characteristic polynomials of the k -step method, i.e. $\zeta^k p_k(\zeta^{-1}; \tau)$ and $\zeta^k q_k(\zeta^{-1}; \tau)$, are a Von Neumann and a Schur polynomial, respectively.*

Proof. From Remark 3.3, one obtains

$$(4.8) \quad \tilde{R}_k(1 - \zeta; \tau) = \tau^{\alpha-1} \tilde{R}_k \left(\frac{1 - \zeta}{\tau}; 1 \right),$$

since

$$\frac{d^l}{d\zeta^l} (1 - \zeta)^{\alpha-1} \Big|_{\zeta=1-\tau} = \tau^{\alpha-1} \frac{d^l}{d\zeta^l} \tilde{R}_k \left(\frac{1 - \zeta}{\tau}; 1 \right) \Big|_{\zeta=1-\tau}, \quad l = 0, 1, \dots, 2k - 1.$$

In addition, using the Gauss hypergeometric functions, in [9, Theorem 4.1] it has been proved that

$$\tilde{R}_k \left(\frac{1 - \zeta}{\tau}; 1 \right) = \frac{{}_2F_1(1 - k, 1 - \alpha - k; 1 - 2k; (\tau - 1 + \zeta)/\tau)}{{}_2F_1(-k, \alpha - k; 1 - 2k; (\tau - 1 + \zeta)/\tau)},$$

or equivalently, by denoting with $\mathcal{P}_r^{(\gamma, \beta)}$ the Jacobi polynomial of degree r and by using [20, eq. 142, p. 464] and the symmetry of such polynomials,

$$(4.9) \quad \tilde{R}_k \left(\frac{1 - \zeta}{\tau}; 1 \right) = \tau \frac{(\tau - 1 + \zeta)^{k-1} \mathcal{P}_{k-1}^{(1-\alpha, \alpha)}(2\tau/(\tau - 1 + \zeta) - 1)}{(\tau - 1 + \zeta)^k \mathcal{P}_k^{(\alpha-1, -\alpha)}(2\tau/(\tau - 1 + \zeta) - 1)}.$$

From (4.2) and (4.7)–(4.9), one therefore gets

$$\begin{aligned} p_k(\zeta; \tau) &= (1 - \zeta) \tau^\alpha (\tau - 1 + \zeta)^{k-1} \mathcal{P}_{k-1}^{(1-\alpha, \alpha)}(2\tau/(\tau - 1 + \zeta) - 1), \\ q_k(\zeta; \tau) &= (\tau - 1 + \zeta)^k \mathcal{P}_k^{(\alpha-1, -\alpha)}(2\tau/(\tau - 1 + \zeta) - 1). \end{aligned}$$

It follows that, if we denote with θ_i the i th root of $\mathcal{P}_{k-1}^{(1-\alpha, \alpha)}$ then the roots of $p_k(\zeta; \tau)$ are given by

$$(4.10) \quad \zeta_i = 1 + \tau \frac{1 - \theta_i}{1 + \theta_i} > 1, \quad i = 1, \dots, k - 1, \quad \zeta_k = 1,$$

where the inequality follows from the fact that the roots of the Jacobi polynomials belong to $(-1, 1)$. Similarly, by denoting with ϑ_i the i th root of $\mathcal{P}_k^{(\alpha-1, -\alpha)}$, one deduces that the roots of $q_k(\zeta; \tau)$ read

$$(4.11) \quad \zeta_i = 1 + \tau \frac{1 - \vartheta_i}{1 + \vartheta_i} > 1, \quad i = 1, \dots, k.$$

From (4.10)–(4.11) the statement follows immediately. \square

A first important consequence of the previous result is that the finite recurrence scheme is always 0-stable independently of the stepnumber k and $\tau \in (0, 1]$. More precisely, in the case $g \equiv 0$ the zero solution of (4.6) is stable with respect to perturbations of the initial values. A second very important consequence of Theorem 4.2 will be discussed in Section 4.2.

4.1. Consistency. In this section we examine the consistency of the method. While, theoretically, it is only exact for constant solutions (see Remark 4.1), numerically one observes that the consistency is rather well simulated if k is large enough. The analysis will also provide some hints about the choice of the memory length m . We restrict our consideration to the case $p = 1$ ($m = k$) but the generalization is immediate.

For a given $y(t)$, the FBDF of order 1 yields the approximation

$${}_0D_t^\alpha y(t) = \frac{1}{h^\alpha} \sum_{j=0}^N (-1)^j \binom{\alpha}{j} (y(t-jh) - y(0)) + O(h), \quad t = Nh.$$

Let

$$\Delta_h^\alpha y(t) := \sum_{j=0}^N (-1)^j \binom{\alpha}{j} (y(t-jh) - y(0)).$$

Writing a rational approximation of degree k to $\omega_1^{(\alpha)}(\zeta) = (1-\zeta)^\alpha$ as

$$R_k(\zeta) = \sum_{j=0}^{\infty} \gamma_j \zeta^j,$$

the corresponding method produces an approximation of the type

$${}_0D_t^\alpha y(t) \approx \frac{1}{h^\alpha} \sum_{j=0}^N \gamma_j (y(t-jh) - y(0)).$$

Denoting by

$$R_{k,h}^\alpha y(t) := \sum_{j=0}^N \gamma_j (y(t-jh) - y(0)),$$

we obtain

$$\begin{aligned} & {}_0D_t^\alpha y(t) - \frac{1}{h^\alpha} R_{k,h}^\alpha y(t) = \\ &= {}_0D_t^\alpha y(t) - \frac{1}{h^\alpha} \Delta_h^\alpha y(t) + \frac{1}{h^\alpha} \Delta_h^\alpha y(t) - \frac{1}{h^\alpha} R_{k,h}^\alpha y(t) \\ &= O(h) + \frac{1}{h^\alpha} \sum_{j=0}^N \left[(-1)^j \binom{\alpha}{j} - \gamma_j \right] (y(t-jh) - y(0)). \end{aligned}$$

The consistency of the method is ensured if

$$(4.12) \quad \frac{1}{h^\alpha} \sum_{j=0}^N \left[(-1)^j \binom{\alpha}{j} - \gamma_j \right] (y(t-jh) - y(0)) \rightarrow 0$$

as $h \rightarrow 0$ (cf. [8]). While this cannot be true for a fixed $k < \infty$, in what follows we show that numerically, i.e., for $h \geq h_0 > 0$, the consistency is well simulated if k is large enough and if the rational approximation to A_p^α is reliable.

As pointed out in [16], a certain method for FDEs with generating function $\omega^{(\alpha)}(\zeta)$ is consistent of order p if

$$h^{-\alpha} \omega^{(\alpha)}(e^{-h}) = 1 + O(h^p).$$

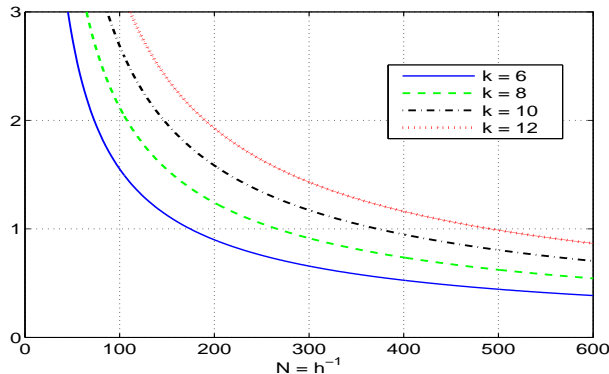


FIG. 4. Plot of the function $q_k(h)$ in (4.13) for $\alpha = 1/2$, $\tau = 1/10$ and different values of k .

In this setting, in order to understand the numerical consistence of our method, we consider the above relation by replacing $\omega^{(\alpha)}(e^{-h})$ with $\omega_1^{(\alpha)}(e^{-h})$ and $R_k(e^{-h})$. In particular, if we set

$$(4.13) \quad q_k(h) = \log_h \left(h^{-\alpha} \left| R_k(e^{-h}) - \omega_1^{(\alpha)}(e^{-h}) \right| \right),$$

then we obtain

$$\begin{aligned} h^{-\alpha} |R_k(e^{-h})| &\leq h^{-\alpha} \left| \omega_1^{(\alpha)}(e^{-h}) \right| + h^{-\alpha} \left| R_k(e^{-h}) - \omega_1^{(\alpha)}(e^{-h}) \right|. \\ &= 1 + O(h) + h^{q_k(h)} \end{aligned}$$

This implies that the consistency of the FBDF of the first order is well simulated as long as $q_k(h)$ is larger than 1. In Figure 4, we plot such function for $\alpha = 1/2$, $\tau = 1/10$, and different values of k .

The previous experiment does not take care of the perturbation introduced in the approximation of the fractional derivative of fractional powers of the independent variable which may be present in the solution of the FDE. In order to control such perturbations, we therefore consider the following second experiment. Going back to formula (4.12), we let $N = 1/h$ and $y(t) = E_\alpha(-t^\alpha)$ where $E_\alpha(x)$ denotes the one-parameter Mittag-Leffler function (see e.g. [21, Chapter 1])

$$(4.14) \quad E_\alpha(x) = \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(k\alpha + 1)}.$$

In Figure 5, we then consider the behavior of the function

$$(4.15) \quad \tilde{q}_k(h) = \log_h \left(h^{-\alpha} \left| \sum_{j=0}^N \left((-1)^j \binom{\alpha}{j} - \gamma_{j,k} \right) (y(t_{N-j}) - y(0)) \right| \right)$$

which, similarly to $q_k(h)$, has to be compared with 1. The values of $y(t)$ have been computed using the `Matlab` function `m1f` from [22] that implements the Mittag-Leffler function together with the Schur-Parlett algorithm.

We conclude this section by considering what happens with the general assumption

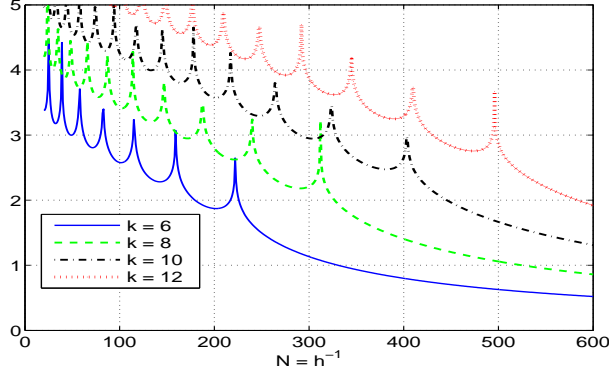


FIG. 5. Plot of the function $\bar{q}_k(h)$ in (4.15) for $\alpha = 1/2$, $\tau = 1/10$ and different values of k .

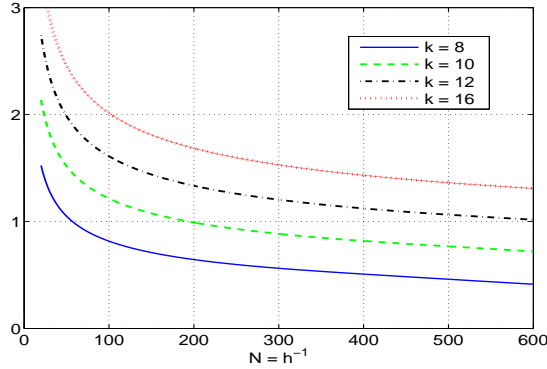


FIG. 6. Plot of the function $\bar{q}_k(h)$ in (4.16), for $\alpha = 1/2$, $\tau = 1/10$ and different values of k .

$|y(t)| \leq M$. Using this bound, by (4.12) we consider the function

$$(4.16) \quad \bar{q}_k(h) = \log_h \left(h^{-\alpha} \sum_{j=0}^N \left| (-1)^j \binom{\alpha}{j} - \gamma_{j,k} \right| \right),$$

whose behavior is reported in Figure 6.

As already mentioned, the numerical analysis reported in this section can also be used to select a proper value for k for a fixed time stepping h or viceversa. Figures 4 and 6 are in fact independent of the problem and can be easily used to this aim.

4.2. Absolute stability analysis. Let us consider the usual scalar test problem

$$(4.17) \quad {}_0D_t^\alpha y(t) = \lambda y(t), \quad \lambda \in \mathbb{C}, \quad y(0) = y_0,$$

whose exact solution is given by

$$y(t) = E_\alpha(\lambda t^\alpha) y_0.$$

It is known that $y(t) \rightarrow 0$ as $t \rightarrow +\infty$ whenever

$$|\arg(\lambda) - \pi| < \left(1 - \frac{\alpha}{2}\right) \pi,$$

(see (4.14) and [17]). The corresponding numerical approximation provided by an FBDF depends on the parameter $\zeta = h^\alpha \lambda$ and the absolute stability region of the method is defined as

$$\mathcal{D} \equiv \{\zeta \in \mathbb{C} : y_n = y_n(\zeta) \rightarrow 0 \text{ as } n \rightarrow +\infty\}.$$

A suitable reformulation of the theory developed in [17], which actually concerns methods for solving Abel-Volterra integral equations of the second kind, allows to prove that \mathcal{D} coincides with the complement of $\{\omega_p^{(\alpha)}(\zeta) : |\zeta| \leq 1\}$ so that a good approximation of the generating function should lead to similar stability domains and hence good linear stability properties. We consider the behavior of methods based on the Gauss-Jacobi rule whose corresponding absolute stability regions are given by, see (2.10)-(2.11),

$$\mathcal{D} = \mathbb{C} \setminus \left\{ \frac{p_m(\zeta)}{q_m(\zeta)} : |\zeta| \leq 1 \right\}.$$

REMARK 4.3. *From a theoretical point of view, from Theorem 4.2 one deduces that for $p = 1$, $q_m(\zeta)$ has no roots inside or on the unit circle of the complex plane for each $m = k$ and $\tau \in (0, 1]$. The absolute stability regions of such methods are therefore always unbounded since they are the complement of bounded regions.*

As shown in Figure 7, the methods simulate the behavior of the FBDFs rapidly, i.e. already for k and therefore m small. In particular, see the frame on the top of such figure, the stability domain of the method of degree $k = m = 12$ is very close to the one of the FBDF of order 1 (FBDF1).

5. Numerical examples. As first example, we consider the one-dimensional Nigmatullin's type equation

$$\begin{aligned} {}_0D_t^\alpha u(x, t) &= \frac{\partial^2 u(x, t)}{\partial x^2}, \quad t > 0, \quad x \in (0, \pi), \\ u(0, t) &= u(\pi, t) = 0, \\ u(x, 0) &= \sin x. \end{aligned}$$

If we discretize the spatial derivative by applying the classical central differences on a uniform mesh of meshsize $\delta = \pi/(s+1)$, we obtain the s -dimensional FDE

$$(5.1) \quad {}_0D_t^\alpha y(t) = Ly(t), \quad y(0) = y_0,$$

where $L = \delta^{-2} \cdot \text{tridiag}(1, -2, 1)$, and y_0 is the sine function evaluated at the interior grid points. It is known that y_0 is the eigenvector of L corresponding to its largest eigenvalue $\lambda = -4 \sin^2(\delta/2)/\delta^2$. This implies that the exact solution of (5.1) is given by, see (4.14),

$$y(t) = E_\alpha(t^\alpha \lambda) y_0.$$

In Figure 8 some results are reported. We compare the maximum norm of the error at each step of the FBDF of order 1 and the method based on the Gauss-Jacobi rule for some values of k and α . The initial values for the k -step schemes are defined according to the strategy described in Section 4. The reference solutions have been computed using the already mentioned `Matlab` function `m1f` from [22]. The dimension of the

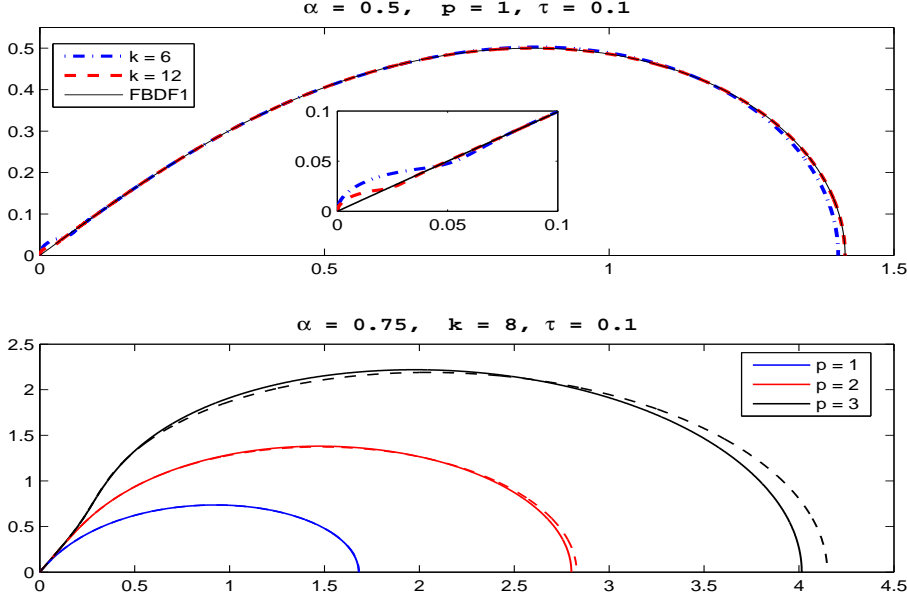


FIG. 7. Top: boundary of the stability domains of the methods based on the Gauss-Jacobi rule with $k = 6, 12$ and of the FBDF1. Down: boundary of the stability domains of the same methods of degree $k = 8$ for $p = 1, 2, 3$ (solid lines) and of the corresponding FBDFs (dashed lines). In each case the boundary is symmetric about the real axis.

problem is $s = 50$, and we consider a uniform time step $h = 1/N$ with $N = 250$ so that $h \approx \delta^2$. As one can see, if we set $\tau = 1$, i.e. if we use the classical Padé approximation of $(1 - \zeta)^{\alpha-1}$ (see Remark 3.3), the k -step methods simulate quite well the FBDF1 initially and an improvement of the results can be obtained by slightly increasing (considering to the total number of integration steps) the stepnumber k . A noticeable improvement can be obtained by choosing a different value of τ . In particular, if we set $\tau = \hat{\tau} = 4k/N$ (see (3.17)) then the 6-step method provides a numerical solution with the same accuracy of the one provided by the FBDF1 over the entire integration interval.

As second example we consider the following nonlinear problem

$$\begin{aligned} {}_0D_t^\alpha u(x, t) &= \frac{\partial(p(x)u(x, t))}{\partial x} + K_\alpha \frac{\partial^2 u(x, t)}{\partial x^2} + ru(x, t) \left(1 - \frac{u(x, t)}{K}\right), \\ u(0, t) &= u(5, t) = 0, \quad t \in [0, 1], \\ u(x, 0) &= x^2(5 - x)^2, \quad x \in [0, 5]. \end{aligned}$$

This is a particular instance of the time fractional Fokker-Planck equation with a nonlinear source term [24]. In population biology, its solution $u(x, t)$ represents the population density at location x and time t and the nonlinear source term in the equation is known as Fisher's growth term.

The application of the classical second order semi-discretization in space with stepsize $\delta = 5/(s + 1)$ leads to the following initial value problem

$$(5.2) \quad {}_0D_t^\alpha y(t) = Jy(t) + g(y(t)), \quad t \in (0, 1], \quad y(0) = y_0,$$

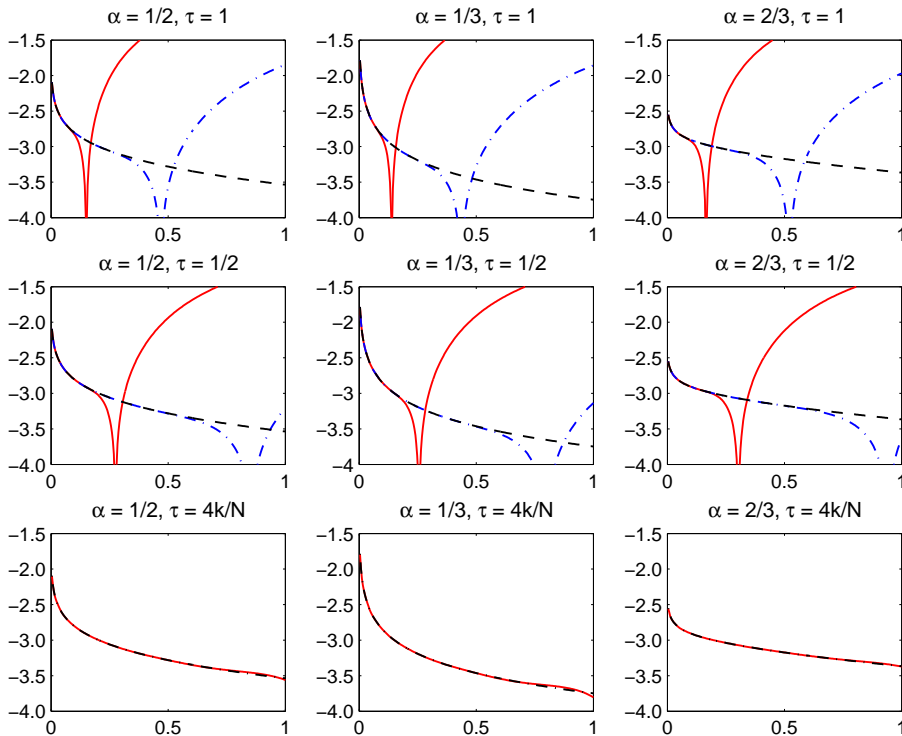


FIG. 8. Step by step error (in logarithmic scale) for the solution of (5.1) for the FBDF of order 1 (dashed line) and the method based on the Gauss-Jacobi rule with $k = 6$ (solid line) and $k = 12$ (dash-dotted line).

where, for each $i = 1, \dots, s$, $(y(t))_i \equiv y_i(t) \approx u(i\delta, t)$, $y_i(0) = u(i\delta, 0)$, $(g(y))_i = ry_i(1 - y_i/K)$, and J is a tridiagonal matrix whose significant entries are

$$J_{ii} = p'(x_i) - \frac{2K_\alpha}{\delta^2}, \quad i = 1, \dots, s,$$

$$J_{i,i-1} = -\frac{p(x_i)}{2\delta} + \frac{K_\alpha}{\delta^2}, \quad J_{i-1,i} = \frac{p(x_{i-1})}{2\delta} + \frac{K_\alpha}{\delta^2}, \quad i = 1, \dots, s-1.$$

In our experiment, we set $\alpha = 0.8$, $p(x) = -1$, $r = 0.2$, $K_\alpha = K = 1$ (see [24, Example 5.4]) and $s = 90$. We solved (5.2) over a uniform meshgrid with stepsize $h = 1/256$ by using the FBDF1 and the 6-step method with $\tau = 24/256$. The so-obtained numerical solutions have the same qualitative behavior as shown in Figure 9 for different times $t = 1/8, 1/2, 1$. This is confirmed by the step by step maximum norm of the difference between them reported in Figure 10.

In addition, in the same figure the step by step maximum norm of the difference between the numerical solutions provided by the FBDF1 and by the polynomial approximation with memory length $k = 12$ is drawn. As one can see, the advantage of our approach with respect to the classical short memory principle is undeniable.

6. Conclusion. In this paper we have presented a new approach for the construction of m -step formulas for the solution of FDEs. The method shows encouraging results in the discrete approximation of the FDE solution especially if we consider the computational saving with respect to the attainable accuracy. Indeed good results

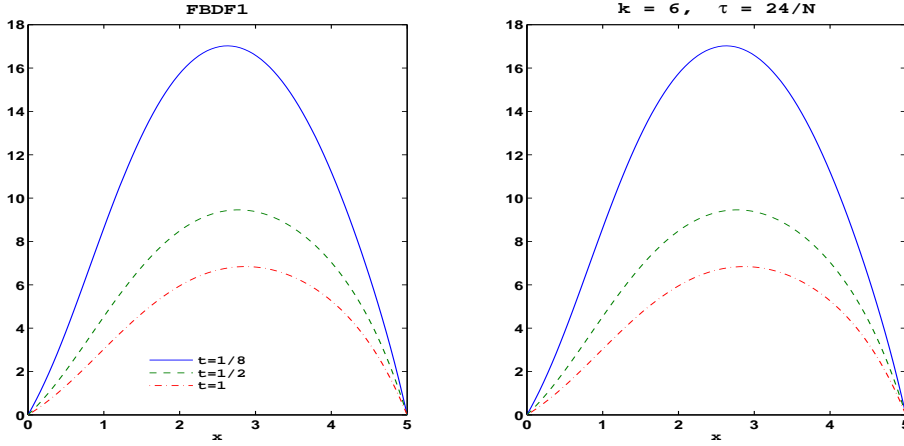


FIG. 9. Numerical solution of (5.2) with $\alpha = 0.8$ provided by the FBDF1 and the method based on the Gauss-Jacobi rule with $k = 6$ at $t = 1/8, 1/2, 1$.

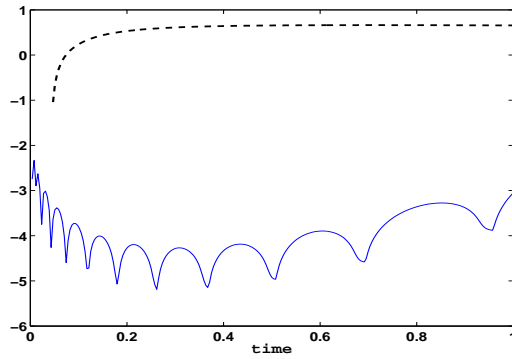


FIG. 10. Step by step norm of the difference (in logarithmic scale) between the numerical solutions provided by the FBDF1 and that given by the Gauss-Jacobi rule with $k = 6$ (solid line) or by the polynomial approximation with $k = 12$ (dashed line).

are attainable with short memory length. Theoretically the method is 0-stable and the consistency is well simulated. The absolute linear stability is preserved.

With respect to the standard linear multistep techniques, the basic drawback of this class of methods is that they need to be properly set in order to avoid poor accuracy. In particular the number of nodes k of the Gauss-Jacobi rule (with respect to the number of points N of the time-discretization) may have a great influence on the accuracy (cf. Figure 8). A proper setting should be done by considering the numerical consistency (Section 4.1). Also the choice of the parameter τ of the transform (2.5) is crucial for the accuracy (see again Figure 8).

We remark that even if the paper only deals with the approximation of FBDFs, the ideas can easily be extended to other Fractional Linear Multistep Methods (FLMMs) such as the well known Fractional Adams type methods. It is just necessary to detect the generating function or the corresponding Toeplitz matrix and then apply the technique presented in the paper.

We further remark that even if the fractional derivative approximation is confined

here to the case $0 < \alpha < 1$, the extension to $\alpha > 1$ is immediate. For instance, the Grünwald-Letnikov discretization prescribes to take as coefficients of the backward differentiation, the Taylor coefficients of $(1 - \xi)^\alpha$, that is, the first column of A_1^α . In this view, each definition of fractional derivative that can be written as $[p(\xi)]^\alpha$ where $p(\xi)$ is a polynomial representing an integer order derivative can be suitably approximated by a rational form that finally leads to a short memory method.

REFERENCES

- [1] D.A. Bini, N.J. Higham, B. Meini, Algorithms for the matrix p th root, *Numer. Algorithms* 39 (2005), 349–378.
- [2] H. Brunner, P.J. van der Houwen, *The numerical solution of Volterra equations*. CWI Monographs, 3. North-Holland Publishing Co., Amsterdam, 1986.
- [3] W. Deng, Short memory principle and a predictor-corrector approach for fractional differential equations, *J. Comput. Appl. Math.* 206 (2007), 174–188.
- [4] K. Diethelm, N.J. Ford, Analysis of fractional differential equations, *J. Math. Anal. Appl.* 265 (2002), 229–248.
- [5] K. Diethelm, J.M. Ford, N.J. Ford, M. Weilbeer, Pitfalls in fast numerical solvers for fractional differential equations, *J. Comput. Appl. Math.* 186 (2006), 482–503.
- [6] N.J. Ford, A.C. Simpson, The numerical solution of fractional differential equations: speed versus accuracy, *Numer. Algorithms* 26 (2001), 333–346.
- [7] A. Frommer, S. Güttel, M. Schweitzer, Efficient and stable Arnoldi restarts for matrix functions based on quadrature, Preprint BUW-IMACM 13/17 (2013) <http://www.math.uni-wuppertal.de>.
- [8] L. Galeone, R. Garrappa, On multistep methods for differential equations of fractional order, *Mediterr. J. Math.* 3 (2006), 565–580.
- [9] O. Górník, F. Greco, K. Ziętak, A Padé family of iterations for the matrix sign function and related problems, *Numerical Linear Algebra with Applications* 19 (2012), 585–605.
- [10] E. Hairer, S.P. Norsett, G. Wanner, *Solving ordinary differential equations. I. Nonstiff problems*. Second edition. Springer Series in Computational Mathematics, 8. Springer-Verlag, Berlin, 1993.
- [11] N. Hale, N.J. Higham, L.N. Trefethen, Computing A^α , $\log(A)$, and related matrix functions by contour integrals, *SIAM J. Numer. Anal.* 46 (2008), 2505–2523.
- [12] N.J. Higham, *Functions of matrices. Theory and computation*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.
- [13] F.B. Hildebrand, *Introduction to Numerical Analysis*. New York: McGraw-Hill, 1956.
- [14] J.R. Li, A fast time stepping method for evaluating fractional integrals, *SIAM J. Sci. Comput.* 31 (2010), 4696–4714.
- [15] M. López-Fernández, C. Lubich, A. Schädle, Adaptive, fast, and oblivious convolution in evolution equations with memory, *SIAM J. Sci. Comput.* 30 (2008), 1015–1037.
- [16] C. Lubich, Discretized fractional calculus, *SIAM J. Math. Anal.* 17 (1986), 704–719.
- [17] C. Lubich, A stability analysis of convolution quadratures for Abel-Volterra integral equations, *IMA J. Numer. Anal.* 6 (1986), 87–101.
- [18] W. McLean, Fast summation by interval clustering for an evolution equation with memory, *SIAM J. Sci. Comput.* 34 (2012), A3039–A3056.
- [19] P. Novati, Numerical approximation to the fractional derivative operator, *Numer. Math.* 127 (2014), 539–566.
- [20] A.P. Prudnikov, Yu.A. Brychkov, O.I. Marichev, *Integrals and Series, vol.3: More Special Functions*, Gordon and Breach Science Publishers, 1990.
- [21] I. Podlubny, *Fractional differential equations*. Mathematics in Science and Engineering, 198. Academic Press, Inc., San Diego, CA, 1999.
- [22] I. Podlubny, Mittag-Leffler Function, <http://www.mathworks.com/matlabcentral/fileexchange/8738>, 2009.
- [23] A. Schädle, M. López-Fernández, C. Lubich, Fast and oblivious convolution quadrature, *SIAM J. Sci. Comput.* 28 (2006), 421–438.
- [24] Q. Yang, F. Liu, I. Turner, Stability and convergence of an effective numerical method for the Time-Space Fractional Fokker-Planck Equation with a nonlinear source term, *Int. J. Differ. Equ.* (2010), Art. ID 464321, 22 pp.