Well-posedness of \(hp\)-version discontinuous Galerkin methods for fractional diffusion wave equations

Kassem Mustapha†
Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, 31261, Saudi Arabia
AND
Dominik Schötzau‡
Department of Mathematics, University of British Columbia, Vancouver, BC, V6T 1Z2, Canada

[Received on January 2013]


We establish the well-posedness of an \(hp\)-version time-stepping discontinuous Galerkin (DG) method for the numerical solution of fractional super-diffusion evolution problems. In particular, we prove the existence and uniqueness of approximate solutions for generic \(hp\)-version finite element spaces featuring non-uniform time-steps and variable approximation degrees. We then derive new \(hp\)-version error estimates in a non-standard norm, which are completely explicit in the local discretization and regularity parameters. As a consequence, we show that by employing geometrically refined time-steps and linearly increasing approximation orders, exponential rates of convergence in the number of temporal degrees of freedom are achieved for solutions with singular (temporal) behavior near \(t = 0\) caused by the weakly singular kernel. Moreover, we show optimal algebraic convergence rates for \(h\)-version approximations on graded meshes. We present a series of numerical tests where we verify experimentally that our theoretical convergence properties also hold true in the stronger \(L_\infty\)-norm.

Keywords: Fractional diffusion, \(hp\)-version discontinuous Galerkin methods, convergence analysis

1. Introduction

We propose and analyze an \(hp\)-version discontinuous Galerkin (DG) method for the temporal discretization of fractional wave equations of the form

\[
 u'(t) + B_\alpha A u(t) = f(t), \quad t \in (0, T), \quad \text{with } u(0) = u_0, \tag{1.1}
\]

where \(u' = \frac{du}{dt}\), \(A\) is a self-adjoint linear elliptic spatial operator (independent of time), and \(B_\alpha\) is the Riemann–Liouville fractional integral operator of order \(\alpha \in (0, 1)\), given by

\[
 B_\alpha v(t) = \int_0^t \omega_\alpha(t-s)v(s)\,ds \quad \text{with} \quad \omega_\alpha(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}, \tag{1.2}
\]

†Email: kassem@kfupm.edu.sa
‡Email: schoetzau@math.ubc.ca

© The author 2014. Published by Oxford University Press on behalf of the Institute of Mathematics and its Applications. All rights reserved.
and in the following also referred to as memory or convolution term.

Note that with the choice \( A = -\Delta \), problem (1.1) is sometimes called a fractional wave equation \([29]\), since in the limit case \( \alpha = 1 \) we obtain the wave equation \( u'' - \Delta u = f'(t) \) (after differentiation with respect to \( t \)). Problem (1.1) can be thought of as a model problem occurring in the theory of anomalous diffusion processes, or in wave propagation in viscoelastic materials. The notions and concepts of anomalous dynamical properties have been predicted and observed in numerous systems from various disciplines, see for example, \([6, 7, 13, 14]\).

Features of anomalous diffusion include history dependence (memory term), long-range (or non-local) correlation in time and heavy tail characteristics. If the anomalous diffusion exponent \( \alpha \) lies in the interval \(( -1, 0 )\), the underlying diffusion process will be categorized as sub-diffusive and the fractional operator \( B_\alpha \) in (1.1) corresponds to the Riemann–Liouville fractional derivative operator \( \partial_t^{-\alpha} \), see, for example, \([25]\) and references therein. However, if \( \alpha \in (0, 1) \), the process is called super-diffusive, and we shall focus on this case. For an extensive study of anomalous diffusion problems, we refer the reader to \([21]\) and the references therein.

A number of low-order methods for numerically solving (1.1) can be found in the literature. For example, the combination of finite difference methods with numerical quadrature rules for the memory term has been studied in \([12, 17, 19, 27]\) and the references therein. If convolution quadrature techniques are employed as in \([3]\), fast summation can be invoked; see \([28]\). In the recent papers \([24, 25]\), piecewise constant and linear DG methods in time have been proposed and analyzed for (1.1). The error analysis relies on the observation that on each time interval, the DG solution takes its maximum values at one of the end points. However, this is no longer true in the case of DG methods of higher order. For other low-order numerical methods (including adaptive discretizations, finite difference and ADI backward Euler) for problems of the form (1.1), we refer the reader to \([1, 9, 34]\), and also to \([5, 37]\) for positive-type and completely monotonic kernels.

Let us also digress by mentioning that in the case \( A = -\Delta u \), various numerical methods have been applied for the following alternative representation of the fractional diffusion wave problem:

\[
\int_0^t \omega_{1-\alpha}(t - s)u''(s) \, ds - \Delta u(t) = f(t), \quad t \in (0, T);
\]

see \([4, 22, 33, 36, 38]\) and references therein. By applying the Riemann–Liouville fractional integral operator \( B_\alpha \) to both sides of (1.3), problem (1.3) can be cast into the general form of the fractional wave problem (1.1). A similar remark is also applicable for the fractional diffusion wave problem:

\[
u''(t) - \frac{\partial}{\partial t} \int_0^t \omega_\alpha(t - s)u(s) \, ds = f(t), \quad t \in (0, T) .\]

The formulations in (1.1), (1.3), and (1.4) are equivalent under suitable (but reasonable) assumptions on the source terms and the initial data. However, the numerical methods obtained for each representation are formally different. Here we shall focus on \( hp\)-DG methods for problem (1.1); similar \( hp\)-DG approaches might be feasible for equations (1.3) and (1.4).

For problem (1.1) and to the best of our knowledge, the only class of methods which achieve spectral accuracy are based on inverting the Laplace transform of the solutions; see, e.g., \([10, 11, 20]\) and the references therein. A difficulty with such schemes is their need to compute an analytic continuation of the Laplace transform of the inhomogeneous source term \( f'(t) \) in (1.1). In addition, it is unclear whether this approach can be modified to handle effectively non-linear versions of (1.1). Here, we pursue an alternative approach for achieving high-order accuracy based on \( hp\)-version DG time-stepping. This approach easily allows for the incorporation of non-uniform time-steps and locally varying polynomial
degrees, as is needed for proper $hp$-version refinement strategies. Indeed, in [23, 30], it has been shown in the context of parabolic problems that temporal singularities caused by the weakly singular kernel $\omega_\alpha$ or by incompatible initial data can be resolved at exponential rates of convergence in the number of temporal degrees of freedom. We mention that $hp$-version continuous Galerkin (CG) methods might be feasible and more appropriate for wave-type problems. Extensions of the analysis presented herein to $hp$-CG methods might be possible by following along the lines of this paper and those of [23], in conjunction with the $hp$-analysis in [35] for CG methods for initial-value problems. This will be a topic for future research.

The nonlocal nature of the operator $B_\alpha$ means that on each time subinterval, one must efficiently evaluate a sum of integrals over all previous time subintervals. Thus, it is important to maintain high accuracy of the discrete solutions and to reduce the number of time-steps as much as possible. Due to their spectral and exponential convergence properties, $hp$-DG methods allow one to achieve these requirements to a large extent, both for smooth and singular solutions. In addition, the sum involving all previous times levels can be evaluated via fast algorithms; see for example [16].

The main purpose of the present paper is to establish the well-posedness of the $hp$-version DG method. We point out that this is a non-trivial task since the DG method allows one to only control the jumps of the approximate solutions. While this is enough for low-order DG schemes as mentioned above, this is the major difficulty for the analysis of $hp$-version DG methods for (1.1). Notice also that this in contrast to the parabolic setting considered in [23, 30], where well-posedness of the DG methods follows relatively straightforwardly from the different nature of the equations. To overcome this difficulty and to extend the piecewise linear approach of [24, 25] to higher-order approximations, we show the positivity and coercivity of the memory term in a non-standard norm. This will be the key to establish the existence, uniqueness and the convergence of discrete solutions.

We derive generic $hp$-version error estimates, which are completely explicit in the local step sizes, the local polynomial degrees, and the local regularity of the analytical solution. Consequently, we then proceed along the lines of [23, 30] and investigate two refinement strategies in the case where the solution $u$ of (1.1) lacks regularity as $t \to 0$ (due to the weakly singular kernel $\omega_\alpha(t)$). First, we show that $hp$-version DG schemes based on geometrically refined time-steps and on linearly increasing approximation orders achieve exponential convergence rates in the number of time degrees of freedom. Second, we prove optimal algebraic convergence rates of the $h$-version DG method over non-uniform meshes that are properly graded towards $t = 0$.

We also present a series of numerical tests which indicate the validity of our theoretical convergence properties in the stronger $L_\infty(0, T)$-norm; this is not covered by our theoretical analysis and is the subject of ongoing research. Our tests include a scalar problem, as well as a problem in one space dimension where we combine our $hp$-version time-stepping method with a standard (continuous) finite element discretization in space.

The outline of the paper is as follows. In Section 2, we fix some technical assumptions and notations, and introduce the $hp$-version time-stepping DG method. In Section 3, we prove our main results regarding the well-posedness of the approximate solutions. Section 4 is devoted to the error analysis of the methods and to the proof of various convergence results. In Section 5, we present a series of numerical examples.

2. DG time discretization

In this section, we specify our abstract setting used for problem (1.1), and introduce the $hp$-version time-stepping DG method for its discretization in time.
Hence, the above eigenfunction expansions imply the following positivity property:

\[ \mathbb{X} = \{ v \in \mathbb{H} : ||v||_{\mathbb{X}} < \infty \} \quad \text{where} \quad ||v||_{\mathbb{X}}^2 = ||A^{1/2}v||^2 = \sum_{m=1}^{\infty} \lambda_m (v, \phi_m)^2, \]

and associate with the linear operator \( A \) the bilinear form, denoted by the same symbol:

\[ A(u,v) = \sum_{m=1}^{\infty} A_m u_m v_m \quad \text{where} \quad u_m = \langle u, \phi_m \rangle \quad \text{and} \quad v_m = \langle v, \phi_m \rangle \quad \text{for} \quad u, v \in \mathbb{X}. \]

As in [17, 18], we have the following positivity property of \( B_\alpha \) for scalar functions:

\[ \int_0^T B_\alpha v(t) v(t) dt \geq 0. \]

Hence, the above eigenfunction expansions imply the following positivity property:

\[ \int_0^T A(B_\alpha v,v) dt = \sum_{j=1}^{\infty} \lambda_j (B_\alpha v_m, v_m)_T \geq 0, \]

where from now on we use the notation \((\cdot, \cdot)_T\) and \( || \cdot ||_T \) for the inner product and the associated norm in the Lebesgue space \( L_2(0,t) \), respectively.

As a concrete example of our setting, one may take \( \mathbb{H} = L_2(\Omega) \) for a bounded, Lipschitz domain \( \Omega \subseteq \mathbb{R}^d \) and \( A = -\Delta \) subject to homogeneous Dirichlet boundary conditions. In this case, \( \mathbb{X} = H_0^1(\Omega) \) and \( ||u||_{\mathbb{X}} = ||\nabla u||_{L^2(\Omega)} \).

### 2.1 Time-stepping

To describe the \( hp \)-DG method, we introduce a (possibly non-uniform) partition \( \mathcal{M} \) of the time interval \([0,T]\) given by the points \( 0 = t_0 < t_1 < \cdots < t_N = T \). We set \( t_n = (t_{n-1}, t_n) \) and \( k_n = t_n - t_{n-1} \) for \( 1 \leq n \leq N \). The maximum step-size is defined as \( k = \max_{1 \leq n \leq N} k_n \). We also introduce the space

\[ C(\mathcal{M};\mathbb{X}) = \{ v \in L_2((0,T);\mathbb{X}) : v|_{t_n} \in C(I_k;\mathbb{X}), \ 1 \leq n \leq N \}. \]

For a function \( v \in C(\mathcal{M};\mathbb{X}) \), we write \( v_n^- = v(t_n^-), \ v_n^+ = v(t_n^+) \), and set \( |v|^n = v_n^+ - v_n^- \), for \( n = 0, \cdots, N \) with \( v_0^0 = v_0^+ \) and \( v_N^0 = v_N^- \).

With each subinterval \( I_n \) we associate a polynomial degree \( p_n \in \mathbb{N}_0 \). These degrees are then stored in the degree vector \( \mathbf{p} = (p_1, p_2, \ldots, p_N) \). Next, we introduce the discontinuous finite element space

\[ \mathcal{W}(\mathcal{M},\mathbf{p};\mathbb{X}) = \{ v \in C(\mathcal{M};\mathbb{X}) : v|_{t_n} \in \mathbb{P}_{p_n}(\mathbb{X}), \ 1 \leq n \leq N \}, \]

where \( \mathbb{P}_{p_n}(\mathbb{X}) \) denotes the space of polynomials of degree \( \leq p_n \) with coefficients in \( \mathbb{X} \).

Following [24], the \( hp \)-version DG approximation \( U \in \mathcal{W}(\mathcal{M},\mathbf{p};\mathbb{X}) \) is now defined as follows: Given \( U(t) \) for \( 0 \leq t \leq t_{n-1} \), then \( U \in \mathbb{P}_{p_n}(\mathbb{X}) \) on the next time-step \( I_n \) is determined by requesting that

\[ \langle U(t_n^-), X(n-1) \rangle + \int_{t_{n-1}}^{t_n} \left( \langle U', X \rangle + A(B_\alpha U, X) \right) dt = \langle U(t_{n-1}^-), X(n-1) \rangle + \int_{t_{n-1}}^{t_n} (f, X) dt \]

(2.4)
for all \( X \in \mathbb{P}_p(X) \). The time-stepping procedure is initialized with a suitable approximation \( U^0 \) to \( u_0 \). After \( N \) steps, it yields the approximate solution \( U \in \mathcal{W}(\mathcal{M}, \mathbb{P}; X) \) for \( 0 \leq t \leq T_N \).

Note that, in the scheme (2.4), we have used that the operators \( \mathcal{B}_\alpha \) and \( A \) commute. Indeed, this is due to the fact that \( \mathcal{B}_\alpha \) is a purely time-dependent operator while the spatial \( A \) is assumed to be independent of time.

### 3. Existence and uniqueness of DG solutions

In this section, we show the well-posedness of the discrete solutions by diagonalizing (1.1) based on eigenfunction expansions. We start by showing the following crucial properties of the memory term.

**Lemma 3.1** Let \( v, w \in C(\mathcal{M}; \mathbb{R}) \), and \( c_\alpha = \cos(\alpha \pi/2) \) for \( 0 < \alpha < 1 \). Then:

(i) If we have \( \max_{n=0}^N \{ |v^n_\alpha|^2, |v^n_{\alpha'}|^2 \} + (\mathcal{B}_\alpha v, v)_T = 0 \), then \( v \equiv 0 \) on \((0, T)\).

(ii) The memory term satisfies the coercivity property: \( (\mathcal{B}_\alpha v, v)_T \geq c_\alpha \| \mathcal{B}_\alpha v \|_T^2 \).

(iii) The memory term satisfies the continuity property: \( |(\mathcal{B}_\alpha w, v)_T|^2 \leq \frac{1}{c_\alpha} (\mathcal{B}_\alpha v, v)_T (\mathcal{B}_\alpha w, w)_T \).

**Proof.** First, we recall the positivity property of \( \mathcal{B}_\alpha \) in (2.1) for any \( v \in C(\mathcal{M}; \mathbb{R}) \). Thus, from the given identity in \((i)\), we have

\[
\max_{n=0}^N \{ |v^n_\alpha|^2, |v^n_{\alpha'}|^2 \} = 0 \quad \text{and} \quad (v, \mathcal{B}_\alpha v)_T = 0.
\]

The first identity ensures that \( v \) has no jumps over the interior nodes \( t_n \) for \( n = 1, \ldots, N - 1 \). Without loss of generality, we may thus assume that \( v \in C([0, T]; \mathbb{R}) \).

The rest of our proof of \((i)\) is based on using Laplace transforms. To that end, we extend \( v \) by zero outside the interval \((0, T)\) and define the Laplace transform \( \hat{v} \) of \( v \) by

\[
\hat{v}(iy) = \int_0^\infty e^{-iyt} v(t) \, dt.
\]

By Plancherel’s theorem, the fact that \( \overline{\hat{v}(iy)} = \hat{v}(-iy) \) (\( v \) is a real-valued function) and since \( \hat{\mathcal{B}_\alpha}(iy) = \hat{v}(iy)^{-\alpha} \), we find that

\[
\int_0^\infty v(t) \mathcal{B}_\alpha v(t) \, dt = \int_{-\infty}^{\infty} \hat{v}(t) \int_{-\infty}^{\infty} \hat{\omega}(t-s) \hat{v}(s) \, ds \, dt
\]

\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{\hat{v}(iy)} \hat{\omega}(iy) \hat{v}(iy) \, dy
\]

\[
= \frac{1}{\pi} \int_0^\infty \text{Re}(iy)^{-\alpha} |\hat{v}(iy)|^2 \, dy = \frac{1}{\pi} c_\alpha \int_0^\infty y^{-\alpha} |\hat{v}(iy)|^2 \, dy.
\]

Since we have \( \int_0^\infty v(t) \mathcal{B}_\alpha v(t) \, dt = 0 \) and \( y^{-\alpha} c_\alpha > 0 \) for \( y > 0 \) and \( 0 < \alpha < 1 \), it follows that \( \hat{v} = 0 \) almost everywhere in \((0, \infty)\). Using this and because \( v \in C([0, T]; \mathbb{R}) \), we obtain \( v \equiv 0 \) on \((0, T)\). Hence, the proof of property \((i)\) is completed.

Next, we show the coercivity result \((ii)\) for the memory term. Using (3.1), we have that

\[
(\mathcal{B}_\alpha v, v)_T = \frac{1}{\pi} c_\alpha \int_0^\infty y^{-\alpha} |\hat{v}(iy)|^2 \, dy = \frac{1}{\pi} c_\alpha \int_0^\infty |(iy)^{-\alpha} \hat{v}(iy)|^2 \, dy.
\]
But,

\[
\int_0^\infty |(iy - \frac{q}{2}\hat{\nu}(iy)|^2 dy = \int_0^\infty |\mathcal{B}_q^2\hat{\nu}(iy)|^2 dy = \pi \int_0^\infty |\mathcal{B}_q^2 v(t)|^2 dt \geq \pi \int_0^T |\mathcal{B}_q^2 v(t)|^2 dt,
\]

and hence the coercivity result \((ii)\) now follows.

To show inequality \((iii)\), we follow again \((3.1)\) and obtain

\[
|\langle \mathcal{B}_q w, v \rangle_T| \leq \frac{1}{2\pi} \int_{-\infty}^\infty |\overline{\nu}(iy)(iy-\alpha\hat{\nu}(iy)| dy.
\]

Noting that, for any \(\varepsilon > 0\), we have

\[
\left( \int_{-\varepsilon}^\varepsilon |\overline{\nu}(iy)(iy-\alpha\hat{\nu}(iy)| dy \right)^2 \leq \int_{-\infty}^\infty |y^{-\frac{q}{2}}\overline{\nu}(iy)|^2 dy \int_{-\infty}^\infty |y^{-\frac{q}{2}}\hat{\nu}(iy)|^2 dy
\]

\[
= 4 \int_0^\infty y^{-\alpha}|\overline{\nu}(iy)|^2 dy \int_0^\infty y^{-\alpha}|\hat{\nu}(iy)|^2 dy
\]

\[
= \frac{4\pi^2}{c_{\alpha}^2} \langle \mathcal{B}_q v, v \rangle_T \langle \mathcal{B}_q w, w \rangle_T
\]

and therefore the proof of inequality \((iii)\) follows.

We are now ready to prove the existence and uniqueness of DG solutions.

**THEOREM 3.1** The discrete solution \(U\) of \((2.4)\) exists and is unique.

**Proof.** Since \(A\) possesses a complete orthonormal eigensystem \(\{\lambda_m, \phi_m\}_{m \geq 1}\), problem \((2.4)\) can be reduced to a linear system of equations on each subinterval \(I_n\). Indeed, if we take \(X = \phi_m w\) with \(w \in \mathbb{P}_{p_1}(\mathbb{R})\) in \((2.4)\), then we find that

\[
U_{m,n-1}^{n-1} w_{n-1}^+ + \int_{i_{m-1}}^{i_n} \left( \nu_m' w + \lambda_m \mathcal{B}_q U_m w \right) dt = U_{m,n-1}^{n-1} w_{n-1}^+ + \int_{i_{m-1}}^{i_n} f_m w dt
\]

for all \(m \geq 1\) and \(w \in \mathbb{P}_{p_1}(\mathbb{R})\), with \(U_m = \langle U, \phi_m \rangle \in \mathbb{P}_{p_1}(\mathbb{R})\), and \(f_m = \langle f, \phi_m \rangle\).

Because of the finite dimensionality of system \((3.2)\) \(((p_n+1) \times (p_n+1)\) equations), the existence of the scalar function \(U_m\) follows from its uniqueness. Since the DG solution is constructed element by element, it is enough to show the uniqueness on the first time interval \((0, t_1)\). That is, it is enough to consider \(n = 1\) in \((3.2)\) (for \(n \geq 2\) the proof is completely analogous). To this end, let \(U_{m_1}\) and \(U_{m_2}\) be two DG solutions on \(I_1\). By linearity, the difference \(V_m := (U_{m_1} - U_{m_2})|_{t_1}\) then satisfies

\[
V_{m,+}^0 w_{m,+}^0 + \int_0^{t_1} \left( V_{m,+}^0 w + \lambda_m \mathcal{B}_q V_m w \right) dt = 0 \quad \forall w \in \mathbb{P}_{p_1}(\mathbb{R}), \quad \forall m \geq 1.
\]

Choosing \(w = V_m\) and integrating yield

\[
\frac{1}{2} \left( |V_{m,-}^0|^2 + |V_{m,+}^0|^2 \right) + \lambda_m (\mathcal{B}_q V_m, V_m)_{t_1} = 0.
\]

Therefore, from the positivity \(\lambda_m (\mathcal{B}_q V_m, V_m)_{t_1} \geq 0\), and Lemma 3.1 \((i)\) (with \(t_1\) in place of \(T\)), we conclude that \(V_m \equiv 0\) on \(I_1\). 

\(\square\)
4. Error analysis

This section is devoted to deriving error estimates for the $hp$-DG method. Our first result are $hp$-error estimates that are explicit in all the parameters of interest. Then we discuss several implications of these results, both for smooth and singular solutions.

4.1 Error estimates

We shall show errors bounds in the space $C(\mathcal{M}; \mathcal{X})$, endowed with the non-standard norm

$$\|v\|^2_{\mathcal{M}, \mathcal{X}} := \max_{n=0}^{N} \left\{ \|v^n\|^2, \|v^n_{\mathcal{A}}\|^2 \right\} + c_\alpha \|\mathcal{B}v\|^2_{L_2(\mathcal{X})}$$

for each fixed $0 < \alpha < 1$,

where for any Sobolev spaces $V$ in time and $W$ in space, we abbreviate the Bochner norm $\|\cdotp\|_{((0,T);W)}$ by $\|\cdotp\|_{V(W)}$.

**Remark 4.1** For each fixed $0 < \alpha < 1$, the expression $\|\cdotp\|_{\mathcal{M}, \mathcal{X}}$ defines a norm on $C(\mathcal{M}; \mathcal{X})$. This follows from eigenfunction expansions, and the first two properties in Lemma 3.1.

For convenience, we reformulate the DG scheme (2.4) in terms of the global bilinear form

$$G_N(U, X) = \langle U^0, X^0 \rangle + \sum_{n=1}^{N-1} \langle [U]^n, [X]^n \rangle + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left( \langle U', X \rangle + A(\mathcal{B}U, X) \right) dt. \tag{4.1}$$

By summing (2.4) over all time-steps, the DG method can now equivalently be written as: Find $U \in \mathcal{H}(\mathcal{M}, \mathcal{X})$ such that

$$G_N(U, X) = \langle U^0, X^0 \rangle + \int_0^{t_N} \langle f, X \rangle dt \quad \forall X \in \mathcal{H}(\mathcal{M}, \mathcal{X}). \tag{4.2}$$

**Remark 2.** Integration by parts yields the following alternative expression for the bilinear form $G_N$:

$$G_N(U, X) = \langle U^N, X^N \rangle - \sum_{n=1}^{N-1} \langle U^n, [X]^n \rangle + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left( -\langle U', X \rangle + A(\mathcal{B}U, X) \right) dt. \tag{4.4}$$

Since $G_N(u, X) = \langle u_0, X_0^0 \rangle + \int_0^{t_N} \langle f, X \rangle dt$ (the solution $u$ is continuous with values in $\mathcal{X}$), we have the Galerkin orthogonality property:

$$G_N(U^0 - u, X) = \langle U^0 - u_0, X_0^0 \rangle \quad \forall X \in \mathcal{H}(\mathcal{M}, \mathcal{X}). \tag{4.5}$$

4.1.1 An $hp$-version projection operator. For a continuous function $u : [0, T] \to \mathcal{X}$ we define the piecewise $hp$-interpolant $\Pi u : [0, T] \to \mathcal{H}(\mathcal{M}, \mathcal{X})$ by setting

$$\Pi u(t_n) = u(t_n) \in \mathcal{X} \quad \text{and} \quad \int_{t_{n-1}}^{t_n} \langle u - \Pi u, v \rangle dt = 0 \quad \forall v \in \mathcal{P}_{p-1}(\mathcal{X}). \tag{4.6}$$

Note that for $p = 0$, the second set of conditions is not required.

Then the following $hp$-version approximation properties of $\Pi$ holds true $[23, 30]$.

**Theorem 4.1** Let $1 \leq n \leq N$ and $0 \leq q_n \leq p_n$. Then:
Remark 4.2 and the fact that

By choosing \( u \) with the task of estimating \( \theta \), we decompose the error

Moreover, we have that \( \eta^n = 0 \) for all \( 1 \leq n \leq N \). Hence, using the alternative expression for \( G_N \) in Remark 4.2 yields

Moreover, \( \int_{t_{n-1}}^{t_n} \langle \eta, X' \rangle \, dt = 0 \) by definition of the operator \( \Pi \) (note that for \( p_n = 0 \), we have \( X' \equiv 0 \)). Therefore, we conclude that

First, we show the following bound.

Lemma 4.1 We have

\[
\| \theta \|_{\alpha, X} \leq 2\| U^0 - u_0 \| + \frac{\sqrt{2}}{c_\alpha} \omega_{\alpha+1}(T) \| \eta \|_{L^2(X)}.
\]

Proof. By choosing \( X = \theta \) in (4.6), using the alternative definition of \( G_N \) (with \( n \) in place of \( N \)) in Remark 4.2 and the fact that \( \langle \theta', \theta \rangle = \frac{1}{2} \frac{d}{dt} \| \theta \|^2 \), we observe after elementary manipulations that

\[
\| \theta^n \|^2 + \| \theta^n_+ \|^2 + \sum_{j=1}^{n-1} \| \theta^j \|^2 + 2 \int_{t_{n-1}}^{t_n} A(\partial_\alpha \theta, \theta) \, dt = 2\langle U^0 - u_0, \theta^n_+ \rangle - 2 \int_{t_{n-1}}^{t_n} A(\partial_\alpha \eta, \theta) \, dt \quad \text{for } n \geq 1.
\]
Therefore, the use of the inequality $2(U_0^0 - u_0, \theta_0^0) \leq 2\|U_0^0 - u_0\|^2 + \frac{1}{2} \|\theta_0^0\|^2$, we notice that

$$2\|\theta^n\|^2 + \|\theta_0^0\|^2 + 2\sum_{j=1}^{n-1} \|\theta_j\|^2 + 2\int_0^{t_n} A(\mathcal{B}_\alpha \theta, \theta) dt \leq 4\|U_0^0 - u_0\|^2 + \frac{2}{c_\alpha} \int_0^{t_n} |A(\mathcal{B}_\alpha \eta, \eta)| dt.$$ 

Now, by the coercivity result in Lemma 3.1 (ii) (with $t_n$ in place of $T$), we have

$$\int_0^{t_n} A(\mathcal{B}_\alpha \theta, \theta) dt = \sum_{m=1}^{\infty} \lambda_m \int_0^{t_n} \mathcal{B}_\alpha \theta_m \theta_m dt \geq c_\alpha \sum_{m=1}^{\infty} \lambda_m \|\mathcal{B}_\alpha \theta_m\|_{\mathcal{T}}^2 = c_\alpha \|\mathcal{B}_\alpha \theta\|_{L_2(0,t_n;\mathcal{X})}^2.$$ 

Therefore, the use of the inequality $\|\theta_n\|^2 \leq 2\|\theta_{n-1}\|^2 + 2\|\theta\|^2$ for $n \geq 2$, yields

$$\max_{n=0} \{\|\theta_n\|^2, \|\theta_{n-1}\|^2\} + 2c_\alpha \|\mathcal{B}_\alpha \theta\|_{L_2(\mathcal{X})}^2 \leq 4\|U_0^0 - u_0\|^2 + \frac{2}{c_\alpha} \int_0^{T} |A(\mathcal{B}_\alpha \eta, \eta)| dt,$$

and consequently,

$$\|\theta\|_{\alpha,\mathcal{X}}^2 \leq 4\|U_0^0 - u_0\|^2 + \frac{2}{c_\alpha} \int_0^{T} \|\mathcal{B}_\alpha \eta(t)\|_{\mathcal{X}} \|\eta(t)\|_{\mathcal{X}} dt.$$ (4.7)

To bound the integral term on the right-hand side of (4.7), we first recall the following inequality from [8, Lemma 6.3]: for $g \in L_2(0,T)$, there holds

$$\|\mathcal{B}_\alpha g\|^2_T \leq \omega_{\alpha+1}(T) \int_0^T \omega_\alpha(T-t) \int_0^t g^2(s) ds dt \leq \omega_{\alpha+1}(T)\|g\|^2_T$$ for $0 < \alpha < 1$. (4.8)

Then we employ the Cauchy-Schwarz inequality and inequality (4.8). This gives

$$\int_0^{T} \|\mathcal{B}_\alpha \eta(t)\|_{\mathcal{X}} \|\eta(t)\|_{\mathcal{X}} dt \leq \|\mathcal{B}_\alpha \eta\|_{L_2(\mathcal{X})} \|\eta\|_{L_2(\mathcal{X})} \leq \omega_{\alpha+1}(T)\|\eta\|_{L_2(\mathcal{X})}^2.$$ 

Combining this bound with (4.7) completes the proof. \hfill \Box

**Theorem 4.2** Let $u$ be the solution of (1.1) and $U$ be the DG solution defined by (2.4). Then we have

$$\|U - u\|_{\alpha,\mathcal{X}} \leq C \left(\|U_0^0 - u_0\| + \|\eta\|_{L_\infty(\mathcal{E})} + \omega_{\alpha+1}(T) \left(c_\alpha^{-1} + \sqrt{c_\alpha}\right) \|\eta\|_{L_2(\mathcal{X})}\right).$$
Proof. This bound follows from the decomposition of $U - u$ in (4.5), the triangle inequality, Lemma 4.1, the bound (4.8) (with $\|\mathcal{R}_{\tau} \eta(t)\|$ in place of $|\mathcal{R}_{\nu} g(t)|$), and finally the inequality $\|\eta(t)\| \leq C\|\eta(t)\|_{X}$. □

Next, let us combine Theorem 4.2 and Theorem 4.1 into the following $hp$-version error estimates. For the sake of simplicity, we assume here that $U^{0} = u_{0}$.

**Corollary 4.1** Let $U^{0} = u_{0}$. For $0 \leq q_{j} \leq p_{j}$ and $u \in H^{q_{j}+1}(I_{j};\mathbb{X})$, $1 \leq j \leq N$, we then have the error estimates

$$\|U - u\|_{\alpha;\mathbb{X}} \leq C_{\text{max}} \sum_{j=1}^{N} \left( \frac{k_{j}}{2} \right)^{2q_{j}+1} \frac{(p_{j} - q_{j})!}{(p_{j} + q_{j})!} \|u^{q_{j}+1}\|_{L_{2}(I_{j};\mathbb{X})}^{2} + C_{\omega_{\alpha+1}}(T)(c_{\alpha} + c_{\alpha}) \sum_{j=1}^{N} p_{j}^{-2} \left( \frac{k_{j}}{2} \right)^{2q_{j}+2} \frac{(p_{j} - q_{j})!}{(p_{j} + q_{j})!} \|u^{q_{j}+1}\|_{L_{2}(I_{j};\mathbb{X})}^{2}.$$  

The constant $C > 0$ is independent of $\alpha, T, q_{j}, p_{j}, k_{j}$, and $u$.

For uniform parameters $k$, $p$ and $q$ (i.e., $k_{j} = k$, $p_{j} = p$ and $q_{j} = q$), the bounds in Corollary 4.1 and the fact that $(p - q)!/(p + q)! \sim C_{q} p^{-2q}$ for $p \to \infty$ (by Stirling’s formula) result in the following estimates:

$$\|U - u\|_{\alpha;\mathbb{X}} \leq C \frac{k^{\min[p,q]+1}}{p^{q}} \left( \|u^{q+1}\|_{L_{\infty}(\Omega)} + \|u^{q+1}\|_{L_{2}(\Omega)} \right).$$

These estimates show that the time-stepping DG scheme converges either as the time-steps are decreased (i.e., $k \to 0$), or as $p$ is increased (i.e., $p \to \infty$). For a large values of $q$, we note that it is more advantageous to increase $p$ and keep $k$ fixed ($p$-version of the DG method) rather than to reduce $k$ for $p$ fixed ($h$-version of the DG method).

4.2 Consequences

We discuss the convergence rates that are obtained from our error estimates above. We consider a pure $hp$-approach based on geometrically refined time-steps and linearly increasing approximation orders, as well as an $h$-version approach on algebraically graded meshes.

4.2.1 $hp$-version on geometric time-steps. Following [23], we now consider the $hp$-version DG method for problems with solutions that have start-up singularities as $t \to 0$, but are analytic for $t > 0$.

More precisely, we stipulate that the solution $u$ has the analytic regularity:

$$\|u^{(j)}(t)\| \leq C_{0} d^{j}(j + 1)! t^{\sigma - j} \quad \forall t \in (0, T), \forall j \geq 1,$$

for positive constants $\sigma$, $C_{0}$ and $d$.

**Remark 4.3** Proving the regularity statement (4.9) remains an open issue, which is beyond the scope of the present paper. However, let us make the following comments. First, for finitely many derivatives this result has been proved in [17, 24]; see equation (4.12) below. Second, for parabolic integro-differential equations, a regularity property of the form (4.9) has been proved [23, Theorem 4.1] under suitable analyticity and smoothness assumptions on $f(t)$ and $u_{0}$, respectively. In that sense, the above analytic regularity property is natural assumption.
To resolve the singular behavior of the solution, we shall make use of geometrically refined time-steps and linearly increasing degree vectors, and apply the $hp$-techniques that were developed in [2, 23, 30]. To describe this, we first partition $(0, T)$ into (coarse) time intervals $\mathcal{I}_i = (0, T_i)$ for $i = 1, 2, \ldots, K$. The first interval $\mathcal{I}_1 = (0, T_1)$ near $t = 0$ is then further subdivided geometrically into $L + 1$ subintervals $\{I_n\}_{n=1}^{L+1}$ by using the geometric time-steps

$$t_0 = 0, \quad t_n = \delta^{L+1-n}T_1 \quad \text{for} \ 1 \leq n \leq L + 1. \quad (4.10)$$

As usual, we call $\delta \in (0, 1)$ the geometric refinement factor, and $L$ is the number of refinement levels. Let $\mathcal{M}_{L,\delta}$ be the mesh on $(0, T)$ defined in this way. The polynomial degree distribution $p$ on $\mathcal{M}_{L,\delta}$ is defined as follows. On the first coarse interval $\mathcal{I}_1$ the degrees are chosen to be linearly increasing:

$$p_n = [\mu n] \quad \text{for} \ 1 \leq n \leq L + 1, \quad (4.11)$$

for a slope parameter $\mu > 0$. On the coarse time intervals $\mathcal{I}_i = (0, T_i)$ away from $t = 0$, we set the approximation degrees uniformly to $p_{L+1} = [\mu(L+1)]$. The resulting $hp$-version finite element space is denoted by $\mathcal{W}(\mathcal{M}_{L,\delta}, \mathbf{p}; \mathbb{X})$. As before, we let $U^0 = u_0$, and then from Theorem 4.2 we have

$$\|U - u\|_{\alpha, \mathbb{X}} \leq C(\|u - \Pi u\|_{L^\infty(\mathbb{X})} + \|u - \Pi u\|_{L^2(\mathbb{X})}),$$

where $C$ only depends on $\alpha$ and $T$. Hence, it remains to estimate the interpolation bound on the right-hand side. This is now done by proceeding along the lines of [23, Theorem 4.2], see also [2, 30], taking into account the analytic regularity (4.9) and the specific construction of the sequences of $hp$-spaces. We readily obtain exponential convergence in the number of the degrees of freedom in time.

**Theorem 4.3** Let the solution $u$ of problem (1.1) satisfy the analytic regularity (4.9). Let $U \in \mathcal{W}(\mathcal{M}_{L,\delta}, \mathbf{p}; \mathbb{X})$ be the $hp$-version DG approximation with $U^0 = u_0$. Then there exists a slope $\mu_0 > 0$ depending on $\delta$ and the constants $\sigma$ and $d$ in (4.9) such that for linearly increasing polynomial degree vectors $\mathbf{p}$ with slope $\mu \geq \mu_0$ we have the error estimate

$$\|U - u\|_{\alpha, \mathbb{X}} \leq C \exp\left(-b_\gamma t^{1/4}\right),$$

with positive constants $C$ and $b$ that are independent of the number $\mathcal{N} = \dim(\mathcal{W}(\mathcal{M}_{L,\delta}, \mathbf{p}; \mathbb{X}))$, but depending on the problem parameters $T$ and $\alpha$, the regularity parameters $C_0, d$ and $\sigma$ in (4.9), and the mesh parameters $\delta, T_1$ and $\mu$.

4.2.2 **$h$-version on graded time-steps.** Next, we focus on the $h$-version DG method where the order is uniformly equal to $p$, (i.e., $p_j = p$ for all $j \geq 1$), and where $p$ is typically low. Following [24], we assume that the solution $u$ of (1.1) satisfies the finite regularity assumption:

$$\|u^{(j)}(t)\|_{\mathbb{X}} \leq C_1 t^{\sigma - j} \quad \forall 1 \leq j \leq p + 1, \quad (4.12)$$

for some positive constants $C_1$ and $\sigma$; for a proof we refer the reader to [15, 17].

To capture the singular behaviour of $u$ near $t = 0$, we employ a family of non-uniform meshes denoted by $\mathcal{M}_{T}$, where the time-steps are graded towards $t = 0$; see [17, 24]. More precisely, we assume that for a fixed parameter $\gamma \geq 1$, there holds

$$c_\gamma t^{\gamma} \leq k_1 \leq c_\gamma t^{\gamma}, \quad k_n \leq C_\gamma t^{1-1/\gamma} \quad \text{and} \quad t_n \leq C_\gamma t_{n-1} \quad \text{for} \ 2 \leq n \leq N. \quad (4.13)$$
For instance, one may choose
\[ t_n = (n/N)^T \quad \text{for } 0 \leq n \leq N. \] (4.14)

We derive next the error estimate for the \( h \)-version DG solution, giving rise to optimal rates of convergence. Indeed, these results are high-order extensions of the ones shown in [24] for \( p = 0 \) and 1.

**Theorem 4.4** Let the solution \( u \) of problem (1.1) satisfy (4.12), and assume that the time mesh satisfies assumptions (4.13). Let \( U \in \mathcal{W}(\mathcal{M}, p; \mathcal{X}) \) be the DG approximation with \( p = (p, \cdots, p) \), for \( p \geq 1 \) and \( U^0 = u_0 \). Then we have the error estimate:
\[
\|U - u\|_{\alpha, \mathcal{X}} \leq C_k \min\{\gamma^p, p+1\} \quad \text{for } \gamma \geq 1,
\]
where \( C > 0 \) is a constant that depends only on \( T, \alpha, \gamma, \sigma \) and \( p \).

**Proof.** Theorem 4.2 yields
\[
\|U - u\|_{\alpha, \mathcal{X}}^2 \leq C \|\eta\|_{L_\infty(\Omega)}^2 + C\|\eta\|_{L_2(\Omega)}^2.
\]
Using Theorem 4.1, the regularity assumption (4.12) (with \( \sigma > 1/2 \)), and (4.13), we obtain
\[
\|\eta\|_{\mathcal{L}(l_1; \mathcal{X})}^2 + \|\eta\|_{L_2(l_1; \mathcal{X})}^2 \leq C_k \|u'(t)\|_{L_2(l_1; \mathcal{X})}^2 \leq C_k \int_0^{\bar{t}} t^{2\sigma-2} \, dt \leq C_k^{2\sigma} \leq C_k^{2\gamma},
\]
and for \( n \geq 2 \), we have
\[
\|\eta\|_{\mathcal{L}(l_1; \mathcal{X})}^2 + \|\eta\|_{L_2(l_1; \mathcal{X})}^2 \leq C_k^{2p+1} \|u^{(p+1)}\|_{L_2(l_1; \mathcal{X})}^2 + C \sum_{n=2}^N k_n^{2p+2} \|u^{(p)}\|_{L_2(l_1; \mathcal{X})}^2 \leq C_k^{2p+2} \left( \max_{n=2}^N t_n^{2\sigma-2(p+1)/\gamma} + \int_0^{\bar{t}} t^{2\sigma-2(p+1)/\gamma} \, dt \right).
\]
Therefore, integrating and then inserting the obtained estimates in (4.15) completes the proof. \(\square\)

**5. Numerical examples**

We present a series of numerical tests. We show results for a scalar example, as well as for a problem in one space dimension. Our main aim will be to show numerically that the convergence behavior for the errors measured in \( L_\infty(0, T) \) is nearly identical to the theoretical predictions made in the weaker norm \( \| \cdot \|_{\alpha, \mathcal{X}} \).

**5.1 A scalar example**

In order to focus on the time discretization in isolation, we first consider the scalar problem:
\[
u' + \int_0^t \omega(t-s) u(s) \, ds = f(t), \quad t \in [0, T] \quad \text{with } T = 1.
\] (5.1)
Choosing \( u(0) = 0 \) and the source term \( f(t) = (\alpha + 1)t^\alpha \), we find that

\[
u(t) = \Gamma(\alpha + 2) \left( 1 - \sum_{p=0}^{\infty} \frac{(-t^\alpha + 1)^p}{\Gamma(1 + (\alpha + 1)p)} \right); \tag{5.2}\]

see [17, 24]. Now, to evaluate the \( L_\infty \)-norm time errors, we introduce the finer grid

\[
g N,m = \{ t_{j-1} + nk_j/m : 1 \leq j \leq N, \ 0 \leq n \leq m \}, \tag{5.3}\]

and define the quantity \( \|v\|_m = \max_{t \in g N,m} |v(t)| \). Thus, for large values of \( m \), the error measure \( \|U - u\|_m \) approximates the uniform time error \( \|U - u\|_{L_\infty(0,T)} \).

### 5.1.1 h-version on graded time-steps

We first test the performance of the h-version DG method with uniform polynomial degree \( p \) on the graded time-steps defined in (4.14). Since the exact solution (5.2) behaves like \( t^{\alpha+1} \) as \( t \to 0^+ \), we see that the regularity conditions (4.12) holds for \( \sigma = \alpha + 1 \). Thus, in agreement with Theorem 4.4 (for \( \| \cdot \|_{\alpha,X} \)-norm), we expect \( \|U - u\|_{L_\infty(0,T)} \) to converge of order \( O(k^{\min(p,\sigma,p+1)}{\gamma}) \) for \( \gamma \geq 1 \). This is demonstrated in Table 1, where we computed the errors and the experimental rates of convergence for various values of \( \gamma \) and \( \alpha \).

<table>
<thead>
<tr>
<th>( N )</th>
<th>( \gamma = 1, \ \gamma\sigma = 1.2 )</th>
<th>( \gamma = 4/3, \ \gamma\sigma = 1.6 )</th>
<th>( \gamma = 2, \ \gamma\sigma = 2.4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>3.06e-04</td>
<td>5.92e-05</td>
<td>4.42e-05</td>
</tr>
<tr>
<td>128</td>
<td>1.33e-04</td>
<td>1.200</td>
<td>1.93e-05</td>
</tr>
<tr>
<td>256</td>
<td>5.80e-05</td>
<td>1.201</td>
<td>6.33e-06</td>
</tr>
<tr>
<td>512</td>
<td>2.52e-05</td>
<td>1.201</td>
<td>2.08e-06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( N )</th>
<th>( \gamma = 1, \ \gamma\sigma = 1.2 )</th>
<th>( \gamma = 2, \ \gamma\sigma = 2.4 )</th>
<th>( \gamma = 2.5, \ \gamma\sigma = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>8.78e-05</td>
<td>1.65e-06</td>
<td>7.70e-07</td>
</tr>
<tr>
<td>64</td>
<td>3.72e-05</td>
<td>1.229</td>
<td>3.00e-07</td>
</tr>
<tr>
<td>128</td>
<td>1.60e-05</td>
<td>1.219</td>
<td>5.57e-08</td>
</tr>
<tr>
<td>256</td>
<td>6.91e-06</td>
<td>1.209</td>
<td>1.04e-08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( N )</th>
<th>( \gamma = 1, \ \gamma\sigma = 1.5 )</th>
<th>( \gamma = 2, \ \gamma\sigma = 3 )</th>
<th>( \gamma = 8/3, \ \gamma\sigma = 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1.33e-04</td>
<td>7.15e-06</td>
<td>1.09e-06</td>
</tr>
<tr>
<td>16</td>
<td>4.22e-05</td>
<td>1.650</td>
<td>7.27e-07</td>
</tr>
<tr>
<td>32</td>
<td>1.42e-05</td>
<td>1.571</td>
<td>8.24e-08</td>
</tr>
<tr>
<td>64</td>
<td>4.90e-06</td>
<td>1.535</td>
<td>9.81e-09</td>
</tr>
</tbody>
</table>

Table 1. The errors \( \|U - u\|_{2,0} \) with different mesh gradings and values of \( \alpha \) for the h-version DG method of degree \( p \). We observe convergence of order \( O(k^{\min(p,\sigma,p+1)}{\gamma}) \) for \( \gamma \geq 1 \).

### 5.1.2 hp-version on geometric time-steps

To demonstrate exponential convergence in time, we use the geometrically refined time-step and linearly increasing polynomial degrees as introduced in Section 4.2.1 for the exact solution in (5.2) with \( \alpha = 0.5 \). We choose \( T_1 = 1 \) and \( \mu = 1 \). As before, notice that the analytic regularity property (4.12) holds true with \( \sigma = \alpha + 1 \). In accordance with Theorem 4.3, we
Fig. 1. The errors $|||U - u|||_{L^{1/2}}$ plotted against $\mathcal{N}^{1/2}$ for three grading factors $\delta$ and with $\alpha = 0.5$.

Fig. 2. The errors $|||U - u|||_{L^{1/2}}$ plotted against $\mathcal{N}^{1/2}$ and $\delta$ and with $\alpha = 0.5$. We observe the best error when $\delta$ is in the neighborhood of 0.25.

We expect the uniform error to converge exponentially as well, that is, $||U - u||_{L^\infty(0,T)} \leq C\exp(-bN^{1/2})$. To calculate the coefficient $b$ in the exponent, we employ the formula:

$$
\log(\text{error}(\mathcal{N}_{L-1})/\text{error}(\mathcal{N}_{L})) / (\mathcal{N}_{L}^{1/2} - \mathcal{N}_{L-1}^{1/2}),
$$

(5.4)

where $\mathcal{N}_{L} = \dim(W(\mathcal{M}_{L,\delta}, p; X))$ and $\text{error}(\mathcal{N}_{L})$ is the corresponding error in $L^\infty(0,T)$. The numerical values of $b$ should be approximately the same for different values of geometric gradings $\mathcal{N}$. This is confirmed in Table 2 for three values of the grading factor $\delta$ close to 0.25. The results are also displayed graphically in Figure 1, where we show the errors against $\mathcal{N}^{1/2}$, denoted by "dofs $1/2$" in the plot. In the semi-logarithmic scale, the curves are roughly straight lines, which indicates exponential convergence rates. Indeed, for a fixed $\alpha$, from the 3d-plot in Figure 2 of the errors against the parameters $\mathcal{N}^{1/2}$ and $\delta$, we observe that values of $\delta$ in the neighborhood of the interval $[0.2, 0.25]$ yields the best results. To demonstrate that this remains valid for different values of $\alpha$, we present in Figure 3 the errors achieved as a function of $\delta$ for different values of $\alpha$, but for a fixed $\mathcal{N} = 35$. 
In this section, we test the $hp$-DG time-stepping scheme for the one-dimensional problem:

$$u_t(x,t) - \int_0^t \omega_\alpha(t-s) u_{xx}(x,s) \, ds = f(x,t), \quad \text{in } \Omega \times (0,T)$$  \hfill (5.5)

with $u(x,0) = u_0$ where $\Omega = (0,1)$ and $T = 1$. We impose homogeneous Dirichlet boundary conditions.

To discretize (5.5), we will employ our $hp$-DG time discretization combined with a standard continuous finite element (FE) discretization in space. To this end, we construct a family of quasi-uniform partitions of the domain $\Omega$ into subintervals with maximum step-size $h$, and let $S_h \subset H_0^1(\Omega)$ denote the space of continuous, piecewise polynomial functions of degree $\leq r$ with $r \geq 1$ and typically low. For a partition $\mathcal{M} = \{I_n\}_{n=1}^N$ of the time interval $(0,T)$ and a degree vector $p = (p_1,p_2,\cdots,p_N)$, the DG space (2.3) is now modified to the fully discrete space

$$\mathcal{W}(\mathcal{M},p;S_h) = \{ U_h : [0,T] \to S_h : U_h|_{I_n} \in \mathbb{P}_p(S_h), \ 1 \leq n \leq N \}. \hfill (5.6)$$

Here, $\mathbb{P}_p(S_h)$ is the space of polynomials of degree $\leq p$ in the time variable with coefficients in $S_h$.

We arrive at the following fully-discrte $hp$-DG FE scheme: find $U_h \in \mathcal{W}(\mathcal{M},p;S_h)$ such that

$$G_N(U_h,X) = \langle U_h^0,X_+^0 \rangle + \int_0^T \langle f(t),X(t) \rangle \, dt \quad \forall X \in \mathcal{W}(\mathcal{M},p;S_h), \hfill (5.7)$$

with $U_h(0) = R_h u(0)$, where $G_N$ is the global bilinear form defined as in (4.1), and $R_h : H_0^1(\Omega) \to S_h$ is the Ritz projection given by $A(R_h v, \chi) = A(v,\chi)$ for all $\chi \in S_h$.

### Table 2. The errors $|||U - u|||_{51}$ and the number $b$ for different choices of $\delta$ for $\alpha = 0.5$.

<table>
<thead>
<tr>
<th>$L$</th>
<th>$N_L$</th>
<th>$\delta = 0.2$</th>
<th>$\delta = 0.25$</th>
<th>$\delta = 0.3$</th>
<th>$\delta = 0.33$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>20</td>
<td>2.25e-05</td>
<td>2.39</td>
<td>4.36e-05</td>
<td>2.84</td>
</tr>
<tr>
<td>5</td>
<td>27</td>
<td>6.13e-06</td>
<td>1.79</td>
<td>5.46e-06</td>
<td>2.87</td>
</tr>
<tr>
<td>6</td>
<td>35</td>
<td>1.67e-06</td>
<td>1.81</td>
<td>6.82e-07</td>
<td>2.89</td>
</tr>
<tr>
<td>7</td>
<td>44</td>
<td>5.55e-07</td>
<td>1.54</td>
<td>1.75e-07</td>
<td>1.90</td>
</tr>
</tbody>
</table>

Fig. 3. The errors $|||U - u|||_{51}$ plotted against $\delta$ for different values of $\alpha$ and fixed $N = 35$. 
One may easily show fully discrete versions of Theorems 4.3 and 4.4, respectively, by appropriately modifying the analysis given in [23, Section 7] for parabolic integro-differential equations in conjunction with the results shown in Sections 4.2.1 and 4.2.2. More precisely, for a sufficiently regular solution \( u \), one may obtain estimates for \( ||u - U_h||_{L_2(\Omega)} \), which is bounded by a term that converges exponentially or of optimal algebraic order in time plus a term of order \( O(h^{r+1}) \) in space.

5.2.1 Numerical results. We choose the initial datum such that the exact solution of (5.5) is given by \( u(x,t) = \sin(\pi x) - t^{1+\alpha} \exp(-t) \sin(2\pi x) \). As before, it can be seen that the regularity conditions (4.9) and (4.12) hold for \( \sigma \leq \alpha + 1 \). To approximate the norm \( ||v||_{L_2(\Omega)} \), we use the quantity \( ||v||_{m} := \max_{t \in [0,T]} ||v(t)|| \). To compute it, we apply a composite Gauss quadrature rule with \( r + 1 \) points on each interval of the finest spatial mesh.

We first test the \( h \)-version scheme on the non-uniformly graded meshes \( M = M_{\gamma} \) in (4.14) for various choices of \( \gamma \geq 1 \) and for \( \alpha = 0.5 \). In space, we consider a mesh sequence consisting of \( N_{x} \) uniform subintervals, each of length \( h = 1/N_{x} \) (we have the same number of subintervals in both time and space). We thus expect to see that the global error is bounded by:

\[
||U_h - u||_{L_2(\Omega)} \leq C(h^{r+1} + k^{\min(\gamma(\alpha+1), p+1)}) = O(h^{\min(r+1, \gamma(1+\alpha), p+1)}) \quad \text{for} \quad 1 \leq \gamma \leq (p+1)/(\alpha+1).
\]

The results shown in Table 3 are in full agreement with these bounds.

<table>
<thead>
<tr>
<th>( N )</th>
<th>( \gamma = 1, \gamma \sigma = 1.5 )</th>
<th>( \gamma = 4/3, \gamma \sigma = 2 )</th>
<th>( \gamma = 2, \gamma \sigma = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>1.041e-03</td>
<td>3.212e-04</td>
<td>5.299e-04</td>
</tr>
<tr>
<td>32</td>
<td>4.082e-04</td>
<td>1.351</td>
<td>8.939e-05</td>
</tr>
<tr>
<td>64</td>
<td>1.499e-04</td>
<td>1.445</td>
<td>2.316e-05</td>
</tr>
<tr>
<td>128</td>
<td>5.385e-05</td>
<td>1.714</td>
<td>5.854e-06</td>
</tr>
</tbody>
</table>

Table 3. The errors \( ||U_h - u||_{L_2(\Omega)} \) for the \( h \)-version time-stepping DG, spatial FE method for different mesh gradings. We almost observe convergence of order \( h^{r+1} + k^{\min(\gamma(\alpha+1), p+1)} = O(h^{\min(r+1, \gamma(1+\alpha), p+1)}) \) for \( \gamma \geq 1 \).

Next, we test the performance of the \( hp \)-version time-stepping and use the geometric time partition \( M_{L, \delta} \) defined in (4.10)–(4.11), again on a uniform spatial mesh with \( N_{x} \) subintervals. As before, we take

<table>
<thead>
<tr>
<th>( L )</th>
<th>( N_{x} )</th>
<th>( r = 1 )</th>
<th>( r = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>32</td>
<td>1.147e-03</td>
<td>1.998</td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>2.870e-04</td>
<td>1.998</td>
</tr>
<tr>
<td>6</td>
<td>128</td>
<td>7.176e-05</td>
<td>2.000</td>
</tr>
<tr>
<td>7</td>
<td>256</td>
<td>1.794e-05</td>
<td>2.000</td>
</tr>
</tbody>
</table>

Table 4. The errors \( ||U_h - u||_{L_2(\Omega)} \) and the order of convergence with respect to \( N_{x} \) for \( \alpha = 0.5 \).
\[ \alpha = 0.5, \] and set \( T_1 = 1 \) and \( \mu = 1 \). The regularity assumption (4.9) holds for \( \sigma = \alpha + 1 \), and thus we expect the global error to be bounded by:

\[ \| U_h - u \|_{L^\infty (L^2)} \leq C h^{r+1} + C \exp(-b N_1^{1/2}) \text{ where } N = \dim(\mathcal{W}(\mathcal{M}_{L,\delta}, p; S_h)). \]

We approximate the \( L^\infty (L^2) \)-errors as before.

We first demonstrate the convergence order \( O(h^{r+1}) \) in space. Indeed, in Table 4 we select \( \delta = 0.3 \), and compute the errors and the numerical orders of convergence with respect to the change in the number of subintervals. For \( r = 1 \), we observe that the convergence rate is of optimal order \( O(h^2) \) and the spatial error dominates the temporal error, while for \( r = 2 \) the orders are now suboptimal due to the influence of the errors of the time discretization.

To demonstrate exponential convergence in time, we choose \( r = 2 \) and take a relatively large number of subintervals in space so that the temporal errors are dominating. Then we use the formula given by (5.4) to calculate the coefficient \( b \) in the exponential convergence bound. Again, the computed values of \( b \) should be approximately the same for different values of geometric refinements \( L \). This is illustrated tabularly in Table 5 and graphically in Figure 4, where we plot the errors against \( N^{1/2} \). In the semi-logarithmic scale, the curves are roughly straight lines, which indicates exponential convergence rates.

### References

<table>
<thead>
<tr>
<th>( L )</th>
<th>( N_1 )</th>
<th>( \delta = 0.27 )</th>
<th>( \delta = 0.3 )</th>
<th>( \delta = 0.33 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>20</td>
<td>4.92e-05</td>
<td>9.20e-05</td>
<td>1.61e-04</td>
</tr>
<tr>
<td>5</td>
<td>27</td>
<td>6.98e-06</td>
<td>1.53e-05</td>
<td>3.12e-05</td>
</tr>
<tr>
<td>6</td>
<td>35</td>
<td>9.81e-07</td>
<td>2.53e-06</td>
<td>5.96e-06</td>
</tr>
<tr>
<td>7</td>
<td>44</td>
<td>1.43e-07</td>
<td>4.17e-07</td>
<td>1.13e-06</td>
</tr>
</tbody>
</table>

Table 5. The errors \( \| U_h - u \|_{L^1} \) and the calculated exponent \( b \) for different choices of \( \delta \), with \( \alpha = 0.5, r = 2 \) and \( N_t = 200 \).