DISCONTINUOUS GALERKIN METHODS AND THEIR ADAPTIVITY FOR THE TEMPERED FRACTIONAL (CONVECTION) DIFFUSION EQUATIONS*

Xudong Wang   and   Weihua Deng

School of Mathematics and Statistics, Gansu Key Laboratory of Applied Mathematics and Complex Systems, Lanzhou University, Lanzhou 730000, China
Email: xdwang14@lzu.edu.cn, dengwh@lzu.edu.cn

Abstract

This paper focuses on the adaptive discontinuous Galerkin (DG) methods for the tempered fractional (convection) diffusion equations. The DG schemes with interior penalty for the diffusion term and numerical flux for the convection term are used to solve the equations, and the detailed stability and convergence analyses are provided. Based on the derived posteriori error estimates, the local error indicator is designed. The theoretical results and the effectiveness of the adaptive DG methods are, respectively, verified and displayed by the extensive numerical experiments. The strategy of designing adaptive schemes presented in this paper works for the general PDEs with fractional operators.

Mathematics subject classification: 26A33, 65M60, 65M12.
Key words: Adaptive DG methods, Tempered fractional equations, Posteriori error estimate.

1. Introduction

Fractional calculus [8] is a popular mathematical tool for modelling anomalous diffusions [27], being ubiquitous in nature. Microscopically, anomalous diffusion can be described by continuous time random walk (CTRW), defined by the waiting time and jump length; generally the first moment of the waiting time and/or the second moment of the jump length diverge(s). Sometimes, it is better to temper the broad distribution(s) of the waiting time and/or the jump length [4, 19, 25, 44], because of the boundedness of physical space or the finite lifespan of the biological particles or the slow transition of different diffusion types. Based on the tempered CTRW, the partial differential equations (PDEs) characterizing the evolution of the functional distribution of the trajectories of the particles are derived [41], which reduce to the PDEs describing the distribution of the positions of the particles if taking the parameter p over there as 0, called tempered fractional PDEs; here, we discuss their (adaptive) discontinuous Galerkin (DG) methods.

There are already some works for numerically solving (tempered) fractional PDEs by variational methods [19,22,26,29,31,33,40,42,46]. Ervin and Roop [22] firstly present the variational formulation for the fractional advection dispersion equation. The DG methods are particularly applied to fractional problems with their majority of characteristics [7,13,24,34,39,43], naturally being formulated for any order of accuracy in any element, being flexible in choosing element sizes in any place, suitable for adaptivity, being local and easy to invert for mass matrix, leading

to an explicit formulation for time dependent problems, etc. Cockburn and Mustapha [14] provide a hybridizable DG method for fractional diffusion problems; McLean and Mustapha [29] discuss the superconvergence of the DG method for the fractional diffusion and wave equations; Xu and Hesthaven [42], and Wang et al [40], respectively, consider DG and hybridized DG methods for the fractional convection-diffusion equations; Zayernouri and Karniadakis [46] design discontinuous spectral element methods for the time and space fractional differential equations. Du et al [20] give a convergent adaptive finite element algorithm for nonlocal diffusion and peridynamic models. Ainsworth and Glusa [3], and Chen et al [11], discuss the adaptive algorithms about fractional Laplacian with integral definition and spectral definition, respectively. Zhao et al [45] design an adaptive algorithm for Riesz fractional derivative with a posteriori error estimator based on gradient recovery approach. It seems that there are not works for investigating the potential advantages of DG methods in adaptivity for fractional problems, by deriving posteriori error estimates and providing the local error indicators.

The model we consider in this paper is the two dimensional space tempered fractional differential equation with absorbing boundary conditions [17, 18, 21], i.e.,

$$
\begin{cases}
\partial_t u + \mathbf{b} \cdot \nabla u - \kappa_1 \nabla_x^{\alpha,\lambda} u - \kappa_2 \nabla_y^{\beta,\lambda} u = f, & (\mathbf{x}, t) \in \Omega \times J, \\
u(\mathbf{x}, t) = u_{\text{in}}, & \Gamma_{\text{in}} \times J, \\
u(\mathbf{x}, 0) = u_0(\mathbf{x}), & \mathbf{x} \in \Omega, \\
u(\mathbf{x}, t) = 0, & (\mathbf{x}, t) \in \mathbb{R}^2 \backslash \bar{\Omega} \times J,
\end{cases}
\tag{1.1}
$$

where $\alpha, \beta \in (0, 1)$, $\lambda > 0$, and $\kappa_1, \kappa_2 > 0$ in the domain $\Omega = [a, b] \times [c, d]$ and $J = [0, T]$. The boundary of the domain Ω is decomposed into two parts: the inflow part Γ_{in} and outflow part Γ_{out} defined by

$$
\Gamma_{\text{in}} = \{\mathbf{x} \in \partial\Omega : \mathbf{b} \cdot \mathbf{n} < 0\}, \qquad \Gamma_{\text{out}} = \partial\Omega \backslash \Gamma_{\text{in}},
\tag{1.2}
$$

where $\mathbf{n}$ is the unit outward normal vector on the boundary. The model (1.1) describes a convection-diffusion problem with convection term $\mathbf{b} \cdot \nabla u$ and diffusion term $\kappa_1 \nabla_x^{\alpha,\lambda} u + \kappa_2 \nabla_y^{\beta,\lambda} u$ in horizontal and vertical directions respectively [17, 28]. Because of the existence of nonlocal operators $\nabla_x^{\alpha,\lambda}$ and $\nabla_y^{\beta,\lambda}$, the local boundary $\partial\Omega$ itself cannot be hit by the majority of discontinuous sample trajectories; based on this physical implication, this problem should be specified the generalized Dirichlet boundary conditions, i.e., in the complementary of Ω [17, 18, 21]. Note that if $\kappa_1 = \kappa_2$ and $\alpha = \beta$, the diffusion term will not reduce to a two-dimensional fractional Laplacian [17]. The function $f \in L^2(J; L^2(\Omega))$ is a source term; the convection coefficient $\mathbf{b}$ is assumed to be continuous and satisfy $\nabla \cdot \mathbf{b} = 0$, and the initial function $u_0 \in L^2(\Omega)$. The tempered fractional operator $\nabla_x^{\alpha,\lambda}$ is defined from [9] and will be shown in the next section. Compared with non-tempered case, the tempered operator $\nabla_x^{\alpha,\lambda}$ characterizes the physical reality that the jump length of a particle will not be arbitrarily large [41]. As for the discussion of the adaptivity of the fractional problems, we start from the steady state version of (1.1) with $\mathbf{b} = \mathbf{0}$. The first part of the paper focuses on designing the DG scheme of (1.1) with genuinely triangular grids, and offering explicit theoretical analyses. Being different from [33], which constructs the LDG scheme by rewriting the fractional equation as a first order system, we adopt the primal DG methods, namely interior penalty (IP) method, still keeping the advantages over the classical continuous Galerkin method in facilitating hp-adaptivity and yielding block diagonal mass matrices in time-dependent problems. Generally, the non-ignorable drawback of the IP method is to specify sufficient large penalty parameter for guaranteeing numerical stability, which degrades the performance of the iterative solver of the linear system [36]. Fortunately, for the (tempered) fractional equations, this drawback disappears, since the schemes are stable

for any value of the penalty parameter, say, simply taking as 1. For the convection term, the upwind flux [13, 34] is used for ensuring numerical stability. High-order accuracy of $h^{N-\epsilon}$ can be obtained by only assuming that the weak solution has sufficient regularity in each element.

Mesh adaption is the basic technique of balancing the computational cost and accuracy, which introduces extra points near the singularities or the high gradient part of the solution to be computed. The key ingredient of adaptivity is a posteriori error estimators [7, 10], which are computable quantities depending on the computed solution and data. We derive a posteriori error estimators for fractional operators and obtain the local error indicators, being used to dynamically and locally refine or coarsen meshes. To show the effectiveness of the local error indicators, we adaptively solve the fractional differential equations with singularities, including both the steady state and time dependent ones. For the steady state equations, two schemes are presented. One is based on energy norm, while another one is based on dual weighted residual (DWR). It is observed that the latter performs better than the former. For the time dependent one, both the space mesh and time-step size are adapted, simultaneously.

The outline of this paper is as follows. Section 2 is composed of five subsections. The first subsection reviews the definitions and properties of tempered fractional calculus. The notations and the variational formulations of DG schemes are, respectively, proposed in the second and third subsections. In the fourth subsection, we perform the stability analysis and error estimates for the two dimensional tempered fractional convection-diffusion equations. The numerical results are provided in the last subsection. Section 3 is discussing adaptivity, composed of two parts, which are, respectively, for the stationary equation and the evolution equation. We conclude the paper with some remarks in the last section.

2. DG for the Tempered Fractional Convection-Diffusion Equation

In this section, we design the DG scheme for the tempered fractional convection-diffusion equation (1.1), provide the detailed stability and convergence proof, and numerically verify the theoretical results.

2.1. Tempered fractional operators

We firstly introduce some preliminary definitions of tempered fractional calculus [9, 25].

Definition 2.1. *For any $\alpha > 0, \lambda > 0$, the left and right tempered Riemann-Liouville fractional integrals of function $u(x)$ defined on $\mathbb{R}$ are given by*

$$_{-\infty}I_x^{\alpha,\lambda}u(x) = \mathrm{e}^{-\lambda x}{}_{-\infty}I_x^{\alpha}[\mathrm{e}^{\lambda x}u(x)] = \frac{1}{\Gamma(\alpha)}\int_{-\infty}^x (x-\xi)^{\alpha-1}\mathrm{e}^{-\lambda(x-\xi)}u(\xi)d\xi,$$

$$_xI_\infty^{\alpha,\lambda}u(x) = \mathrm{e}^{\lambda x}{}_xI_\infty^{\alpha}[\mathrm{e}^{-\lambda x}u(x)] = \frac{1}{\Gamma(\alpha)}\int_x^\infty (\xi-x)^{\alpha-1}\mathrm{e}^{-\lambda(\xi-x)}u(\xi)d\xi.$$

Definition 2.2. *For any $\alpha > 0, n - 1 < \alpha < n, n \in \mathbb{N}^+, \lambda > 0$, the left and right tempered Riemann-Liouville fractional derivatives of function $u(x)$ defined on $\mathbb{R}$ are given by*

$$_{-\infty}D_x^{\alpha,\lambda}u(x) = \mathrm{e}^{-\lambda x}{}_{-\infty}D_x^{\alpha}[\mathrm{e}^{\lambda x}u(x)] = \frac{\mathrm{e}^{-\lambda x}}{\Gamma(n-\alpha)}\frac{d^n}{dx^n}\int_{-\infty}^x (x-\xi)^{n-\alpha-1}\mathrm{e}^{\lambda\xi}u(\xi)d\xi,$$

$$_xD_\infty^{\alpha,\lambda}u(x) = \mathrm{e}^{\lambda x}{}_xD_\infty^{\alpha}[\mathrm{e}^{-\lambda x}u(x)] = \frac{(-1)^n\mathrm{e}^{\lambda x}}{\Gamma(n-\alpha)}\frac{d^n}{dx^n}\int_x^\infty (\xi-x)^{n-\alpha-1}\mathrm{e}^{-\lambda\xi}u(\xi)d\xi.$$

Definition 2.3. *The Riesz tempered fractional derivative is $\nabla_x^{\alpha,\lambda}$ of $u(x)$ with $\alpha, \beta \in (0,1)$, $\lambda > 0$, is defined as*

$$\nabla_x^{\alpha,\lambda} u(x) = -\kappa_\alpha(\mathbb{D}_x^{\alpha,\lambda} - 2\lambda^\alpha)u(x), \tag{2.1}$$

where $\kappa_\alpha = \frac{1}{2\cos(\alpha\pi/2)}$ and $\mathbb{D}_x^{\alpha,\lambda} = {}_{-\infty}D_x^{\alpha,\lambda} + {}_xD_\infty^{\alpha,\lambda}$. The definition of $\nabla_y^{\beta,\lambda}$ is similar.

Lemma 2.1. *For $u(x) \in L^2(\mathbb{R})$ and $\mu, \lambda > 0$, it holds that*

$$\mathscr{F}[{}_{-\infty}I_x^{\mu,\lambda}u(x)](\omega) = (\lambda + i\omega)^{-\mu}\hat{u}(\omega),$$
$$\mathscr{F}[{}_xI_\infty^{\mu,\lambda}u(x)](\omega) = (\lambda - i\omega)^{-\mu}\hat{u}(\omega).$$

If $u(x) \in C_0^\infty(\mathbb{R})$ further, then

$$\mathscr{F}[{}_{-\infty}D_x^{\mu,\lambda}u(x)](\omega) = (\lambda + i\omega)^\mu\hat{u}(\omega),$$
$$\mathscr{F}[{}_xD_\infty^{\mu,\lambda}u(x)](\omega) = (\lambda - i\omega)^\mu\hat{u}(\omega),$$

where $\hat{u}(\omega) := \mathscr{F}[u(x)](\omega) = \int_{-\infty}^\infty e^{-i\omega x}u(x)dx$.

Based on the Fourier transform in Lemma 2.1, the equivalence of the fractional derivative space with fractional order Hilbert spaces in one dimension has been established in [22]. Here, we extend some results to two dimensions since the problem (1.1) contains two single-directional fractional operators. For $A, B > 0$, let $A \sim B$, which means that there exist positive constants c_1, c_2 independent on B satisfying $c_1 B \leq A \leq c_2 B$.

Lemma 2.2 ([19]). *Let $p, q > 0$ and $\alpha > 0$. Then*

$$(p + q)^\alpha \sim p^\alpha + q^\alpha. \tag{2.2}$$

More specifically,

$$2^{\alpha-1}(p^\alpha + q^\alpha) \leq (p + q)^\alpha \leq (p^\alpha + q^\alpha) \quad \text{for} \quad 0 < \alpha \leq 1;$$
$$(p^\alpha + q^\alpha) \leq (p + q)^\alpha \leq 2^{\alpha-1}(p^\alpha + q^\alpha) \quad \text{for} \quad \alpha > 1.$$

Let $\mu > 0$. We define the norms for the functions in $H^{\mu,\lambda}(\mathbb{R}^2)$ and $H^\mu(\mathbb{R}^2)$ in terms of the Fourier transform, i.e.,

$$|u(x,y)|_{H_x^{\mu,\lambda}(\mathbb{R}^2)}^2 := \int_{\mathbb{R}^2} (\lambda^2 + \omega_1^2)^\mu |\hat{u}|^2 d\omega,$$

$$|u(x,y)|_{H_y^{\mu,\lambda}(\mathbb{R}^2)}^2 := \int_{\mathbb{R}^2} (\lambda^2 + \omega_2^2)^\mu |\hat{u}|^2 d\omega,$$

$$|u(x,y)|_{H^{\mu,\lambda}(\mathbb{R}^2)}^2 := \int_{\mathbb{R}^2} (\lambda^2 + |\omega|^2)^\mu |\hat{u}|^2 d\omega,$$

$$\|u(x,y)\|_{H^\mu(\mathbb{R}^2)}^2 := \int_{\mathbb{R}^2} (1 + |\omega|^{2\mu}) |\hat{u}|^2 d\omega,$$

where ω_1 and ω_2 are two components of ω. Noticing that $|\omega|^2 = \omega_1^2 + \omega_2^2$, together with (2.2), we have

$$1 + |\omega|^{2\mu} \sim (\lambda^2 + |\omega|^2)^\mu \sim (\lambda^2 + \omega_1^2)^\mu + (\lambda^2 + \omega_2^2)^\mu.$$

Therefore, for a fixed λ,

$$\|u\|_{H^\mu(\mathbb{R}^2)}^2 \sim |u|_{H^{\mu,\lambda}(\mathbb{R}^2)}^2 \sim |u|_{H_x^{\mu,\lambda}(\mathbb{R}^2)}^2 + |u|_{H_y^{\mu,\lambda}(\mathbb{R}^2)}^2, \tag{2.3}$$

which implies that the semi-norm $|\cdot|_{H^{\mu,\lambda}(\mathbb{R}^2)}$ is a norm and can be presented as the sum of two semi-norms $|\cdot|_{H_x^{\mu,\lambda}(\mathbb{R}^2)} + |\cdot|_{H_y^{\mu,\lambda}(\mathbb{R}^2)}$.

By Plancherel's theorem, the following Lemma has been proved in [19, Theorem 3.2].

Lemma 2.3. Let $0 < \mu < 1$, and $\lambda > 0$. If $u \in H^{\mu/2}(\mathbb{R}^2)$, then

$$\left({}_{-\infty}D_x^{\mu/2,\lambda}u, \, {}_xD_\infty^{\mu/2,\lambda}u \right) \sim \| {}_{-\infty}D_x^{\mu/2,\lambda}u\|_{L^2(\mathbb{R}^2)}^2 \sim \| {}_xD_\infty^{\mu/2,\lambda}u\|_{L^2(\mathbb{R}^2)}^2 \sim |u|_{H_x^{\mu/2,\lambda}(\mathbb{R}^2)}^2, \quad (2.4)$$

$$\left({}_{-\infty}D_y^{\mu/2,\lambda}u, \, {}_yD_\infty^{\mu/2,\lambda}u \right) \sim \| {}_{-\infty}D_y^{\mu/2,\lambda}u\|_{L^2(\mathbb{R}^2)}^2 \sim \| {}_yD_\infty^{\mu/2,\lambda}u\|_{L^2(\mathbb{R}^2)}^2 \sim |u|_{H_y^{\mu/2,\lambda}(\mathbb{R}^2)}^2. \quad (2.5)$$

Combining (2.3) and Lemma 2.3, we have

$$\|u\|_{H^{\mu/2}(\mathbb{R}^2)}^2 \sim \left({}_{-\infty}D_x^{\mu/2,\lambda}u, \, {}_xD_\infty^{\mu/2,\lambda}u \right) + \left({}_{-\infty}D_y^{\mu/2,\lambda}u, \, {}_yD_\infty^{\mu/2,\lambda}u \right). \quad (2.6)$$

Remark 2.1. The formula (2.6) shows the coercivity of the tempered fractional operator for $0 < \mu < 1$. If $1 < \mu < 2$, the tempered fractional operator is also coercive; for the details of the proof, see [19, Theorem 3.2].

In the following, we use $H^\mu(\Omega)$ to denote the space of functions on Ω that admits extensions to $H^\mu(\mathbb{R}^2)$, equipped with the quotient norm $\|u\|_{H^\mu(\Omega)} := \inf_{\tilde{u}} \|\tilde{u}\|_{H^\mu(\mathbb{R}^2)}$, where the infimum extends over all possible $\tilde{u} \in H^\mu(\mathbb{R}^2)$ such that $\tilde{u} = u$ on Ω (in the sense of distributions). $H_0^\mu(\Omega)$ denotes the closure of $C_0^\infty(\Omega)$ w.r.t. $\|\cdot\|_{H^\mu(\Omega)}$. Then for $u \in H_0^{\mu/2}(\Omega)$, the result in domain Ω of (2.6) is also valid:

$$\|u\|_{H^{\mu/2}(\Omega)}^2 \sim \left({}_aD_x^{\mu/2,\lambda}u, \, {}_xD_b^{\mu/2,\lambda}u \right) + \left({}_cD_y^{\mu/2,\lambda}u, \, {}_yD_d^{\mu/2,\lambda}u \right). \quad (2.7)$$

We also list the fractional Poincaré-Friedrichs inequality for $H_0^\mu(\Omega)$ [19,22].

Lemma 2.4. Let $0 < s < \mu$, and $s \neq n - \frac{1}{2}$ $(n \in \mathbb{N}^+)$. If $u \in H_0^\mu(\Omega)$, then

$$\|u\|_{L^2(\Omega)} \leq C|u|_{H_0^\mu(\Omega)} \quad and \quad |u|_{H_0^s(\Omega)} \leq C|u|_{H_0^\mu(\Omega)}.$$

2.2. Notations for DG methods

Denote Ω_h as a conforming subdivision of $\Omega \subset \mathbb{R}^2$ with non-overlapping triangles. For an element $T \in \Omega_h$, h_T denotes its diameter, and $h = \max_{T \in \Omega_h} h_T$. The family of meshes Ω_h is assumed to be shape-regular, i.e., there exists a constant c such that for all $T \in \Omega_h$, $\frac{h_T}{\rho_T} \leq c$, where ρ_T denotes the radius of the largest inscribed ball in T. Denote Γ as the union of the boundaries of the elements T of Ω_h, and Γ_i the set of interior faces of Ω_h, i.e., the set of faces that are not included in the boundary $\partial\Omega = \Gamma_{\text{in}} \cup \Gamma_{\text{out}}$, where Γ_{in} is introduced in (1.1). Associated with the mesh Ω_h, we define the broken Sobolev spaces with $s \geq 0$,

$$H^s(\Omega_h) := \{v : \Omega \to \mathbb{R} \big| \ v|_T \in H^s(T), \ \forall T \in \Omega_h\}, \quad (2.8)$$

$$H_0^s(\Omega_h) := \{v \in H^s(\Omega_h), \ \ v|_{\partial\Omega} = 0\}, \quad (2.9)$$

equipped with the broken Sobolev norm

$$\|v\|_{H^s(\Omega_h)}^2 := \sum_{T \in \Omega_h} \|v\|_{H^s(T)}^2. \quad (2.10)$$

If $s = k + \sigma \notin \mathbb{N}_0$, then $\| \cdot \|_{H^s(T)}$ is the Aronszajn-Slobodeckij norm [32] defined as

$$\|v\|^2_{H^s(T)} = \|v\|^2_{H^k(T)} + \int_T \int_T \frac{|\partial^k v(\mathbf{x}) - \partial^k v(\mathbf{y})|^2}{|\mathbf{x} - \mathbf{y}|^{2+2\sigma}} d\mathbf{x} d\mathbf{y}. \tag{2.11}$$

The DG finite element space is defined as follows:

$$V_h = \{v : v \in L^2(\Omega) \big| \ v|_T \in P_N(T), \ \forall T \in \Omega_h\}, \tag{2.12}$$

where $P_N(T)$ denotes the set of polynomials of degree less than or equal to N. The numerical solution $u_h(\mathbf{x}, t)$ can be expressed by

$$u_h(\mathbf{x}, t)|_T = \sum_{j=1}^{N_p} u_h(\mathbf{x}_j, t) l_j(\mathbf{x}) \qquad \mathbf{x} \in T, \tag{2.13}$$

where $l_j(\mathbf{x})$ denotes the two-dimensional multivariate Lagrange interpolation basis function, and $N_p = (N + 1)(N + 2)/2$ is degree of freedom in one element.

Next, we introduce some notations to manipulate numerical fluxes. If two elements T_e^1 and T_e^2 are neighbors and share one common side e, there are two traces of the function v along e. We assume that the normal vector $\mathbf{n}_e$ is oriented from T_e^1 to T_e^2, and denote

$$\{v\} = \frac{1}{2}(v|_{T_e^1}) + \frac{1}{2}(v|_{T_e^2}), \quad [v] = (v|_{T_e^1}) - (v|_{T_e^2}), \qquad \forall e = \partial T_e^1 \cap \partial T_e^2.$$

The definition of jump and average to sides that belong to the boundary $\partial\Omega$ is:

$$\{v\} = [v] = (v|_{T_e^1}), \qquad \forall e = \partial T_e^1 \cap \partial\Omega.$$

We also introduce the bilinear form $J_0 : H^s(\Omega_h) \times H^s(\Omega_h) \to \mathbb{R}$ that penalizes the jump of the function values:

$$J_0(v, w) = \sum_{e \in \Gamma_i} \int_e [v][w].$$

Though the solutions to fractional equations typically have quite low regularity, the low regularity is commonly located at the boundary, not the interior [2]. In addition, the estimates for higher order Hölder norms of solution can be obtained by technically considering a weighted norm which vanishes at the boundary [2]. Here for simplicity, we only assume the solution to have sufficient regularity in each elements for DG method. Then the traces in numerical fluxes, the penalty J_0 and (2.9) are all well-defined with respect to the element in which they are located in the broken Sobolev space $H^s(\Omega_h)$.

2.3. Variational formulation

From the definition of the tempered fractional derivatives (2.1), we rewrite (1.1) as

$$\begin{cases} \dfrac{\partial u}{\partial t} + \mathbf{b} \cdot \nabla u + \kappa_1 \kappa_\alpha (\mathbb{D}_x^{\alpha,\lambda} - 2\lambda^\alpha)u + \kappa_2 \kappa_\beta (\mathbb{D}_y^{\beta,\lambda} - 2\lambda^\beta)u = f & \Omega \times J, \\ u(\mathbf{x}, t) = u_{\text{in}} & \Gamma_{\text{in}} \times J, \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}) & \Omega, \\ u(\mathbf{x}, t) = 0 & \mathbb{R}^2 \backslash \bar{\Omega} \times J. \end{cases} \tag{2.14}$$

Considering the convection term $\mathbf{b}\cdot\nabla u$, we assume the weak solution $u \in H_0^1(\Omega_h)$ and formulate the DG variational formulation as: find $u \in V := H^1(0, T; H_0^1(\Omega_h))$, such that

$$\left(\frac{\partial u}{\partial t}, v\right) + b_h(u, v) + a_h(u, v) - \kappa(u, v) = (f, v) - \sum_{e \in \Gamma_{\text{in}}} \int_e \mathbf{b} \cdot \mathbf{n_e} u_{\text{in}} v \quad \forall v \in V, \tag{2.15}$$

where $\kappa = 2\lambda^\alpha \kappa_1 \kappa_\alpha + 2\lambda^\beta \kappa_2 \kappa_\beta > 0$,

$$b_h(u, v) = -(\mathbf{b}u, \nabla_h v) + \sum_{e \in \Gamma_{\text{i}}} \int_e \mathbf{b} \cdot \mathbf{n_e} \hat{u}[v] + \sum_{e \in \Gamma_{\text{out}}} \int_e \mathbf{b} \cdot \mathbf{n_e} uv, \tag{2.16}$$

$$a_h(u, v) = \kappa_1 \kappa_\alpha a_x^\alpha(u, v) + \kappa_2 \kappa_\beta a_y^\beta(u, v) + J_0(u, v), \tag{2.17}$$

and

$$a_x^\alpha(u, v) = ({}_aD_x^{\frac{\alpha}{2},\lambda}u, {}_xD_b^{\frac{\alpha}{2},\lambda}v) + ({}_aD_x^{\frac{\alpha}{2},\lambda}v, {}_xD_b^{\frac{\alpha}{2},\lambda}u),$$

$$a_y^\beta(u, v) = ({}_cD_y^{\frac{\beta}{2},\lambda}u, {}_yD_d^{\frac{\beta}{2},\lambda}v) + ({}_cD_y^{\frac{\beta}{2},\lambda}v, {}_yD_d^{\frac{\beta}{2},\lambda}u).$$

∇_h denotes the piecewise gradient operator.

The numerical flux $\hat{u}$ in $b_h(u, v)$ stems from integration by part. It is a single valued function defined on the faces and should be designed based on different guiding principles for different PDEs to guarantee stability and convergence. Here we choose an upwind flux: $\hat{u} := u^{\text{up}}$. We recall that $\mathbf{n_e}$ is a unit normal vector pointing from T_e^1 to T_e^2:

$$u^{\text{up}} = \begin{cases} u|_{T_e^1}, & \text{if } \mathbf{b} \cdot \mathbf{n_e} \geq 0, \\ u|_{T_e^2}, & \text{if } \mathbf{b} \cdot \mathbf{n_e} < 0. \end{cases} \tag{2.18}$$

We discretize the time derivative with Euler backward difference scheme. Let N_T be a positive integer and $\tau = T/N_T$ denote the time step. We also use the notations:

$$t_n = n \cdot \tau, \quad u^n(\mathbf{x}) = u(\mathbf{x}, t_n) \quad \forall n \geq 0.$$

Define the L^2 projection operator $P_h : L^2(\Omega) \to V_h$, i.e., for each element T_i,

$$(P_h u - u, v)_{T_i} = 0 \qquad \forall v \in P_N(T_i). \tag{2.19}$$

The fully discrete DG scheme is as follows: find $u_h^{n+1} \in V_h$, such that $\forall v \in V_h$,

$$\left(\frac{u_h^{n+1} - u_h^n}{\tau}, v\right) + b_h(u_h^{n+1}, v) + a_h(u_h^{n+1}, v) - \kappa(u_h^{n+1}, v)$$

$$= (f^{n+1}, v) - \sum_{e \in \Gamma_{\text{in}}} \int_e \mathbf{b} \cdot \mathbf{n_e} u_{\text{in}} v \tag{2.20}$$

with known u_h^n, and if $n = 0$, $u_h^0 = P_h u_0(x)$.

2.4. Stability analysis and error estimates

In this subsection, we discuss the stability and error estimates for the fully discrete scheme (2.20). First, we present three lemmas for preparation.

Lemma 2.5 (Discrete Grönwall inequality [34]). *Let $\tau, B, C > 0$ and $(a_n)_n$, $(b_n)_n$, $(c_n)_n$ be sequences of nonnegative numbers satisfying*

$$a_n + \tau \sum_{i=0}^{n} b_i \leq B + C\tau \sum_{i=0}^{n} a_i + \tau \sum_{i=0}^{n} c_i, \quad \forall n \geq 0.$$

Then, if $C\tau < 1$,

$$a_n + \tau \sum_{i=0}^{n} b_i \leq e^{C(n+1)\tau}(B + \tau \sum_{i=0}^{n} c_i), \quad \forall n \geq 0.$$

Lemma 2.6. *For any function $f(x) \in L^2(\Omega)$ and $g(x,t) \in L^1(0,T;L^2(\Omega))$, if $|f(x)| \leq \int_0^T g(x,t)dt$ for any $x \in \Omega$, then*

$$\|f(x)\|_{L^2(\Omega)}^2 \leq T \int_0^T \|g(x,t)\|_{L^2(\Omega)}^2 dt.$$

This inequality can be easily obtained by using Hölder's inequality. Here we omit the proof.

Lemma 2.7. *Let Ω_h be the conforming subdivision of Ω with non-overlapping triangles for some given h. Assume that $u \in H^1(\Omega_h)$ defined in (2.8) in two dimensions. Then for arbitrary small $\epsilon > 0$, $u \in H^{1/2-\epsilon}(\Omega)$ defined in (2.11).*

Proof. Considering the additive property of L^2 norm, we can easily obtain $u \in L^2(\Omega)$ from $u \in H^1(\Omega_h)$. The rest part is to prove $|u|_{H^s(\Omega)}^2 < \infty$ $\forall s \in (0, 1/2)$. Splitting the double integral of Aronszajn-Slobodeckij semi-norm in (2.11) into each element $T_i \in \Omega_h$ yields

$$\begin{aligned}
|u|_{H^s(\Omega)}^2 &= \int_\Omega \int_\Omega \frac{|u(\mathbf{x}) - u(\mathbf{y})|^2}{|\mathbf{x} - \mathbf{y}|^{2+2s}} d\mathbf{x} d\mathbf{y} \\
&= \sum_{i,j} I_{i,j} := \sum_{i,j} \int_{T_i} \int_{T_j} \frac{|u_i(\mathbf{x}) - u_j(\mathbf{y})|^2}{|\mathbf{x} - \mathbf{y}|^{2+2s}} d\mathbf{x} d\mathbf{y},
\end{aligned} \tag{2.21}$$

where $u_i(\mathbf{x}) = u(\mathbf{x})|_{T_i}$. The integrals in (2.21) can be divided into four categories: $T_i = T_j$; $\bar{T}_i \bigcap \bar{T}_j = \emptyset$; $\bar{T}_i \bigcap \bar{T}_j =$ a edge; $\bar{T}_i \bigcap \bar{T}_j =$ a vertex. For the fixed mesh Ω_h, the number of summations in (2.21) is finite. As a result, the following we want to show is that $I_{ij} < \infty$ for each i and j.

Since $u(\mathbf{x}) \in H^1(\Omega_h)$, then for each element T_i, we have $u(\mathbf{x})|_{T_i} \in H^1(T)$. Therefore, for the first case $T_i = T_j$, considering the embedding theorem, we have

$$I_{i,i} = |u|_{H^s(T_i)}^2 \leq C|u|_{H^1(T_i)}^2 < \infty.$$

For the second case $\bar{T}_i \bigcap \bar{T}_j = \emptyset$, the singular term $|\mathbf{x} - \mathbf{y}|^{-2-2s}$ can be bounded below by a positive constant, and then

$$I_{i,j} \leq C \int_{T_i} \int_{T_j} |u_i(\mathbf{x}) - u_j(\mathbf{y})|^2 d\mathbf{x} d\mathbf{y} \leq C \int_{T_i} \int_{T_j} u_i^2(\mathbf{x}) + u_j^2(\mathbf{y}) d\mathbf{x} d\mathbf{y} < \infty.$$

When T_i and T_j share a common edge $\mathbf{x}^*$ in the third case, we use a basic and well known technique in the boundary element method community relying on Duffy-type transforms for numerically computing such integrals [1,37]. Here we simply explain this method. First consider the affine mappings

$$\chi_i : \hat{T} \to T_i, \qquad \chi_i(\hat{\mathbf{x}}) = B_i \hat{\mathbf{x}} + \mathbf{x}_i^{(1)},$$

and we have

$$I_{i,j} = 4|T_i||T_j| \int_{\hat{T}} \int_{\hat{T}} \frac{|u_i(\chi_i(\hat{\mathbf{x}})) - u_j(\chi_j(\hat{\mathbf{y}}))|^2}{|\chi_i(\hat{\mathbf{x}}) - \chi_j(\hat{\mathbf{y}})|^{2+2s}} d\hat{\mathbf{x}} d\hat{\mathbf{y}}$$

$$= 4|T_i||T_j| \iiiint_{\hat{T} \times \hat{T}} F_{ij}(\hat{x}_1, \hat{x}_2, \hat{y}_1, \hat{y}_2) d\hat{x}_1 d\hat{x}_2 d\hat{y}_1 d\hat{y}_2.$$

Since the case of a common edge is considered, we assume that for the two adjacent elements T_i, T_j, the affine mappings satisfy

$$\chi_i(t, 0) = \chi_j(t, 0) \quad \text{for all} \ \ t \in [0, 1]. \tag{2.22}$$

Therefore, if we define $\hat{\mathbf{z}} = (\hat{y}_1 - \hat{x}_1, \hat{y}_2, \hat{x}_2)$, the singularity of the integrand is localized at $\hat{\mathbf{z}} = \mathbf{0}$:

$$I_{i,j} = 4|T_i||T_j| \int_0^1 \int_{-\hat{x}_1}^{1-\hat{x}_1} \int_0^{\hat{z}_1 + \hat{x}_1} \int_0^{\hat{x}_1} F_{ij}(\hat{x}_1, \hat{z}_3, \hat{x}_1 + \hat{z}_1, \hat{z}_2) d\hat{\mathbf{z}} d\hat{x}_1. \tag{2.23}$$

To deal with the singularity at $\hat{\mathbf{z}} = \mathbf{0}$, we decompose the domain of integration as five parts $D_k, 1 \leq k \leq 5$. Here for simplicity, we only list the first part since the others will be handled similarly (see [1, 37] for the other parts):

$$D_1 = \{(\hat{x}_1, \hat{\mathbf{z}}) : -1 \leq \hat{z}_1 \leq 0, 0 \leq \hat{z}_2 \leq 1 + \hat{z}_1, 0 \leq \hat{z}_3 \leq \hat{z}_2 - \hat{z}_1, \hat{z}_2 - \hat{z}_1 \leq \hat{x}_1 \leq 1\}. \tag{2.24}$$

Then let $\xi \in [0, 1]$ and $\eta = (\eta_1, \eta_2, \eta_3) \in [0, 1]^3$. For the integral in D_1, consider the mapping $T_1 : [0, 1] \times [0, 1]^3 \to D_1$,

$$T_1(\xi, \eta) = (\xi, -\xi\eta_1\eta_2, \xi\eta_1(1 - \eta_2), \xi\eta_1\eta_3)$$

with the Jacobian determinants $|JT_1| = \xi^3 \eta_1^2$. By now, we finally transform the singularity $\hat{\mathbf{z}} = \mathbf{0}$ into $\xi = 0$ and $\eta_1 = 0$:

$$\int_{D_1} F_{ij} = \int_{[0,1]} \int_{[0,1]^3} \frac{|u_i(\chi_i(\xi, \xi\eta_1\eta_3)) - u_j(\chi_j(\xi(1 - \eta_1\eta_2), \xi\eta_1(1 - \eta_2)))|^2}{|d_1(\eta)|^{2+2s}} \cdot \xi^{1-2s} \eta_1^2 d\eta d\xi, \tag{2.25}$$

where

$$d_1(\eta) = \chi_i(1, \eta_1\eta_3) - \chi_j(1 - \eta_1\eta_2, \eta_1(1 - \eta_2)).$$

Note that $d_1(\eta) = 0$ when $\eta_1 = 0$ due to the condition (2.22). It is easy to find $d_1(\eta) \sim \eta_1$ as $\eta_1 \to 0$ since the affine mappings χ_i, χ_j are linear, implying

$$\eta_1^2 |d_1(\eta)|^{-2-2s} \sim \eta_1^{-2s} \quad as \ \ \eta_1 \to 0.$$

Therefore, if u_i and u_j are bounded, $\int_{D_1} F_{ij} < \infty$ when $s < 1/2$.

However, $H^1(T_i)$ cannot be embedded into $C^0(\bar{T}_i)$ in two dimensions, which implies the numerator in (2.25) may be singular at $\eta_1 = 0$. We assume the singularity of u_i or u_j at the common vertex $\mathbf{x}^*$ is $|\mathbf{x} - \mathbf{x}^*|^{-p}(p > 0)$, i.e.,

$$|u_{i(j)}(\mathbf{x})| \sim |\mathbf{x} - \mathbf{x}^*|^{-p} \quad \text{as} \quad \mathbf{x} \to \mathbf{x}^*.$$

So if $p < 1/2 - s$, the integral $I_{i,j}$ converges while $I_{i,j}$ diverges when $p \geq 1/2 - s$. But in the latter case, we can show a contradiction that $u_i \notin H^1(T_i)$. More precisely, for any given $s \in (0, 1/2)$, since $p \geq 1/2 - s > 0$, there exists a positive constant ϵ satisfying $\epsilon < p$ and $|\mathbf{x} - \mathbf{x}^*|^{-p} > |\mathbf{x} - \mathbf{x}^*|^{-\epsilon}$. Then it can be easily verified that $|\mathbf{x} - \mathbf{x}^*|^{-\epsilon} \notin H^1(T_i)$, which shows

$|\mathbf{x} - \mathbf{x}^*|^{-p} \notin H^1(T_i)$. So for $u_i \in H^1(T_i)$, the worst case is $u \sim |\mathbf{x} - \mathbf{x}^*|^{-\epsilon}$ as $\mathbf{x} \to \mathbf{x}^*$ with arbitrarily small $\epsilon > 0$. In this case, $I_{i,j}$ converges for any fixed s satisfying $0 < s < \frac{1}{2}$ where $\epsilon < \frac{1}{2} - s$.

Similarly, the integrals $I_{i,j}$ in $D_k, 2 \le k \le 5$ also converge. The fourth case is also valid with the same technique as the one used in the third case, but with different domain decomposition and Duffy-type transform; see details in [1, 37]. Finally, for any $s \in (0, \frac{1}{2})$,

$$|u|^2_{H^s(\Omega)} = \sum_{i,j} I_{i,j} < \infty,$$

which implies for arbitrary small $\epsilon > 0$, $u \in H^{1/2-\epsilon}(\Omega)$.   □

Following Lemma 2.7, one simple example is that the piecewise constant function in two dimensions belongs to $H^{1/2-\epsilon}(\Omega)$, which is the same as the one in one dimension. So for the weak solution u in DG method, it still has weak regularity $H^{1/2-\epsilon}(\Omega)$ in the whole domain. Then for any $\alpha, \beta \in (0,1)$, we have $u \in H^{\alpha/2}(\Omega) \bigcap H^{\beta/2}(\Omega)$, which is the keystone of the subsequent stability and convergence analyses. In the following, C denotes a generic constant independent of h and τ, which takes different values in different occurrences. First, we prove the coercivity and continuity of the bilinear form $a_h(u, v)$. Define the energy norm on V as

$$\|v\|^2_{E(\Omega_h)} = |v|^2_{H_x^{\alpha/2,\lambda}(\Omega)} + |v|^2_{H_y^{\beta/2,\lambda}(\Omega)} + \sum_{e \in \Gamma_i} \int_e [v]^2, \tag{2.26}$$

where

$$|v|^2_{H_x^{\alpha/2,\lambda}(\Omega)} = \|_a D_x^{\alpha/2,\lambda} v\|^2_{L^2(\Omega)}, \quad |v|^2_{H_y^{\beta/2,\lambda}(\Omega)} = \|_c D_y^{\beta/2,\lambda} v\|^2_{L^2(\Omega)}.$$

Then from (2.7) and (2.17), there exists a positive constant $\gamma < 1$ such that

$$\begin{aligned} a_h(v, v) &= 2\kappa_1 \kappa_\alpha \left({}_a D_x^{\frac{\alpha}{2},\lambda} v, {}_x D_b^{\frac{\alpha}{2},\lambda} v \right) + 2\kappa_2 \kappa_\beta \left({}_c D_y^{\frac{\beta}{2},\lambda} v, {}_y D_d^{\frac{\beta}{2},\lambda} v \right) + \sum_{e \in \Gamma} \int_e [v]^2 \\ &\ge \gamma |v|^2_{H_x^{\alpha/2,\lambda}(\Omega)} + \gamma |v|^2_{H_y^{\beta/2,\lambda}(\Omega)} + \sum_{e \in \Gamma} \int_e [v]^2 \\ &\ge \gamma \|v\|^2_{E(\Omega_h)}. \end{aligned} \tag{2.27}$$

On the other hand, from the definition of $a_h(u, v)$, and Cauchy-Schwarz's inequality, we have

$$\begin{aligned} |a_h(u, v)| &\le |\kappa_1 \kappa_\alpha| \left(\left| \left({}_a D_x^{\frac{\alpha}{2},\lambda} u, {}_x D_b^{\frac{\alpha}{2},\lambda} v \right) \right| + \left| \left({}_a D_x^{\frac{\alpha}{2},\lambda} v, {}_x D_b^{\frac{\alpha}{2},\lambda} u \right) \right| \right) \\ &\quad + |\kappa_2 \kappa_\beta| \left(\left| \left({}_c D_y^{\frac{\beta}{2},\lambda} u, {}_y D_d^{\frac{\beta}{2},\lambda} v \right) \right| + \left| \left({}_c D_y^{\frac{\beta}{2},\lambda} v, {}_y D_d^{\frac{\beta}{2},\lambda} u \right) \right| \right) + \sum_{e \in \Gamma} \int_e [u][v] \\ &\le C |u|_{H_x^{\alpha/2,\lambda}(\Omega)} \cdot |v|_{H_x^{\alpha/2,\lambda}(\Omega)} + C |u|_{H_y^{\beta/2,\lambda}(\Omega)} \cdot |v|_{H_y^{\beta/2,\lambda}(\Omega)} + \sum_{e \in \Gamma} \int_e [u][v] \\ &\le C \|u\|_{E(\Omega_h)} \cdot \|v\|_{E(\Omega_h)}. \end{aligned} \tag{2.28}$$

Next we deal with another bilinear form $b_h(u, v)$. For the upwind flux, there is

$$
\begin{aligned}
b_h(v, v) &= -(\mathbf{b}v, \nabla_h v) + \sum_{e \in \Gamma_i} \int_e \mathbf{b} \cdot \mathbf{n_e} v^{\mathrm{up}}[v] + \sum_{e \in \Gamma_{\mathrm{out}}} \int_e \mathbf{b} \cdot \mathbf{n_e} v^2 \\
&= \sum_{e \in \Gamma_i} \int_e \mathbf{b} \cdot \mathbf{n_e} \left(v^{\mathrm{up}}[v] - \frac{1}{2}[v^2] \right) - \frac{1}{2} \sum_{e \in \partial\Omega} \int_e \mathbf{b} \cdot \mathbf{n}_e v^2 + \sum_{e \in \Gamma_{\mathrm{out}}} \int_e \mathbf{b} \cdot \mathbf{n_e} v^2 \\
&= \sum_{e \in \Gamma_i} \int_e \mathbf{b} \cdot \mathbf{n_e} \left(v^{\mathrm{up}}[v] - \{v\}[v] \right) - \frac{1}{2} \sum_{e \in \Gamma_{\mathrm{in}}} \int_e \mathbf{b} \cdot \mathbf{n}_e v^2 + \frac{1}{2} \sum_{e \in \Gamma_{\mathrm{out}}} \int_e \mathbf{b} \cdot \mathbf{n_e} v^2 \\
&= \frac{1}{2} \sum_{e \in \Gamma_i} \int_e |\mathbf{b} \cdot \mathbf{n_e}| [v]^2 + \frac{1}{2} \sum_{e \in \Gamma_{\mathrm{in}}} \int_e |\mathbf{b} \cdot \mathbf{n}_e| v^2 + \frac{1}{2} \sum_{e \in \Gamma_{\mathrm{out}}} \int_e |\mathbf{b} \cdot \mathbf{n_e}| v^2 \geq 0. \quad (2.29)
\end{aligned}
$$

Now we examine the stability property of the scheme (2.20).

Theorem 2.1. *For absorbing boundary conditions, the fully discrete scheme* (2.20) *is uncon-ditionally stable, and there exists a positive constant C independent of h and τ, such that for all $m > 0$,*

$$
\|u_h^m\|_{L^2(\Omega)}^2 + 2\tau\gamma \sum_{n=1}^m \|u_h^n\|_{E(\Omega_h)}^2
$$
$$
\leq C \left(\|u_h^0\|_{L^2(\Omega)}^2 + \tau \sum_{n=1}^m \|f^n\|_{L^2(\Omega)}^2 + \tau h^{-1} \sum_{e \in \Gamma_{\mathrm{in}}} \|u_{\mathrm{in}}\|_{L^2(e)}^2 \right). \quad (2.30)
$$

Proof. Taking $v = u_h^{n+1}$ in (2.20), using (2.27), (2.29), and Cauchy-Schwarz's inequality,

$$
\frac{1}{\tau}(u_h^{n+1} - u_h^n, u_h^{n+1}) + \gamma\|u_h^{n+1}\|_{E(\Omega_h)}^2 - \kappa\|u_h^{n+1}\|_{L^2(\Omega)}^2
$$
$$
\leq \|f^{n+1}\|_{L^2(\Omega)} \cdot \|u_h^{n+1}\|_{L^2(\Omega)} + \sum_{e \in \Gamma_{\mathrm{in}}} \|\mathbf{b}\|_{L^\infty(e)} \|u_{\mathrm{in}}\|_{L^2(e)}^2 \|u_h^{n+1}\|_{L^2(e)}.
$$

Noting that the inequality:

$$
(x^2 - y^2) \leq (x^2 - y^2 + (x - y)^2) = 2(x - y)x \quad \forall x, y \in \mathbb{R},
$$

together with Young's inequality and trace theorem, we obtain

$$
\frac{1}{2\tau}(\|u_h^{n+1}\|_{L^2(\Omega)}^2 - \|u_h^n\|_{L^2(\Omega)}^2) + \gamma\|u_h^{n+1}\|_{E(\Omega_h)}^2
$$
$$
\leq \frac{1}{2}\|f^{n+1}\|_{L^2(\Omega)}^2 + (\kappa + 1)\|u_h^{n+1}\|_{L^2(\Omega)}^2 + \frac{C}{h} \sum_{e \in \Gamma_{\mathrm{in}}} \|u_{\mathrm{in}}\|_{L^2(e)}^2. \quad (2.31)
$$

Multiplying (2.31) by 2τ and summing from $n = 0$ to $n = m - 1$ lead to

$$
\|u_h^m\|_{L^2(\Omega)}^2 - \|u_h^0\|_{L^2(\Omega)}^2 + 2\tau\gamma \sum_{n=1}^m \|u_h^n\|_{E(\Omega_h)}^2
$$
$$
\leq \tau \sum_{n=1}^m \|f^n\|_{L^2(\Omega)}^2 + (2\kappa + 2)\tau \sum_{n=1}^m \|u_h^n\|_{L^2(\Omega)}^2 + \frac{C\tau}{h} \sum_{e \in \Gamma_{\mathrm{in}}} \|u_{\mathrm{in}}\|_{L^2(e)}^2.
$$

Then the desired result is obtained by discrete Grönwall inequality of Lemma 2.5. $\qquad\square$

Remark 2.2. It can be seen that the stability is unconditional for the penalty parameter when $\alpha, \beta \in (0, 1)$. The essential reason is that for classical problem integration by part yields the boundary term $\{\nabla v \cdot \mathbf{n_e}\}[v]$ [34, pp. 38-39], which makes coercivity valid only when penalty parameter is sufficiently large. But for (tempered) fractional problem, this boundary term does not exist in the weak form (2.15).

The solutions to fractional equations typically have low regularity, but it is not too low everywhere. Commonly, the solutions have low regularity at the boundary and the higher regularity in the interior [2]. We assume that the solution of (2.15) is sufficiently regular in each element for $\alpha, \beta \in (0, 1)$ here, and we prove the convergence order in L^2 norm to be $\mathcal{O}(h^{N-\epsilon} + \tau)$. On the contrary, for the case of the solution with low regularity, the adaptive DG scheme will be considered in the next section.

Theorem 2.2. *Let u^n be the exact solution of* (2.15), *u_h^n the numerical solution of the fully discrete scheme* (2.20). *Then*

$$\|u^{N_T} - u_h^{N_T}\|_{L^2(\Omega)}^2 \leq Ch^{2N-\epsilon}\|u\|_{W^{1,\infty}(0,T;H^{N+1}(\Omega_h))}^2 + C\tau^2\|u\|_{H^2(0,T;L^2(\Omega))}^2. \tag{2.32}$$

Proof. As usual, we denote the error $e^n = u^n - u_h^n$ by two parts $\rho^n = u^n - P_h u^n$ and $\theta^n = u_h^n - P_h u^n$. Subtracting (2.20) from (2.15) , we have

$$\left(\frac{\partial u^{n+1}}{\partial t} - \frac{u^{n+1} - u^n}{\tau}, v\right) + \left(\frac{e^{n+1} - e^n}{\tau}, v\right) + b_h(e^{n+1}, v) + a_h(e^{n+1}, v) - \kappa(e^{n+1}, v) = 0.$$

Noting that $e^n = \rho^n - \theta^n$, we get

$$\left(\frac{\theta^{n+1} - \theta^n}{\tau}, v\right) + b_h(\theta^{n+1}, v) + a_h(\theta^{n+1}, v) - \kappa(\theta^{n+1}, v)$$

$$= \left(\frac{\rho^{n+1} - \rho^n}{\tau}, v\right) + b_h(\rho^{n+1}, v) + a_h(\rho^{n+1}, v) - \kappa(\rho^{n+1}, v) + \left(\frac{\partial u^{n+1}}{\partial t} - \frac{u^{n+1} - u^n}{\tau}, v\right).$$

Taking $v = \theta^{n+1}$, similarly to the proof of stability, we obtain

$$\frac{1}{2\tau}(\|\theta^{n+1}\|_{L^2(\Omega)}^2 - \|\theta^n\|_{L^2(\Omega)}^2) + \gamma\|\theta^{n+1}\|_{E(\Omega_h)}^2 - \kappa\|\theta^{n+1}\|_{L^2(\Omega)}^2 \leq \sum_{i=1}^{5} |T_i|, \tag{2.33}$$

where

$$T_1 = (\frac{\rho^{n+1} - \rho^n}{\tau}, \theta^{n+1}), \quad T_2 = b_h(\rho^{n+1}, \theta^{n+1}), \quad T_3 = a_h(\rho^{n+1}, \theta^{n+1}),$$

$$T_4 = \kappa(\rho^{n+1}, \theta^{n+1}), \quad T_5 = (\frac{\partial u^{n+1}}{\partial t} - \frac{u^{n+1} - u^n}{\tau}, \theta^{n+1}).$$

Since $\rho^{n+1} - \rho^n = \int_{t_n}^{t_{n+1}} \rho_t dt$, by Lemma 2.6, we have

$$\|\rho^{n+1} - \rho^n\|_{L^2(\Omega)}^2 \leq \tau \int_{t_n}^{t_{n+1}} \|\rho_t\|_{L^2(\Omega)}^2 dt.$$

Hence, with Hölder's, Young's inequalities, we obtain

$$|T_1| \leq \left\|\frac{\rho^{n+1} - \rho^n}{\tau}\right\|_{L^2(\Omega)} \cdot \|\theta^{n+1}\|_{L^2(\Omega)}$$

$$\leq \frac{3}{4\epsilon_1} \cdot \frac{1}{\tau} \int_{t_n}^{t_{n+1}} \|\rho_t\|_{L^2(\Omega)}^2 dt + \frac{\epsilon_1}{3}\|\theta^{n+1}\|_{L^2(\Omega)}^2$$

$$\leq C\tau^{-1}h^{2N+2} \int_{t_n}^{t_{n+1}} |u_t|_{H^{N+1}(\Omega_h)}^2 dt + \frac{\epsilon_1}{3}\|\theta^{n+1}\|_{L^2(\Omega)}^2.$$

Considering the property of the projection P_h (2.19) and the piecewise constant $\mathbf{b}$, together with trace inequalities, we have

$$|T_2| \leq |(\mathbf{b}\rho^{n+1}, \nabla_h \theta^{n+1})| + \left| \sum_{e \in \Gamma} \int_e \mathbf{b} \cdot \mathbf{n} \hat{\rho}^{n+1} [\theta^{n+1}] \right|$$

$$\leq \|\mathbf{b}\|_{L^\infty(\Omega)} \sum_{e \in \Gamma} \|\hat{\rho}^{n+1}\|_{L^2(e)} \cdot \|[\theta^{n+1}]\|_{L^2(e)}$$

$$\leq \|\mathbf{b}\|_{L^\infty(\Omega)}^2 \frac{1}{4\epsilon_2} \sum_{e \in \Gamma} \|\hat{\rho}^{n+1}\|_{L^2(e)}^2 + \epsilon_2 \sum_{e \in \Gamma} \|[\theta^{n+1}]\|_{L^2(e)}^2$$

$$\leq Ch^{2N+1} |u^{n+1}|_{H^{N+1}(\Omega_h)}^2 + \epsilon_2 \sum_{e \in \Gamma} \|[\theta^{n+1}]\|_{L^2(e)}^2.$$

From the continuity of $a(u, v)$ (2.28), we obtain

$$|T_3| \leq C\|\rho^{n+1}\|_{E(\Omega_h)} \cdot \|\theta^{n+1}\|_{E(\Omega_h)}$$

$$\leq \frac{C^2}{4\epsilon_3} \|\rho^{n+1}\|_{E(\Omega_h)}^2 + \epsilon_3 \|\theta^{n+1}\|_{E(\Omega_h)}^2$$

$$\leq Ch^{2N+1-\epsilon} |u^{n+1}|_{H^{N+1}(\Omega_h)}^2 + \epsilon_3 \|\theta^{n+1}\|_{E(\Omega_h)}^2.$$

In the last inequality, using embedding theorem, the first two terms in the energy norm $\|\rho^{n+1}\|_{E(\Omega_h)}$ (2.26) both can be bounded by $\|\rho^{n+1}\|_{H^{1/2-\epsilon}(\Omega)}$ since $\alpha, \beta < 1$. Then, basing on Lemma 2.7, we have the approximation property [6],

$$\|u - P_h u\|_{H^{1/2-\epsilon}(\Omega)} \leq C\|u - P_h u\|_{H^1(\Omega_h)} \leq Ch^N |u|_{H^{N+1}(\Omega_h)}. \tag{2.34}$$

For the last term in the energy norm $\|\rho^{n+1}\|_{E(\Omega_h)}$, using trace theorem,

$$\sum_{e \in \Gamma} \int_e [\rho^{n+1}]^2 \leq Ch^{-1} \|\rho^{n+1}\|_{L^2(\Omega)}^2 + h\|\nabla_h \rho^{n+1}\|_{L^2(\Omega)}^2 \leq Ch^{2N+1} |u^{n+1}|_{H^{N+1}(\Omega_h)}^2.$$

Similarly, for the fourth term, we have

$$|T_4| \leq \kappa \|\rho^{n+1}\|_{L^2(\Omega)} \cdot \|\theta^{n+1}\|_{L^2(\Omega)}$$

$$\leq \kappa^2 \frac{3}{4\epsilon_1} \|\rho^{n+1}\|_{L^2(\Omega)}^2 + \frac{\epsilon_1}{3} \|\theta^{n+1}\|_{L^2(\Omega)}^2$$

$$\leq Ch^{2N+2} |u^{n+1}|_{H^{N+1}(\Omega_h)}^2 + \frac{\epsilon_1}{3} \|\theta^{n+1}\|_{L^2(\Omega)}^2.$$

Note that the Taylor expansion with integral remainder has the form

$$u^n = u^{n+1} - \tau u_t^{n+1} + \int_{t_{n+1}}^{t_n} (t_n - t) u_{tt}(t) dt.$$

Thus we have

$$\left| \frac{\partial u^{n+1}}{\partial t} - \frac{u^{n+1} - u^n}{\tau} \right| \leq \frac{1}{\tau} \int_{t_n}^{t_{n+1}} (t - t_n) u_{tt}(t) dt \leq \int_{t_n}^{t_{n+1}} |u_{tt}(t)| dt,$$

and then

$$|T_5| \leq \left\| \frac{\partial u^{n+1}}{\partial t} - \frac{u^{n+1} - u^n}{\tau} \right\|_{L^2(\Omega)} \cdot \|\theta^{n+1}\|_{L^2(\Omega)}$$

$$\leq \frac{3}{4\epsilon_1} \left\| \frac{\partial u^{n+1}}{\partial t} - \frac{u^{n+1} - u^n}{\tau} \right\|_{L^2(\Omega)}^2 + \frac{\epsilon_1}{3} \|\theta^{n+1}\|_{L^2(\Omega)}^2$$

$$\leq C\tau \int_{t_n}^{t_{n+1}} \|u_{tt}\|_{L^2(\Omega)}^2 dt + \frac{\epsilon_1}{3} \|\theta^{n+1}\|_{L^2(\Omega)}^2.$$

Substituting T_i, $1 \leq i \leq 5$, into (2.33), we have

$$\frac{1}{2\tau}(\|\theta^{n+1}\|_{L^2(\Omega)}^2 - \|\theta^n\|_{L^2(\Omega)}^2) + (\gamma - \epsilon_3)\|\theta^{n+1}\|_{E(\Omega_h)}^2$$

$$\leq (\kappa + \epsilon_1)\|\theta^{n+1}\|_{L^2(\Omega)}^2 + \epsilon_2 \sum_{e \in \Gamma} \|[\theta^{n+1}]\|_{L^2(e)}^2 + C\tau \int_{t_n}^{t_{n+1}} \|u_{tt}\|_{L^2(\Omega)}^2 dt$$

$$+ C\tau^{-1}h^{2N+2} \int_{t_n}^{t_{n+1}} |u_t|_{H^{N+1}(\Omega_h)}^2 dt + Ch^{2N-\epsilon}|u^{n+1}|_{H^{N+1}(\Omega_h)}^2. \qquad (2.35)$$

From the definition of energy norm (2.26), we know

$$\|\theta^{n+1}\|_{E(\Omega_h)}^2 \geq \sum_{e \in \Gamma} \|[\theta^{n+1}]\|_{L^2(e)}^2.$$

Therefore, choosing ϵ_2 and ϵ_3 small enough such that $\epsilon_2 + \epsilon_3 \leq \gamma$ leads to

$$(\gamma - \epsilon_3)\|\theta^{n+1}\|_{E(\Omega_h)}^2 - \epsilon_2 \sum_{e \in \Gamma} \|[\theta^{n+1}]\|_{L^2(e)}^2 \geq 0.$$

Multiplying (2.35) by 2τ, summing over n from 0 to $N_T - 1$, and $\theta^0 = 0$, we obtain

$$\|\theta^{N_T}\|_{L^2(\Omega)}^2 \leq 2\tau(\kappa + \epsilon_1) \sum_{n=1}^{N_T} \|\theta^n\|_{L^2(\Omega)}^2 + Ch^{2N+2} \int_0^T \|u_t\|_{L^2(\Omega)}^2 dt$$

$$+ C\tau^2 \int_0^T \|u_{tt}\|_{L^2(\Omega)}^2 dt + Ch^{2N-\epsilon}\|u\|_{L^\infty(0,T;H^{N+1}(\Omega_h))}^2$$

$$\leq 2\tau(\kappa + \epsilon_1) \sum_{n=1}^{N_T} \|\theta^n\|_{L^2(\Omega)}^2 + Ch^{2N-\epsilon}\|u\|_{W^{1,\infty}(0,T;H^{N+1}(\Omega_h))}^2 + C\tau^2\|u\|_{H^2(0,T;L^2(\Omega))}^2.$$

By the discrete Grönwall inequality of Lemma 2.5 and the triangle inequality, we obtain the desired result.

2.5. Numerical experiments

In this section, we offer the numerical performance of the proposed schemes with three examples to validate the preceding theoretical analysis. We use the backward Euler discretization to solve the method-of-line fractional PDE, i.e., the classical ODE system. We take the sufficiently small time steps $\tau \sim h^{N+1}$ to observe the spatial convergence order. As to the spatial approximation, we adopt the interpolation bases [24].

We define the local mass matrix M^k and the local spatial stiffness matrix S_x^k, S_y^k at element T^k as

$$M_{ij}^k = (l_i^k(\mathbf{x}), l_j^k(\mathbf{x}))_{T^k}, \qquad (S_{x/y}^k)_{ij} = \left(\frac{\partial l_j^k(\mathbf{x})}{\partial x/y}, l_i^k(\mathbf{x})\right)_{T^k},$$

where the notation x/y denotes x or y. It is a little complicated to build the tempered fractional spatial stiffness matrix, since tempered fractional operators are nonlocal and we need all the information of the related elements in x direction or y direction when generating any stiffness matrix of an element. The algorithm of fractional spatial stiffness matrix in two dimensions was introduced in [33], where fractional diffusion equations was solved by local DG method. We use the part of discretzing fractional operator in [33] and get the global tempered fractional spatial stiffness matrices $_lG_x, _lG_y, _rG_x, _rG_y$, where '$l/r$' denote left/right tempered fractional derivative. With the above notations, we get the matrix form of the fully discrete form (2.20) as

$$M\frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^n}{\Delta t} + b_1 S_x \mathbf{u}_h^{n+1} + \kappa_1 \kappa_\alpha (_lG_x + _rG_x)\mathbf{u}_h^{n+1}$$
$$+ b_2 S_y \mathbf{u}_h^{n+1} + \kappa_2 \kappa_\beta (_lG_y + _rG_y)\mathbf{u}_h^{n+1} - \kappa M \mathbf{u}_h^{n+1} = M\mathbf{F}_h^{n+1},$$

where M, S_x, S_y are global mass and stiffness matrices, and their non-zero diagonal blocks are constructed by M^k, S_x^k, and S_y^k respectively.

Example 2.1. Consider the problem

$$\frac{\partial u}{\partial t} + \mathbf{b} \cdot \nabla u - \kappa_1 \nabla_x^{\alpha,\lambda} u - \kappa_2 \nabla_y^{\beta,\lambda} u = f, \tag{2.36}$$

where $\mathbf{b} = (0.5, 0.5)$, $\kappa_1 = 0.1$, $\kappa_2 = 0.2$, $\lambda = 0.2$, $T = 1$ and $\alpha, \beta \in (0,1)$ on the computational domain $\Omega = (0,2) \times (0,2)$. Its exact solution is $u = \mathrm{e}^{-t} x^{3.8}(2-x)^{3.8} y^{3.8}(2-y)^{3.8}$ with appropriate initial and boundary conditions.

Example 2.2. Consider the same problem (2.36), with a different source term f so that the exact solution $u = \mathrm{e}^{-t} \sin\frac{\pi}{2}x \sin\frac{\pi}{2}y$.

Example 2.3. Consider the problem (2.36) with the same parameters. But now the exact solution u is unknown. We assume that the smooth source term $f = 1$ and the initial condition $u_{01} = \sin\frac{\pi}{2}x \sin\frac{\pi}{2}y$ or $u_{02} = xy$.

Since the exact solution is unknown, the rate of convergence is calculated by means of mesh-nesting technique for this example. More precisely, there is

$$\text{rate} = \frac{\ln(e_h/e_{2h})}{\ln(2)}, \tag{2.37}$$

where $e_h = \|u_h - u_{2h}\|$, and u_h denotes the numerical solution under mesh size h. The corresponding results for different initial condition u_{01} and u_{02} are presented in Tables 2.3 and 2.4, respectively.

Table 2.1 and Table 2.2 list the L^2 errors and convergence orders for different parameters (α, β) in different DG finite element space P_N, and K the total number of triangle elements. It can be seen that the convergence order is better than the theoretical result $h^{N-\epsilon}$, for any of $(\alpha, \beta) \in (0,1)$. For the solution with relatively low regularity in the first example, the

Table 2.1: Numerical errors (L^2) and orders of convergence on triangular meshes for Example 2.1.

N	(α,β)	K 72 error	200 error	order	392 error	order	648 error	order
1	(0.2,0.2)	7.03e-2	2.60e-2	1.95	1.31e-2	2.03	7.83e-3	2.04
	(0.5,0.5)	6.45e-2	2.35e-2	1.98	1.18e-2	2.05	6.99e-3	2.07
	(0.7,0.2)	6.67e-2	2.44e-2	1.97	1.23e-2	2.05	7.29e-3	2.06
		K 72	200		392		648	
2	(0.2,0.2)	8.90e-3	1.68e-3	3.27	7.02e-4	2.59	3.86e-4	2.38
	(0.5,0.5)	7.73e-3	1.45e-3	3.27	6.23e-4	2.52	3.39e-4	2.42
	(0.7,0.2)	8.07e-3	1.50e-3	3.30	6.33e-4	2.56	3.43e-4	2.43
		K 72	200		392		648	
3	(0.2,0.2)	1.05e-3	1.49e-4	3.82	4.78e-5	3.38	1.95e-5	3.57
	(0.5,0.5)	9.64e-4	1.32e-4	3.89	4.00e-5	3.55	1.56e-5	3.75
	(0.7,0.2)	9.85e-4	1.35e-4	3.89	4.10e-5	3.53	1.59e-5	3.77

Table 2.2: Numerical errors (L^2) and orders of convergence on triangular meshes for Example 2.2.

N	(α,β)	K 72 error	200 error	order	392 error	order	648 error	order
1	(0.2,0.2)	4.61e-2	1.67e-2	1.98	8.51e-3	2.01	5.12e-3	2.02
	(0.5,0.5)	4.41e-2	1.60e-2	1.99	8.11e-3	2.01	4.88e-3	2.02
	(0.7,0.2)	4.49e-2	1.63e-2	1.99	8.26e-3	2.02	4.96e-3	2.03
		K 72	200		392		648	
2	(0.2,0.2)	3.13e-3	6.88e-4	2.96	2.70e-4	2.78	1.35e-4	2.77
	(0.5,0.5)	2.80e-3	6.20e-4	2.95	2.43e-4	2.78	1.21e-4	2.77
	(0.7,0.2)	2.89e-3	6.35e-4	2.97	2.48e-4	2.79	1.23e-4	2.79
		K 72	200		392		648	
3	(0.2,0.2)	3.12e-4	4.06e-5	4.00	1.09e-5	3.91	4.11e-6	3.87
	(0.5,0.5)	3.00e-4	3.88e-5	4.01	1.03e-5	3.93	3.87e-6	3.91
	(0.7,0.2)	3.05e-4	3.93e-5	4.01	1.05e-5	3.93	3.92e-6	3.92

Table 2.3: Numerical errors (L^2 and L^∞) and orders of convergence on nesting triangular meshes for Example 2.3 with initial condition u_{01} and source term $f = 1$.

(α,β)	Penalty	K Norm	8 and 32 error	32 and 128 error	order	128 and 512 error	order
(0.5,0.5)	Yes	L^2	1.63e-1	2.37e-2	2.78	2.97e-3	2.99
		L^∞	4.26e-1	1.21e-1	1.81	3.71e-2	1.71
	No	L^2	1.63e-1	2.38e-2	2.78	3.01e-3	2.98
		L^∞	4.60e-1	1.29e-1	1.84	3.96e-2	1.70
		K	8 and 32	32 and 128		128 and 512	
(0.2,0.7)	Yes	L^2	1.62e-1	2.35e-2	2.79	2.91e-3	3.01
		L^∞	4.30e-1	1.45e-1	1.57	3.84e-2	1.92
	No	L^2	1.63e-1	2.36e-2	2.79	3.01e-3	2.97
		L^∞	4.60e-1	1.53e-1	1.59	4.43e-2	1.79

Table 2.4: Numerical errors (L^2 and L^∞) and orders of convergence on nesting triangular meshes for Example 2.3 with initial condition u_{02} and source term $f = 1$.

(α, β)	Penalty	K Norm	8 and 32 error	32 and 128 error	order	128 and 512 error	order
(0.5,0.5)	Yes	L^2	7.06e-2	1.25e-2	2.50	1.65e-3	2.92
		L^∞	2.67e-1	8.58e-2	1.64	3.02e-2	1.51
	No	L^2	7.18e-2	1.53e-2	2.23	1.72e-3	3.15
		L^∞	2.51e-1	9.60e-2	1.39	4.42e-2	1.12
		K	8 and 32	32 and 128		128 and 512	
(0.2,0.7)	Yes	L^2	7.04e-2	1.24e-2	2.51	1.31e-3	3.24
		L^∞	2.65e-1	8.52e-2	1.64	3.19e-2	1.42
	No	L^2	7.16e-2	1.53e-2	2.23	1.66e-3	3.20
		L^∞	2.52e-1	9.58e-2	1.40	4.71e-2	1.02

convergence order is shown to be relatively low for higher order polynomial $N = 2, 3$. One of the advantage of DG method is its flexibility and ability to support high-order accuracy while maintaining a large degree of locality in the formulation, we choose the nice analytic solutions in the two examples to verify the high convergence order of the DG scheme [16].

In Tables 2.3 and 2.4, we compare the numerical results for the schemes with or without the penalty terms, denoted as "Yes" and "No" in the table. We only use the linear polynomial $N = 1$ for simplicity in this example. The exact solution should be continuous for a smooth source term $f = 1$. No obvious differences can be found in L^2 errors and L^∞ ones whether the penalty terms exist or not. The errors are slightly smaller for the scheme with penalty terms than those of the scheme without penalty terms, especially for L^∞ norm.

3. Adaptive DG Algorithm

This section focuses on the adaptive DG scheme for the fractional diffusion equations. We derive posteriori error estimates, and design the local error indicators. The numerical experiments are performed to show the performances of the adaptive schemes.

3.1. Stationary Equation

Through the analyses in Section 2, there is no essential difference between tempered fractional derivatives and fractional derivatives. So we consider the fractional stationary equation for simplicity,

$$\begin{cases} \mathbb{D}_x^\alpha u + \mathbb{D}_y^\beta u = f & \Omega, \\ u = 0 & \mathbb{R}^2 \backslash \Omega, \end{cases} \tag{3.1}$$

and also assume the weak solution $u \in V = H^{1/2-\epsilon}(\Omega)$ for arbitrary small ϵ.

A posteriori error estimators are the essential ingredient of adaptivity, which are computable quantities depending on the computed solution and data that provide information about the quality of approximation and may thus be used to make judicious mesh modifications. The ultimate purpose is to construct the estimator of meshes that will eventually be equivalent to the exact error. We present two kinds of methods to construct the estimator: energy norm method and DWR method.

3.1.1. Scheme 1 – Energy norm method [6]

In this scheme, we assume $\alpha = \beta$ for simplicity. Then for each element $T \in \Omega_h$, we define the local error estimator η_T by

$$\eta_T^2 = \eta_{1,T}^2 + \eta_{2,T}^2 = h_T^\alpha \|R\|_{L^2(T)}^2 + \|[u_h]\|_{L^2(\partial T)}^2,$$

where $R := f - (\mathbb{D}_x^\alpha u_h + \mathbb{D}_y^\alpha u_h)$. We will prove that the exact error can be bounded by η_T, where the constants in these inequalities depend only on the properties of the triangulation. The upper estimate shows that η_T can be used as a reliable stopping criterion for the adaptive algorithm. While the lower estimate cannot be easily obtained due to the nonlocal properties of the fractional operator [3].

Lemma 3.1. *There exists a constant C_1, depending only on the minimum angle of Ω_h, such that*

$$\|u - u_h\|_{E(\Omega_h)}^2 \leq C_1 \sum_{T \in \Omega_h} \eta_T^2, \tag{3.2}$$

where the energy norm $\| \cdot \|_{E(\Omega_h)}$ is defined in (2.26) but with $\lambda = 0$ and $\alpha = \beta$.

Proof. The weak form reads as follows: find $u \in V \subset H^{\alpha/2}(\Omega)$, such that

$$a(u, v) = (f, v), \qquad \forall v \in V, \tag{3.3}$$

where the symmetry bilinear form

$$a(u, v) = a_x^\alpha(u, v) + a_y^\alpha(u, v) + \sum_{e \in \Gamma} \int_e [u][v].$$

The notations $a_x^\alpha(u, v)$ and $a_y^\alpha(u, v)$ are defined in (2.17) with $\lambda = 0$.

Let $u_h \in V_h$ (defined in (2.12)) be the numerical solution, and $e_h := u - u_h$. Then e_h satisfies the residual equation

$$a(e_h, v) = (R, v) - \sum_{e \in \Gamma} \int_e [u_h][v], \qquad \forall v \in V. \tag{3.4}$$

Now we resort to the Scott-Zhang interpolation operator $\Pi : V \to V_h$ [38], which satisfies the approximation property, for any $v \in V$,

$$\|v - \Pi v\|_{L^2(T)} \leq C h_T^{\alpha/2} |v|_{H^{\alpha/2}(\tilde{T})}, \tag{3.5}$$

where $\tilde{T}$ is the union of all elements having nonempty intersection with T. Then we have

$$|a(e_h, v)| = |a(e_h, v - \Pi v)| \leq \sum_{T \in \Omega_h} |(R, v - \Pi v)_T| + \sum_{e \in \Gamma} \int_e \big|[u_h][v - \Pi v]\big|$$

$$\leq \sum_{T \in \Omega_h} \|R\|_{L^2(T)} \cdot \|v - \Pi v\|_{L^2(T)} + \sum_{e \in \Gamma} \|[u_h]\|_{L^2(e)} \cdot \|v\|_{L^2(e)}$$

$$\leq C \left(\sum_{T \in \Omega_h} \eta_T^2 \right)^{1/2} \cdot \left(\sum_{T \in \Omega_h} h_T^{-\alpha} \|v - \Pi v\|_{L^2(T)}^2 + \|v\|_{L^2(\partial T)}^2 \right)^{1/2}$$

$$\leq C \left(\sum_{T \in \Omega_h} \eta_T^2 \right)^{1/2} \cdot \left(\sum_{T \in \Omega_h} |v|_{H^{\alpha/2}(\tilde{T})}^2 + \|v\|_{L^2(\partial T)}^2 \right)^{1/2}$$

$$\leq C \left(\sum_{T \in \Omega_h} \eta_T^2 \right)^{1/2} \cdot \|v\|_{E(\Omega_h)}, \tag{3.6}$$

where we use the following inequality in the last step, from the definition on Aronszajn-Slobodeckij seminorm (2.11) [23]

$$\sum_{T \in \Omega_h} |v|^2_{H^{\alpha/2}(\tilde{T})} \leq C|v|^2_{H^{\alpha/2}(\Omega)}.$$

Taking $v = e_h \in H^{\alpha/2}(\Omega)$, we obtain

$$\|e_h\|_{E(\Omega_h)} \leq C \left(\sum_{T \in \Omega_h} \eta_T^2 \right)^{1/2}. \tag{3.7}$$

This completes the proof of the lemma. □

3.1.2. Scheme 2 – DWR method

Different from the scheme above, which aims at estimating the error with respect to the generic energy norm, the DWR method [5] for goal-oriented error estimation aims at economical computation of arbitrary quantities of physical interest. This is typically required in the design cycles of technical applications. 'Goal-oriented' adaptivity is designed to achieve these tasks with minimal cost.

When solving the fractional problems, the DWR method is significantly better than the traditional approach since the fractional operator and energy norm is nonlocal. In detail, for left Riemann-Liouville fractional operator, the numerical solution on one element T is affected by all the elements on the left. Therefore, the lower bound is not local absolutely and the over refinement may occur. While DWR multiplies every local error indicator on one element by a weight, consisting of the dual solution. It has the feature of a 'generalized' Green function $G(T, T')$, which describes the dependence of the target error quantity $J(e_h)$ concentrated at some element T [5].

Without assuming $\alpha = \beta$, let $z \in V$ be the solution of the associated dual problem

$$a(\varphi, z) = J(\varphi) \quad \forall \varphi \in V, \tag{3.8}$$

and $z_h \in V_h \subset V$ be discontinuous finite element approximation defined by

$$a(\varphi_h, z_h) = J(\varphi_h) \quad \forall \varphi_h \in V_h, \tag{3.9}$$

where the bilinear form $a(\cdot, \cdot)$ is defined as

$$a(u, v) = a_x^\alpha(u, v) + a_y^\beta(u, v) + \sum_{e \in \Gamma} \int_e [u][v].$$

Using this construction together with Galerkin orthogonality, we obtain

$$
\begin{aligned}
J(e_h) = a(e_h, z) &= a(e_h, z - z_h) \\
&= \sum_{T \in \Omega_h} (R, z - z_h) + \sum_{e \in \Gamma} \int_e [u_h][z - z_h] \\
&\leq \sum_{T \in \Omega_h} \Big(\|R\|_{L^2(T)} \cdot \|z - z_h\|_{L^2(T)} + \|[u_h]\|_{L^2(\partial T)} \|[z - z_h]\|_{L^2(\partial T)} \Big). \tag{3.10}
\end{aligned}
$$

Thus, we define the local error indicator

$$\eta_T := \|R\|_{L^2(T)} \cdot \|z - z_h\|_{L^2(T)} + \|[u_h]\|_{L^2(\partial T)}\|[z - z_h]\|_{L^2(\partial T)}. \tag{3.11}$$

Taking

$$J(\varphi) = \Big((_a D_x^{\alpha/2}\varphi, _x D_b^{\alpha/2}e_h) + (_x D_b^{\alpha/2}\varphi, _a D_x^{\alpha/2}e_h) + (_c D_y^{\beta/2}\varphi, _y D_d^{\beta/2}e_h)$$

$$+ (_y D_d^{\beta/2}\varphi, _c D_y^{\beta/2}e_h) + \sum_{e \in \Gamma}[\varphi][e_h] \Big) \cdot \|e_h\|_{E(\Omega_h)}^{-1}$$

such that $J(e_h) = \|e_h\|_{E(\Omega_h)}$, we have the global upper bound in energy norm,

$$\|e_h\|_{E(\Omega_h)} = J(e_h) \le \sum_{T \in \Omega_h} \eta_T. \tag{3.12}$$

If we take a rough estimate to $J(e_h)$ by $\|z - z_h\|_{L^2(T)} \le Ch_T^{\alpha/2}|z|_{H^{\alpha/2}(\tilde{T})}$ and $\|[z - z_h]\|_{L^2(\partial T)} \le C\|z\|_{L^2(\partial T)}$, then substituting them into (3.10), similarly to (3.6), we have, with a priori analysis in forms of bounds for z,

$$\|e_h\|_{E(\Omega_h)} \le C \Big(\sum_{T \in \Omega_h} h_T^\alpha \|R\|_{L^2(\Omega)}^2 + \|[u_h]\|_{L^2(\partial T)}^2 \Big)^{1/2}.$$

Here, we get a global posteriori error estimate based on energy norm that is consistent with Scheme 1. In this sense, a posteriori error estimate based on DWR is more meticulous. In order to evaluate the posteriori error representation (3.12), we need information about the discontinuous dual solution z. Since in practice, z is not explicitly known, but it can be obtained by solving the dual problem numerically. Here, we approximate z by a high-order method. We take V_h as the linear discontinuous finite element and solve the dual problem by using quadratic discontinuous finite element on the current mesh yielding an approximation $z_h^{(2)} \in V_h^{(2)}$ to z in (3.11), and z_h can be got by linear interpolation of $z_h^{(2)}$ [5]. This yields the approximate local error indicator

$$\eta_T \approx \|R\|_{L^2(T)} \cdot \|z_h^{(2)} - \Pi_h z_h^{(2)}\|_{L^2(T)} + \|[u_h]\|_{L^2(\partial T)}\|[z_h^{(2)} - \Pi_h z_h^{(2)}]\|_{L^2(\partial T)}. \tag{3.13}$$

Remark 3.1. Besides the above mentioned points that a posteriori error estimate based on DWR is better than a global posteriori error estimate based on energy norm, the former also has advantages when the fractional derivative operators are anisotropic, i.e., $\alpha \ne \beta$. For Scheme 1 of energy norm, we take $\alpha = \beta$ in (3.1) for simplicity. Actually, if $\alpha < \beta$, we must take the exponent of h_T in the indicator $\eta_T := h_T^{\alpha/2}\|R\|_{L^2(T)}$ to be $\alpha/2$ to guarantee the upper bound. But this bound looks too large for β. A posteriori error estimate based on DWR avoids this problem, and it is still effective for complex problems.

3.1.3. Marking Strategy

The goal of adaptive methods is the generation of a mesh which is adapted to the problem. An optimal mesh should be as coarse as possible while meeting the criterion, in order to save computational time and memory requirements. A global refinement of the mesh would lead to the best error reduction, but the amount of new unknowns might be much larger than needed to reduce the error below given tolerance. We use the so-called Döfler marking strategy

[15] to refine the mesh. More strategies can be found in [30, 35], like maximum strategy and equidistribution strategy.

Marking Strategy: Given a parameter $0 < \theta < 1$, construct a subset $\hat{\Omega}_h \subset \Omega_h$ such that

$$\sum_{T \in \hat{\Omega}_h} \eta_T^2 \geq \theta^2 \sum_{T \in \Omega_h} \eta_T^2. \tag{3.14}$$

Then we shall refine the mesh Ω_h by this marking strategy and the convergence will be obtained similarly to [12, Theorem 4.5].

3.1.4. Numerical experiments

In this section we will present some numerical experiments using the two schemes given above. We compare them with uniformly refinement and with each other.

Example 3.1. Consider the 1D fractional equation on the domain $\Omega = [0, 2]$ with $\alpha = 0.8$,

$$\begin{cases} {}_0D_x^\alpha u = f, & \Omega, \\ u = 0, & \mathbb{R} \backslash \Omega, \end{cases}$$

and its dual problem is

$$\begin{cases} {}_xD_2^\alpha u = f, & \Omega, \\ u = 0, & \mathbb{R} \backslash \Omega. \end{cases}$$

The source term f is chosen such that the exact solution is $u = \left(1 - (x-1)^2\right)^\gamma$, $\gamma = 0.7$, which has poor regularity near the boundary, and we use discontinuous piecewise linear function for approximation.
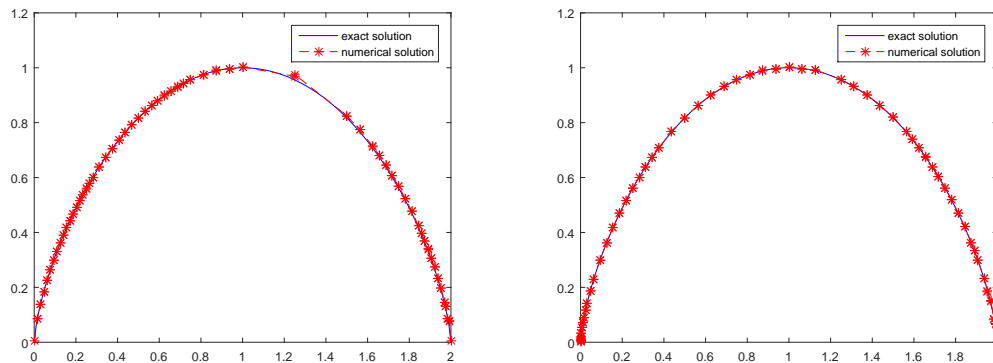

Fig. 3.1. Adaptive refinement of Example 3.1: energy norm indicator on 60 elements (left) and weighted indicator on 58 elements (right).

This is a boundary layer problem, i.e., the solution has less regularity around the endpoints of domain $[0, 2]$, where the mesh should be finer. We initially divide the interval $[0, 2]$ into 8 cells uniformly, then refine the mesh based on the energy norm indicator in Fig. 3.1 (left) and the weighted indicator in Fig. 3.1 (right). The numerical solution approximates the exact solution both very well obviously. And from an intuitive point of view, the right one is better than the left one. Besides, we compare the two kinds of refinements with uniform refinement, and show
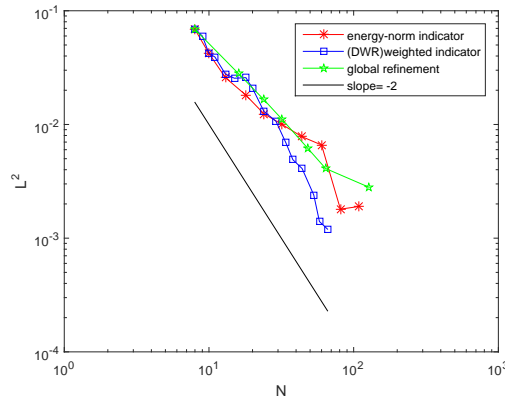
Fig. 3.2. L^2 error versus number of elements N of Example 3.1, for uniform refinement, the energy norm indicator, the weighted indicator obtained by the DWR approach.

their convergence rates. A reference line is provided which shows the optimal convergence rate N^{-2}. As shown in Fig. 3.2, the uniform refinement is the worst while the refinement based on the weighted indicator is the best, almost achieving the optimal convergence rate.

Example 3.2. Consider the 2D fractional equation on the domain by $\Omega := [0,2] \times [0,2]$ with $\alpha = 0.2$, $\beta = 0.8$,

$$\begin{cases} {}_0D_x^\alpha u + {}_xD_2^\alpha u + {}_0D_y^\beta u + {}_yD_2^\beta u = f, & \Omega, \\ u = 0, & \mathbb{R}^2 \backslash \Omega. \end{cases} \tag{3.15}$$

and its dual problem is itself. The source term f is chosen such that the exact solution writes

$$u(x,y) = x(x-2)y(y-2)\arctan\left(\frac{\sqrt{x^2+y^2}-2}{0.05}\right).$$

Being the same as the case of 1D, we compare the three kinds of refinements: uniform refinement, the refinement based on energy norm indicator, and weighted indicator. This exact solution has less regularity inside the domain, i.e., interior layer. Fig. 3.3 is the shape of the
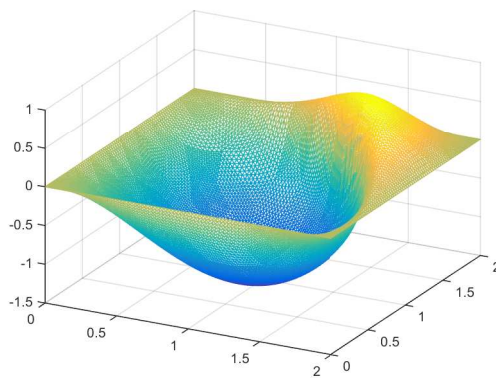

Fig. 3.3. Surface plot of the exact solution of Experiment 3.2, and the steep region around the arc $x^2 + y^2 = 4$.

Table 3.1: Example 3.2: Uniform refinement.

K	$\|u - u_k\|_{L^2}$	$\|u - u_k\|_E$
32	0.3164	1.1565
72	0.1908	0.9010
128	0.1257	0.6794
200	0.0935	0.5264
288	0.0755	0.4357
392	0.0637	0.3803
512	0.0528	0.3362

Table 3.2: Example 3.2: Refinement based on the energy norm indicator.

k	K	$\|u - u_k\|_{L^2}$	$\|u - u_k\|_E$	η	I_{eff}
1	8	0.6752	1.2608	3.5025	2.7781
2	15	0.3930	1.0819	2.0108	1.8586
3	28	0.2614	0.7705	1.6311	2.1168
4	42	0.1652	0.6573	1.1697	1.7796
5	86	0.1169	0.4383	0.9193	2.0973
6	152	0.0872	0.4074	0.6599	1.6198
7	201	0.0684	0.3023	0.5550	1.8361

Table 3.3: Example 3.2: Refinement based on the weighted indicator.

k	K	$\|u - u_k\|_{L^2}$	$\|u - u_k\|_E$	η	I_{eff}
1	8	0.6752	1.2608	3.2037	2.5410
2	15	0.3930	1.0819	2.2824	2.1096
3	28	0.2614	0.7705	1.4480	1.8792
4	36	0.1650	0.6408	1.0232	1.5968
5	63	0.1242	0.4852	0.9040	1.8632
6	110	0.0933	0.4174	0.6412	1.5360
7	166	0.0551	0.2766	0.4584	1.6572

exact solution. The adaptive mesh will be finer in the steep region. The experiment datum are presented in Table 3.1, Table 3.2, and Table 3.3. We denote the total error indicator by η, and the effectiveness index $I_{eff} = \eta/\|e\|_{E(\Omega_h)}$. If the effectiveness index I_{eff} remains roughly constant in different meshes, the indicator η approximates the true error e_h well. By comparing these tables, adaptive refinement is significantly better than uniform refinement. The weighted indicator is slightly better than the energy norm indicator since the L^2 norm error and energy norm error of the former is smaller than the latter, and the effectiveness index I_{eff} of the former changes more moderately than the latter. In Fig. 3.4, two kinds of adaptive meshes are presented, and both refine the steep region. The right one (based on the weighted indicator) is slightly better than the left one (based on the energy norm indicator). Although the advantage here is not obvious, maybe the problem is a little simple or the mesh is not fine enough, DWR method is still an important approach for complex problem and arbitrary quantities of physical interest.
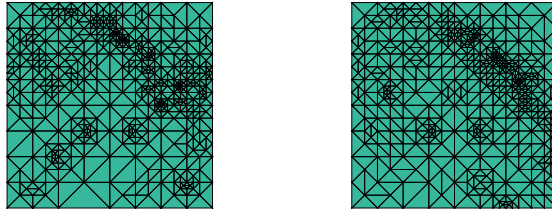
Fig. 3.4. Mesh after 11 iterations (973 cells) based on energy norm indicator on the left and mesh after 11 iterations (903 cells) based on weighted indicator on the right.

3.2. Evolution Equation

In this section, we consider the time dependent tempered fractional equation:

$$\begin{cases} \partial_t u + \mathbf{b} \cdot \nabla u - \kappa_1 \nabla_x^{\alpha,\lambda} u - \kappa_2 \nabla_y^{\beta,\lambda} u = f, & (\mathbf{x},t) \in \Omega \times J, \\ u(\mathbf{x},0) = u_0(\mathbf{x}), & \mathbf{x} \in \Omega, \\ u(\mathbf{x},t) = 0, & (\mathbf{x},t) \in \mathbb{R}^2 \backslash \Omega \times J. \end{cases} \tag{3.16}$$

Different from the stationary equation, the mesh here is adapted to the solution in every time step using a posteriori error indicator, i.e., the adaptive algorithm solving the evolution equation at the n-th time step reads as

$$\text{Solve} \rightarrow \text{Estimate} \rightarrow \text{Refine/Coarsen}.$$

Here the refinement/coarsening procedure includes both the mesh and time step size modifications. In this paper, we propose the following algorithm, similar to [10], to modify the time step size τ_n and mesh Ω_h^n starting from the initial time step size $\tau_{n,0} = \tau_{n-1}$ and initial mesh $\Omega_h^{n,0} = \Omega_h^{n-1}$:

1. Refine the time step size $\tau_{n,0}$ to the final time step size τ_n such that the associated time error indicators are less than the prescribed tolerances.

2. Refine/Coarsen the mesh $\Omega_h^{n,0}$ to the final mesh Ω_h^n such that the associated space error indicators are less than the prescribed tolerances.

3. Enlarge the initial time step size $\tau_{n+1,0}$ for next time step if the current time error indicator is much less than the tolerances.

3.2.1. A posteriori error analyses

While the previous section is concentrated on the spatial mesh refinement based on the local error indicator, this section describes the process of evolution of the solution and the time step size based on the time error indicator. Then we introduce the time and space local error indicator, and show their upper bound property. For convenience, we take $\mathbf{b} = \mathbf{0}$, $\alpha = \beta$. Let $u \in V := H^1(0,T; H^{\alpha/2}(\Omega))$ be the solution of

$$(\partial_t u, v) + a(u,v) - \kappa(u,v) = (f,v), \qquad \forall v \in V, \tag{3.17}$$

and u_h^n be the fully discrete discontinuous finite element approximation defined by

$$\left(\frac{u_h^n - u_h^{n-1}}{\tau_n}, v_h \right) + a(u_h^n, v_h) - \kappa(u_h^n, v_h) = (\bar{f}^n, v_h), \qquad \forall v_h \in V_h, \tag{3.18}$$

where $\bar{f}^n = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} f(\mathbf{x},t) dt$.

Theorem 3.1. *For any integer* $1 \leq m \leq N$, *there exists a positive constant* C *depending only on the minimum angle of meshes* Ω_h^n, $n = 1, \cdots, m$, *such that the posteriori error estimate*

$$\|u^m - u_h^m\|_{L^2(\Omega)}^2 + \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \|u^m - u_h^m\|_{E(\Omega_h^n)}^2 dt$$

$$\leq \|u_0 - u_h^0\|_{L^2(\Omega)}^2 + C \sum_{n=1}^{m} \tau_n (\eta_{\text{time1}}^n + \eta_{\text{time2}}^n) + C \sum_{n=1}^{m} \tau_n \eta_{\text{space}}^n,$$

holds, where the time error indicator and space error indicator are given by

$$\eta_{\text{time1}}^n = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \|f - \bar{f}^n\|_{L^2(\Omega)}^2 dt, \qquad \eta_{\text{time2}}^n = \|u_h^n - u_h^{n-1}\|_{L^2(\Omega)}^2, \qquad (3.19a)$$

$$\eta_{\text{space}}^n = \sum_{T \in \Omega_h^n} \eta_T^n, \qquad \eta_T^n = h_T^\alpha \|R^n\|_{L^2(T)}^2 + \|[u_h]\|_{L^2(\partial T)}^2, \qquad (3.19b)$$

$$R^n = \bar{f}^n - \frac{u_h^n - u_h^{n-1}}{\tau_n} - \kappa_1 \kappa_\alpha \mathbb{D}_x^{\alpha,\lambda} u_h^n - \kappa_2 \kappa_\alpha \mathbb{D}_y^{\alpha,\lambda} u_h^n + \kappa u_h^n. \qquad (3.19c)$$

Proof. The proof is similar to Lemma 3.1, but it is needed to deal with the temporal derivative carefully. From the definition of R^n, we have, for any $v \in V$,

$$\left(\frac{u_h^n - u_h^{n-1}}{\tau_n}, v \right) + a(u_h^n, v) - \kappa(u_h^n, v) = (\bar{f}^n, v) - (R^n, v) + \sum_{e \in \Gamma} [u_h^n][v]. \qquad (3.20)$$

Then for any $t \in (t_{n-1}, t_n]$, we denote $u_h(t)$ as the linear interpolation of u_h^{n-1} and u_h^n such that $\partial_t u_h(t) = (u_h^n - u_h^{n-1})/\tau_n$. Then from (3.17) and (3.20), for any $v \in V$ and its Scott-Zhang interpolator $\Pi^n v$, we have

$$(\partial_t (u - u_h), v) + a(u - u_h^n, v) - \kappa(u - u_h^n, v)$$

$$= (f - \bar{f}^n, v) + (R^n, v) - \sum_{e \in \Gamma} [u_h^n][v]$$

$$= (f - \bar{f}^n, v) + (R^n, v - \Pi^n v) - \sum_{e \in \Gamma} \int_e [u_h^n][v - \Pi v],$$

where the last equality uses the Galerkin orthogonality.

Taking $v = u - u_h := e_h$, similarly to (3.6), using the property of Scott-Zhang interpolation operator (3.5) and the formulae

$$a(u - u_h^n, u - u_h) = \frac{1}{2} \|u - u_h^n\|_{E(\Omega_h)}^2 + \frac{1}{2} \|u - u_h\|_{E(\Omega_h)}^2 - \frac{1}{2} \|u_h - u_h^n\|_{E(\Omega_h)}^2,$$

$$(u - u_h^n, u - u_h) = \frac{1}{2} \|u - u_h^n\|_{L^2(\Omega)}^2 + \frac{1}{2} \|u - u_h\|_{L^2(\Omega)}^2 - \frac{1}{2} \|u_h - u_h^n\|_{L^2(\Omega)}^2$$

$$\leq \|u - u_h\|_{L^2(\Omega)}^2,$$

we have

$$\frac{1}{2}\frac{d}{dt}\|e_h\|^2_{L^2(\Omega)} + \frac{1}{2}\|e_h\|^2_{E(\Omega_h)} + \frac{1}{2}\|u - u_h^n\|^2_{E(\Omega_h)}$$

$$\leq \kappa\|e_h\|^2_{L^2(\Omega)} + \frac{1}{2}\|u_h - u_h^n\|^2_{E(\Omega_h)} + (f - \bar{f}^n, e_h) + (R^n, e_h - \Pi^n e_h) + \sum_{e \in \Gamma}\int_e |[u_h][e_h - \Pi e_h]|$$

$$\leq C(\|e_h\|^2_{L^2(\Omega)} + \|f - \bar{f}^n\|^2_{L^2(\Omega)} + \|u_h^n - u_h^{n-1}\|^2_{L^2(\Omega)}) + C\sum_{T \in \Omega_h^n}\left(h_T^\alpha\|R^n\|^2_{L^2(\Omega)} + \|[u_h]\|^2_{L^2(\partial T)}\right)$$

$$+ \frac{1}{4}\|e^n\|^2_{E(\Omega_h)}, \tag{3.21}$$

where we use the following inequality in the last step, since u_h is the linear interpolation of u_h^{n-1} and u_h^n in $(t_{n-1}, t_n]$,

$$\|u_h - u_h^n\|^2_{E(\Omega_h)} \leq C\|u_h^n - u_h^{n-1}\|^2_{E(\Omega_h)}.$$

Integrating (3.21) in time from t_{n-1} to t_n and summing over n from 1 to m, with the discrete Grönwall inequality, we complete the proof. $\square$

3.2.2. Numerical experiment

In this part, we present the numerical results using the adaptive algorithm of evolution equations given above and focus on observing the evolution of exact solution by taking it with different regularity at different time, which will indicate the effectiveness of the indicator.

Example 3.3. Consider the equation (3.16) on the domain $\Omega = [0, 2] \times [0, 2]$, with $\mathbf{b} = \mathbf{0}, \kappa_1 = 0.1, \kappa_2 = 0.2, \lambda = 0.2$, and $\alpha = \beta = 0.8$. The source term f is chosen such that the exact solution is

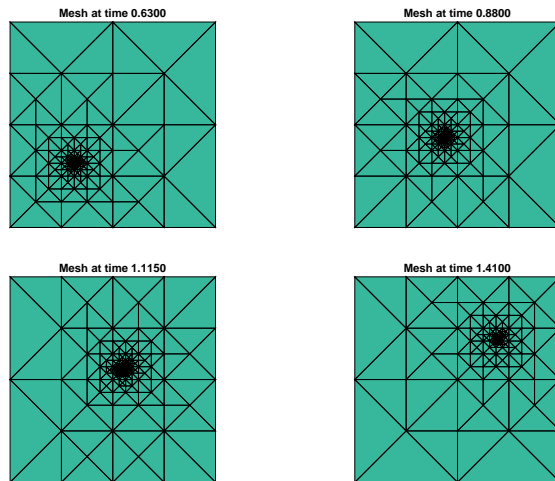$$u(x, y) = x(x - 2)y(y - 2)e^{-((x-t)^2 + (y-t)^2)/0.005}.$$

Fig. 3.5. Mesh refinement at different time. It shows that the local refinement of meshes change as the singularity point moves on the line $y = x$.

Due to the small denominator 0.005 in this example, the solution is steep at the point (t, t), where we expect to observe more refinements. The adaptation process yields the meshes shown in Fig. 3.5, which is consistent with our expectation completely.

4. Conclusion

We discuss the DG methods for the two dimensional time-dependent space tempered fractional convection-diffusion equations, especially their adaptivity; and the provided strategy for designing adaptive schemes works for the general PDEs with fractional operators. DG methods are superior to finite element method in many ways, but the stability of the discrete schemes needs more attention. The interior boundaries are usually connected by two ways: interior penalty and numerical flux. We use the former to deal with the diffusion term and the latter to the convection term. The stability and convergence analyses are explicitly provided. The theoretical results are confirmed by numerical experiments. For the adaptivity of the DG methods, we consider two schemes of the stationary problem, i.e., a posteriori error estimator based on the traditional energy norm and another one based on the dual weighted residual. The numerical experiments confirm the advantage of the DWR method. Finally, we consider the fractional evolution problem. A posteriori error estimate is provided and the indicator is designed; and its effectiveness is displayed by numerical experiments.

References

[1] G. Acosta, F. Bersetche, J.P. Borthagaray, A short FE implementation for a 2d homogeneous Dirichlet problem of a fractional Laplacian, *Comput. Math. Appl.*, **74** (2017), 784–816.

[2] G. Acosta, J.P. Borthagaray, A Fractional Laplace Equation: Regularity of Solutions and Finite Element Approximations, *SIAM J. Numer. Anal.*, **55** (2017), 472–495.

[3] M. Ainsworth, C. Glusa, Aspects of an adaptive finite element method for the fractional Laplacian: A priori and a posteriori error estimates, efficient implementation and multigrid solver, *Comput. Methods Appl. Mech. Engrg.*, **327** (2017), 4–35.

[4] B. Baenmera, M.M. Meerschaert, Tempered stable Lévy motion and transient super-diffusion, *J. Comput. Appl. Math.*, **233** (2010), 2438–2448.

[5] W. Bangerth, R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Birkhäuser-Verlag, Basel, 2003.

[6] S.C. Brenner, L.R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer Verlag, New York, 1994.

[7] E. Burman, A posteriori error estimation for interior penalty finite element approximations of the advection-reaction equation, *SIAM J. Numer. Anal.*, **47** (2009), 3584–3607.

[8] P.L. Butzer, U. Westphal, *An Introduction to Fractional Calculus*, World Scientific, Singapore, 2000.

[9] M.H. Chen, W.H. Deng, High order algorithms for the fractional substantial diffusion equation with truncated Lévy flights, *SIAM J. Sci. Comput.*, **19** (2014), 1431–1458.

[10] Z.M. Chen, J. Feng, An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems, *Math. Comp.*, **73** (2004), 1167–1193.

[11] L. Chen, R.H. Nochetto, E. Otárola, A.J. Salgado, A PDE approach to fractional diffusion: a posteriori error analysis, *J. Comput. Phys.*, **293** (2015), 339–358.

[12] Z.M. Chen, H.J. Wu, *Selected Topics in Finite Element Methods*, Science Press, Beijing, 2010.

[13] B. Cockburn, C.-W. Shu, The local discontinuous Galerkin finite element method for the time dependent convection-diffusion systems, *SIAM J. Numer. Anal.*, **35** (1998), 2440–2463.

[14] B. Cockburn, K. Mustapha, A hybridizable discontinuous Galerkin method for fractional diffusion problems, *Numer. Math.*, **130** (2015), 293–314.

[15] W. Dörfler, A convergent adaptive algorithm for Poissons equation, *SIAM J. Numer. Anal.*, **33** (1996), 1106–1124.

[16] W.H. Deng, J.S. Hesthaven, Local discontinuous Galerkin methods for fractional diffusion equations, *ESAIM: Math. Model. Numer. Anal.*, **47** (2013), 1845–1864.

[17] W.H. Deng, B.Y. Li, W.Y. Tian, P.W. Zhang, Boundary problems for the fractional and tempered fractional operators, *Multiscale Model. Simul.*, **16** (2018), 125–149.

[18] W.H. Deng, Z.J. Zhang, *High Accuracy Algorithm for the Differential Equations Governing Anomalous Diffusion*, World Scientific, Singapore, 2019.

[19] W.H. Deng, Z.J. Zhang, Variational formulation and efficient implementation for solving the tempered fractional problems, *Numer. Methods Partial Differential Equations*, **34** (2018), 1224–1257.

[20] Q. Du, L. Tian, X. Zhao, A convergent adaptive finite element algorithm for nonlocal diffusion and peridynamic models, *SIAM J. Numer. Anal.*, **51** (2013), 1211–1234.

[21] B. Dybiec, E. Gudowska-Nowak, P. Hänggi, Lévy-Brownian motion on finite intervals: Mean first passage time analysis, *Phys. Rev. E*, **73** (2006), 046104.

[22] V.J. Ervin, J.E. Roop, Variational formulation for the stationary fractional advection dispersion equation, *Numer. Methods Partial Differential Equations*, **22** (2005), 558–576.

[23] B. Faermann, Localization of the Aronszajn-Slobodeckij norm and application to adaptive boundary element methods. Part I. The two-dimensional case, *IMA J. Numer. Anal.*, **20** (2000), 203–234.

[24] J.S. Hesthaven, T. Warburton, *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*, Springer Verlag, New York, 2008.

[25] C. Li, W.H. Deng, High order schemes for the tempered fractional diffusion equations, *Adv. Comput. Math.*, **42** (2016), 543–572.

[26] X.J. Li, C.J. Xu, A space-time spectral method for the time fractional diffusion equation, *SIAM J. Numer. Anal.*, **47** (2009), 2108–2131.

[27] R. Metzler, E. Barkai, J. Klafter, Anomalous diffusion and relaxation close to thermal equilibrium: a fractional Fokker-Planck equation approach, *Phys. Rev. Lett.*, **82** (1999), 3563–3567.

[28] R. Metzler, J. Klafter, The random walk's guide to anomalous diffusion: a fractional dynamics approach, *Phys. Rep.*, **339** (2000), 1–77.

[29] W. McLean, K. Mustapha, Superconvergence of a discontinuous Galerkin method for fractional diffusion and wave equations, *SIAM J. Numer. Anal.*, **51** (2013), 491–515.

[30] P. Morin, R.H. Nochetto, K.G. Siebert, Convergence of adaptive finite element methods, *SIAM Rev.*, **44** (2002), 631–658.

[31] R.H. Nochetto, E. Otárola, A.J. Salgado, A PDE approach to fractional diffusion in general domains: a priori error analysis, *Found. Comput. Math.*, **15** (2015), 733–791.

[32] E.D. Nezza, G. Palatucci, E. Valdinoci, Hitchhiker's guide to the fractional Sobolev spaces, *Bull. Sci. Math.*, **136** (2012), 521–573.

[33] L.L. Qiu, W.H.Deng, J.S. Hesthaven, Nodal discontinuous Galerkin methods for fractional diffusion equations on 2D domain with triangular meshes, *J. Comput. Phys.*, **298** (2015), 678–694.

[34] B. Rivière, *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implemantation*, SIAM Philadelphia, 2008.

[35] A. Schmidt, K.G. Siebert, *Design of Adaptive Finite Element Software*, Springer-Verlag, Berlin Heidelberg, 2005.

[36] K. Shahbazi, An explicit expression for the penalty parameter of the interior penalty method, *J. Comput. Phys.*, **205** (2005), 401–407.

[37] S.A. Sauter, C. Schwab, *Boundary element methods*, Springer-Verlag, Berlin, 2011.

[38] L.R. Scott, S.Y. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, *Math. Comp.*, **54** (1990), 483–493.

[39] K.X. Wang, H. Wang, M. Al-Lawatia, H.X. Rui, A family of characteristic discontinuous Galerkin methods for transient advection-diffusion equations and their optimal-order L2 error estimates, *Commun. Comput. Phys.*, **6** (2009), 203–230.

[40] S.Q. Wang, J.Y. Yuan, W.H. Deng, Y.J. Wu, A hybridized discontinuous Galerkin method for 2D fractional convection-diffusion equations, *J. Sci. Comput.*, **68** (2016), 826–847.

[41] X.C. Wu, W.H. Deng, E. Barkai, Tempered fractional Feynman-Kac equation: Theory and examples, *Phys. Rev. E*, **93** (2016), 032151.

[42] Q. Xu, J.S. Hesthaven, Discontinuous Galerkin method for fractional convection-diffusion equations, *SIAM J. Numer. Anal.*, **52** (2014), 405–423.

[43] J. Yan, C.-W.Shu, A local discontinuous Galerkin method for KdV type equations, *SIAM J. Numer. Anal.*, **40** (2002), 769–791.

[44] Z.J. Zhang, W.H. Deng, G.E. Karniadakis, A Riesz basis Galerkin method for the tempered fractional Laplacian, *SIAM J. Numer. Anal.*, **56** (2018), 3010–3039.

[45] X. Zhao, X.Z. Hu, W. Cai, G.E. Karniadakis, Adaptive finite element method for fractional differential equations using hierarchical matrices, *Comput. Methods Appl. Mech. Engrg.*, **325** (2017), 56–76.

[46] M. Zayernouri, G. E. Karniadakis, Discontinuous spectral element methods for time-and-space-fractional advection equations, *SIAM J. Sci. Comput.*, **36** (2014), B684–B707.