SIAM J. NUMER. ANAL.
Vol. 49, No. 4, pp. 1369–1396

AN hp-VERSION DISCONTINUOUS GALERKIN METHOD FOR INTEGRO-DIFFERENTIAL EQUATIONS OF PARABOLIC TYPE*

K. MUSTAPHA†, H. BRUNNER‡, H. MUSTAPHA§, AND D. SCHÖTZAU¶

Abstract. We study the numerical solution of a class of parabolic integro-differential equations with weakly singular kernels. We use an hp-version discontinuous Galerkin (DG) method for the discretization in time. We derive optimal hp-version error estimates and show that exponential rates of convergence can be achieved for solutions with singular (temporal) behavior near $t = 0$ caused by the weakly singular kernel. Moreover, we prove that by using nonuniformly refined time steps, optimal algebraic convergence rates can be achieved for the h-version DG method. We then combine the DG time-stepping method with a standard finite element discretization in space, and present an optimal error analysis of the resulting fully discrete scheme. Our theoretical results are numerically validated in a series of test problems.

Key words. parabolic Volterra integro-differential equation, weakly singular kernel, hp-version DG time-stepping, exponential convergence, finite element method, fully discrete scheme

AMS subject classifications. 45D05, 45J05, 65M15, 65M50, 65M60, 65M70, 65M99

DOI. 10.1137/100797114

1. Introduction.

We study the discretization in time and space of parabolic Volterra integro-differential equations of the form

$$
\begin{aligned}
u'(t) + Au(t) + \mathcal{B}Au(t) &= f(t), \qquad 0 < t < T,\\
u(0) &= u_0.
\end{aligned} \tag{1.1}
$$

Here, A is a self-adjoint linear elliptic operator and $\mathcal{B}$ is the Volterra operator given by the weakly singular kernel

$$
\mathcal{B}v(t) = \int_0^t (t-s)^{\alpha-1}\, b(s)v(s)\, ds \quad \text{for} \quad 0 < \alpha < 1, \tag{1.2}
$$

where b is a continuous function on $[0, T]$. In section 2.1, we shall set out precise technical assumptions. Problems of type (1.1) can be thought of as a model problem occurring in the theory of heat conduction in materials with memory, population dynamics, and visco-elasticity; see, for example, [6, 7, 19] and the references therein.

Over the last few decades various numerical discretization methods have been proposed and analyzed for linear and semilinear problems of the form (1.1) (including smooth and weakly singular kernels), both for semidiscrete and fully discrete schemes;

*Received by the editors June 1, 2010; accepted for publication (in revised form) April 21, 2011; published electronically July 5, 2011. This work was supported by the KFUPM through the project SB101020, and the Natural Sciences and Engineering Research Council of Canada (NSERC).

http://www.siam.org/journals/sinum/49-4/79711.html

†Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia (kassem@kfupm.edu.sa).

‡Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, NL, A1C 5S7, Canada, and Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong (hbrunner@math.hkbu.edu.hk).

§Department of Mining and Materials Engineering, McGill University, Montreal, QC, Canada (hussein.mustapha@mcgill.ca).

¶Mathematics Department, University of British Columbia, Vancouver, BC, V6T 1Z2, Canada (schoetzau@math.ubc.ca).

see, for example, [9, 12, 16, 23, 24, 27, 28] and the references therein. It is well known that the presence of a weakly singular kernel in the memory term typically leads to a temporal singularity in the solution u of (1.1) or for problems of the form $\partial u/\partial t$+memory term $= f(t)$; see, for example, [9, 11, 12, 13, 15, 16] and the references therein. In order to avoid suboptimal convergence rates of the time discretization, the lack of regularity has to be compensated by using locally refined time-steps near $t = 0$.

In this work, we shall study how to overcome this issue by means of the *hp*-version discontinuous Galerkin (DG) time-stepping method. The origins of the DG methods can be traced back to the 1970's where they were proposed as variational methods for numerically solving initial-value problems and transport problems [10, 18]; see also [3, 5, 8] and the references therein. In the 1980's, DG time-stepping methods were successfully applied to purely parabolic problems (that is, for problems of the form (1.1) without memory terms); see, for example, [4, 25] and the references therein. In these papers, only low-order and constant approximation orders have been considered, thereby giving rise to at most algebraic rates of convergence in the number of degrees of freedom (dofs) in time. Subsequently, in the recent paper [9], piecewise constant and linear DG methods in time have been proposed and studied for the Volterra integro-differential equation (1.1). The error analysis there is based on the fact that on each time interval, the DG solution takes its maximum values on one of the endpoints. However, this is not true in the case of DG methods of higher order.

The *hp*-version DG (*hp*-DG) method for the time discretization of linear parabolic problems has been introduced in [21, 26]. We also refer the reader to [20] for an analysis of this approach applied to nonlinear initial-value problems in $\mathbb{R}^d$. The main feature of the *hp*-DG method is that it allows for locally varying time-steps and approximation orders. In the above-mentioned papers, it has been shown, both theoretically and numerically, that the *hp*-DG method, based on geometrically refined time-steps and linearly increasing approximation orders, is capable of resolving temporal start-up singularities near $t = 0$ at exponential rates of convergence in the number of degrees of freedom (in time). In the recent paper [1], the *hp*-DG method has been applied to a scalar version of the model problem in (1.1). It has been proved and verified numerically that the temporal singularities near $t = 0$ induced by the weakly singular kernel (1.2) can be approximated at exponential rates of convergence.

The present paper has two purposes. First, we extend the *hp*-version analysis of [1] to the abstract parabolic problem (1.1). We introduce the *hp*-DG time-stepping method and derive optimal *hp*-version error estimates that are completely explicit in all the parameters of interest. These results imply spectral convergence for the *p*-version DG method for problems with smooth solutions. Next, as for the scalar case considered in [1], we prove that by using geometrically refined time-steps and linearly increasing approximation orders, start-up singularities near $t = 0$ can be resolved at exponential rates of convergence. Moreover, we show that the *h*-version DG method on nonuniformly refined time-steps, but with a fixed approximation order, yields optimal algebraic convergence rates.

Notice that in our analysis we will consider sufficiently regular initial data. Thus, we will be concerned only with singularities caused by the weakly singular kernel (1.2), and not by incompatible initial data, which has been the main motivation in the work [21, 26] for purely parabolic problems. We believe that our convergence results can be extended to nonsmooth initial data provided that regularity results as in Theorem 4.1 hold. How to establish this regularity remains an open question and is the subject of ongoing research.

Second, we combine the time-stepping method with standard (continuous) finite elements in space in the case where $A = -\Delta$ with homogeneous Dirichlet boundary conditions. We carry out the error analysis for the resulting fully discrete scheme and show that, for smooth solutions, we achieve spectral convergence rates in time and space.

The outline of this paper is as follows. In section 2, we introduce the *hp*-DG time-stepping method. In section 3, we derive *hp*-version error bounds that are explicit in all the parameters of interest and discuss several consequences of these estimates. Section 4 is devoted to establishing exponential rates of convergence for the *hp*-DG method on geometrically refined time-steps and linearly increasing approximation orders. In section 5, we consider the *h*-version method with a fixed approximation order on nonuniformly refined time-steps. In section 6, we proceed to consider and analyze a fully discrete scheme. In section 7, we present a series of numerical examples to validate our theoretical results. Finally, we end the paper with some concluding remarks in section 8.

2. Discontinuous Galerkin time-stepping.

In this section, we review the weak formulation of (1.1), and introduce the *hp*-DG time-stepping method.

2.1. Weak formulation.

To formulate the initial-boundary value problem (1.1) in an abstract setting, let $\mathbb{H}$ be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. We suppose that A is a linear, self-adjoint, positive-definite operator with domain $D(A) \subseteq \mathbb{H}$. We further assume that A possesses a complete orthonormal eigensystem $\{\phi_m\}_{m=1}^{\infty}$ with

$$A\phi_m = \lambda_m \phi_m \quad \text{and} \quad \langle \phi_m, \phi_{m'} \rangle = \delta_{m,m'} \quad \text{for } m,\, m' \geq 1, \tag{2.1}$$

for real eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots$. When $\mathbb{H}$ is infinite-dimensional we also require that $\lambda_m \to \infty$ as $m \to \infty$. We set $\mathbb{X} = D(A^{1/2})$, and endow it with the norm $\|v\|_{\mathbb{X}} = \|A^{1/2}v\|$. Then, we associate with A the bilinear form $A : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ defined in terms of eigenfunction expansions by the following: for $u,\, v \in \mathbb{X}$,

$$A(u, v) := \sum_{m=1}^{\infty} \lambda_m\, u_m\, v_m, \quad \text{where } u_m = \langle u, \phi_m \rangle \text{ and } v_m = \langle v, \phi_m \rangle \tag{2.2}$$

(u_m and v_m are the Fourier coefficients of u and v, respectively).

By construction, the bilinear form $A(u, v)$ is symmetric, continuous, and coercive, with continuity and coercivity constants equal to one. That is, we have

$$\begin{aligned} A(u, v) &= A(v, u) && \forall u, v \in \mathbb{X}, \\ |A(u, v)| &\leq \|u\|_{\mathbb{X}} \|v\|_{\mathbb{X}} && \forall u, v \in \mathbb{X}, \\ A(u, u) &\geq \|u\|_{\mathbb{X}}^2 && \forall u \in \mathbb{X}. \end{aligned}$$

Taking the inner product of $\mathcal{B}Au(t)$ with the function $v \in \mathbb{X}$ and using (2.2) and (2.1), we notice that

$$\begin{aligned}
\langle \mathcal{B}Au(t), v\rangle &= \Big\langle \int_0^t (t-s)^{\alpha-1}\, b(s) \sum_{m=1}^{\infty} \lambda_m u_m(s)\phi_m \, ds, \sum_{j=1}^{\infty} v_j \phi_j \Big\rangle \\
&= \sum_{m=1}^{\infty} \lambda_m \int_0^t (t-s)^{\alpha-1}\, b(s) u_m(s)\, ds \Big\langle \phi_m, \sum_{j=1}^{\infty} v_j\phi_j \Big\rangle \\
&= \sum_{m=1}^{\infty} \int_0^t (t-s)^{\alpha-1}\, b(s)\lambda_m\, u_m(s)\, v_m\, ds \\
&= \int_0^t (t-s)^{\alpha-1}\, b(s) \sum_{m=1}^{\infty} \lambda_m\, u_m(s)\, v_m\, ds = \int_0^t (t-s)^{\alpha-1}\, b(s) A(u(s), v)\, ds\,.
\end{aligned}$$

Thus, the weak formulation of the abstract parabolic problem (1.1) now consists in finding $u(t)$ such that $u(0) = u_0$ and for $t \in (0, T)$,

$$\langle u'(t), v\rangle + A(u(t), v) + \mathcal{B}\left[A(u(\cdot), v)\right](t) = \langle f(t), v\rangle \quad \forall v \in X, \tag{2.3}$$

where

$$\mathcal{B}\left[A(u(\cdot), v)\right](t) = \int_0^t (t-s)^{\alpha-1}\, b(s) A(u(s), v)\, dt\,.$$

Following the derivation given in [2, Theorem 1], we observe that the variational problem (2.3) has a unique solution $u \in C([0,T]; D(A))$ and $u' \in C([0,T]; \mathbb{H})$, provided that $f \in H^1(0,T;\mathbb{H})$ and $u_0 \in D(A)$. Since we will restrict our analysis to smooth initial data, this regularity property is sufficient for our purpose.

Remark 2.1. As the standard example of a problem of the form (1.1), one may take $A = -\Delta$, subject to homogeneous Dirichlet boundary conditions, on a bounded and convex Lipschitz domain in $\mathbb{R}^d$, $d \geq 1$. In this case, we have $\mathbb{H} = L_2(\Omega)$, $D(A) = H^2(\Omega) \cap H_0^1(\Omega)$, $D(A^{1/2}) = H_0^1(\Omega)$, and $\|u\|_{\mathbb{X}} = \|\nabla u\|_{L_2(\Omega)}$. The bilinear form $A(u,v)$ is given by $A(u,v) = \int_\Omega \nabla u \cdot \nabla v\, dx$. By the standard Poincaré inequality, the norm $\|u\|_{\mathbb{X}}$ is equivalent to full H^1-norm $\|u\|_{H^1(\Omega)}$.

2.2. Time discretization. To describe the *hp*-DG method, we introduce a (possibly nonuniform) partition $\mathcal{M}$ of the time interval $[0,T]$ given by the points

$$0 = t_0 < t_1 < \cdots < t_N = T. \tag{2.4}$$

We set $I_n = (t_{n-1}, t_n]$ and $k_n = t_n - t_{n-1}$ for $1 \leq n \leq N$. The maximum step-size is defined as $k = \max_{1\leq n\leq N} k_n$. 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

$$\mathbf{p} := (p_1, p_2, \ldots, p_N). \tag{2.5}$$

We now introduce the discontinuous finite element space

$$\mathcal{W}(\mathcal{M}, \mathbf{p}) = \left\{ v : [0,T] \to \mathbb{X} \,:\, v|_{I_n} \in \mathbb{P}_{p_n},\ 1 \leq n \leq N \right\}, \tag{2.6}$$

where $\mathbb{P}_{p_n}$ denotes the space of polynomials of degree $\leq p_n$ with coefficients in $\mathbb{X}$. We follow the usual convention that a function $v \in \mathcal{W}(\mathcal{M}, \mathbf{p})$ is left-continuous at each time level t_n, writing

$$v^n = v(t_n) = v(t_n^-), \qquad v_+^n = v(t_n^+), \qquad [v]^n = v_+^n - v^n\,.$$

The hp-DG approximation $U \in \mathcal{W}(\mathcal{M}, \mathbf{p})$ is now obtained as follows: Given $U(t)$ for $0 \le t \le t_{n-1}$, the approximation $U \in \mathbb{P}_{p_n}$ on the next time-step I_n is determined by requesting that

$$\begin{aligned} \langle U_+^{n-1}, X_+^{n-1}\rangle + \int_{t_{n-1}}^{t_n} \Big[\langle U', X\rangle + A(U,X) + \mathcal{B}\left[A(U(\cdot),X)\right]\Big]\, dt \\ = \langle U^{n-1}, X_+^{n-1}\rangle + \int_{t_{n-1}}^{t_n} \langle f, X\rangle\, dt \end{aligned} \tag{2.7}$$

for all test functions $X \in \mathbb{P}_{p_n}$. This time-stepping procedure starts from a suitable approximation U^0 to u_0, and after N steps it yields the approximate solution $U \in \mathcal{W}(\mathcal{M}, \mathbf{p})$ for $0 \le t \le t_N$.

Remark 2.2. Using the eigenspaces of A on each subinterval I_n, problem (2.7) can be reduced to a linear system of $(p_n + 1) \times (p_n + 1)$ equations. Because of the finite dimensionality of this system, the existence of the DG solution U follows from it uniqueness. To this end, if U_1 and U_2 are two DG solutions of (1.1) that satisfy (2.7) on I_n, then from (3.10) we observe that $G_n(\theta, X) = 0$, where $\theta = U_1 - U_2$ on I_n and zero on $(0, t_{n-1}]$. Hence, for k sufficiently small (see condition (3.8)), an application of Lemma 3.9 yields that $U_1 - U_2 = 0$. Thus the DG solution U defined by (2.7) is uniquely solvable for k sufficiently small.

3. Error analysis.

This section is devoted to deriving error estimates for the hp-DG method. Our main results are error estimates that are explicit in all parameters of interest. They imply that the DG method yields spectral accuracy for smooth solutions and exponential rates of convergence for analytic solutions. Our analysis relies on the techniques introduced in [20, 21] for initial-value ODEs and parabolic problems.

3.1. Global formulation and Galerkin orthogonality.

For our error analysis, it will be convenient to reformulate the DG scheme (2.7) in terms of the global bilinear form

$$\begin{aligned} G_N(U,X) = \langle U_+^0, X_+^0\rangle + \sum_{n=1}^{N-1} \langle [U]^n, X_+^n\rangle \\ + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big[\langle U', X\rangle + A(U,X) + \mathcal{B}\left[A(U(\cdot),X)\right]\Big]\, dt. \end{aligned} \tag{3.1}$$

By summing up (2.7) over all the time-steps, the DG method can now, equivalently, be written as follows: Find $U \in \mathcal{W}(\mathcal{M}, \mathbf{p})$ such that

$$G_N(U,X) = \langle U^0, X_+^0\rangle + \int_0^{t_N} \langle f, X\rangle\, dt \quad \forall X \in \mathcal{W}(\mathcal{M}, \mathbf{p}). \tag{3.2}$$

Remark 3.1. Integration by parts yields the following alternative expression for the bilinear form G_N in (3.1):

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

Since the solution u is continuous with values in $\mathbb{X}$, it follows that

$$G_N(u,X) = \langle u_0, X_+^0\rangle + \int_0^{t_N} \langle f, X\rangle\,dt.$$

Thus, the following Galerkin orthogonality property holds:

$$G_N(U-u,X) = \langle U^0 - u_0, X_+^0\rangle \quad \forall X \in \mathcal{W}(\mathcal{M},\mathbf{p}); \tag{3.3}$$

see also [21, Proposition 2.6].

3.2. An *hp*-version projection operator. We introduce a projection operator that has been used various times in the analysis of DG time-stepping methods; see [25]. In our Hilbert space setting, it is given as follows. For a continuous function $\widehat{u} : [-1,1] \to \mathbb{X}$, we define $\widehat{\Pi}^p\widehat{u} : [-1,1] \to \mathbb{P}_p$ by

$$\widehat{\Pi}^p\widehat{u}(1) = \widehat{u}(1) \in \mathbb{X} \quad \text{and} \quad \int_{-1}^{1} \langle \widehat{u} - \widehat{\Pi}^p\widehat{u}, v\rangle\,dt = 0 \quad \forall v \in \mathbb{P}_{p-1}. \tag{3.4}$$

Note that for $p = 0$, the second conditions are not required. From [21, Lemma 3.2] it follows that $\widehat{\Pi}^p$ is well defined.

For any continuous function $u : [0,T] \to \mathbb{X}$ we now define the piecewise *hp*-interpolant $\Pi u : [0,T] \to \mathcal{W}(\mathcal{M},\mathbf{p})$ by setting

$$(\Pi u)|_{I_n} = \widehat{\Pi}^{p_n}(u \circ F_n) \circ F_n^{-1}, \qquad 1 \le n \le N, \tag{3.5}$$

where $F_n : [-1,1] \to \overline{I}_n$ is the affine mapping given by $F_n(\hat{t}) = (k_n\hat{t} + t_n + t_{n-1})/2$. To state the *hp*-version approximation properties of Π, we set

$$\Gamma_{p,q} = \frac{\Gamma(p+1-q)}{\Gamma(p+1+q)}, \tag{3.6}$$

and further introduce the notation

$$\|\phi\|_{I_n} = \sup_{t\in I_n} \|\phi(t)\|.$$

Then, the following result holds true.

THEOREM 3.2. *For $1 \le n \le N$, let u be in $C([t_{n-1},t_n];\mathbb{X})$. Then we have the following:*

(i) *If u is on $[t_{n-1},t_n]$ analytic with values in $\mathbb{X}$, there holds*

$$\int_{t_{n-1}}^{t_n} \|\Pi u - u\|_{\mathbb{X}}^2\,dt + k_n\|\Pi u - u\|_{I_n}^2 \le Ck_n\exp(-\tilde{b}\,p_n).$$

(ii) *For any $0 \le q_n \le p_n$ and $u|_{I_n} \in H^{q_n+1}(I_n;\mathbb{X})$, there holds*

$$\int_{t_{n-1}}^{t_n} \|\Pi u - u\|_{\mathbb{X}}^2\,dt \le \frac{C}{\max\{1,p_n^2\}}\left(\frac{k_n}{2}\right)^{2q_n+2} \Gamma_{p_n,q_n} \int_{t_{n-1}}^{t_n} \|u^{(q_n+1)}\|_{\mathbb{X}}^2\,dt.$$

(iii) *For any $0 \le q_n \le p_n$ and $u|_{I_n} \in H^{q_n+1}(I_n;\mathbb{H})$, there holds*

$$\|\Pi u - u\|_{I_n}^2 \le C\left(\frac{k_n}{2}\right)^{2q_n+1} \Gamma_{p_n,q_n} \int_{t_{n-1}}^{t_n} \|u^{(q_n+1)}\|^2\,dt,$$

where the constants C and $\tilde{b}$ are independent of k_n and p_n.

Proof. The first and second bounds have been established in [21, section 3]. The third bound has been shown in [20, Theorem 3.9 and Corollary 3.10] for functions with values in $\mathbb{R}^d$. A careful inspection of the proofs there shows that it also holds for functions with values in the Hilbert space $\mathbb{H}$. ∎

Remark 3.3. Due to the continuous embedding of $\mathbb{X}$ in $\mathbb{H}$, we also have

$$\|\Pi u - u\|_{I_n}^2 \le C \left(\frac{k_n}{2}\right)^{2q_n+1} \Gamma_{p_n,q_n} \int_{t_{n-1}}^{t_n} \|u^{(q_n+1)}\|_{\mathbb{X}}^2 \, dt, \tag{3.7}$$

provided that $u|_{I_n} \in H^{q_n+1}(I_n; \mathbb{X})$.

To derive error estimates in the norm $\|\cdot\|_{I_n}$, we shall make use of the following inverse estimate from [20, Lemma 3.1]. While it has been proved for functions with values in $\mathbb{R}^d$ there, it can be readily seen that the same result also holds true for functions with values in $\mathbb{H}$.

LEMMA 3.4. *Let $\phi \in \mathcal{W}(\mathcal{M}, \mathbf{p})$. Then for $1 \le n \le N$, we have*

$$\|\phi\|_{I_n}^2 \le C \left(\log(p_n + 2) \int_{t_{n-1}}^{t_n} \|\phi'\|^2 (t - t_{n-1}) \, dt + \|\phi^n\|^2 \right).$$

3.3. Error bounds. We begin by stating two technical lemmas that are needed for the subsequent derivation of the error estimates. The first lemma has been proved in [9, Lemma 6.3].

LEMMA 3.5. *If $g \in L_2(0,T)$ and $\alpha \in (0,1)$, then*

$$\int_0^T \left(\int_0^t (t-s)^{\alpha-1} g(s) \, ds \right)^2 dt \le \frac{T^\alpha}{\alpha} \int_0^T (T-t)^{\alpha-1} \int_0^t g^2(s) \, ds \, dt.$$

We shall need the discrete Gronwall inequality from [9, Lemma 6.4].

LEMMA 3.6. *Let $\{a_j\}_{j=1}^N$ and $\{b_j\}_{j=1}^N$ be sequences of nonnegative numbers with $0 \le b_1 \le b_2 \le \cdots \le b_N$. Assume that there exists a constant $K \ge 0$ such that*

$$a_n \le b_n + K \sum_{j=1}^n a_j \int_{t_{j-1}}^{t_j} (t_n - t)^{\alpha-1} \, dt \quad \text{for} \quad 1 \le n \le N \quad \text{and} \quad \alpha \in (0,1).$$

Assume further that $\kappa = \frac{K\, k^\alpha}{\alpha} < 1$. Then for $n = 1, \ldots, N$, we have $a_n \le C b_n$, where C is a constant that depends on K, T, α, and κ.

Throughout the rest of this paper, we shall always implicitly assume that the *maximum step-size k is sufficiently small* so that the condition $\kappa < 1$ in Lemma 3.6 is satisfied. More precisely, we shall require that

$$\frac{3}{4} \frac{T^\alpha}{\alpha^2} k^\alpha < 1; \tag{3.8}$$

see Lemma 3.7. Let us point out the fact that this condition is independent of the polynomial degrees p_n.

We are now ready to derive our error estimates. Let u be the solution of (1.1), and let U be the DG approximation defined in (3.2). We assume that $u : [0,T] \to \mathbb{X}$ is continuous. To bound the error $U - u$, we decompose it into two terms:

$$U - u = (U - \Pi u) + (\Pi u - u) =: \theta + \eta, \tag{3.9}$$

where Π is the hp-version interpolation operator in (3.5). Theorem 3.2 can be used to bound η and the main task now reduces to estimate the first term $\theta \in \mathcal{W}(\mathcal{M}, \mathbf{p})$. The Galerkin orthogonality relation (3.3) implies that

$$G_N(\theta, X) = \langle U^0 - u_0, X_+^0 \rangle - G_N(\eta, X) \quad \forall X \in \mathcal{W}(\mathcal{M}, \mathbf{p}).$$

By construction of the interpolant Π we have that $\eta^n = 0$ for all $1 \le n \le N$. Hence, using the alternative expression for G_N in Remark 3.1 yields that

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

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

(3.10)
$$G_N(\theta, X) = \langle U^0 - u_0, X_+^0 \rangle - \int_0^{t_N} \Big[A(\eta, X) + \mathcal{B}\left[A(\eta(\cdot), X)\right] \Big] \, dt \quad \forall X \in \mathcal{W}(\mathcal{M}, \mathbf{p}).$$

First, we show the following bound.

LEMMA 3.7. *For $1 \le n \le N$, we have*

$$\|\theta^n\|^2 + \int_0^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt \le C \left(\|U^0 - u_0\|^2 + \int_0^{t_n} \|\eta\|_{\mathbb{X}}^2 \, dt \right).$$

Proof. By choosing $X = \theta$ in (3.10), then using the alternative definition of G_N in Remark 3.1 and the fact that $\langle \theta', \theta \rangle = (d/dt)\|\theta\|^2/2$, we observe that

$$\|\theta^n\|^2 + \|\theta_+^0\|^2 + \sum_{j=1}^{n-1} \|[\theta]^j\|^2 + 2\int_0^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt = 2\langle U^0 - u_0, \theta_+^0 \rangle$$
$$- 2\int_0^{t_n} \Big[A(\eta, \theta) + \mathcal{B}\left[A(\eta(\cdot), \theta)\right] + \mathcal{B}\left[A(\theta(\cdot), \theta)\right] \Big] \, dt.$$

Due to the inequality

$$2\langle U^0 - u_0, \theta_+^0 \rangle \le \|U^0 - u_0\|^2 + \|\theta_+^0\|^2,$$

we obtain

$$\|\theta^n\|^2 + 2\int_0^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt \le \|U^0 - u_0\|^2 + 2\left(|Q_1^n| + |Q_2^n| + |Q_3^n|\right), \tag{3.11}$$

where

$$Q_1^n = \int_0^{t_n} A(\eta, \theta) \, dt, \quad Q_2^n = \int_0^{t_n} \mathcal{B}\left[A(\eta(\cdot), \theta)\right] \, dt, \quad \text{and} \quad Q_3^n = \int_0^{t_n} \mathcal{B}\left[A(\theta(\cdot), \theta)\right] \, dt.$$

To bound $|Q_1^n|$, we use the geometric-arithmetic mean inequality $|ab| \le \frac{\varepsilon a^2}{2} + \frac{b^2}{2\varepsilon}$, valid for any $\varepsilon > 0$. We find that

$$|Q_1^n| \le \int_0^{t_n} \|\eta\|_{\mathbb{X}} \|\theta\|_{\mathbb{X}} \, dt \le \frac{3}{4} \int_0^{t_n} \|\eta\|_{\mathbb{X}}^2 \, dt + \frac{1}{3} \int_0^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt.$$

To estimate $|Q_2^n|$, we employ the Cauchy–Schwarz inequality, again the geometric-arithmetic mean inequality, and Lemma 3.5 (with $T = t_n$):

$$\begin{aligned}
|Q_2^n| &\le \int_0^{t_n} \int_0^t (t-s)^{\alpha-1} \|\eta(s)\|_{\mathbb{X}} \|\theta(t)\|_{\mathbb{X}} \, ds \, dt \\
&\le \left(\int_0^{t_n} \left(\int_0^t (t-s)^{\alpha-1} \|\eta(s)\|_{\mathbb{X}} \, ds \right)^2 dt \right)^{1/2} \left(\int_0^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt \right)^{1/2} \\
&\le \frac{3}{4} \int_0^{t_n} \left(\int_0^t (t-s)^{\alpha-1} \|\eta(s)\|_{\mathbb{X}} \, ds \right)^2 dt + \frac{1}{3} \int_0^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt \\
&\le \frac{3t_n^\alpha}{4\alpha} \int_0^{t_n} (t_n - t)^{\alpha-1} \int_0^t \|\eta(s)\|_{\mathbb{X}}^2 \, ds \, dt + \frac{1}{3} \int_0^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt \\
&\le \frac{3t_n^{2\alpha}}{4\alpha^2} \int_0^{t_n} \|\eta\|_{\mathbb{X}}^2 \, dt + \frac{1}{3} \int_0^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt.
\end{aligned}$$

Similarly, we notice that

$$|Q_3^n| \le \frac{3t_n^\alpha}{4\alpha} \int_0^{t_n} (t_n - t)^{\alpha-1} \int_0^t \|\theta(s)\|_{\mathbb{X}}^2 \, ds \, dt + \frac{1}{3} \int_0^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt.$$

Inserting the above bounds for $|Q_1^n|$, $|Q_2^n|$, and $|Q_3^n|$ in (3.11) implies that

$$\begin{aligned}
\|\theta^n\|^2 + \int_0^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt \le \|U^0 - u_0\|^2 &+ \frac{3}{4} \left(\frac{T^{2\alpha}}{\alpha^2} + 1 \right) \int_0^{t_n} \|\eta\|_{\mathbb{X}}^2 \, dt \\
&+ \frac{3T^\alpha}{4\alpha} \sum_{j=1}^n \int_{t_{j-1}}^{t_j} (t_n - t)^{\alpha-1} \, dt \int_0^{t_j} \|\theta\|_{\mathbb{X}}^2 \, dt \, .
\end{aligned}$$

Thus, an application of the Gronwall inequality in Lemma 3.6 completes the proof. ∎

Next, we prove the subsequent bound.

LEMMA 3.8. *For $1 \le n \le N$, we have*

$$\int_{t_{n-1}}^{t_n} \|\theta'\|^2 (t - t_{n-1}) \, dt \le C p_n^2 \left(\|U^0 - u_0\|^2 + \int_0^{t_n} \|\eta\|_{\mathbb{X}}^2 \, dt \right).$$

Proof. We choose $X = (t - t_{n-1})\theta' \in \mathbb{P}_{p_n}$ on I_n and zero elsewhere in (3.10), and refer to the definition of G_N given by (3.1) to obtain

$$\begin{aligned}
\int_{t_{n-1}}^{t_n} &\left[\|\theta'\|^2 (t - t_{n-1}) + A(\theta, (t - t_{n-1})\theta') + \mathcal{B}\left[A(\theta(\cdot), (t - t_{n-1})\theta') \right] \right] dt \\
&= -\int_{t_{n-1}}^{t_n} \left[A(\eta, (t - t_{n-1})\theta') + \mathcal{B}\left[A(\eta(\cdot), (t - t_{n-1})\theta') \right] \right] dt.
\end{aligned}$$

Simple manipulations show that

$$\begin{aligned}
\int_{t_{n-1}}^{t_n} A(\theta, (t - t_{n-1})\theta') \, dt &= \frac{1}{2} \int_{t_{n-1}}^{t_n} (t - t_{n-1}) \frac{d}{dt} A(\theta, \theta) \, dt \\
&= \frac{k_n}{2} \|\theta^n\|_{\mathbb{X}}^2 - \frac{1}{2} \int_{t_{n-1}}^{t_n} \|\theta\|_{\mathbb{X}}^2 \, dt.
\end{aligned}$$

Hence,

$$\int_{t_{n-1}}^{t_n} \|\theta'\|^2(t-t_{n-1})\,dt \le \frac{1}{2}\int_{t_{n-1}}^{t_n} \|\theta\|_{\mathbb{X}}^2\,dt + |Q_4^n| + |Q_5^n| + |Q_6^n|, \tag{3.12}$$

where

$$Q_4^n = \int_{t_{n-1}}^{t_n} A(\eta,(t-t_{n-1})\theta')\,dt, \quad Q_5^n = \int_{t_{n-1}}^{t_n} \mathcal{B}\left[A(\eta(\cdot),(t-t_{n-1})\theta')\right]\,dt,$$
$$Q_6^n = \int_{t_{n-1}}^{t_n} \mathcal{B}\left[A(\theta(\cdot),(t-t_{n-1})\theta')\right]\,dt.$$

To bound the term $|Q_4^n|$, we use the geometric-arithmetic mean inequality and a standard inverse inequality to obtain

$$\begin{aligned}
|Q_4^n| &\le \int_{t_{n-1}}^{t_n} \|\eta\|_{\mathbb{X}}\,\|\theta'\|_{\mathbb{X}}(t-t_{n-1})\,dt \\
&\le \frac{k_n^2p_n^{-2}}{2}\int_{t_{n-1}}^{t_n} \|\theta'\|_{\mathbb{X}}^2\,dt + \frac{p_n^2}{2}\int_{t_{n-1}}^{t_n} \|\eta\|_{\mathbb{X}}^2\,dt \\
&\le \frac{p_n^2}{2}\int_{t_{n-1}}^{t_n} \|\theta\|_{\mathbb{X}}^2\,dt + \frac{p_n^2}{2}\int_{t_{n-1}}^{t_n} \|\eta\|_{\mathbb{X}}^2\,dt.
\end{aligned}$$

To bound $|Q_5^n|$ we use Lemma 3.5 (with $T = t_n$), the standard inverse inequality, and proceed as follows:

$$\begin{aligned}
|Q_5^n| &\le \int_{t_{n-1}}^{t_n}\int_0^t (t-s)^{\alpha-1}\|\eta(s)\|_{\mathbb{X}}\,ds\,k_n\|\theta'(t)\|_{\mathbb{X}}\,dt \\
&\le \left(\int_0^{t_n}\left(\int_0^t (t-s)^{\alpha-1}\|\eta(s)\|_{\mathbb{X}}\,ds\right)^2 dt\right)^{1/2}\left(k_n^2\int_{t_{n-1}}^{t_n}\|\theta'\|_{\mathbb{X}}^2\,dt\right)^{1/2} \\
&\le \left(\frac{t_n^{2\alpha}}{\alpha^2}\int_0^{t_n}\|\eta(t)\|_{\mathbb{X}}^2\,dt\right)^{1/2}\left(p_n^4\int_{t_{n-1}}^{t_n}\|\theta\|_{\mathbb{X}}^2\,dt\right)^{1/2} \\
&\le \frac{T^{2\alpha}p_n^2}{\alpha^2}\int_0^{t_n}\|\eta(s)\|_{\mathbb{X}}^2\,ds + \frac{p_n^2}{2}\int_{t_{n-1}}^{t_n}\|\theta\|_{\mathbb{X}}^2\,dt.
\end{aligned}$$

Similarly,

$$|Q_6^n| \le \left(\frac{t_n^{2\alpha}}{\alpha^2}\int_0^{t_n}\|\theta\|_{\mathbb{X}}^2\,dt\right)^{1/2}\left(p_n^4\int_{t_{n-1}}^{t_n}\|\theta\|_{\mathbb{X}}^2\,dt\right)^{1/2} \le \frac{T^\alpha p_n^2}{\alpha}\int_0^{t_n}\|\theta\|_{\mathbb{X}}^2\,dt.$$

Using the obtained bounds of $|Q_4^n|$, $|Q_5^n|$, and $|Q_6^n|$ in (3.12), we get

$$\int_{t_{n-1}}^{t_n}\|\theta'\|^2(t-t_{n-1})\,dt \le Cp_n^2\int_0^{t_n}\|\eta\|_{\mathbb{X}}^2\,dt + Cp_n^2\int_0^{t_n}\|\theta\|_{\mathbb{X}}^2\,dt,$$

and hence, by Lemma 3.7 we complete the proof. ∎

In the following, we introduce the norms

$$\|\phi\|_{J_n} = \sup_{t\in(0,t_n]} \|\phi(t)\| \qquad \text{and} \qquad \|\phi\|_J = \sup_{t\in(0,T]} \|\phi(t)\|, \tag{3.13}$$

and define

$$|\mathbf{p}|_n := \max\left\{ \max_{j=1}^{n} p_j, 1 \right\}.$$

We are now ready establish the following bound for $\theta = U - \Pi u$.

LEMMA 3.9. *For $1 \le n \le N$, we have*

$$\|\theta\|_{J_n}^2 \le C \log(|\mathbf{p}|_n + 2)|\mathbf{p}|_n^2 \left(\|U^0 - u_0\|^2 + \int_0^{t_n} \|\eta\|_{\mathbb{X}}^2 \, dt \right).$$

Proof. From the inverse inequality in Lemma 3.4 and the results of Lemmas 3.7 and 3.8, we obtain, for $1 \le j \le n \le N$,

$$\begin{aligned} \|\theta\|_{I_j}^2 &\le C \left(\log(p_j + 2) \int_{t_{j-1}}^{t_j} \|\theta'\|^2 (t - t_{j-1}) \, dt + \|\theta^j\|^2 \right) \\ &\le C \log(p_j + 2) \left(p_j^2 \|U^0 - u_0\|^2 + p_j^2 \int_0^{t_j} \|\eta\|_{\mathbb{X}}^2 \, dt \right) \\ &\le C \log(|\mathbf{p}|_n + 2)|\mathbf{p}|_n^2 \left(\|U^0 - u_0\|^2 + \int_0^{t_n} \|\eta\|_{\mathbb{X}}^2 \, dt \right). \end{aligned}$$

Since the right-hand side in the bound above is independent of the time level j, the desired estimate follows. ∎

The following abstract error bounds in $L^2(0, t_n; \mathbb{X})$ and $C([0, t_n]; \mathbb{H})$ present our first main result.

THEOREM 3.10. *Let u be the solution of* (1.1)*, and let U be the DG solution defined by* (2.7)*. Then we have the error estimates*

$$\int_0^{t_n} \|U - u\|_{\mathbb{X}}^2 \, dt + \|(U - u)^n\|^2 \le C \left(\|U^0 - u_0\|^2 + \int_0^{t_n} \|u - \Pi u\|_{\mathbb{X}}^2 \, dt \right)$$

and

$$\|U - u\|_{J_n}^2 \le C\|u - \Pi u\|_{J_n}^2 + C \log(|\mathbf{p}|_n + 2)|\mathbf{p}|_n^2 \left(\|U^0 - u_0\|^2 + \int_0^{t_n} \|u - \Pi u\|_{\mathbb{X}}^2 \, dt \right).$$

Proof. To prove the first bound, we start from the decomposition of $U - u$ in (3.9), then employ the triangle inequality, Lemma 3.7, and the fact that $\eta^n = 0$ for $1 \le n \le N$. The second bound follows similarly using the result of Lemma 3.9. ∎

Let us now combine Theorems 3.10 and 3.2 to obtain hp-version error estimates that are completely explicit in the step-sizes k_j, the polynomial degree p_j, and the regularity parameters q_j.

COROLLARY 3.11. *For $1 \le n \le N$, $0 \le q_j \le p_j$, and $u \in H^{q_j+1}(I_j; \mathbb{X})$, we have the error estimates*

$$\int_0^{t_n} \|U - u\|_{\mathbb{X}}^2 \, dt + \|(U - u)^n\|^2 \le C \sum_{j=1}^{n} \hat{p}_j^{-2} \left(\frac{k_j}{2} \right)^{2q_j+2} \Gamma_{p_j,q_j} \int_{t_{j-1}}^{t_j} \|u^{(q_j+1)}\|_{\mathbb{X}}^2 \, dt$$

and

$$\|U-u\|_{J_n}^2 \le C \max_{j=1}^{n} \left(\frac{k_j}{2}\right)^{2q_j+1} \Gamma_{p_j,q_j} \int_{t_{j-1}}^{t_j} \|u^{(q_j+1)}\|_X^2\, dt$$
$$+C\log(|\mathbf{p}|_n+2)|\mathbf{p}|_n^2 \sum_{j=1}^{n} \hat{p}_j^{-2} \left(\frac{k_j}{2}\right)^{2q_j+2} \Gamma_{p_j,q_j} \int_{t_{j-1}}^{t_j} \|u^{(q_j+1)}\|_{\mathbb{X}}^2\, dt,$$

where we define $\hat{p}_j := \max\{1, p_j\}$.

Proof. These bounds follow immediately from Theorem 3.10 and the approximation properties in Theorem 3.2. For the second bound, we have also used (3.7). ∎

For uniform parameters k, p, and q (i.e., $k_j = k$, $p_j = p$, and $q_j = q$), the bounds in Corollary 3.11 result in the following error estimates.

COROLLARY 3.12. *For* $1 \le n \le N$, $0 \le q \le p$, *and* $u \in H^{q+1}(0, t_n; \mathbb{X})$, *we have the error bounds*

$$\int_0^{t_n} \|U-u\|_{\mathbb{X}}^2\, dt + \|(U-u)^n\|^2 \le C\, \frac{k^{2\min\{p,q\}+2}}{p^{2q+2}} \int_0^{t_n} \|u^{(q+1)}\|_{\mathbb{X}}^2\, dt$$

and

$$\|U-u\|_{J_n}^2 \le C\, \frac{k^{2\min\{p,q\}+2}}{p^{2q}} \left(\max_{j=1}^{n} \max_{t\in I_j} \|u^{(q+1)}(t)\|_{\mathbb{X}} + \log(p+2) \int_0^{t_n} \|u^{(q+1)}\|_{\mathbb{X}}^2\, dt \right).$$

Proof. This follows from Corollary 3.11 and the fact that $\Gamma_{p,q} \sim p^{-2q}$ for $p \to \infty$, which is a consequence of Stirling's formula or Jordan's lemma [17, 22]. ∎

The estimates in Corollary 3.12 show that the DG time-stepping scheme converges either as the time-steps are decreased (i.e., $k \to 0$,) or as p is increased (i.e., $p \to \infty$). We observe that the first estimate is optimal in both k and p, while the second one falls short by one power from being optimal in p. For a large 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). For a smooth solution u, arbitrarily high order convergence rates are possible if the polynomials degree p is increased. This is referred to as spectral convergence. In fact, if u is analytic on $[0, t_n]$ with values in $\mathbb{X}$, we obtain exponential rates of convergence for the p-version (with fixed step-size k):

$$\int_0^{t_n} \|U-u\|_{\mathbb{X}}^2\, dt + \|U-u\|_{J_n}^2 \le C\exp(-\tilde{b}\, p), \tag{3.14}$$

which follows readily from the first approximation result in Theorem 3.2.

4. Exponential convergence. Next, we consider the *hp*-DG method for solutions that have start-up singularities at time $t = 0$, but are analytic for $t > 0$. In our (regularity) analysis, we will restrict ourselves to smooth initial data. Thus, we will only be concerned with singularities caused by the weakly singular kernel (1.2) and not by incompatible initial data. We believe that our exponential convergence results in Theorem 4.2 can be extended to nonsmooth initial data provided that analytic regularity results as in Theorem 4.1 hold. How to establish this regularity remains an open question and is the subject of ongoing research.

Let $\mathcal{A}(0,T;\mathbb{H})$ denote the space of the functions which are analytic on $[0,T]$ with values in $\mathbb{H}$. Thus, a function $g \in \mathcal{A}(0,T;\mathbb{H})$ can be characterized by the analyticity constants C_g and d_g such that

$$\|g^{(j)}(t)\|_{\mathbb{H}} \le C_g d_g^j \Gamma(j+1) \quad \text{for} \quad t \in [0,T] \quad \text{with} \quad j \ge 0.$$

Next, we state the following regularity properties of the solution u of (1.1) where a brief sketch of the proof will be provided. Full details can be found in [14].

THEOREM 4.1. *Assume that $f(t) = f_1(t) + t^\rho f_2(t)$, where $\rho \in \mathbb{R}^+\mathbb{N}$. Let b be real-analytic, and assume that $f_1, f_2 \in \mathcal{A}(0,T;D(A^{3/2}))$ and that $u_0 \in D(A^{3/2})$. Then there exist constants C_0 and d depending on $\|Au_0\|_{\mathbb{X}}$ and the analyticity constants of b, f_1, and f_2 such that*

$$\|u^{(j)}(t)\|_{\mathbb{X}} \le C_0 d^j \Gamma(j+2) t^{\sigma-j} \quad \textit{for} \quad t \in (0,T] \quad \textit{and} \quad j \ge 1, \tag{4.1}$$

where $\sigma \ge 1$ with $\sigma := \min\{\alpha, \rho\} + 1$ for $j \ge 2$.

Proof. For the sake of simplicity, we restrict ourselves to the case $b(s) = 1$ and for convenience, we introduce the following notation: Given a function v defined on $[0,T]$, we set $F_0 v(t) := v(t)$ and for $j \ge 1$,

$$\begin{aligned} F_j v(t) := (tF_{j-1}v(t))' &= v + (2^j - 1)t\, v' + \sum_{\ell=2}^{j-1} t^\ell v^{(\ell)} \sum_{i=0}^{j-\ell} (\ell+1)^i (2^{j-\ell-i+1} - 1) + t^j v^{(j)} \\ &=: G_j v(t) + t^j v^{(j)}. \end{aligned}$$

Multiplying both sides of (1.1) by t and rearranging the terms, we obtain

$$tu' + tAu + \int_0^t (t-s)^{\alpha-1}\, sAu(s)\, ds + \int_0^t (t-s)^\alpha\, Au(s)\, ds = tf.$$

Differentiation yields

$$F_1 u'(t) + F_1 Au(t) + \int_0^t (t-s)^{\alpha-1} \left[F_1 Au(s) + \alpha\, F_0 Au(s)\right] ds = F_1 f(t).$$

Repeating the above two steps j-times, tedious calculations show that

$$F_j u'(t) + F_j Au(t) + \sum_{i=0}^{j} \alpha^{j-i} \binom{j}{i} \int_0^t (t-s)^{\alpha-1}\, F_i Au(s)\, ds = F_j f(t).$$

Therefore,

$$\begin{aligned} t^j u^{(j+1)}(t) + t^j Au^{(j)}(t) + \int_0^t (t-s)^{\alpha-1}\, s^j Au^{(j)}(s)\, ds &= F_j f(t) - G_j u'(t) - G_j Au(t) \\ - \int_0^t (t-s)^{\alpha-1}\, G_j Au(s)\, ds &- \sum_{i=0}^{j-1} \alpha^{j-i} \binom{j}{i} \int_0^t (t-s)^{\alpha-1}\, F_i Au(s)\, ds. \end{aligned}$$

We proceed in our proof by induction with respect to j and obtain, after lengthy but straightforward calculations,

$$t^j \|u^{(j+1)}(t)\|_{\mathbb{X}} + t^j \|Au^{(j)}(t)\|_{\mathbb{X}} \le C_0 d^{j+1} \Gamma(j+3) t^{\sigma-1} \quad \text{for} \quad t \in (0,T] \quad \text{and} \quad j \ge 0,$$

where $\sigma \geq 1$ with $\sigma = \min\{\alpha, \rho\} + 1$ for $j \geq 1$. This completes the proof. ∎

To resolve the singular behavior of the solution, we shall make use of geometrically refined time-steps and linearly increasing degree vectors [1, 21]. To that end, we first partition $(0,T)$ into (coarse) time intervals $\{\mathfrak{J}_i\}_{i=1}^K$. The first interval $\mathfrak{J}_1 = (0, T_1)$ near $t = 0$ is then further subdivided geometrically into $L+1$ subintervals $\{I_n\}_{n=1}^{L+1}$ by using the time-steps

$$t_0 = 0, \qquad t_n = \delta^{L+1-n} T_1 \quad \text{for } 1 \leq n \leq L+1. \tag{4.2}$$

As usual, we call $\delta \in (0,1)$ the geometric refinement factor and L is the number of refinement levels.

From (4.2), we observe that the subintervals $\{I_n\}_{n=1}^{L+1}$ satisfy

$$k_n = t_n - t_{n-1} = \lambda t_{n-1} \quad \text{with} \quad \lambda = (1-\delta)/\delta. \tag{4.3}$$

Let $\mathcal{M}_{L,\delta}$ be a geometric mesh of $(0,T)$ with $\{\mathfrak{J}_i\}_{i=1}^K$ denoting the underlying quasi-uniform partition of $(0,T)$, and let $\{I_n\}_{n=1}^{L+1}$ be the geometric refinement of $\mathfrak{J}_1$ defined by (4.2). Let $\mathcal{W}(\mathcal{M}_{L,\delta}, \mathbf{p})$ be the corresponding finite dimensional discrete space where the polynomial degrees p_n on the first interval $\mathfrak{J}_1$ are chosen to be linearly increasing:

$$p_n = \lfloor \mu n \rfloor \qquad \text{for } 1 \leq n \leq L+1, \tag{4.4}$$

for a parameter $\mu > 0$, and on the time intervals $\{\mathfrak{J}_i\}_{i=2}^K$ away from $t = 0$, we set the approximation degrees uniformly to $p_{L+1} = \lfloor \mu(L+1) \rfloor$.

Our main result of this section states that nonsmooth solutions satisfying (4.1) can be approximated at exponential rates convergence on the *hp*-version discretizations introduced above.

THEOREM 4.2. *Let the solution u of problem* (1.1) *satisfy the regularity property* (4.1). *Let $U \in \mathcal{W}(\mathcal{M}_{L,\delta}, \mathbf{p})$ be the hp-DG approximation obtained on a geometrically refined partition $\mathcal{M}_{L,\delta}$. Assuming that $U^0 = u_0$, then there exists a slope $\mu_0 > 0$ depending on δ and the constants σ and d in* (4.1) *such that for linearly increasing polynomial degree vectors $\mathbf{p}$ with slope $\mu \geq \mu_0$ we have the error estimate*

$$\|U - u\|_J + \|U - u\|_{L_2(0,T,\mathbb{X})} \leq C_1 \exp(-C_2 \mathcal{N}^{\frac{1}{2}}),$$

with constants C_1 and C_2 that are independent of the number $\mathcal{N} = \dim(\mathcal{W}(\mathcal{M}_{L,\delta}, \mathbf{p}))$.

Proof. We proceed in several steps.

Step 1. Setting $e = U - u$, we obtain from Theorem 3.10

$$\|e\|_J^2 + \int_0^T \|e\|_{\mathbb{X}}^2\, dt \leq C \max\{E_1, E_2\} + C \log(p_{L+1} + 1) p_{L+1}^2 (E_3 + E_4), \tag{4.5}$$

where

$$\begin{aligned}
E_1 &= \max_{n=1}^{L+1} \|\Pi u - u\|_{I_n}^2, \\
E_2 &= \max_{i=2}^{K} \|\Pi u - u\|_{\mathfrak{J}_i}^2, \\
E_3 &= \int_0^{T_1} \|\Pi u - u\|_{\mathbb{X}}^2\, dt, \\
E_4 &= \int_{T_1}^{T} \|\Pi u - u\|_{\mathbb{X}}^2\, dt.
\end{aligned}$$

On the coarse elements $\mathfrak{J}_i$, $2 \le i \le K$, away from $t = 0$ the solution u is analytic. Hence, from the first bound in Theorem 3.2, we readily find that

$$E_2 + E_4 \le C_1 \exp(-b_1 L). \tag{4.6}$$

It remains to bound the error on the element $\{I_n\}_{n=1}^{L+1}$ in $\mathfrak{J}_1$, i.e., the errors E_1 and E_3.

Step 2. On the first subinterval I_1 adjacent to $t = 0$, we set $q_n = 0$ and obtain, using Theorem 3.2, (3.7), and the regularity assumption (4.1),

$$\begin{aligned} \|\Pi u - u\|_{I_1}^2 &\le C k_1 \int_0^{t_1} \|u'\|_{\mathbb{X}}^2 \, dt \le C k_1 \int_0^{t_1} t^{2\sigma-2} \, dt \\ &= C \frac{k_1^{2\sigma}}{2\sigma - 1} \le C_2 \exp(-b_2 L). \end{aligned} \tag{4.7}$$

Similarly, we see that

$$\int_0^{t_1} \|\Pi u - u\|_{\mathbb{X}}^2 \, dt \le C k_1^2 \int_0^{t_1} \|u'\|_{\mathbb{X}}^2 \, dt \le C_3 \exp(-b_3 L). \tag{4.8}$$

Step 3. On the subintervals I_n away from the singular point $t = 0$ we start from Theorem 3.2 and (3.7) to get that, for $2 \le n \le L + 1$, $0 \le q_n \le p_n$,

$$\|\Pi u - u\|_{I_n}^2 \le C \left(\frac{k_n}{2}\right)^{2q_n+1} \Gamma_{p_n,q_n} \int_{t_{n-1}}^{t_n} \|u^{(q_n+1)}\|_{\mathbb{X}}^2 \, dt.$$

Then, from the regularity property (4.1), we readily conclude that

$$\begin{aligned} \|\Pi u - u\|_{I_n}^2 &\le C\, \Gamma_{p_n,q_n} \left(\frac{k_n}{2}\right)^{2q_n+1} d^{2q_n+2} \Gamma(q_n+3)^2 \int_{t_{n-1}}^{t_n} t^{2\sigma-2q_n-2} \, dt \\ &\le C\, \Gamma_{p_n,q_n} \left(\frac{k_n}{2}\right)^{2q_n+2} d^{2q_n+2} \Gamma(q_n+3)^2 t_{n-1}^{2\sigma-2q_n-2}. \end{aligned}$$

From (4.3) and (4.2), we have $k_n^{2q_n+2} = \lambda^{2q_n+2} t_{n-1}^{2q_n+2}$ with $t_{n-1} \le \delta^{L+2-n} T_1$ and hence

$$\begin{aligned} \|\Pi u - u\|_{I_n}^2 &\le C\, \Gamma_{p_n,q_n} d^{2q_n} \left(\frac{\lambda}{2}\right)^{2q_n+2} \Gamma(q_n+3)^2 t_{n-1}^{2\sigma} \\ &\le C \Gamma_{p_n,q_n} \left(\frac{d\lambda}{2}\right)^{2q_n} \Gamma(q_n+3)^2 \delta^{2\sigma L} \delta^{2\sigma(2-n)}. \end{aligned}$$

Since $\Gamma(q_n + 3) = (q_n + 2)(q_n + 1)\Gamma(q_n + 1) \le C q_n^2 \Gamma(q_n + 1)$,

$$\|\Pi u - u\|_{I_n}^2 \le C\, q_n^4 \Gamma_{p_n,q_n} \left(\frac{d\lambda}{2}\right)^{2q_n} \Gamma(q_n+1)^2 \delta^{2\sigma L} \delta^{2\sigma(2-n)}. \tag{4.9}$$

Using interpolation arguments analogous to [22, Lemma 3.39], it can be seen that property (4.9) also holds for any noninteger regularity parameter q_n with $0 \le q_n \le p_n$. Thus, we take $q_n = c_n p_n$ with $c_n \in (0, 1)$ and proceed as in [22, Theorem 3.36]. We obtain

$$\Gamma_{p_n,q_n} \left(\frac{d\lambda}{2}\right)^{2q_n} \Gamma(q_n+1)^2 \le C p_n \left(\left(\frac{\lambda d c_n}{2}\right)^{2c_n} \frac{(1-c_n)^{1-c_n}}{(1+c_n)^{1+c_n}} \right)^{p_n}.$$

Noting that

$$\inf_{0<c_n<1}\left(\frac{\lambda\, dc_n}{2}\right)^{2c_n}\frac{(1-c_n)^{1-c_n}}{(1+c_n)^{1+c_n}} =: \ell_{\lambda,d}(c_{\min})<1 \quad \text{with} \quad c_{\min}=\frac{1}{\sqrt{1+(\lambda d/2)^2}},$$

and thus, choosing $c_n = c_{\min}$ and using that $q_n \le p_n$, we conclude that

$$\|\Pi u-u\|_{I_n}^2 \le Cp_n^5(\ell_{\lambda,d}(c_{\min}))^{p_n}\delta^{2\sigma L}\delta^{-2\sigma n}.$$

Let now

$$\mu_0=\frac{2\sigma\log(\delta)}{\log(\ell_{\lambda,\mathrm{d}}(\mathrm{c}_{\min}))}>0.$$

Then, for $\mu \ge \mu_0$ and $p_n = \lfloor \mu n\rfloor \ge \mu_0 n$, we have

$$(\ell_{\lambda,d}(c_{\min}))^{p_n} \le \ell_{\lambda,d}(c_{\min})^{\mu_0 n} \le \delta^{2\sigma n},$$

and hence,

$$\|\Pi u-u\|_{I_n}^2 \le Cp_n^5\delta^{2\sigma L} \le Cp_{L+1}^5\delta^{2\sigma L} \le C_4\exp(-b_4 L) \quad \text{for} \quad 2\le n\le L+1, \tag{4.10}$$

where we have absorbed the factor p_{L+1}^5 into the constants C_4 and b_4.

Using similar arguments readily shows that

$$\begin{aligned}
\sum_{j=2}^{L+1}\int_{t_{j-1}}^{t_j}\|\Pi u-u\|_{\mathbb{X}}^2\,dt &\le \sum_{j=2}^{L+1} p_j^{-2}\,\Gamma_{p_j,q_j}\left(\frac{k_j}{2}\right)^{2q_j+2}\int_{t_{j-1}}^{t_j}\|u^{(q_j+1)}\|_{\mathbb{X}}^2\,dt\\
&\le C\sum_{j=2}^{L+1} p_j^{-2}\,\Gamma_{p_j,q_j}\left(\frac{k_j}{2}\right)^{2q_j+2} d^{2q_j}\Gamma(q_j+3)^2\int_{t_{j-1}}^{t_j} t^{2(\sigma-1-q_j)}\,dt\\
&\le C\sum_{j=2}^{L+1} p_j^{2}\,\Gamma_{p_j,q_j}\left(\frac{k_j}{2}\right)^{2q_j+3} d^{2q_j}\Gamma(q_j+1)^2 t_{j-1}^{2(\sigma-1-q_j)}\\
&\le C\sum_{j=2}^{L+1} p_j^{2}\,\Gamma_{p_j,q_j}\left(\frac{d\lambda}{2}\right)^{2q_j}\Gamma(q_j+1)^2 t_{j-1}^{2\sigma+1}\\
&\le C\delta^{L(2\sigma+1)}\sum_{j=2}^{L+1} p_j^3(\ell_{\lambda,d}(c_{\min}))^{p_j}\delta^{(2\sigma+1)(1-j)}\\
&\le C\delta^{L(2\sigma+1)}p_{L+1}^3\sum_{j=2}^{L+1}\left((\ell_{\lambda,d}(c_{\min}))^{p_j}\delta^{-2\sigma j}\right)\delta^{-j} \le C\delta^{L(2\sigma+1)}p_{L+1}^3\,.
\end{aligned}$$

Thus, we obtain

$$\sum_{j=2}^{L+1}\int_{t_{j-1}}^{t_j}\|\Pi u-u\|_{\mathbb{X}}^2\,dt \le C_5\exp(-b_5 L). \tag{4.11}$$

Step 4. We are now ready to complete the proof. From (4.7) and (4.10), we conclude that

$$E_1=\max_{j=1}^{L+1}\|\Pi u-u\|_{I_j} \le C_6\exp(-b_6 L). \tag{4.12}$$

Similarly, from (4.8) and (4.11) we get that

$$E_3 = \int_0^{T_1} \|\Pi u - u\|_{\mathbb{X}}\, dt \le C_7 \exp(-b_7 L). \tag{4.13}$$

Referring to (4.5), (4.6), (4.12), and (4.13) yields

$$\|e\|_J^2 + \int_0^T \|e\|_{\mathbb{X}}^2\, dt \le C\exp(-\tilde{b}L),$$

where we have absorbed the term $\log(p_{L+1}+1)p_{L+1}^2$ in (4.5) into the constants C and b. Since $\mathcal{N} = \dim(\mathcal{W}(\mathcal{M}_{L,\delta}, \mathbf{p})) \le CL^2$ for L sufficiently large, we obtain the desired result. ▯

5. Algebraic convergence. In this section, we study the convergence analysis of the h-version DG method assuming that the order of the DG solution U defined by (2.7) is p (i.e., $p_j = p \ge 0$ for all $j \ge 1$), and $U^0 = u_0$. Furthermore, we assume that the solution u of (1.1) satisfies the regularity assumption

$$\|u^{(j)}(t)\|_{\mathbb{X}} \le C_p\, t^{\sigma-j} \quad \text{for} \quad 1 \le j \le p+1, \quad \text{where} \quad \sigma \ge 1. \tag{5.1}$$

As before, the singular behavior of u near $t = 0$ may lead to suboptimal convergence rates if we work with quasi-uniform time meshes. Therefore, we employ a family of non-uniform meshes denoted by $\mathcal{M}_\gamma$, where the time-steps are concentrated near $t = 0$. To this end, we assume that, for a fixed $\gamma \ge 1$,

$$k_n \le C_\gamma k t_n^{1-1/\gamma} \quad \text{and} \quad t_n \le C_\gamma t_{n-1} \quad \text{for } 2 \le n \le N, \tag{5.2}$$

with

$$c_\gamma k^\gamma \le k_1 \le C_\gamma k^\gamma. \tag{5.3}$$

For instance, one may choose

$$t_n = (n/N)^\gamma T \quad \text{for } 0 \le n \le N. \tag{5.4}$$

In the next theorem we derive the following error estimate of the h-version DG solution, giving rise to optimal algebraic rates of convergence.

THEOREM 5.1. *Let the solution u of problem* (1.1) *satisfy the regularity property* (5.1). *Let $U \in \mathcal{W}(\mathcal{M}_\gamma, \mathbf{p})$ be the DG approximation with $\mathbf{p} = (p, \ldots, p)$ with $p \ge 0$, and assume that $U^0 = u_0$. Then we have the error estimate*

$$\|U - u\|_J \le C \times \begin{cases} k^{\gamma\sigma}, & 1 \le \gamma < (p+1)/\sigma, \\ k^{p+1}, & \gamma \ge (p+1)/\sigma, \end{cases}$$

where C is a constant that depends on T, γ, σ, and p.

Proof. Theorem 3.10 yields

$$\|U - u\|_J^2 \le C \left(\|u - \Pi u\|_J^2 + \int_0^T \|u - \Pi u\|_{\mathbb{X}}^2\, dt \right). \tag{5.5}$$

Using (3.7), the regularity assumption (5.1), and (5.3), we get

$$\|u - \Pi u\|_{I_1}^2 \le Ck_1 \int_0^{t_1} \|u'(t)\|_{\mathbb{X}}^2\, dt \le Ck_1 \int_0^{t_1} t^{2\sigma-2}\, dt = C\frac{t_1^{2\sigma}}{2\sigma - 1} \le Ck^{2\gamma\sigma}.$$

For $n \geq 2$, we use (5.2) and obtain

$$\begin{aligned}\|u - \Pi u\|_{I_n}^2 &\leq C \left(\frac{k_n}{2}\right)^{2p+1} \Gamma_{p,p} \int_{t_{n-1}}^{t_n} \|u^{(p+1)}(t)\|_{\mathbb{X}}^2 \, dt \\ &\leq C k_n^{2p+1} \int_{t_{n-1}}^{t_n} t^{2\sigma-2p-2} \, dt \\ &\leq C \, k_n^{2p+2} t_n^{2\sigma-2p-2} \\ &\leq C \, k^{2p+2} t_n^{2\sigma-(2p+2)/\gamma}.\end{aligned}$$

Thus, we may bound the interpolation error over $(0, T]$ as follows:

$$\|u - \Pi u\|_J^2 = \max_{n=1}^{N} \|u - \Pi u\|_{I_n}^2 \leq C \times \begin{cases} k^{2\gamma\sigma}, & 1 \leq \gamma \leq (p+1)/\sigma, \\ k^{2p+2}, & \gamma \geq (p+1)/\sigma. \end{cases} \tag{5.6}$$

Similar to the above derivations and using Theorem 3.2,

$$\int_0^{t_1} \|u - \Pi u\|_{\mathbb{X}}^2 \, dt \leq C k_1^2 \int_0^{t_1} \|u'(t)\|_{\mathbb{X}}^2 \, dt \leq C k_1^{2\sigma+1} \leq C k^{\gamma(2\sigma+1)}$$

and

$$\begin{aligned}\sum_{n=2}^{N} \int_{t_{n-1}}^{t_n} \|u - \Pi u\|_{\mathbb{X}}^2 \, dt &\leq C \Gamma_{p,p} \sum_{n=2}^{N} \left(\frac{k_n}{2}\right)^{2p+2} \int_{t_{n-1}}^{t_n} \|u^{(p+1)}(t)\|_{\mathbb{X}}^2 \, dt \\ &\leq C \sum_{n=2}^{N} k_n^{2p+2} \int_{t_{n-1}}^{t_n} t^{2(\sigma-1-p)} \, dt \\ &\leq C \, k^{2p+2} \sum_{j=2}^{N} t_n^{(1-1/\gamma)(2p+2)} \int_{t_{n-1}}^{t_n} t^{2(\sigma-1-p)} \, dt \\ &\leq C \, k^{2p+2} \int_{t_1}^{T} t^{2\sigma-(2p+2)/\gamma} \, dt.\end{aligned}$$

Therefore,

$$\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|u - \Pi u\|_{\mathbb{X}}^2 \, dt \leq C \left(k^{\gamma(2\sigma+1)} + k^{2p+2} \int_{t_1}^{T} t^{2\sigma-(2p+2)/\gamma} \, dt \right),$$

and the result follows from (5.5) and (5.6), after noting that

$$\begin{aligned}\int_{t_1}^{T} t^{2\sigma-(2p+2)/\gamma} \, dt &\leq C \times \begin{cases} t_1^{2\sigma-(2p+2)/\gamma+1}, & 1 \leq \gamma < (p+1)/(\sigma+1/2), \\ \log(T/t_1), & \gamma = (p+1)/(\sigma+1/2), \\ T^{2\sigma+1-(2p+2)/\gamma}, & \gamma > (p+1)/(\sigma/2), \end{cases} \\ &\leq C \times \begin{cases} k^{2\gamma\sigma-2p-2}, & 1 \leq \gamma < (p+1)/\sigma, \\ T^{2\sigma+1-(2p+2)/\gamma}, & \gamma \geq (p+1)/\sigma. \end{cases}\end{aligned}$$

This finishes the proof. $\square$

Remark 5.2. For the piecewise-constant case $p = 0$, since $U'(t) = 0$ and $U(t) = U^n = U_+^{n-1}$ for $t \in I_n$, the DG method (2.7) amounts to a generalized backward-Euler scheme

$$\left\langle \frac{U^n - U^{n-1}}{k_n}, \chi \right\rangle + A(U^n, \chi) + \omega_{nn} k_n A(U^n, \chi) = \langle \bar{f}^n, \chi \rangle - \sum_{j=1}^{n-1} \omega_{nj} k_j A(U^j, \chi)$$

for all $\chi \in \mathbb{X}$, where

$$\bar{f}^n = \frac{1}{k_n} \int_{t_{n-1}}^{t_n} f(t)\, dt \quad \text{and} \quad \omega_{nj} = \frac{1}{k_n k_j} \int_{t_{n-1}}^{t_n} \int_{t_{j-1}}^{\min(t,t_j)} (t-s)^{\alpha-1} b(s)\, ds\, dt.$$

In this case, we observe from Theorem 5.1 that an optimal convergence rate can be achieved over a uniform time mesh.

6. Fully discrete scheme and error estimates.

In this section we introduce and analyze a fully discrete scheme for numerically solving the following parabolic integro-differential equation: Find $u(x,t)$ such that

$$u_t - \Delta u - \mathcal{B}\Delta u = f(x,t) \qquad \text{in } \Omega \times (0,T), \tag{6.1}$$

$$u = 0 \qquad \text{on } \partial\Omega \times (0,T), \tag{6.2}$$

$$u|_{t=0} = u_0 \qquad \text{in } \Omega. \tag{6.3}$$

Here, Ω is a bounded and convex Lipschitz domain in $\mathbb{R}^d$ for $d \geq 1$. As pointed out in Remark 2.1, problem (6.1)–(6.3) fits into the framework of section 2.1 with the spaces $\mathbb{H} = L_2(\Omega)$ and $\mathbb{X} = H_0^1(\Omega)$. The spatial operator is $A = -\Delta$, and the associated spatial bilinear form is given by

$$A(u,v) = \int_\Omega \nabla u \cdot \nabla v\, dx. \tag{6.4}$$

6.1. Discretization.

To discretize (6.1)–(6.3), we will employ the hp-DG time discretization combined with a standard continuous finite element discretization in space.

We construct a partition of the domain Ω into (families of shape-regular) triangular or quadrilateral finite elements with maximum diameter h, and let $S_h \subset H_0^1(\Omega)$ denote the space of continuous, piecewise polynomial functions of degree $\leq r$ with $r \geq 1$.

For a partition $\mathcal{M} = \{I_n\}_{n=1}^N$ of $(0,T)$ and a degree vector $\mathbf{p} = (p_1, p_1, \ldots, p_N)$, the trial space is now given by

$$\mathcal{W}(\mathcal{M}, \mathbf{p}, S_h) = \{ U_h : [0,T] \to S_h \,:\, U_h|_{I_n} \in \mathbb{P}_{p_n}(S_h),\ 1 \leq n \leq N \}. \tag{6.5}$$

Here, we denote by $\mathbb{P}_p(S_h)$ the space of polynomials of degree $\leq p$ in the time variable with coefficients in S_h. Thus, a function $U_h(x,t)$ in $\mathcal{W}(\mathcal{M}, \mathbf{p}, S_h)$ is continuous in x but may be discontinuous over $t = t_n$.

Applying the hp-DG time-stepping method and standard finite elements in space, we arrive at the following fully-discrete hp-DG finite element scheme: Find $U_h \in \mathcal{W}(\mathcal{M}, \mathbf{p}, S_h)$ such that

$$\begin{aligned} G_N(U_h, X) &= \langle U_h^0, X_+^0 \rangle + \int_0^{t_N} \langle f(t), X(t) \rangle\, dt \quad \forall\ X \in \mathcal{W}(\mathcal{M}, \mathbf{p}, S_h), \\ U_h(0) &= U_h^0 \end{aligned} \tag{6.6}$$

for a suitable approximation $U_h^0 \in S_h$ to u_0.

6.2. Error estimates. To analyze the formulation (6.6), in place of (3.9) we now decompose the error as

$$U_h - u = (U_h - \Pi R_h u) + \Pi\xi + \eta, \tag{6.7}$$

with $\xi = R_h u - u$ and η defined in (3.9). The operator $R_h : H_0^1(\Omega) \to S_h$ is the Ritz projection associated with the bilinear form $A(u, v)$. It is given by

$$A(R_h v, \chi) = A(v, \chi) \quad \text{for } v \in H_0^1(\Omega) \text{ and } \chi \in S_h. \tag{6.8}$$

In what follows, we denote by $H^{s+1}(\Omega)$ the standard Sobolev space of order $s+1$ and write $\|u\|_{s+1}$ for its norm. The standard $L_2(\Omega)$-norm is denoted by $\|u\|$. The projection R_h satisfies the following approximation property.

LEMMA 6.1. *For $r \geq 1$ and $s \geq 0$, we have*

$$\|u - R_h u\|^2 \leq C\frac{h^{2\min\{s,r\}+2}}{r^{2s+2}}\|u\|_{s+1}^2. \tag{6.9}$$

Then, the following result holds.

THEOREM 6.2. *If u is the solution of problem* (6.1)–(6.3), *and $U_h \in \mathcal{W}(\mathcal{M}, \mathbf{p}, S_h)$ is the approximate solution defined by* (6.6), *then*

$$\begin{aligned} G_N(U_h - \Pi R_h u, X) &= \langle U_h^0 - R_h u_0, X_+^0\rangle - \int_0^{t_N} \langle \xi', X\rangle \, dt \\ &\quad - \int_0^{t_N} \Big[A(\eta, X) + \mathcal{B}\left[A(\eta(\cdot), X)\right]\Big]\, dt \quad \forall X \in \mathcal{W}(\mathcal{M}, \mathbf{p}, S_h). \end{aligned} \tag{6.10}$$

Proof. We first note that the Galerkin orthogonality property (3.3) now takes the form

$$G_N(U_h - u, X) = \langle U_h^0 - u_0, X_+^0\rangle \quad \forall X \in \mathcal{W}(\mathcal{M}, \mathbf{p}, S_h).$$

Hence, from the decomposition (6.7) we see that

$$G_N(U_h - \Pi R_h u, X) = \langle U_h^0 - u_0, X_+^0\rangle - G_N(\Pi\xi + \eta, X) \quad \forall X \in \mathcal{W}(\mathcal{M}, \mathbf{p}, S_h). \tag{6.11}$$

Since $(\Pi\xi)^n = \xi^n$ and $\eta^n = 0$, using the alternative expression for G_N in Remark 3.1 yields

$$\begin{aligned} G_N(\Pi\xi + \eta, X) &= \langle \xi^N, X^N\rangle - \sum_{n=1}^{N-1}\langle \xi^n, [X]^n\rangle \\ &\quad + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \Big[- \langle \Pi\xi + \eta, X'\rangle + A\big(\Pi\xi + \eta, X\big) + \mathcal{B}\left[A(\Pi\xi(\cdot) + \eta(\cdot), X)\right]\Big]\, dt. \end{aligned}$$

With the aid of the equality $\int_{t_{n-1}}^{t_n} \langle \Pi\xi, X'\rangle\, dt = \int_{t_{n-1}}^{t_n} \langle \xi, X'\rangle\, dt$, integration by parts shows that

$$\int_{t_{n-1}}^{t_n} \langle \Pi\xi, X'\rangle\, dt = \int_{t_{n-1}}^{t_n} \langle \xi, X'\rangle\, dt = \langle \xi^n, X^n\rangle - \langle \xi^{n-1}, X_+^{n-1}\rangle - \int_{t_{n-1}}^{t_n} \langle \xi', X\rangle\, dt.$$

Therefore, since $A\big(\Pi\xi, X\big) = A\big(\Pi(R_h u - u), X\big) = A\big(R_h \Pi u - \Pi u, X\big) = 0$ (from the definition of the Ritz projector), we observe that

$$G_N(\Pi\xi + \eta, X) = \langle \xi^0, X_+^0\rangle - \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\langle \eta, X'\rangle\,dt + \int_0^{t_N}\langle \xi', X\rangle\,dt \\ + \int_0^{t_N}\Big[A(\eta, X) + \mathcal{B}\left[A(\eta(\cdot), X)\right]\Big]\,dt.$$

Finally, we insert this expression into (6.11) and use that $\int_{t_{n-1}}^{t_n}\langle \eta, X'\rangle\,dt = 0$, which completes the proof. □

For brevity, we set $\psi = U_h - \Pi R_h u$ and prove the following two auxiliary estimates.

LEMMA 6.3. *For $1 \le n \le N$, we have*

$$\|\psi^n\|^2 \;+\; \int_0^{t_n}\|\psi\|_1^2\,dt \;\;\le\;\; C\|U_h^0 \;-\; R_h u_0\|^2 \;+\; C\int_0^{t_n}\|\xi'\|^2\,dt \;+\; C\int_0^{t_n}\|\eta\|_1^2\,dt.$$

Proof. We choose $X = \psi$ in (6.10), follow the proof of Lemma 3.7 with $\langle \xi', \psi\rangle + A(\eta, \psi)$ in place of $A(\eta, \theta)$ in Q_1^n, and use the inequality $\|\psi\| \le \|\psi\|_1$. The desired result then readily follows. □

LEMMA 6.4. *For $1 \le n \le N$, we have*

$$\int_{t_{n-1}}^{t_n}\|\psi'\|^2(t - t_{n-1})\,dt \;\le\; Cp_n^2\left(\|U_h^0 - R_h u_0\| + \int_0^{t_n}\|\xi'\|^2\,dt + \int_0^{t_n}\|\eta\|_1^2\,dt\right).$$

Proof. We choose $X = (t - t_{n-1})\psi'$ on I_n and zero elsewhere in (6.10), and then following the steps given in the proof of Lemma 3.8 with

$$Q_4^n = \int_{t_{n-1}}^{t_n}\left[\langle \xi'(t), (t - t_{n-1})\psi'\rangle + A(\eta, (t - t_{n-1})\psi')\right]dt,$$

we readily obtain the required result. □

Next, we estimate the first term on the right-hand side of (6.7).

LEMMA 6.5. *If $U_h \in \mathcal{W}(\mathcal{M}, \mathbf{p}, S_h)$ is the approximate solution defined by* (6.6), *then, for $1 \le n \le N$,*

$$\|U_h - \Pi R_h u\|_{J_n}^2 \le C\log(|\mathbf{p}|_n + 2)|\mathbf{p}|_n^2\left(\|U_h^0 - R_h u_0\|^2 + \int_0^{t_n}\|\xi'\|^2 dt + \int_0^{t_n}\|\eta\|_1^2 dt\right).$$

Proof. Adapting the proof of Lemma 3.9 and using Lemmas 6.3 and 6.4 instead of Lemmas 3.7 and 3.8, respectively, we complete the proof. □

We are now ready to show the following error estimates for the fully discrete scheme. For the rest of this paper, let u be the solution of (6.1)–(6.3), and let U_h be the approximate solution defined by (6.6) with $U_h^0 = R_h u_0$.

THEOREM 6.6. *For $1 \le n \le N$, we have the error estimates*

$$\int_0^{t_n}\|U_h - u\|^2\,dt + \|(U_h - u)^n\|^2 \le C\left(\|\Pi\xi\|_{J_n}^2 + \int_0^{t_n}(\|\xi'\|^2 + \|u - \Pi u\|_1^2)\,dt\right)$$

and

$$\|U_h - u\|_{J_n}^2 \le C(\|u - \Pi u\|_{J_n}^2 + \|\Pi\xi\|_{J_n}^2) + C\log(|\mathbf{p}|_n + 2)|\mathbf{p}|_n^2\int_0^{t_n}(\|\xi'\|^2 + \|u - \Pi u\|_1^2)\,dt.$$

Proof. To prove the first bound, we start from the decomposition of $U_h - u$ in (6.7), then employ the triangle inequality, Lemma 6.3, and the fact that $\eta^n = 0$ for $1 \le n \le N$. The second bound follows similarly using Lemma 6.5. ◻

In the remainder of this paper we assume that u and the corresponding initial condition u_0 satisfy the regularity assumptions:

$$u_0 \in H^{s+1}(\Omega), \quad u|_{I_n} \in H^{q_n+1}(t_{n-1}, t_n; H^1(\Omega)) \cap H^1(t_{n-1}, t_n; H^{s+1}(\Omega))$$

for $1 \le n \le N$ and $1 \le s \le r$.

THEOREM 6.7. *For $1 \le n \le N$ and for $0 \le q_j \le p_j$, we have*

$$\int_0^{t_n} \|U_h - u\|^2 \, dt + \|(U_h - u)^n\|^2 \le C \frac{h^{2\min\{s,r\}+2}}{r^{2s+2}} \left(\|u_0\|_{s+1}^2 + \int_0^{t_n} \|u'\|_{s+1}^2 \, dt \right) + C \sum_{j=1}^n \hat{p}_j^{-2} k_j e_1(k_j, p_j, q_j)$$

and

$$\|U_h - u\|_{J_n}^2 \le C \frac{h^{2\min\{s,r\}+2}}{r^{2s+2}} \log(|\mathbf{p}|_n + 2)|\mathbf{p}|_n^2 \left(\|u_0\|_{s+1}^2 + \int_0^{t_n} \|u'\|_{s+1}^2 \, dt \right) + C \left(\max_{j=1}^n e_1(k_j, p_j, q_j) + \log(|\mathbf{p}|_n + 2)|\mathbf{p}|_n^2 \sum_{j=1}^n \hat{p}_j^{-2} k_j e_1(k_j, p_j, q_j) \right),$$

where $\hat{p}_j = \max\{1, p_j\}$ and

$$e_1(k_j, p_j, q_j) = \left(\frac{k_j}{2} \right)^{2q_j+1} \Gamma_{p_j, q_j} \int_{t_{j-1}}^{t_j} \|u^{(q_j+1)}\|_1^2 \, dt \, .$$

Proof. Using Theorems 6.6 and 3.2 reduced our task to bound $\|\Pi\xi\|_{J_n}$ and $\int_0^{t_n} \|\xi'\|^2 \, dt$. The triangle inequality yields

$$\|\Pi\xi\|_{J_n} \le \|\Pi\xi - \xi\|_{J_n} + \|\xi\|_{J_n} \le \|\Pi\xi - \xi\|_{J_n} + \|\xi(0)\| + \int_0^{t_n} \|\xi'\| \, dt.$$

To bound the first term on the right-hand side, we use Theorem 3.2 for $q_n = 0$ with $\Pi\xi - \xi$ in place of $\Pi u - u$ and get

$$\|\Pi\xi - \xi\|_{J_n}^2 = \max_{j=1}^n \left(\|\Pi\xi - \xi\|_{I_j}^2 \right) \le C \max_{j=1}^n \left(k_j \int_{t_{j-1}}^{t_j} \|\xi'\|^2 \, dt \right),$$

and thus, with the help of the Cauchy–Schwarz inequality for integrals, we obtain

$$\|\Pi\xi\|_{J_n}^2 \le C \left(\|\xi(0)\|^2 + \int_0^{t_n} \|\xi'\|^2 \, dt \right).$$

Therefore, after noting from the approximation property (6.9) that

$$\|\xi(0)\|^2 + \int_0^{t_n} \|\xi'\|^2 \, dt \le C \frac{h^{2\min\{s,r\}+2}}{r^{2s+2}} \|u_0\|_{s+1}^2 + C \int_0^{t_n} \frac{h^{2\min\{s,r\}+2}}{r^{2s+2}} \|u'\|_{s+1}^2 \, dt,$$

the assertion follows. ◻

For uniform parameters k, p, and q (i.e., $k_j = k$, $p_j = p$, and $q_j = q$), the bounds in Theorem 6.7 result in the following error estimates.

COROLLARY 6.8. *For $1 \le n \le N$, we have*

$$\int_0^{t_n} \|U_h - u\|^2 dt + \|(U_h - u)^n\|^2 \le C\frac{h^{2\min\{r,s\}+2}}{r^{2s+2}} \left(\|u_0\|_{s+1}^2 + \int_0^{t_n} \|u'\|_{s+1}^2 \, dt \right) + C\,\frac{k^{2\min\{p,q\}+2}}{p^{2q+2}} \int_0^{t_n} \|u^{(q+1)}\|_1^2 dt$$

and

$$\|U_h - u\|_{J_n}^2 \le C\frac{p^2 h^{2\min\{r,s\}+2}}{r^{2s+2}} \log(p+2) \left(\|u_0\|_{s+1}^2 + \int_0^{t_n} \|u'\|_{s+1}^2 \, dt \right) + C\frac{k^{2\min\{p,q\}+2}}{p^{2q}} \left(\max_{j=1}^{n} \max_{t\in I_j} \|u^{(q+1)}(t)\|_1 + \log(p+2) \int_0^{t_n} \|u^{(q+1)}\|_1^2 \, dt \right).$$

Proof. These estimates follow readily from Theorem 6.7 and the fact that $\Gamma_{p,q}$ behaves like p^{-2q} for $p \to \infty$. ∎

The estimates in Corollary 6.8 show that the discrete scheme converges either as k, $h \to 0$, or as p, $r \to \infty$. We observe that the first estimate is optimal in the four parameters k, h, p, and r, while the second one falls short by one power from being optimal in p. For a smooth solution u, spectral convergence rates are achieved if the polynomial degrees p and r are increased on fixed partitions.

COROLLARY 6.9. *We assume the regularity estimates* (4.1) *and* (5.1) *for $\mathcal{M} = \mathcal{M}_{L,\delta}$ and $\mathcal{M} = \mathcal{M}_\gamma$, respectively. Also, we assume that $\|u_0\|_{s+1}^2 + \int_0^{t_n} \|u'(t)\|_{s+1}^2 \, dt \le C$ for some $1 \le s \le r$. Then*

$$\|U_h - u\|_J \le C \log(p_{L+1} + 2) p_{L+1}^2 \frac{h^{2\min\{s,r\}+2}}{r^{2s+2}} + C \exp(-\tilde{b}\mathcal{N}^{\frac{1}{2}}) \text{ for } \mathcal{M} = \mathcal{M}_{L,\delta},$$

where b is a constant independent of the number $\mathcal{N} = \dim(\mathcal{W}(\mathcal{M}_{L,\delta}, \mathbf{p}))$, and

$$\|U_h - u\|_J \le C\frac{h^{2\min\{s,r\}+2}}{r^{2s+2}} + C \begin{cases} k^{\gamma\sigma}, & 1 \le \gamma < (p+1)/\sigma \\ k^{p+1}, & \gamma \ge (p+1)/\sigma \end{cases} \quad \textit{for } \mathcal{M} = \mathcal{M}_\gamma\,.$$

Proof. These results follow immediately from Theorem 6.7 and the already bounded term $e_1(k_j, p_j, q_j)$ for $1 \le j \le N$ inside Theorems 4.2 and 5.1. ∎

7. Numerical examples.

We now apply the hp-DG method (2.7) and its spatially discrete version (6.6) to some problems of the form (1.1) and (6.1)–(6.3). In all our examples, we consider $T = 1$.

7.1. Scalar examples.

To demonstrate the effect of the time discretization by itself, with no additional errors arising from a spatial discretization, we first consider the scalar Volterra integro-differential equation

$$u'(t) + u(t) + \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} u(s) \, ds = f(t) \quad \text{for } 0 < t < T \text{ with } u(0) = u_0. \tag{7.1}$$

We choose u_0 and $f(t)$ such that the solution u of (7.1) is given by

$$u(t) = t^{\alpha+1} \exp(-t). \tag{7.2}$$

Table 7.1

The errors $\|U-u\|_{J,11}$ with different mesh gradings for the h-version DG method of order p and $\alpha=0.5$. We observe numerical convergence order $O(k^{(\alpha+1)\gamma})$ for $1\le\gamma<(p+1)/(\alpha+1)$, and $O(k^{p+1})$ for $\gamma>(p+1)/(\alpha+1)$.

$p=1$	i	$\gamma=1$		$\gamma=4/3$		$\gamma=2$	
	6	2.133e-04		3.279e-05		4.797e-05	
	7	7.626e-05	1.477	8.283e-06	1.976	1.189e-05	2.003
	8	2.710e-05	1.486	2.090e-06	1.978	2.959e-06	1.997
	9	9.609e-06	1.489	5.261e-07	1.981	7.383e-07	1.994
$p=2$	i	$\gamma=1$		$\gamma=2$		$\gamma=2.2$	
	5	4.399e-05		8.022e-07		8.663e-07	
	6	1.512e-05	1.540	1.010e-07	2.989	1.089e-07	2.991
	7	5.269e-06	1.521	1.266e-08	2.995	1.366e-08	2.995
	8	1.848e-06	1.511	1.585e-09	2.998	1.710e-09	2.998
$p=3$	i	$\gamma=1$		$\gamma=2$		$\gamma=8/3$	
	3	1.184e-04		4.881e-06		5.591e-06	
	4	4.018e-05	1.559	6.060e-07	3.009	3.566e-07	3.970
	5	1.393e-05	1.528	7.562e-08	3.002	2.241e-08	3.992
	6	4.881e-06	1.513	9.449e-09	3.001	1.405e-09	3.995

For $\alpha\in(0,1)$, we notice that near $t=0$ the second derivative $u''(t)$ is unbounded, while u is real-analytic away from $t=0$.

For scalar problems of this type, the *hp*-DG method (including *h*- and *p*-versions) has been extensively tested in [1], for smooth and nonsmooth solutions. Here we illustrate the results of section 5 (which have not been demonstrated in [1], neither theoretically nor numerically). To do so, we employ a time mesh of the form (5.4) with $N=2^i$ subintervals for various choices of the mesh grading parameter $\gamma\ge 1$. To tabulate our numerical results, we introduce the finer grid

$$\mathcal{G}^{N,m}=\{\,t_{i-1}+\ell k_i/m : 1\le i\le N \text{ and } 0\le \ell\le m\,\}, \tag{7.3}$$

and setting $\|v\|_{J,m}:=\max_{t\in\mathcal{G}^{N,m}}|v(t)|$. Thus, for large values of m, $\|U-u\|_{J,m}$ can be viewed as an approximation of the uniform error $\|U-u\|_J$.

For $0<\alpha<1$, since the solution u in (7.2) behaves like $t^{\alpha+1}$ as $t\to 0^+$, the regularity condition (5.1) holds for $\sigma=\alpha+1$. Thus, from Theorem 5.1 we expect $\|U-u\|_J$ to converge of order $O(k^{\gamma\sigma})$ for $1\le\gamma<(p+1)/(\alpha+1)$, and of order $O(k^{p+1})$ for $\gamma\ge(p+1)/(\alpha+1)$. The numerical results shown in Table 7.1 are consistent with these error bounds.

7.2. A problem in one space dimension. In this section, we verify the theoretical results of section 6 for the following parabolic integro-differential equation in one space dimension:

$$u_t-u_{xx}-\int_0^t\frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}u_{xx}(x,s)\,ds=f(x,t),\qquad (x,t)\in\Omega\times(0,1),$$
$$u(x,0)=u_0(x),\qquad x\in\Omega.$$

Here, we take $\Omega=(0,1)$, and assume that $u=u(x,t)$ satisfies the homogeneous Dirichlet boundary conditions $u(0,t)=0=u(1,t)$ for all $t\in(0,1)$. The initial datum is chosen so that the exact solution is given by

$$u(x,t)=\sin(\pi x)-t^{1+\alpha}\exp(-t)\sin(2\pi x).$$

Table 7.2
The errors $\|U_h - u\|_{J,11}$ for the h-version DG method of spatial order $r = 2$ for different mesh gradings and $\alpha = 0.5$. We observe convergence of order $h^{\min\{r+1,(\alpha+1)\gamma\}}$ for $1 \le \gamma \le (p+1)/(\alpha+1)$.

	i	$\gamma = 1$		$\gamma = 4/3$		$\gamma = 2$	
$p = 1$	4	7.224e-04		2.617e-04		3.917e-04	
	5	3.220e-04	1.166	7.975e-05	1.714	1.107e-04	1.823
	6	1.314e-04	1.292	2.199e-05	1.858	3.008e-05	1.879
	7	5.027e-05	1.386	5.717e-06	1.944	7.899e-06	1.929
$p = 2$	4	1.044e-04		9.548e-05		9.555e-05	
	5	3.343e-05	1.643	1.195e-05	2.998	1.195e-05	2.998
	6	1.126e-05	1.570	1.494e-06	2.999	1.495e-06	2.999
	7	3.871e-06	1.540	3.436e-07	2.121	1.868e-07	2.999

It can be readily seen that the regularity conditions (4.1) and (5.1) hold for $\sigma \le \alpha+1$.

We apply the fully discrete scheme (6.6) with the space $S_h \subset H_0^1(\Omega)$ of continuous piecewise polynomials of degree r. We choose U_h^0 to be the L_2-projection of the initial datum u_0 into the space S_h. We measure the error in the norm

$$\|v\|_{J,m} := \max_{t\in\mathcal{G}^{N,m}} \|v(t)\|.$$

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 nonuniformly graded meshes $\mathcal{M} = \mathcal{M}_\gamma$ in (5.4) for various choices of $\gamma \ge 1$. In space, we consider a mesh sequence consisting of $N_x = 2^i$ uniform subintervals, each of length $h = 1/N_x$. This means that there is a constant c_γ such that $c_\gamma k \le h \le k$. From Corollary 6.9, we see that the global error is bounded by

$$\|U_h - u\|_J \le Ch^{r+1} + Ck^{\gamma(\alpha+1)} \quad \text{for } 1 \le \gamma \le (p+1)/(\alpha+1).$$

Hence, we expect to see convergence of order $h^{\min\{r+1,\gamma(1+\alpha)\}}$. The results shown in Table 7.2 are in full agreement with these error bounds. Next, we test the performance of the hp-version time-stepping and use the geometric time partition $\mathcal{M}_{L,\delta}$ defined in (4.2)–(4.4), again on a uniform spatial mesh with N_x subintervals. We set $T_1 = 1$ and $\mu = 1$, so that we have a geometric time-mesh consisting of $L+1$ subintervals with a refinement factor equal to δ. The regularity assumption (4.1) holds for $\sigma = \alpha + 1$, and thus from Corollary 6.9 the global error is bounded by

$$\|U_h - u\|_J \le Ch^{r+1} + C\exp(-\tilde{b}\mathcal{N}^{1/2}), \text{ where } \mathcal{N} = \dim(\mathcal{W}(\mathcal{M}_{L,\delta}, \mathbf{p})).$$

We approximate the norm $\|v\|_{J,m} = \max_{t\in\mathcal{G}^{L+1,m}} \|v(t)\|$ as before.

In Table 7.3, we set $\delta = 0.3$ and compute the error and the numerical order of convergence with respect to the change in the number of subintervals in the spatial mesh by using the following formula:

$$\frac{\log(\text{error}(N_x(i-1))/\text{error}(N_x(i)))}{\log(N_x(i)/N_x(i-1))} \quad \text{for} \quad i \ge 1,$$

where $N_x(i) = 2^{i+4}$ and $\text{error}(N_x(i))$ is the corresponding error with $L = i + 3$. For $r = 1$, we observe that the convergence rate is of the optimal order 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 error of the time discretization.

TABLE 7.3
The errors $\|U_h - u\|_{J,51}$ and the order of convergence with respect to N_x for $\alpha = 0.5$.

L	N_x	$r = 1$		$r = 2$	
3	16	4.4061e-03		4.5383e-04	
4	32	1.1117e-03	1.9867	8.6172e-05	2.3969
5	64	2.7729e-04	2.0033	1.4845e-05	2.5372
6	128	6.9422e-05	1.9979	2.4743e-06	2.5849
7	256	1.7357e-05	1.9998	4.0829e-07	2.5994

TABLE 7.4
The errors $\|U_h - u\|_{J,51}$ and the number $\tilde{b}$ for different choices of δ for $\alpha = 0.5$, $r = 2$, and $N_x = 200$.

L	$\mathcal{N}(L)$	$\delta = 0.25$		$\delta = 0.3$		$\delta = 0.35$	
3	14	2.1701e-04		4.5280e-04		8.1656e-04	
4	20	2.9864e-05	2.7151	8.6086e-05	2.2726	2.0525e-04	1.8904
5	27	3.8272e-06	2.8377	1.4837e-05	2.4284	4.6033e-05	2.0647
6	35	4.8163e-07	2.8790	2.4736e-06	2.4884	9.8118e-06	2.1471
7	44	8.4694e-08	2.4236	4.0852e-07	2.5111	2.0525e-06	2.1815

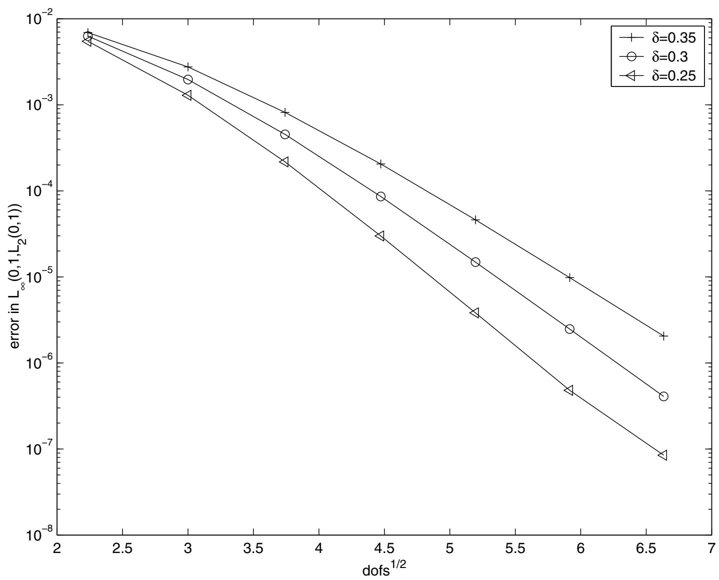

FIG. 7.1. *The errors $\|U_h - u\|_{J,51}$ plotted against $\mathcal{N}^{1/2}$ for different refinement factors δ for $\alpha = 0.5$, $r = 2$, and $N_x = 200$.*

To demonstrate exponential convergence in time, we choose $r = 2$ and take N_x relatively large so that the time errors are dominating. Then we use the formula

$$\frac{\log(\text{error}(\mathcal{N}(L-1))/\text{error}(\mathcal{N}(L)))}{\sqrt{\mathcal{N}(L)} - \sqrt{\mathcal{N}(L-1)}}$$

to calculate the coefficient $\tilde{b}$ in the expected exponential error estimates $\exp(-\tilde{b}\mathcal{N}^{1/2})$, where $\mathcal{N}(L) = \dim(\mathcal{W}(\mathcal{M}_{L,\delta}, \mathbf{p}))$ and $\text{error}(\mathcal{N}(L))$ is the corresponding error. These values of $\tilde{b}$ should be approximately the same for different values of L. The results in Table 7.4 illustrate the expected convergence rates for various values of the grading factor δ. These results are also displayed graphically in Figure 7.1, where we plot the error against $\mathcal{N}^{1/2}$, denoted by dofs$^{1/2}$ in the plot. In the semi-logarithmic plot, the curves are roughly straight lines, which indicates exponential convergence rates in excellent agreement with our theoretical results.

8. Concluding remarks. In this paper, we have studied the numerical solution of a class of integro-differential equations of parabolic type of the form (1.1), where the kernel is weakly singular. The first part of this work has focused on the *hp*-DG time-stepping method in the absence of a spatial discretization. We have derived error estimates that are fully explicit in all the parameters of interests. Our estimates show that spectral and exponential convergence can be achieved for smooth and analytic solutions, respectively. We have also shown that exponential convergence rates of convergence can be achieved when temporal singularities near $t = 0$ caused by the weakly singular kernel are resolved using geometrically refined time-steps and linearly increasing polynomial degrees.

In the second part of this paper, we have introduced and analyzed a fully discrete scheme for (6.1)–(6.3); in space we have employed a standard continuous Galerkin finite element method. We have proved that spectral convergence in time and space can be achieved for smooth solutions provided that the approximation orders in time and space are increased. We have also presented fully discrete error estimates on geometrically and nonuniformly graded time-steps.

On each time interval I_n, the *hp*-DG method (2.7) reduces the problem (1.1) to a coupled elliptic system of $p_n + 1$ equations, which is very costly to solve numerically, particularly for large approximation orders. For purely parabolic differential equations, this problem was overcome by the use of complex diagonalization techniques; see [21]. Extensions of these results to problems of the form (6.1)–(6.3) are the subject of ongoing work.

Notice that in this paper, we have only looked at time singularities caused by the weakly singular kernel (1.2), and assumed that u_0 and f are (sufficiently) smooth. The extension of the regularity bounds in (4.1) to the case of nonsmooth initial data remains an open problem.

REFERENCES

[1] H. Brunner and D. Schötzau, *hp-discontinuous Galerkin time-stepping for Volterra integrodifferential equations*, SIAM J. Numer. Anal., 44 (2006), pp. 224–245.

[2] C. Chen, V. Thomée, and L. B. Wahlbin, *Finite element approximation of a parabolic integro-differential equation with a weakly singular kernel*, Math. Comp., 58 (1992), pp. 587–602.

[3] M. Delfour, W. Hager, and F. Trochu, *Discontinuous Galerkin methods for ordinary differential equations*, Math. Comp., 36 (1981), pp. 455–473.

[4] K. Eriksson, C. Johnson, and V. Thomée, *Time discretization of parabolic problems by the discontinuous Galerkin method*, RAIRO Modél. Math. Anal. Numér., 19 (1985), pp. 611–643.

[5] D. Estep, *A posteriori error bounds and global error control for approximation of ordinary differential equations*, SIAM J. Numer. Anal., 32 (1995), pp. 1–48.

[6] A. Friedman and M. Shinbrot, *Volterra integral equations in Banach space*, Trans. Amer. Math. Soc., 126 (1967), pp. 131–179.

[7] M. L. Heard, *An abstract parabolic Volterra integrodifferential equation*, SIAM J. Math. Anal., 13 (1982), pp. 81–105.

[8] C. Johnson, *Error estimates and adaptive time-step control for a class of one-step methods for stiff ordinary differential equations*, SIAM J. Numer. Anal., 25 (1988), pp. 908–926.

[9] S. Larsson, V. Thomée, and L. Wahlbin, *Numerical solution of parabolic integro-differential equations by the discontinuous Galerkin method*, Math. Comp., 67 (1998), pp. 45–71.

[10] P. Lasaint and P.-A. Raviart, *On a finite element method for solving the neutron transport equation*, in Mathematical Aspects of Finite Elements in Partial Differential Equations (Madison, 1974), Academic Press, New York, 1974, pp. 89–123.

[11] W. McLean and K. Mustapha, *A second-order accurate numerical method for a fractional wave equation*, Numer. Math., 105 (2007), pp. 481–510.

[12] W. McLean, I. H. Sloan, and V. Thomée, *Time discretization via Laplace transformation of an integro-differential equation of parabolic type*, Numer. Math., 102 (2006), pp. 497–522.

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

[14] K. Mustapha, *Regularity of solutions to parabolic integro-differential equations*, in preparation.

[15] K. Mustapha and W. McLean, *Discontinuous Galerkin method for an evolution equation with a memory term of positive type*, Math. Comp., 78 (2009), pp. 1975–1995.

[16] K. Mustapha and H. Mustapha, *A second-order accurate numerical method for a semilinear integro-differential equation with a weakly singular kernel*, IMA J. Numer. Anal., 30 (2010), pp. 555–578.

[17] P. W. J. Oliver, *Asymptotics and Special Functions*, Academic Press, San Diego, 1974.

[18] W. Reed and T. Hill, *Triangular Mesh Methods for the Neutron Transport Equation*, Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, NM, 1973.

[19] M. Renardy, W. J. Hrusa, and J. A. Nohel, *Mathematical Problems in Viscoelasticity*, Pitman Monogr. Surveys Pure Appl. Math. 35, Longman Science and Technical, John Wiley and Sons, New York, 1987.

[20] D. Schötzau and C. Schwab, *An hp a-priori error analysis of the DG time-stepping method for initial value problems*, Calcolo, 37 (2000), pp. 207–232.

[21] D. Schötzau and C. Schwab, *Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method*, SIAM J. Numer. Anal., 38 (2000), pp. 837–875.

[22] C. Schwab, *p and hp-Finite Element Methods. Theory and Applications in Solid and Fluid Mechanics*, Oxford University Press, New York, 1998.

[23] R. K. Sinha, R. E. Ewing, and R. D. Lazarov, *Mixed finite element approximations of parabolic integro-differential equations with nonsmooth initial data*, SIAM. J. Numer. Anal., 47 (2009), pp. 3269–3292.

[24] I. H. Sloan and V. Thomée, *Time discretization of an integro-differential equation of parabolic type*, SIAM J. Numer. Anal., 23 (1986), pp. 1052–1061.

[25] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer Ser. Comput. Math. 25, Springer-Verlag, Berlin, 2006.

[26] T. Werder, K. Gerdes, D. Schötzau, and C. Schwab, *hp-discontinuous Galerkin time stepping for parabolic problems*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 6685–6708.

[27] E. G. Yanik and G. Fairweather, *Finite element methods for parabolic and hyperbolic partial integro-differential equations*, Nonlinear Anal., 12 (1988), pp. 785–809.

[28] N. Y. Zhang, *On fully discrete Galerkin approximations for partial integro-differential equations of parabolic type*, Math. Comp., 60 (1993), pp. 133–166.