

A PDE Approach to Numerical Fractional Diffusion*

Ricardo H. Nochetto, Enrique Otárola, Abner J. Salgado

Abstract. Fractional diffusion has become a fundamental tool for the modeling of multiscale and heterogeneous phenomena. However, due to its nonlocal nature, its accurate numerical approximation is delicate. We survey our research program on the design and analysis of efficient solution techniques for problems involving fractional powers of elliptic operators. Starting from a localization PDE result for these operators, we develop local techniques for their solution: a priori and a posteriori error analyses, adaptivity and multilevel methods. We show the flexibility of our approach by proposing and analyzing local solution techniques for a space-time fractional parabolic equation.

Mathematics Subject Classification (2000). Primary 26A33, 65N12, 65N30, 65N50, 65N55, 65F10, 65M15, 65M60; Secondary 35J70, 65R10. 65J08,

Keywords. Fractional diffusion, fractional derivatives and integrals, nonlocal operators, Muckenhoupt weights, finite elements, anisotropic elements, a posteriori error estimates, adaptive algorithm, multilevel methods, stability, fully-discrete methods.

1. Introduction

Diffusion is the tendency of a substance to evenly spread into surrounding space, and is one of the most common physical processes. The classical models of diffusion lead to local and thoroughly studied equations. However, in recent times, it has become evident that many of the assumptions that lead to these models are not always satisfactory or not even realistic in practice. Consequently, different models of diffusion have been proposed, fractional diffusion being one of them.

Capturing the essential behavior of fractional diffusion with the simplest and crudest models is of paramount importance in science and engineering. This allows for understanding of physical applications where long range or anomalous diffusion is considered [1], complex phenomena in mechanics [5], biophysics [22], turbulence [31], image processing [39], nonlocal electrostatics [46], finance [50] and the control of devices [2, 3]. To understand the behavior of fractional diffusion, computational science is fundamental. It is one of the pillars, together with theory and experiments, of scientific inquiry. A carefully crafted computational model

*RHN has been supported in part by NSF grants DMS-1109325 and DMS-1411808. EO has been supported in part by NSF grants DMS-1109325 and DMS-1411808 and by CONICYT through project Anillo ACT1106. AJS has been supported in part by NSF grant DMS-1418784.

can replace a very expensive or unrealizable experimental setting, and it can give new insight into the theoretical developments of a specific discipline. The analysis of such computational schemes is the realm of numerical analysis, which offers a rigorous mathematical description of the extent to which the computer's output approximates the process of interest. These crucial aspects of modern research are blended together in this paper, which deals with the design and analysis of efficient numerical techniques for problems involving fractional diffusion.

The mathematical structure of fractional diffusion is shared by a wide class of *nonlocal operators* [34, 54]. These operators have a strong connection with real-world problems, since they constitute a fundamental part of the modeling and simulation of complex phenomena. It is evident that the particular type of mathematical operator appearing in applications can widely vary and that a unified analysis might be well beyond reach. A more modest, but nevertheless quite ambitious, goal is to develop computational tools and their analysis for a representative of a particular class. We discuss below our contributions to fractional diffusion.

Exploiting the breakthrough by L. Caffarelli and L. Silvestre [25], we have made a decisive advance in numerical fractional diffusion. Although the analysis of our method is intricate, *its implementation is done using standard components of finite element analysis* [30, 29, 61, 60, 62]. This is the main advantage of our scheme, since alternative approaches require less traditional techniques (such as special quadrature) to cope with the mathematical difficulties inherent to fractional diffusion. Even in 1D [45, 66], the integral formulation of fractional diffusion is notoriously difficult from the numerical standpoint due to the presence of a nonintegrable kernel; see [67] for a discussion. In contrast, our PDE approach to fractional diffusion, as well as the recent method of [12], can handle multidimensions easily and efficiently – a highly desirable feature.

Here we are interested in the design and analysis of numerical techniques to solve problems involving fractional powers of the Dirichlet Laplace operator $(-\Delta)^s$, $s \in (0, 1)$, the so-called fractional Laplacian. Let Ω be a bounded Lipschitz domain of \mathbb{R}^n ($n \geq 1$), with boundary $\partial\Omega$. Given a function f , we seek u such that

$$(-\Delta)^s u = f \quad \text{in } \Omega. \quad (1.1)$$

Our approach, however, is not particular to the fractional Laplacian and it can be applied to any second order, symmetric and uniformly elliptic operator [60, §7].

One of the main difficulties to study (1.1) is that the fractional Laplacian is a nonlocal operator [24, 26, 23, 25, 49]. To localize it, Caffarelli and Silvestre showed in [25] that any power $s < 1$ of the fractional Laplacian in \mathbb{R}^n can be realized as an operator that maps a Dirichlet boundary condition to a Neumann-type condition via an extension problem on the upper half-space \mathbb{R}_+^{n+1} . This result was adapted in [18, 26, 72] to bounded domains Ω , thus obtaining an extension problem posed on the semi-infinite cylinder $\mathcal{C} := \Omega \times (0, \infty)$. This extension is the mixed boundary value problem:

$$\operatorname{div}(y^\alpha \nabla \mathcal{U}) = 0 \text{ in } \mathcal{C}, \quad \mathcal{U} = 0 \text{ on } \partial_L \mathcal{C}, \quad \frac{\partial \mathcal{U}}{\partial \nu^\alpha} = d_s f \text{ on } \Omega \times \{0\}, \quad (1.2)$$

where $\partial_L \mathcal{C} := \partial\Omega \times (0, \infty)$ is the lateral boundary of \mathcal{C} , and $d_s := 2^{1-2s} \frac{\Gamma(1-s)}{\Gamma(s)}$ is a positive normalization constant that depends only on s ; see [23, 25] for details. The parameter α is defined as $\alpha = 1 - 2s \in (-1, 1)$, and the so-called conormal exterior derivative of \mathcal{U} at $\Omega \times \{0\}$ is

$$\frac{\partial \mathcal{U}}{\partial \nu^\alpha} = - \lim_{y \rightarrow 0^+} y^\alpha \mathcal{U}_y. \quad (1.3)$$

We call y the *extended variable* and the dimension $n + 1$ in \mathbb{R}_+^{n+1} the *extended dimension* of problem (1.2). The limit in (1.3) must be understood in the distributional sense; see [23, 25, 26]. As noted in [18, 25, 26, 72], the fractional Laplacian and the Dirichlet-to-Neumann operator of problem (1.2) are related by

$$d_s (-\Delta)^s u = \frac{\partial \mathcal{U}}{\partial \nu^\alpha} \quad \text{in } \Omega. \quad (1.4)$$

We propose, and analyze in Section 2, the following strategy to solve (1.1): given f we solve (1.2), thus obtaining a function \mathcal{U} ; setting $u : x' \in \Omega \mapsto u(x') = \mathcal{U}(x', 0) \in \mathbb{R}$, we obtain the solution of (1.1). The main advantage of this algorithm is that we are solving the local problem (1.2) instead of dealing with the nonlocal operator $(-\Delta)^s$ of (1.1). However, this comes at the expense of incorporating one more dimension to the problem, thus raising the question of computational efficiency. This motivates the use of anisotropic meshes to compensate for the singular behavior of $\mathcal{U}(\cdot, y)$ as $y \rightarrow 0$, and well as the development of a posteriori error estimators and multilevel methods. These are reviewed in Sections 3 and 4, respectively. To show the flexibility of our approach, in Section 5 we consider a parabolic equation with fractional diffusion and fractional time derivative.

Throughout this work Ω is an open, bounded and connected domain of \mathbb{R}^n , $n \geq 1$, with polyhedral boundary $\partial\Omega$. Given the semi-infinite cylinder $\mathcal{C} = \Omega \times (0, \infty)$ and $\mathcal{Y} > 0$, the truncated cylinder with base Ω and height \mathcal{Y} is defined by $\mathcal{C}_\mathcal{Y} := \Omega \times (0, \mathcal{Y})$ with lateral boundary $\partial_L \mathcal{C}_\mathcal{Y} := \partial\Omega \times (0, \mathcal{Y})$.

We will be dealing with objects defined in \mathbb{R}^{n+1} and it will be convenient to distinguish the extended dimension. A vector $x \in \mathbb{R}^{n+1}$, will be denoted by

$$x = (x_1, \dots, x_n, x_{n+1}) = (x', x_{n+1}) = (x', y),$$

with $x_i \in \mathbb{R}$ for $i = 1, \dots, n + 1$, $x' \in \mathbb{R}^n$ and $y \in \mathbb{R}$.

By $a \lesssim b$ we mean $a \leq Cb$, with a constant C that neither depends on a, b or the discretization parameters. The value of C might change at each occurrence.

2. A PDE approach: formulation and FEM

For a bounded domain there are several ways, not necessarily equivalent, to define the fractional Laplacian; see [11, 60] for a discussion. As in [60] we adopt the definition based on spectral theory [10]. The operator $(-\Delta)^{-1} : L^2(\Omega) \rightarrow L^2(\Omega)$ is compact, symmetric and positive, whence its spectrum $\{\lambda_k^{-1}\}_{k \in \mathbb{N}}$ is discrete, real,

positive and accumulates at zero. Moreover, there exists $\{\varphi_k\}_{k \in \mathbb{N}}$, which is an orthonormal basis of $L^2(\Omega)$ and satisfies $-\Delta \varphi_k = \lambda_k \varphi_k$ in Ω and $\varphi_k = 0$ on $\partial\Omega$. Fractional powers of the Dirichlet Laplacian can be defined for $w \in C_0^\infty(\Omega)$ by

$$(-\Delta)^s w = \sum_{k=1}^{\infty} \lambda_k^s w_k \varphi_k, \quad (2.1)$$

where $w_k = \int_{\Omega} w \varphi_k$. By density $(-\Delta)^s$ can be extended to the space

$$\mathbb{H}^s(\Omega) = \left\{ w = \sum_{k=1}^{\infty} w_k \varphi_k : \sum_{k=1}^{\infty} \lambda_k^s w_k^2 < \infty \right\} \quad (2.2)$$

For $s \in (0, 1)$ we denote by $\mathbb{H}^{-s}(\Omega)$ the dual of $\mathbb{H}^s(\Omega)$.

2.1. The Caffarelli-Silvestre extension problem. The Caffarelli-Silvestre extension leads to a *nonuniformly* elliptic equation [25, 18, 24, 26]. We consider weighted Sobolev spaces with the weight $|y|^\alpha$, $\alpha \in (-1, 1)$. If $D \subset \mathbb{R}^{n+1}$, $L^2(|y|^\alpha, D)$ is the space of measurable functions on D such that

$$\|w\|_{L^2(|y|^\alpha, D)}^2 = \int_D |y|^\alpha w^2 < \infty.$$

Similarly we define the weighted Sobolev space

$$H^1(|y|^\alpha, D) = \{w \in L^2(|y|^\alpha, D) : |\nabla w| \in L^2(|y|^\alpha, D)\},$$

where ∇w is the distributional gradient of w . We equip $H^1(|y|^\alpha, D)$ with the norm

$$\|w\|_{H^1(|y|^\alpha, D)} = \left(\|w\|_{L^2(|y|^\alpha, D)}^2 + \|\nabla w\|_{L^2(|y|^\alpha, D)}^2 \right)^{1/2}. \quad (2.3)$$

Since $\alpha \in (-1, 1)$ then $|y|^\alpha$ belongs to the Muckenhoupt class $A_2(\mathbb{R}^{n+1})$ [35, 38, 40, 57, 73]. This, in particular, implies that $H^1(|y|^\alpha, D)$ with norm (2.3) is Hilbert and $C^\infty(D) \cap H^1(|y|^\alpha, D)$ is dense in $H^1(|y|^\alpha, D)$ (cf. [73, Proposition 2.1.2, Corollary 2.1.6], [48] and [40, Theorem 1]). We recall the definition of Muckenhoupt classes.

Definition 2.1 (Muckenhoupt class A_2 , [57, 73]). Let ω be a weight and $N \geq 1$. We say that $\omega \in A_2(\mathbb{R}^N)$ if

$$C_{2,\omega} = \sup_B \left(\int_B \omega \right) \left(\int_B \omega^{-1} \right) < \infty, \quad (2.4)$$

where the supremum is taken over all balls B in \mathbb{R}^N .

If $\omega \in A_2(\mathbb{R}^N)$, we call it an A_2 -weight, $C_{2,\omega}$ in (2.4) is the A_2 -constant of ω .

To study problem (1.2) we define the weighted Sobolev space

$$\mathring{H}_L^1(y^\alpha, \mathcal{C}) = \{w \in H^1(y^\alpha, \mathcal{C}) : w = 0 \text{ on } \partial_L \mathcal{C}\}. \quad (2.5)$$

As [60, (2.21)] shows, the following *weighted Poincaré inequality* holds:

$$\|w\|_{L^2(y^\alpha, \mathcal{C})} \lesssim \|\nabla w\|_{L^2(y^\alpha, \mathcal{C})}, \quad \forall w \in \dot{H}_L^1(y^\alpha, \mathcal{C}), \quad (2.6)$$

then the seminorm on $\dot{H}_L^1(y^\alpha, \mathcal{C})$ is equivalent to (2.3). For $w \in H^1(y^\alpha, \mathcal{C})$, $\text{tr}_\Omega w$ denotes its trace onto $\Omega \times \{0\}$ which satisfies [60, Proposition 2.5]

$$\text{tr}_\Omega \dot{H}_L^1(y^\alpha, \mathcal{C}) = \mathbb{H}^s(\Omega), \quad \|\text{tr}_\Omega w\|_{\mathbb{H}^s(\Omega)} \leq C_{\text{tr}_\Omega} \|w\|_{\dot{H}_L^1(y^\alpha, \mathcal{C})}. \quad (2.7)$$

Let us now describe the Caffarelli-Silvestre extension result for bounded domains; [25, 72]. Given $f \in \mathbb{H}^{-s}(\Omega)$, let $u \in \mathbb{H}^s(\Omega)$ solve (1.1). If $\mathcal{U} \in \dot{H}_L^1(y^\alpha, \mathcal{C})$ solves (1.2), then $u = \mathcal{U}(\cdot, 0)$ and (1.4) holds.

2.2. A priori error analysis. We now review the main results of [60] about the a priori error analysis of discretizations of problem (1.1). This will also serve to make clear the limitations of this theory, thereby justifying the quest for an a posteriori error analysis. In this section we assume that

$$\|w\|_{H^2(\Omega)} \lesssim \|\Delta_{x'} w\|_{L^2(\Omega)}, \quad \forall w \in H^2(\Omega) \cap H_0^1(\Omega). \quad (2.8)$$

This holds if, for instance, the domain Ω is convex [42].

Since \mathcal{C} is unbounded, problem (1.2) cannot be approximated with finite-element-like techniques. However, since the solution of problem (1.2) decays exponentially in y [60, Proposition 3.1], by truncating \mathcal{C} to $\mathcal{C}_\mathcal{Y}$ and setting a homogeneous Dirichlet condition on $y = \mathcal{Y}$, we only incur in an exponentially small error in terms of \mathcal{Y} [60, Theorem 3.5]. If

$$\dot{H}_L^1(y^\alpha, \mathcal{C}_\mathcal{Y}) := \{v \in H^1(y^\alpha, \mathcal{C}_\mathcal{Y}) : v = 0 \text{ on } \partial_L \mathcal{C}_\mathcal{Y} \cup \Omega \times \{\mathcal{Y}\}\},$$

then the aforementioned problem reads: find $v \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\mathcal{Y})$ such that

$$\int_{\mathcal{C}_\mathcal{Y}} y^\alpha \nabla v \nabla \phi = d_s \langle f, \text{tr}_\Omega \phi \rangle, \quad \forall v \in \dot{H}_L^1(y^\alpha, \mathcal{C}_\mathcal{Y}). \quad (2.9)$$

Here $\langle \cdot, \cdot \rangle$ is the duality pairing between $\mathbb{H}^{-s}(\Omega)$ and $\mathbb{H}^s(\Omega)$; see (2.7).

If \mathcal{U} and v denote the solution of (1.2) and (2.9), respectively, then [60, Theorem 3.5] provides the following exponential estimate

$$\|\nabla(\mathcal{U} - v)\|_{L^2(y^\alpha, \mathcal{C})} \lesssim e^{-\sqrt{\lambda_1} \mathcal{Y}/4} \|f\|_{\mathbb{H}^{-s}(\Omega)}.$$

To study the finite element discretization of (2.9) we must understand the regularity of \mathcal{U} and v . As [60, Theorem 2.7] reveals, the second order regularity of \mathcal{U} is much worse in the extended direction, namely

$$\|\Delta_{x'} \mathcal{U}\|_{L^2(y^\alpha, \mathcal{C})} + \|\partial_y \nabla_{x'} \mathcal{U}\|_{L^2(y^\alpha, \mathcal{C})} \lesssim \|f\|_{\mathbb{H}^{1-s}(\Omega)}, \quad (2.10)$$

$$\|\mathcal{U}_{yy}\|_{L^2(y^\beta, \mathcal{C})} \lesssim \|f\|_{L^2(\Omega)}, \quad (2.11)$$

with $\beta > 2\alpha + 1$. Thus, *graded* meshes in the extended variable y play a fundamental role. Estimates (2.10)–(2.11) motivate the construction of a mesh over \mathcal{C}_γ as follows. We consider a partition \mathcal{I}_γ of the interval $[0, \gamma]$ with points

$$y_k = k^\gamma M^{-\gamma} \gamma, \quad k = 0, \dots, M, \quad (2.12)$$

where $\gamma > 3/(1 - \alpha) = 3/(2s)$. Let $\mathcal{T}_\Omega = \{K\}$ be a conforming and shape regular mesh of Ω , where $K \subset \mathbb{R}^n$ is an element that is isoparametrically equivalent either to the unit cube $[0, 1]^n$ or the unit simplex in \mathbb{R}^n . The collection of these triangulations is denoted by \mathbb{T}_Ω . We construct the mesh \mathcal{T}_γ of \mathcal{C}_γ as the tensor product of \mathcal{T}_Ω and \mathcal{I}_γ . By construction \mathcal{T}_γ is anisotropic in the extended variable (cf. [36, 60]). The set of all triangulations of \mathcal{C}_γ obtained with this procedure is \mathbb{T} .

For $\mathcal{T}_\gamma \in \mathbb{T}$, we define the finite element space

$$\mathbb{V}(\mathcal{T}_\gamma) = \{W \in C^0(\bar{\mathcal{C}}_\gamma) : W|_T \in \mathcal{P}_1(K) \otimes \mathbb{P}_1(I) \ \forall T \in \mathcal{T}_\gamma, W|_{\Gamma_D} = 0\}, \quad (2.13)$$

where $\Gamma_D = \partial_L \mathcal{C}_\gamma \cup \Omega \times \{\gamma\}$ is the Dirichlet boundary. The set $\mathcal{P}_1(K)$ is either the space $\mathbb{P}_1(K)$ of polynomials of total degree at most 1, when the base K of an element $T = K \times I$ is simplicial, or the space of polynomials $\mathbb{Q}_1(K)$ of degree not larger than 1 in each variable provided K is an n -rectangle. We also define $\mathbb{U}(\mathcal{T}_\Omega) = \text{tr}_\Omega \mathbb{V}(\mathcal{T}_\Omega)$. Note that $\#\mathcal{T}_\gamma = M \#\mathcal{T}_\Omega$, and that $\#\mathcal{T}_\Omega \approx M^n$ implies $\#\mathcal{T}_\gamma \approx M^{n+1}$.

The Galerkin approximation of (2.9) is the function $V_{\mathcal{T}_\gamma} \in \mathbb{V}(\mathcal{T}_\gamma)$ such that

$$\int_{\mathcal{C}_\gamma} y^\alpha \nabla V_{\mathcal{T}_\gamma} \nabla W = d_s(f, \text{tr}_\Omega W), \quad \forall W \in \mathbb{V}(\mathcal{T}_\gamma). \quad (2.14)$$

Existence and uniqueness of $V_{\mathcal{T}_\gamma}$ immediately follows from $\mathbb{V}(\mathcal{T}_\gamma) \subset \hat{H}_L^1(y^\alpha, \mathcal{C}_\gamma)$ and the Lax-Milgram Lemma. It is trivial also to obtain a best approximation result *à la* Cea. This reduces the numerical analysis of (2.14) to a question in approximation theory which in turn can be answered by the study of piecewise polynomial interpolation in Muckenhoupt weighted Sobolev spaces; see [60, 63]. Exploiting the Cartesian structure of the mesh we can handle the anisotropy, construct a quasi interpolant $\Pi_{\mathcal{T}_\gamma} : L^1(\mathcal{C}_\gamma) \rightarrow \mathbb{V}(\mathcal{T}_\gamma)$, thus extending [36], and obtain

$$\begin{aligned} \|v - \Pi_{\mathcal{T}_\gamma} v\|_{L^2(y^\alpha, T)} &\lesssim h_K \|\nabla_{x'} v\|_{L^2(y^\alpha, S_T)} + h_I \|\partial_y v\|_{L^2(y^\alpha, S_T)}, \\ \|\partial_{x_j} (v - \Pi_{\mathcal{T}_\gamma} v)\|_{L^2(y^\alpha, T)} &\lesssim h_K \|\nabla_{x'} \partial_{x_j} v\|_{L^2(y^\alpha, S_T)} + h_I \|\partial_y \partial_{x_j} v\|_{L^2(y^\alpha, S_T)}, \end{aligned}$$

with $j = 1, \dots, n + 1$; see [60, Theorems 4.6–4.8] and [63]. However, since (2.11) implies $\mathcal{U} \notin H^2(y^\alpha, \mathcal{C}_\gamma)$ the second estimate is not meaningful for $j = n + 1$. We must measure the regularity of \mathcal{U}_{yy} with a stronger weight and compensate with a graded mesh. This makes anisotropic estimates essential. These considerations allow us to obtain ([60, Theorem 5.4] and [60, Corollary 7.11]):

Theorem 2.2 (a priori error estimate). *Let $\mathcal{T}_\gamma \in \mathbb{T}$ and $\mathbb{V}(\mathcal{T}_\gamma)$ be defined by (2.13). If $V_{\mathcal{T}_\gamma} \in \mathbb{V}(\mathcal{T}_\gamma)$ solves (2.14), then we have*

$$\|\mathcal{U} - V_{\mathcal{T}_\gamma}\|_{\hat{H}_L^1(y^\alpha, \mathcal{C})} \lesssim |\log(\#\mathcal{T}_\gamma)|^s (\#\mathcal{T}_\gamma)^{-1/(n+1)} \|f\|_{\mathbb{H}^{1-s}(\Omega)},$$

where $\mathcal{Y} \approx \log(\#\mathcal{T}_{\mathcal{Y}})$. Alternatively, if u denotes the solution of (1.1), then

$$\|u - \text{tr}_{\Omega} V_{\mathcal{T}_{\mathcal{Y}}}\|_{\mathbb{H}^s(\Omega)} \lesssim |\log(\#\mathcal{T}_{\mathcal{Y}})|^s (\#\mathcal{T}_{\mathcal{Y}})^{-1/(n+1)} \|f\|_{\mathbb{H}^{1-s}(\Omega)}.$$

Remark 2.3 (domain and data regularity). The results of Theorem 2.2 hold only if $f \in \mathbb{H}^{1-s}(\Omega)$ and the domain Ω is such that (2.8) holds.

Remark 2.4 (quasi-uniform meshes). Let \mathcal{U} solve (1.2), and $V_{\mathcal{T}_{\mathcal{Y}}}$ be the solution of (2.14), constructed over a quasi-uniform mesh of size $h_{\mathcal{T}}$. If $f \in \mathbb{H}^{1-s}(\Omega)$ and $\mathcal{Y} \approx |\log h_{\mathcal{T}}|$, then for all $\varepsilon > 0$

$$\|\nabla(\mathcal{U} - V_{\mathcal{T}_{\mathcal{Y}}})\|_{L^2(y^{\alpha}, \mathcal{C}_{\mathcal{Y}})} \lesssim h_{\mathcal{T}}^{s-\varepsilon} \|f\|_{\mathbb{H}^{1-s}(\Omega)}, \quad (2.15)$$

where the hidden constant blows up if ε tends to 0.

Remark 2.5 (Case $s = \frac{1}{2}$). If $s = \frac{1}{2}$, then there is no weight in (2.15) and we obtain the optimal estimate

$$\|\nabla(\mathcal{U} - V_{\mathcal{T}_{\mathcal{Y}}})\|_{L^2(\mathcal{C}_{\mathcal{Y}})} \lesssim h_{\mathcal{T}} \|f\|_{H_{00}^{1/2}(\Omega)}.$$

2.3. Numerical illustrations. Let us present several numerical examples. These have been implemented with the `deal.II` library [8, 9]. Integrals are evaluated with Gaussian quadratures of sufficiently high order and linear systems are solved using CG with ILU preconditioner and the exit criterion being that the ℓ^2 -norm of the residual is less than 10^{-12} .

2.3.1. Quasi-uniform meshes. According to [60], $\partial_y \mathcal{U} \approx y^{-\alpha}$, which formally implies $\partial_y \mathcal{U} \in H^{s-\varepsilon}(y^{\alpha}, \mathcal{C})$ provided $f \in \mathbb{H}^{1-s}(\Omega)$. In this sense (2.15) seems sharp with respect to regularity even though it does not exhibit the optimal rate. We verify this with a simple numerical illustration in Figure 2.1.

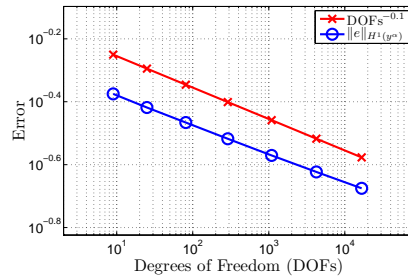


Figure 2.1. Rate of convergence $\#(\mathcal{T}_{\mathcal{Y}})^{-s/(n+1)}$ for quasi-uniform meshes: $s = 0.2$, $n = 1$.

Let $n = 1$, $\Omega = (0, 1)$, $s = 0.2$, $f = \pi^{2s} \sin(\pi x)$, then $u(x) = \sin(\pi x)$, and the solution \mathcal{U} to (1.2) is $\mathcal{U}(x, y) = 2^{1-s} \pi^s \Gamma(s)^{-1} \sin(\pi x) K_s(\pi y)$, where by K_s we denote the modified Bessel function of the second kind; [60, §2.4]. Figure 2.1 shows the rate of convergence for the $H^1(y^{\alpha}, \mathcal{C}_{\mathcal{Y}})$ -seminorm. Estimate (2.15) predicts a rate of $h_{\mathcal{T}}^{0.2-\varepsilon}$. In other words the rate of convergence, when measured in terms of degrees of freedom, is $(\#\mathcal{T}_{\mathcal{Y}})^{-0.1-\varepsilon}$, which is what Figure 2.1 displays.

2.3.2. Graded meshes: a square domain. Let $\Omega = (0, 1)^2$. Then

$$\varphi_{m,n}(x_1, x_2) = \sin(m\pi x_1) \sin(n\pi x_2), \quad \lambda_{m,n} = \pi^2 (m^2 + n^2), \quad m, n \in \mathbb{N}.$$

If $f(x_1, x_2) = (2\pi^2)^s \sin(\pi x_1) \sin(\pi x_2)$, then $u(x_1, x_2) = \sin(\pi x_1) \sin(\pi x_2)$, by (2.1) and $\mathcal{U}(x_1, x_2, y) = 2^{1-s/2} \pi^s \Gamma(s)^{-1} \sin(\pi x_1) \sin(\pi x_2) y^s K_s(\sqrt{2}\pi y)$ [60, (2.24)].

We construct a sequence of meshes $\{\mathcal{T}_{\mathcal{Y}_k}\}_{k \geq 1}$, where \mathcal{T}_Ω is obtained by uniform refinement and $\mathcal{T}_{\mathcal{Y}_k}$ is given by (2.12) with parameter $\gamma > 3/(1 - \alpha)$. On the basis of Theorem 2.2, the truncation parameter \mathcal{Y}_k is chosen by

$$\mathcal{Y}_k \geq \frac{2}{\sqrt{\lambda_1}} (\log C - \log(\#\mathcal{T}_{\mathcal{Y}_{k-1}}))^{-1/3}.$$

With this type of meshes,

$$\|u - \text{tr}_\Omega V_{\mathcal{T}_{\mathcal{Y}_k}}\|_{\mathbb{H}^s(\Omega)} \lesssim \|\mathcal{U} - V_{\mathcal{T}_{\mathcal{Y}_k}}\|_{\hat{H}_L^1(y^\alpha, \mathcal{C})} \lesssim |\log(\#\mathcal{T}_{\mathcal{Y}_k})|^s \cdot (\#\mathcal{T}_{\mathcal{Y}_k})^{-1/3},$$

which is near-optimal in \mathcal{U} but suboptimal in u , since we should expect (see [21])

$$\|u - \text{tr}_\Omega V_{\mathcal{T}_{\mathcal{Y}_k}}\|_{\mathbb{H}^s(\Omega)} \lesssim h_{\mathcal{T}_{\mathcal{Y}_k}}^{2-s} \lesssim (\#\mathcal{T}_{\mathcal{Y}_k})^{-(2-s)/3}.$$

Figure 2.3.2 shows the rates of convergence for $s = 0.2$ and $s = 0.8$ respectively. In both cases, we obtain the rate given by Theorem 2.2.

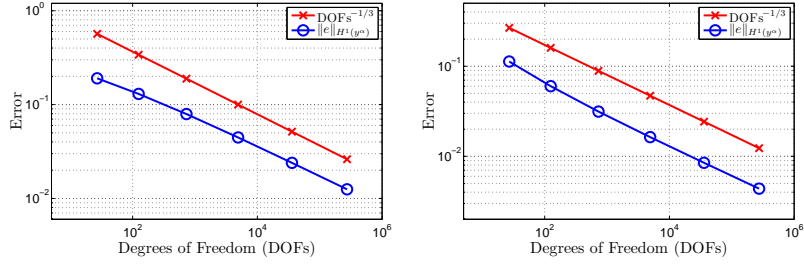


Figure 2.2. Computational rate of convergence for a square and graded meshes. The left panel shows the rate for $s = 0.2$ and the right one for $s = 0.8$. The experimental rate of convergence is $(\#\mathcal{T}_{\mathcal{Y}_k})^{-1/3}$, in agreement with Theorem 2.2.

2.3.3. Graded meshes: a circular domain. If $\Omega = \{|x'| \in \mathbb{R}^2 : |x'| < 1\}$, then

$$\varphi_{m,n}(r, \theta) = J_m(j_{m,n}r) (A_{m,n} \cos(m\theta) + B_{m,n} \sin(m\theta)), \quad \lambda_{m,n} = j_{m,n}^2, \quad (2.16)$$

where J_m is the m -th Bessel function of the first kind; $j_{m,n}$ is the n -th zero of J_m and $A_{m,n}, B_{m,n}$ are normalization constants to ensure $\|\varphi_{m,n}\|_{L^2(\Omega)} = 1$.

If $f = (\lambda_{1,1})^s \varphi_{1,1}$, then (2.1) and [60, (2.24)] show that $u = \varphi_{1,1}$ and

$$\mathcal{U}(r, \theta, y) = 2^{1-s} \Gamma(s)^{-1} (\lambda_{1,1})^{s/2} \varphi_{1,1}(r, \theta) y^s K_s(\sqrt{2}\pi y).$$

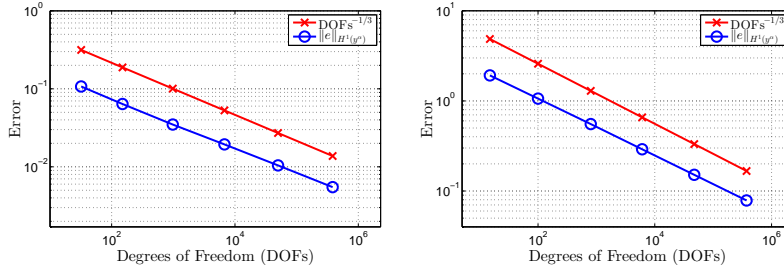


Figure 2.3. Computational rate of convergence for a circle with graded meshes. The left panel shows the rate for $s = 0.3$ and the right one for $s = 0.7$. The experimental rate of convergence is $(\#\mathcal{T}_{y_k})^{-1/3}$, in agreement with Theorem 2.2.

We construct a sequence of meshes $\{\mathcal{T}_{y_k}\}_{k \geq 1}$ as in §2.3.2. With these meshes

$$\|\mathcal{U} - V_{\mathcal{T}_{y_k}}\|_{\hat{H}_T^1(\mathcal{C}, y^\alpha)} \lesssim |\log(\#\mathcal{T}_{y_k})|^s (\#\mathcal{T}_{y_k})^{-1/3}, \quad (2.17)$$

which is near-optimal. Figure 2.3 shows the errors of $\|\mathcal{U} - V_{\mathcal{T}_{y_k}}\|_{H^1(y^\alpha, \mathcal{C}_{y_k})}$ for $s = 0.3$ and $s = 0.7$. The results, again, are in agreement with Theorem 2.2.

3. A posteriori error analysis and adaptivity

Starting with the pioneering works [6, 7] in the late 1970's, adaptive finite element methods (AFEM) have been the subject of intense research. This is because AFEM yields optimal performance in situations where classical FEM cannot. In our problem we are incorporating one extra dimension and so AFEM is essential to retain computational efficiency. In addition, the a priori theory requires $f \in \mathbb{H}^{1-s}(\Omega)$ and (2.8). If either of these does not hold \mathcal{U} may have singularities in the x' -variables and exhibit fractional regularity. This would not allow us to attain the almost optimal rate of convergence of Theorem 2.2: a quasi-uniform refinement of Ω would not result in an efficient solution technique. An adaptive loop driven by an a posteriori error estimator is *fundamental* to recover optimal rates of convergence. Let us explore this now and develop a new estimator.

3.1. Residual estimators. A residual error estimator uses the strong form of the local residual to measure the error. Let $T \in \mathcal{T}_y$ and ν be the unit outer normal to T . Integration by parts yields

$$\int_T y^\alpha \nabla V_{\mathcal{T}_y} \cdot \nabla W = \int_{\partial T} W y^\alpha \nabla V_{\mathcal{T}_y} \cdot \nu - \int_T \operatorname{div}(y^\alpha \nabla V_{\mathcal{T}_y}) W.$$

Since $\alpha \in (-1, 1)$ the boundary integral is meaningless for $y = 0$. Even the very first step in the derivation of a residual a posteriori error estimator fails! There is nothing left to do but to consider a different type of estimator.

3.2. Local problems on stars over isotropic refinements. Following [6, 56] we can construct, over shape regular meshes, an error estimator based on the solution of small problems on stars. Its construction and analysis is similar to the developments of §3.3, so we skip details. Under some assumptions this estimator is equivalent to the error up to data oscillation; see [62]. We designed an adaptive algorithm driven by this error estimator on shape regular meshes [27, 56], and we illustrate its performance with a simple but revealing numerical example.

Let $\Omega = (0, 1)$ and $s \in (0, 1)$. If $f(x') = \pi^{2s} \sin(\pi x')$, then $u(x') = \sin(\pi x')$, and $\mathcal{W}(x', y) = 2^{1-s} \pi^s \Gamma(s)^{-1} \sin(\pi x') K_s(\pi y)$ solves (1.2). We point out that we are solving a two dimensional problem so the optimal rate we expect is $\mathcal{O}(\#\mathcal{T}_y^{-0.5})$. Figure 3.1 shows the experimental rate of convergence for $s = 0.2$ and $s = 0.6$ which, as we see, is $\mathcal{O}(\#\mathcal{T}_y^{-s/2})$ and coincides with the suboptimal one obtained with quasi-uniform refinement [60, §5.1]. These numerical experiments show that adaptive isotropic refinement cannot be optimal, thus justifying the need to introduce *cylindrical stars* together with a new anisotropic error estimator, which will treat the x' -coordinates and the extended direction y separately.

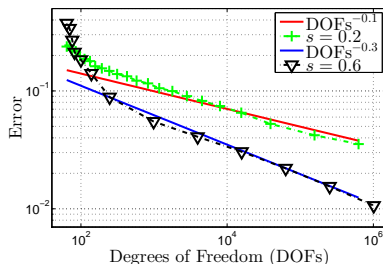


Figure 3.1. Computational rate of convergence $\#\mathcal{T}_y^{-s/2}$ for an isotropic adaptive algorithm for $n = 1$, $s = 0.2$ and $s = 0.6$.

3.3. Local problems over cylindrical stars. Inspired by [6, 56], we deal with the anisotropy of the mesh in the extended variable y and the coefficient y^α by considering local problems on cylindrical stars. The solutions of these local problems allow us to define an anisotropic a posteriori error estimator which, under certain assumptions, is equivalent to the error up to data oscillation terms.

Given a node z on the mesh \mathcal{T}_y , we exploit the tensor product structure of \mathcal{T}_y , and we write $z = (z', z'')$ where z' and z'' are nodes on the meshes \mathcal{T}_Ω and \mathcal{I}_y respectively. For $K \in \mathcal{T}_\Omega$, we denote by $\mathcal{N}(K)$ and $\mathring{\mathcal{N}}(K)$ the set of nodes and interior nodes of K , respectively. We set

$$\mathcal{N}(\mathcal{T}_\Omega) = \bigcup_{K \in \mathcal{T}_\Omega} \mathcal{N}(K), \quad \mathring{\mathcal{N}}(\mathcal{T}_\Omega) = \bigcup_{K \in \mathcal{T}_\Omega} \mathring{\mathcal{N}}(K).$$

The *star* around z' is $S_{z'} = \bigcup_{K \ni z'} K \subset \Omega$, and the *cylindrical star* around z' is

$$\mathcal{C}_{z'} := \bigcup \{T \in \mathcal{T}_y : T = K \times I, K \ni z'\} = S_{z'} \times (0, \mathcal{Y}) \subset \mathcal{C}_y.$$

For each node $z' \in \mathcal{N}(\mathcal{T}_\Omega)$ we define the discrete space

$$\mathbb{W}(\mathcal{C}_{z'}) = \{W \in C(\bar{\mathcal{C}}_{z'}) : W|_T \in \mathcal{P}_2(K) \otimes \mathbb{P}_2(I) \ \forall T = K \times I \in \mathcal{C}_{z'}, \\ W|_{\partial\mathcal{C}_{z'} \setminus \Omega \times \{0\}} = 0\},$$

where, if K is a quadrilateral, $\mathcal{P}_2(K) = \mathbb{Q}_2(K)$ — the space of polynomials of degree not larger than 2 in each variable. If K is a simplex, $\mathcal{P}_2(K) = \mathbb{P}_2(K) \oplus \mathbb{B}(K)$ where $\mathbb{P}_2(K)$ is the space of polynomials of total degree at most 2, and $\mathbb{B}(K)$ is the space spanned by a local cubic bubble function.

For each cylindrical star $\mathcal{C}_{z'}$ we define $\eta_{z'} \in \mathbb{W}(\mathcal{C}_{z'})$ to be the solution of

$$\int_{\mathcal{C}_{z'}} y^\alpha \nabla \eta_{z'} \nabla W = d_s \langle f, \text{tr}_\Omega W \rangle - \int_{\mathcal{C}_{z'}} y^\alpha \nabla V_{\mathcal{T}_y} \nabla W, \quad \forall W \in \mathbb{W}(\mathcal{C}_{z'}).$$

We finally define the local indicators $\mathcal{E}_{z'}$ and global error estimators $\mathcal{E}_{\mathcal{T}_y}$ as follows:

$$\mathcal{E}_{z'} = \|\eta_{z'}\|_{\mathcal{C}_{z'}} \quad \mathcal{E}_{\mathcal{T}_y}^2 = \sum_{z' \in \mathcal{N}(\mathcal{T}_\Omega)} \mathcal{E}_{z'}^2. \quad (3.1)$$

We can obtain a local lower bound for the error without any oscillation term and free of any constant.

Theorem 3.1 (localized lower bound). *Let $v \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_y)$ and $V_{\mathcal{T}_y} \in \mathbb{V}(\mathcal{T}_y)$ solve (2.9) and (2.14) respectively. Then, for any $z' \in \mathcal{N}(\mathcal{T}_\Omega)$, we have*

$$\mathcal{E}_{z'} \leq \|\nabla(v - V_{\mathcal{T}_y})\|_{L^2(y^\alpha, \mathcal{C}_{z'})}.$$

For $z' \in \mathcal{N}(\mathcal{T}_\Omega)$, we define the local *data oscillation* as

$$\text{osc}_{z'}(f)^2 := d_s h_{z'}^{2s} \|f - f_{z'}\|_{L^2(S_{z'})}^2, \quad f_{z'}|_K := \int_K f, \quad (3.2)$$

whence the global data oscillation reads

$$\text{osc}_{\mathcal{T}_\Omega}(f)^2 := \sum_{z' \in \mathcal{N}(\mathcal{T}_\Omega)} \text{osc}_{z'}(f)^2.$$

To mark elements for refinement we use the *total error indicator*

$$\tau_{\mathcal{T}_\Omega}(V_{\mathcal{T}_y}, S_{z'}) := (\mathcal{E}_{z'}^2 + \text{osc}_{z'}(f)^2)^{1/2} \quad \forall z' \in \mathcal{N}(\mathcal{T}_\Omega). \quad (3.3)$$

Let $\mathcal{H}_{\mathcal{T}_\Omega} = \{S_{z'} : z' \in \mathcal{N}(\mathcal{T}_\Omega)\}$ and, for any $\mathcal{M} \subset \mathcal{H}_{\mathcal{T}_\Omega}$, we set

$$\tau_{\mathcal{T}_\Omega}(V_{\mathcal{T}_y}, \mathcal{M}) := \left(\sum_{S_{z'} \in \mathcal{M}} \tau_{\mathcal{T}_\Omega}(V_{\mathcal{T}_y}, S_{z'})^2 \right)^{1/2}. \quad (3.4)$$

Under certain assumptions we can bound the error by the estimator, up to oscillation terms; see [30] for details.

Theorem 3.2 (global upper bound). *Let $v \in \mathring{H}_L^1(\mathcal{C}_y, y^\alpha)$ and $V_{\mathcal{T}_y} \in \mathbb{V}(\mathcal{T}_y)$ solve (2.9) and (2.14), respectively. The total error estimator $\tau_{\mathcal{T}_\Omega}(V_{\mathcal{T}_y}, \mathcal{H}_{\mathcal{T}_\Omega})$, defined in (3.4) satisfies*

$$\|\nabla(v - V_{\mathcal{T}_y})\|_{L^2(y^\alpha, \mathcal{C}_y)} \lesssim \tau_{\mathcal{T}_\Omega}(V_{\mathcal{T}_y}, \mathcal{H}_{\mathcal{T}_\Omega}).$$

3.4. Numerical experiments. We now illustrate the performance of the a posteriori error estimator (3.1). We use an almost standard adaptive loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}. \quad (3.5)$$

where the modules in (3.5) are as follows:

- **SOLVE:** Given a mesh \mathcal{T}_y we compute $V_{\mathcal{T}_y}$, the solution of (2.14).
- **ESTIMATE:** Given $V_{\mathcal{T}_y}$ we calculate the local error indicators (3.1) and the local oscillations (3.2) to construct the total error indicator (3.3):

$$\{\tau_{\mathcal{T}_\Omega}(V_{\mathcal{T}_y}, S_{z'})\}_{S_{z'} \in \mathcal{K}_{\mathcal{T}_\Omega}} = \text{ESTIMATE}(V_{\mathcal{T}_y}, \mathcal{T}_y).$$

- **MARK:** Using Dörfler marking [33] with parameter $0 < \theta \leq 1$ we select a set

$$\mathcal{M} = \text{MARK}(\{\tau_{\mathcal{T}_\Omega}(V_{\mathcal{T}_y}, S_{z'})\}_{S_{z'} \in \mathcal{K}_{\mathcal{T}_\Omega}}, V_{\mathcal{T}_y}) \subset \mathcal{K}_{\mathcal{T}_\Omega}$$

of minimal cardinality that satisfies $\tau_{\mathcal{T}_\Omega}(V_{\mathcal{T}_y}, \mathcal{M}) \geq \theta \tau_{\mathcal{T}_\Omega}(V_{\mathcal{T}_y}, \mathcal{K}_{\mathcal{T}_\Omega})$.

- **REFINE:** We generate a new mesh \mathcal{T}'_Ω by bisecting all the elements $K \in \mathcal{T}_\Omega$ contained in \mathcal{M} based on newest-vertex bisection method; [65, 64]. We choose the truncation parameter as $\mathcal{Y} = 1 + \frac{1}{3} \log(\#\mathcal{T}'_\Omega)$ [60, Remark 5.5]. We set $M \approx (\#\mathcal{T}'_\Omega)^{1/n}$ and construct \mathcal{I}'_y by the rule (2.12). The new mesh $\mathcal{T}'_y = \text{REFINE}(\mathcal{M})$ is obtained as the tensor product of \mathcal{T}'_Ω and \mathcal{I}'_y .

3.4.1. Smooth but incompatible data. The example of [60, §6.3] shows that Theorem 2.2 is sharp: $f \in \mathbb{H}^{1-s}(\Omega)$ is necessary to obtain an optimal rate of convergence with a quasioptimal mesh in the x' -direction. A certain compatibility between the data and the boundary condition is necessary. Moreover, [60, §6.3] shows how, in some simple cases, one can guess the singularity and a priori design a mesh that captures it and recover the optimal rate of convergence. This is not always possible and here we show that the estimator (3.1) automatically produces a sequence of meshes that yield the optimal rate of convergence. Consider $\Omega = (0, 1)^2$ and $f = 1$. From (2.2) we see that functions in $\mathbb{H}^{1-s}(\Omega)$, for $1 - s > \frac{1}{2}$, must have a vanishing trace. Therefore, if $s < \frac{1}{2}$, $f \notin \mathbb{H}^{1-s}(\Omega)$ and Theorem 2.2 cannot be invoked. Nevertheless, as Figure 3.2 shows, we recover the optimal rate.

3.4.2. L-shaped domain with incompatible data. We now combine the singularity introduced by the data incompatibility of § 3.4.1 and the effect of a reentrant corner. Consider $\Omega = (-1, 1)^2 \setminus (0, 1) \times (-1, 0)$ and $f = 1$. As Figure 3.3 displays, we recover the optimal rate of convergence for all possible cases of s .

4. Multilevel methods

Since we increase the dimension by one in the approximation of the nonlocal operator $(-\Delta)^s$ by the local problem (1.2), it becomes essential to develop efficient

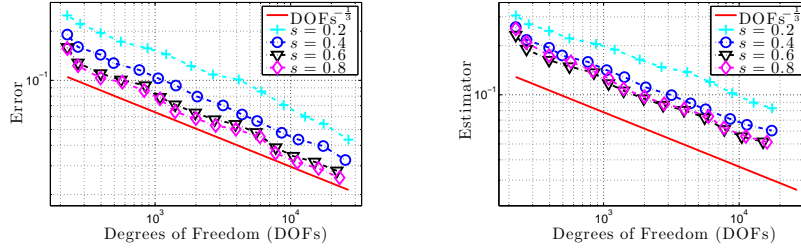


Figure 3.2. Computational rate of convergence of AFEM for § 3.4.1, $n = 2$ and $s = 0.2, 0.4, 0.6$ and $s = 0.8$. The left panel shows the the error vs. the number of degrees of freedom, the right one the total error indicator. We recover the optimal rate $\#(\mathcal{T}_\gamma)^{-1/3}$. For $s < \frac{1}{2}$, the right hand side $f = 1 \notin \mathbb{H}^{1-s}(\Omega)$ and a quasiuniform mesh in Ω does not deliver the optimal rate of convergence [60, §6.3].

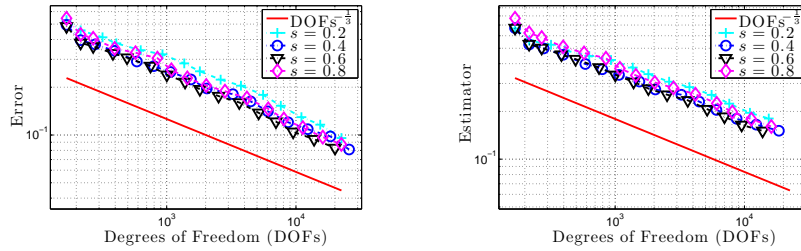


Figure 3.3. Computational rate of convergence of AFEM on the smooth but incompatible right hand side over an L-shaped domain of § 3.4.2 for $n = 2$ and $s = 0.2, 0.4, 0.6$ and $s = 0.8$. The left panel shows the error vs. the number of degrees of freedom, whereas the right one that for the total error indicator. In all cases we recover the optimal rate $\#(\mathcal{T}_\gamma)^{-1/3}$.

linear algebra methods. It is known that multilevel methods are the most efficient techniques for the solution of discretizations of PDEs [19, 20, 43, 75]. Multigrid methods for equations of the type (1.2), however, have not been explored.

Using the multilevel framework of [14, 15, 75], the Xu-Zikatanov identity [77] and exploiting the fact that $|y|^\alpha \in A_2(\mathbb{R}^{n+1})$, which turns out to be critical, we derived in [29] an almost uniform convergent multilevel method to solve (2.14). As we have shown the mesh in the extended dimension must be graded towards the bottom of the cylinder thus becoming anisotropic. We apply line smoothers over vertical lines in the extended domain and prove that the corresponding multigrid V-cycle converges nearly uniformly, i.e., the contraction factor depends linearly on the number of levels, and thus logarithmically on the problem size.

4.1. Multilevel decomposition and multigrid algorithm. We follow [13, 14] to present a multilevel decomposition of $\mathbb{V}(\mathcal{T}_\gamma)$. To simplify the

analysis and implementation, we consider a sequence of nested discretizations constructed as follows: We introduce a sequence of nested uniform partitions of the unit interval $\{\hat{\mathcal{I}}_k\}$, with mesh points $\hat{y}_{l,k}$, for $l = 0, \dots, M_k$ and $k = 0, \dots, J$. For $\gamma > 3/(1 - \alpha)$ the family $\{\mathcal{I}_k\}$ is given by

$$y_{l,k} = \mathcal{Y}\hat{y}_{l,k}^\gamma, \quad l = 0, \dots, M_k. \quad (4.1)$$

For $k = 0, \dots, J$, $\mathcal{T}_{\Omega,k}$ is obtained by uniform refinement. The mesh $\mathcal{T}_{\mathcal{Y},k}$ is the tensor product of $\mathcal{T}_{\Omega,k}$ and \mathcal{I}_k , given by (4.1).

If $\mathbb{V}_k := \mathbb{V}(\mathcal{T}_{\mathcal{Y},k})$ we have the sequence $\mathbb{V}_0 \subset \mathbb{V}_1 \subset \dots \subset \mathbb{V}_J = \mathbb{V}$, and a space macro-decomposition $\mathbb{V} = \sum_{k=0}^J \mathbb{V}_k$. We now introduce a space micro-decomposition: For $j = 1, \dots, \mathcal{M}_k$ let $\mathcal{I}_{k,j}$ be a subset of $\{1, 2, \dots, \mathcal{N}_k\}$, and assume that $\cup_{j=1}^{\mathcal{M}_k} \mathcal{I}_{k,j} = \{1, 2, \dots, \mathcal{N}_k\}$. The sets $\mathcal{I}_{k,j}$ may not be disjoint but the cardinality of their intersection is bounded independently of J and \mathcal{N}_k . If the nodal basis of \mathbb{V}_k is $\phi_{k,i}$, we define $\mathbb{V}_{k,j} = \text{span}\{\phi_{k,i} : i \in \mathcal{I}_{k,j}\}$ and we have the micro-decomposition

$$\mathbb{V} = \sum_{k=0}^J \sum_{j=1}^{\mathcal{M}_k} \mathbb{V}_{k,j}. \quad (4.2)$$

With this notation we define a symmetric V-cycle multigrid method as in [16, Algorithm 3.1], with $m \geq 1$ pre and post smoothing steps. When $m = 1$, it is equivalent to the application of successive subspace corrections to the decomposition (4.2) with exact sub-solvers so that the V-cycle multigrid method has a smoother at each level of block Gauss-Seidel type [14, 75]. In particular, the nodal decomposition $\mathcal{I}_{k,j} = \{j\}$ yields a point-wise Gauss-Seidel smoother. If the indices in $\mathcal{I}_{k,j}$ are such that the corresponding vertices lie on a vertical line, we obtain *line smoothers*, which are essential to handle anisotropy.

4.2. Analysis of the multigrid method. The success of multigrid methods for uniformly elliptic operators is because smoothers are effective in reducing the high frequency components of the error while coarse grid corrections reduce the low frequency ones. However, their effectiveness depends on several factors such as the shape of the mesh. A key ingredient in the design and analysis of a multigrid method on anisotropic meshes is the use of line smoothers [4, 17, 44, 71].

When solving (2.14) on graded meshes, the approximation on the coarse grid is dominated by the larger mesh size in the x' -direction and thus the coarse grid correction cannot capture the smaller scale in the y -direction. This is why line smoothing, i.e., solving sub-problems restricted to one vertical line is necessary. Owing to the nature of the decomposition, the smoother requires the evaluation of the inverse of the operator over a vertical line. This can be efficiently realized since the corresponding matrix is tridiagonal.

Our analysis hinges on [77], which requires stability of the micro-decomposition.

Lemma 4.1 (nodal stability and anisotropic inverse inequalities). *Let $\mathcal{T}_{\mathcal{Y}} \in \mathcal{T}$ with \mathcal{T}_{Ω} quasiuniform and $\mathcal{I}_{\mathcal{Y}}$ graded so that (4.1) holds. If $v = \sum_{j=1}^{\mathcal{M}_J} v_j \in \mathbb{V}_J$,*

then

$$\sum_{j=1}^{\mathcal{M}_J} \|v_j\|_{L^2(y^\alpha, \mathcal{C}_y)}^2 \lesssim \|v\|_{L^2(y^\alpha, \mathcal{C}_y)}^2 \lesssim \sum_{j=1}^{\mathcal{M}_J} \|v_j\|_{L^2(y^\alpha, \mathcal{C}_y)}^2. \quad (4.3)$$

Moreover, for $T = K \times I \in \mathcal{T}_y$ we have the following local inverse inequalities

$$\|\nabla_{x'} v\|_{L^2(y^\alpha, T)} \lesssim h_K^{-1} \|v\|_{L^2(y^\alpha, T)}, \quad \|\partial_y v\|_{L^2(y^\alpha, T)} \lesssim h_I^{-1} \|v\|_{L^2(y^\alpha, T)}. \quad (4.4)$$

We now examine the V-cycle multigrid method applied to the decomposition (4.2) with exact sub-solvers on $\mathbb{V}_{k,j}$, i.e., with line smoothers; see [16, §III.12] and [74]. A key observation in favor of subspaces $\{\mathbb{V}_{k,j}\}_{j=1}^{\mathcal{M}_k}$ follows.

Lemma 4.2 (nodal stability of y -derivatives). *In the setting of Lemma 4.1 we have*

$$\sum_{j=1}^{\mathcal{M}_J} \|\partial_y v_j\|_{L^2(y^\alpha, \mathcal{C}_y)}^2 \lesssim \|\partial_y v\|_{L^2(y^\alpha, \mathcal{C}_y)}^2 \lesssim \sum_{j=1}^{\mathcal{M}_J} \|\partial_y v_j\|_{L^2(y^\alpha, \mathcal{C}_y)}^2. \quad (4.5)$$

We use the quasi-interpolant of [63], Lemmas 4.1 and 4.2 and obtain nearly uniform convergence of the V-cycle multigrid method. We follow [74, 76].

Theorem 4.3 (convergence of multigrid with line smoothers). *The symmetric V-cycle multigrid method with line smoothing converges with a contraction rate*

$$\delta \leq 1 - \frac{1}{1 + CJ},$$

where C is independent of \mathcal{N}_J and depends on y^α only through C_{2,y^α} .

4.3. Numerical illustrations. We consider two cases:

- $n = 1$, $\Omega = (0, 1)$, $u = \sin(3\pi x)$,
- $n = 2$, $\Omega = (0, 1)^2$, $u = \sin(2\pi x_1) \sin(2\pi x_2)$,

and $\mathcal{Y} = 1$, which has been chosen, as discussed in [60], to capture the exponential decay of the solution. All of our algorithms are implemented based on the MATLAB[©] software package *iFEM* [28].

Table 4.1 shows the number of iterations for the one and two dimensional problems, respectively. As we see, the method converges almost uniformly with respect to the number of degrees of freedom.

5. Space-time fractional diffusion problems

We now review the numerical approximation of an initial boundary value problem for a space-time fractional parabolic equation developed in [62]. Given $s \in (0, 1)$, $\gamma \in (0, 1]$, a function f , and an initial datum u_0 , we seek u such that

$$\partial_t^\gamma u + (-\Delta)^s u = f \quad \text{in } \Omega \times (0, T), \quad u(0) = u_0 \quad \text{in } \Omega. \quad (5.1)$$

$h_{\mathcal{T}_\Omega}$	DOFs	$s = 0.15$	$s = 0.3$	$s = 0.6$	$s = 0.8$
$\frac{1}{16}$	289	7	6	5	5
$\frac{1}{32}$	1,089	9	9	6	6
$\frac{1}{64}$	4,225	10	10	6	6
$\frac{1}{128}$	16,641	10	11	6	6
$\frac{1}{256}$	66,049	11	10	6	6
$\frac{1}{512}$	263,169	11	10	6	7

$h_{\mathcal{T}_\Omega}$	DOFs	$s = 0.15$	$s = 0.3$	$s = 0.6$	$s = 0.8$
$\frac{1}{16}$	4,913	8	7	6	5
$\frac{1}{32}$	35,937	11	8	6	6
$\frac{1}{64}$	274,625	12	9	6	6
$\frac{1}{128}$	2,146,689	13	9	6	6

Table 4.1. Number of iterations for a multigrid method for $(-\Delta)^s u = f$ with $n = 1$ (top) and $n = 2$ (bottom) using a line smoother in the extended direction. The mesh in Ω is uniform with mesh size $h_{\mathcal{T}_\Omega}$, whereas that in the extended direction $(0, \mathcal{Y})$ is graded according to (4.1).

If $\gamma \in (0, 1)$, ∂_t^γ denotes the left Caputo derivative of order γ , defined by

$$\partial_t^\gamma u(x, t) := \frac{1}{\Gamma(1-\gamma)} \int_0^t \frac{1}{(t-r)^\gamma} \frac{\partial u(x, r)}{\partial r} dr, \quad (5.2)$$

where Γ is the Gamma function. For $\gamma = 1$, we consider the usual derivative ∂_t .

Fractional derivatives and integrals, like the Caputo derivative, are a powerful tool to describe memory and hereditary properties of materials [41]. Problem (5.1) is motivated by anomalous diffusion processes (subdiffusion) and models dynamical systems with chaotic motion [69], fractional PID controllers [68], subdiffusion phenomena in highly heterogeneous aquifers [59] and plasma turbulence [32].

Problem (5.1) is complicated due to the nonlocality of the fractional time derivative [47, 70] besides the presence of the nonlocal space operator. We overcome the latter with the Caffarelli-Silvestre extension of § 2.1 and rewrite (5.1) as an elliptic problem with dynamic boundary condition:

$$\begin{cases} \operatorname{div}(y^\alpha \nabla \mathcal{U}) = 0 \text{ in } \mathcal{C} \times (0, T), & \mathcal{U} = u_0, \text{ on } \Omega \times \{0\}, t = 0 \\ \mathcal{U} = 0 \text{ on } \partial_L \mathcal{C} \times (0, T) & d_s \partial_t^\gamma \mathcal{U} + \partial_\nu^\alpha \mathcal{U} = d_s f \text{ on } (\Omega \times \{0\}) \times (0, T). \end{cases} \quad (5.3)$$

There are several approaches via finite differences, finite elements and spectral methods to treat the Caputo derivative of order γ . We refer to [58, §1] for an overview. In [52, 53] a finite difference scheme is proposed, which has a consistency error $\mathcal{O}(\tau^{2-\gamma})$, where τ denotes the time step. This estimate, however, requires a rather strong regularity assumption in time [55, 62]. In [62], we examined the behavior of $\partial_t u$ and $\partial_{tt} u$ when $t \downarrow 0$ and derived realistic time-regularity estimates

for u ; see also [55, 58]. With these refined results we analyzed the truncation error and showed discrete stability. The latter leads to an energy estimate.

5.1. The Caffarelli-Silvestre extension.

We define

$$\begin{aligned}\mathcal{W} &:= \{w \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; \mathbb{H}^s(\Omega)) : \partial_t^\gamma w \in L^2(0, T; \mathbb{H}^{-s}(\Omega))\}, \\ \mathcal{V} &:= \{w \in L^2(0, T; \dot{H}_L^1(y^\alpha, \mathcal{C})) : \partial_t^\gamma \text{tr}_\Omega w \in L^2(0, T; \mathbb{H}^{-s}(\Omega))\}.\end{aligned}$$

Thus, given $f \in L^2(0, T; \mathbb{H}^{-s}(\Omega))$, a function $u \in \mathcal{W}$ solves (5.1) if and only if the α -harmonic extension $\mathcal{U} \in \mathcal{V}$ solves (5.3). A weak formulation of (5.3) reads: Find $\mathcal{U} \in \mathcal{V}$ such that $\text{tr}_\Omega \mathcal{U}(0) = u_0$ and, for a.e. $t \in (0, T)$,

$$\langle \text{tr}_\Omega \partial_t^\gamma \mathcal{U}, \text{tr}_\Omega \phi \rangle + \frac{1}{d_s} \int_{\mathcal{C}} y^\alpha \nabla \mathcal{U} \nabla \phi = \langle f, \text{tr}_\Omega \phi \rangle \quad \forall \phi \in \dot{H}_L^1(y^\alpha, \mathcal{C}). \quad (5.4)$$

Remark 5.1 (dynamic boundary condition). Problem (5.4) is an elliptic problem with a dynamic boundary condition: $\partial_\nu^\alpha \mathcal{U} = d_s f - d_s \text{tr}_\Omega \partial_t^\gamma \mathcal{U}$ on $\Omega \times \{0\}$. Its analysis is slightly different from the standard theory for parabolic equations.

Given $s \in (0, 1)$, $\gamma \in (0, 1]$, a forcing term $f \in L^2(0, T; \mathbb{H}^{-s}(\Omega))$ and $u_0 \in L^2(\Omega)$, problems (5.1) and (5.3) have a unique solution; see [62, Theorem 2.6].

5.2. Time regularity. As in the elliptic case, in what follows we assume that (2.8) holds. We refer to [62] for a complete space regularity analysis of problem (5.1). We focus here on time regularity.

For $\gamma = 1$, we demand sufficient time regularity of the right-hand side together with compatibility conditions for the initial datum u_0 . We express this as

$$\text{tr}_\Omega \partial_{tt} \mathcal{U} \in L^2(0, T; \mathbb{H}^{-s}(\Omega)). \quad (5.5)$$

For $\gamma \in (0, 1)$, (5.5) is inconsistent with the solution of (5.1). Properties of the Mittag-Leffler function show that (5.5) never holds if $u_0 \neq 0$. Time derivatives of u are unbounded as $t \downarrow 0$. In particular, $\partial_{tt} u(x', t) \notin L^2(0, T; \mathbb{H}^{-s}(\Omega))$. However,

$$\int_{0^+} t^\sigma \|\partial_{tt} u(\cdot, t)\|_{\mathbb{H}^{-s}(\Omega)}^2 dt$$

is finite provided $\sigma > 3 - 2\gamma$. For this reason, when $\gamma \in (0, 1)$, we assume

$$t^{\sigma/2} \text{tr}_\Omega \partial_{tt} \mathcal{U} \in L^2(0, T; \mathbb{H}^{-s}(\Omega)) \quad \sigma > 3 - 2\gamma,$$

which is a valid assumption provided $\mathcal{A}(u_0, f) < \infty$, where

$$\mathcal{A}(u_0, f) = \|u_0\|_{\mathbb{H}^s(\Omega)} + \|f\|_{H^2(0, T; \mathbb{H}^{-s}(\Omega))}. \quad (5.6)$$

Theorem 5.2 (time regularity: $\gamma \in (0, 1)$). *If $f \in H^2(0, T; \mathbb{H}^{-s}(\Omega))$ and $u_0 \in \mathbb{H}^s(\Omega)$, then for $t \in (0, T]$, the solution of (5.1) satisfies*

$$\|\partial_t u(\cdot, t) - \delta^1 u(\cdot, t)\|_{\mathbb{H}^{-s}(\Omega)} \lesssim t^{\gamma-1} \mathcal{A}(u_0, f),$$

where $\delta^1 u(\cdot, t) = t^{-1}(u(\cdot, t) - u(\cdot, 0))$. Moreover,

$$\|t^{\sigma/2} \partial_{tt} u\|_{L^2(0, T; \mathbb{H}^{-s}(\Omega))} \lesssim \mathcal{A}(u_0, f),$$

where $\sigma > 3 - 2\gamma$. The hidden constant is independent of τ but blows up as $\gamma \downarrow 0$.

5.3. Time discretization. Let $\mathcal{K} \in \mathbb{N}$ denote the number of time steps. The time step is $\tau = T/\mathcal{K}$, and, for $0 \leq k \leq \mathcal{K}$, $t_k = k\tau$ and $I_k = (t_k, t_{k+1}]$. For $\phi \in C([0, T], \mathcal{X})$ we denote $\phi^k = \phi(t_k)$ and $\phi^\tau = \{\phi^k\}_{k=0}^{\mathcal{K}}$. Moreover,

$$\|\phi^\tau\|_{\ell^\infty(\mathcal{X})} = \max_{0 \leq k \leq \mathcal{K}} \|\phi^k\|_{\mathcal{X}}, \quad \|\phi^\tau\|_{\ell^2(\mathcal{X})}^2 = \sum_{k=1}^{\mathcal{K}} \tau \|\phi^k\|_{\mathcal{X}}^2.$$

For $W^\tau \subset \mathcal{X}$ we define, for $k = 0, \dots, \mathcal{K} - 1$, $\delta^1 W^{k+1} = \tau^{-1}(W^{k+1} - W^k)$.

5.3.1. Time discretization for $\gamma = 1$. We apply the backward Euler scheme to (5.4) for $\gamma = 1$: determine $V^\tau = \{V^k\}_{k=0}^{\mathcal{K}} \subset \dot{H}_L^1(y^\alpha, \mathcal{C})$ such that

$$\text{tr}_\Omega V^0 = u_0, \quad (5.7)$$

and, if $f^{k+1} = f(t^{k+1})$ for $k = 0, \dots, \mathcal{K} - 1$, then $V^{k+1} \in \dot{H}_L^1(y^\alpha, \mathcal{C})$ solves

$$(\delta^1 \text{tr}_\Omega V^{k+1}, \text{tr}_\Omega W)_{L^2(\Omega)} + \frac{1}{d_s} \int_{\mathcal{C}} y^\alpha \nabla V^{k+1} \nabla W = \langle f^{k+1}, \text{tr}_\Omega W \rangle, \quad \forall W \in \dot{H}_L^1(y^\alpha, \mathcal{C}).$$

Define U^τ by $U^k = \text{tr}_\Omega V^k \in \mathbb{H}^s(\Omega)$, which is a piecewise constant (in time) approximation of u , solution to problem (5.1). Note that (5.7) does not require an extension of u_0 . The stability of this scheme is elementary [62, Lemma 3.3]:

$$\|\text{tr}_\Omega V^\tau\|_{\ell^\infty(L^2(\Omega))}^2 + \|V^\tau\|_{\ell^2(\dot{H}_L^1(y^\alpha, \mathcal{C}))}^2 \lesssim \|u_0\|_{L^2(\Omega)}^2 + \|f^\tau\|_{\ell^2(\mathbb{H}^{-s}(\Omega))}^2.$$

5.3.2. Time discretization for $\gamma \in (0, 1)$. We now discretize ∂_t^γ for $\gamma \in (0, 1)$. We consider the scheme proposed in [52, 53] but resort to the regularity results of Theorem 5.2. Using (5.2) and the Taylor formula with integral remainder yields

$$\partial_t^\gamma u(\cdot, t_{k+1}) = \frac{1}{\Gamma(2-\gamma)} \sum_{j=0}^k a_j \frac{u(\cdot, t_{k+1-j}) - u(\cdot, t_{k-j})}{\tau^\gamma} + \mathbf{r}_\gamma^{k+1}(\cdot), \quad (5.8)$$

for $0 \leq k \leq \mathcal{K} - 1$, where

$$a_j = (j+1)^{1-\gamma} - j^{1-\gamma}, \quad \mathbf{r}_\gamma^{k+1} = \frac{1}{\Gamma(1-\gamma)} \sum_{j=0}^k \int_{I_j} \frac{1}{(t_{k+1}-t)^\gamma} R(\cdot, t) dt, \quad (5.9)$$

and

$$R(\cdot, t) = \partial_t u(\cdot, t) - \frac{1}{\tau} (u(\cdot, t_{j+1}) - u(\cdot, t_j)) \quad \forall t \in I_j. \quad (5.10)$$

Notice that from (5.9) we deduce that $a_j > 0$ for all $j \geq 0$ and $1 = a_0 > a_1 > a_2 > \dots > a_j$, $\lim_{j \rightarrow \infty} a_j = 0$.

5.3.3. Consistency. We estimate \mathbf{r}_γ^τ using a cancellation property: R , defined in (5.10), has vanishing mean in I_j whence we can write

$$\mathbf{r}_\gamma^{k+1} = \frac{1}{\Gamma(1-\gamma)} \sum_{j=0}^k \int_{I_j} (\psi_\gamma(t) - \bar{\psi}_\gamma^j) R(\cdot, t) dt,$$

with $\psi_\gamma(t) = (t_{k+1} - t)^{-\gamma}$ and $\bar{\psi}_\gamma^j = \int_{I_j} \psi_\gamma(t) dt$. Applying [62, Lemma 2.1] yields

$$\|r_\gamma^\tau\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))} \lesssim \|\psi_\gamma - \bar{\psi}_\gamma^\tau\|_{L^1(0,T)} \|R^\tau\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))},$$

which reduces the estimation of the residual to deriving suitable bounds for each term on the right hand side of this expression. The regularity results of Theorem 5.2, combined with the special structure of the kernel ψ_γ , yield an estimate for r_γ^τ . This estimate, although giving lower rates of convergence than [52, (3.4)], takes into account the correct behavior of the solution and the singularity of its derivatives as $t \downarrow 0$ [62, Proposition 3.6]:

$$\|r_\gamma^\tau\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))} \lesssim \tau^\theta \mathcal{A}(u_0, f) \quad 0 < \theta < 1/2.$$

The hidden constant is independent of the data and τ but blows up as $\theta \uparrow \frac{1}{2}$.

5.3.4. Stability and energy estimates. We apply (5.8). If $\gamma \in (0, 1)$ and $\phi^\tau \in L^2(\Omega)$ we define the discrete fractional derivative, for $k = 0, \dots, \mathcal{K} - 1$, by

$$\Gamma(2 - \gamma)\delta^\gamma \phi^{k+1} := \sum_{j=0}^k \frac{a_j}{\tau^{\gamma-1}} \delta^1 \phi^{k+1-j} = \frac{\phi^{k+1}}{\tau^\gamma} - \sum_{j=0}^{k-1} \frac{a_j - a_{j+1}}{\tau^\gamma} \phi^{k-j} - \frac{a_k}{\tau^\gamma} \phi^0.$$

The implicit semi-discrete scheme to solve (5.1) reads: Set V^0 as in (5.7) and, for $k = 0, \dots, \mathcal{K} - 1$, $V^{k+1} \in \hat{H}_L^1(y^\alpha, \mathcal{C})$ satisfies, for every $W \in \hat{H}_L^1(y^\alpha, \mathcal{C})$,

$$(\delta^\gamma \operatorname{tr}_\Omega V^{k+1}, \operatorname{tr}_\Omega W)_{L^2(\Omega)} + \frac{1}{d_s} \int_{\mathcal{C}} y^\alpha \nabla V^{k+1} \nabla W = \langle f^{k+1}, W \rangle. \quad (5.11)$$

Let $I^\sigma w$ be the Riemann-Liouville integral of order σ of w evaluated at $t = T$.

Theorem 5.3 (stability for $\gamma \in (0, 1)$). *The implicit semi-discrete scheme (5.11) is unconditionally stable and satisfies*

$$I^{1-\gamma} \|\operatorname{tr}_\Omega V^\tau\|_{L^2(\Omega)}^2 + \|V^\tau\|_{\ell^2(\hat{H}_L^1(y^\alpha, \mathcal{C}))}^2 \leq I^{1-\gamma} \|u_0\|_{L^2(\Omega)}^2 + \|f^\tau\|_{\ell^2(\mathbb{H}^{-s}(\Omega)')}^2 \quad (5.12)$$

Deducing an energy estimate for problem (5.1) is cumbersome due to the non-locality of the fractional time derivative. A key ingredient in deriving such a result is an integration by parts formula, which for a function not vanishing at $t = 0$ and $t = T$ involves boundary terms that need to be estimated; for a step in this direction see [37, 51]. The energy estimate (5.12) yields:

Corollary 5.4 (energy estimate for u). *Let $\gamma \in (0, 1)$. Then,*

$$I^{1-\gamma} \|u\|_{L^2(\Omega)}^2 + \|u\|_{L^2(0,T;\mathbb{H}^s(\Omega))}^2 \leq I^{1-\gamma} \|u_0\|_{L^2(\Omega)}^2 + \|f\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))}^2. \quad (5.13)$$

Remark 5.5 (limiting case). Given $g \in L^p(0, T)$, we have $I^\sigma g \rightarrow g$ in $L^p(0, T)$ as $\sigma \downarrow 0$; see [70, Theorem 2.6]. This implies that, taking the limit as $\gamma \uparrow 1$ in (5.13), we recover the well known stability result for a parabolic equation, i.e.,

$$\|u\|_{L^\infty(0,T;L^2(\Omega))}^2 + \|u\|_{L^2(0,T;\mathbb{H}^s(\Omega))}^2 \leq \|u_0\|_{L^2(\Omega)}^2 + \|f\|_{L^2(0,T;\mathbb{H}^{-s}(\Omega))}^2. \quad (5.14)$$

This allows us to unify the estimate of Corollary 5.4 for all $\gamma \in (0, 1]$.

5.4. Truncation. We now study the space discretization of (5.4). Since $\mathcal{U}(t)$ decays exponentially in the extended direction y , for a.e. $t \in (0, T)$, we truncate \mathcal{C} to $\mathcal{C}_\mathcal{Y}$ for a suitable \mathcal{Y} and seek solutions in $\mathcal{C}_\mathcal{Y}$. Define, for $\gamma \in (0, 1]$,

$$\Lambda_\gamma^2(u_0, f) := I^{1-\gamma} \|u_0\|_{L^2(\Omega)}^2 + \|f\|_{L^2(0, T; \mathbb{H}^{-s}(\Omega))}^2, \quad (5.15)$$

where, by Remark 5.5, I^0 is the identity. For $\gamma \in (0, 1]$ and $s \in (0, 1)$, we have

$$\|\nabla \mathcal{U}\|_{L^2(0, T; L^2(y^\alpha, \Omega \times (\mathcal{Y}, \infty)))} \lesssim e^{-\sqrt{\lambda_1} \mathcal{Y}/2} \Lambda_\gamma(u_0, f), \quad (5.16)$$

where $\mathcal{Y} > 1$ and \mathcal{U} denotes the solution to (5.4).

As a consequence of (5.16), we can consider the truncated problem

$$\begin{cases} \operatorname{div}(y^\alpha \nabla v) = 0 \text{ in } \mathcal{C}_\mathcal{Y} \times (0, T), & v = 0 \text{ on } (\partial_L \mathcal{C}_\mathcal{Y} \cup \Omega_\mathcal{Y}) \times (0, T) \\ d_s \partial_t^\gamma \operatorname{tr}_\Omega v + \partial_\nu^\alpha v = d_s f \text{ on } (\Omega \times \{0\}) \times (0, T), \end{cases} \quad (5.17)$$

with the initial condition $v = u_0$ on $\Omega \times \{0\}$ and $t = 0$, and where $\Omega_\mathcal{Y} = \Omega \times \{\mathcal{Y}\}$ with $\mathcal{Y} \geq 1$ sufficiently large. Upon defining

$$\mathcal{V}_\mathcal{Y} = \left\{ w \in L^2(0, T; \mathring{H}_L^1(y^\alpha, \mathcal{C}_\mathcal{Y})) : \partial_t^\gamma \operatorname{tr}_\Omega w \in L^2(0, T; \mathbb{H}^{-s}(\Omega)) \right\},$$

we understand (5.17) as: seek $v \in \mathcal{V}_\mathcal{Y}$ such that $\operatorname{tr}_\Omega v(0) = u_0$ and for a.e. $t \in (0, T)$,

$$\langle \partial_t^\gamma \operatorname{tr}_\Omega v, \operatorname{tr}_\Omega \phi \rangle + \frac{1}{d_s} \int_{\mathcal{C}_\mathcal{Y}} y^\alpha \nabla v \nabla \phi = \langle f, \operatorname{tr}_\Omega \phi \rangle, \quad \forall \phi \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_\mathcal{Y}). \quad (5.18)$$

If \mathcal{U} solves (5.4), v solves (5.18) and $\Lambda_\gamma(u_0, f)$ is given in (5.15), then we have the following exponential convergence result for every $\gamma \in (0, 1]$ and $\mathcal{Y} \geq 1$:

$$I^{1-\gamma} \|\operatorname{tr}_\Omega(\mathcal{U} - v)\|_{L^2(\Omega)}^2 + \|\nabla(\mathcal{U} - v)\|_{L^2(0, T; L^2(y^\alpha, \mathcal{C}_\mathcal{Y}))}^2 \lesssim e^{-\sqrt{\lambda_1} \mathcal{Y}} \Lambda_\gamma^2(u_0, f). \quad (5.19)$$

5.5. Fully discrete scheme. We describe a fully discrete scheme to solve (5.18). The space discretization hinges on the FEM discussed in Section 2. The discretization in time uses the schemes proposed in §5.3.1 for $\gamma = 1$ and in §5.3.2 for $\gamma \in (0, 1)$.

The fully discrete scheme computes the sequence $V_{\mathcal{T}_\mathcal{Y}}^\tau \subset \mathbb{V}(\mathcal{T}_\mathcal{Y})$, an approximation of the solution to (5.18) at each time step. We first initialize the scheme:

$$V_{\mathcal{T}_\mathcal{Y}}^0 = \mathcal{I}_{\mathcal{T}_\Omega} u_0, \quad (5.20)$$

where $\mathcal{I}_{\mathcal{T}_\Omega} = G_{\mathcal{T}_\mathcal{Y}} \circ \mathcal{H}_\alpha$, \mathcal{H}_α denotes the α -harmonic extension onto $\mathcal{C}_\mathcal{Y}$ and $G_{\mathcal{T}_\mathcal{Y}}$ the weighted elliptic projection of [62, §4.3]; notice that $\operatorname{tr}_\Omega V_{\mathcal{T}_\mathcal{Y}}^0 = \operatorname{tr}_\Omega G_{\mathcal{T}_\mathcal{Y}} v(0)$. For $k = 0, \dots, \mathcal{K} - 1$, let $V_{\mathcal{T}_\mathcal{Y}}^{k+1} \in \mathbb{V}(\mathcal{T}_\mathcal{Y})$ satisfy for all $W \in \mathbb{V}(\mathcal{T}_\mathcal{Y})$

$$(\delta^\gamma \operatorname{tr}_\Omega V_{\mathcal{T}_\mathcal{Y}}^{k+1}, \operatorname{tr}_\Omega W)_{L^2(\Omega)} + \frac{1}{d_s} \int_{\mathcal{C}_\mathcal{Y}} y^\alpha \nabla V_{\mathcal{T}_\mathcal{Y}}^{k+1} \nabla W = \langle f^{k+1}, \operatorname{tr}_\Omega W \rangle. \quad (5.21)$$

An approximate solution to problem (5.1) is given by $U_{\mathcal{F}_\Omega}^\tau = \text{tr}_\Omega V_{\mathcal{F}_\Omega}^\tau$.

The discrete scheme (5.20)–(5.21) is unconditionally stable for all $\gamma \in (0, 1]$:

$$I^{1-\gamma} \|\text{tr}_\Omega V_{\mathcal{F}_\Omega}^\tau\|_{L^2(\Omega)}^2 + \|V_{\mathcal{F}_\Omega}^\tau\|_{\ell^2(\hat{H}_L^1(y^\alpha, \mathcal{C}_y))}^2 \lesssim I^{1-\gamma} \|\text{tr}_\Omega V_{\mathcal{F}_\Omega}^0\|_{L^2(\Omega)}^2 + \|f^\tau\|_{\ell^2(\mathbb{H}^{-s}(\Omega))}^2,$$

where I^0 is the identity according to Remark 5.5 (case $\gamma = 1$).

We define $\mathcal{B}(u_0, f) := \|u_0\|_{\mathbb{H}^{1+3s}(\Omega)} + \|f|_{t=0}\|_{\mathbb{H}^{1+s}(\Omega)} + \|f\|_{W_\infty^1(0, T; \mathbb{H}^{1-(1-2\mu)s}(\Omega))}$ for $\mu > 0$. The error estimates for (5.20)–(5.21) read as follows.

Theorem 5.6 (error estimates for $\gamma \in (0, 1)$). *Let $\gamma \in (0, 1)$, v and $V_{\mathcal{F}_\Omega}^\tau$ solve (5.18) and (5.20)–(5.21), respectively. If $\mathcal{A}(u_0, f), \mathcal{B}(u_0, f) < \infty$ and \mathcal{F}_Ω satisfies (2.12), then*

$$[I^{1-\gamma} \|\text{tr}_\Omega(v^\tau - V_{\mathcal{F}_\Omega}^\tau)\|_{L^2(\Omega)}^2]^{\frac{1}{2}} \lesssim \tau^\theta \mathcal{A}(u_0, f) + |\log N|^{2s} N^{-\frac{(1+s)}{n+1}} \mathcal{B}(u_0, f),$$

and

$$\|v^\tau - V_{\mathcal{F}_\Omega}^\tau\|_{\ell^2(\hat{H}_L^1(y^\alpha, \mathcal{C}_y))} \lesssim \tau^\theta \mathcal{A}(u_0, f) + |\log N|^s N^{-\frac{1}{n+1}} \mathcal{B}(u_0, f),$$

where $0 < \theta < \frac{1}{2}$, and \mathcal{A} is defined in (5.6).

We finally observe that exploiting the regularity (5.5), we can derive optimal (linear) error estimates in time for $\gamma = 1$.

References

- [1] S. Abe and S. Thurner. Anomalous diffusion in view of Einstein’s 1905 theory of Brownian motion. *Physica A: Statistical Mechanics and its Applications*, 356(24):403–407, 2005.
- [2] H. Antil and E. Otárola. A FEM for an optimal control problem of fractional powers of elliptic operators. Submitted to SIAM J. Control Optim. arXiv:1406.7460, 2014.
- [3] H. Antil, E. Otárola, and A.J. Salgado. A fractional space-time optimal control problem: analysis and discretization. Submitted to SIAM J. Control Optim. arXiv:1504.00063, 2015.
- [4] T. Apel and J. Schöberl. Multigrid methods for anisotropic edge refinement. *SIAM J. Numer. Anal.*, 40(5):1993–2006, 2002.
- [5] T.M. Atanackovic, S. Pilipovic, B. Stankovic, and D. Zorica. *Fractional Calculus with Applications in Mechanics: Vibrations and Diffusion Processes*. 2014.
- [6] I. Babuška and A. Miller. A feedback finite element method with a posteriori error estimation. I. The finite element method and some basic properties of the a posteriori error estimator. *Comput. Methods Appl. Mech. Engrg.*, 61(1):1–40, 1987.
- [7] I. Babuška and W.C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15(4):736–754, 1978.
- [8] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II—differential equations analysis library. *Technical Reference: <http://dealii.org>*.

- [9] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II—a general-purpose object-oriented finite element library. *ACM Trans. Math. Software*, 33(4):Art. 24, 27, 2007.
- [10] M.Š. Birman and M.Z. Solomjak. *Spektralnaya teoriya samosopryazhennykh operatorov v gilbertovom prostranstve*. Leningrad. Univ., Leningrad, 1980.
- [11] M. Bonforte, Y. Sire, and J.L. Vázquez. Existence, uniqueness and asymptotic behaviour for fractional porous medium equations on bounded domains. arXiv:1311.5121, 2013.
- [12] A. Bonito and J. E. Pasciak. Numerical approximation of fractional powers of elliptic operators. *Math. Comp.*, 84(295):2083–2110, 2015.
- [13] J.H. Bramble and J.E. Pasciak. New convergence estimates for multigrid algorithms. *Math. Comp.*, 49(180):311–329, 1987.
- [14] J.H. Bramble, J.E. Pasciak, J.P. Wang, and J. Xu. Convergence estimates for multigrid algorithms without regularity assumptions. *Math. Comp.*, 57(195):23–45, 1991.
- [15] J.H. Bramble, J.E. Pasciak, J.P. Wang, and J. Xu. Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comp.*, 57(195):1–21, 1991.
- [16] J.H. Bramble and X. Zhang. The analysis of multigrid methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, pages 173–415. North-Holland, Amsterdam, 2000.
- [17] J.H. Bramble and X. Zhang. Uniform convergence of the multigrid V -cycle for an anisotropic problem. *Math. Comp.*, 70(234):453–470, 2001.
- [18] C. Brändle, E. Colorado, A. de Pablo, and U. Sánchez. A concave–convex elliptic problem involving the fractional laplacian. *Proceedings of the Royal Society of Edinburgh, Section: A Mathematics*, 143:39–71, 2013.
- [19] A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Math. Comp.*, 31:333–390, 1977.
- [20] A. Brandt. *Multigrid techniques: 1984 guide with applications to fluid dynamics*. Ges. für Mathematik u. Datenverarbeitung, 1984.
- [21] S.C. Brenner and L.R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.
- [22] A. Bueno-Orovio, D. Kay, V. Grau, B. Rodriguez, and K. Burrage. Fractional diffusion models of cardiac electrical propagation: role of structural heterogeneity in dispersion of repolarization. *J. R. Soc. Interface*, 11(97), 2014.
- [23] X. Cabré and Y. Sire. Nonlinear equations for fractional Laplacians II: Existence, uniqueness, and qualitative properties of solutions. *Trans. Amer. Math. Soc.*, 367(2):911–941, 2015.
- [24] X. Cabré and J. Tan. Positive solutions of nonlinear problems involving the square root of the Laplacian. *Adv. Math.*, 224(5):2052–2093, 2010.
- [25] L. Caffarelli and L. Silvestre. An extension problem related to the fractional Laplacian. *Comm. Part. Diff. Eqs.*, 32(7-9):1245–1260, 2007.
- [26] A. Capella, J. Dávila, L. Dupaigne, and Y. Sire. Regularity of radial extremal solutions for some non-local semilinear equations. *Comm. Part. Diff. Eqs.*, 36(8):1353–1384, 2011.

- [27] J.M. Cascón and R.H. Nochetto. Quasioptimal cardinality of AFEM driven by nonresidual estimators. *IMA J. Numer. Anal.*, 32(1):1–29, 2012.
- [28] L. Chen. *iFEM: An integrated finite element methods package in matlab*. Technical report, University of California at Irvine, 2009.
- [29] L. Chen, R.H. Nochetto, E. Otárola, and A.J. Salgado. Multilevel methods for nonuniformly elliptic operators. Submitted to *Math. Comp.* arXiv:1403.4278, 2014.
- [30] L. Chen, R.H. Nochetto, E. Otárola, and A.J. Salgado. A PDE approach to fractional diffusion: A posteriori error analysis. *J. Comput. Phys.*, 293:339–358, 2015.
- [31] W. Chen. A speculative study of 2/3-order fractional laplacian modeling of turbulence: Some thoughts and conjectures. *Chaos*, 16(2):1–11, 2006.
- [32] D. del Castillo-Negrete, B. A. Carreras, and V. E. Lynch. Fractional diffusion in plasma turbulence. *Physics of Plasmas*, 11(8):3854–3864, 2004.
- [33] W. Dörfler. A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.
- [34] Q. Du, M. Gunzburger, R. B. Lehoucq, and K. Zhou. Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Rev.*, 54(4):667–696, 2012.
- [35] J. Duoandikoetxea. *Fourier analysis*. American Mathematical Society, 2001.
- [36] R.G. Durán and A.L. Lombardi. Error estimates on anisotropic Q_1 elements for functions in weighted Sobolev spaces. *Math. Comp.*, 74(252):1679–1706, 2005.
- [37] V.J. Ervin and J.P. Roop. Variational formulation for the stationary fractional advection dispersion equation. *Numer. Methods Partial Differential Equations*, 22(3):558–576, 2006.
- [38] E. B. Fabes, C.E. Kenig, and R.P. Serapioni. The local regularity of solutions of degenerate elliptic equations. *Comm. Part. Diff. Eqs.*, 7(1):77–116, 1982.
- [39] P. Gatto and J. Hesthaven. Numerical approximation of the fractional laplacian via hp-finite elements, with an application to image denoising. *J. Sci. Comp.*, pages 1–22, 2014.
- [40] V. Gol’dshstein and A. Ukhlov. Weighted Sobolev spaces and embedding theorems. *Trans. Amer. Math. Soc.*, 361(7):3829–3850, 2009.
- [41] R. Gorenflo, F. Mainardi, D. Moretti, and P. Paradisi. Time fractional diffusion: a discrete random walk approach. *Nonlinear Dynam.*, 29(1-4):129–143, 2002. Fractional order calculus and its applications.
- [42] P. Grisvard. *Elliptic problems in nonsmooth domains*. Pitman, 1985.
- [43] W. Hackbusch. *Multigrid methods and applications*. Springer-Verlag, 1985.
- [44] W. Hackbusch. The frequency decomposition multi-grid method. I. Application to anisotropic equations. *Numer. Math.*, 56(2-3):229–245, 1989.
- [45] Y. Huang and A. Oberman. Numerical methods for the fractional Laplacian: A finite difference-quadrature approach. *SIAM J. Numer. Anal.*, 52(6):3056–3084, 2014.
- [46] R. Ishizuka, S.-H. Chong, and F. Hirata. An integral equation theory for inhomogeneous molecular fluids: The reference interaction site model approach. *J. Chem. Phys.*, 128(3), 2008.

- [47] A.A. Kilbas, H.M. Srivastava, and J.J. Trujillo. *Theory and applications of fractional differential equations*. Elsevier Science B.V., Amsterdam, 2006.
- [48] A. Kufner and B. Opic. How to define reasonably weighted Sobolev spaces. *Comment. Math. Univ. Carolin.*, 25(3):537–554, 1984.
- [49] N.S. Landkof. *Foundations of modern potential theory*. Springer-Verlag, 1972.
- [50] S. Z. Levendorskiĭ. Pricing of the American put under Lévy processes. *Int. J. Theor. Appl. Finance*, 7(3):303–335, 2004.
- [51] X. Li and C. Xu. A space-time spectral method for the time fractional diffusion equation. *SIAM J. Numer. Anal.*, 47(3):2108–2131, 2009.
- [52] Y. Lin, X. Li, and C. Xu. Finite difference/spectral approximations for the fractional cable equation. *Math. Comp.*, 80(275):1369–1396, 2011.
- [53] Y. Lin and C. Xu. Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys.*, 225(2):1533–1552, 2007.
- [54] R. Lipton. Dynamic brittle fracture as a small horizon limit of peridynamics. *J. Elasticity*, 117(1):21–50, 2014.
- [55] W. McLean. Regularity of solutions to a time-fractional diffusion equation. *ANZIAM J.*, 52(2):123–138, 2010.
- [56] P. Morin, R.H. Nochetto, and K.G. Siebert. Local problems on stars: a posteriori error estimators, convergence, and performance. *Math. Comp.*, 72(243):1067–1097, 2003.
- [57] B. Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.*, 165:207–226, 1972.
- [58] K. Mustapha and W. McLean. Superconvergence of a discontinuous Galerkin method for fractional diffusion and wave equations. *SIAM J. Numer. Anal.*, 51(1):491–515, 2013.
- [59] R. R. Nigmatullin. The realization of the generalized transfer equation in a medium with fractal geometry. *Physica Status Solidi (b)*, 133(1):425–430, 1986.
- [60] R. H. Nochetto, E. Otárola, and A. J. Salgado. A PDE approach to fractional diffusion in general domains: A priori error analysis. *Found. Comput. Math.*, 15(3):733–791, 2015.
- [61] R.H. Nochetto, E. Otárola, and A.J. Salgado. Convergence rates for the classical, thin and fractional elliptic obstacle problems. *Philos. Transact. A Math. Phys. Eng. Sci.*, 2014. accepted for publication.
- [62] R.H. Nochetto, E. Otárola, and A.J. Salgado. A PDE approach to space-time fractional parabolic problems. Submitted to *SIAM J. Numer. Anal.* arXiv:1404.0068, 2014.
- [63] R.H. Nochetto, E. Otárola, and A.J. Salgado. Piecewise polynomial interpolation in Muckenhoupt weighted Sobolev spaces and applications. *Numer. Math.*, 2015. DOI: 10.1007/s00211-015-0709-6.
- [64] R.H. Nochetto, K.G. Siebert, and A. Veiser. Theory of adaptive finite element methods: an introduction. In *Multiscale, nonlinear and adaptive approximation*, pages 409–542. Springer, Berlin, 2009.
- [65] R.H. Nochetto and A. Vesser. Primer of adaptive finite element methods. In *Multiscale and Adaptivity: Modeling, Numerics and Applications, CIME Lectures*. Springer, 2011.

- [66] R.H. Nochetto, T. von Petersdorff, and C.-S. Zhang. A posteriori error analysis for a class of integral equations and variational inequalities. *Numer. Math.*, 116(3):519–552, 2010.
- [67] E. Otárola. *A PDE approach to numerical fractional diffusion*. PhD thesis, University of Maryland, College Park, May 2014.
- [68] I. Podlubny. Fractional-order systems and $PI^\lambda D^\mu$ -controllers. *IEEE Trans. Automat. Control*, 44(1):208–214, 1999.
- [69] A.I. Saichev and G.M. Zaslavsky. Fractional kinetic equations: solutions and applications. *Chaos*, 7(4):753–764, 1997.
- [70] S.G. Samko, A.A. Kilbas, and O.I. Marichev. *Fractional integrals and derivatives*. Gordon and Breach Science Publishers, Yverdon, 1993.
- [71] R. Stevenson. Robustness of multi-grid applied to anisotropic equations on convex domains and on domains with re-entrant corners. *Numer. Math.*, 66(3):373–398, 1993.
- [72] P.R. Stinga and J.L. Torrea. Extension problem and Harnack’s inequality for some fractional operators. *Comm. Part. Diff. Eqs.*, 35(11):2092–2122, 2010.
- [73] B.O. Turesson. *Nonlinear potential theory and weighted Sobolev spaces*. Springer-Verlag, 2000.
- [74] Y. Wu, L. Chen, X. Xie, and J. Xu. Convergence analysis of V-cycle multigrid methods for anisotropic elliptic equations. *IMA J. Numer. Anal.*, 32(4):1329–1347, 2012.
- [75] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, pages 581–613, 1992.
- [76] J. Xu, L. Chen, and R.H. Nochetto. Optimal multilevel methods for $H(\text{grad})$, $H(\text{curl})$, and $H(\text{div})$ systems on graded and unstructured grids. In *Multiscale, nonlinear and adaptive approximation*, pages 599–659. Springer, Berlin, 2009.
- [77] J. Xu and L. Zikatanov. The method of alternating projections and the method of subspace corrections in Hilbert spaces. *Math. Comp.*, 15:573–597, 2002.

Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA.

E-mail: rhn@math.umd.edu

Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile.

E-mail: enriqueotarola@gmail.com

Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA.

E-mail: asalgad1@utk.edu