Adv Comput Math (2016) 42:377–393
DOI 10.1007/s10444-015-9428-x

Convergence and superconvergence analyses of HDG methods for time fractional diffusion problems

Kassem Mustapha[1] · **Maher Nour**[1] · **Bernardo Cockburn**[2]

Received: 17 November 2014 / Accepted: 1 October 2015 /
Published online: 26 October 2015

Abstract We study the hybridizable discontinuous Galerkin (HDG) method for the spatial discretization of time fractional diffusion models with Caputo derivative of order $0 < \alpha < 1$. For each time $t \in [0, T]$, when the HDG approximations are taken to be piecewise polynomials of degree $k \geq 0$ on the spatial domain Ω, the approximations to the exact solution u in the $L_\infty(0, T; L_2(\Omega))$-norm and to ∇u in the $L_\infty(0, T; \mathbf{L}_2(\Omega))$-norm are proven to converge with the rate h^{k+1} provided that u is sufficiently regular, where h is the maximum diameter of the elements of the mesh. Moreover, for $k \geq 1$, we obtain a superconvergence result which allows us to compute, in an elementwise manner, a new approximation for u converging with a rate h^{k+2} (ignoring the logarithmic factor), for quasi-uniform spatial meshes. Numerical experiments validating the theoretical results are displayed.

Keywords Anomalous diffusion · Time fractional · Discontinuous Galerkin methods · Hybridization · Convergence analysis

Communicated by: Jan Hesthaven

Support of the King Fahd University of Petroleum and Minerals (KFUPM) through the project FT131011 is gratefully acknowledged.

✉ Kassem Mustapha
kassem@kfupm.edu.sa

Maher Nour
mnoor@kfupm.edu.sa

Bernardo Cockburn
cockburn@math.umn.edu

[1] Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

[2] School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA

Mathematics Subject Classification (2010) 26A33 · 65M12 · 65M15 · 65N30

1 Introduction

In this paper, we study the method resulting from using exact integration in time and a hybridizable discontinuous Galerkin (HDG) method for the spatial discretization of the following time fractional diffusion model problem:

$$
\begin{aligned}
{}^c\mathrm{D}^{1-\alpha}u(x,t) - \Delta u(x,t) &= f(x,t) \quad \text{for } (x,t) \in \Omega \times (0,T], && \text{(1a)}\\
u(x,t) &= g(x) \quad \text{for } (x,t) \in \partial\Omega \times (0,T], && \text{(1b)}
\end{aligned}
$$

with $u(x,0) = u_0(x)$ for $x \in \Omega$, where Ω is a convex polyhedral domain of $\mathbb{R}^d (d = 1, 2, 3)$ with boundary $\partial\Omega$, f, g and u_0 are given functions assumed to be sufficiently regular such that the solution u of Eq. 1 is in the space $W^{1,1}(0,T; H^2(\Omega))$, (further regularity assumptions will be imposed later), and $T > 0$ is a fixed but arbitrary value. Here, ${}^c\mathrm{D}^{1-\alpha}$ denotes time fractional Caputo derivative of order α defined by

$$
{}^c\mathrm{D}^{1-\alpha}v(t) := \mathcal{I}^\alpha v'(t) := \int_0^t \omega_\alpha(t-s)v'(s)\,ds \quad \text{with } 0 < \alpha < 1, \tag{2}
$$

where v' denotes the time derivative of the function v and $\mathcal{I}^\alpha$ is the Riemann–Liouville (time) fractional integral operator; with $\omega_\alpha(t) := \frac{t^{\alpha-1}}{\Gamma(\alpha)}$ and Γ being the gamma function.

In this work, we investigate a high-order accurate numerical method for the space discretization for problem (1). Using exact integration in time, we propose to deal with the accuracy issue by developing a high-order HDG method that allows for locally varying spatial meshes and approximation orders which are beneficial to handle problems with low regularity. The HDG methods were introduced in [4] in the framework of steady-sate diffusion which share with the classical (hybridized version of the) mixed finite element methods their remarkable convergence and superconvergence properties, [7], as well as the way in which they can be efficiently implemented, [13]. They provide approximations that are more accurate than the ones given by any other DG method for second-order elliptic problems [25]. In [6], a similar method was studied for the fractional subdiffusion problem:

$$
u'(x,t) - \mathrm{D}^{1-\alpha}\Delta u(x,t) = f(x,t) \quad \text{for } (x,t) \in \Omega \times (0,T], \tag{3}
$$

where $\mathrm{D}^{1-\alpha}$ is the Riemann–Liouville fractional time derivative operator,

$$
\mathrm{D}^{1-\alpha}v(t) := (\mathcal{I}^\alpha v(t))' = \frac{\partial}{\partial t}\int_0^t \omega_\alpha(t-s)\,v(s)\,ds\,. \tag{4}
$$

(For other numerical methods of Eq. 3, see [2, 8, 9, 17–19, 22, 23, 29] and related references therein.) When $f \equiv 0$ (that is, homogeneous case), the two representations (1a) and (3) are different ways of writing the same equation, as they are equivalent under reasonable assumptions on the initial data. However, the numerical methods obtained for each representation are formally different. In [6], the authors extended the approach of the error analysis used in [1] for the heat equation by using several important properties of $D^{1-\alpha}$. Indeed, a duality argument was applied (where delicate regularity estimates were required) to prove the superconvergence properties of the method.

We start our work by introducing the spatial semi-discrete HDG method for the model problem (1) in the next section. In order to actually implement the HDG scheme, we discretize in time using a generalized Crank-Nicolson scheme [20]. The existence and uniqueness of the resulting fully discrete scheme will be shown. In Section 3, we prove the main optimal convergence results of the HDG method. Indeed, for each time $t \in [0, T]$, we prove that the error of the HDG approximation to the solution u of problem (1) in the $L_\infty(0, T; L_2(\Omega))$-norm and to the flux $\mathbf{q} := -\nabla u$ in the $L_\infty(0, T; \mathbf{L}_2(\Omega))$-norm converge with order h^{k+1} where k is the polynomial degree and h is the maximum diameter of the elements of the spatial mesh; see Theorem 2. Some important properties of the fractional integral operator $\mathcal{I}^\alpha$ are used in our a priori error analysis. In Section 4, for quasi-uniform meshes and whenever $k \geq 1$, by a simple elementwise postprocessing with a computation cost that is negligible in comparison with that of obtaining the HDG approximate solution, we obtain a better approximation to u converging in the $L_\infty(0, T; L_2(\Omega))$-norm with a rate of order $\sqrt{\log(T/h^{2/(\alpha+1)})}h^{k+2}$; see Theorem 3. Here, we partially rely on the superconvergence analysis of the postprocessed HDG scheme in [6, Section 5]. In Section 5, we present some numerical tests which indicate the validity of our theoretical optimal convergence rates of the HDG scheme as well as the superconvergence rates of the postprocessed HDG scheme.

Here is a brief history of the numerical methods for problem (1) in the existing literature. For the *one dimensional case*, a box-type scheme based on combining order reduction approach and an L_1-discretization was considered in [32]. An explicit finite difference (FD) method was studied in [26]. For an implicit FD scheme in time and Legendre spectral methods, we refer the reader to [15]. An extension of this work was considered in [14], where a time-space spectral method has been proposed and analyzed. An implicit Crank–Nicolson had been considered in [27] where the stability of the proposed scheme was proven. Two finite difference/element approaches were developed in [30]. Therein, the time direction was approximated by the fractional linear multistep method and the space direction was approximated by the standard finite element method (FEM). A compact difference scheme (fourth order in space) was proposed in [33] for solving problem (1) but with a variable diffusion parameter. The unconditional stability and the global convergence of the scheme were shown. In [28], a high-order local DG (LDG) method for space discretization was studied. Optimal convergence rates was proved.

For the *two- (or three-) dimensional cases*, a standard second-order central difference approximation was used in space, and, for the time stepping, two alternating direction implicit (ADI) schemes (L_1-approximation and backward Euler method) were investigated in [31]. A fractional ADI scheme for problem (1a) in 3D was proposed in [3]. Unique solvablity, unconditional stablity and convergence in H^1-norm were shown. A compact fourth order FD method (in space) with operator-splitting techniques was considered in [10]. The Caputo derivative was evaluated by the L_1

approximation, and the second order spatial derivatives were approximated by the fourth-order, compact (implicit) FDs. In [12], the authors developed two simple fully discrete schemes based on piecewise linear Galerkin FEMs in space and implicit backward differences for the time discretizations. Finally, a high-order accurate (variable) time-stepping discontinuous Petrov-Galerkin that allows low regularity combined with standard finite elements in space was investigated recently in [20]. Stability and error analysis were rigourously studied.

2 The HDG method

This section is devoted to defining a scalar approximation $u_h(t)$ to $u(t)$, a vector approximation $\boldsymbol{q}_h(t)$ to the flux $\boldsymbol{q}(t)$, and a scalar approximation $\widehat{u}_h(t)$ to the trace of $u(t)$ on element boundaries for each time $t \in [0, T]$, using a spatial HDG method. We begin by discretizing the domain Ω by a conforming triangulation (for simplicity) $\mathcal{T}_h$ made of simplexes K; we denote by $\partial\mathcal{T}_h$ the set of all the boundaries ∂K of the elements K of $\mathcal{T}_h$. We denote by $\mathcal{E}_h$ the union of faces F of the simplexes K of the triangulation $\mathcal{T}_h$.

Next, we introduce the discontinuous finite element spaces:

$$W_h = \{w \in L^2(\Omega)\colon w|_K \in \mathcal{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\}, \tag{5a}$$
$$\boldsymbol{V}_h = \{\boldsymbol{v} \in [L_2(\Omega)]^d\colon\ \boldsymbol{v}|_K \in [\mathcal{P}_k(K)]^d \quad \forall\, K \in \mathcal{T}_h\}, \tag{5b}$$
$$M_h = \{\mu \in L^2(\mathcal{E}_h)\,:\ \mu|_F \in \mathcal{P}_k(F) \quad \forall\, F \in \mathcal{E}_h\}, \tag{5c}$$

where $\mathcal{P}_k(K)$ is the space of polynomials of total degree at most k in the spatial variable.

To describe our scheme, we rewrite (1a) as a first order system as follows: $\boldsymbol{q} + \nabla u = 0$ and ${}^c\mathrm{D}^{1-\alpha}u + \nabla\cdot\boldsymbol{q} = f$. So, $\boldsymbol{q}$ and u satisfy: for $t \in (0, T]$,

$$(\boldsymbol{q}, \boldsymbol{\phi}) - (u, \nabla\cdot\boldsymbol{\phi}) + \langle u, \boldsymbol{\phi}\cdot\boldsymbol{n}\rangle = 0 \quad \forall\boldsymbol{\phi} \in H(\mathrm{div}, \Omega), \tag{6a}$$
$$\left({}^c\mathrm{D}^{1-\alpha}u, \chi\right) - (\boldsymbol{q}, \nabla\chi) + \langle\boldsymbol{q}\cdot\boldsymbol{n}, \chi\rangle = (f, \chi) \quad \forall\,\chi \in H^1(\Omega)\,. \tag{6b}$$

where $(v, w) := \sum_{K\in\mathcal{T}_h}(v, w)_K$ and $\langle v, w\rangle := \sum_{K\in\mathcal{T}_h}\langle v, w\rangle_{\partial K}$. Throughout the paper, for any domain D in $\mathbb{R}^d$, by $(u, v)_D = \int_D uv\ \ dx$ we denote the L_2-inner product on D. However, we use instead $\langle u, v\rangle_D$ for the L_2-inner product when D is a domain of $\mathbb{R}^{d-1}$. We use $\|\cdot\|_D$ to denote the $L^2(D)$-norm where we drop D when $D = \Omega$. For vector functions $\boldsymbol{v}$ and $\boldsymbol{w}$, the notation is similarly defined with the integrand being the dot product $\boldsymbol{v}\cdot\boldsymbol{w}$. For later use, the norm and semi-norm on any Sobolev space X are denoted by $\|\cdot\|_X$ and $|\cdot|_X$, respectively. We also denote $\|\cdot\|_{X(0,T;Y(\Omega))}$ by $\|\cdot\|_{X(Y)}$.

For each $t > 0$, the HDG method provides approximations $u_h(t) \in W_h$, $\boldsymbol{q}_h(t) \in \boldsymbol{V}_h$, and $\widehat{u}_h(t) \in M_h$ of $u(t)$, $\boldsymbol{q}(t)$, and the trace of $u(t)$, respectively. These are determined by requiring that

$$(\boldsymbol{q}_h, \boldsymbol{r}) - (u_h, \nabla\cdot\boldsymbol{r}) + \langle\widehat{u}_h, \boldsymbol{r}\cdot\boldsymbol{n}\rangle = 0, \quad \forall\,\boldsymbol{r} \in\boldsymbol{V}_h, \tag{7a}$$

$$\left({}^c\mathrm{D}^{1-\alpha}u_h, w\right) - (\boldsymbol{q}_h, \nabla w) + \langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, w\rangle = (f, w), \quad \forall\, w \in W_h, \tag{7b}$$

$$\langle \widehat{u}_h, \mu\rangle_{\partial\Omega} = \langle g, \mu\rangle_{\partial\Omega}, \quad \forall\, \mu \in M_h, \tag{7c}$$

$$\langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, \mu\rangle - \langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, \mu\rangle_{\partial\Omega} = 0, \quad \forall\, \mu \in M_h, \tag{7d}$$

and take the numerical trace for the flux as

$$\widehat{\boldsymbol{q}}_h = \boldsymbol{q}_h + \tau\left(u_h - \widehat{u}_h\right)\boldsymbol{n} \quad \text{on } \partial\mathcal{T}_h, \tag{7e}$$

for some nonnegative stabilization function τ defined on $\partial\mathcal{T}_h$; we assume that, for each element $K \in \mathcal{T}_h$, $\tau|_{\partial K}$ is constant on each of its faces. At $t = 0$, $u_h(0) \in W_h$ approximates the initial solution u_0.

The first two equations are inspired by the weak form of the fractional differential equations satisfied by the exact solution, Eq. 6a. The form of the numerical trace given by Eq. 7d allows us to express $(u_h, \boldsymbol{q}_h, \widehat{\boldsymbol{q}}_h)$ elementwise in terms of $\widehat{u}_h$ and f by using Eqs. 7a, 7b and 7e. Then, the numerical trace $\widehat{u}_h$ is determined by as the solution of the transmission condition (7d), which enforces the single-valuedness of the normal component of the numerical trace $\widehat{\boldsymbol{q}}_h$, and the boundary condition (7c). Thus, the only globally-coupled degrees of freedom are those of $\widehat{u}_h$.

In our experiments, to implement our spatial semi-discrete HDG scheme (7), we use for simplicity a generalized Crank-Nicolson (CN) scheme for time discretization, see [20]. Formally, the CN scheme is second-order accurate provided that the continuous solution is sufficiently regular. To this end, we introduce a uniform partition of the time interval $[0, T]$ given by the points: $t_i = i\delta$ for $i = 0, \cdots, N$, with $\delta = T/N$ being the time-step size. We take δ to be sufficiently small so that the spatial discretizations errors are dominant.

The time-stepping CN combined with the HDG method provides approximations $u_h^j \in W_h$, $\boldsymbol{q}_h^j \in \boldsymbol{V}_h^j$, and $\widehat{u}_h^j \in M_h$ of $u(t_j)$, $\boldsymbol{q}(t_j)$, and the trace of $u(t_j)$, respectively, for $j = 1, \cdots, N$. Starting from $u_h^0 = u_h(0) \approx u_0$, and with appropriate choices of $\boldsymbol{q}_h^0$ and $\widehat{u}_h^0$, our fully discrete scheme is defined by:

$$\begin{aligned}
\left(\boldsymbol{q}_h^{j-\frac{1}{2}}, \boldsymbol{r}\right) - (u_h^{j-\frac{1}{2}}, \nabla\cdot\boldsymbol{r}) + \langle \widehat{u}_h^{j-\frac{1}{2}}, \boldsymbol{r}\cdot\boldsymbol{n}\rangle &= 0, \quad \forall\, \boldsymbol{r} \in \boldsymbol{V}_h,\\
(\mathcal{J}_\alpha \overline{u}_h(t_j), w) - \left(\boldsymbol{q}_h^{j-\frac{1}{2}}, \nabla w\right) + \langle \widehat{\boldsymbol{q}}_h^{j-\frac{1}{2}} \cdot \boldsymbol{n}, w\rangle &= \left(f^{j-\frac{1}{2}}, w\right), \quad \forall\, w \in W_h,\\
\langle \widehat{u}_h^j, \mu\rangle_{\partial\Omega} &= \langle g, \mu\rangle_{\partial\Omega}, \quad \forall\, \mu \in M_h,\\
\langle \widehat{\boldsymbol{q}}_h^j \cdot \boldsymbol{n}, \mu_1\rangle - \langle \widehat{\boldsymbol{q}}_h^j \cdot \boldsymbol{n}, \mu_1\rangle_{\partial\Omega} &= 0, \quad \forall\, \mu_1 \in M_h,
\end{aligned} \tag{8}$$

where $f^{j-\frac{1}{2}} := \frac{1}{2}(f(t_{j-1}) + f(t_j))$, $\widehat{\boldsymbol{q}}_h^j = \boldsymbol{q}_h^j + \tau\left(u_h^j - \widehat{u}_h^j\right)\boldsymbol{n}$ on $\partial\mathcal{T}_h$,

$$\mathcal{J}_\alpha \overline{u}_h(t_j) = \int_{t_{j-1}}^{t_j} \int_0^t \omega_\alpha(t-s)\overline{u}_h(s)\,ds\,dt,$$

with $\overline{u}_h(s) := \delta^{-1}(u_h^i - u_h^{i-1})$ for $s \in (t_{i-1}, t_i)$, $\boldsymbol{q}_h^{j-\frac{1}{2}} := \frac{1}{2}(\boldsymbol{q}_h^j + \boldsymbol{q}_h^{j-1})$, and the functions $u_h^{j-\frac{1}{2}}$, $\widehat{u}_h^{j-\frac{1}{2}}$, and $\widehat{\boldsymbol{q}}_h^{j-\frac{1}{2}}$ are similarly defined.

For each $1 \le j \le N$, Eq. 8 amounts to a square linear system. Thus the existence of the CN HDG solution follows from its uniqueness. We prove the uniqueness by induction hypothesis on j. We let $f^{i-\frac{1}{2}}$(for $1 \le i \le j$) and g be identically zero in Eq. 8, we assume that $(u_h^i, \boldsymbol{q}_h^i, \widehat{u}_h^i) \equiv (0, \boldsymbol{0}, 0)$ for $1 \le i \le j-1$ and the task is to show that this holds true for $i = j$. To do so, choose $\boldsymbol{r} = \boldsymbol{q}_h^j$, $w = u_h^j$, $\mu = \widehat{\boldsymbol{q}}_h^j \cdot \boldsymbol{n}$ and $\mu_1 = \widehat{u}_h^j$ in Eq. 8 and then simplify, yield

$$\begin{aligned} \|\boldsymbol{q}_h\|^2 - \Big(u_h^j, \nabla \cdot \boldsymbol{q}_h^j\Big) + \langle \widehat{u}_h^j, \boldsymbol{q}_h^j \cdot \boldsymbol{n}\rangle &= 0, \\ 2\Big(\mathcal{J}_\alpha \overline{u}_h(t_j), u_h^j\Big) - \Big(\boldsymbol{q}_h^j, \nabla u_h^j\Big) + \langle \widehat{\boldsymbol{q}}_h^j \cdot \boldsymbol{n}, u_h^j\rangle &= 0, \\ \langle \widehat{\boldsymbol{q}}_h^j \cdot \boldsymbol{n}, \widehat{u}_h^j\rangle &= 0. \end{aligned}$$

Since $\Big(u_h^j, \nabla \cdot \boldsymbol{q}_h^j\Big) = \langle u_h^j, \boldsymbol{q}_h^j \cdot \boldsymbol{n}\rangle - \Big(\boldsymbol{q}_h^j, \nabla u_h^j\Big)$, adding the above equations give

$$2\Big(\mathcal{J}^\alpha \overline{u}_h(t_j), u_h^j\Big) + \|\boldsymbol{q}_h^j\|^2 + \langle \widehat{u}_h^j - u_h^j, \Big(\boldsymbol{q}_h^j - \widehat{\boldsymbol{q}}_h^j\Big) \cdot \boldsymbol{n}\rangle = 0\,.$$

Hence, by the induction hypothesis and the identity $\Big(\boldsymbol{q}_h^j - \widehat{\boldsymbol{q}}_h^j\Big) \cdot \boldsymbol{n} = \tau\,\big(u_h^j - \widehat{u}_h^j\big)$ on $\partial\mathcal{T}_h$, we notes that

$$2\int_0^{t_j} (\mathcal{I}^\alpha \overline{u}_h(t), \overline{u}_h(t))\,dt + \|\boldsymbol{q}_h^j\|^2 + \|\sqrt{\tau}(\widehat{u}_h^j - u_h^j)\|_{\partial\mathcal{T}_h}^2 = 0,$$

and therefore, the use of the coercivity property of $\mathcal{I}^\alpha$, Eq. 9, completes the proof.

3 Error estimates

In this section, we carry our a priori error analysis of the HDG method. First, we state the coercivity and continuity properties of $\mathcal{I}^\alpha$ [24, Lemma 3.1] that will be used throughout the paper: with $c_\alpha := \cos(\frac{\alpha\pi}{2})$,

$$\int_0^T (\mathcal{I}^\alpha v(t), v(t))\,dt \ge c_\alpha \int_0^T \|\mathcal{I}^{\frac{\alpha}{2}} v(t)\|^2\,dt \quad \text{for } v \in \mathcal{C}(0,T; L_2(\Omega)), \tag{9}$$

and for $v,\ w \in \mathcal{C}(0,T; L_2(\Omega))$, we have

$$2\Big|\int_0^T (v, \mathcal{I}^\alpha w)\,dt\Big| \le \int_0^T \Big(\frac{1}{c_\alpha^2}(v, \mathcal{I}^\alpha v) + (w, \mathcal{I}^\alpha w)\Big)\,dt\,. \tag{10}$$

We also use the following integral inequality [6, Lemma 4]:

Lemma 1 *For any $t \ge 0$, suppose that $E^2(t) \le A(t) + 2\int_0^t B(s)\,E(s)\,ds$, for some nonnegative functions A and B. Then,*

$$E(T) \le \max_{t\in(0,T)} A^{1/2}(t) + \int_0^T B(s)\,ds \ \text{ for any } \ T > 0.$$

Next, we define the projections which play the comparison function role in the error analysis. For each $t \in (0, T]$, we assume that $\boldsymbol{q}(t) \in [H^1(\mathcal{T}_h)]^d$ and

$u(t) \in H^1(\mathcal{T}_h)$, where $H^1(\mathcal{T}_h) = \prod_{K\in\mathcal{T}_h} H^1(K)$, the projections $\boldsymbol{\Pi}_V \boldsymbol{q}(t) \in \boldsymbol{V}_h$ and $\Pi_W u(t) \in W_h$ are defined by: on each simplex $K \in \mathcal{T}_h$ and for all faces F of K,

$$(\boldsymbol{\Pi}_V \boldsymbol{q}(t), \boldsymbol{v})_K = (\boldsymbol{q}(t), \boldsymbol{v})_K, \tag{11a}$$

$$(\Pi_W u(t), w)_K = (u(t), w)_K, \tag{11b}$$

$$\langle \boldsymbol{\Pi}_V \boldsymbol{q}(t) \cdot \boldsymbol{n} + \tau \Pi_W u(t), \mu \rangle_F = \langle \boldsymbol{q}(t) \cdot \boldsymbol{n} + \tau u(t), \mu \rangle_F, \tag{11c}$$

for al $\boldsymbol{v} \in [\mathcal{P}_{k-1}(K)]^d$, $w \in \mathcal{P}_{k-1}(K)$ and $\mu \in \mathcal{P}_k(F)$. This projection introduced in [5] to study HDG methods for the steady-state diffusion problem and also used in the error analyses of HDG methods for classical diffusion [1] as well as for fractional subdiffusion [6] problems. As mentioned in [5], the projection $\boldsymbol{\Pi}_V$ depends not only on **q**, but rather on both **q** and u. Similarly for the projection Π_W. Hence the notations $\boldsymbol{\Pi}_V$ and Π_W are somewhat misleading but convenient.

Its approximation properties are described in the following result.

Theorem 1 *([5]) Suppose* $\tau|_{\partial K}$ *is nonnegative and* $\tau_K^{\max} := \max \tau|_{\partial K} > 0$. *Then the system (11) is uniquely solvable for* $\boldsymbol{\Pi}_V \boldsymbol{q}$ *and* $\Pi_W u$. *Furthermore, there is a constant* C *independent of K and* τ *such that for each* $t \in (0, T]$,

$$\|e_{\boldsymbol{q}}(t)\|_K \le C\, h_K^{k+1} \Big(|\boldsymbol{q}(t)|_{\boldsymbol{H}^{k+1}(K)} + \tau_K^* \, |u(t)|_{H^{k+1}(K)} \Big),$$

$$\|e_u(t)\|_K \le C\, h_K^{k+1} \Big(|u(t)|_{H^{k+1}(K)} + |\nabla \cdot \boldsymbol{q}(t)|_{H^k(K)} / \tau_K^{\max} \Big),$$

where $e_{\boldsymbol{q}} := \boldsymbol{\Pi}_V \boldsymbol{q} - \boldsymbol{q}$, $e_u := \Pi_W u - u$, *and* h_K *is the diameter of the spatial mesh element* K. *Here* $\tau_K^* := \max \tau|_{\partial K \setminus F^*}$, *where* F^* *is a face of* K *at which* $\tau|_{\partial K}$ *is maximum.*

Note that the approximation error of the projection is of order $k+1$ provided that the stabilization function is such that both τ_K^* and $1/\tau_K^{\max}$ are uniformly bounded and the exact solution is sufficiently regular.

Thus, the main task now is to estimate the terms $\varepsilon_u := \Pi_W u - u_h$ and $\boldsymbol{\varepsilon}_{\mathbf{q}} := \boldsymbol{\Pi}_V \boldsymbol{q} - \boldsymbol{q}_h$. For convenience, we further introduce the following notations: $\varepsilon_{\widehat{u}} := P_M u - \widehat{u}_h$ and $\boldsymbol{\varepsilon}_{\widehat{q}} := \boldsymbol{P}_M \boldsymbol{q} - \widehat{\boldsymbol{q}}_h$ where P_M denotes the L^2-orthogonal projection onto M_h, and $\boldsymbol{P}_M$ denotes the vector-valued projection each of whose components are equal to P_M. For later use, for each $t \in (0, T]$, Eq. 11c is equivalent to

$$\langle \boldsymbol{\Pi}_V \boldsymbol{q}(t) \cdot \boldsymbol{n} + \tau \Pi_W u(t) - P_M(\boldsymbol{q}(t) \cdot \boldsymbol{n}) - \tau P_M u(t), \mu \rangle_F \quad \forall\, \mu \in \mathcal{P}_k(F)\,.$$

Since $\boldsymbol{\Pi}_V \boldsymbol{q}(t) \cdot \boldsymbol{n} + \tau \Pi_W u(t) - P_M(\boldsymbol{q}(t) \cdot \boldsymbol{n}) - \tau P_M u(t) \in \mathcal{P}_k(F)$,

$$\boldsymbol{\Pi}_V \boldsymbol{q}(t) \cdot \boldsymbol{n} + \tau \Pi_W u(t) - P_M(\boldsymbol{q}(t) \cdot \boldsymbol{n}) - \tau P_M u(t) = 0 \quad \text{for each } t \in (0, T]\,. \tag{13}$$

The projection of the errors satisfy the equations stated in the next lemma.

Lemma 2 *For each $t > 0$, we have*

$$(\boldsymbol{\varepsilon_q}, \boldsymbol{r}) - (\varepsilon_u, \nabla \cdot \boldsymbol{r}) + \langle \varepsilon_{\widehat{u}}, \boldsymbol{r} \cdot \boldsymbol{n} \rangle = (e_{\boldsymbol{q}}, \boldsymbol{r}), \quad \forall \boldsymbol{r} \in \boldsymbol{V}_h \tag{14a}$$

$$(\mathcal{I}^\alpha \varepsilon_u', w) - (\boldsymbol{\varepsilon_q}, \nabla w) + \langle \boldsymbol{\varepsilon}_{\widehat{q}} \cdot \boldsymbol{n}, w \rangle = (\mathcal{I}^\alpha e_u', w), \quad \forall w \in W_h \tag{14b}$$

$$\langle \varepsilon_{\widehat{u}}, \mu \rangle_{\partial\Omega} = 0, \quad \forall \mu \in M_h \tag{14c}$$

$$\langle \boldsymbol{\varepsilon}_{\widehat{q}} \cdot \boldsymbol{n}, \mu \rangle - \langle \boldsymbol{\varepsilon}_{\widehat{q}} \cdot \boldsymbol{n}, \mu \rangle_{\partial\Omega} = 0, \quad \forall \mu \in M_h \tag{14d}$$

and we also have

$$\boldsymbol{\varepsilon}_{\widehat{q}} \cdot \boldsymbol{n} := \boldsymbol{\varepsilon_q} \cdot \boldsymbol{n} + \tau(\varepsilon_u - \varepsilon_{\widehat{u}}) \quad \text{on } \partial\mathcal{T}_h. \tag{14e}$$

Proof From (6), we recall that $\boldsymbol{q}$ and u satisfy the equations

$$(\boldsymbol{q}, \boldsymbol{r}) - (u, \nabla \cdot \boldsymbol{r}) + \langle u, \boldsymbol{r} \cdot \boldsymbol{n} \rangle = 0 \quad \forall \boldsymbol{r} \in \boldsymbol{V}_h,$$

$$(\mathcal{I}^\alpha u', w) - (\boldsymbol{q}, \nabla w) + \langle \boldsymbol{q} \cdot \boldsymbol{n}, w \rangle = (f, w) \quad \forall w \in W_h\,.$$

By the equalities $\boldsymbol{q} = \boldsymbol{\Pi}_V \boldsymbol{q} - e_{\boldsymbol{q}}$ and $u = \Pi_W u - e_u$, the fact that P_M is the L^2-projection into M_h and (11), we get

$$(\boldsymbol{\Pi}_V \boldsymbol{q}, \boldsymbol{r}) - (u, \nabla \cdot \boldsymbol{r}) + \langle P_M u, \boldsymbol{r} \cdot \boldsymbol{n} \rangle = (e_{\boldsymbol{q}}, \boldsymbol{r}),$$

$$(\mathcal{I}^\alpha (\Pi_W u)', w) - (\boldsymbol{q}, \nabla w) + \langle \boldsymbol{\Pi}_V \boldsymbol{q} \cdot \boldsymbol{n} + \tau(\Pi_W u - P_M u), w \rangle = (f + \mathcal{I}^\alpha e_u', w),$$

$\forall \boldsymbol{r} \in \boldsymbol{V}_h$ and $\forall w \in W_h$, given that, for each element $K \in \mathcal{T}_h$, τ is constant on each face F of K. Hence, by Eq. 11a and Eq. 11b, we observe that

$$(\boldsymbol{\Pi}_V \boldsymbol{q}, \boldsymbol{r}) - (\Pi_W u, \nabla \cdot \boldsymbol{r}) + \langle P_M u, \boldsymbol{r} \cdot \boldsymbol{n} \rangle = (e_{\boldsymbol{q}}, \boldsymbol{r}), \quad \forall \boldsymbol{r} \in \boldsymbol{V}_h \tag{16}$$

$$\begin{aligned}(\mathcal{I}^\alpha (\Pi_W u)', w) - (\boldsymbol{\Pi}_V \boldsymbol{q}, \nabla w) + \langle \boldsymbol{\Pi}_V \boldsymbol{q} \cdot \boldsymbol{n} + \tau(\Pi_W u - P_M u), w \rangle \\ = (f + \mathcal{I}^\alpha e_u', w), \quad \forall w \in W_h\,.\end{aligned} \tag{17}$$

Subtracting the Eqs. 7a and 7b from Eqs. 16 and 17, respectively, we obtain Eqs. 14a and 14b, respectively. The Eq. 14c follows directly from Eqs. 7c and 1b

By the definition of $\boldsymbol{\varepsilon}_{\widehat{q}}$ and since P_M is the L^2-projection into M_h, we have

$$\begin{aligned}\langle \boldsymbol{\varepsilon}_{\widehat{q}} \cdot \boldsymbol{n}, \mu \rangle - \langle \boldsymbol{\varepsilon}_{\widehat{q}} \cdot \boldsymbol{n}, \mu \rangle_{\partial\Omega} &= \langle (\boldsymbol{P}_M \boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n}, \mu \rangle - \langle (\boldsymbol{P}_M \boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n}, \mu \rangle_{\partial\Omega} \\ &= \langle (\boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n}, \mu \rangle - \langle (\boldsymbol{q} - \widehat{\boldsymbol{q}}_h) \cdot \boldsymbol{n}, \mu \rangle_{\partial\Omega} \\ &= [\langle \boldsymbol{q} \cdot \boldsymbol{n}, \mu \rangle - \langle \boldsymbol{q} \cdot \boldsymbol{n}, \mu \rangle_{\partial\Omega}] - [\langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, \mu \rangle - \langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, \mu \rangle_{\partial\Omega}] = 0\end{aligned}$$

where in the last equality we used that $\boldsymbol{q}$ is in $\mathbf{H}(\text{div}, \Omega)$ and Eq. 7d. Thus, the identity (14d) holds. For the proof of Eq. 14e,

$$\begin{aligned}\boldsymbol{\varepsilon}_{\widehat{q}} \cdot \boldsymbol{n} &= P_M(\boldsymbol{q} \cdot \boldsymbol{n}) - (\boldsymbol{q}_h \cdot \boldsymbol{n} + \tau\,(u_h - \widehat{u}_h)) && \text{by (7e)}, \\ &= (\boldsymbol{\Pi}_V \boldsymbol{q} \cdot \boldsymbol{n} + \tau\,(\Pi_W u - P_M u)) - (\boldsymbol{q}_h \cdot \boldsymbol{n} + \tau\,(u_h - \widehat{u}_h)) && \text{by (13)}, \\ &= \boldsymbol{\varepsilon}_{\mathbf{q}} \cdot \boldsymbol{n} + \tau(\varepsilon_u - \varepsilon_{\widehat{u}})\,.\end{aligned}$$

□

Lemma 3 *Let* $S_h := \|\sqrt{\tau}(\varepsilon_u - \varepsilon_{\widehat{u}})\|_{\partial\mathcal{T}_h}$. *For* $T > 0$,

$$\int_0^T \left(\mathcal{I}^\alpha \varepsilon_u', \varepsilon_u'\right)\, dt + \|\boldsymbol{\varepsilon_q}(T)\|^2 \;+\; S_h^2(T) \le \|\boldsymbol{\varepsilon_q}(0)\|^2 + S_h^2(0) + \frac{1}{c_\alpha^2}\int_0^T (\mathcal{I}^\alpha e_u', e_u')\, dt + 2\int_0^T (e_{\boldsymbol{q}}', \boldsymbol{\varepsilon_q})\, dt\,.$$

Proof Since $(\varepsilon_u, \nabla\cdot\boldsymbol{r}) = -(\nabla\varepsilon_u, \boldsymbol{r}) + \langle\varepsilon_u, \boldsymbol{r}.\boldsymbol{n}\rangle$, Eq. 14a can be rewritten as:

$$(\boldsymbol{\varepsilon}_{\mathbf{q}}, \boldsymbol{r}) + (\nabla\varepsilon_u, \boldsymbol{r}) + \langle\varepsilon_{\widehat{u}} - \varepsilon_u, \boldsymbol{r}\cdot\boldsymbol{n}\rangle = (e_{\boldsymbol{q}}, \boldsymbol{r})\,.$$

A time differentiation of both sides yields,

$$\left(\boldsymbol{\varepsilon}_{\mathbf{q}}', \boldsymbol{r}\right) + \left(\nabla\varepsilon_u', \boldsymbol{r}\right) + \langle\varepsilon_{\widehat{u}}' - \varepsilon_u', \boldsymbol{r}\cdot\boldsymbol{n}\rangle = \left(e_{\boldsymbol{q}}', \boldsymbol{r}\right)\,.$$

Setting $\boldsymbol{r} = \boldsymbol{\varepsilon}_{\mathbf{q}}$ and choosing $w = \varepsilon_u'$ in Eq. 14b, we observe that

$$\begin{aligned}\left(\boldsymbol{\varepsilon}_{\mathbf{q}}', \boldsymbol{\varepsilon}_{\mathbf{q}}\right) + \left(\nabla\varepsilon_u', \boldsymbol{\varepsilon}_{\mathbf{q}}\right) + \langle\varepsilon_{\widehat{u}}' - \varepsilon_u', \boldsymbol{\varepsilon}_{\mathbf{q}}\cdot\boldsymbol{n}\rangle &= \left(e_{\boldsymbol{q}}', \boldsymbol{\varepsilon}_{\mathbf{q}}\right)\,,\\ \left(\mathcal{I}^\alpha\varepsilon_u', \varepsilon_u'\right) - \left(\boldsymbol{\varepsilon}_{\mathbf{q}}, \nabla\varepsilon_u'\right) + \langle\boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}, \varepsilon_u'\rangle &= \left(\mathcal{I}^\alpha e_u', \varepsilon_u'\right)\,.\end{aligned}$$

Combining the above two equations and using $(\boldsymbol{\varepsilon}_{\mathbf{q}}', \boldsymbol{\varepsilon}_{\mathbf{q}}) = \frac{1}{2}\frac{d}{dt}\|\boldsymbol{\varepsilon}_{\mathbf{q}}\|^2$, we obtain

$$\left(\mathcal{I}^\alpha\varepsilon_u', \varepsilon_u'\right) + \frac{1}{2}\frac{d}{dt}\|\boldsymbol{\varepsilon}_{\mathbf{q}}\|^2 + \psi_h = \left(\mathcal{I}^\alpha e_u', \varepsilon_u'\right) + \left(e_{\boldsymbol{q}}', \boldsymbol{\varepsilon}_{\mathbf{q}}\right)\,, \tag{18}$$

where

$$\psi_h = \langle\varepsilon_{\widehat{u}}' - \varepsilon_u', \boldsymbol{\varepsilon}_{\mathbf{q}}\cdot\boldsymbol{n}\rangle + \langle\boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}, \varepsilon_u'\rangle\,.$$

A time differentiation of Eq. 14c followed by choosing $\mu = \boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}$ and then using Eq. 14d yields $\langle\boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}, \varepsilon_{\widehat{u}}'\rangle_{\partial\Omega} = \langle\boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}, \varepsilon_{\widehat{u}}'\rangle = 0$. Thus, by Eq. 14e,

$$\psi_h = \langle\varepsilon_{\widehat{u}}' - \varepsilon_u', (\boldsymbol{\varepsilon}_{\mathbf{q}} - \boldsymbol{\varepsilon}_{\widehat{q}})\cdot\boldsymbol{n}\rangle = \langle\tau(\varepsilon_u' - \varepsilon_{\widehat{u}}'), (\varepsilon_u - \varepsilon_{\widehat{u}})\rangle = \frac{1}{2}\frac{d}{dt}S_h^2(t)\,. \tag{19}$$

Now, integrating Eq. 18 over the time interval $[0, T]$ and using Eq. 19, we get

$$\int_0^T(\mathcal{I}^\alpha\varepsilon_u', \varepsilon_u')\,dt + \frac{1}{2}\int_0^T\frac{d}{dt}\left[\|\boldsymbol{\varepsilon}_{\mathbf{q}}\|^2 + S_h^2\right]dt = \int_0^T(\mathcal{I}^\alpha e_u', \varepsilon_u')\,dt + \int_0^T(e_{\boldsymbol{q}}', \boldsymbol{\varepsilon}_{\mathbf{q}})\,dt\,.$$

Therefore,

$$\begin{aligned}&2\int_0^T\left(\mathcal{I}^\alpha\varepsilon_u', \varepsilon_u'\right)\,dt + \|\boldsymbol{\varepsilon}_{\mathbf{q}}(T)\|^2 + S_h^2(T)\\ &\quad= \|\boldsymbol{\varepsilon}_{\mathbf{q}}(0)\|^2 + S_h^2(0) + 2\int_0^T\left(\mathcal{I}^\alpha e_u', \varepsilon_u'\right)dt + 2\int_0^T\left(e_{\boldsymbol{q}}', \boldsymbol{\varepsilon}_{\mathbf{q}}\right)dt\,.\end{aligned} \tag{20}$$

An application of the continuity property of the fractional derivative operator $\mathcal{I}^\alpha$, Eq. 10, yields

$$2\left|\int_0^T\left(\mathcal{I}^\alpha e_u', \varepsilon_u'\right)\,dt\right| \le \frac{1}{c_\alpha^2}\int_0^T\left(\mathcal{I}^\alpha e_u', e_u'\right)\,dt + \int_0^T\left(\mathcal{I}^\alpha\varepsilon_u', \varepsilon_u'\right)\,dt\,.$$

Finally, inserting this in Eq. 20 and simplifying completes the proof. □

Next, we show the main error bounds of the HDG method. For the remaining part of the paper, we choose $u_h(0) = \Pi_W u_0$ and so, $\varepsilon_u(0) = 0$.

Theorem 2 *Assume that* $u \in \mathcal{C}^1(0, T; H^{k+1}(\Omega))$ *and* $\boldsymbol{q} \in \mathcal{C}^1(0, T; \boldsymbol{H}^{k+1}(\Omega))$. *Assume also that* τ_K^* *and* $1/\tau_K^{\max}$ *are bounded by* C. *Then we have that*

$$\|(u - u_h)(T)\| + \|(\boldsymbol{q} - \boldsymbol{q}_h)(T)\| \le C_1(1 + T)\, h^{k+1}.$$

The constant C_1 *only depends on* C, α, $\|u\|_{\mathcal{C}^1(H^{k+1})}$, *and on* $\|\boldsymbol{q}\|_{\mathcal{C}^1(\boldsymbol{H}^{k+1})}$.

Proof From the decompositions: $u - u_h = \varepsilon_u - e_u$ and $\boldsymbol{q} - \boldsymbol{q}_h = \boldsymbol{\varepsilon}_{\mathbf{q}} - e_{\boldsymbol{q}}$, and the error projection in Theorem 1, we have

$$\|(u - u_h)(T)\| + \|(\boldsymbol{q} - \boldsymbol{q}_h)(T)\| \le C_1\, h^{k+1} + \|\varepsilon_u(T)\| + \|\boldsymbol{\varepsilon}_{\mathbf{q}}(T)\|\,.$$

The task now is to estimate $\|\varepsilon_u(T)\|$ and $\|\boldsymbol{\varepsilon}_{\mathbf{q}}(T)\|$. From Lemma 3, for $t \ge 0$, we have $E^2(t) \le A(t) + 2\int_0^t B(s)\, E(s)\, ds$ where

$$A(t) := \|\boldsymbol{\varepsilon}_{\mathbf{q}}(0)\|^2 + S_h^2(0) + \frac{1}{c_\alpha^2}\int_0^t (\mathcal{I}^\alpha e_u', e_u')\, ds, \quad B(t) := \|e_{\boldsymbol{q}}'(t)\|,$$

$$E(t) := \left(\|\boldsymbol{\varepsilon}_{\mathbf{q}}(t)\|^2 + \int_0^t (\mathcal{I}^\alpha \varepsilon_u', \varepsilon_u')\, ds\right)^{\frac{1}{2}}.$$

Thus, an application of the integral inequality in Lemma 1, yields

$$E(T) \le \max_{t\in(0,T)} A^{\frac{1}{2}}(t) + \int_0^T B(s)\, ds \ \text{ for any } \ T > 0.$$

Hence,

$$\|\boldsymbol{\varepsilon}_{\mathbf{q}}(T)\|^2 + \int_0^T (\mathcal{I}^\alpha \varepsilon_u', \varepsilon_u')\, ds \le C\Bigg(\|\boldsymbol{\varepsilon}_{\mathbf{q}}(0)\|^2 + S_h^2(0) + \int_0^T \left(\frac{1}{c_\alpha^2}\|\mathcal{I}^\alpha e_u'\|\,\|e_u'\| + T\|e_{\boldsymbol{q}}'\|^2\right) ds\Bigg)\,. \tag{21}$$

However, since $\varepsilon_u(t) = \int_0^t \varepsilon_u'(s)\, ds = \mathcal{I}^{1-\frac{\alpha}{2}}(\mathcal{I}^{\frac{\alpha}{2}}\varepsilon_u')(t)$ because $\varepsilon_u(0)$, by the Cauchy-Schwarz inequality and the coercivity property of the operator $\mathcal{I}^\alpha$, Eq. 9,

$$\begin{aligned}
\|\varepsilon_u(t)\|^2 &\le \left(\int_0^t \omega_{1-\frac{\alpha}{2}}(t-s)\, \|\mathcal{I}^{\frac{\alpha}{2}}\varepsilon_u'(s)\|\, ds\right)^2 \\
&\le \int_0^t \omega_{1-\frac{\alpha}{2}}^2(s)\, ds \int_0^t \|\mathcal{I}^{\frac{\alpha}{2}}\varepsilon_u'(s)\|^2\, ds \\
&= \frac{t^{1-\alpha}}{(1-\alpha)\Gamma^2(1-\frac{\alpha}{2})}\int_0^t \|\mathcal{I}^{\frac{\alpha}{2}}\varepsilon_u'(s)\|^2\, ds \\
&\le \frac{t^{1-\alpha}}{(1-\alpha)\Gamma^2(1-\frac{\alpha}{2})\, c_\alpha}\int_0^t (\mathcal{I}^\alpha \varepsilon_u', \varepsilon_u')\, ds\,.
\end{aligned}$$

Therefore, combining Eq. 21 with the above bound, and apply Theorem 1 for the time derivative error projections e_u' and $e_{\boldsymbol{q}}'$, we obtain

$$\|\boldsymbol{\varepsilon}_{\mathbf{q}}(t)\|^2 + \|\varepsilon_u(t)\|^2 \leq C_1^2(1+T)^2\Big(\|\boldsymbol{\varepsilon}_{\mathbf{q}}(0)\|^2 + S_h^2(0) + h^{2k+2}\Big).$$

To complete the proof, we need to bound $\|\boldsymbol{\varepsilon}_{\mathbf{q}}(0)\|^2 + S_h^2(0)$.

Since $(\varepsilon_u, \nabla\cdot\boldsymbol{r}) = -(\nabla\varepsilon_u, \boldsymbol{r}) + \langle\varepsilon_u, \boldsymbol{r}.\boldsymbol{n}\rangle$, setting $\boldsymbol{r} = \boldsymbol{\varepsilon}_{\mathbf{q}}$ in Eq. 14a and $w = \varepsilon_u'$ in Eq. 14b yield

$$\|\boldsymbol{\varepsilon}_{\mathbf{q}}\|^2 + (\nabla\varepsilon_u, \boldsymbol{\varepsilon}_{\mathbf{q}}) + \langle\varepsilon_{\widehat{u}} - \varepsilon_u, \boldsymbol{\varepsilon}_{\mathbf{q}}\cdot\boldsymbol{n}\rangle = (e_{\boldsymbol{q}}, \boldsymbol{\varepsilon}_{\mathbf{q}}),$$

$$\big(\mathcal{I}^\alpha\varepsilon_u', \varepsilon_u\big) - (\boldsymbol{\varepsilon}_{\mathbf{q}}, \nabla\varepsilon_u) + \langle\boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}, \varepsilon_u\rangle = \big(\mathcal{I}^\alpha e_u', \varepsilon_u\big).$$

Adding the above equations, and using $\langle\boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}, \varepsilon_{\widehat{u}}\rangle = 0$ (this follows by choosing $\mu = \boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}$ in Eq. 14c and $\mu = \varepsilon_{\widehat{u}}$ in Eq. 14d) and (14e), we obtain

$$\big(\mathcal{I}^\alpha\varepsilon_u', \varepsilon_u\big) + \|\boldsymbol{\varepsilon}_{\mathbf{q}}\|^2 + S_h^2 = \big(\mathcal{I}^\alpha e_u', \varepsilon_u\big) + (e_{\boldsymbol{q}}, \boldsymbol{\varepsilon}_{\mathbf{q}}).$$

Now, integrating over the time interval $[0, t]$, observing that

$$\int_0^t \big(\mathcal{I}^\alpha\varepsilon_u', \varepsilon_u\big)\,ds = \int_0^t \Big(\mathrm{D}^{1-\alpha}\varepsilon_u, \varepsilon_u\Big)\,ds \geq 0,$$

(in the first equality we used $\varepsilon_u(0) = 0$ and the last inequality follows from the nonnegativity property of the Riemann–Liouville fractional derivative operator $\mathrm{D}^{1-\alpha}$, see [16, Section 2]) and using the inequality $(e_{\boldsymbol{q}}, \boldsymbol{\varepsilon}_{\mathbf{q}}) \leq \frac{1}{2}\|e_{\boldsymbol{q}}\|^2 + \frac{1}{2}\|\boldsymbol{\varepsilon}_{\mathbf{q}}\|^2$, we get

$$\int_0^t [\frac{1}{2}\|\boldsymbol{\varepsilon}_{\mathbf{q}}\|^2 + S_h^2]\,ds \leq \int_0^t \Big(\|\mathcal{I}^\alpha e_u'\|\,\|\varepsilon_u\| + \frac{1}{2}\|e_{\boldsymbol{q}}\|^2\Big)\,ds.$$

Therefore, by the mean value theorem for integrals, there exist $t^*, \tilde{t} \in (0, t)$ such that

$$t\Big(\frac{1}{2}\|\boldsymbol{\varepsilon}_{\mathbf{q}}(t^*)\|^2 + S_h^2(t^*)\Big) \leq t\Big(\|\varepsilon_u(\tilde{t})\| \max_{s\in(0,t)} \|\mathcal{I}^\alpha e_u'(s)\| + \frac{1}{2}\max_{s\in(0,t)} \|e_{\boldsymbol{q}}(s)\|^2\Big).$$

Finally, dividing by t, taking the limit when t goes to zero, and using again the fact that $\varepsilon_u(0) = 0$, we observe that $\|\boldsymbol{\varepsilon}_{\mathbf{q}}(0)\|^2 + S_h^2(0) \leq \|e_{\boldsymbol{q}}(0)\|^2 \leq C_1 h^{2k+2}$ by the error estimate of $e_{\boldsymbol{q}}$ given in Theorem 1. The proof is now complete. □

4 Superconvergence and post-processing

In this section, we seek a better approximation to u by means of an element-by-element postprocessing. We begin by describing such approximation, then we show how to get our superconvergence result by a duality argument.

Following [1, 11], for each $t \in [0, T]$, we define the postprocessed HDG solution $u_h^\star(t) \in \mathcal{P}_{k+1}(K)$ to $u(t)$ for each simplex $K \in \mathcal{T}_h$, as follows:

$$(u_h^\star(t), 1)_K = (u_h(t), 1)_K \tag{22a}$$

$$(\nabla u_h^\star(t), \nabla w)_K = -(\boldsymbol{q}_h(t), \nabla w)_K \qquad \forall\, w \in \mathcal{P}_{k+1}(K). \tag{22b}$$

Since (22) amounts to a square linear system (for each fixed $t \in (0, T]$), the existence of the postprocessed HDG solution follows from its uniqueness. To this end, we let

$u_h(t)$ and $\boldsymbol{q}_h(t)$ to be identically zero in (22). The task is to show that $u_h^\star(t) \equiv 0$ for each $t \in (0, T]$. We choose $w = u_h^\star(t)$ in Eq. 22b and observe that $u_h^\star(t)$ is equal to a constant c_0 on K. Hence, by Eq. 22a, it is easy to see that $c_0 = 0$.

For showing the superconvergence property of $u_h^\star$, splitting the postprocessed error as: $u - u_h^\star = (u - P_{k+1}u) + P_0\zeta + (\zeta - P_0\zeta)$ where $\zeta = P_{k+1}u - u_h^\star$ and P_ℓ (for $\ell \geq 0$) be the $L^2(\Omega)$-projection into the space of functions which are polynomials of total degree $\leq \ell$ on each element $K \in \mathcal{T}_h$. Hence, by the triangle inequality and the error properties of the projection P_ℓ,

$$\|u - u_h^\star\| \leq C\, h_K^{k+2}\, |u|_{H^{k+2}(K)} + \|P_0\zeta\|_K + Ch\|\nabla\zeta\|_K\,. \tag{23}$$

By the definition of $u_h^\star$, Eqs. 22a and 22b, and the definition of the projection operator Π_W, Eq. 11b, we have

$$\|P_0\zeta\|_K^2 = (P_{k+1}u - u_h, P_0\zeta)_K = (P_{k+1}u - u, P_0\zeta)_K + (P_0\varepsilon_u, P_0\zeta)_K$$

$$\|\nabla\zeta\|_K^2 = (\nabla P_{k+1}u + \boldsymbol{q}_h, \nabla\zeta)_K = (\nabla(P_{k+1}u - u), \nabla\zeta)_K - (\mathbf{q} - \boldsymbol{q}_h, \nabla\zeta)_K\,.$$

We combine the above three equations and then we apply the Cauchy-Schwarz and simplify to observe that

$$\|(u - u_h^\star)(T)\|_K \leq C\, h_K^{k+2}\, |u(T)|_{H^{k+2}(K)} + \|P_0\varepsilon_u(T)\|_K + C\,h\,\|(\mathbf{q} - \boldsymbol{q}_h)(T)\|_K\,. \tag{24}$$

The main task now is to show that the term $\|P_0\varepsilon_u(T)\|$ is of order $O(h^{k+2})$. Then the postprocessed approximation $u_h^\star$ would converge faster than the original approximation u_h. Noting that $\|P_0\varepsilon_u(T)\| = \sup_{\Theta \in C_0^\infty(\Omega)} \frac{(P_0\varepsilon_u(T), \Theta)}{\|\Theta\|}$. To estimate the expression $(P_0\varepsilon_u(T), \Theta)$, we use the traditional duality approach by using the solution of the dual problem

$$\boldsymbol{\Phi} + \nabla\Psi = 0 \;\text{ and }\; (\mathcal{I}^{\alpha*}\Psi)' - \nabla\cdot\boldsymbol{\Phi} = 0 \;\text{ on }\; \Omega \times (0, T), \tag{25}$$

with $\Psi = 0$ on $\partial\Omega \times (0, T)$ and $\Psi(T) = \Theta$ on Ω, where $\mathcal{I}^{\alpha*}$ is the adjoint operator of $\mathcal{I}^\alpha$ defined by [23]:

$$\mathcal{I}^{\alpha*}\psi(t) = \int_t^T \omega_\alpha(s - t)\,\psi(s)\,ds\,.$$

Integrating $(\mathcal{I}^{\alpha*}\Psi)' - \nabla\cdot\boldsymbol{\Phi} = 0$ over the time interval (t, T), we obtain

$$\mathcal{I}^{\alpha*}\Psi(t) + \int_t^T \nabla\cdot\boldsymbol{\Phi}(s)\,ds = 0\,. \tag{26}$$

We define now the adjoint $\mathrm{D}^{\alpha*}$ of the Riemann–Liouville fractional derivative operator D^α(see (4) for the definition of D^α) as follows [23]: for $t \in (0, T)$,

$$\mathrm{D}^{\alpha*}v(t) = -\frac{\partial}{\partial t}\int_t^T \omega_{1-\alpha}(s - t)\,v(s)\,ds \quad \text{for any } v \in \mathcal{C}^1(0, T)\,.$$

Since $\int_t^q \omega_\alpha(s - t)\,\omega_{1-\alpha}(q - s)\,ds = 1$, it is easy to see that $\mathcal{I}^{\alpha*}$ is the *right-inverse* of $\mathrm{D}^{\alpha*}$, that is, $\mathrm{D}^{\alpha*}(\mathcal{I}^{\alpha*}\Psi)(t) = \Psi(t)$. Hence, using this after applying the operator $\mathrm{D}^{\alpha*}$ to both sides of Eq. 26, yields

$$\Psi(t) + \mathrm{D}^{\alpha*}\left(\int_t^T \nabla\cdot\boldsymbol{\Phi}(q)\,dq\right) = 0\,. \tag{27}$$

However, since

$$\begin{aligned}
\mathrm{D}^{\alpha*}\left(\int_t^T \nabla\cdot\boldsymbol{\Phi}(s)\,ds\right) &= -\frac{\partial}{\partial t}\int_t^T \omega_{1-\alpha}(s-t)\int_s^T \nabla\cdot\boldsymbol{\Phi}(q)\,dq\,ds\\
&= -\frac{\partial}{\partial t}\int_t^T \nabla\cdot\boldsymbol{\Phi}(q)\int_t^q \omega_{1-\alpha}(s-t)\,ds\,dq\\
&= -\frac{\partial}{\partial t}\int_t^T \omega_{2-\alpha}(q-t)\,\nabla\cdot\boldsymbol{\Phi}(q)\,dq\\
&= \int_t^T \omega_{1-\alpha}(s-t)\,\nabla\cdot\boldsymbol{\Phi}(s)\,ds,
\end{aligned}$$

differentiating both sides of Eq. 27 with respect to t, yield $\Psi' - \nabla\cdot\mathrm{D}^{\alpha*}\boldsymbol{\Phi} = 0$. Therefore, an alternative formulation of the dual problem (25) is given by:

$$\begin{aligned}
\boldsymbol{\Phi} + \nabla\Psi &= 0 \text{ on } \Omega\times(0,T), && \text{(28a)}\\
\Psi' - \nabla\cdot\mathrm{D}^{\alpha*}\boldsymbol{\Phi} &= 0 \text{ on } \Omega\times(0,T), && \text{(28b)}\\
\Psi &= 0 \text{ on } \partial\Omega\times(0,T), && \text{(28c)}\\
\Psi(T) &= \Theta \text{ on } \Omega. && \text{(28d)}
\end{aligned}$$

In the next lemma, an expression for the quantity $(P_0\varepsilon_u(T), \Theta)$ in terms of the errors ε_u', $\boldsymbol{\varepsilon}_{\mathbf{q}}$, the projection errors $e_{\boldsymbol{q}}$ and e_u', and the solution of the dual problem will be given. In it, I_h is any interpolation operator from $L^2(\Omega)$ into $W_h\cap H_0^1(\Omega)$, P_W is the L^2-projection into W_h and $\boldsymbol{\Pi}^{\text{BDM}}$ is the well-known projection associated to the lowest-order Brezzi-Douglas-Marini (BDM) space.

Lemma 4 *Assume that $k \geq 1$. Then, for any $T > 0$,*

$$\begin{aligned}
(P_0\varepsilon_u(T), \Theta) = \int_0^T &[(\boldsymbol{\varepsilon}_{\boldsymbol{q}}, \mathrm{D}^{\alpha*}(\nabla\mathrm{I}_h\Psi) - \boldsymbol{\Pi}^{\text{BDM}}\nabla\Psi)\\
&+(e_{\boldsymbol{q}}, \mathrm{D}^{\alpha*}(\boldsymbol{\Pi}^{\text{BDM}}\nabla\Psi - \nabla P_W\Psi)) + (\varepsilon_u' - e_u', P_0\Psi - \mathrm{I}_h\Psi)]\,dt.
\end{aligned}$$

Proof Since $\Psi(T) = \Theta$ by Eq. 28d and $\varepsilon_u(0) = 0$, we have

$$\begin{aligned}
(P_0\varepsilon_u(T), \Theta) &= \int_0^T [((P_0\varepsilon_u)', \Psi) + (P_0\varepsilon_u, \Psi')]\,dt\\
&= \int_0^T [(\varepsilon_u', P_0\Psi) + (\varepsilon_u, P_0\nabla\cdot\mathrm{D}^{\alpha*}\boldsymbol{\Phi})]\,dt
\end{aligned}$$

by the definition of the L^2-projection P_0 and by Eq. 28b.

By the commutativity property $P_0\nabla\cdot = \nabla\cdot\boldsymbol{\Pi}^{\text{BDM}}$ and the first error equation (14a) with $\boldsymbol{r} := \mathrm{D}^{\alpha*}\boldsymbol{\Pi}^{\text{BDM}}\boldsymbol{\Phi}$ (since $k\geq 1$), we get for each $t\in(0,T]$,

$$\begin{aligned}
(\varepsilon_u, P_0\nabla\cdot\mathrm{D}^{\alpha*}\boldsymbol{\Phi}) &= \left(\varepsilon_u, \nabla\cdot\mathrm{D}^{\alpha*}\boldsymbol{\Pi}^{\text{BDM}}\boldsymbol{\Phi}\right),\\
&= \left(\boldsymbol{\varepsilon}_{\mathbf{q}}, \mathrm{D}^{\alpha*}\boldsymbol{\Pi}^{\text{BDM}}\boldsymbol{\Phi}\right) + \langle\varepsilon_{\widehat{u}}, \mathrm{D}^{\alpha*}\boldsymbol{\Pi}^{\text{BDM}}\boldsymbol{\Phi}\cdot\boldsymbol{n}\rangle - \left(e_{\boldsymbol{q}}, \mathrm{D}^{\alpha*}\boldsymbol{\Pi}^{\text{BDM}}\boldsymbol{\Phi}\right)
\end{aligned}$$

$$= \left(\boldsymbol{\varepsilon}_{\mathbf{q}}, \mathrm{D}^{\alpha *}\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{\Phi}\right) - \left(e_{\boldsymbol{q}}, \mathrm{D}^{\alpha *}\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{\Phi}\right)$$

$$= \left(\boldsymbol{\varepsilon}_{\mathbf{q}}, \mathrm{D}^{\alpha *}(-\boldsymbol{\Pi}^{\mathrm{BDM}}\nabla\Psi + \nabla \mathrm{I}_h\Psi)\right) - (\boldsymbol{\varepsilon}_{\mathbf{q}}, \mathrm{D}^{\alpha *}(\nabla \mathrm{I}_h \Psi)) - \left(e_{\boldsymbol{q}}, \mathrm{D}^{\alpha *}\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{\Phi}\right).$$

Noting that, in the second last equality we used

$$\langle \varepsilon_{\widehat{u}}, \mathrm{D}^{\alpha *}\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{\Phi}\cdot\boldsymbol{n}\rangle = \langle \varepsilon_{\widehat{u}}, \mathrm{D}^{\alpha *}\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{\Phi}\cdot\boldsymbol{n}\rangle_{\partial\Omega} = 0$$

which follows from Eq. 14d (because $\mathrm{D}^{\alpha *}\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{\Phi} \in \mathbf{H}(\mathrm{div}, \Omega)$) and the fact that $\varepsilon_{\widehat{u}} = 0$ on $\partial\Omega$ by Eq. 14c.

But, by the error (14b) with $w := \mathrm{D}^{\alpha *}(\mathrm{I}_h\Psi)$,

$$(\boldsymbol{\varepsilon}_{\mathbf{q}}, \mathrm{D}^{\alpha *}(\nabla \mathrm{I}_h\Psi)) = \left(\mathcal{I}^{\alpha}\left(\varepsilon_u' - e_u'\right), \mathrm{D}^{\alpha *}\left(\mathrm{I}_h\Psi\right)\right) + \langle \boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}, \mathrm{D}^{\alpha *}(\mathrm{I}_h\Psi)\rangle.$$

Now, putting together all the above intermediate steps,

$$\begin{aligned}(P_0\varepsilon_u(T), \Theta) = \int_0^T &\Big[\left(\varepsilon_u', P_0\Psi\right) + (\boldsymbol{\varepsilon}_{\mathbf{q}}, \mathrm{D}^{\alpha *}(\nabla \mathrm{I}_h\Psi) - \boldsymbol{\Pi}^{\mathrm{BDM}}\nabla\Psi) \\ &- \left(\mathrm{D}^{\alpha}\mathcal{I}^{\alpha}\left(\varepsilon_u' - e_u'\right), \mathrm{I}_h\Psi\right) - \langle \boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}, \mathrm{D}^{\alpha *}(\mathrm{I}_h\Psi)\rangle \\ &-(e_{\boldsymbol{q}}, \mathrm{D}^{\alpha *}\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{\Phi})\Big]\, dt. \end{aligned} \tag{29}$$

But, $\langle \boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}, \mathrm{D}^{\alpha *}(\mathrm{I}_h\Psi)\rangle = \langle \boldsymbol{\varepsilon}_{\widehat{q}}\cdot\boldsymbol{n}, \mathrm{D}^{\alpha *}(\mathrm{I}_h\Psi)\rangle_{\partial\Omega} = 0$ by Eq. 14d and the identity $\mathrm{I}_h\Psi = 0$ on $\partial\Omega$ by the boundary condition of the dual problem (28c). Using this and the identity $\mathrm{D}^{\alpha}(\mathcal{I}^{\alpha}(\varepsilon_u' - e_u')) = (\varepsilon_u - e_u)'$ in Eq. 29, we observe

$$(P_0\varepsilon_u(T), \Theta) = \int_0^T [(\boldsymbol{\varepsilon}_{\mathbf{q}}, \mathrm{D}^{\alpha *}(\nabla \mathrm{I}_h\Psi - \boldsymbol{\Pi}^{\mathrm{BDM}}\nabla\Psi))$$

$$-(e_{\boldsymbol{q}}, \mathrm{D}^{\alpha *}\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{\Phi}) + (\varepsilon_u', P_0\Psi - \mathrm{I}_h\Psi) + (e_u', \mathrm{I}_h\Psi)]\, dt.$$

Therefore, the desired result now follows after noting that

$$-\int_0^T \left(e_{\boldsymbol{q}}, \mathrm{D}^{\alpha *}\boldsymbol{\Pi}^{\mathrm{BDM}}\boldsymbol{\Phi}\right) dt = \int_0^T \left(e_{\boldsymbol{q}}, \mathrm{D}^{\alpha *}(\boldsymbol{\Pi}^{\mathrm{BDM}}\nabla\Psi - \nabla P_W\Psi)\right) dt,$$

(by Eq. 28a, the fact that P_W is the L^2-projection into W_h, and the orthogonality property of the projection $\boldsymbol{\Pi}_V$, Eq. 11a) and that $(e_u', \mathrm{I}_h\Psi) = (e_u', \mathrm{I}_h\Psi - P_0\Psi)$ (by the fact that $P_0\Psi$ is constant on each element $K \in \mathcal{T}_h$, and the orthogonality property of the projection Π_W, Eq. 11b). The proof is completed now. □

In the next theorem we state the superconvergence estimate of the postprocessed HDG approximation. For the proof, we follow the derivation in [6, Section 5] step-by-step and use Lemma 4 instead of [6, Lemma 7], and we also use the achieved HDG error estimates in Theorem 2.

Theorem 3 *Assume that* $u \in \mathcal{C}^1(0, T; H^{k+2}(\Omega))$ *and* $\boldsymbol{q} \in \mathcal{C}^1(0, T; \boldsymbol{H}^{k+1}(\Omega))$. *Assume also that* τ_K^* *and* $1/\tau_K^{\max}$ *are bounded by* C. *Then, we have*

$$\|(u - u_h^*)(T)\| \le C_2 \max\left\{1, \sqrt{\log\left(Th^{-2/(\alpha+1)}\right)}\right\} h^{k+2} \quad \text{for } k \ge 1,$$

where the constant C_2, only depends on C*, α, T, $\|u\|_{\mathcal{C}^1(H^{k+2})}$, and on $\|\boldsymbol{q}\|_{\mathcal{C}^1(\boldsymbol{H}^{k+1})}$.*

5 Numerical experiments

In this section, we present numerical experiments devised to validate our theoretical predictions from HDG spatial discretizations. To do so, we use the fully discrete CN HDG scheme (8). We take the (uniform) time steps δ to be sufficiently small so that the HDG and postprocessed HDG spatial discretizations errors are dominant. This is achieved by fixing the ratio $\frac{\delta^2}{h^{k+2}}$ to a given number less than the unit because the time stepping CN scheme is second-order accurate provided that the exact solution is sufficiently regular.

We choose the spatial domain Ω to be the unit interval (0, 1) and $T = 1$ in (1). We impose homogenous Dirichlet boundary conditions and choose the source term f and the initial data u_0 so that the exact solution is $u(x, t) = t^{3-\alpha} sin(\pi x)$. For different values of α, we obtain the history of convergence of the errors $\|(u - u_h)(T)\|$, $\|(\boldsymbol{q} - \boldsymbol{q}_h)(T)\|$ and $\|(u - u_h^\star)(T)\|$ for different values of the polynomial degree,

Table 1 The errors $\|(u_h - u)(T)\|$, $\|(\boldsymbol{q}_h - \boldsymbol{q})(T)\|$ and $\|(u_h^\star - u)(T)\|$, and the corresponding rates of convergence for $\alpha = 0.5$ with HDG solutions of degree $k = 0, 1, 2$

N						
	$k = 0$					
4	5.269e-01		7.899e-01		5.048e-01	
8	3.027e-01	0.799	4.028e-01	0.972	2.922e-01	0.788
16	1.616e-01	0.905	2.025e-01	0.992	1.566e-01	0.899
32	8.342e-02	0.954	1.014e-01	0.997	8.098e-02	0.951
64	4.237e-02	0.977	5.072e-02	0.999	4.117e-02	0.976
128	2.135e-02	0.989	2.537e-02	0.999	2.076e-02	0.988
	$k = 1$					
4	6.031e-02		5.936e-02		7.401e-03	
8	1.502e-02	2.005	1.321e-02	2.165	8.835e-04	3.066
16	4.144e-03	1.858	3.487e-03	1.924	1.142e-04	2.951
32	1.048e-03	1.983	8.649e-04	2.011	1.420e-05	3.008
64	2.697e-04	1.958	2.199e-04	1.976	1.812e-06	2.970
	$k = 2$					
4	3.960e-03		4.596e-03		8.902e-04	
8	5.059e-04	2.969	4.868e-04	3.239	5.497e-05	4.017
16	6.352e-05	2.993	5.652e-05	3.107	3.416e-06	4.008

We observe optimal convergence of order h^{k+1} for the errors in u_h and $\boldsymbol{q}_h$, and superconvergence rates of order h^{k+2} (when $k \geq 1$) for the error from the postprocessed HDG solution $u_h^\star$

Table 2 The errors $\|(u_h - u)(T)\|$, $\|(\boldsymbol{q}_h - \boldsymbol{q})(T)\|$ and $\|(u_h^\star - u)(T)\|$, and the corresponding rates of convergence for $\alpha = 0.7$ with HDG solutions of degree $k = 0, 1, 2$

N						
	$k = 0$					
4	5.455e-01		7.705e-01		5.240e-01	
8	3.122e-01	0.805	3.088e-01	9.884	3.020e-01	0.795
16	1.661e-01	0.910	1.939e-01	1.002	1.612e-01	0.905
32	8.558e-02	0.957	9.674e-02	1.003	8.320e-02	0.954
64	4.342e-02	0.979	4.830e-02	1.002	4.225e-02	0.978
128	2.187e-02	0.989	2.413e-02	1.001	2.128e-02	0.989
	$k = 1$					
4	6.081e-02		6.005e-02		7.898e-03	
8	1.501e-02	2.018	1.321e-02	2.185	9.403e-04	3.070
16	4.154e-03	1.854	3.485e-03	1.922	1.218e-04	2.949
32	1.048e-03	1.987	8.434e-04	2.047	1.506e-05	3.015
64	2.682e-04	1.966	2.143e-04	1.977	1.953e-06	2.947
	$k = 2$					
4	4.025e-03		4.978e-03		1.079e-03	
8	5.088e-04	2.984	5.014e-04	3.312	6.698e-05	4.010
16	6.367e-05	2.998	5.701e-05	3.137	4.167e-06	4.007

$k = 0, 1, 2$. To compute the spatial L_2-norm, we apply a composite Gauss quadrature rule with 4 points on each interval of the finest spatial mesh. The numerical results (errors and convergence rates) of the experiments are presented in Tables 1 and 2. In full agreement with our theoretical results, we obtain optimal convergence rates for the HDG scheme and $O(h^{k+2})$ superconvergence rates for the postprocessed HDG scheme.

References

1. Chabaud, B., Cockburn, B.: Uniform-in-time superconvergence of HDG methods for the heat equation. Math. Comp. **81**, 107–129 (2012)
2. Chen, C.-M., Liu, F., Anh, V., Turner, I.: Numerical methods for solving a two-dimensional variable-order anomalous sub-diffusion equation. Math. Comp. **81**, 345–366 (2012)
3. Chen, J., Liu, F., Liu, Q., Anh, V., Turner, I.: Numerical simulation for the three-dimension fractional sub-diffusion equation. Appl. Math. Model. **38**, 3695–3705 (2014)
4. Cockburn, B., Gopalakrishnan, J., Lazarov, R.: Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems. SIAM J. Numer. Anal. **47**, 1319–1365 (2009)
5. Cockburn, B., Gopalakrishnan, J., Sayas, F.-J.: A projection-based error analysis of HDG methods. Math. Comp. **79**, 1351–1367 (2010)
6. Cockburn, B., Mustapha, K.: A hybridizable discontinuous Galerkin method for fractional diffusion problems. Numer. Math. **130**, 293–314 (2015)

7. Cockburn, B., Qiu, W., Shi, K.: Conditions for superconvergence of HDG methods for second-order elliptic problems. Math. Comp. **81**, 1327–1353 (2012)
8. Cuesta, E., Lubich, C., Palencia, C.: Convolution quadrature time discretization of fractional diffusive-wave equations. Math. Comp. **75**, 673–696 (2006)
9. Cui, M.: Compact finite difference method for the fractional diffusion equation. J. Comput. Phys. **228**, 7792–7804 (2009)
10. Cui, M.: Convergence analysis of high-order compact alternating direction implicit schemes for the two-dimensional time fractional diffusion equation. Numer. Algor. **62**, 383–409 (2013)
11. Gastaldi, L., Nochetto, R.H.: Sharp maximum norm error estimates for general mixed finite element approximations to second order elliptic equations, RAIRO Modél. Math. Anal. Numér. **23**, 103–128 (1989)
12. Jin, B., Lazarov, R., Zhou, Z.: Two schemes for fractional diffusion and diffusion-wave equations with nonsmooth data, arXiv:1404.3800 (2014)
13. Kirby, R.M., Sherwin, S.J., Cockburn, B.: To HDG or to CG: A comparative study. J. Sci. Comput. **51**, 183–212 (2012)
14. Li, X., Xu, C.: A space-time spectral method for the time fractional diffusion equation, SIAM. J. Numer. Anal. **47**, 2108–2131 (2009)
15. Lin, Y., Xu, C.: Finite differnce/spectral approximations for the time-fractional diffusion equation. J. Comput. Phys. **225**, 1533–1552 (2007)
16. McLean, W., Mustapha, K.: Convergence analysis of a discontinuous Galerkin method for a sub-diffusion equation. Numer. Algor. **52**, 69–88 (2009)
17. McLean, W., Mustapha, K.: Error analysis of a discontinuous Galerkin method for a fractional diffusion equation with a non-smooth initial data. J. Comput. Phys. **293**, 201–217 (2015)
18. Mustapha, K.: Time-stepping discontinuous Galerkin methods for fractional diffusion problems. Numer. Math. **130**, 497–516 (2015)
19. Mustapha, K., AlMutawa, J.: A finite difference method for an anomalous sub-diffusion equation, theory and applications. Numer. Algor. **61**, 525–543 (2012)
20. Mustapha, K., Abdallah, B., Furati, K.: A discontinuous Petrov-Galerkin method for time-fractional diffusion equations. SIAM J. Numer. Numer. Anal. **52**, 2512–2529 (2014)
21. Mustapha, K., McLean, W.: Piecewise-linear, discontinuous Galerkin method for a fractional diffusion equation. Numer. Algor. **56**, 159–184 (2011)
22. Mustapha, K., McLean, W.: Uniform convergence for a discontinuous Galerkin, time stepping method applied to a fractional diffusion equation. IMA J. Numer. Anal. **32**, 906–925 (2012)
23. Mustapha, K., McLean, W.: Superconvergence of a discontinuous Galerkin method for fractional diffusion and wave equations. SIAM J. Numer. Anal. **51**, 491–515 (2013)
24. Mustapha, K., Schötzau, D.: Well-posedness of $hp-$version discontinuous Galerkin methods for fractional diffusion wave equations. IMA J. Numer. Anal. **34**, 1426–1446 (2014)
25. Nguyen, N.C., Peraire, J., Cockburn, B.: Hybridizable discontinuous Galerkin methods, Proceedings of the International Conference on Spectral and High Order Methods (Trondheim, Norway), Lect. Notes Comput. Sci. Engrg., Springer Verlag (2009)
26. Quintana-Murillo, J., Yuste, S.B.: An explicit difference method for solving fractional diffusion and diffusion-wave equations in the Caputo form. J. Comput. Nonlin. Dyn. **6**, 021014 (2011)
27. Sweilam, N.H., Khader, M.M., Mahdy, A.M.S.: Crank-Nicolson finite difference method for solving time-fractional diffusion equation. J. Fract. Cal. Appl. **2**, 1–9 (2012)
28. Xu, Q., Zheng, Z.: Discontinuous Galerkin method for time fractional diffusion equation. J. Informat. Comput. Sci. **10**, 3253–3264 (2013)
29. Yuste, S.B., Quintana-Murillo, J.: On Three Explicit Difference Schemes for Fractional Diffusion and Diffusion-Wave Equations. Phys. Scripta T **136**, 014025 (2009)
30. Zengo, F., LI, C., Liu, F., Turner, I.: The use of finite difference/element approaches for solving the time-fractional subdiffusion equation. SIAM J. Sci. Comput. **35**, A2976—A3000 (2013)
31. Zhang, Y.-N., Sun, Z.-Z.: Alternating direction implicit schemes for the two-dimensional fractional sub-diffusion equation. J. Comput. Phys. **230**, 8713–8728 (2011)
32. Zhao, X., Sun, Z.-z.: A box-type scheme for fractional sub-diffusion equation with Neumann boundary conditions. J. Comput. Phys. **230**, 6061–6074 (2011)
33. Zhao, X., Xu, Q.: Efficient numerical schemes for fractional sub-diffusion equation with the spatially variable coefficient. Appl. Math. Model. **38**, 3848–3859 (2014)