Tensor FEM for spectral fractional diffusion

L. Banjai and J. Melenk and R. Nochetto and E. Otarola and A. Salgado

and Ch. Schwab

Research Report No. 2017-36

July 2017
Latest revision: August 2018

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Funding: NSF grants DMS-1109325 and DMS-1411808, Austrian Science Fund (FWF) project F 65

TENSOR FEM FOR SPECTRAL FRACTIONAL DIFFUSION*

LEHEL BANJAI†, JENS M. MELENK‡, RICARDO H. NOCHETTO§, ENRIQUE OTÁROLA¶, ABNER J. SALGADO‖, AND CHRISTOPH SCHWAB**

Abstract. We design and analyze several Finite Element Methods (FEMs) applied to the Caffarelli-Silvestre extension that localizes the fractional powers of symmetric, coercive, linear elliptic operators in bounded domains with Dirichlet boundary conditions. We consider open, bounded, polytopal but not necessarily convex domains $\Omega \subset \mathbb{R}^d$ with $d = 1, 2$. For the solution to the Caffarelli-Silvestre extension, we establish analytic regularity with respect to the extended variable $y \in (0, \infty)$. Specifically, the solution belongs to countably normed, power–exponentially weighted Bochner spaces of analytic functions with respect to y, taking values in corner-weighted Kondat'ev type Sobolev spaces in Ω. In $\Omega \subset \mathbb{R}^2$, we discretize with continuous, piecewise linear, Lagrangian FEM (P_1-FEM) with mesh refinement near corners, and prove that first order convergence rate is attained for compatible data $f \in \mathbb{H}^{1-s}(\Omega)$ with $0 < s < 1$ denoting the fractional power.

We also prove that tensorization of a P_1-FEM in Ω with a suitable *hp*-FEM in the extended variable achieves log-linear complexity with respect to $\mathcal{N}_\Omega$, the number of degrees of freedom in the domain Ω. In addition, we propose a novel, *sparse tensor product FEM* based on a multilevel P_1-FEM in Ω and on a P_1-FEM on radical–geometric meshes in the extended variable. We prove that this approach also achieves log-linear complexity with respect to $\mathcal{N}_\Omega$. Finally, under the stronger assumption that the data be analytic in $\overline{\Omega}$, and *without compatibility at* $\partial\Omega$, we establish *exponential rates of convergence of hp-FEM for spectral fractional diffusion operators* in energy norm. This is achieved by a combined tensor product *hp*-FEM for the Caffarelli-Silvestre extension in the truncated cylinder $\Omega \times (0, \mathcal{Y})$ with anisotropic geometric meshes that are refined towards $\partial\Omega$. We also report numerical experiments for model problems which confirm the theoretical results. We indicate several extensions and generalizations of the proposed methods to other problem classes and to other boundary conditions on $\partial\Omega$.

Key words. Fractional diffusion, nonlocal operators, weighted Sobolev spaces, regularity estimates, finite elements, anisotropic *hp*–refinement, corner refinement, sparse grids, exponential convergence.

AMS subject classifications. 26A33, 65N12, 65N30.

1. Introduction.

We are interested in the design and analysis of a variety of efficient numerical techniques to solve problems involving certain fractional powers of the linear, elliptic, self-adjoint, second order, differential operator $\mathcal{L}w = -\mathrm{div}(A\nabla w) + cw$, supplemented with homogeneous Dirichlet boundary conditions. The coefficient $A \in$

*The results in this paper were obtained when the authors met at the MFO Oberwolfach during the WS 1711 in March 2017. The research of RHN was supported in part by NSF grants DMS-1109325 and DMS-1411808. The research of EO was supported in part by CONICYT through project FONDECYT 3160201. The research of AJS was supported in part by NSF grant DMS-1418784. The research of JMM was supported by the Austrian Science Fund (FWF) project F 65. The research of CS was supported by the Swiss National Science Foundation (SNSF) under grant No. 159940.

†Maxwell Institute for Mathematical Sciences, School of Mathematical & Computer Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK (l.banjai@hw.ac.uk).

‡Institut für Analysis und Scientific Computing, Technische Universität Wien, A-1040 Vienna, Austria (melenk@tuwien.ac.at).

§Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA (rhn@math.umd.edu).

¶Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile (enrique.otarola@usm.cl).

‖Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA (asalgad1@utk.edu).

**Seminar for Applied Mathematics, ETH Zürich, ETH Zentrum, HG G57.1, CH8092 Zürich, Switzerland (christoph.schwab@sam.math.ethz.ch).

$L^\infty(\Omega, \mathsf{GL}(\mathbb{R}^d))$ is symmetric and uniformly positive definite and $0 \le c \in L^\infty(\Omega, \mathbb{R})$ (additional regularity requirements will be imposed in the course of our convergence rate analysis ahead). We denote by Ω a bounded domain of $\mathbb{R}^d$ ($d = 1, 2$), with Lipschitz boundary $\partial\Omega$ and further properties imposed as required: the FEM convergence theory in Section 5 will focus on polygonal domains $\Omega \subset \mathbb{R}^2$, the hp-FEM results in Section 7 require analytic $\partial\Omega$.

The Dirichlet problem for the fractional Laplacian is as follows: Given a function f and $s \in (0, 1)$, we seek u such that

$$\mathcal{L}^s u = f \quad \text{in } \Omega \,. \tag{1.1}$$

An essential difficulty in the analysis of (1.1) and in the design of efficient numerical methods for this problem is that $\mathcal{L}^s$ is a nonlocal operator [15, 16, 17, 19, 36]. In the case of the Dirichlet Laplacian $\mathcal{L} = -\Delta$, Caffarelli and Silvestre in [17] localize it by using a nonuniformly elliptic PDE posed in one more spatial dimension. They showed that any power $s \in (0, 1)$ of the fractional Laplacian in $\mathbb{R}^d$ can be realized as the Dirichlet-to-Neumann map of an extension to the upper half-space $\mathbb{R}^{d+1}_+$. This result was extended by Cabré and Tan [16] and by Stinga and Torrea [63] to bounded domains Ω and more general operators, thereby obtaining an extension posed on the semi–infinite cylinder $\mathcal{C} := \Omega \times (0, \infty)$; we also refer to [19]. This extension is the following local boundary value problem

$$\begin{cases} \mathfrak{L}\mathscr{U} = -\mathrm{div}\,(y^\alpha \boldsymbol{A} \nabla \mathscr{U}) + c y^\alpha \mathscr{U} = 0 & \text{in } \mathcal{C}, \\ \mathscr{U} = 0 & \text{on } \partial_L \mathcal{C}, \\ \partial_{\nu^\alpha} \mathscr{U} = d_s f & \text{on } \Omega \times \{0\}, \end{cases} \tag{1.2}$$

where $\boldsymbol{A} = \mathrm{diag}\{A, 1\} \in L^\infty(\bar{\mathcal{C}}, \mathsf{GL}(\mathbb{R}^{d+1}))$, $\partial_L \mathcal{C} := \partial\Omega \times (0, \infty)$ signifies the lateral boundary of $\mathcal{C}$, $d_s := 2^{1-2s}\Gamma(1-s)/\Gamma(s)$ is a positive normalization constant and the parameter α is defined as $\alpha = 1 - 2s \in (-1, 1)$ [17, 63]. The so–called conormal exterior derivative of $\mathscr{U}$ at $\Omega \times \{0\}$ is

$$\partial_{\nu^\alpha} \mathscr{U} = -\lim_{y \to 0^+} y^\alpha \mathscr{U}_y. \tag{1.3}$$

We shall refer to y as the *extended variable* and to the dimension $d + 1$ in $\mathbb{R}^{d+1}_+$ the *extended dimension* of problem (1.2). Throughout the text, points $x \in \mathcal{C}$ will be written as $x = (x', y)$ with $x' \in \Omega$ and $y > 0$. The limit in (1.3) must be understood in the distributional sense [16, 17, 63]. With the extension $\mathscr{U}$ at hand, the fractional powers of $\mathcal{L}$ in (1.1) and the Dirichlet-to-Neumann operator of problem (1.2) are related by

$$d_s \mathcal{L}^s u = \partial_{\nu^\alpha} \mathscr{U} \quad \text{in } \Omega \,. \tag{1.4}$$

In [48] the extension problem (1.2) was first used as a way to obtain a numerical technique to approximate the solution to (1.1). A piecewise linear finite element method (P_1-FEM) was proposed and analyzed. In this work, we extend the results of [48] in several directions:

a) In Theorem 5.10, we generalize the error analysis of [48], based on the localization of $\mathcal{L}^s$ given by (1.2), to nonconvex polygonal domains $\Omega \subset \mathbb{R}^2$, under the requirement of Lipschitz regularity in Ω for A and c, and for $f \in \mathbb{H}^{1-s}(\Omega)$ (see (2.2) ahead).

b) In Theorem 4.7 we prove, again under Lipschitz regularity in Ω for A and c, *weighted H^2 (with respect to the extended variable y) regularity estimates* for the solution $\mathscr{U}$ of (1.2). We use these to propose a novel, sparse tensor product P_1-FEM in $\mathcal{C}$ which is realized by invoking (in parallel) $\mathcal{O}(\log \mathcal{N}_\Omega)$ many instances of anisotropic tensor product P_1-FEM in $\mathcal{C}$. We prove, in Theorem 5.13, that, when the base of the cylinder $\mathcal{C}$ is a polygonal domain $\Omega \subset \mathbb{R}^2$, this approach yields a method with $\mathcal{O}(\mathcal{N}_\Omega \log \mathcal{N}_\Omega)$ degrees of freedom realizing the (optimal) asymptotic convergence rate of $\mathcal{N}_\Omega^{-1/2}$ for $f \in \mathbb{H}^{1-s}(\Omega)$.

c) We show, in Theorem 5.16, that a full tensor product approach of an *hp*-FEM in the extended variable y with P_1-FEM in Ω yields the same rate. To achieve this, we establish *weighted analytic regularity of $\mathscr{U}$ with respect to the extended variable y*, in terms of countably normed weighted Bochner-Sobolev spaces. This extends, in the case $d = 2$, recent work [40] to a general diffusion operator $\mathcal{L}$ in (1.1) and to nonconvex, polygonal domains, under the requirement of Lipschitz regularity in Ω for A and c.

d) We propose in Section 6 a novel *diagonalization technique* which decouples the degrees of freedom introduced by a Galerkin (semi-)discretization in the extended variable. It reduces the y-semidiscrete Caffarelli-Stinga extension to the solution of independent, singularly perturbed *second order* reaction-diffusion equations in Ω. This decoupling allows us to establish exponential convergence for *analytic data f without boundary compatibility* as discussed in the following item e). The diagonalization also permits to block-diagonalize the stiffness matrix of the fully discrete problem with corresponding benefits for the solver complexity of the linear system of equations.

e) We establish an *exponential convergence* rate (7.7) of a local *hp*-FEM for the fractional differential operator $\mathfrak{L}$ in (1.2). This requires, however, the data A, c and f to be *analytic* in $\bar{\Omega}$ and the boundary $\partial\Omega$ to be analytic as well. Here, no boundary compatibility of f at $\partial\Omega$ is required. For brevity of exposition, we detail the mathematical argument in intervals $\Omega \subset \mathbb{R}^1$ and in bounded domains $\Omega \subset \mathbb{R}^2$ with analytic boundary $\partial\Omega$, and for constant coefficients A and c, and only outline the necessary extensions, with references, for polygons $\Omega \subset \mathbb{R}^2$; see Theorems 7.3, 7.7 and Remark 7.8.

f) We present numerical experiments in each of the previous cases which illustrate our results, and indicate their sharpness.

g) We indicate how the presently developed discretizations and error bounds extend in several directions, in particular to three dimensional polyhedral domains Ω, to Neumann or mixed Dirichlet-Neumann boundary conditions on $\partial\Omega$, etc.

To close the introduction, we comment on other numerical approaches to fractional PDEs. In addition to [48], numerical schemes that deal with spectral fractional powers of elliptic operators have been proposed in [40] and [13]. The very recent work [40] adopts the same Galerkin framework as [48] and the present article and, independently, proposes to use high order discretizations in the extended variable to exploit analyticity. The starting point of [13] is the so-called Balakrishnan formula, a contour integral representation of the inverse $\mathcal{L}^{-s}$. Upon discretizing the integral by a suitable quadrature formula, the numerical scheme of [13] results in a collection of (decoupled) singularly perturbed reaction diffusion problems in Ω. This connects [13] with our approach in Section 7. However, the decoupled reaction diffusion problems in Ω which arise in our approach result from a Galerkin discretization in the extended variable. For the integral definition of the fractional Laplacian in several dimensions we mention, in particular, the analysis of [2, 24]. We refer the reader to [12] for a detailed account of all the approaches mentioned above.

2. Notation and preliminaries. We adopt the notation of [48, 52]: For $\mathscr{Y} > 0$ the truncated cylinder with base Ω and height $\mathscr{Y}$ is $\mathcal{C}_{\mathscr{Y}} = \Omega \times (0, \mathscr{Y})$, its lateral boundary is $\partial_L \mathcal{C}_{\mathscr{Y}} = \partial\Omega \times (0, \mathscr{Y})$. If $x \in \mathcal{C}$, we set $x = (x', y)$ with $x' \in \Omega$ and $y \in (0, \infty)$. By $a \lesssim b$ we mean $a \leq Cb$, with a constant C that does not depend on a, b, or the discretization parameters. The notation $a \sim b$ signifies $a \lesssim b \lesssim a$. The value of C might change at each occurrence. The spaces $H^1(D)$, $H_0^1(D)$, and more generally $H^\sigma(D)$, $\sigma \in \mathbb{R}$, denote the usual Sobolev spaces (see, e.g., [3]).

2.1. Fractional powers of elliptic operators. To define $\mathcal{L}^s$, as in [48], we invoke spectral theory [11]. The operator $\mathcal{L}$ induces an inner product $a_\Omega(\cdot,\cdot)$ on $H_0^1(\Omega)$

$$a_\Omega(w, v) = \int_\Omega (A\nabla w \cdot \nabla v + cwv) \, \mathrm{d}x', \tag{2.1}$$

and $\mathcal{L}$ is an isomorphism $H_0^1(\Omega) \to H^{-1}(\Omega)$ given by $u \mapsto a_\Omega(u, \cdot)$. The eigenvalue problem: Find $(\lambda, \phi) \in \mathbb{R} \times H_0^1(\Omega) \setminus \{0\}$ such that

$$a_\Omega(\phi, v) = \lambda(\phi, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega)$$

has a countable collection of solutions $\{\lambda_k, \varphi_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^+ \times H_0^1(\Omega)$, with the real eigenvalues enumerated in increasing order, counting multiplicities, and such that $\{\varphi_k\}_{k\in\mathbb{N}}$ is an orthonormal basis of $L^2(\Omega)$ and an orthogonal basis of $(H_0^1(\Omega), a_\Omega(\cdot,\cdot))$. In terms of these eigenpairs, we introduce, for $s \geq 0$, the spaces

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

We denote by $\mathbb{H}^{-s}(\Omega)$ the dual space of $\mathbb{H}^s(\Omega)$. The duality pairing between $\mathbb{H}^s(\Omega)$ and $\mathbb{H}^{-s}(\Omega)$ will be denoted by $\langle \cdot, \cdot \rangle$. Through this duality pairing, we identify elements of $f \in \mathbb{H}^{-s}(\Omega)$ with sequences $\{f_k\}_k$ with $\sum_k f_k^2 \lambda_k^{-2s} = \|f\|^2_{\mathbb{H}^{-s}(\Omega)}$, which allows us to extend the definition of the norm in (2.2) to $s < 0$. We have the isometries $\|w\|^2_{L^2(\Omega)} = \|w\|^2_{\mathbb{H}^0}$ and $a_\Omega(w, w) = \|w\|^2_{\mathbb{H}^1}$; by (real) interpolation between $L^2(\Omega)$ and $H_0^1(\Omega)$, we infer for $s \in (0, 1)$ that $\mathbb{H}^s(\Omega) = [L^2(\Omega), H_0^1(\Omega)]_s$. We notice that, if $s \in (0, \frac{1}{2})$, then $\mathbb{H}^s(\Omega) = H^s(\Omega) = H_0^s(\Omega)$, while, for $s \in (\frac{1}{2}, 1)$, $\mathbb{H}^s(\Omega)$ can be characterized by [37, 39, 64]

$$\mathbb{H}^s(\Omega) = \{ w \in H^s(\Omega) : w = 0 \text{ on } \partial\Omega \}. \tag{2.3}$$

If $s = \frac{1}{2}$, we have that $\mathbb{H}^{\frac{1}{2}}(\Omega)$ is the so–called Lions–Magenes space $H_{00}^{\frac{1}{2}}(\Omega)$ [37, 64]. Finally, we notice that since the domain Ω is Lipshitz and the coefficients of $\mathcal{L}$ are assumed to be smooth, we have that $\mathbb{H}^s(\Omega) = H^s(\Omega) \cap H_0^1(\Omega)$ for $s \in (1, 3/2)$.

For functions $w = \sum_k w_k \varphi_k \in \mathbb{H}^1(\Omega)$, the operator $\mathcal{L} : \mathbb{H}^1(\Omega) \to \mathbb{H}^{-1}(\Omega)$ takes the form $\mathcal{L}w = \sum_k \lambda_k w_k \varphi_k$. For $s \in (0, 1)$ and $w = \sum_k w_k \varphi_k \in \mathbb{H}^s(\Omega)$, the operator $\mathcal{L}^s : \mathbb{H}^s(\Omega) \to \mathbb{H}^{-s}(\Omega)$ is defined by

$$\mathcal{L}^s w = \sum_{k=1}^\infty \lambda_k^s w_k \varphi_k. \tag{2.4}$$

In the scale of spaces $\mathbb{H}^\sigma(\Omega)$, problem (1.1) admits the following shift theorem.

LEMMA 2.1 (shift theorem). *Let $s \in (0,1)$ and $\sigma \geq 0$. If $f \in \mathbb{H}^{-s+\sigma}(\Omega)$, then the solution u of* (1.1) *satisfies $u \in \mathbb{H}^{s+\sigma}(\Omega)$ and $\|u\|_{\mathbb{H}^{s+\sigma}(\Omega)} \lesssim \|f\|_{\mathbb{H}^{-s+\sigma}(\Omega)}$.*

Proof. The assertion follows from the definition of $\mathbb{H}^s(\Omega)$ and from the fact that the solution u can be written as $u = \sum u_k \varphi_k$ with $u_k = f_k \lambda_k^{-s}$ and $k \in \mathbb{N}$. ◻

REMARK 2.2 (compatibility conditions). *Frequently, as in [48], in the present work we will impose the regularity assumption $f \in \mathbb{H}^{1-s}(\Omega)$. We notice that, for $s \in (0,1/2)$, and in view of* (2.3), *this assumption requires f to have a vanishing trace on $\partial\Omega$. If this is the case, we will thus say that f has to satisfy a* (boundary) compatibility *condition. Finally, we notice that functions $f \in C^\infty(\overline{\Omega})$ are (generically) only in $\mathbb{H}^{1/2-\delta}(\Omega)$ with $\delta > 0$.* ∎

REMARK 2.3 (nonconvexity of the domain). *Except for Proposition 3.1 below, the analysis of the numerical schemes proposed in this work does not require the convexity of Ω. A key ingredient that will allow for such an analysis on properly refined meshes in Ω is a regularity shift result in weighted Sobolev spaces in both Ω and the extended domain $(0,\infty)$; see Theorem 5.5 below.*

2.2. The extension property. Both extensions, the one by Caffarelli–Silvestre for $\Omega = \mathbb{R}^d$ [17] and that of Cabré–Tan [16] and Stinga–Torrea for Ω bounded and general elliptic operators [63] require us to deal with the *nonuniformly* (but local) linear, second order elliptic equation (1.2). Here, Lebesgue and Sobolev spaces with the weight y^α for $\alpha \in (-1,1)$ [14, 16, 17, 19] naturally arise. If $D \subset \mathbb{R}^{d+1}$, we define $L^2(y^\alpha, D)$ as the Lebesgue space for the measure $|y|^\alpha \,\mathrm{d}x$. We also define the weighted Sobolev space

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

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

$$\|w\|_{H^1(y^\alpha,D)} = \left(\|w\|^2_{L^2(y^\alpha,D)} + \|\nabla w\|^2_{L^2(y^\alpha,D)} \right)^{\frac{1}{2}}. \tag{2.5}$$

In view of the fact that $\alpha \in (-1,1)$, the weight y^α belongs to the Muckenhoupt class $A_2(\mathbb{R}^{d+1})$ [27, 28, 31, 46, 65]. This, in particular, implies that $H^1(y^\alpha, D)$ with norm (2.5) is Hilbert and $C^\infty(D) \cap H^1(y^\alpha, D)$ is dense in $H^1(y^\alpha, D)$ (cf. [65, Proposition 2.1.2, Corollary 2.1.6], [35] and [31, Theorem 1]).

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

$$\mathring{H}^1(y^\alpha, \mathcal{C}) = \left\{ w \in H^1(y^\alpha, \mathcal{C}) : w = 0 \text{ on } \partial_L \mathcal{C} \right\}. \tag{2.6}$$

In [48, inequality (2.21)] the following *weighted Poincaré inequality* is shown:

$$\|w\|_{L^2(y^\alpha,\mathcal{C})} \lesssim \|\nabla w\|_{L^2(y^\alpha,\mathcal{C})} \quad \forall w \in \mathring{H}^1(y^\alpha, \mathcal{C}). \tag{2.7}$$

Consequently, the seminorm on $\mathring{H}^1(y^\alpha, \mathcal{C})$ is equivalent to (2.5). For $w \in H^1(y^\alpha, \mathcal{C})$, $\operatorname{tr}_\Omega w$ denotes its trace on $\Omega \times \{0\}$, which satisfies (see [48, Proposition 2.5])

$$\operatorname{tr}_\Omega \mathring{H}^1(y^\alpha, \mathcal{C}) = \mathbb{H}^s(\Omega), \qquad \|\operatorname{tr}_\Omega w\|_{\mathbb{H}^s(\Omega)} \leq C_{\operatorname{tr}_\Omega} \|w\|_{\mathring{H}^1(y^\alpha,\mathcal{C})}. \tag{2.8}$$

Define the bilinear form $a_\mathcal{C} : \mathring{H}^1(y^\alpha, \mathcal{C}) \times \mathring{H}^1(y^\alpha, \mathcal{C}) \to \mathbb{R}$ by

$$a_\mathcal{C}(v, w) = \int_\mathcal{C} y^\alpha (\boldsymbol{A} \nabla v \cdot \nabla w + c v w) \,\mathrm{d}x' \,\mathrm{d}y, \tag{2.9}$$

and note that it is continuous and, owing to (2.7), also coercive. Consequently, it induces an inner product on $\mathring{H}^1(y^\alpha, \mathcal{C})$ and the *energy norm* $\|\cdot\|_{\mathcal{C}}$:

$$\|v\|_{\mathcal{C}}^2 := a_{\mathcal{C}}(v,v) \sim \|\nabla v\|_{L^2(y^\alpha,\mathcal{C})}^2 . \tag{2.10}$$

Occasionally, we will restrict the integration to the truncated cylinder $\mathcal{C}_{\mathcal{Y}}$. The corresponding bilinear form and norm are denoted by

$$a_{\mathcal{C}_{\mathcal{Y}}}(v,w) := \int_{\mathcal{C}_{\mathcal{Y}}} y^\alpha (\boldsymbol{A}\nabla v \cdot \nabla w + cvw)\,\mathrm{d}x'\,\mathrm{d}y, \qquad \|v\|_{\mathcal{C}_{\mathcal{Y}}}^2 = a_{\mathcal{C}_{\mathcal{Y}}}(v,v) . \tag{2.11}$$

With these definitions the weak formulation of (1.2) reads: Find $\mathscr{U} \in \mathring{H}^1(y^\alpha, \mathcal{C})$ such that

$$a_{\mathcal{C}}(\mathscr{U}, v) = d_s \langle f, \operatorname{tr}_\Omega v\rangle \qquad \forall v \in \mathring{H}^1(y^\alpha, \mathcal{C}). \tag{2.12}$$

The fundamental result of Caffarelli and Silvestre [17] then reads as follows (see also [16, Proposition 2.2] and [63, Theorem 1.1] for bounded domains and for general elliptic operators): given $f \in \mathbb{H}^{-s}(\Omega)$, let $u \in \mathbb{H}^s(\Omega)$ solve (1.1). If $\mathscr{U} \in \mathring{H}^1(y^\alpha, \mathcal{C})$ solves (2.12), then $u = \operatorname{tr}_\Omega \mathscr{U}$ and

$$d_s \mathcal{L}^s u = \partial_{\nu^\alpha} \mathscr{U} \quad \text{in } \Omega. \tag{2.13}$$

3. A first order FEM for fractional diffusion. The first work that, in a numerical setting, exploits the identity (2.13) for the design and analysis of a finite element approximation of solutions to (1.1) is [48]; see also [52]. Let us briefly review the main results of [48].

First, [48] truncates $\mathcal{C}$ to $\mathcal{C}_{\mathcal{Y}}$ and places homogeneous Dirichlet boundary conditions on $y = \mathcal{Y}$, thus obtaining an approximation $\mathcal{U}$ (which, by slight abuse of notation, is understood to coincide with its extension by zero to $\mathcal{C} \setminus \mathcal{C}_{\mathcal{Y}}$. The error committed in this approximation is exponentially small: There holds with λ_1 being the first eigenvalue of the operator $\mathcal{L}$ (see [48, Theorem 3.5])

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

Second, since $\partial_y^2 \mathscr{U}$ has a blow-up singularity at $y = 0$, [48] develops a regularity theory for $\mathscr{U}$ in weighted Sobolev spaces; the reader will find a generalization of that regularity theory in Theorem 4.7 below. Consequently, *graded* meshes in the extended variable y play a fundamental role. In the notation of the present work, with a mesh $\mathcal{T}$ on Ω and a mesh $\mathcal{G}^M$ on $(0, \mathcal{Y})$ that is graded towards $y = 0$, the truncated cylinder $\mathcal{C}_{\mathcal{Y}}$ is partitioned by tensor product elements $K \times I$ with $K \in \mathcal{T}$ and $I \in \mathcal{G}^M$. On this mesh, the tensor product space $\mathbb{V}^{1,1}_{h,M}(\mathcal{T}, \mathcal{G}^M)$ of piecewise bilinears in $\Omega \times (0, \mathcal{Y})$ (see (5.2) for the precise definition) is used in a Galerkin method. The Galerkin approximation $\mathscr{U}_{h,M} \in \mathbb{V}^{1,1}_{h,M}(\mathcal{T}, \mathcal{G}^M)$ of $\mathcal{U}$ satisfies a best approximation property à la Céa. From there, upon studying piecewise polynomial interpolation in Muckenhoupt weighted Sobolev spaces [48, 49] error estimates were obtained under the assumption that $f \in \mathbb{H}^{1-s}(\Omega)$ and that Ω is convex (see [48, Theorem 5.4] and [48, Corollary 7.11]):

PROPOSITION 3.1 (a priori error estimate). *Let $\mathcal{G}^M$ be suitably graded towards $y = 0$ and $\mathbb{V}^{1,1}_{h,M}$ be constructed with tensor product elements and $\mathscr{U}_{h,M} \in \mathbb{V}^{1,1}_{h,M}$*

denote the Galerkin approximation to $\mathcal{U}$. Then, for suitable truncation parameter $\mathcal{Y} \sim \log \mathcal{N}_{\Omega,\mathcal{Y}}$ we have that

$$
\begin{aligned}
\|u - \operatorname{tr}_\Omega \mathscr{U}_{h,M}\|_{\mathbb{H}^s(\Omega)} &\lesssim \|\nabla(\mathscr{U} - \mathscr{U}_{h,M})\|_{L^2(y^\alpha,\mathcal{C})} \\
&\lesssim |\log \mathcal{N}_{\Omega,\mathcal{Y}}|^s (\mathcal{N}_{\Omega,\mathcal{Y}})^{-1/(d+1)} \|f\|_{\mathbb{H}^{1-s}(\Omega)},
\end{aligned}
$$

where $\mathcal{N}_{\Omega,\mathcal{Y}} := \#\mathcal{T}\#\mathcal{G}^M$ corresponds to the total number of unknowns.

REMARK 3.2 (complexity). *Up to logarithmic factors, Proposition 3.1 yields rates of convergence of $(\mathcal{N}_{\Omega,\mathcal{Y}})^{-1/(d+1)}$. In terms of error versus work, this P_1-FEM is sub-optimal as a method to compute in Ω. In this paper we propose and study P_1-FE methods in Ω that afford an error decay $(\mathcal{N}_{\Omega,\mathcal{Y}})^{-1/d}$ (up to possibly logarithmic terms).* ■

4. Analytic regularity. We obtain regularity results for the solution of (1.2) that will underlie the analysis of the various FEMs in Section 5 and 7. We begin by recalling that if $u = \sum_{k=1}^{\infty} u_k \varphi_k$ solves (1.1), then the unique solution $\mathscr{U}$ of problem (1.2) admits the representation [48, formula (2.24)]

$$
\mathscr{U}(x', y) = \sum_{k=1}^{\infty} u_k \varphi_k(x') \psi_k(y), \quad u_k := \lambda_k^{-s} f_k. \tag{4.1}
$$

We also recall that $\{\lambda_k, \varphi_k\}_{k\in\mathbb{N}}$ is the set of eigenpairs of the elliptic operator $\mathcal{L}$, supplemented with homogeneous Dirichlet boundary conditions. The functions ψ_k solve

$$
\begin{cases}
\dfrac{\mathrm{d}^2}{\mathrm{d}y^2}\psi_k(y) + \dfrac{\alpha}{y}\dfrac{\mathrm{d}}{\mathrm{d}y}\psi_k(y) - \lambda_k \psi_k(y) = 0, & y \in (0,\infty), \\
\psi_k(0) = 1, & \lim\limits_{y\to\infty} \psi_k(y) = 0.
\end{cases} \tag{4.2}
$$

Thus, if $s = \frac{1}{2}$, we have $\psi_k(y) = \exp(-\sqrt{\lambda_k} y)$ [16, Lemma 2.10]; more generally, if $s \in (0,1) \setminus \{\frac{1}{2}\}$, then [19, Proposition 2.1]

$$
\psi_k(y) = c_s (\sqrt{\lambda_k} y)^s K_s(\sqrt{\lambda_k} y),
$$

where $c_s = 2^{1-s}/\Gamma(s)$ and K_s denotes the modified Bessel function of the second kind. We refer the reader to [1, Chapter 9.6] for a comprehensive treatment of the Bessel function K_s and recall the following properties.

LEMMA 4.1 (properties of K_ν). *The modified Bessel function of the second kind K_ν satisfies:*

(i) For $\nu > -1$ and $z > 0$, $K_\nu(z)$ is real and positive [1, Chapter 9.6].

(ii) For $\nu \in \mathbb{R}$, $K_\nu(z) = K_{-\nu}(z)$ [1, Chapter 9.6].

(iii) For $\nu > 0$, [1, estimate (9.6.9)]

$$
\lim_{z\downarrow 0} \frac{K_\nu(z)}{\frac{1}{2}\Gamma(\nu)\left(\frac{1}{2}z\right)^{-\nu}} = 1. \tag{4.3}
$$

(iv) For $\ell \in \mathbb{N}$, [1, formula (9.6.28)]

$$
\left(\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\right)^{\ell} (z^\nu K_\nu(z)) = (-1)^\ell z^{\nu-\ell} K_{\nu-\ell}(z). \tag{4.4}
$$

(v) For $z > 0$, $z^{\min\{\nu,1/2\}} e^z K_\nu(z)$ is a decreasing function [45, Theorem 5].
(vi) For $\nu > 0$, [1, estimate (9.7.2)]

$$K_\nu(z) \sim \sqrt{\frac{\pi}{2z}} e^{-z}, \qquad z \to \infty, \quad |\arg z| \leq 3\pi/2 - \delta, \quad \delta > 0.$$

REMARK 4.2 (consistency for $s = \frac{1}{2}$). *A combination of formulas (9.2.10) and (9.6.10) in [1] yields $K_{\frac{1}{2}}(z) = \sqrt{\frac{\pi}{2z}} e^{-z}$. Since $c_{\frac{1}{2}} = \sqrt{\frac{2}{\pi}}$, we have arrived at*

$$\lim_{s \to \frac{1}{2}} \psi_k(y) = \exp(-\sqrt{\lambda_k} y) \quad \forall y > 0.$$

■

In order to understand the nature of the y-dependence of the solution $\mathscr{U}$ given by the representation formula (4.1), we derive regularity estimates for the solution ψ_k of problem (4.2). To this end, define the function $\psi(z) = c_s z^s K_s(z)$ and notice that

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\psi(z) - \psi(z) + \frac{\alpha}{z}\frac{\mathrm{d}}{\mathrm{d}z}\psi(z) = 0, \quad z \in (0,\infty), \quad \psi(0) = 1, \quad \lim_{z\to\infty} \psi(z) = 0. \quad (4.5)$$

The differential equation and Leibniz' formula imply, for any $\ell \in \mathbb{N}_0$,

$$\begin{aligned}\frac{\mathrm{d}^{\ell+2}}{\mathrm{d}z^{\ell+2}}\psi(z) &= \frac{\mathrm{d}^{\ell}}{\mathrm{d}z^{\ell}}\psi(z) - \alpha \frac{\mathrm{d}^{\ell}}{\mathrm{d}z^{\ell}}\left(z^{-1}\frac{\mathrm{d}}{\mathrm{d}z}\psi(z)\right) \\ &= \frac{\mathrm{d}^{\ell}}{\mathrm{d}z^{\ell}}\psi(z) - \alpha \sum_{j=0}^{\ell} \binom{\ell}{j} \frac{\mathrm{d}^{j}}{\mathrm{d}z^{j}}(z^{-1}) \frac{\mathrm{d}^{\ell-j}}{\mathrm{d}z^{\ell-j}}\psi'(z).\end{aligned}$$

We thus have arrived at the bound

$$\left|\frac{\mathrm{d}^{\ell+2}}{\mathrm{d}z^{\ell+2}}\psi(z)\right| \leq \left|\frac{\mathrm{d}^{\ell}}{\mathrm{d}z^{\ell}}\psi(z)\right| + |\alpha| \sum_{j=0}^{\ell} \frac{\ell!}{(\ell-j)!} z^{-(1+j)} \left|\frac{\mathrm{d}^{\ell+1-j}}{\mathrm{d}z^{\ell+1-j}}\psi(z)\right|, \quad (4.6)$$

which is essential to derive the following asymptotic result.

LEMMA 4.3 (behavior of ψ near $z = 0$). *Let ψ solve (4.5). Let $s \in (0,1)$ and set $d_s = 2^{1-2s}\Gamma(1-s)/\Gamma(s)$. Then there is $C_s > 0$ independent of $z \in (0,1)$ and $\ell \in \mathbb{N}$ such that*

$$\left|\frac{\mathrm{d}^{\ell}}{\mathrm{d}z^{\ell}}\psi(z)\right| \leq C_s d_s \ell! z^{2s-\ell}. \quad (4.7)$$

Proof. We proceed by induction, starting with the case $\ell = 1$. The differentiation formula (4.4) with $\ell = 1$ yields that

$$\psi'(z) = c_s(z^s K_s(z))' = -c_s z^s K_{s-1}(z) = -c_s z^s K_{1-s}(z), \quad (4.8)$$

where we used Lemma 4.1 (ii). The asymptotic formula (4.3) shows that there is $\tilde{C}_s$ independent of $z \in (0,1)$ such that

$$\left|\frac{K_{1-s}(z)}{\frac{1}{2}\Gamma(1-s)(\frac{1}{2}z)^{-(1-s)}} - 1\right| \leq \tilde{C}_s.$$

Set $C_s := \tilde{C}_s + 1$ to arrive at the fact that we have, for all $z \in (0,1)$,

$$\left|\frac{\mathrm{d}}{\mathrm{d}z}\psi(z)\right| \le \left|\frac{K_{1-s}(z)}{\frac{1}{2}\Gamma(1-s)(\frac{1}{2}z)^{-(1-s)}}\right| \left(\frac{1}{2}\Gamma(1-s)\left(\frac{1}{2}z\right)^{-(1-s)}\right) c_s z^s \le C_s d_s z^{2s-1}.$$

This shows (4.7) for $\ell = 1$. For the induction step, we assume that (4.7) holds for all differentiation orders up to $\ell + 1 \ge 1$. Together with (4.6), we then get

$$\begin{aligned}\left|\frac{\mathrm{d}^{\ell+2}}{\mathrm{d}z^{\ell+2}}\psi(z)\right| &\le C_s d_s \ell! z^{2s-\ell} + C_s d_s z^{2s-\ell-2}\sum_{j=0}^{\ell}\frac{\ell!}{(\ell-j)!}(\ell+1-j)! \\ &\le C_s d_s \ell! z^{2s-\ell-2}\left[1 + \sum_{i=1}^{\ell+1} i\right] = C_s d_s \ell! z^{2s-\ell-2}\left[1 + \frac{1}{2}(\ell+1)(\ell+2)\right],\end{aligned}$$

because $z \in (0,1)$. Noting $1 + \frac{1}{2}(\ell+1)(\ell+2) \le (\ell+1)(\ell+2)$ gives

$$\left|\frac{\mathrm{d}^{\ell+2}}{\mathrm{d}z^{\ell+2}}\psi(z)\right| \le C_s d_s (\ell+2)! z^{2s-\ell-2},$$

which concludes the proof. □

We now analyze the behavior of ψ for large, positive values of z. We show that ψ and all its derivatives decay exponentially as $z \to \infty$.

LEMMA 4.4 (behavior of ψ for z large). *Let ψ solve (4.5). Fix $\epsilon \in (0,1)$ and $s \in (0,1)$. Then there is a constant $C_{\epsilon,s}$ depending solely on ϵ and s such that*

$$\left|\frac{\mathrm{d}^{\ell}}{\mathrm{d}z^{\ell}}\psi(z)\right| \le C_{\epsilon,s}\ell!\epsilon^{-\ell} z^{s-\ell-\frac{1}{2}} e^{-(1-\epsilon)z} \quad \forall z \ge 1, \quad \forall \ell \in \mathbb{N}_0. \tag{4.9}$$

Proof. The proof is a consequence of Cauchy's integral formula for derivatives [4, 20] and Lemma 4.1 (vi). Let $B_\sigma(\zeta) \subset \mathbb{C}$ denote the ball with center ζ and radius σ. For any fixed $z \ge 1$, $\ell \in \mathbb{N}_0$ we then have

$$\left|\frac{\mathrm{d}^{\ell}}{\mathrm{d}z^{\ell}}\psi(z)\right| = \left|\frac{\ell!}{2\pi i}\int_{\zeta \in \partial B_{\epsilon z}(z)}\frac{\psi(\zeta)}{(\zeta - z)^{\ell+1}}\,\mathrm{d}\zeta\right| \le \ell!\epsilon^{-\ell}z^{-\ell}\max_{\zeta \in \partial B_{\epsilon z}(z)}|\psi(\zeta)|.$$

Recalling $\psi(z) = c_s z^s K_s(z)$ and invoking Lemma 4.1 (vi) we conclude that

$$\left|\frac{\mathrm{d}^{\ell}}{\mathrm{d}z^{\ell}}\psi(z)\right| \le C_{\epsilon,s} c_s \ell! \epsilon^{-\ell} z^{s-\ell-\frac{1}{2}} e^{-(1-\epsilon)z},$$

with $C_{\epsilon,s} = C' \max\{(1+\epsilon)^{s-\frac{1}{2}}, (1-\epsilon)^{s-\frac{1}{2}}\}$ and C' such that $|K_s(z)| \le C'|z|^{-\frac{1}{2}}e^{-\operatorname{Re} z}$ for $\operatorname{Re} z \ge 0$. We remark in passing that the constant C' can be chosen independent of $s \in (0,1)$ since, for ζ in the right half plane, the function $\mu \mapsto |K_\mu(\zeta)|$ for $\mu > 0$ is monotone increasing, [26, (10.37.1)]. □

REMARK 4.5 (Cauchy's integral formula). *Cauchy's integral formula can also be invoked to analyze the ℓ-th derivative of ψ near $z = 0$. However, the resulting estimate is not quite as sharp as (4.7) since it include a term $\epsilon^{-\ell}$ with $\epsilon \in (0,1)$, as it appears in the estimate (4.9).* ■

To analyze global regularity properties of the α–harmonic extension $\mathscr{U}$ of a function u on Ω we define the weight

$$\omega_{\beta,\gamma}(y) = y^\beta e^{\gamma y}, \qquad 0 \leq \gamma < 2\sqrt{\lambda_1}, \tag{4.10}$$

with a parameter $\beta \in \mathbb{R}$ that will be specified later. We recall that the parameter $\lambda_1 > 0$ is the smallest eigenvalue of $\mathcal{L}$. In terms of the weight (4.10), we define the weighted norm

$$\|v\|_{L^2(\omega_{\beta,\gamma},\mathcal{C})} := \left(\int_0^\infty \int_\Omega \omega_{\beta,\gamma}(y) |v(x',y)|^2 \, \mathrm{d}x' \, \mathrm{d}y \right)^{\frac{1}{2}}. \tag{4.11}$$

Our analysis will require control of certain weighted integrals of derivatives of ψ. We define, for γ satisfying (4.10) and β, $\delta \in \mathbb{R}$, $\ell \in \mathbb{N}$, $\lambda > 0$, the integrals

$$\Phi(\delta,\gamma,\lambda) = \int_0^\infty z^\delta e^{\gamma z/\sqrt{\lambda}} \left|\psi(z)\right|^2 \, \mathrm{d}z, \tag{4.12}$$

$$\Psi_\ell(\beta,\gamma,\lambda) = \int_0^\infty z^{\beta+2\ell} e^{\gamma z/\sqrt{\lambda}} \left| \frac{\mathrm{d}^\ell}{\mathrm{d}z^\ell} \psi(z) \right|^2 \, \mathrm{d}z. \tag{4.13}$$

Let us bound the integrals $\Phi(\delta,\gamma,\lambda)$ and $\Psi_\ell(\beta,\gamma,\lambda)$.

Lemma 4.6 (bounds on Φ and Ψ_ℓ). *Let $\delta > -1$, $\beta > -1-4s$, $\ell \in \mathbb{N}$, and let γ satisfy $0 \leq \gamma < 2\sqrt{\lambda_1}$. If $\lambda \geq \lambda_1$, then we have that*

$$\Phi(\delta,\gamma,\lambda) \lesssim 1, \tag{4.14}$$

where the hidden constant is independent of λ. In addition, there exists $\kappa > 1$ such that for every $\ell \in \mathbb{N}$ we have the following bound

$$\Psi_\ell(\beta,\gamma,\lambda) \lesssim \kappa^{2\ell} (\ell!)^2, \tag{4.15}$$

where the hidden constant is independent of ℓ and λ.

Proof. We derive (4.15). As a first step, we write $\Psi_\ell = \Psi_\ell(\beta,\gamma,\lambda)$ as follows:

$$\Psi_\ell = \int_0^1 z^{\beta+2\ell} e^{\frac{\gamma z}{\sqrt{\lambda}}} \left| \frac{\mathrm{d}^\ell}{\mathrm{d}z^\ell} \psi(z) \right|^2 \, \mathrm{d}z + \int_1^\infty z^{\beta+2\ell} e^{\frac{\gamma z}{\sqrt{\lambda}}} \left| \frac{\mathrm{d}^\ell}{\mathrm{d}z^\ell} \psi(z) \right|^2 \, \mathrm{d}z =: \mathrm{I} + \mathrm{II}, \tag{4.16}$$

and estimate each term separately. We start by bounding term I. Since $0 \leq \gamma < 2\sqrt{\lambda_1}$ and $\lambda \geq \lambda_1$ we have that

$$\sup_{z \in (0,1)} e^{\frac{\gamma z}{\sqrt{\lambda}}} < \sup_{z \in (0,1)} e^{2z} \leq e^2.$$

Consequently, an application of Lemma 4.3 yields

$$\mathrm{I} = \int_0^1 z^{\beta+2\ell} e^{\frac{\gamma z}{\sqrt{\lambda}}} \left| \frac{\mathrm{d}^\ell}{\mathrm{d}z^\ell} \psi(z) \right|^2 \, \mathrm{d}z \lesssim d_s^2 (\ell!)^2 \int_0^1 z^{\beta+2\ell+2(2s-\ell)} \, \mathrm{d}z \lesssim d_s^2 (\ell!)^2,$$

where last integral converges because $\beta > -1-4s$. Notice that the hidden constant blows up when $\beta \downarrow -1-4s$.

We now estimate the term II in (4.16). To do this we utilize the estimate (4.9) of Lemma 4.4 as follows:

$$\mathrm{II} \leq C_\epsilon^2 c_s^2 (\ell!)^2 \epsilon^{-2\ell} \int_1^\infty z^{\beta+2\ell} z^{2s-2\ell-1} e^{\frac{\gamma z}{\sqrt{\lambda}}} e^{-2(1-\epsilon)z} \,\mathrm{d}z.$$

Define

$$\hat{\gamma} := \sup_{\lambda \geq \lambda_1} \left(\frac{\gamma}{\sqrt{\lambda}} - 2(1-\epsilon) \right).$$

Notice that, since $0 \leq \frac{\gamma}{\sqrt{\lambda_1}} < 2$ by (4.10), the parameter $\epsilon \in (0,1)$ can be selected such that $\hat{\gamma} < 0$. Consequently

$$\mathrm{II} \lesssim C_\epsilon^2 c_s^2 (\ell!)^2 \epsilon^{-2\ell} \int_1^\infty z^{\beta+2s-1} e^{\hat{\gamma} z} \,\mathrm{d}z \lesssim C_\epsilon^2 c_s^2 (\ell!)^2 \epsilon^{-2\ell}.$$

Inserting the estimates for the terms I and II in (4.16) and selecting $\kappa = \epsilon^{-1} > 1$ we arrive at the desired estimate (4.15). The estimate (4.14) is obtained in a similar way: We decompose Φ as in (4.16) and use that, as estimate (4.3) shows, ψ is bounded as $z \downarrow 0^+$ and decays exponentially to zero as $z \uparrow \infty$; see Lemma 4.1 (v) and (vi). For brevity, we skip the details. ∎

On the basis of Lemma 4.6, we provide global regularity results for the α-harmonic extension $\mathscr{U}$ in weighted Sobolev spaces.

Theorem 4.7 (global regularity of $\mathscr{U}$). *Let $\mathscr{U} \in \mathring{H}^1(y^\alpha, \mathcal{C})$ solve (1.2) with $s \in (0,1)$. Let $0 \leq \tilde{\nu} < s$ and $0 \leq \nu < 1+s$. Then there exists $\kappa > 1$ such that the following holds for all $\ell \in \mathbb{N}_0$ with the weight $w_{\beta,\gamma}$ given by (4.10):*

$$\|\partial_y^{\ell+1} \mathscr{U}\|_{L^2(\omega_{\alpha+2\ell-2\tilde{\nu},\gamma},\mathcal{C})} \lesssim \kappa^{\ell+1} (\ell+1)! \|f\|_{\mathbb{H}^{-s+\tilde{\nu}}(\Omega)}, \tag{4.17}$$

$$\|\nabla_{x'} \partial_y^{\ell+1} \mathscr{U}\|_{L^2(\omega_{\alpha+2(\ell+1)-2\nu,\gamma},\mathcal{C})} \lesssim \kappa^{\ell+1} (\ell+1)! \|f\|_{\mathbb{H}^{-s+\nu}(\Omega)}, \tag{4.18}$$

$$\|\mathcal{L} \partial_y^{\ell+1} \mathscr{U}\|_{L^2(\omega_{\alpha+2(\ell+1)-2\nu,\gamma},\mathcal{C})} \lesssim \kappa^{\ell+1} (\ell+1)! \|f\|_{\mathbb{H}^{1-s+\nu}(\Omega)}. \tag{4.19}$$

In all these inequalities, the hidden constants are independent of ℓ, $\mathscr{U}$, and f. In addition, if $0 \leq \nu' < 1-s$ then

$$\|\mathcal{L} \mathscr{U}\|_{L^2(\omega_{\alpha-2\nu',\gamma},\mathcal{C})} \lesssim \|f\|_{\mathbb{H}^{1-s+\nu'}(\Omega)}, \tag{4.20}$$

$$\|\nabla_{x'} \mathscr{U}\|_{L^2(\omega_{\alpha-2\nu',\gamma},\mathcal{C})} \lesssim \|f\|_{\mathbb{H}^{-s+\nu'}(\Omega)}, \tag{4.21}$$

$$\|\mathscr{U}\|_{L^2(\omega_{\alpha-2\nu',\gamma},\mathcal{C})} \lesssim \|f\|_{\mathbb{H}^{-1-s+\nu'}(\Omega)}, \tag{4.22}$$

where the constant entailed in $\lesssim$ is independent of $\mathscr{U}$ and f.

Proof. Following [48, Theorem 2.7] we start from the representation formula (4.1) to arrive at

$$\|\partial_y^{\ell+1} \mathscr{U}\|^2_{L^2(\omega_{\alpha+2\ell-2\sigma,\gamma},\mathcal{C})} = \sum_{k=1}^\infty f_k^2 \lambda_k^{-2s} \int_0^\infty y^{\alpha+2\ell-2\sigma} e^{\gamma y} \left| \frac{\mathrm{d}^{\ell+1}}{\mathrm{d}y^{\ell+1}} \psi_k(y) \right|^2 \,\mathrm{d}y\ .$$

With the change of variable $z = \sqrt{\lambda_k} y$ and recalling $\psi(z) = c_s z^s K_s(z)$ and $\psi_k(y) = \psi(\sqrt{\lambda_k} y)$ as well as the definition of Ψ_ℓ given in (4.13), to obtain

$$\begin{aligned}\|\partial_y^{\ell+1} \mathscr{U}\|^2_{L^2(\omega_{\alpha+2\ell-2\sigma,\gamma},\mathcal{C})} &= \sum_{k=1}^\infty f_k^2 \lambda_k^{-2s+(\ell+1)-\left(\frac{\alpha+2\ell-2\sigma}{2}\right)-\frac{1}{2}} \Psi_{\ell+1}(\alpha-2\sigma-2,\gamma,\lambda_k) \\ &\lesssim (\ell+1)!^2 \kappa^{2(\ell+1)} \sum_{k=1}^\infty f_k^2 \lambda_k^{\sigma-s} = (\ell+1)!^2 \kappa^{2(\ell+1)} \|f\|^2_{\mathbb{H}^{-s+\sigma}(\Omega)},\end{aligned}$$

where the last inequality follows from the estimate (4.15) with $\beta = \alpha - 2\sigma - 2 = 1 - 2s - 2\sigma - 2 > -1 - 4s$. This shows (4.17).

We now derive (4.19); the proof of the estimate (4.18) follows by using similar arguments. As before, we estimate

$$\begin{aligned} &\|\mathcal{L}\partial_y^{\ell+1}\mathscr{U}\|^2_{L^2(\omega_{\alpha+2(\ell+1)-2\nu,\gamma},\mathcal{C})} \\ &= \sum_{k=1}^{\infty} f_k^2 \lambda_k^{2(1-s)} \int_0^\infty y^{\alpha+2(\ell+1)-2\nu} e^{\gamma y} \left| \frac{\mathrm{d}^{\ell+1}}{\mathrm{d}y^{\ell+1}} \psi_k(y) \right|^2 \mathrm{d}y \\ &= \sum_{k=1}^{\infty} f_k^2 \lambda_k^{2(1-s)+(\ell+1)-\left(\frac{\alpha+2(\ell+1)-2\nu}{2}\right)-\frac{1}{2}} \Psi_{\ell+1}(\alpha - 2\nu, \gamma, \lambda_k), \end{aligned}$$

where we applied again the change of variable $z = \sqrt{\lambda_k} y$ and used the definition of Ψ_ℓ given by (4.13). We now notice that $\alpha - 2\nu > 1 - 2s - 4 - 2s = -1 - 4s$. Thus an application of the estimate (4.15) with $\beta = \alpha - 2\nu$ reveals that

$$\|\mathcal{L}\,\partial_y^{\ell+1}\mathscr{U}\|^2_{L^2(\omega_{\alpha+2(\ell+1)-2\nu,\gamma},\mathcal{C})} \lesssim \kappa^{2(\ell+1)}(\ell+1)!^2 \|f\|^2_{\mathbb{H}^{1-s+\nu}(\Omega)}.$$

This yields (4.19). The proofs of (4.20), (4.21), (4.22) rely on similar arguments using that $\nu' < 1 - s$ implies $\delta := \alpha - 2\nu' = 1 - 2s - 2\nu' > 1 - 2s - 2(1-s) = -1$, and thus, as a consequence of (4.14), that $\Phi(\delta, \gamma, \lambda) \lesssim 1$. This concludes the proof. ◻

5. h-FE discretization in Ω. In this section we present and analyze three discretizations of (2.12) that rely on P_1-FEM in Ω, and structure it as follows: Section 5.1 introduces the FE approximation in Ω and fixes notation on Finite Element spaces. Section 5.2 presents the FE discretization in $\mathcal{C}$ in abstract form. Section 5.3 discusses a basic decomposition of the FE discretization error into two parts: a semidiscretization error with respect to $x' \in \Omega$, and a corresponding error with respect to $y \in (0, \mathcal{Y})$, where $0 < \mathcal{Y} < \infty$ denotes a truncation parameter of the cylinder $(0, \infty)$. Section 5.4 then deals with two first order tensor product FEMs in $\mathcal{C}$. The first one, as in [48], is a full tensor product FEM for which we show (under sufficient regularity of the solution and under compatibility assumptions on the data) convergence in Ω with rate 1 (in terms of the mesh size), but with superlinear complexity in terms of the number $\mathcal{N}_\Omega$ of degrees of freedom in Ω. To reduce the complexity, we propose the second, novel approach: a sparse tensor product of P_1 finite elements for the extended problem in $\mathcal{C}$, for which we show the same convergence rate, but with (essentially) linear complexity in terms of $\mathcal{N}_\Omega$ requiring only marginally more regularity of the data f in Ω. Section 5.5 addresses a third method, namely, the use of an hp-FEM in the extended variable y, combined with a P_1-FEM in Ω.

5.1. Notation and FE spaces. For a truncation parameter $\mathcal{Y} > 0$ (which is fixed, and which will be selected ahead), we denote by $\mathcal{G}^M$ a generic partition of $[0, \mathcal{Y}]$ into M intervals. In particular, the following two types of partitions that are refined towards $y = 0$ will be used:

- *Graded* meshes $\mathcal{G}^k_{gr,\eta}$. Here k indicates the mesh size near $y = 1$ and η characterizes the mesh grading towards $y = 0$; see Section 5.4.2 ahead for details.
- *Geometric* meshes $\mathcal{G}^M_{geo,\sigma}$. This mesh has M elements and $\sigma \in (0, 1)$ is the subdivision ratio; see Section 5.5.1 ahead for details.

Given a mesh $\mathcal{G}^M = \{I_m\}_{m=1}^M$ in $[0, \mathcal{Y}]$, where $I_m = [y_{m-1}, y_m]$, $y_0 = 0$ and $y_M = \mathcal{Y}$, we associate to $\mathcal{G}^M$ a *polynomial degree distribution* $\boldsymbol{r} = (r_1, r_2, \ldots, r_M) \in \mathbb{N}^M$.

With these ingredients at hand we define the finite element space

$$S^{\boldsymbol{r}}((0,\mathcal{Y}),\mathcal{G}^M) = \left\{ v_M \in C[0,\mathcal{Y}] : v_M|_{I_m} \in \mathbb{P}_{r_m}(I_m), I_m \in \mathcal{G}^M, m = 1,\ldots,M \right\}.$$

We also define the subspace of $S^{\boldsymbol{r}}((0,\mathcal{Y}),\mathcal{G}^M)$ of functions that vanish at $y=\mathcal{Y}$:

$$S^{\boldsymbol{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}),\mathcal{G}^M) = \left\{ v_M \in S^{\boldsymbol{r}}((0,\mathcal{Y}),\mathcal{G}^M) : v_M(\mathcal{Y}) = 0 \right\}. \tag{5.1}$$

In the particular case that $r_i = r$ for $i = 1,\ldots,M$, we write $S^r((0,\mathcal{Y}),\mathcal{G}^M)$ or $S^r_{\{\mathcal{Y}\}}((0,\mathcal{Y}),\mathcal{G}^M)$ as appropriate. In Ω, we consider Lagrangian FEM of polynomial degree $q \geq 1$ based on shape-regular, simplicial triangulations denoted by $\mathcal{T}$. Denote by $h(\mathcal{T}) = \max\{\operatorname{diam}(K) : K \in \mathcal{T}\}$ the meshwidth of $\mathcal{T}$. We introduce

$$S^q_0(\Omega,\mathcal{T}) = \left\{ v_h \in C(\bar{\Omega}) : v_h|_K \in \mathbb{P}_q(K) \quad \forall K \in \mathcal{T},\ v_h|_{\partial\Omega} = 0 \right\}.$$

Later, we will also consider nested sequences $\{\mathcal{T}^\ell\}_{\ell\geq 0}$ of triangulations of Ω that are generated by bisection–tree refinement of a coarse, regular initial triangulation $\mathcal{T}^0$ of Ω. Then, we denote by $h_\ell = \max\{\operatorname{diam}(K) : K \in \mathcal{T}^\ell\}$ the meshwidth of $\mathcal{T}^\ell$.

We define the finite–dimensional *tensor product space*

$$\mathbb{V}^{q,\boldsymbol{r}}_{h,M}(\mathcal{T},\mathcal{G}^M) := S^q_0(\Omega,\mathcal{T}) \otimes S^{\boldsymbol{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}),\mathcal{G}^M) \subset \mathring{H}^1(y^\alpha,\mathcal{C}), \tag{5.2}$$

and write $\mathbb{V}_{h,M}$ if the arguments are clear from the context. In the ensuing error analysis, we also require semidiscretizations which are based on the following (infinite–dimensional) Hilbertian tensor product spaces

$$\begin{aligned} \mathbb{V}^q_h(\mathcal{C}_\mathcal{Y}) &:= S^q_0(\Omega,\mathcal{T}_h) \otimes H^1_\mathcal{Y}(y^\alpha,(0,\mathcal{Y})) \subset \mathring{H}^1(y^\alpha,\mathcal{C}), \\ \mathbb{V}^{\boldsymbol{r}}_M(\mathcal{C}_\mathcal{Y}) &:= H^1_0(\Omega) \otimes S^{\boldsymbol{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}),\mathcal{G}^M) \subset \mathring{H}^1(y^\alpha,\mathcal{C}), \end{aligned} \tag{5.3}$$

where

$$H^1_\mathcal{Y}(y^\alpha,(0,\mathcal{Y})) = \{v \in H^1(y^\alpha,(0,\mathcal{Y})) \colon v(\mathcal{Y}) = 0\}. \tag{5.4}$$

Both of them are closed subspaces of $\mathring{H}^1(y^\alpha,\mathcal{C})$, so that Galerkin projections with respect to the inner product given by the bilinear form $a_{\mathcal{C}_\mathcal{Y}}$ in (2.11) are well defined. We denote these projections by G^q_h and $G^{\boldsymbol{r}}_M$, respectively. To the space $\mathbb{V}^{q,\boldsymbol{r}}_{h,M}(\mathcal{T},\mathcal{G}^M)$, defined in (5.2), we can also associate a Galerkin projection with respect to $a_{\mathcal{C}_\mathcal{Y}}$. We remark that this projector is the composition of the semidiscrete projections:

$$G^{q,\boldsymbol{r}}_{h,M} = G^q_h \circ G^{\boldsymbol{r}}_M = G^{\boldsymbol{r}}_M \circ G^q_h :\ \mathring{H}^1(y^\alpha,\mathcal{C}) \to \mathbb{V}^{q,\boldsymbol{r}}_{h,M}(\mathcal{T},\mathcal{G}^M). \tag{5.5}$$

5.2. FE discretization and quasioptimality. The *FE approximation* $\mathscr{U}_{h,M}$ is defined as $\mathscr{U}_{h,M} = G^{q,\boldsymbol{r}}_{h,M}\mathscr{U} \in \mathbb{V}_{h,M}$, i.e., it satisfies

$$a_{\mathcal{C}_\mathcal{Y}}(\mathscr{U}_{h,M},\phi) = d_s\langle f, \operatorname{tr}_\Omega \phi\rangle \quad \forall \phi \in \mathbb{V}_{h,M}. \tag{5.6}$$

Coercivity of $a_{\mathcal{C}_\mathcal{Y}}$ immediately implies existence and uniqueness of $\mathscr{U}_{h,M}$. In addition, Galerkin orthogonality gives quasioptimality of $\mathscr{U}_{h,M}$. More precisely, as in [48, Section 4], we have the following result.

LEMMA 5.1 (Céa and truncation). *Let $\mathscr{U}$ be the solution to problem* (2.12), *and let $\mathscr{U}_{h,M} = G^{q,\boldsymbol{r}}_{h,M}\mathscr{U}$ be its finite element approximation that solves* (5.6). *Then we have*

$$\begin{aligned} \|\nabla(\mathscr{U} - \mathscr{U}_{h,M})\|_{L^2(y^\alpha,\mathcal{C})} \lesssim{}& \min_{v_{h,M}\in\mathbb{V}_{h,M}} \|\nabla(\mathscr{U} - v_{h,M})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \\ &+ \|\nabla\mathscr{U}\|_{L^2(y^\alpha,\mathcal{C}\setminus\mathcal{C}_\mathcal{Y})}, \end{aligned} \tag{5.7}$$

where the hidden constant does not depend on $\mathbb{V}_{h,M}$.

As already noted in [48, Prop. 3.1], the second term on the right-hand side of (5.7) is exponentially small in $\mathcal{Y}$. More precisely, using (4.17) and (4.21) we get, with the selection $\gamma < 2\sqrt{\lambda_1}$, that

$$\|\nabla \mathscr{U}\|_{L^2(y^\alpha, \mathcal{C}\setminus\mathcal{C}_\mathcal{Y})} \lesssim \exp(-\gamma\mathcal{Y}/2)\|f\|_{\mathbb{H}^{-s}(\Omega)}. \tag{5.8}$$

5.3. FE error splitting. As (5.8) shows, the second term on the right-hand side of of (5.7) decays exponentially in $\mathcal{Y}$. Thus, we now concentrate on estimating the first one.

As in [48, 40], we separate the errors incurred by discretizations with respect to x' and y as follows.

Lemma 5.2 (dimensional error splitting). *Let* $\mathscr{U}$ *be the solution to problem* (2.12) *and let* $\Pi^q_{x'} : L^2(\Omega) \to S^q_0(\Omega, \mathcal{T})$ *be a linear operator that is simultaneously stable in* L^2 *and* H^1*, i.e. there exist constants* c_{L^2}, c_{H^1} *such that*

$$\|\Pi^q_{x'} v\|_{L^2(\Omega)} \leq c_{L^2}\|v\|_{L^2(\Omega)} \;\; \forall v \in L^2(\Omega), \quad |\Pi^q_{x'} v|_{H^1(\Omega)} \leq c_{H^1}|v|_{H^1(\Omega)} \;\; \forall v \in H^1_0(\Omega).$$

Then

$$\begin{aligned} \min_{v_{h,M}\in\mathbb{V}_{h,M}} \|\nabla(\mathscr{U} - v_{h,M})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} &\leq \|\nabla(\mathscr{U} - \Pi^q_{x'}\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \\ &\quad + \sqrt{c^2_{L^2} + c^2_{H^1}}\|\nabla(\mathscr{U} - \pi^{\boldsymbol{r}}_y\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})}. \end{aligned} \tag{5.9}$$

Proof. The desired estimate follows from the tensor-product structure of the finite element space defined in (5.2) and the triangle inequality, upon choosing in (5.9) the function $v_{h,M} := \Pi^q_{x'} \otimes \pi^{\boldsymbol{r}}_y \mathscr{U}$. □

5.4. h-FE error analysis. In the present subsection we analyze convergence rates and complexity for two particular instances of the FE-space $\mathbb{V}^{q,\boldsymbol{r}}_{h,M}(\mathcal{T}, \mathcal{G}^M)$:

(a) The case when $\boldsymbol{r} = (1, 1, \dots, 1)$ on a graded mesh $\mathcal{G}^M$ and $q = 1$. A particular instance of this was first introduced in [48]; see Section 3. Generalizing the results of [48, 40], we allow $\Omega \subset \mathbb{R}^2$ to be a polygon with finitely many straight sides and corners $\{\boldsymbol{c}\}$. This will mandate the use of a sequence of nested triangulations $\{\mathcal{T}^\ell\}_{\ell\geq 0}$ of the domain Ω with, in general, local refinement towards the corners $\boldsymbol{c} \in \partial\Omega$.

(b) The case $\boldsymbol{r} = (1, 1, \dots, 1)$ on a nested sequence $\{\mathcal{G}^{\ell'}\}_{\ell'\geq 1}$ of graded meshes in $(0, \mathcal{Y})$. At the same time, we also consider multilevel approximations in Ω on a sequence $\{\mathcal{T}^\ell\}_{\ell\geq 0}$ of nested triangulations with appropriate corner refinement in Ω, a particular instance being the so-called bisection–tree refinements.

In all cases, we bound the first term on the right-hand side of (5.7).

5.4.1. P_1-FEM in Ω with mesh refinement at $\boldsymbol{c}$. We start with the case $s = 1$. In a bounded polygon $\Omega \subset \mathbb{R}^2$ with straight sides and corners $\{\boldsymbol{c}\}$ we consider the Dirichlet problem

$$\mathcal{L}w = g \text{ in } \Omega\,, \quad w = 0 \text{ on } \partial\Omega\,, \tag{5.10}$$

for $g \in H^{-1}(\Omega)$. It is immediate that problem (5.10) has a unique solution $w \in H^1_0(\Omega)$. However, in general the solution w does not belong to $H^2(\Omega)$. Under additional regularity assumptions on A, c and on g, it rather belongs to weighted Sobolev spaces of Kondrat'ev type in Ω which we now define.

For the finite set $\{\boldsymbol{c}\}$ of corners $\boldsymbol{c}$ of Ω define $\Omega \ni x \mapsto \Phi(x) = \prod_{\boldsymbol{c}} |x - \boldsymbol{c}|$. For $0 \leq \beta \in \mathbb{R}$, we denote $L^2_\beta(\Omega) = L^2(\Phi^{2\beta}, \Omega)$. We also define the space $H^2_\beta(\Omega)$ as the closure of $H^2(\Omega) \cap H^1_0(\Omega)$ with respect to the norm

$$\|w\|_{H^2_\beta(\Omega)} = \|w\|_{H^1(\Omega)} + \|D^2 w\|_{L^2(\Phi^{2\beta},\Omega)} \, . \tag{5.11}$$

With this setting at hand, we present the following result on regularity shift in weighted Sobolev spaces for the solution of problem (5.10).

PROPOSITION 5.3 (weighted regularity estimate). *Let $\Omega \subset \mathbb{R}^2$ be a polygon. Let $A \in W^{1,\infty}(\Omega, \mathsf{GL}(\mathbb{R}^2))$ be uniformly positive definite, $c \in W^{1,\infty}(\Omega, \mathbb{R})$. Then there exists $\beta \in [0, 1)$ (depending only on Ω, A, c) such that for every $g \in L^2_\beta(\Omega)$ the solution w of* (5.10) *belongs to $H^2_\beta(\Omega)$ and*

$$\|w\|_{H^2_\beta(\Omega)} \lesssim \|\mathcal{L}w\|_{L^2_\beta(\Omega)} = \|g\|_{L^2_\beta(\Omega)} \, , \tag{5.12}$$

where the hidden constant is independent of g.

Proof. This result is a particular case of [9, Theorem 1.1]. It suffices to set, in the notation of that reference, $m = 1$, $b_j = 0$, and $\beta = 1 - a$. □

REMARK 5.4 (Laplacian). *In the special case that $\mathcal{L} = -\Delta$, i.e., when* (5.10) *corresponds to the Dirichlet Poisson problem in a polygon Ω, the parameter β must satisfy $\beta > 1 - \min_{\boldsymbol{c}} \pi/\omega_{\boldsymbol{c}}$, where $0 < \omega_{\boldsymbol{c}} < 2\pi$ is the interior opening angle of Ω at the vertex $\boldsymbol{c}$. If Ω is convex, the choice $\beta = 0$ is admissible, and then* (5.12) *reduces to the classical regularity shift for the Dirichlet problem of the Poisson equation in convex domains. We refer the reader to the discussion in [9, equations (2) and (3)] for more details.* ▪

Proposition 5.3 and the regularity of $\mathscr{U}$ given in Theorem 4.7 imply the following regularity result for $\mathscr{U}$ in weighted norms in Ω.

THEOREM 5.5 (global regularity of $\mathscr{U}$: weighted estimates in Ω). *Let $\Omega \subset \mathbb{R}^2$ be a bounded polygon, and let A, c satisfy the assumptions of Proposition 5.3. Let $\mathscr{U} \in \mathring{H}^1(y^\alpha, \mathcal{C})$ solve* (1.2) *with $s \in (0, 1)$. Then there exists $\beta \in [0, 1)$ (depending only on Ω, A, c) such that the following regularity assertions hold with the weight $\omega_{\delta,\gamma}$ given by* (4.10)*:*

(i) For $0 \leq \nu' < 1 - s$, we have that

$$\|\mathscr{U}\|_{L^2(\omega_{\alpha-2\nu',\gamma},(0,\infty);H^2_\beta(\Omega))} \lesssim \|f\|_{\mathbb{H}^{1-s+\nu'}(\Omega)}. \tag{5.13}$$

(ii) For $0 \leq \tilde{\nu} < 1 + s$, there exists $\kappa > 1$ such that

$$\|\partial_y^{\ell+1}\mathscr{U}\|_{L^2(\omega_{\alpha+2(\ell+1)-2\tilde{\nu},\gamma},(0,\infty);H^2_\beta(\Omega))} \lesssim \kappa^{\ell+1}(\ell+1)!\|f\|_{\mathbb{H}^{1-s+\tilde{\nu}}(\Omega)}, \tag{5.14}$$

for all $\ell \in \mathbb{N}_0$.

In both estimates, the hidden constants are independent of $\mathscr{U}$ and f.

Proof. The proof of (5.14) follows with the aid of (4.20) and that of (5.13) with (4.19) by using the weighted regularity shift (5.12). In more detail, for a fixed $y > 0$ and $m \in \mathbb{N}_0$, set $w = \partial_y^m \mathscr{U}(\cdot, y)$ in (5.10). Notice that $g = \partial_y^m \mathcal{L}\mathscr{U}(\cdot, y)$. Since $\beta \geq 0$ we have that $g \in L^2_\beta(\Omega)$ and estimate (5.12) holds. Square it and multiply it by either $\omega_{\alpha-2\nu',\gamma}$ if $m = 0$, or $\omega_{\alpha+2m-2\nu,\gamma}$ when $m \geq 1$. Integration with respect to y over $(0, \infty)$ allows us then to conclude. □

The regularity result provided in Theorem 5.5 will be the basis for the analysis of a P_1-FEM on properly refined meshes in Ω that will allow us to recover in Theorem 5.10 below the full first order convergence rate. The proof will require meshes and approximation operators suitable for the approximation of functions with $H^2_\beta(\Omega)$ regularity. This is achieved with appropriate refinement towards the vertices of Ω. Since the sparse grids approach ahead will require nested spaces, it is expedient to impose the additional constraint that the meshes be obtained by repeated bisection from an initial triangulation. Lemma 5.2 indicates that it is desirable to have approximation operators that are uniformly bounded with respect to ℓ in both, $L^2(\Omega)$ and $H^1(\Omega)$. The following lemma shows that these requirements can be fulfilled:

Lemma 5.6 (meshes $(\mathcal{T}^\ell_\beta)_{\ell\geq 0}$ and operators Π^ℓ_β). *Let $\Omega \subset \mathbb{R}^2$ be a bounded polygon with straight sides and corners $\{\mathbf{c}\}$ and $\beta \in [0,1)$. Then there is a sequence $\{\mathcal{T}^\ell_\beta\}_{\ell\geq 0}$ of nested, regular bisection-tree meshes with associated quasi-interpolation operators $\Pi^\ell_\beta : L^2(\Omega) \to S^1_0(\Omega, \mathcal{T}^\ell_\beta)$ such that the following properties hold.*

(i) $N_\ell := \dim S^1_0(\Omega, \mathcal{T}^\ell_\beta) \lesssim h_\ell^{-2}$.

(ii) Simultaneous stability:

$$\|\Pi^\ell_\beta v\|_{L^2(\Omega)} \lesssim \|v\|_{L^2(\Omega)} \qquad \forall v \in L^2(\Omega), \tag{5.15}$$

$$\|\Pi^\ell_\beta v\|_{H^1(\Omega)} \lesssim \|v\|_{H^1(\Omega)} \qquad \forall v \in H^1_0(\Omega). \tag{5.16}$$

(iii) Projection property: $\Pi^\ell_\beta v = v$ *for all* $v \in S^1_0(\Omega, \mathcal{T}^\ell_\beta)$.

(iv) Optimal approximation rates for $H^1_0(\Omega)$ and $H^2_\beta(\Omega)$-functions:

$$N_\ell \|w - \Pi^\ell_\beta w\|^2_{L^2(\Omega)} \lesssim \|w\|^2_{H^1(\Omega)} \tag{5.17}$$

$$N_\ell \|w - \Pi^\ell_\beta w\|^2_{L^2(\Omega)} + \|\nabla_{x'}(w - \Pi^\ell_\beta w)\|^2_{H^1(\Omega)} \lesssim N_\ell^{-1} \|w\|^2_{H^2_\beta(\Omega)}, \tag{5.18}$$

for all $w \in H^1_0(\Omega)$ and all $w \in H^1_0(\Omega) \cap H^2_\beta(\Omega)$, respectively.
In (ii) and (iv), constants hidden in $\lesssim$ are independent of ℓ.

Proof. We divide the proof in two steps.

Step 1: The meshes $\mathcal{T}^\ell_\beta$ are constructed as described in [30] with an appropriate refinement towards the corners $\{\boldsymbol{c}\}$ of Ω. By construction, property (i) holds. Since such a sequence is obtained by "newest vertex bisection" the meshes $\mathcal{T}^\ell_\beta$ are uniformly (in ℓ) shape-regular (see, e.g., [50, Lemma 1]). Additionally, they exhibit the following approximation property

$$\inf_{v\in S^1_0(\Omega,\mathcal{T}^\ell_\beta)} \|w - v\|_{L^2(\Omega)} \lesssim h_\ell \|w\|_{H^1(\Omega)} \lesssim N_\ell^{-1/2} \|w\|_{H^1(\Omega)} \quad \forall w \in H^1_0(\Omega), \tag{5.19}$$

with hidden constants independent of ℓ. In view of the the continuous embedding $H^2_\beta(\Omega) \hookrightarrow C^0(\bar{\Omega})$, the nodal interpolant $Iw \in S^1_0(\Omega, \mathcal{T}^\ell_\beta)$ is well defined. By construction, we thus have that:

$$h_\ell^{-1} \|w - Iw\|_{L^2(\Omega)} + \|\nabla_{x'}(w - Iw)\|_{L^2(\Omega)} \lesssim h_\ell \|w\|_{H^2_\beta(\Omega)}; \tag{5.20}$$

see [47, Section 5].

Step 2: Step 1 shows that the spaces $S^1_0(\Omega, \mathcal{T}^\ell_\beta)$ have the desired approximation properties. The operator Π^ℓ_β can be taken as the $L^2(\Omega)$-projection onto $S^1_0(\Omega, \mathcal{T}^\ell_\beta)$. Then Π^ℓ_β is stable in $L^2(\Omega)$ and, by [29], also stable in $H^1(\Omega)$ so that (ii) holds. The approximation properties (iii), (iv) now follow from the fact that Π^ℓ_β reproduces the space $S^1_0(\Omega, \mathcal{T}^\ell_\beta)$, the simultaneous stability in $L^2(\Omega)$ and $H^1(\Omega)$ as well as the approximation properties (5.19) and (5.20). ∎

REMARK 5.7 (other quasi-interpolants). *In the proof of Lemma 5.6, the L^2-projection may be replaced with Scott-Zhang type quasi-interpolants that are projections onto $S_0^1(\Omega, \mathcal{T}_\beta^\ell)$ and have suitable local stability properties in both L^2 and H^1. Such operators are constructed, e.g., in [7, Lemma 4] by dropping in the classical Scott-Zhang operator [62] the degrees of freedom associated with nodes on $\partial\Omega$ and noting that the remaining operator is well-defined and (locally) stable in $L^2(\Omega)$.* ∎

5.4.2. Linear interpolant π_η^1 on radical-geometric meshes in $[0, \mathcal{Y}]$. To approximate the solution $\mathscr{U}$ with respect to the extended variable y, we shall use a continuous, piecewise linear interpolant on suitably refined meshes $\mathcal{G}_{gr,\eta}^k$ in $[0, \mathcal{Y}]$. The mesh is radical on $[0, 1]$ and geometric on $[1, \mathcal{Y}]$, and the parameter k indicates the mesh size near the point 1. Specifically, for $\mathcal{Y} > 1$, $\eta > 0$, and $k = 1/N$ for an integer $N \in \mathbb{N}$, the mesh $\mathcal{G}_{gr,\eta}^k$ is given by

$$\mathcal{G}_{gr,\eta}^k := \{I_i \,|\, i = 1, \dots, N\} \cup \{J_j \,|\, j = 1, \dots, N'\}, \tag{5.21a}$$

$$I_i = \big[((i-1)k)^\eta, (ik)^\eta\big], \qquad i = 1, \dots, N, \tag{5.21b}$$

$$J_j = \big[\exp((j-1)k), \exp(jk)\big], \qquad j = 1, \dots, N' - 1 := \lfloor N \log \mathcal{Y} \rfloor - 1, \tag{5.21c}$$

$$J_{N'} = \big[\exp((N'-1)k), \mathcal{Y}\big]. \tag{5.21d}$$

Given η and $\mathcal{Y}$, we denote by $\pi_\eta^1 : C((0, \mathcal{Y}]) \to S^1((0, \mathcal{Y}), \mathcal{G}_{gr,\eta}^k)$ the piecewise linear interpolation operator over all the elements of the mesh $\mathcal{G}_{gr,\eta}^k$ with the exception of the first one, i.e., I_1. On that element, π_η^1 corresponds to the linear interpolant in the midpoint of I_1 and the right endpoint of I_1. The operator

$$\pi_{\eta,\{\mathcal{Y}\}}^1 : C((0, \mathcal{Y}]) \to S_{\{\mathcal{Y}\}}^1\left((0, \mathcal{Y}), \mathcal{G}_{gr,\eta}^k\right) \tag{5.22}$$

is obtained from π_η^1 by subtracting a linear function on the element abutting at $\mathcal{Y}$ so as to satisfy $(\pi_{\eta,\{\mathcal{Y}\}}^1 u)(\mathcal{Y}) = 0$. These operators naturally extend to Hilbert space valued functions. The approximation properties of these operators are as follows.

LEMMA 5.8 (interpolation error estimates). *Let X be a Hilbert space, $\alpha \in (-1, 1)$, $\theta \in (0, 1]$, and $0 \le \gamma' < \gamma$. Let the mesh grading parameter η that defines the radical mesh $\mathcal{G}_{gr,\eta}^k$ satisfy $\eta\theta \ge 1$. In this setting the following assertions hold:*

(i) The number of elements in $\mathcal{G}_{gr,\eta}^k$ is bounded by $k^{-1}(1 + \log \mathcal{Y})$.

(ii) For every $u \in C((0, \infty); X)$ with $u' \in L^2(\omega_{\alpha+2(1-\theta),\gamma}, (0, \infty); X)$ we have

$$\|u - \pi_\eta^1 u\|_{L^2(\omega_{\alpha,\gamma'}, (0,\mathcal{Y}); X)} \lesssim k \|u'\|_{L^2(\omega_{\alpha+2(1-\theta),\gamma}, (0,\mathcal{Y}); X)}, \tag{5.23}$$

$$\begin{aligned} \|u - \pi_{\eta,\{\mathcal{Y}\}}^1 u\|_{L^2(\omega_{\alpha,\gamma'}, (0,\mathcal{Y}); X)} &\lesssim k \|u'\|_{L^2(\omega_{\alpha+2(1-\theta),\gamma}, (0,\mathcal{Y}); X)} \\ &\quad + \sqrt{\mathcal{Y} k}\, \mathcal{Y}^{\alpha/2} \exp(\mathcal{Y}\gamma'/2) \|u(\mathcal{Y})\|_X. \end{aligned} \tag{5.24}$$

Furthermore, under the assumption that $\lim_{y\to\infty} u(y) = 0$ in X and the constraint

$$\mathcal{Y}^{-1/2+\theta} \exp(-\mathcal{Y}\gamma/2) \le k^{1/2} \tag{5.25}$$

the following estimate holds:

$$\|u - \pi_{\eta,\{\mathcal{Y}\}}^1 u\|_{L^2(y^\alpha, (0,\mathcal{Y}); X)} \lesssim k \|u'\|_{L^2(\omega_{\alpha+2(1-\theta),\gamma}, (0,\infty); X)}. \tag{5.26}$$

(iii) For $u \in C((0,\infty);X)$ with $u'' \in L^2(\omega_{\alpha+2(1-\theta),\gamma},(0,\infty);X)$ and $j \in \{0,1\}$

$$\|(u-\pi_\eta^1 u)^{(j)}\|_{L^2(\omega_{\alpha,\gamma'},(0,\mathcal{Y});X)} \lesssim k^{2-j}\|u''\|_{L^2(\omega_{\alpha+2(1-\theta),\gamma}(0,\mathcal{Y});X)}, \tag{5.27}$$

$$\begin{aligned}\|(u-\pi_{\eta,\{\mathcal{Y}\}}^1 u)^{(j)}\|_{L^2(\omega_{\alpha,\gamma'},(0,\mathcal{Y});X)} &\lesssim k^{2-j}\|u''\|_{L^2(\omega_{\alpha+2(1-\theta),\gamma}(0,\mathcal{Y});X)} \\ &\quad + (\mathcal{Y}k)^{1/2-j}\mathcal{Y}^{\alpha/2}\exp(\mathcal{Y}\gamma'/2)\|u(\mathcal{Y})\|_X.\end{aligned} \tag{5.28}$$

Furthermore, under the assumption that, for $j \in \{0,1\}$, we have that

$$\lim_{y\to\infty} u^{(j)}(y) = 0 \qquad \text{in } X,$$

and, under the constraint

$$\mathcal{Y}^{-1/2+\theta}\exp(-\mathcal{Y}\gamma/2) \le k^{3/2} \tag{5.29}$$

the following estimate holds for $j \in \{0,1\}$:

$$\|(u-\pi_{\eta,\{\mathcal{Y}\}}^1 u)^{(j)}\|_{L^2(y^\alpha,(0,\mathcal{Y});X)} \lesssim k^{2-j}\|u''\|_{L^2(\omega_{\alpha+2(1-\theta),\gamma},(0,\infty);X)}. \tag{5.30}$$

Proof. We present the details for the proof of (ii), as that of (iii) is similar. The technique used to obtain interpolation error estimates on the radical mesh on $[0,1]$ is well-established; see, for instance, [59, Example 3.47]. We introduce the mesh points $y_i := (ik)^\eta$, $i = 0,\dots,N$ so that $I_i = [y_{i-1},y_i]$.

For the first element $I_1 = [y_0,y_1] = [0,k^\eta]$, we invoke the estimate (A.3) with the choice $\delta = 1-\theta \in [0,1)$ and a scaling argument to conclude that

$$\|u-\pi_\eta^1 u\|^2_{L^2(y^\alpha,I_1;X)} \lesssim k_1^{2\theta}\|u'\|^2_{L^2(y^{\alpha+2(1-\theta)},I_1;X)}, \tag{5.31}$$

where $k_1 = |I_1| = k^\eta$; we recall that $\theta\eta \ge 1$.

For the remaining elements I_i, $i = 2,\dots,N$, of $[0,1]$, we use that $k_i \lesssim k y_{i-1}^{(\eta-1)/\eta}$, where $k_i = |I_i| = y_i - y_{i-1}$ and η defines the radical mesh on $[0,1]$ in (5.21b). Recalling the standard interpolation estimate

$$\|u-\pi_\eta^1 u\|^2_{L^2(I_i)} \lesssim k_i^2\|u'\|^2_{L^2(I_i)}, \tag{5.32}$$

we obtain, upon using that $\max_{y\in I_i} y^\alpha \lesssim \min_{y\in I_i} y^\alpha$ and tensorization with X, the bound

$$\begin{aligned}\|u-\pi_\eta^1 u\|^2_{L^2(y^\alpha,I_i;X)} &\lesssim k_i^2\|u'\|^2_{L^2(y^\alpha,I_i;X)} \lesssim k^2 y_{i-1}^{2(\eta-1)/\eta}\|u'\|^2_{L^2(y^\alpha,I_i;X)} \\ &\lesssim k^2\|u'\|^2_{L^2(y^{\alpha+2(\eta-1)/\eta},I_i,;X)} \lesssim k^2\|u'\|^2_{L^2(y^{\alpha+2(1-\theta)},I_i;X)}\,.\end{aligned} \tag{5.33}$$

The last relation holds because $\eta\theta \ge 1$.

For the elements beyond $y = 1$, we abbreviate $J_j := [\widetilde{y}_{j-1},\widetilde{y}_j] := \big[\exp((j-1)k,\exp(jk)\big]$ for $j = 1,\dots,N'$. Notice that, since $k \le 1$,

$$|J_j| = \exp((j-1)k)(1-e^k) \sim \widetilde{y}_{j-1}k, \qquad j = 1,\dots,N'-1\,. \tag{5.34}$$

Using that the weight functions $\omega_{\alpha,\gamma'}$ and $\omega_{\alpha,\gamma}$, defined in (4.10), are slowly varying over the intervals J_j, i.e.,

$$\max_{y\in J_j}\omega_{\alpha,\gamma'}(y) \lesssim \min_{y\in J_j}\omega_{\alpha,\gamma'}(y) \quad \text{and} \quad \max_{y\in J_j}\omega_{\alpha,\gamma}(y) \lesssim \min_{y\in J_j}\omega_{\alpha,\gamma}(y), \tag{5.35}$$

we obtain

$$\sum_j \|u-\pi^1_\eta u\|^2_{L^2(\omega_{\alpha,\gamma'},J_j;X)} \overset{(5.32)}{\lesssim} \sum_j |J_j|^2 \|u'\|^2_{L^2(\omega_{\alpha,\gamma'},J_j;X)}$$
$$\overset{(5.34),(5.35)}{\lesssim} k^2 \sum_j \widetilde{y}^2_{j-1} e^{-(\gamma-\gamma')\widetilde{y}_{j-1}} \|u'\|^2_{L^2(\omega_{\alpha,\gamma},J_j;X)}.$$

The estimate $\widetilde{y}^2_{j-1} e^{-(\gamma-\gamma')\widetilde{y}_{j-1}} \lesssim 1$ implies

$$\sum_j \|u-\pi^1_\eta u\|^2_{L^2(\omega_{\alpha,\gamma'},J_j;X)} \lesssim k^2 \sum_j \|u'\|^2_{L^2(\omega_{\alpha,\gamma},J_j;X)}. \tag{5.36}$$

Combining (5.31), (5.33), and (5.36) finishes the proof of the approximation properties of π^1_η. The correction on the last element to obtain (5.24) for the operator $\pi^1_{\eta,\{\mathcal{Y}\}}$ is straightforward in view of (5.34). The estimate (5.26) follows from (5.24) by controlling $\|u(\mathcal{Y})\|_X$ with the aid of Lemma A.2. □

It is worth stressing that the choices $k = 2^{-\ell}$, $\ell = 0, 1, \ldots$, lead to nested meshes:

COROLLARY 5.9 (nested meshes). *For every fixed $\eta \geq 0$, $\mathcal{Y} \geq 1$ and for $k_\ell = 2^{-\ell}$, the sequence $\{\mathcal{G}^{k_\ell}_{gr,\eta}\}_{\ell=0}^\infty$ of graded meshes in $(0,\mathcal{Y})$ is nested and each $\mathcal{G}^{k_\ell}_{gr,\eta}$ has $\mathcal{O}(2^\ell(1+\log\mathcal{Y}))$ elements.*

Proof. For fixed $\mathcal{Y} > 0$, it follows directly from the definition of the mesh points (5.21), in terms of k, that the meshes are nested. □

5.4.3. Tensor P_1-FEM in $\mathcal{C}$ with corner mesh refinement in Ω. We now derive a rate of convergence for corner-refined meshes in, not necessarily convex, polygons.

THEOREM 5.10 (error estimates). *Let $\Omega \subset \mathbb{R}^2$ be a bounded polygon with a finite set of corners $\{\boldsymbol{c}\}$. Let $u \in \mathbb{H}^s(\Omega)$ and $\mathscr{U} \in \mathring{H}^1(y^\alpha,\mathcal{C})$ solve (1.1) and (1.2), respectively, with $f \in \mathbb{H}^{1-s}(\Omega)$ and $s \in (0,1)$. Let $\beta \in [0,1)$ be such that (5.12) holds. Let further $\{\mathcal{T}^\ell_\beta\}_{\ell\geq 0}$ denote a sequence of uniformly shape-regular meshes in Ω that satisfy (5.17) and (5.18). Let $\mathcal{G}^k_{gr,\eta}$ be the graded–exponential mesh in $(0,\mathcal{Y})$ defined in (5.21) with the parameter η chosen to satisfy $\eta s > 1$, $k = 1/N$, and with $N \in \mathbb{N}$ chosen so that $h_\ell/2 \leq k \leq h_\ell$, and with the cut-off parameter $\mathcal{Y} > 0$ chosen as*

$$\mathcal{Y} \sim |\log h_\ell|\ . \tag{5.37}$$

Assume that the constant hidden in (5.37) is fixed sufficiently large, independent of ℓ. For every $\ell, M \in \mathbb{N}$, denote by $\mathscr{U}_{h_\ell,M} \in \mathbb{V}^{1,1}_{h_\ell,M}(\mathcal{T}^\ell_\beta,\mathcal{G}^k_{gr,\eta})$ the solution of (5.6). Then we have the following error estimate

$$\|u - \mathrm{tr}_\Omega\, \mathscr{U}_{h_\ell,M}\|_{\mathbb{H}^s(\Omega)} \lesssim \|\nabla(\mathscr{U} - \mathscr{U}_{h_\ell,M})\|_{L^2(y^\alpha,\mathcal{C})} \lesssim h_\ell \|f\|_{\mathbb{H}^{1-s}(\Omega)} \tag{5.38}$$

The total number of degrees of freedom satisfies

$$\mathcal{N}_{\Omega,\mathcal{Y}} := \dim \mathbb{V}^{1,1}_{h_\ell,M}(\mathcal{T}^\ell_\beta,\mathcal{G}^k_{gr,\eta}) = \mathcal{O}(h_\ell^{-3}\log|\log h_\ell|) = \mathcal{O}(\mathcal{N}_\Omega^{1+1/2}\log\log\mathcal{N}_\Omega)\ , \tag{5.39}$$

where $\mathcal{N}_\Omega = \#\mathcal{T}^\ell_\beta$.

Before proving Theorem 5.10, we note a corollary that follows from a simple interpolation argument.

COROLLARY 5.11 (reduced regularity). *Assume that the meshes satisfy the conditions of Theorem 5.10 and that $f \in \mathbb{H}^{-s+t}(\Omega)$ with $t \in [0,1]$. Then we have*

$$\|u - \mathrm{tr}_\Omega\, \mathscr{U}_{h_\ell,M}\|_{\mathbb{H}^s(\Omega)} \lesssim \|\nabla(\mathscr{U} - \mathscr{U}_{h_\ell,M})\|_{L^2(y^\alpha,\mathcal{C})} \lesssim h^t_\ell \|f\|_{\mathbb{H}^{t-s}(\Omega)}, \tag{5.40}$$

where the hidden constant additionally depends on t.

Proof of Theorem 5.10: The proof relies on the stability and approximation properties (5.18) of Π_ℓ^β, using arguments similar to those employed in [48], [40, Section 4.1]. For completeness we provide the details.

A sequence $\{\mathcal{T}_\beta^\ell\}_{\ell \geq 0}$ (and associated approximation operators Π_β^ℓ) as required in the statement of Theorem 5.10 has been constructed in Lemma 5.6. That sequence of meshes is even nested. However, nestedness is not essential in the ensuing arguments. The essential ingredients are the approximation properties (5.17), (5.18) and the existence of an operator Π_β^ℓ that is simultaneously stable in $L^2(\Omega)$ and $H^1(\Omega)$; shape-regularity of the meshes ensures the existence of such operators by Remark 5.7. For a given choice of k, η, and $\mathcal{Y}$, we denote by $\pi_{\eta,\{\mathcal{Y}\}}^{1,\ell}$ the nodal interpolation operator on the mesh (5.21), which we analyzed in Lemma 5.8. By Lemmas 5.1 and 5.2, and by the choice (5.37) (recall (5.8)) it suffices to bound

$$\|\nabla(\mathscr{U} - \pi_{\eta,\{\mathcal{Y}\}}^{1,\ell}\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} + \|\nabla(\mathscr{U} - \Pi_\beta^\ell\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} =: I + II\ .$$

Recalling that $\nabla = (\nabla_{x'}, \partial_y)$ we split the first term I into

$$I \lesssim \|\partial_y(\mathscr{U} - \pi_{\eta,\{\mathcal{Y}\}}^{1,\ell}\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} + \|\nabla_{x'}(\mathscr{U} - \pi_{\eta,\{\mathcal{Y}\}}^{1,\ell}\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} =: I_a + I_b\ .$$

In view of (5.37) (with implied constant sufficiently large) we immediately obtain that the conditions (5.25) and (5.29) of Lemma 5.8 are satisfied. Since $\eta s > 1$, we can therefore bound the term I_a using Lemma 5.8, item (iii), with $j = 1$ and $X = L^2(\Omega)$ and the term I_b using Lemma 5.8, item (ii) with $X = H_0^1(\Omega)$. Together with the regularity estimates of Theorem 4.7 we have thus arrived at

$$I \lesssim I_a + I_b \lesssim h_\ell \|f\|_{\mathbb{H}^0(\Omega)}.$$

We apply the same splitting to the term II to obtain

$$II \lesssim \|\partial_y(\mathscr{U} - \Pi_\beta^\ell\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} + \|\nabla_{x'}(\mathscr{U} - \Pi_\beta^\ell\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} =: II_a + II_b\ .$$

Since $N_\ell = \mathcal{O}(h_\ell^{-2})$ we have, from (5.17), that

$$II_a = \|\partial_y\mathscr{U} - \Pi_\beta^\ell(\partial_y\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \lesssim h_\ell \|\partial_y\mathscr{U}\|_{L^2(\omega_{\alpha,0},(0,\mathcal{Y});H^1(\Omega))}\ ,$$

which can be controlled with the aid of (4.18) with $\nu = 1$ there. To bound II_b we use (5.18) and get

$$II_b \lesssim h_\ell \|\mathscr{U}\|_{L^2(\omega_{\alpha,0},(0,\mathcal{Y});H_\beta^2(\Omega))}\ .$$

Using the regularity estimate (5.13) with $\nu' = 0$ we conclude the proof of (5.38).

To obtain (5.39), we first note that by Lemma 5.8 item (i), the number of elements in $\mathcal{G}_{gr,\eta}^k$ with $h_\ell = 2^{-\ell}$ and with the choice $\mathcal{Y} \simeq |\log h_\ell| \simeq \ell$ is $\mathcal{O}(2^\ell \log \ell)$. We finally observe that the total number of degrees of freedom in the tensor product space is the product of the dimensions of the component spaces, i.e., $\mathcal{O}(h_\ell^{-2} h_\ell^{-1} \log|\log h_\ell|)$. $\square$

5.4.4. Sparse grid P_1-FEM with corner mesh refinement. For the approximation with piecewise linears and for $f \in \mathbb{H}^{1-s}(\Omega)$, the convergence order in (5.38) is optimal. However, the complexity of the method implied by (5.39) is superlinear with respect to the number of degrees of freedom $\mathcal{N}_\Omega$ in Ω. To reduce the complexity to nearly linear, we develop a sparse tensor product approach in what follows. It is based on the subspace hierarchies

$$\{S^1_0(\Omega, \mathcal{T}^\ell_\beta)\}_{\ell \geq 0}\ , \quad \{S^1_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^{2^{-\ell'}}_{gr,\eta})\}_{\ell' \geq 0},$$

where $\{\mathcal{T}^\ell_\beta\}_{\ell\geq 0}$ is a nested sequence of bisection–tree meshes in Ω which are β-graded toward the corners $\{\boldsymbol{c}\}$ in such a way that first-order convergence in $h_\ell = \mathcal{O}(2^{-\ell})$ is achieved for H^2_β functions; such meshes are provided by Lemma 5.6. The sequence $\{\mathcal{G}^{2^{-\ell'}}_{gr,\eta}\}_{\ell'\geq 0}$ consists of nested graded meshes on $[0,\mathcal{Y}]$ that achieve, for functions belonging to weighted H^2-spaces in $(0,\mathcal{Y})$, as introduced in Theorem 4.7, first order convergence (cf. the precise statements in Lemma 5.8 and in Corollary 5.9).

For $\ell, \ell' \geq 0$, we denote by

$$\Pi^\ell_\beta : L^2(\Omega) \to S^1_0(\Omega, \mathcal{T}^\ell_\beta) \text{ and } \pi^{1,\ell'}_{\eta,\{\mathcal{Y}\}} : C((0,\mathcal{Y}]) \to S^1_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^{2^{-\ell'}}_{gr,\eta})$$

the corresponding (quasi)interpolatory projections introduced in Lemma 5.6 and formula (5.22), respectively. Set in addition $\Pi^{-1}_\beta := 0$ and $\pi^{1,-1}_{\eta,\{\mathcal{Y}\}} := 0$. Then, for $L \in \mathbb{N}_0$, we define the *sparse tensor product space* as

$$\hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathcal{Y}) = \sum_{\ell,\ell'\geq 0, \ell+\ell'\leq L} S^1_0(\Omega, \mathcal{T}^\ell_\beta) \otimes S^1_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^{2^{-\ell'}}_{gr,\eta})\ . \tag{5.41}$$

We immediately flag that the sum in (5.41) is not direct (cf. Remark 5.12 below). By zero extension we have $\hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathcal{Y}) \subset \mathring{H}^1(y^\alpha, \mathcal{C})$.

We define the approximation $\hat{\mathscr{U}}_L \in \hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathcal{Y})$ as the solution to (5.6) with $\hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathcal{Y})$ taking the role of $\mathbb{V}_{h,M}$ there.

Remark 5.12 (implementation). *The computation of the sparse tensor FE approximation $\hat{\mathscr{U}}_L \in \hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathcal{Y})$ by directly evaluating* (5.6) *would require an explicit representation of the sparse tensor product subspace $\hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathcal{Y})$ and therefore, in particular, an explicit basis for the "increment spaces" in* (5.41)*, i.e., for the complements of $S^1_0(\Omega, \mathcal{T}^{\ell-1}_\beta)$ in $S^1_0(\Omega, \mathcal{T}^\ell_\beta)$ and the complements of $S^1_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^{2^{-(\ell'-1)}}_{gr,\eta})$ in $S^1_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^{2^{-\ell'}}_{gr,\eta})$. Construction of bases for the increment spaces is possible, based on ideas from multiresolution analyses. We opt, instead, to compute $\hat{\mathscr{U}}_L \in \hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathcal{Y})$ from the so-called* combination formula *(see, e.g., [33, Section 4.2, Equation (4.6)]). It is based on anisotropic $\mathring{H}^1(y^\alpha, \mathcal{C})$-Galerkin projections*

$$G^{1,1}_{\ell,\ell'} := G^1_{h_\ell} \circ G^1_{2^{-\ell'}} : \mathring{H}^1(y^\alpha; \mathcal{C}) \to \mathbb{V}^{1,1}_{h,M}(\mathcal{T}_\ell, \mathcal{G}^{2^{-\ell'}}_{gr,\eta})\ , \tag{5.42}$$

with the semidiscrete projections defined in (5.5)*. The projectors $G^{1,1}_{\ell,\ell'}$ in* (5.42) *can be realized with standard FE bases in Ω and in $(0,\mathcal{Y})$. The combination formula then takes the form*

$$\hat{\mathscr{U}}_L \quad = \quad \sum_{\ell=0}^{L} (\mathscr{U}_{\ell,L-\ell} - \mathscr{U}_{\ell-1,L-\ell})\ ,$$

where $\mathscr{U}_{\ell,\ell'} := G^{1,1}_{\ell,\ell'}\mathscr{U}$ *and* $\mathscr{U}_{-1,j} = 0$ *for* $j \in \mathbb{N}_0$,

The convergence of our sparse grids scheme is the content of the next result.

Theorem 5.13 (convergence for sparse grids). *Let* $\Omega \subset \mathbb{R}^2$ *be a polygon and let* $\beta \in [0,1)$ *be such that* (5.12) *holds. Let* $s \in (0,1)$ *and* $1 < \nu < 1+s$. *Let* $f \in \mathbb{H}^{-s+\nu}(\Omega)$ *and* $\mathscr{U} \in \mathring{H}^1(y^\alpha,\mathcal{C})$ *solve* (1.2). *Let* $\eta(\nu-1) \geq 1$ *and select* $\mathscr{Y} \sim |\log h_L|$ *with a sufficiently large hidden constant that is independent of* ℓ. *Then the sparse tensor product space* $\hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathscr{Y})$ *of* (5.41) *and the corresponding Galerkin approximation* $\hat{\mathscr{U}}_L \in \hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathscr{Y})$ *to* $\mathscr{U}$ *satisfy*

$$\|\nabla(\mathscr{U} - \hat{\mathscr{U}}_L)\|_{L^2(y^\alpha,\mathcal{C})} \lesssim h_L|\log h_L|\|f\|_{\mathbb{H}^{-s+\nu}(\Omega)}, \tag{5.43}$$

$$\dim \hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathscr{Y}) \lesssim \mathcal{N}_\Omega \log\log \mathcal{N}_\Omega. \tag{5.44}$$

Proof. Before proving the theorem, we recall that a sequence $\{\mathcal{T}^\ell_\beta\}_{\ell\geq 0}$ of meshes with corresponding approximation operators Π^ℓ_β is constructed in Lemma 5.6. We begin by proving (5.44). From the condition $\mathscr{Y} \sim |\log h_L| \sim L$, we have, by Lemma 5.8, item (i), that $\#(\mathcal{G}^{2^{-\ell'}}_{gr,\eta}) \lesssim 2^{\ell'}|\log h_L| \sim 2^{\ell'}\log L$. Consequently,

$$\dim \hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathscr{Y}) \lesssim \sum_{\ell,\ell'\geq 0, \ell+\ell'\leq L} 2^{2\ell+\ell'}|\log L| \lesssim 2^{2L}\log L \sim \mathcal{N}_\Omega \log\log \mathcal{N}_\Omega\ , \tag{5.45}$$

where we have also used that $N_\ell = \dim(S^1_0(\Omega,\mathcal{T}^\ell_\beta)) \sim 2^{2\ell}$.

To prove (5.43), we next study the error of our method. From Lemma 5.1 and (5.8) it suffices to study the best approximation error in $\mathbb{V}^{1,1}_L(\mathcal{C}_\mathscr{Y})$. To do so, we introduce the *sparse tensor product interpolation* projector

$$\hat{\Pi}^L_\mathscr{Y} : C((0,\mathscr{Y}];L^2(\Omega)) \to \hat{\mathbb{V}}^{1,1}_L(\mathcal{C}_\mathscr{Y})$$

which is defined by

$$\hat{\Pi}^L_\mathscr{Y} w := \sum_{\ell,\ell'\geq 0, \ell+\ell'\leq L} (\Pi^\ell_\beta - \Pi^{\ell-1}_\beta) \otimes (\pi^{1,\ell'}_{\eta,\{\mathscr{Y}\}} - \pi^{1,\ell'-1}_{\eta,\{\mathscr{Y}\}})w\ . \tag{5.46}$$

It is convenient to introduce the operators

$$Q^\ell_\beta := \Pi^\ell_\beta - \Pi^{\ell-1}_\beta\ , \quad q^{1,\ell'}_\eta := \pi^{1,\ell'}_{\eta,\{\mathscr{Y}\}} - \pi^{1,\ell'-1}_{\eta,\{\mathscr{Y}\}}\ .$$

As in the proof of Theorem 5.10, we split the error into

$$\begin{aligned} \min_{\hat{v}_L\in\hat{\mathbb{V}}_L} \|\nabla(\mathscr{U} - \hat{v}_L)\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} &\lesssim \|\partial_y(\mathscr{U} - \hat{\Pi}^L_\mathscr{Y}\mathscr{U})\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} \\ &\quad + \|\nabla_{x'}(\mathscr{U} - \hat{\Pi}^L_\mathscr{Y}\mathscr{U})\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} =: I + II\ . \end{aligned} \tag{5.47}$$

Each one of these terms can now be bounded in the usual sparse grid fashion, *provided* that $\mathscr{U}$ has so-called mixed regularity, which is indeed the case by Theorems 4.7, 5.5.

Let us bound term I in (5.47). From the estimate (4.18) of Theorem 4.7 we infer

$$\|\partial^2_y\mathscr{U}\|_{L^2(\omega_{\alpha+2(2-\nu),\gamma},(0,\infty);H^1(\Omega))} \lesssim \|f\|_{\mathbb{H}^{-s+\nu}(\Omega)}, \qquad 0 \leq \nu < 1+s. \tag{5.48}$$

Of interest to us is the case $1 < \nu < 1+s < 2$. Then, with the mesh grading parameter η satisfying $\eta(-1+\nu) \geq 1$ and upon assuming that $\mathcal{Y} \geq CL$ for $C > 0$ sufficiently large so that the condition (5.29) is satisfied we estimate

$$\begin{aligned} I &\leq \sum_{\ell+\ell'>L} \|\partial_y (Q_\beta^\ell \otimes q_y^{1,\ell'} \mathscr{U})\|_{L^2(y^\alpha, \mathcal{C}_\mathcal{Y})} \\ &\leq \sum_{\ell+\ell'>L} \|\partial_y [((I_{x'} \otimes q_y^{1,\ell'}) \circ (Q_\beta^\ell \otimes I_y)\mathscr{U}]\|_{L^2(y^\alpha, \mathcal{C}_\mathcal{Y})} \\ &\lesssim \sum_{\ell+\ell'>L} 2^{-\ell} \|\partial_y [(I_{x'} \otimes q_y^{1,\ell'})\mathscr{U}]\|_{L^2(y^\alpha,(0,\mathcal{Y});H^1(\Omega))}, \end{aligned}$$

where, in the last step, we used the approximation property (5.17). We now apply the estimate (5.30) with $j = 0$, $\theta = \nu - 1$ and $X = H^1(\Omega)$, to arrive at

$$I \lesssim \sum_{\ell+\ell'>L} 2^{-\ell-\ell'} \|\partial_y^2 \mathscr{U}\|_{L^2(\omega_{\alpha+2(2-\nu),\gamma},(0,\mathcal{Y});H^1(\Omega))} \lesssim L 2^{-L} \|f\|_{\mathbb{H}^{-s+\nu}(\Omega)} ,$$

where in the last step we have used the regularity estimate (5.48).

Let us now bound, using similar arguments, the term II in (5.47). From (5.13) and (5.14) we obtain, for $1 \leq \nu < 2 - s$, the regularity estimate

$$\|\partial_y \mathscr{U}\|_{L^2(\omega_{\alpha+2(2-\nu),\gamma},(0,\infty);H^2_\beta(\Omega))} \lesssim \|f\|_{\mathbb{H}^{-s+\nu}(\Omega)}. \tag{5.49}$$

Hence, for $\eta(-1+\nu) \geq 1$, and again under the condition that $\mathcal{Y} \geq CL$ so that (5.25) is satisfied, we can estimate

$$\begin{aligned} II &\leq \sum_{\ell+\ell'>L} \|\nabla_{x'} (Q_\beta^\ell \otimes q_y^{1,\ell'} \mathscr{U})\|_{L^2(y^\alpha, \mathcal{C}_\mathcal{Y})} \\ &\leq \sum_{\ell+\ell'>L} \|\nabla_{x'} [((I_{x'} \otimes q_y^{1,\ell'}) \circ (Q_\beta^\ell \otimes I_y)\mathscr{U}]\|_{L^2(y^\alpha, \mathcal{C}_\mathcal{Y})} \\ &\lesssim \sum_{\ell+\ell'>L} 2^{-\ell} \|(I_{x'} \otimes q_y^{1,\ell'})\mathscr{U}\|_{L^2(y^\alpha,(0,\mathcal{Y});H^2_\beta(\Omega))}, \end{aligned}$$

where in the last step we used the approximation properties of Π_β^ℓ, as stated in (5.18). The approximation properties of $\pi_{\eta,\{\mathcal{Y}\}}^{1,\ell'}$ given in (5.26) with the regularity estimate of (5.49) allow us to conclude that

$$II \lesssim \sum_{\ell+\ell'>L} 2^{-\ell-\ell'} \|\partial_y \mathscr{U}\|_{L^2(\omega_{\alpha+2(2-\nu),\gamma},(0,\mathcal{Y});H^2_\beta(\Omega))} \lesssim L 2^{-L} \|f\|_{\mathbb{H}^{-s+\nu}(\Omega)} .$$

Collecting the bounds obtained for I and II yields the result. □

Theorem 5.13 shows that it is possible to obtain near optimal order convergence for fractional diffusion in Ω, by using only P_1-FEM in both Ω and the extended dimension. An alternative approach is based on exploiting *analytic regularity* of the solution of the extended problem. In this case, *exponentially convergent hp-FEM with respect to the extended variable y* will achieve near optimal order for conforming P_1-FEM in Ω, as observed recently in [40], and, as we show (by a different argument) in Theorem 5.16 of Section 5.5.

5.5. *hp*-FEM in $(0,\infty)$ and P_1-FEM in Ω. The discretizations in the preceding Sections 5.4.4 and 5.4.3 were of first order in x' and y. We showed that full tensor product FEM can lead to first order convergence in Ω at the expense of superlinear complexity (5.39). Here, we address the use of the so-called *hp*-FEM in $(0,\mathcal{Y})$; the analytic regularity estimates derived in Section 4 allow us to prove *exponential convergence estimates* for corresponding high-order discretizations in $(0,\mathcal{Y})$. We consider two situations:

a) The case where $\boldsymbol{r}$ is a so-called linear degree vector in $(0,\mathcal{Y})$, which will imply exponential convergence with respect to y (cf. Lemma 6.2 below). If fixed order FEM on a sequence $\{\mathcal{T}_\beta^\ell\}_{\ell\geq 0}$ of regular, simplicial corner-refined meshes in Ω are used, near optimal, algebraic convergence rates (with respect to the number $\mathcal{N}_\Omega$ of degrees of freedom in Ω) result for the solution of (1.1) in Ω (Theorem 5.16). We mention [34] where, in a structurally similar context, analyticity in the extended variable y is also exploited by an *hp*-FEM.
b) The case where $\boldsymbol{r}$ is a linear degree vector in $(0,\mathcal{Y})$, and where we use the *hp*-FEM in Ω; in this case, and under the additional assumption (7.1) of analyticity on the data c, f, A, *exponential convergence in terms of the number $\mathcal{N}_{\Omega,\mathcal{Y}}$ of degrees of freedom in $\mathcal{C}_\mathcal{Y}$* can be achieved. We confine the exposition to the case $\Omega=(0,1)$ and to $\Omega\subset\mathbb{R}^2$ with analytic boundary. This will be the content of Section 7.

5.5.1. Univariate *hp*-interpolation operator. We present here the construction of a univariate interpolation operator that leads to exponential convergence for analytic functions that may have a singularity at $y=0$. The construction is essentially taken from the work by Babuška and collaborators, [32, 8] and discussed in the literature on *hp*-FEM (see, e.g., [59, Sec. 4.4.1], [6, Thm. 8] and also [40]).

To make matters precise, we consider geometric meshes $\mathcal{G}^M_{geo,\sigma}$ on $[0,\mathcal{Y}]$ with M elements and grading factor $\sigma\in(0,1)$: $\{I_i\,|\,i=1,\dots,M\}$ with $I_1=[0,\mathcal{Y}\sigma^{M-1}]$ and $I_i=[\mathcal{Y}\sigma^{M-i+1},\mathcal{Y}\sigma^{M-i}]$ for $i=2,\dots,M$. On such meshes, we consider a *linear degree vector* $\boldsymbol{r}$ with slope $\mathfrak{s}$ given by

$$r_i := 1+\lfloor\mathfrak{s}(i-1)\rfloor\,,\quad i=1,2,...,M\,. \tag{5.50}$$

We denote by $\widehat{K}=(-1,1)$ the reference interval. We will require a base interpolation operator $\widehat{\Pi}_r$ that allows for exponential convergence in r for analytic functions:

LEMMA 5.14 (polynomial approximation operator $\widehat{\Pi}_r$). *There exists a linear operator $\widehat{\Pi}_r : H^1(\widehat{K})\to\mathbb{P}_r(\widehat{K})$ with the following properties:*

1. *$(\widehat{\Pi}_r\widehat{u})(\pm1)=\widehat{u}(\pm1)$ for all $\widehat{u}\in H^1(\widehat{K})$.*
2. *For every $K_u>0$ there exist $C'=C(K_u)$, $b=b(K_u)>0$ such that if, for all $\ell\in\mathbb{N}_0$, we have $\|\widehat{u}^{(\ell)}\|_{L^2(\widehat{K})}\leq C_uK_u^{\ell+1}(\ell+1)!$ then*

$$\|\widehat{u}-\widehat{\Pi}_r\widehat{u}\|_{H^1(\widehat{K})}\leq C'C_ue^{-br}\qquad\forall r\in\mathbb{N}.$$

Proof. Classical examples of such operators include the Gauss-Lobatto interpolation operator and the "Babuška-Szabó operator" Π_r^{BS} as described, e.g., in the survey [6, Example 13] or in [59, Theorem 3.14]. ☐

With the aid of $\widehat{\Pi}_r$ we introduce the operators $\pi_y^{\boldsymbol{r}}$ and $\pi_{y,\{\mathcal{Y}\}}^{\boldsymbol{r}}$ on an arbitrary mesh $\mathcal{G}^M$ on $[0,\mathcal{Y}]$ with M elements and polynomial degree distribution $\boldsymbol{r}\in\mathbb{N}^M$ in an element-by-element fashion in the usual way below. However, for $\pi_y^{\boldsymbol{r}}$ we modify the approximation on the first element $I_1=[0,y_1]$ by interpolating in the points $y_1/2$ and y_1 instead of the endpoints. The operator $\pi_{y,\{\mathcal{Y}\}}^{\boldsymbol{r}}$ is obtained by a further modification that enforces $\pi^{\boldsymbol{r}}_{y,\{\mathcal{Y}\}}(\mathcal{Y}) = 0$. Specifically, with $F_{I_i} : \widehat{K} \to I_i$ denoting the affine, orientation-preserving element maps for element $I_i \in \mathcal{G}^M$ we have

$$
\begin{aligned}
((\pi^{\boldsymbol{r}}_y u)|_{I_1} \circ F_{I_1})(\xi) &= 2(u \circ F_{I_1})(1)(\xi - 1/2) + 2(u \circ F_{I_1})(1/2)(1-\xi),\\
((\pi^{\boldsymbol{r}}_y u)|_{I_i} \circ F_{I_i})(\xi) &= \widehat{\Pi}_{\boldsymbol{r}_m}(u \circ F_{I_i}), \qquad i = 2,\dots,M,\\
(\pi^{\boldsymbol{r}}_{y,\{\mathcal{Y}\}} u)|_{I_i} &= (\pi^{\boldsymbol{r}}_y u)|_{I_i}, \qquad i = 1,\dots,M-1,\\
((\pi^{\boldsymbol{r}}_{y,\{\mathcal{Y}\}} u)|_{I_M} \circ F_{I_M})(\xi) &= ((\pi^{\boldsymbol{r}}_y u)|_{I_M} \circ F_{I_M})(\xi) - (u \circ F_{I_M})(1)(\xi+1)/2 \,.
\end{aligned}
$$

The definition of $\pi^{\boldsymbol{r}}_y$, $\pi^{\boldsymbol{r}}_{y,\{\mathcal{Y}\}}$ is naturally extended for functions $u \in C^0((0,\mathcal{Y}];X)$, where X denotes a Hilbert space. We will apply these operators to functions from the following two classes of analytic functions of the extended variable y:

$$
\begin{aligned}
\mathcal{B}^1_{\beta,\gamma}(C_u,K_u;X) := \big\{&u \in C^\infty((0,\infty);X) \,:\, \|u\|_{L^2(\omega_{\alpha,\gamma},(0,\infty);X)} < C_u,\\
&\|u^{(\ell+1)}\|_{L^2(\omega_{\alpha+2(\ell+1)-2\beta,\gamma},(0,\infty);X)} < C_u K_u^{\ell+1}(\ell+1)! \quad \forall \ell \in \mathbb{N}_0\big\}
\end{aligned}
\tag{5.51}
$$

and

$$
\begin{aligned}
\mathcal{B}^2_{\beta,\gamma}(C_u,K_u,X) := \big\{&u \in C^\infty((0,\infty);X) \,:\\
&\|u\|_{L^2(\omega_{\alpha,\gamma},(0,\infty);X)} + \|u'\|_{L^2(\omega_{\alpha,\gamma},(0,\infty);X)} \le C_u,\\
&\|u^{(\ell+2)}\|_{L^2(\omega_{\alpha+2(\ell+1)-2\beta},\gamma,(0,\infty);X)} \le C_u K_u^{\ell+2}(\ell+2)! \quad \forall \ell \in \mathbb{N}_0\big\}.
\end{aligned}
\tag{5.52}
$$

We recall that the weight $\omega_{\beta,\gamma}$ is defined as in (4.10). In the case that $X = \mathbb{R}$, we omit the tag X in (5.51), (5.52).

The approximation properties of the operators $\pi^{\boldsymbol{r}}_y$ and $\pi^{\boldsymbol{r}}_{y,\{\mathcal{Y}\}}$ are given in the following lemma.

LEMMA 5.15 (exponential interpolation error estimates). *Let $\beta \in (0,1]$, $\gamma > 0$, C_u, $K_u \ge 0$. Let $\sigma \in (0,1)$. Then there exists a slope $\mathfrak{s}_{min} > 0$ such that on the geometric mesh $\mathcal{G}^M_{geo,\sigma}$ the following estimates hold for any polynomial degree distribution $\boldsymbol{r} = (r_i)_{i=1}^M$ with $r_i \ge 1 + \mathfrak{s}_{min}(i-1)$:*

(i) If $u \in \mathcal{B}^1_{\beta,\gamma}(C_u,K_u;X)$ and $\sigma^M \mathcal{Y} \le 1$, then

$$
\|u - \pi^{\boldsymbol{r}}_y u\|_{L^2(\omega_{\alpha,\gamma},(0,\mathcal{Y});X)} \lesssim C_u \mathcal{Y}^\beta e^{-bM}, \tag{5.53}
$$

$$
\|u - \pi^{\boldsymbol{r}}_{y,\{\mathcal{Y}\}} u\|_{L^2(\omega_{\alpha,\gamma},(0,\mathcal{Y});X)} \lesssim C_u \left(\mathcal{Y}^\beta e^{-bM} + \mathcal{Y}^{-1/2+\beta} e^{-\gamma\mathcal{Y}/2}\right). \tag{5.54}
$$

(ii) If $u \in \mathcal{B}^2_{\beta,\gamma}(C_u,K_u;X)$ and $\sigma^M \mathcal{Y} \le 1$, then

$$
\|(u - \pi^{\boldsymbol{r}}_y u)'\|_{L^2(\omega_{\alpha,\gamma},(0,\mathcal{Y});X)} \lesssim C_u \mathcal{Y}^\beta e^{-bM}, \tag{5.55}
$$

$$
\|(u - \pi^{\boldsymbol{r}}_{y,\{\mathcal{Y}\}} u)'\|_{L^2(\omega_{\alpha,\gamma},(0,\mathcal{Y});X)} \lesssim C_u \left(\mathcal{Y}^\beta e^{-bM} + \mathcal{Y}^{-3/2+\beta} e^{-\gamma\mathcal{Y}/2}\right). \tag{5.56}
$$

In all the estimates, the hidden constant and $b > 0$ depend only on β, γ, α, σ, and K_u.

Proof. See Appendix A. ∎

5.5.2. hp-discretization in y and P_1 FEM in Ω. With the hp-approximation operator $\pi^{\boldsymbol{r}}_y$ of the previous section at hand, we can analyze the properties of the space $\mathbb{V}^{1,\boldsymbol{r}}_{h,M}(\mathcal{T}^\ell_\beta, \mathcal{G}^M_{geo,\sigma})$. The following result generalizes [40] in that we allow for a general elliptic operator $\mathcal{L}$ and in that the appropriate mesh grading in Ω is included to compensate for the lack of full elliptic regularity in the Sobolev scale $H^s(\Omega)$.

Theorem 5.16 (error estimates). *Let $\Omega \subset \mathbb{R}^2$ be a bounded polygon with (a finite set of) corners $\{\boldsymbol{c}\}$. Let $\beta \in [0,1)$ be such that (5.12) holds. Let $u \in \mathbb{H}^s(\Omega)$ and $\mathscr{U} \in \mathring{H}^1(y^\alpha, \mathcal{C})$ solve (1.1) and (1.2), respectively, with $f \in \mathbb{H}^{1-s}(\Omega)$ and $s \in (0,1)$. Let $\beta \in [0,1)$ be such that (5.12) holds and let $\{\mathcal{T}^\ell_\beta\}_{\ell \geq 0}$ be a sequence of uniformly shape-regular meshes of meshwidth h_ℓ that satisfy (5.17) and (5.18). For arbitrary, fixed $0 < \sigma < 1$, denote $\mathcal{G}^M_{geo,\sigma}$ a geometric mesh on $(0,\mathcal{Y})$ with $\mathcal{Y} \sim |\log h_\ell|$ with a sufficiently large implied constant that is independent of ℓ, and assume that the number M of elements in $\mathcal{G}^M_{geo,\sigma}$ satisfies $c_1 M \leq \mathcal{Y} \leq c_2 M$ with absolute constants c_1 and c_2. Let $\mathscr{U}_{h_\ell,M}$ be the Galerkin projection (5.6) onto the space $\mathbb{V}^{1,\boldsymbol{r}}_{h_\ell,M}(\mathcal{T}^\ell_\beta, \mathcal{G}^M_{geo,\sigma})$. Then there exists a minimal slope $\mathfrak{s}_{min}$ independent of h_ℓ and f such that for linear polynomial degree vectors $\boldsymbol{r}$ with slope $\mathfrak{s} \geq \mathfrak{s}_{min}$ there holds*

$$\|u - \operatorname{tr}_\Omega \mathscr{U}_{h_\ell,M}\|_{\mathbb{H}^s(\Omega)} \lesssim \|\nabla(\mathscr{U} - \mathscr{U}_{h_\ell,M})\|_{L^2(y^\alpha,\mathcal{C})} \lesssim h_\ell \|f\|_{\mathbb{H}^{1-s}(\Omega)}. \tag{5.57}$$

In addition, the total number of degrees of freedom satisfies, with $\mathcal{N}_\Omega = \#\mathcal{T}^\ell_\beta$,

$$\dim \mathbb{V}^{1,\boldsymbol{r}}_{h_\ell,M}(\mathcal{T}^\ell_\beta, \mathcal{G}^M_{geo,\sigma}) \sim \mathcal{N}_{\Omega,\mathcal{Y}} \sim M^2 h_\ell^{-2} \sim h_\ell^{-2}(\log h_\ell)^2 \sim \mathcal{N}_\Omega \log \mathcal{N}_\Omega.$$

More generally, if $f \in \mathbb{H}^{t-s}(\Omega)$ for $t \in [0,1]$, then the bound (5.57) takes the form

$$\|u - \operatorname{tr}_\Omega \mathscr{U}_{h_\ell,M}\|_{\mathbb{H}^s(\Omega)} \lesssim \|\nabla(\mathscr{U} - \mathscr{U}_{h_\ell,M})\|_{L^2(y^\alpha,\mathcal{C})} \lesssim h_\ell^t \|f\|_{\mathbb{H}^{t-s}(\Omega)}.$$

Proof. The starting point is again the error decomposition (5.9). The univariate hp-interpolation operator $\pi^{\boldsymbol{r}}_y$ constructed in Section 5.5.1 makes the semidiscretization error $\mathscr{U} - \pi^{\boldsymbol{r}}_y \mathscr{U}$ in y exponentially small in M (see Lemma 6.2 below for details). In turn, the assumption $M \sim |\log h_\ell|$ implies any desired algebraic convergence in h_ℓ by suitably selecting the hidden constant. On the other hand, the error $\mathscr{U} - \Pi^q_{x'} \mathscr{U}$ in (5.9) is controlled as in the proof of Theorem 5.10.

Finally, the estimate for $f \in \mathbb{H}^{\sigma-s}(\Omega)$ follows by interpolation. $\square$

6. Diagonalization: semidiscretization in y. We now explore the possibilities offered by a semidiscretization in y. We will observe, among other things, that this leads to a sequence of decoupled singularly perturbed, *linear second order* elliptic problems in Ω. We remark that this decoupling idea dates back at least to [38], in the context of the Laplacian in rectangular domains.

For an arbitrary mesh $\mathcal{G}^M$ on $[0,\mathcal{Y}]$ and for a polynomial degree distribution $\boldsymbol{r}$, we consider the following y-semidiscrete problem: Find $\mathscr{U}_M \in \mathbb{V}^{\boldsymbol{r}}_M(\mathcal{C}_\mathcal{Y})$ such that

$$a_\mathcal{C}(\mathscr{U}_M, \phi) = d_s\langle f, \operatorname{tr}_\Omega \phi\rangle \qquad \forall \phi \in \mathbb{V}^{\boldsymbol{r}}_M(\mathcal{C}_\mathcal{Y}), \tag{6.1}$$

where $\mathbb{V}^{\boldsymbol{r}}_M(\mathcal{C}_\mathcal{Y})$ is defined in (5.3) and is a closed subspace of $\mathring{H}^1(y^\alpha, \mathcal{C})$. In what follows we obtain an explicit formula for $\mathscr{U}_M$. To accomplish this, we consider the following eigenvalue problem: Find $(v,\mu) \in S^{\boldsymbol{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^M) \setminus \{0\} \times \mathbb{R}$ such that

$$\mu \int_0^\mathcal{Y} y^\alpha v'(y) w'(y)\, \mathrm{d}y = \int_0^\mathcal{Y} y^\alpha v(y) w(y)\, \mathrm{d}y \qquad \forall w \in S^{\boldsymbol{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^M), \tag{6.2}$$

where $S^{\boldsymbol{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^M)$ is defined in (5.1). All eigenvalues μ are positive, and the space $S^{\boldsymbol{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^M)$ has an eigenbasis $(v_i)_{i=1}^{\mathcal{M}}$, with $\mathcal{M} := \dim S^{\boldsymbol{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^M)$, such that, for i, $j \in \{1, \dots, \mathcal{M}\}$,

$$\int_0^{\mathscr{Y}} y^\alpha v_i'(y) v_j'(y) \, \mathrm{d}y = \delta_{i,j}, \qquad \int_0^{\mathscr{Y}} y^\alpha v_i(y) v_j(y) \, \mathrm{d}y = \mu_i \delta_{i,j}. \tag{6.3}$$

We write $\mathscr{U}_M(x', y) := \sum_{j=1}^{\mathcal{M}} U_j(x') v_j(y)$ and consider $\phi(x', y) = V(x') v_i(y)$, with $V \in H_0^1(\Omega)$ as a test function, in (6.1). This yields the following system of *decoupled* problems for $i = 1, \dots, \mathcal{M}$: Find $U_i \in H_0^1(\Omega)$ such that

$$a_{\mu_i,\Omega}(U_i, V) = d_s v_i(0) \langle f, V \rangle \qquad \forall V \in H_0^1(\Omega), \tag{6.4}$$

where

$$a_{\mu_i,\Omega}(U, V) := \mu_i a_\Omega(U, V) + \int_\Omega UV \, \mathrm{d}x',$$

and a_Ω is introduced in (2.1). An important observation is that, for functions of the form $Z(x', y) = \sum_{i=1}^{\mathcal{M}} V_i(x') v_i(y)$ with $V_i \in H_0^1(\Omega)$, we have the equality

$$a_{\mathcal{C}}(Z, Z) = a_{\mathcal{C}_{\mathscr{Y}}}(Z, Z) = \sum_{i=1}^{\mathcal{M}} \|V_i\|_{\mu_i,\Omega}^2, \qquad \|V\|_{\mu_i,\Omega}^2 := a_{\mu_i,\Omega}(V, V). \tag{6.5}$$

To obtain a fully discrete scheme, select a mesh $\mathcal{T}$ on Ω and the corresponding space $S_0^q(\Omega, \mathcal{T})$ and let $\Pi_i : H_0^1(\Omega) \to S_0^q(\Omega, \mathcal{T})$ be the Ritz projectors for the bilinear forms $a_{\mu_i,\Omega}$ defined by

$$a_{\mu_i,\Omega}(u - \Pi_i u, v) = 0 \qquad \forall v \in S_0^q(\Omega, \mathcal{T}). \tag{6.6}$$

With this notation at hand, we can formulate an explicit representation of the Galerkin approximation $\mathscr{U}_{h,M} \in S_0^q(\Omega, \mathcal{T}) \otimes S_{\{\mathscr{Y}\}}^{\boldsymbol{r}}(\mathcal{G}^M)$ to $\mathscr{U}$ as well as an error representation.

Lemma 6.1 (error representation). *Let $(\mu_i, v_i)_{i=1}^{\mathcal{M}}$ be the eigenpairs given by (6.2), (6.3). Let $U_i \in H_0^1(\Omega)$ be the solution to (6.4) and $\Pi_i : H_0^1(\Omega) \to S_0^q(\Omega, \mathcal{T})$ be the Ritz projection defined in (6.6). Let $\mathscr{U}_M$ be the solution to the semidiscrete problem (6.1). Then the Galerkin approximation $\mathscr{U}_{h,M} \in S_0^q(\Omega, \mathcal{T}) \otimes S_{\{\mathscr{Y}\}}^{\boldsymbol{r}}(\mathcal{G}^M)$ to $\mathscr{U}$ satisfies*

$$\mathscr{U}_{h,M}(x', y) = \sum_{i=1}^{\mathcal{M}} \Pi_i U_i(x') v_i(y), \tag{6.7}$$

$$a_{\mathcal{C}}(\mathscr{U}_M - \mathscr{U}_{h,M}, \mathscr{U}_M - \mathscr{U}_{h,M}) = \sum_{i=1}^{\mathcal{M}} \|U_i - \Pi_i U_i\|_{\mu_i,\Omega}^2. \tag{6.8}$$

Proof. Expression (6.7) follows from (6.4) and (6.6), whereas (6.8) is a consequence of (6.5). ◻

We next show that the semidiscretization error $\mathscr{U} - \mathscr{U}_M$ can be made exponentially small on geometric meshes $\mathcal{G}_{geo,\sigma}^M$.

Lemma 6.2 (exponential convergence). *Let $f \in \mathbb{H}^{-s+\nu}(\Omega)$ for $\nu \in (0, s)$. Let $c_1 M \le \mathscr{Y} \le c_2 M$. Consider the geometric mesh $\mathcal{G}_{geo,\sigma}^M$ on $(0, \mathscr{Y})$. Then there exist C, $\mathfrak{s}_{min}$, $b > 0$ (depending solely on s, $\mathcal{L}$, c_1, c_2, σ, ν) such that for any linear degree $\boldsymbol{r}$ with slope $\mathfrak{s} \ge \mathfrak{s}_{min}$ there holds*

$$\|\nabla(\mathscr{U} - \mathscr{U}_M)\|_{L^2(y^\alpha, \mathcal{C})} \le C e^{-bM} \|f\|_{\mathbb{H}^{-s+\nu}(\Omega)}. \tag{6.9}$$

Proof. We begin the proof by invoking Galerkin orthogonality to arrive at

$$\begin{aligned}\|\mathscr{U} - \mathscr{U}_M\|_{\mathcal{C}}^2 &\leq \|\mathscr{U} - \pi_{y,\{\mathscr{Y}\}}^{\boldsymbol{r}}\mathscr{U}\|_{\mathcal{C}}^2 \\ &\lesssim \|\mathscr{U} - \pi_{y,\{\mathscr{Y}\}}^{\boldsymbol{r}}\mathscr{U}\|_{\mathcal{C}_{\mathscr{Y}}}^2 + \|\nabla \mathscr{U}\|_{L^2(y^\alpha,\mathcal{C}\setminus\mathcal{C}_{\mathscr{Y}})}^2,\end{aligned}$$

where $\|\cdot\|_{\mathcal{C}}$ and $\|\cdot\|_{\mathcal{C}_{\mathscr{Y}}}$ are defined by (2.10) and (2.11), respectively. Since (5.8) shows that $\|\nabla\mathscr{U}\|_{L^2(y^\alpha;\mathcal{C}\setminus\mathcal{C}_{\mathscr{Y}})}$ is exponentially small in $\mathscr{Y}$ we may focus on the interpolation error term. To control such a term we first observe that, in view of the definitions of the spaces $\mathcal{B}^j_{\beta,\gamma}$, $j \in \{0,1\}$, given by (5.51), (5.52), the regularity estimates (4.17) and (4.18) of Theorem 4.7, imply that $\mathscr{U}$ viewed as a function in $C^\infty((0,\infty), L^2(\Omega)) \cap C^\infty((0,\infty), H_0^1(\Omega))$ satisfies for $\nu \in (0,s)$ and $K > \kappa$ (with κ as in Theorem 4.7)

$$\mathscr{U} \in \mathcal{B}^1_{\nu,\gamma}(C\|f\|_{\mathbb{H}^{-s+\nu}(\Omega)}, K; H_0^1(\Omega)) \cap \mathcal{B}^2_{\nu,\gamma}(C\|f\|_{\mathbb{H}^{-s+\nu}(\Omega)}, K; L^2(\Omega)). \tag{6.10}$$

From Lemma 5.15 together with the fact that $\mathscr{Y} \sim M$ we conclude that

$$\|\nabla_{x'}(\mathscr{U} - \pi_{y,\{\mathscr{Y}\}}^{\boldsymbol{r}}\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_{\mathscr{Y}})} \leq Ce^{-bM}\|f\|_{\mathbb{H}^{-s+\nu}(\Omega)}, \tag{6.11}$$

$$\|\partial_y(\mathscr{U} - \pi_{y,\{\mathscr{Y}\}}^{\boldsymbol{r}}\mathscr{U})\|_{L^2(y^\alpha,\mathcal{C}_{\mathscr{Y}})} \leq Ce^{-bM}\|f\|_{\mathbb{H}^{-s+\nu}(\Omega)}, \tag{6.12}$$

with $b > 0$ slightly smaller than that in (5.53) – (5.56). This implies the desired estimate (6.9) and concludes the proof. ◻

Finally, for the geometric mesh $\mathcal{G}^M_{geo,\sigma}$ with the linear degree vector $\boldsymbol{r}$ and truncation parameter $\mathscr{Y} \sim M$, we have the following estimates for the eigenvalues μ_i of problem (6.2) and for the point values $v_i(0)$ in (6.4).

Lemma 6.3 (properties of the eigenpairs). *For arbitrary fixed $0 < \sigma < 1$ and for every $M \in \mathbb{N}$, let $\mathcal{G}^M_{geo,\sigma}$ be a geometric mesh on $(0,\mathscr{Y})$ and $\boldsymbol{r}$ a linear degree vector with slope $\mathfrak{s} \geq \mathfrak{s}_{min} > 0$. If $c_1 M \leq \mathscr{Y} \leq c_2 M$ for some constants $0 < c_1 < c_2 < \infty$, then there is $C > 0$ depending only on α, σ, c_1, c_2, $\mathfrak{s}_{min}$ such that for the eigenpairs $(\mu_i, v_i)_{i=1}^{\mathcal{M}}$ given by* (6.2), (6.3) *we have that:*

$$\|v_i\|_{L^\infty(0,\mathscr{Y})} \leq CM^{(1-\alpha)/2}, \qquad C^{-1}\left(\mathscr{Y}\mathfrak{s}^{-2}M^{-2}\sigma^M\right)^2 \leq \mu_i \leq CM^2.$$

Proof. The results follow from Lemmas B.1, B.2, and B.3. ◻

The previously described approach to perform a semidiscretization in y leads to structural insight into the regularity properties of the solution $\mathscr{U}$: it shows that, up to an exponentially small, in $\mathscr{Y}$, error introduced by cutting off at $\mathscr{Y}$, the solution $\mathscr{U}$ can be expressed in terms of solutions of singularly perturbed reaction–diffusion type problems. (A similar structural property for $\mathscr{U}(\cdot,0)$ can also be seen from the Balakrishnan formula, e.g., [13, Equation (4)]). In what follows we will exploit this to design appropriate approximation spaces in the x'-variable. Nevertheless, the diagonalization (6.1)–(6.4) has more far-reaching ramifications:

- The diagonalization technique can be exploited numerically as it is not restricted to the semi-discrete case. It holds for arbitrary, closed tensor product approximation spaces $\mathbb{W}\otimes\mathbb{Q}$, where $\mathbb{W} \subset H_0^1(\Omega)$ and $\mathbb{Q} \subset H^1_{\{\mathscr{Y}\}}(y^\alpha,(0,\mathscr{Y}))$ (cf. (5.4)). It completely decouples the solution of the full Galerkin problem, based on $\mathbb{W}\otimes\mathbb{Q}$, into the (parallel) solution of $\dim\mathbb{Q}$ problems of size $\dim\mathbb{W}$. The numerical experiments in Section 8 exploit this observation; see Remark 8.5 below.

- The observation (6.5) allows one to gauge the impact of solving approximately the $\dim \mathbb{Q}$ problems that are of (singularly perturbed) reaction–diffusion type. For convex domains Ω and spaces $\mathbb{W}$ based on piecewise linears on quasi-uniform meshes, robust (with respect to the singular perturbation parameter) multigrid methods are available (see, e.g., [51]).
- The diagonalization (6.2)–(6.4) also suggests another numerical approach: approximate each solution U_i from a different (closed) space $W_i \subset H^1_0(\Omega)$. This leads to the approximation of $\mathscr{U}$ in the space $\sum_{i=1}^{\mathcal{M}} v^i(y) W_i$. The resulting Galerkin approximation still satisfies (6.7) and (6.8). This approach produces approximation spaces in $\Omega \times (0, \mathcal{Y})$ that do not have tensor product structure but still provides exponential convergence. As in the sparse grids case of Section 5.4.4 this approach allows for reducing the number of degrees of freedom without sacrificing much accuracy; specifically, the exponent $1/4$ in the exponential convergence bound (7.7) that we obtain in the next section could be reduced to $1/3$ if Ω is an interval and the exponent $1/5$ in (7.12) could be reduced to $1/4$ if $\Omega \subset \mathbb{R}^2$ has an analytic boundary, albeit at the expense of breaking the tensor product structure of the discretization.

7. *hp*-FE discretization in Ω. Up to this point, we have exploited the analytic regularity of the solution $\mathscr{U}$ in the extended variable y in order to recover (up to logarithmic terms) optimal complexity of a P_1-FEM, for (1.1) posed in the polygon $\Omega \subset \mathbb{R}^2$, by full tensorization of a hp-FEM with respect to y with the P_1-FEM in Ω

As a final goal, in this section we employ, *in addition*, an hp-FEM in Ω to obtain an *exponentially convergent, local FEM* for the fractional diffusion problem (1.1). Naturally, stronger regularity assumptions on the data f, A and c will be required: in addition to the previously made assumptions on these data, we assume in the present Section 7

$$c, \quad f \in \mathcal{A}(\overline{\Omega}, \mathbb{R}) \,, \quad A \in \mathcal{A}(\overline{\Omega}, \mathsf{GL}(\mathbb{R}^d)) \,. \tag{7.1}$$

Here, $\mathcal{A}(\overline{\Omega}, G)$ denotes the set of functions which are analytic in $\overline{\Omega}$ and take values in the group G.

We describe the setup of tensorized FEM in $\Omega \times (0, \mathcal{Y})$. The choice of the meshes $\mathcal{G}^M$ and $\mathcal{T}$ as well as the degree vector $\boldsymbol{r}$ and the polynomial degree q were not specified in Section 6. Mesh design principles for problems as (6.4) are available in the literature. For meshes, in an h-version context, we mention the so–called Shishkin meshes and refer to [53] for an in-depth discussion of numerical methods for singular perturbation problems. Here, we focus on the hp-version. Appropriate mesh design principles ensuring robust exponential convergence of hp-FEM have been developed in [60, 61, 41, 43, 42]. In these references, linear second order elliptic singular perturbations with a single length scale and exponential boundary layers were considered. As is revealed by the diagonalization (6.4), the y-semidiscrete solution (6.1) contains $\mathcal{M}$ separate length scales μ_i, $i = 1, ..., \mathcal{M}$. These need to be resolved *simultaneously* by the x'-discretization space. To this end, based on [60, 61, 41, 43, 42], we employ a mesh that is geometrically refined towards $\partial\Omega$ such that the smallest length scale $\mu_{\mathcal{M}}$ is resolved. We illustrate the key points in the following Sections 7.1 and 7.2 in dimension $d = 1$, and in dimension $d = 2$ for smooth boundaries.

7.1. Exponential convergence of *hp*-FEM in one dimension. To gain insight into how to discretize the family of problems (6.4), we first consider the following reaction-diffusion problem in $\Omega = (0, 2)$: given $f \in \mathcal{A}(\overline{\Omega}; \mathbb{R})$ and a parameter

$0 < \varepsilon \le 1$, find $u_\varepsilon \in H_0^1(\Omega)$ such that

$$-\varepsilon^2 u_\varepsilon'' + u_\varepsilon = f \quad \text{on } \Omega, \qquad u_\varepsilon(0) = u_\varepsilon(2) = 0\;. \tag{7.2}$$

For (7.2), *hp*-Galerkin FEM afford *robust exponential convergence*. The following result is a particular instance of [41, Proposition 20].

PROPOSITION 7.1 (exponential convergence). *Let $\Omega = (0,2)$. Let $\mathcal{T}^{1D,L}_{geo,\sigma}$ be a mesh on Ω that is geometrically refined towards $\partial\Omega = \{0,2\}$ with L layers and grading factor $\sigma \in (0,1)$:*

$$\mathcal{T}^{1D,L}_{geo,\sigma} := \{(0,\sigma^L),(2-\sigma^L,2)\} \cup \{(\sigma^{L-i+1},\sigma^{L-i}),(2-\sigma^{L-i},2-\sigma^{L-i+1})\}_{i=1}^{L}. \tag{7.3}$$

Select L such that $\sigma^L \le \varepsilon \le 1$. Let f satisfy the analytic regularity estimates

$$\|f^{(\ell)}\|_{L^2(\Omega)} \le C_f K_f^\ell \ell! \qquad \forall \ell \in \mathbb{N}_0, \tag{7.4}$$

for some constants C_f, $K_f > 0$ that depend on f. Then there exist constants C, $b > 0$ independent of $\varepsilon \in (0,1]$ such that for the Galerkin approximation $u_\varepsilon^{q,L} \in S_0^q(\Omega, \mathcal{T}^{1D,L}_{geo,\sigma})$ of the solution u_ε of (7.2) one has exponential convergence in the energy norm, given by $\|w\|^2_{\varepsilon^2,\Omega} := \varepsilon^2\|w'\|^2_{L^2(\Omega)} + \|w\|^2_{L^2(\Omega)}$, i.e.,

$$\|u_\varepsilon - u_\varepsilon^{q,L}\|_{\varepsilon^2,\Omega} \lesssim C_f e^{-bq}.$$

Here the hidden constant and the constant $b > 0$ are independent of ε, but depend on σ and K_f. Furthermore, $L = \mathcal{O}(1 + |\log\varepsilon|)$ so that $\dim S_0^q(\Omega, \mathcal{T}^{1D,L}_{geo,\sigma}) = \mathcal{O}(q^2(1 + |\log\varepsilon|))$.

REMARK 7.2 (exponential convergence). *The discretization described in Proposition 7.1 and its properties warrant the following comments.*

- *The case $\epsilon \ge 1$: Although Proposition 7.1 restricts to $\varepsilon \in (0,1]$, one can check that for $\varepsilon \ge 1$, the mesh degenerates into a fixed mesh with three points $\{0,1,2\}$ and the corresponding approximation result reads*

$$\|u_\varepsilon - u_\varepsilon^{q,L}\|_{\varepsilon^2,\Omega} \lesssim (1+\varepsilon) C_f e^{-bq}\;. \tag{7.5}$$

- *Different length scales: Proposition 7.1 gives* robust exponential convergence *and* does not require explicit knowledge of the singular perturbation parameter ε, *but only a lower bound for it. This is crucial for the presently considered fractional diffusion problem, where the decoupled problems* (6.4) *depend on several length scales characterized by μ_i (which, in turn, depend on the discretization in the extended variable $y \in (0,\mathcal{Y})$). Applying a tensor product hp-FE space directly (i.e., without explicit diagonalization* (6.1)–(6.4)*) to the extended problem* (1.2) *based on the tensor product of the hp-FE space $S_0^q(\Omega, \mathcal{T}^{1D,L}_{geo,\sigma})$ and on the hp-FE space $S^{\mathbf{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}), \mathcal{G}^M_\sigma)$ obviates the numerical solution of the generalized eigenproblem* (6.2). *It requires, however, the hp-space $S_0^q(\Omega, \mathcal{T}^{1D,L}_{geo,\sigma})$ to* concurrently *approximate the solutions of* all singularly perturbed problems (6.4) *in Ω with exponential convergence rates.*
- *Different meshes: If an eigenbasis $(v_i)_{i=1}^{\mathcal{M}}$ satisfying* (6.3) *is available, then for each of the decoupled singularly perturbed problems in Ω, a geometric boundary layer mesh is not mandatory to achieve robust exponential convergence. A coarser mesh, tailored to the specific length scale μ_i in the i-th equation of* (6.4), *will then suffice; we refer to [60, 59] for details.* ∎

Lemma 6.3 asserts that the reaction-diffusion problems (6.4) are singularly perturbed with length scale μ_i ranging from $\mathcal{O}(M^{-2}\sigma^{2M})$ to $\mathcal{O}(M^2)$. Proposition 7.1 implies exponential convergence rates under the analyticity assumption (7.1). In the next result, we combine these two observations to obtain an exponentially convergent *hp*-FEM for the fractional diffusion problem in Ω.

THEOREM 7.3 (exponential convergence). *Let $u \in \mathbb{H}^s(\Omega)$ and $\mathscr{U} \in \mathring{H}^1(y^\alpha, \mathcal{C})$ solve* (1.1) *and* (1.2)*, respectively, with $\Omega = (0,2)$, $A = I$, $c = 0$, and f satisfying* (7.1)*. Given fixed constants c_1, $c_2 > 0$, let $\mathcal{G}^M_{geo,\sigma}$ be a geometric mesh on $[0, \mathscr{Y}]$ with grading factor $\sigma \in (0,1)$ and such that $c_1 M \le \mathscr{Y} \le c_2 M$. Let $\boldsymbol{r}$, on $\mathcal{G}^M_{geo,\sigma}$, be the linear degree vector with slope $\mathfrak{s}$. Let $\mathcal{T}^{1D,L}_{geo,\sigma}$ be a geometric mesh in Ω as described in Proposition 7.1 with an integer L such that*

$$\sigma^{2L} \le \mathscr{Y}^2 (\mathfrak{s} M)^{-4} \sigma^{2M} \, . \tag{7.6}$$

Then, there are constants b, $\mathfrak{s}_{min} > 0$ independent of M such that for $\mathfrak{s} \ge \mathfrak{s}_{min}$ the Galerkin projection $\mathscr{U}_{q,\boldsymbol{r}} \in S^q_0(\Omega, \mathcal{T}^{1D,L}_{geo,\sigma}) \otimes S^{\boldsymbol{r}}_{\{\mathscr{Y}\}}((0,\mathscr{Y}), \mathcal{G}^M_{geo,\sigma})$ of $\mathscr{U}$ satisfies

$$\|u - \operatorname{tr}_\Omega \mathscr{U}_{q,\boldsymbol{r}}\|_{\mathbb{H}^s(\Omega)} \lesssim \|\nabla(\mathscr{U} - \mathscr{U}_{q,\boldsymbol{r}})\|_{L^2(y^\alpha,\mathcal{C})} \lesssim M^2 e^{-bq} + e^{-bM},$$

where the hidden constant is independent of M and of q. In addition, as $M \to \infty$, with L and M related by (7.6)*, we have that, uniformly in $q \in \mathbb{N}$, the total number of degrees of freedom is*

$$\mathcal{N}_{\Omega,\mathscr{Y}} := \dim S^q_0(\Omega, \mathcal{T}^{1D,L}_{geo,\sigma}) \otimes S^{\boldsymbol{r}}_{\{\mathscr{Y}\}}((0,\mathscr{Y}), \mathcal{G}^M_\sigma) = \mathcal{O}(qM^3).$$

Choosing, in particular, $q \sim M$ yields a convergence rate bound in terms of the total number of degrees of freedom $\mathcal{N}_{\Omega,\mathscr{Y}}$ of the form

$$\|u - \operatorname{tr}_\Omega \mathscr{U}_{q,\boldsymbol{r}}\|_{\mathbb{H}^s(\Omega)} \lesssim \|\nabla(\mathscr{U} - \mathscr{U}_{q,\boldsymbol{r}})\|_{L^2(y^\alpha,\mathcal{C})} \lesssim \exp(-b' \mathcal{N}^{1/4}_{\Omega,\mathscr{Y}}) \tag{7.7}$$

for some $b' > 0$ independent of $\mathcal{N}_{\Omega,\mathscr{Y}}$.

Proof. Let $\mathscr{U}_M$ solve (6.1). We proceed in two steps.

Step 1 (Bounds on the semidiscretization error $\mathscr{U} - \mathscr{U}_M$): By the assumption of analyticity of f, there exist constants C_f, K_f such that (7.4) holds. We thus have that $f \in \mathbb{H}^{1/2-\delta}(\Omega)$ for any $\delta > 0$. Consequently, an application of Lemma 6.2 reveals that for a sufficiently large slope $\mathfrak{s}$ of the linear degree vector $\boldsymbol{r}$ (depending on the constants K_f in the analytic regularity bound (7.4) of the data f) there exists $b > 0$ such that

$$\|\nabla(\mathscr{U} - \mathscr{U}_M)\|_{L^2(y^\alpha,\mathcal{C})} \lesssim e^{-bM}.$$

Step 2(Bounds on the errors $\|U_i - \Pi_i U_i\|_{\mu_i,\Omega}$): We first notice that Lemma 6.3 immediately yields $\mathfrak{s}^{-4} M^{-2} \sigma^{2M} \lesssim \mu_i$. This, in view of the assumption (7.6), implies that $\sigma^{2L} \lesssim \mu_i$. Consequently, given that f is analytic on $\overline{\Omega}$, we apply Proposition 7.1 (more precisely, the refinement (7.5)) to obtain that

$$\|U_i - \Pi_i U_i\|_{\mu_i,\Omega} \lesssim \mathscr{Y} e^{-bq} \lesssim M e^{-bq}, \tag{7.8}$$

where we have also used that $\mu_i \lesssim M^2 \lesssim \mathscr{Y}^2$, which follows, again, from Lemma 6.3 and the condition $c_1 M \le \mathscr{Y} \le c_2 M$. We recall that $\|\cdot\|_{\mu_i,\Omega}$ is defined as in (6.5). Finally, combining (7.8) with (6.8) and recalling that $\mathcal{M} \lesssim M^2$ give

$$\|\mathscr{U}_M - \mathscr{U}_{h,M}\|^2_{L^2(y^\alpha,\mathcal{C})} \lesssim \mathcal{M} M^2 e^{-2bq} \lesssim M^4 e^{-bq}.$$

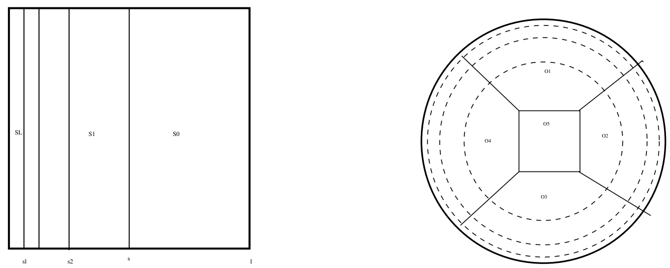

Fig. 7.1. *Anisotropic geometric mesh (see Definition 7.6). Left: geometric refinement of the reference patch with L layers. Right: Example of a mesh with $N = 5$ patches where $n = 4$ and with $L = 3$ layers in the boundary layer patches. Solid lines indicate patches, dashed lines represent mesh lines introduced by refinement of reference patches.*

This concludes the proof. □

Remark 7.4 (other operators). *Theorem 7.3 also holds for* $0 < c \in \mathbb{R}$ *by arguing as in the proof of Theorem 7.7 ahead.* ∎

Remark 7.5 (mesh gradings Ω). *The condition* (7.6) *is a sufficient condition ensuring that the smallest boundary layer length scale (characterized by* $\min_i \mu_i$*) that arises from the diagonalization is resolved by the mesh* $\mathcal{T}^{1D,L}_{geo,\sigma}$*. More generally, if the geometric mesh of* (7.3) *were based on the mesh grading factor* $\sigma_{x'} \in (0,1)$ *(distinct from the factor* σ *in the mesh in the extended variable* y*), then condition* (7.6) *should be replaced with* $\sigma_{x'}^{2L} \leq C\mathcal{Y}^2(\mathfrak{s}M)^{-4}\sigma^{2M}$ *for some chosen* $C > 0$ *independent of* $L, M, \mathcal{Y}$*.* ∎

7.2. Exponential convergence of *hp*-FEM in two dimensions. Let us now discuss the extension of the ideas of Section 7.1 to the two dimensional case. As it is structurally similar to the univariate case, we proceed briefly. For domains $\Omega \subset \mathbb{R}^d$, $d > 1$, with *smooth* boundary, the boundary layers presented in the solutions U_i of the singularly perturbed problems (6.4) can be resolved by meshes that are *anisotropically* refined towards the boundary $\partial\Omega$. A two dimensional analogue of the meshes $\mathcal{T}^{1D}_{geo,L}$ of Proposition 7.1 is presented in [43, Section 3.4.3] and illustrated in Figure 7.1 (right). These *anisotropic geometric meshes* $\mathcal{T}^{2D,L}_{geo,\sigma}$ are created as push-forwards of anisotropically refined geometric meshes on reference patches as detailed in the following definition, where we follow the notation employed in [43, Section 3.4.3].

Definition 7.6 (anisotropic geometric meshes $\mathcal{T}^{2D,L}_{geo,\sigma}$). *Denote by* $S = [0,1]^2$ *the reference element. Let* Ω_i*,* $i = 1,\dots,N$*, be a fixed mesh on* $\Omega \subset \mathbb{R}^2$ *consisting of curvilinear quadrilaterals with bijective element maps* $M_i : S \to \Omega_i$ *satisfying the "usual" conditions for* H^1*-conforming triangulations (see [43, (M1)–(M3) in Section 3.1] for the precise definition). The elements* Ω_i *are called* patches *and the associated maps* M_i patch maps*. Let* Ω_i*,* $i = 1,\dots,n \leq N$*, be such that the left edge* $e := \{0\} \times (0,1)$ *of* S *is mapped to* $\partial\Omega$*, i.e.,* $M_i(e_1) \subset \partial\Omega$*, and that* $M_i(\partial S \setminus e) \cap \partial\Omega = \emptyset$*. Assume that the remaining elements* Ω_i*,* $i = n+1,\dots,N$ *satisfy* $\overline{\Omega}_i \cap \partial\Omega = \emptyset$*.*

Subdivide the reference element S *into* $L+1$ *rectangles* S^ℓ*,* $\ell = 0,\dots,L$*, as follows for chosen grading factor* $\sigma \in (0,1)$*:*

$$S^0 = (0,\sigma^L) \times (0,1), \qquad S^\ell = (\sigma^{L+1-\ell},\sigma^{L-\ell}) \times (0,1), \qquad \ell = 1,\dots,L. \tag{7.9}$$

Define elements Ω_i^ℓ*,* $i = 1,\dots,n$*,* $\ell = 0,\dots,L$*, and the corresponding element maps*

$M_i^\ell : S \to \Omega_i^\ell$ by

$$\begin{aligned}
\Omega_i^0 &:= M_i(S^0), \qquad & M_i^0(\xi,\eta) &:= M_i(\xi\sigma^L,\eta),\\
\Omega_i^\ell &:= M_i(S^\ell), \qquad & M_i^\ell(\xi,\eta) &:= M_i(\sigma^{L+1-\ell}+\xi\sigma^{L-\ell},\eta), \qquad \ell = 1,\dots,L.
\end{aligned}$$

The mesh $\mathcal{T}^{2D,L}_{geo,\sigma}$ given by the elements $\{\Omega_i^\ell : i = 1,\dots,n, \ell = 0,\dots,L\} \cup \{\Omega_j : j = n+1,\dots,N\}$ with corresponding element maps introduced above is a triangulation of Ω that satisfies the "usual" conditions of H^1-conforming triangulations, i.e., conditions [43, (M1)–(M3) in Section 3.1]. For $\mathcal{T}^{2D,L}_{geo,\sigma}$ the FE-space is given by the standard $H^1_0(\Omega)$-conforming space of mapped polynomials of degree q:

$$S_0^q(\mathcal{T}^{2D,L}_{geo,\sigma}) := \{u \in H^1_0(\Omega) : u|_K \circ F_K \in \mathbb{Q}_q(S) \quad \forall K \in \mathcal{T}^{2D,L}_{geo,\sigma}\}, \tag{7.10}$$

where $F_K : S \to K$ is the element map of $K \in \mathcal{T}^{2D,L}_{geo,\sigma}$ and $\mathbb{Q}_q(S)$ is the space of polynomials of degree q in each variable on S.

For such anisotropically refined meshes, we have the following exponential convergence result.

Theorem 7.7 (exponential convergence). *Let $u \in \mathbb{H}^s(\Omega)$ and $\mathscr{U} \in \mathring{H}^1(y^\alpha,\mathcal{C})$ solve* (1.1) *and* (1.2)*, respectively, with $\Omega \subset \mathbb{R}^2$ having an analytic boundary, $A = I$, $0 \le c \in \mathbb{R}$, and f satisfying the regularity requirement* (7.1)*. Given fixed constants c_1, $c_2 > 0$, let $\mathcal{G}^M_{geo,\sigma}$ be a geometric mesh on $[0,\mathcal{Y}]$ with grading factor $\sigma \in (0,1)$ and such that $c_1 M \le \mathcal{Y} \le c_2 M$. Let $\boldsymbol{r}$, on $\mathcal{G}^M_{geo,\sigma}$, be the linear degree vector with slope $\mathfrak{s}$. Assume that L is chosen such that* (7.6) *holds. Let $\mathcal{T}^{2D,L}_{geo,\sigma}$ be an anisotropic geometric mesh with L layers as described in Definition 7.6 where, additionally, the patch maps M_i, $i = 1,\dots,N$, are assumed to be analytic. Then, there are constants C, b, $\mathfrak{s}_{min} > 0$ independent of M such that for $\mathfrak{s} \ge \mathfrak{s}_{min}$ the Galerkin approximation $\mathscr{U}_{q,\boldsymbol{r}} \in S^q_0(\Omega,\mathcal{T}^{2D,L}_{geo,\sigma}) \otimes S^{\boldsymbol{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}),\mathcal{G}^M_{geo,\sigma})$ to $\mathscr{U}$ satisfies*

$$\|u - \operatorname{tr}_\Omega \mathscr{U}_{q,\boldsymbol{r}}\|_{\mathbb{H}^s(\Omega)} \lesssim \|\nabla(\mathscr{U} - \mathscr{U}_{q,\boldsymbol{r}})\|_{L^2(y^\alpha,\mathcal{C})} \le C\left(M^2 e^{-bq} + e^{-bM}\right). \tag{7.11}$$

Furthermore, as $M \to \infty$, with L related to M by (7.6)*, we have that, uniformly in $q \in \mathbb{N}$, the total number of degrees of freedom is*

$$\mathcal{N}_{\Omega,\mathcal{Y}} := \dim S^q_0(\Omega,\mathcal{T}^{2D,L}_{geo,\sigma}) \otimes S^{\boldsymbol{r}}_{\{\mathcal{Y}\}}((0,\mathcal{Y}),\mathcal{G}^M_\sigma) = \mathcal{O}(q^2M^3).$$

Choosing, in particular, $q \sim M$ yields a convergence rate bound in terms of the total number of degrees of freedom $\mathcal{N}_{\Omega,\mathcal{Y}}$ of the form

$$\|u - \operatorname{tr}_\Omega \mathscr{U}_{q,\boldsymbol{r}}\|_{\mathbb{H}^s(\Omega)} \lesssim \|\nabla(\mathscr{U} - \mathscr{U}_{q,\boldsymbol{r}})\|_{L^2(y^\alpha,\mathcal{C})} \lesssim \exp(-b'\mathcal{N}^{1/5}_{\Omega,\mathcal{Y}}) \tag{7.12}$$

for some $b' > 0$ independent of $\mathcal{N}_{\Omega,\mathcal{Y}}$.

Proof. The proof parallels that of Theorem 7.3. We start with the case $c = 0$. By the arguments in [43, Section 3.4.3] the meshes $\mathcal{T}^{2D,L}_{geo,\sigma}$ allow for estimates of the form

$$\inf_{v \in S^q_0(\Omega,\mathcal{T}^{2D,L}_{geo,\sigma})} \|U_i - v\|_{\mu_i,\Omega} \lesssim e^{-bq} \tag{7.13}$$

for the solutions U_i of (6.4), provided L and μ_i satisfy $\sigma^{2L} \lesssim \mu_i$, which is ensured by assumption (7.6) with Lemma 6.3. Here, the hidden constant and $b > 0$ depend on f, $\partial\Omega$, and the analyticity of the patch maps M_i, $i = 1,\dots,N$. The estimates (7.13) then allow us to conclude the proof for $c = 0$ as in Theorem 7.3.

For $c \neq 0$, we observe that the singularly perturbed problems (6.4) in Ω take the form

$$-\mu_i \Delta U_i + (1 + c\mu_i) U_i = f \quad \text{on } \Omega, \qquad U_i|_{\partial\Omega} = 0.$$

This can be transformed to the case $c = 0$ by rewriting it in terms of $\widetilde{\mu}_i := \mu_i/(1+c\mu_i)$ as

$$-\widetilde{\mu}_i \Delta U_i + U_i = \widetilde{f} := \frac{1}{1 + c\mu_i} f \quad \text{on } \Omega, \qquad U_i|_{\partial\Omega} = 0.$$

The approximation result (7.13) holds again (with μ_i replaced with $\widetilde{\mu}_i$ there). □

REMARK 7.8 (limitations and extensions).

(i) *Theorem 7.7 is restricted to $A = I$ and to the coefficient c being constant, as it relies on [43], which in turn builds on the regularity theory developed in [44]. The results of [43] can be generalized to A and c that satisfy* (7.1) *using the results from [42]. In turn, Theorem 7.7 could be generalized to this setting as well.*

(ii) *Theorem 7.7 can be expected to generalize to $\Omega \subset \mathbb{R}^d$ with $d > 2$ if $\partial\Omega$ is analytic. The underlying reason for this is that the boundary layers are structurally a one dimensional phenomenon, which can be resolved with anisotropic refinement transversal to $\partial\Omega$. The approximation result* (7.11) *can therefore be expected to hold. However, the complexity is then $\mathcal{N}_{\Omega,\mathcal{Y}} = \mathcal{O}(qM^{d+2})$, resulting in an exponential convergence bound of* $\exp(-b' \mathcal{N}_{\Omega,\mathcal{Y}}^{1/(d+3)})$.

(iii) *Theorem 7.7 does generalize to so-called "bounded, curvilinear polygonal domains" $\Omega \subset \mathbb{R}^2$. The analogue of Proposition 7.1, i.e., a rigorous convergence analysis of hp-FEM in Ω for the single-scale reaction diffusion problem with the appropriate mesh refinement towards the corners of Ω is available in [42].* ■

8. Numerical experiments. In the numerical experiments, we apply the different discretization techniques of the present paper to the following four problems.

PROBLEM 8.1 (smooth solution). *We consider $A = I$, $c = 0$, and $\Omega = \Omega_L$, where $\Omega_L \subset \mathbb{R}^2$ is the L-shaped polygonal domain determined by the vertices*

$$\boldsymbol{c} \in \{(0,0), (1,0), (1,1), (-1,1), (-1,-1), (0,-1)\}.$$

The solution u with corrresponding right-hand side f are prescribed as follows (recall $x' = (x_1, x_2) \in \Omega$):

$$u(x_1, x_2) = \sin \pi x_1 \sin \pi x_2, \quad f(x_1, x_2) = (2\pi^2)^s \sin \pi x_1 \sin \pi x_2. \tag{8.1}$$

Obviously, $u \in C^\infty(\overline{\Omega})$. Since u is an eigenfunction of the Dirichlet-Laplacian, both u and $f \in \mathbb{H}^t(\Omega)$ for any $t \in \mathbb{R}$. ■

PROBLEM 8.2 (solution with corner singularities). *We consider $A = I$, $c = 0$, $\Omega = \Omega_L$ with the L-shaped domain Ω_L of Problem 8.1. The right-hand side f is*

$$f(x_1, x_2) \equiv 1\ . \tag{8.2}$$

We note $f \in \mathbb{H}^{1-s}(\Omega)$ for $s \in (1/2, 1)$ and that the solution u has corner singularities at vertices of Ω_L, the strongest one at the re-entrant corner. ■

PROBLEM 8.3 (variable coefficients). *We consider $\Omega = (0,1)$ and $A(x') = 1 + \sin(\pi x')$, $c \equiv 0$, and $f \equiv 1$. We note $f \in \mathbb{H}^{1-s}(\Omega)$ for $s \in (1/2, 1)$.* ■

Problem 8.4 (analytic, incompatible data). *We consider* $\Omega = (0,1)$ *and* $A \equiv 1$, $c \equiv 0$, *and* $f \equiv 1$. *We note* $f \in \mathbb{H}^{1-s}(\Omega)$ *for* $s \in (1/2, 1)$. *For* $s \in (0, 1/2)$, *the data* f *is* incompatible *(in the sense defined in Remark 2.2), and we only have* $f \in \mathbb{H}^{1/2-\delta}(\Omega)$, $\delta > 0$. ∎

The error measure in the numerical experiments will always be the energy norm

$$\|u - \operatorname{tr}_\Omega \mathscr{U}_{h,M}\|^2_{\mathbb{H}^s(\Omega)} \lesssim \|\nabla(\mathscr{U} - \mathscr{U}_{h,M})\|^2_{L^2(y^\alpha,\mathcal{C})} = d_s \int_\Omega f(u - \operatorname{tr}_\Omega \mathscr{U}_{h,M}),$$

where $\mathscr{U}_{h,M}$ denotes the discrete solution in $\mathcal{C}_{\mathcal{Y}}$. The exact solution is not known for Problems 8.2, 8.3, 8.4 so that a comparison with a more accurate numerical solution is performed.

Remark 8.5 (algorithmic details). *For the chosen discrete spaces the mass and stiffness matrices in* Ω *and* $(0, \mathcal{Y})$ *are computed. We then numerically solve the generalized eigenvalue problem* (6.3)*, thereby arriving at* $\mathcal{M}$ *decoupled linear systems: Find* $U_i \in S^q_0(\Omega, \mathcal{T})$ *such that*

$$a_{\mu_i,\Omega}(U_i, V) = d_s v^i(0) \int_\Omega f V \,\mathrm{d}x' \qquad \forall V \in S^q_0(\Omega, \mathcal{T}), \tag{8.3}$$

where $a_{\mu_i,\Omega}$ *is defined in* (6.4)*. Following* (6.7)*, the solution is then obtained by*

$$\mathscr{U}_{h,M}(x', y) = \sum_{i=1}^{\mathcal{M}} v^i(y) U_i(x').$$

The implementation was done in Matlab R2017a, with the generalized eigenvalue problem solved with `eig` *and the decoupled linear systems by a direct solver, i.e., Matlab's "backslash" operator.* ∎

8.1. P_1-FEM in Ω with radical meshes in $(0, \mathcal{Y})$. In the following examples we make use of the family of graded meshes $\mathcal{G}^k_{gr,\eta}$ as described in Section 5.4.2 with the particular choices $\eta = 2/s$, $k = h/2$, and $\mathcal{Y} = |\log h|$, where h denotes the meshwidth of the mesh in Ω to be described next.

8.1.1. Smooth solution. In our first experiment we study Problem 8.1 with smooth solution (8.1). We use the P_1-FEM in Ω on a hierarchy of uniformly refined meshes $\mathcal{T}^\ell$. The results are displayed in Figure 8.1. As the theory predicts we see first order convergence in the energy norm with respect to the meshwidth h for both cases $s = 1/4$ and $s = 3/4$.

8.1.2. Mesh refinement at $(0,0)$. In the next experiment we consider Problem 8.2 with $s = 3/4$. We note that $f \in \mathbb{H}^{1-s}(\Omega)$. As in Sec. 8.1.1, we use the graded mesh $\mathcal{G}^k_{gr,\eta}$ in $(0, \mathcal{Y})$. We compare two types of meshes in Ω: uniformly refined meshes and a hierarchy $\{\mathcal{T}^\ell_\beta\}_{\ell \geq 0}$ of bisection–tree meshes in Ω that are suitably refined towards the re-entrant corner at $(0,0)$ as constructed in [30]. Indeed, for this particular problem and approxmiation by P_1-FEM, mesh refinement near the convex corners of Ω is not necessary since in fact the H^2-regularity is lost *only* near the non-convex vertex of Ω. In Figure 8.2 we observe for the corner-refined meshes linear convergence with respect to the meshwidth as predicted by Theorem 5.10 whereas uniform refinement leads to sublinear convergence.

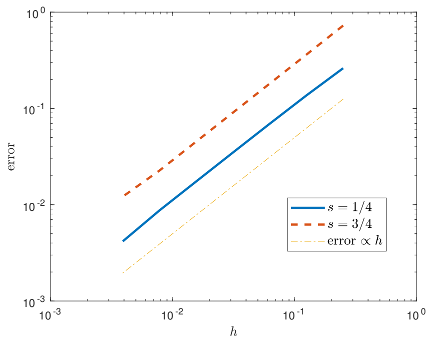

Fig. 8.1. *(cf. Sec. 8.1.1) Error in the energy norm versus the meshwidth in* Ω *with the (smooth) solution given in* (8.1)*. A* P_1*-FEM on uniformly refined meshes in* Ω *and* P_1*-FEM on radical meshes in* $(0, \mathcal{Y})$ *is used.*

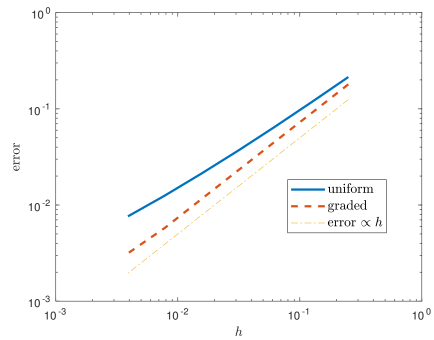

Fig. 8.2. *(cf. Sec. 8.1.2) Error in the energy norm versus meshwidth in* Ω *with the right-hand side* $f \equiv 1$ *and* $s = 3/4$*. Error graphs are shown for a* P_1*-FEM on uniformly and corner-refined refined meshes in* Ω*.*

8.1.3. Sparse grid P_1-FEM with mesh refinement at $(0,0)$. We consider Problem 8.2 with corner-refined meshes in Ω as in Section 8.1.2. While the discrete spaces used in Section 8.1.2 achieve the optimal order convergence with respect to the number of degrees of freedom $\mathcal{N}_\Omega$, the number of degrees of freedom $\mathcal{N}_{\Omega,\mathcal{Y}}$ in the extended problem is of size $\mathcal{O}(\mathcal{N}_\Omega^{1+1/2} \log\log \mathcal{N}_\Omega)$, i.e., it grows superlinearly with respect to $\mathcal{N}_\Omega$. To reduce the complexity to nearly linear, we use sparse grids as explained in Section 5.4.4; see in particular the combination formula described in Remark 5.12. In Figure 8.3, we show the error versus $\mathcal{N}_{\Omega,\mathcal{Y}}$ for a) the full tensor space already described in Section 8.1.2, b) the sparse grids approach of Section 5.4.4 (based on the corner-refined P_1-FEM in Ω and the radial mesh in y), and c) the hp-FEM in the y-variable described in detail in the following Section 8.2. Figure 8.3 shows that the use of sparse grids dramatically reduces the number of degrees of freedom and provides accuracy vs. number of degrees of freedom comparable to hp-FEM in the y-variable.

8.2. P_1-FEM in Ω with hp-FEM in $(0, \mathcal{Y})$. We again start with Problem 8.1 with smooth solution (8.1). P_1-FEM on uniformly refined meshes is used in Ω, whereas in the extended direction y we use hp-discretization on the geometric meshes $\mathcal{G}^M_{geo,\sigma}$ on $[0, \mathcal{Y}]$. We use $\mathcal{Y} = \frac{1}{3}|\log_2 h|$, $M = |\log_2(h/2)|$, $\sigma = 0.05$, and linear degree vector $\boldsymbol{r}$ with slope $\mathfrak{s} = 2$. First order convergence, as predicted by theory, can be seen in Figure 8.4. In Table 8.1 we give more details on the various parameters we used in the numerical experiments: the size of the meshwidth in Ω, the polynomial degrees in the y direction, and the number of decoupled linear systems solved. In Table 8.2 we show how the results change when the grading factor is increased to $\sigma = 0.1$. This has little effect on the error for $s = 0.75$. However, the error for $s = 0.25$ increases significantly, and the convergence order in h is reduced from 1 to about 0.8. Increasing M recovers the optimal convergence order. Namely, further numerical experiments show that the optimal convergence order for $s = 0.25$ is recovered by choosing $M = 2|\log_2(h/2)|$.

We also consider Problem 8.2 with $s = 0.75$. This time we show convergence versus the number of degrees of freedom $\mathcal{N}_{\Omega,\mathcal{Y}}$ in the extended problem and compare with P_1-FEM in Ω on corner-refined meshes as described in Section 8.1.2. We report the computations in Figure 8.3. We obtain nearly optimal complexity as predicted

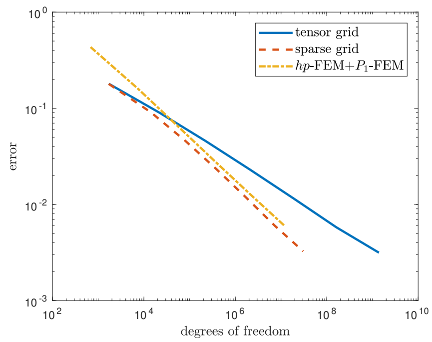

FIG. 8.3. *(cf. Sec. 8.1.3) Error in the energy norm versus the number of degrees of freedom of the extended problem with $f \equiv 1$, $s = 3/4$. A P_1-FEM on corner-refined meshes is used in Ω. We compare a) full tensor grid (radical mesh in $(0, \mathcal{Y})$), b) sparse grids (radical mesh in $(0, \mathcal{Y})$), and c) hp-FEM in $(0, \mathcal{Y})$.*

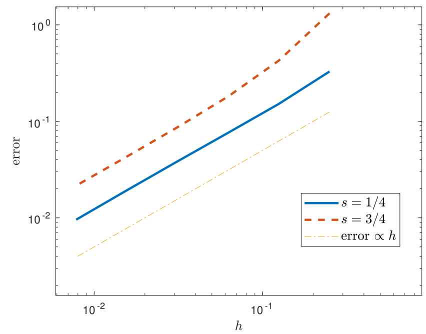

FIG. 8.4. *(cf. Sec. 8.2) Error in the energy norm versus the meshwidth in Ω with the (smooth) solution given in (8.1) for $s = 1/4$ and $s = 3/4$. A P_1-FEM on uniformly refined meshes in Ω and hp-FEM in $(0, \mathcal{Y})$ is used.*

h	p	no. linear sys.	error ($s = 0.25$)	error ($s = 0.75$)
2^{-2}	1	4	0.329	1.32
2^{-3}	2	9	0.152	0.427
2^{-4}	3	16	0.0761	0.181
2^{-5}	4	25	0.0385	0.0869
2^{-6}	5	36	0.0193	0.0431
2^{-7}	6	49	0.00956	0.0216

TABLE 8.1

Meshwidth h of the P_1-FEM on uniform meshes in Ω, highest polynomial degree used in the hp-discretization in the extended direction, number of decoupled linear systems solved and resulting errors in the energy norm for $s = 0.25$ and $s = 0.75$ and with the grading factor $\sigma = 0.05$.

by theory, but interestingly in this example slightly worse behavior compared with sparse grids.

8.3. P_1-FEM in $(0, 1)$ with hp-FEM in $(0, \mathcal{Y})$ and variable A. Next, we investigate the case of a variable diffusion coefficient in Problem 8.3 with $s = 3/4$. In the domain $\Omega = (0, 1)$ we use a uniform mesh and set $\sigma = 0.1$, whereas the remaining parameters for the hp-FEM in the y-direction are chosen as in the previous Section 8.2. As the exact solution is not available, the convergence in the energy norm is estimated numerically with respect to an accurate numerical solution obtained on a finer mesh. The expected first order convergence is seen in Figure 8.5.

8.4. hp-FEM in $(0, 1) \times (0, \mathcal{Y})$. We study the hp-FEM in x and in y for Problem 8.4 with the aim to show that exponential convergence can be achieved even in the case of incompatible data f. We comment that the solution of Problem 8.4 behaves like

$$u(x') \sim \begin{cases} \operatorname{dist}(x', \partial\Omega) + v(x') & \text{for} \quad 1/2 < s < 1\,, \\ \operatorname{dist}(x', \partial\Omega) \log \operatorname{dist}(x', \partial\Omega) + v(x') & \text{for} \quad s = 1/2\,, \\ \operatorname{dist}(x', \partial\Omega)^{2s} + v(x') & \text{for} \quad 0 < s < 1/2\,, \end{cases} \tag{8.4}$$

h	p	no. linear sys.	error ($s = 0.25$)	error ($s = 0.75$)
2^{-2}	1	4	0.353	1.33
2^{-3}	2	9	0.191	0.443
2^{-4}	3	16	0.114	0.184
2^{-5}	4	25	0.0682	0.0872
2^{-6}	5	36	0.0404	0.0431
2^{-7}	6	49	0.0237	0.0215

Table 8.2

Larger grading factor $\sigma = 0.1$ than in Table 8.1, but otherwise the same parameters. Notice little difference in the error for $s = 0.75$, but increased errors for $s = 0.25$.

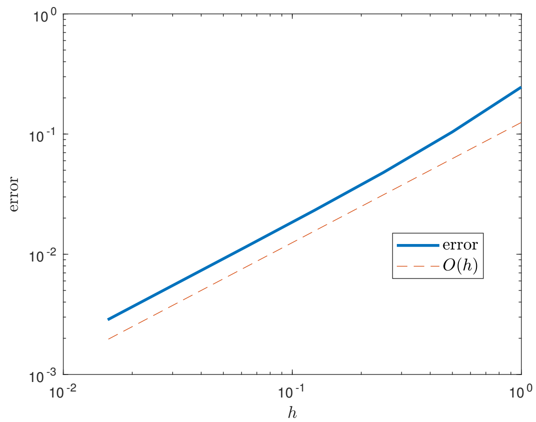

Fig. 8.5. *(cf. Sec. 8.3). Energy norm convergence for P_1-FEM in $(0, 1)$ and hp-FEM in $(0, \mathcal{Y})$ for variable diffusion coefficient A.*

with v denoting a smoother remainder. In (8.4), the first and third case are shown in [18], whereas the "borderline" case $s = 1/2$ follows from [22, Thm1.1 with Eqn. (0.2) and Example 1.6], upon observing that for $s = 1/2$ the differential operator in the Caffarelli-Silvestre extension is the Laplacian in $d + 1$ dimension. Hence, the singular support of u is $\partial\Omega$, i.e., u exhibits an *algebraic boundary singularity* (distinct from the smooth exponential boundary layers arising in linear, elliptic-elliptic singular perturbations) near the boundary of Ω; see Figure 8.6.

In our numerical convergence studies, we compare the numerical solution with an accurate solution obtained on a finer grid.

In $(0, \mathcal{Y})$, we use the same geometric hp-FEM space $\mathcal{G}^M_{geo,\sigma}$ as in Section 8.2. The hp-FEM space $S^q_0(\Omega, \mathcal{T}_L)$ is based on the geometric mesh $\mathcal{T}_L$, which is scaled version of the geometric mesh $\mathcal{T}^{1D}_{geo,L}$ described in Proposition 7.1. We select $q = M$ and $L = M$. Exponential convergence with respect to the polynomial degree q as predicted by the theory is shown in Figure 8.7.

In Figure 8.8 we illustrate the behavior of the solution given by (8.4). We also investigate numerically the borderline case $s = 1/2$ in Figure 8.9. Even if the domain Ω is smooth, u exhibits in general a *boundary singularity* with singular support $\partial\Omega$. For $s = 1/2$ and polygonal Ω, this boundary singularity is the trace, at $y = 0$, of an *edge singularity* of the solution $\mathscr{U}$ of the extended problem (1.2) in $\mathcal{C}$ whose structure is known; see, for instance, [23] and the references therein. For $d = 2$, the hp-FE

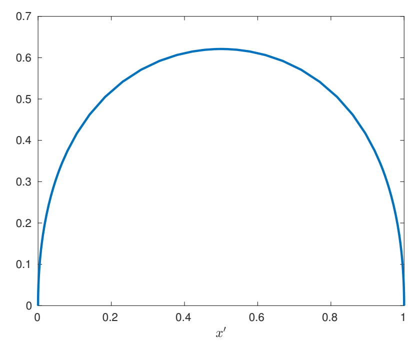

FIG. 8.6. *Solution profile for Problem 8.4 for $s = 0.25$ with algebraic boundary singularity.*

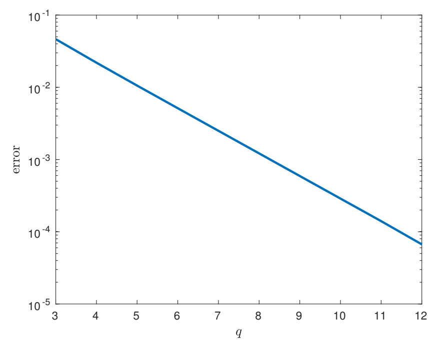

FIG. 8.7. *(cf. Sec. 8.4) Exponential convergence in energy norm of the hp-FEM on $\Omega \times (0, \mathcal{Y})$ versus polynomial order q for $s = 0.25$ and incompatible $f \equiv 1$.*

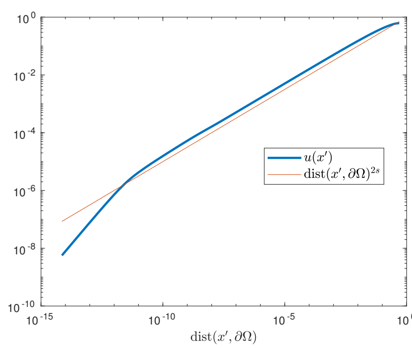

FIG. 8.8. *Numerical verification of the algebraic boundary singularity behavior* (8.4) *of the solution of Problem 8.4 for $s = 1/4$. Note that the change in the slope (from $1/2$ to 1) near the boundary is a numerical artifact – as the approximation is improved, the kink moves to the left.*

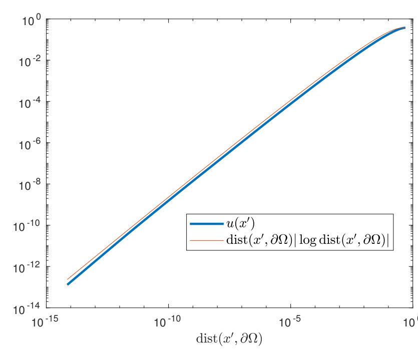

FIG. 8.9. *Solution profile for Problem 8.4 for $s = 1/2$ near the boundary. The numerical solution is compared with* $\mathrm{dist}(x', \partial\Omega)|\log \mathrm{dist}(x', \partial\Omega)|$.

geometric boundary layer meshes in Figure 7.1 appear naturally as traces at $y = 0$ of the 3-dimensional geometric meshes in $\mathcal{C}_{\mathcal{Y}}$ developed in [57, 58].

9. Conclusions and generalizations. In the course of this work, we introduced and analyzed four different types of *local* FEM discretizations for the numerical approximation of the spectral fractional diffusion problem (1.1) in a bounded polygonal domain $\Omega \subset \mathbb{R}^2$ with straight sides (or a bounded interval $\Omega \subset \mathbb{R}$), subject to homogeneous Dirichlet boundary conditions. Our local FEM schemes are based on the Caffarelli-Silvestre extension of (1.1) from Ω to $\mathcal{C}$. Our main contributions are the following.

- **General operators and nonconvex domains.** We proposed a tensor product argument for continuous, piecewise linear FEM in both $(0, \infty)$, and in Ω with proper mesh refinement towards $y = 0$ and the corners $\boldsymbol{c}$ of Ω. Assuming that A and c are as in Proposition 5.3, we showed that the approximate solution to problem (1.1)

exhibits a near optimal asymptotic convergence rate $\mathcal{O}(h_\Omega|\log h_\Omega|)$ subject to the optimal regularity $f \in \mathbb{H}^{1-s}(\Omega)$. However, if $\mathcal{N}_\Omega$ denotes the number of degrees of freedom in the discretization in Ω, then the total number of degrees of freedom grows asymptotically as $\mathcal{O}(\mathcal{N}_\Omega^{3/2})$ (ignoring logarithmic factors).
This result is analogous to the bounds obtained in [48] for convex domains Ω, thus generalizing these results to *nonconvex,* polygonal domains $\Omega \subset \mathbb{R}^2$. The error analysis proceeded by a suitable form of quasi-optimality in Lemma 5.1 and the construction of a tensor product FEM interpolant in the truncated cylinder $\mathcal{C}_\mathcal{Y}$. This interpolant was constructed from a nodal, continuous and piecewise linear interpolant $\pi_\eta^{1,\ell}$ with respect to the extended variable $y \in (0,\mathcal{Y})$ on a radical-geometric mesh, and from an $L^2(\Omega)$ projection Π_β^ℓ in Ω onto the space of continuous, piecewise linears on a suitable sequence $\{\mathcal{T}_\beta^\ell\}_{\ell\geq 0}$ of regular nested, bisection-tree, simplicial meshes with refinement towards the corners $\boldsymbol{c}$ of Ω. A novel result from [29] implies that Π_β^ℓ is also uniformly $H^1(\Omega)$-stable with respect to the refinement level ℓ. The present construction would likewise work with any other concurrently $L^2(\Omega)$ and $H^1(\Omega)$ stable family of quasi-interpolation operators, e.g. those of [62].

- **Sparse tensor grids.** While the regularity requirement $f \in \mathbb{H}^{1-s}(\Omega)$ is, essentially, minimal for first order convergence in Ω, the complexity $\mathcal{O}(\mathcal{N}_\Omega^{3/2})$ due to the extra degrees of freedom in the extended variable results in superlinear work with respect to $\mathcal{N}_\Omega$. We therefore proposed in Section 5.4.4 a *novel, sparse tensor product FE discretization* of the truncated, extended problem. Using novel regularity results for the extended solution in $\mathcal{C}$ in weighted spaces and sparse tensor product constructions of the interpolation operators $\pi_\eta^{1,\ell}$ and Π_β^ℓ in Ω, we proved that this approach still delivers FEM solutions of (1.1) with essentially first order convergence rates (i.e., up to logarithmic factors), under the slightly more stringent regularity $f \in \mathbb{H}^{1-s+\nu}(\Omega)$, $\nu > 0$, while requiring essentially only $\mathcal{O}(\mathcal{N}_\Omega)$ many degrees of freedom.
- ***hp*-FE approximation in the extended variable.** The solution of the extended problem being analytic with respect to the extended variable $y > 0$ allows for designing *hp*-FE approximations with respect to the variable y on geometric meshes and proving *exponential convergence rates* even under finite regularity of A, c, and f as specified in Proposition 5.3. The proof is based on a novel framework of countably normed, weighted Bochner spaces in $(0,\infty)$ to quantify the analytic regularity with respect to y. We also developed a corresponding family of *hp*-interpolation operators *that affords exponential convergence rates* in the extended variable.
Upon tensorization with the projectors Π_β^ℓ onto spaces of continuous, piecewise linear finite elements on simplicial, bisection-tree meshes with corner refinement in Ω, we obtained a class of FE schemes that afford essentially optimal, linear convergence rate in Ω under the regularity $f \in \mathbb{H}^{1-s}(\Omega)$, also for *nonconstant coefficients and nonconvex polygonal domains* Ω*,* thereby generalizing [40]. We remark that the convergence rate bounds essentially equal the results of so-called wavelet Galerkin discretizations for the integral fractional Laplacian (see [55, 54] and the references therein). Wavelet Galerkin methods are based on direct, "nonlocal" Galerkin discretization of integro-differential operators, which entail numerical evaluation of singular integrals and dense stiffness matrices, neither of which occurs in the present local FE approach. However, these methods can also cope with variable exponent $s(x')$, which seems to be beyond reach with the present approach; see [56, 21] and the references therein. We also point out that the *boundary compatibility of* f, which is implicit in the assumption $f \in \mathbb{H}^{1-s}(\Omega)$, is essential in the arguments in

Section 5 as well as in the results of [48, 40, 13].

- **Diagonalization.** We further developed a diagonalization approach, first proposed in [38] for the Laplacian in rectangular domains, which allows us to decouple the second order elliptic system in $\mathcal{C}_{\mathcal{Y}}$, resulting from *any* Galerkin semidiscretization in the extended variable y (either of h-FEM or of hp-FEM type) of the truncated problem, into a finite number of decoupled, singularly perturbed, second order elliptic problems in Ω. This approach is instrumental for both the design of hp-FEMs in Ω in Section 7 as well as the implementation of parallel and inexact solvers in Section 8.
- ***hp*-FEMs.** Exploiting results on *robust exponential convergence of hp-FEMs* for second order, singularly perturbed problems [44, 43, 41, 42], and tensorization with the exponentially convergent hp-FEM in $(0, \mathcal{Y})$ resulted in *exponential convergence for analytic input data A, c, f, and* Ω for *incompatible* forcing f (i.e. $f \in H^{1-s}(\Omega)$ but $f \notin \mathbb{H}^{1-s}(\Omega)$). The boundary incompatibility of f leads to the formation of a strong boundary singularity for $0 < s \leq 1/2$ and a weaker one for $s > 1/2$ with $\partial\Omega$ analytic, which is a genuine fractional diffusion effect. Our analysis in Section 7.2 revealed that for incompatible data f in space dimension $d > 1$, *anisotropic, geometric meshes in* Ω capable of resolving boundary layers over a wide range of length scales, are generally indispensable, even if $\partial\Omega$ is smooth. Section 8 displays an example. For incompatible data f of finite Sobolev regularity in Ω our analysis indicates that optimal, algebraic convergence rates require anisotropic mesh refinement towards $\partial\Omega$, in addition to corner refinement which suffices for compatible data, ie., for $f \in \mathbb{H}^{1-s}(\Omega)$.

The following generalizations of the results of the present work suggest themselves.

- **Boundary conditions.** The present analysis was limited to polygonal domains in two space dimensions and to homogeneous Dirichlet boundary conditions. The extension (1.2) is also available for homogeneous Neumann boundary conditions in [18, Section 7] and for combinations of Dirichlet and Neumann boundary conditions on parts of $\partial\Omega$. Solutions $\mathscr{U}$ of these extensions also admit the representation (4.1), so that the analytic regularity results in Section 4 extend almost verbatim. Likewise, all regularity results in Section 5, being based on [9], extend verbatim to homogeneous Neumann and Dirichlet-Neumann boundary conditions on polygonal domains Ω.
- **Higher dimensions and elements of degree $q \geq 2$ in Ω.** Analogous results as in Section 5 hold for polyhedral domains $\Omega \subset \mathbb{R}^3$ with plane faces, using corresponding regularity results for the Dirichlet Laplacian in weighted spaces in the polyhedron Ω, combined with corresponding FE projections on anisotropically refined FE meshes (with corner and edge-refinements in Ω), as described in [5].
  Returning to polygons, if we consider piecewise polynomials of degree $q \geq 2$ on families of simplicial meshes which are sufficiently refined towards the vertices $\boldsymbol{c}$ of Ω, we expect algebraic convergence rates higher than for linear elements *provided* the compatible forcing $f \in \mathbb{H}^{q-s}(\Omega)$. This implies, in particular, that f should satisfy besides $f \in H^{q-s}_{loc}(\Omega)$ also certain higher-order boundary compatibility on $\partial\Omega$, a consequence of the eigenfunction expansions used in our regularity analysis.

Appendix A. Proof of Lemma 5.15. We will only show (5.53), (5.54) as the estimates (5.55), (5.56) are proved using similar arguments; see, for instance, the proof of [6, Theorem 8]. We distinguish between the first element I_1, the terminal element I_M, and the remaining ones. We write $h_i = |I_i|$. We simplify the exposition by assuming $X = \mathbb{R}$. It is convenient to define, for each interval I_i, $i = 2, \ldots, M$, the quantity C_i by

$$C_i^2 := \sum_{\ell=1}^{\infty} (2K_u)^{-\ell} \frac{1}{\ell!^2} \|u^{(\ell)}\|^2_{L^2(\omega_{\alpha+2\ell-2\beta,\gamma}, I_i)}. \tag{A.1}$$

We observe that, since $u \in \mathcal{B}^1_{\beta,\gamma}(C_u, K_u)$,

$$\sum_{i=2}^{M} C_i^2 \le 2C_u^2, \tag{A.2}$$

where, we recall that the space $\mathcal{B}^1_{\beta,\gamma}(C_u, K_u)$ corresponds to a class of analytic functions and is defined as in (5.51). We begin the proof with an auxiliary result about linear interpolation on the reference element.

Lemma A.1 (linear interpolant). *Let X be a Hilbert space, $\widehat{K} = (0,1)$, and let $\widetilde{\pi}_1$ be the linear interpolant in the points $1/2$, 1. Let $\alpha > -1$ and $\delta \le 1$. Then, for $u \in C((0,1]; X)$ and provided the terms on the right-hand side are finite, we have*

$$\int_{\widehat{K}} y^{\alpha} \|u - \widetilde{\pi}_1 u\|_X^2 \, \mathrm{d}y \lesssim \int_{\widehat{K}} y^{\alpha+2\delta} \|u'\|_X^2 \, \mathrm{d}y, \tag{A.3}$$

$$\int_{\widehat{K}} y^{\alpha} \|(u - \widetilde{\pi}_1 u)'\|_X^2 \, \mathrm{d}y \lesssim \int_{\widehat{K}} y^{\alpha+2\delta} \|u''\|_X^2 \, \mathrm{d}y, \tag{A.4}$$

where the hidden constant is independent of u.

Proof. For notational simplicity, we will prove the lemma only for the case $X = \mathbb{R}$.

We begin with the proof of (A.3). Since $(u - \widetilde{\pi}_1 u)(1) = 0$ we have, for $y \in \widehat{K}$,

$$(u - \widetilde{\pi}_1 u)(y) = \int_1^y (u - \widetilde{\pi}_1 u)'(t) \, \mathrm{d}t,$$

so that

$$\int_0^1 y^{\alpha} |u - \widetilde{\pi}_1 u|^2 \, \mathrm{d}y \le 2 \int_0^1 y^{\alpha} \left| \int_y^1 |u'(t)| \, \mathrm{d}t \right|^2 \mathrm{d}y + 2 \int_0^1 y^{\alpha} \left| \int_y^1 |(\widetilde{\pi}_1 u)'(t)| \, \mathrm{d}t \right|^2 \mathrm{d}y.$$

From Hardy's inequality (e.g., [25, Chapter 2, Theorem 3.1]) we infer

$$\int_0^1 y^{\alpha} \left| \int_y^1 |u'(t)| \, \mathrm{d}t \right|^2 \mathrm{d}y \le (\alpha+1)^{-2} \int_0^1 y^{\alpha+2} |u'(y)|^2 \, \mathrm{d}y.$$

From $(\widetilde{\pi}_1 u)' = 2\int_{1/2}^1 u'(t) \, \mathrm{d}t$ we obtain $|(\widetilde{\pi}_1 u)'|^2 \le C \int_{1/2}^1 t^{\alpha+2\delta} |u'(t)|^2 \, \mathrm{d}t$ and therefore, in view of $\alpha > -1$, the estimate

$$\int_0^1 y^{\alpha} |(\widetilde{\pi}_1 u)'|^2 \, \mathrm{d}y \lesssim \int_0^1 y^{\alpha+2} |u'(y)|^2 \, \mathrm{d}y.$$

This concludes the proof of (A.3) for the case $\delta = 1$. Since the integration range is $y \in \widehat{K} = (0,1)$, we may replace $y^{\alpha+2}$ by $y^{\alpha+2\delta}$.

Next, we prove (A.4) and assume that the right-hand side of (A.4) is finite. Again, it suffices to consider the limiting case $\delta = 1$ and use Hardy's inequality. We mention in passing that (A.5) below actually shows that $u \in C^1((0,1];X)$. We write

$$(u-\widetilde{\pi}_1 u)'(y) = u'(y) - 2\int_{1/2}^1 u'(t)\ \mathrm{d}t = 2\int_{1/2}^1 (u'(y)-u'(t))\ \mathrm{d}t$$
$$= 2\int_{1/2}^1 \int_t^y u''(\tau)\ \mathrm{d}\tau\ \mathrm{d}t. \tag{A.5}$$

Therefore,

$$\begin{aligned}\int_0^1 y^\alpha |(u-\widetilde{\pi}_1 u)'(y)|^2\ \mathrm{d}y &= 4\int_0^1 y^\alpha \left|\int_{1/2}^1\int_t^y u''(\tau)\ \mathrm{d}\tau\ \mathrm{d}t\right|^2\ \mathrm{d}y \\ &\leq 2\int_{1/2}^1\int_0^1 y^\alpha \left|\int_t^y |u''(\tau)|\ \mathrm{d}\tau\right|^2 \mathrm{d}y\,\mathrm{d}t \\ &\lesssim \int_{1/2}^1\int_0^1 y^\alpha \left[\left|\int_y^1 |u''(\tau)|\ \mathrm{d}\tau\right|^2 + \left|\int_t^1 |u''(\tau)|\ \mathrm{d}\tau\right|^2\right] \mathrm{d}y\,\mathrm{d}t \\ &\lesssim \int_0^1 y^{\alpha+2}|u''(y)|^2\ \mathrm{d}y + \int_{1/2}^1 y^{\alpha+2}|u''(y)|^2\ \mathrm{d}y,\end{aligned}$$

where, in the last step we applied Hardy's inequality. Lemma A.1 is thus proved. □

With Lemma A.1 we can estimate the error on the first element I_1 as follows: Scaling the estimate (A.3) gives

$$\|u-\pi_y^{\boldsymbol{r}} u\|_{L^2(\omega_{\alpha,0},I_1)} \leq C h_1^\beta \|u'\|_{L^2(\omega_{\alpha+2-2\beta,0},I_1)}, \tag{A.6}$$

and the assumption $|I_1| = \sigma^M \mathcal{Y} \leq 1$ allows us to insert the weight $e^{\gamma y}$ on both sides of (A.6).

We now bound the error contributions from elements away from the origin, i.e., on the elements I_i, $i = 2,\dots,M$. These elements satisfy $h_i\sigma/(1-\sigma) = \mathrm{dist}(I_i,0)$. For $I_i = (y_{i-1},y_i)$ the pull-back $\widehat{u}_i := u|_{I_i} \circ F_{I_i}$ satisfies

$$\begin{aligned}&\|\widehat{u}_i^{(\ell+1)}\|^2_{L^2(-1,1)} = (h_i/2)^{-1+2(\ell+1)}\|u^{(\ell+1)}\|^2_{L^2(I_i)} \\ &\leq (h_i/2)^{-1+2(\ell+1)} e^{-\gamma y_{i-1}} \max_{y\in I_i} y^{-\alpha-2(\ell+1)+2\beta}\|u^{(\ell+1)}\|^2_{L^2(\omega_{\alpha+2(\ell+1)-2\beta,\gamma},I_i)} \\ &\lesssim e^{-\gamma y_{i-1}} h_i^{-1+2(\ell+1)} h_i^{-\alpha-2(\ell+1)+2\beta}(2(1-\sigma))^{-2(\ell+1)} C_i^2 (2K_u)^{2(\ell+1)}(\ell+1)!^2,\end{aligned}$$

where in the last step we have used (A.1). The operator $\widehat{\Pi}_r$ given by Lemma 5.14 then yields the existence of a $b>0$ that depends solely on K_u and σ for which

$$\|\widehat{u} - \widehat{\Pi}_{\boldsymbol{r}_i}\widehat{u}\|_{L^2(-1,1)} \lesssim C_i e^{-\gamma y_{i-1}} e^{-b\boldsymbol{r}_i} h_i^{-(1+\alpha)/2+\beta}\ .$$

Scaling back to I_i and using again $h_i \sim \mathrm{dist}(I_i,0)$ yields

$$\|u-\pi_y^{\boldsymbol{r}} u\|^2_{L^2(\omega_{\alpha,\gamma},I_i)} \leq C h_i^{2\beta} C_i^2 e^{-2b\boldsymbol{r}_i}.$$

Summation over i and taking the slope of the linear degree vector sufficiently large (see, for instance, the proof of [6, Theorem 8] for details) gives

$$\sum_{i=2}^M \|u-\pi_y^{\boldsymbol{r}} u\|^2_{L^2(\omega_{\alpha,\gamma},I_i)} \lesssim \mathcal{Y}^{2\beta} e^{-2b'M}$$

for suitable $b' > 0$. Combining this with (A.6) gives the desired (5.53).

It remains to prove (5.54). We begin with a preparatory result.

LEMMA A.2 (exponential decay). *Let X be a Hilbert space, and let $\delta \in \mathbb{R}$, $\gamma > 0$, $\mathcal{Y}_0 > 0$. Then the following holds for $u \in C^1((\mathcal{Y}_0, \infty); X)$ in items (i), (ii) and for $u \in C^2((\mathcal{Y}_0, \infty); X)$ in items (iii), (iv) with implied constants depending solely on δ, γ, and $\mathcal{Y}_0$:*

(i) If $\lim_{y\to\infty} u(y) = 0$ and $\|u'\|_{L^2(\omega_{\delta,\gamma},(\mathcal{Y}_0,\infty);X)} < \infty$, then

$$\|u(\mathcal{Y})\|_X \lesssim \mathcal{Y}^{-\delta/2} \exp(-\mathcal{Y}\gamma/2)\|u'\|_{L^2(\omega_{\delta,\gamma},(\mathcal{Y},\infty);X)} \qquad \forall \mathcal{Y} \geq \mathcal{Y}_0. \tag{A.7}$$

(ii) If $\sum_{j=0}^{1} \|u^{(j)}\|_{L^2(\omega_{\delta,\gamma},(\mathcal{Y}_0,\infty);X)} < \infty$, then $\lim_{y\to\infty} u(y) = 0$.

(iii) If $\lim_{y\to\infty} u^{(j)}(y) = 0$ for $j = 0$, 1 and $\|u''\|_{L^2(\omega_{\delta,\gamma},(\mathcal{Y}_0,\infty);X)} < \infty$, then

$$\|u(\mathcal{Y})\|_X \lesssim \mathcal{Y}^{-\delta/2} \exp(-\mathcal{Y}\gamma/2)\|u''\|_{L^2(\omega_{\delta,\gamma},(\mathcal{Y},\infty);X)} \qquad \forall \mathcal{Y} \geq \mathcal{Y}_0. \tag{A.8}$$

(iv) If $\sum_{j=0}^{2} \|u^{(j)}\|_{L^2(\omega_{\delta,\gamma},(\mathcal{Y}_0,\infty);X)} < \infty$, then $\lim_{y\to\infty} u(y) = \lim_{y\to\infty} u'(y) = 0$.

Proof. We will only prove items (i) and (ii) as the remaining two are proved by similar arguments.

We begin the proof with the following observation: There is a constant $c > 0$ (that depends only on δ, $\mathcal{Y}_0$, and γ) such that for $\mathcal{Y} \geq \mathcal{Y}_0$

$$\int_{\mathcal{Y}}^{\infty} y^{-\delta} \exp(-\gamma y)\, \mathrm{d}y \leq c\, \mathcal{Y}^{-\delta} \exp(-\gamma\mathcal{Y}). \tag{A.9}$$

For $\delta \geq 0$, this is immediate. For $\delta < 0$, one integrates by parts once to discover that the leading order asymptotics (as $\mathcal{Y} \to \infty$) of the integral is $\gamma^{-1} \exp(-\gamma\mathcal{Y})\mathcal{Y}^{-\delta}$.

We now proceed with the proof of (A.7): Since $\gamma > 0$, we can write

$$\| - u(\mathcal{Y})\|_X = \left\| \int_{\mathcal{Y}}^{\infty} u'(y)\, \mathrm{d}y \right\|_X \leq \sqrt{\int_{\mathcal{Y}}^{\infty} y^{-\delta} \exp(-\gamma y)\, \mathrm{d}y} \|u'\|_{L^2(\omega_{\delta,\gamma},(\mathcal{Y},\infty))},$$

and (A.7) follows from (A.9). The assertion of item (ii) follows by a similar argument, starting from $u(y) = u(\eta) + \int_\eta^y u'(t)\, \mathrm{d}t$, squaring, multiplying by $\exp(-\gamma\eta)$, and integrating in η from y to ∞. ☐

To prove (5.54) we have to estimate $u(\mathcal{Y})$. Lemma A.2, (i) shows

$$\|u(\mathcal{Y})\|_X \lesssim \mathcal{Y}^{-\alpha/2-(1-\beta)} \exp(-\mathcal{Y}\gamma/2) C_u. \tag{A.10}$$

Since $\pi^{\boldsymbol{r}}_{y,\{\mathcal{Y}\}}$ is obtained from $\pi^{\boldsymbol{r}}_y$ by a correction on the terminal element I_M, the desired (5.54) follows easily from (A.10), if we recall that $|I_M| \sim \mathcal{Y}$.

Appendix B. Analysis of the decoupling eigenvalue problem.

LEMMA B.1 (weighted Poincaré). *Let $\mathcal{Y} > 0$ and $\alpha \in (-1, 1)$. Then, for $v \in C^1((0, \mathcal{Y}])$ with $v(\mathcal{Y}) = 0$ there holds*

$$\|v\|_{L^\infty(0,\mathcal{Y})} \leq \mathcal{Y}^{(1-\alpha)/2}(1 - \alpha)^{-1/2}\|v'\|_{L^2(y^\alpha,(0,\mathcal{Y}))}. \tag{B.1}$$

Proof. From $v(\mathcal{Y}) = 0$ we get $v(y) = -\int_y^{\mathcal{Y}} v'(t)\, \mathrm{d}t$. Hence, for $y \in (0, \mathcal{Y})$,

$$\begin{aligned} |v(y)| &= \left| \int_y^{\mathcal{Y}} v'(t)\, \mathrm{d}t \right| = \left| \int_y^{\mathcal{Y}} t^{-\alpha/2} t^{\alpha/2} v'(t)\, \mathrm{d}t \right| \leq \left(\int_y^{\mathcal{Y}} t^{-\alpha}\, \mathrm{d}t \right)^{1/2} \|v'\|_{L^2(y^\alpha,(0,\mathcal{Y}))} \\ &\leq \mathcal{Y}^{(1-\alpha)/2}(1 - \alpha)^{-1/2}\|v'\|_{L^2(y^\alpha,(0,\mathcal{Y}))}, \end{aligned}$$

which finishes the proof. ◻

Lemma B.2 (eigenvalue upper bound). *Let $\mathcal{Y} > 0$ and $\alpha \in (-1,1)$. Assume that (v,μ) satisfy*

$$\|v'\|^2_{L^2(y^\alpha,(0,\mathcal{Y}))} = 1, \qquad \|v\|^2_{L^2(y^\alpha,(0,\mathcal{Y}))} = \mu, \qquad v(\mathcal{Y}) = 0. \tag{B.2}$$

Then, $0 < \mu \leq \mathcal{Y}^2(1-\alpha^2)^{-1}$.

Proof. We compute, using Lemma B.1

$$\begin{aligned}\mu = \|v\|^2_{L^2(y^\alpha,(0,\mathcal{Y}))} &= \int_0^{\mathcal{Y}} t^\alpha |v(t)|^2 \,\mathrm{d}t \leq \|v\|^2_{L^\infty(0,\mathcal{Y})} \mathcal{Y}^{1+\alpha}(1+\alpha)^{-1} \\ &\leq \mathcal{Y}^{1+\alpha}\mathcal{Y}^{1-\alpha}(1+\alpha)^{-1}(1-\alpha)^{-1}\|v'\|^2_{L^2(y^\alpha,(0,\mathcal{Y}))} = \mathcal{Y}^2(1-\alpha^2)^{-1},\end{aligned}$$

which finishes the proof. ◻

We also need lower bounds for eigenvalues.

Lemma B.3 (eigenvalue lower bound). *Let $\alpha > -1$. Let $\mathcal{G}^M$ be an arbitrary mesh on $(0,\mathcal{Y})$ with the property that for all elements I_i, $i = 2,\dots,M$, not abutting $y = 0$ there holds $|I_i| \leq C_{geo}\operatorname{dist}(I_i,0)$. Let $V_h \subset H^1(y^\alpha,(0,\mathcal{Y}))$ be a subspace of the space of piecewise polynomials of degree q on $\mathcal{G}^M$. Then, with h_{min} denoting the smallest element size,*

$$\|v'\|_{L^2(y^\alpha,(0,\mathcal{Y}))} \lesssim h_{min}^{-1} q^2 \|v\|_{L^2(y^\alpha,(0,\mathcal{Y}))} \qquad \forall v \in V_h, \tag{B.3}$$

where the hidden constant depends solely on C_{geo} and α.

Proof. We emphasize that the condition $h_i \leq C_{geo}\operatorname{dist}(I_i,0)$ is satisfied for all meshes where neighboring elements have comparable size. We also remark that (slightly) sharper estimates (in the dependence on the polynomial degree q) are possible on geometric meshes with linear degree vector. We write $h_i = |I_i|$. We will use the polynomial inverse estimate (B.6) below. For the first element $I_1 = (0,y_1)$ we calculate for $v \in V_h$ and its pull-back $\widehat{v} := v|_{I_1} \circ F_{I_1}$

$$\begin{aligned}\|v'\|^2_{L^2(y^\alpha,\widehat{K})} &= (h_1/2)^{\alpha+1-2}\int_{-1}^1 (1+y)^\alpha |\widehat{v}'(y)|^2 \,\mathrm{d}y \\ &\overset{(B.6)}{\lesssim} h_1^{\alpha+1-2} q^4 \int_{-1}^1 (1+y)^\alpha |\widehat{v}(y)|^2 \,\mathrm{d}y \sim h_1^{-2} q^4 \|v\|^2_{L^2(y^\alpha,I_1)}.\end{aligned} \tag{B.4}$$

For the remaining elements I_i, we exploit that the assumption $h_i \geq C_{geo}\operatorname{dist}(I_i,0)$ ensures that the weight is slowly varying within them, i.e.,

$$\max_{y\in I_i} y^\alpha \leq (1+C_{geo})^{|\alpha|} \min_{y \in I_i} y^\alpha, \qquad i = 2,\dots,M.$$

Hence, the polynomial inverse estimate (B.6) (with $\alpha = \beta = 0$ there) yields by scaling arguments

$$\|v'\|_{L^2(y^\alpha,I_i)} \leq C h_i^{-1} q^2 \|v\|_{L^2(y^\alpha,I_i)}. \tag{B.5}$$

Combining (B.4), (B.5) yields the result. ◻

Lemma B.4 (polynomial inverse estimate). *Let $\widehat{K} = (-1,1)$. For α, $\beta > -1$ there is $C_{\alpha,\beta} > 0$ such that for all $q \in \mathbb{N}_0$ and all $w \in \mathbb{P}_q(\widehat{K})$:*

$$\int_{-1}^1 (1+y)^\alpha (1-y)^\beta |w'(y)|^2 \,\mathrm{d}y \leq C_{\alpha,\beta} q^4 \int_{-1}^1 (1+y)^\alpha (1-y)^\beta |w(y)|^2 \,\mathrm{d}y. \tag{B.6}$$

Proof. Step 1: We assert (B.6) for $\alpha = \beta$. From [10, Chap. III, Props. 6.1, 6.3] we get for $w \in \mathbb{P}_q(\widehat{K})$

$$\int_{-1}^{1} (1-y^2)^{\alpha} |w'(y)|^2 \,\mathrm{d}y \overset{[10, Chap.\ III, Prop.\ 6.1]}{\lesssim} q^2 \int_{-1}^{1} (1-y^2)^{\alpha+1} |w'(y)|^2 \,\mathrm{d}y$$
$$\overset{[10, Chap.\ III, Prop.\ 6.3]}{\lesssim} q^4 \int_{-1}^{1} (1-y^2)^{\alpha} |w(y)|^2 \,\mathrm{d}y.$$

Simple scaling arguments imply for arbitrary (fixed) finite intervals (a,b) and $\alpha' > -1$ for all $w \in \mathbb{P}_q(a,b)$:

$$\int_a^b (y-a)^{\alpha'} (b-y)^{\alpha'} |w'(y)|^2 \,\mathrm{d}y \lesssim q^4 \int_a^b (y-a)^{\alpha'} (b-y)^{\alpha'} |w(y)|^2 \,\mathrm{d}y. \tag{B.7}$$

Step 2: We show (B.6) for $(\alpha,\beta) = (\alpha,0)$. (By symmetry, this also shows the case $(\alpha,\beta) = (0,\beta)$). Since (B.7) implies for $w \in \mathbb{P}_q(\widehat{K})$

$$\int_0^1 (1+y)^{\alpha} |w'(y)|^2 \,\mathrm{d}y \lesssim \int_0^1 |w'(y)|^2 \,\mathrm{d}y \lesssim q^4 \int_0^1 |w(y)|^2 \,\mathrm{d}y,$$

it suffices to prove the bound $\int_{-1}^{0} (1+y)^{\alpha} |w'(y)|^2 \,\mathrm{d}y \lesssim q^4 \int_{-1}^{1} (1+y)^{\alpha} |w(y)|^2 \,\mathrm{d}y$. To that end, define $\widetilde{w}(y) := (1-y)w(y) \in \mathbb{P}_{q+1}(\widehat{K})$ and note $\widetilde{w}'(y) = (1-y)w'(y) - w(y)$. Then

$$\begin{aligned}
\int_{-1}^{0} (1+y)^{\alpha} |w'(y)|^2 \,\mathrm{d}y &\le \int_{-1}^{0} (1+y)^{\alpha} |w'(y)(1-y)|^2 \,\mathrm{d}y \\
&\le 2 \int_{-1}^{0} (1+y)^{\alpha} \left(|\widetilde{w}'(y)|^2 + |w(y)|^2 \right) \mathrm{d}y \\
&\overset{(B.7)}{\lesssim} (q+1)^4 \int_{-1}^{1} (1+y)^{\alpha} \left((1-y)^{\alpha} |\widetilde{w}(y)|^2 + |w(y)|^2 \right) \mathrm{d}y \\
&\lesssim (q+1)^4 \int_{-1}^{1} (1+y)^{\alpha} \left((1-y)^{\alpha} (1-y)^2 |w(y)|^2 + |w(y)|^2 \right) \mathrm{d}y.
\end{aligned}$$

Since $\alpha > -1$, we have $(1-y)^{2+\alpha} \lesssim 1$, which allows us to conclude the proof of the case $\beta = 0$ in (B.6).

Step 3: For arbitary α, $\beta > -1$ we use (scaled versions of) Step 2:

$$\begin{aligned}
&\int_{-1}^{1} (1+y)^{\alpha} (1-y)^{\beta} |w'(y)|^2 \,\mathrm{d}y \lesssim \int_{-1}^{0} (1+y)^{\alpha} |w'(y)|^2 \,\mathrm{d}y + \int_0^1 (1-y)^{\beta} |w'(y)|^2 \,\mathrm{d}y \\
&\overset{\text{Step 2}}{\lesssim} q^4 \int_{-1}^{0} (1+y)^{\alpha} |w(y)|^2 \,\mathrm{d}y + q^4 \int_0^1 (1-y)^{\beta} |w(y)|^2 \,\mathrm{d}y \\
&\lesssim q^4 \int_{-1}^{1} (1+y)^{\alpha} (1-y)^{\beta} |w(y)|^2 \,\mathrm{d}y.
\end{aligned}$$

This concludes the proof. ∎

REFERENCES

[1] M. Abramowitz and I.A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[2] G. Acosta and J.P. Borthagaray. A fractional Laplace equation: regularity of solutions and finite element approximations. *SIAM J. Numer. Anal.*, 55(2):472–495, 2017.

[3] R.A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[4] L.V. Ahlfors. *Complex analysis*. McGraw-Hill Book Co., New York, third edition, 1978. An introduction to the theory of analytic functions of one complex variable, International Series in Pure and Applied Mathematics.

[5] T. Apel. Interpolation of non-smooth functions on anisotropic finite element meshes. *M2AN Math. Model. Numer. Anal.*, 33(6):1149–1185, 1999.

[6] T. Apel and J.M. Melenk. Interpolation and quasi-interpolation in h- and hp-version finite element spaces. In E. Stein, R. de Borst, and T.J.R. Hughes, editors, *Encyclopedia of Computational Mechanics*, pages 1–33. John Wiley & Sons, Chichester, UK, second edition, 2018. extended preprint at `http://www.asc.tuwien.ac.at/preprint/2015/asc39x2015.pdf`.

[7] M. Aurada, M. Feischl, T. Führer, M. Karkulik, and D. Praetorius. Energy norm based error estimators for adaptive BEM for hypersingular integral equations. *Appl. Numer. Math.*, 95:15–35, 2015.

[8] I. Babuška and B.Q. Guo. The *h-p* version of the finite element method for domains with curved boundaries. *SIAM J. Numer. Anal.*, 25(4):837–861, 1988.

[9] Constantin Băcuţă, Hengguang Li, and Victor Nistor. Differential operators on domains with conical points: precise uniform regularity estimates. *Rev. Roumaine Math. Pures Appl.*, 62(3):383–411, 2017.

[10] C. Bernardi, M. Dauge, and Y. Maday. Polynomials in the Sobolev world (version 2). Technical Report 14, IRMAR, 2007. https://hal.archives-ouvertes.fr/hal-00153795.

[11] M.Š. Birman and M.Z. Solomjak. *Spektralnaya teoriya samosopryazhennykh operatorov v gilbertovom prostranstve*. Leningrad. Univ., Leningrad, 1980.

[12] A. Bonito, J.P. Borthagaray, R.H. Nochetto, E. Otárola, and A.J. Salgado. Numerical methods for fractional diffusion. Technical report, 2017. arXiv:1707.01566.

[13] A. Bonito and J.E. Pasciak. Numerical approximation of fractional powers of elliptic operators. *Math. Comp.*, 84(295):2083–2110, 2015.

[14] C. Brändle, E. Colorado, A. de Pablo, and U. Sánchez. A concave-convex elliptic problem involving the fractional Laplacian. *Proc. Roy. Soc. Edinburgh Sect. A*, 143(1):39–71, 2013.

[15] 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.

[16] X. Cabré and J. Tan. Positive solutions of nonlinear problems involving the square root of the Laplacian. *Adv. Math.*, 224(5):2052–2093, 2010.

[17] L. Caffarelli and L. Silvestre. An extension problem related to the fractional Laplacian. *Comm. Part. Diff. Eqs.*, 32(7-9):1245–1260, 2007.

[18] L.A. Caffarelli and P.R. Stinga. Fractional elliptic equations, Caccioppoli estimates and regularity. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 33(3):767–807, 2016.

[19] A. Capella, J. Dávila, L. Dupaigne, and Y. Sire. Regularity of radial extremal solutions for some non-local semilinear equations. *Comm. Partial Differential Equations*, 36(8):1353–1384, 2011.

[20] H. Cartan. *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*. Avec le concours de Reiji Takahashi. Enseignement des Sciences. Hermann, Paris, 1961.

[21] X. Chen, F. Zeng, and G.E. Karniadakis. A tunable finite difference method for fractional differential equations with non-smooth solutions. *Comput. Methods Appl. Mech. Engrg.*, 318:193–214, 2017.

[22] M. Costabel and M. Dauge. General edge asymptotics of solutions of second-order elliptic boundary value problems. I, II. *Proc. Roy. Soc. Edinburgh Sect. A*, 123(1):109–155, 157–184, 1993.

[23] M. Costabel, M. Dauge, and S. Nicaise. Analytic regularity for linear elliptic systems in polygons and polyhedra. *Math. Meths. Appl. Sci.*, 22(8), 2012.

[24] M. D'Elia and M. Gunzburger. The fractional Laplacian operator on bounded domains as a special case of the nonlocal diffusion operator. *Comput. Math. Appl.*, 66(7):1245 – 1260, 2013.

[25] R.A. DeVore and G.G. Lorentz. *Constructive Approximation*. Springer Verlag, 1993.

[26] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.19 of 2018-

06-22. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[27] J. Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Translated and revised from the 1995 Spanish original by David Cruz-Uribe.

[28] 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.

[29] F.D. Gaspoz, C.-J. Heine, and K.G. Siebert. Optimal grading of the newest vertex bisection and H^1-stability of the L_2-projection. *IMA J. Numer. Anal.*, 36(3):1217–1241, 2016.

[30] F.D. Gaspoz and P. Morin. Convergence rates for adaptive finite elements. *IMA J. Numer. Anal.*, 29(4):917–936, 2009.

[31] V. Gol′dshtein and A. Ukhlov. Weighted Sobolev spaces and embedding theorems. *Trans. Amer. Math. Soc.*, 361(7):3829–3850, 2009.

[32] W. Gui and I. Babuška. The h, p and h-p versions of the finite element method in 1 dimension. II. The error analysis of the h- and h-p versions. *Numer. Math.*, 49(6):613–657, 1986.

[33] H. Harbrecht, M. Peters, and M. Siebenmorgen. Combination technique based k-th moment analysis of elliptic problems with random diffusion. *J. Comput. Phys.*, 252:128–141, 2013.

[34] B.N. Khoromskij and J.M. Melenk. Boundary concentrated finite element methods. *SIAM J. Numer. Anal.*, 41(1):1–36, 2003.

[35] A. Kufner and B. Opic. How to define reasonably weighted Sobolev spaces. *Comment. Math. Univ. Carolin.*, 25(3):537–554, 1984.

[36] N.S. Landkof. *Foundations of modern potential theory*. Springer-Verlag, New York-Heidelberg, 1972. Translated from the Russian by A. P. Doohovskoy, Die Grundlehren der mathematischen Wissenschaften, Band 180.

[37] J.-L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I*. Springer-Verlag, New York, 1972.

[38] R. E. Lynch, J. R. Rice, and D. H. Thomas. Direct solution of partial difference equations by tensor product methods. *Numer. Math.*, 6:185–199, 1964.

[39] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, Cambridge, 2000.

[40] D. Meidner, J. Pfefferer, K. Schürholz, and B. Vexler. *hp*-finite elements for fractional diffusion. Technical report, 2017. arxiv:1706.04066v1.

[41] J.M. Melenk. On the robust exponential convergence of *hp* finite element method for problems with boundary layers. *IMA J. Numer. Anal.*, 17(4):577–601, 1997.

[42] J.M. Melenk. *hp-finite element methods for singular perturbations*, volume 1796 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002.

[43] J.M. Melenk and Ch. Schwab. *hp* FEM for reaction-diffusion equations. I. Robust exponential convergence. *SIAM J. Numer. Anal.*, 35(4):1520–1557, 1998.

[44] J.M. Melenk and Ch. Schwab. Analytic regularity for a singularly perturbed problem. *SIAM J. Math. Anal.*, 30(2):379–400, 1999.

[45] K.S. Miller and S.G. Samko. Completely monotonic functions. *Integral Transform. Spec. Funct.*, 12(4):389–402, 2001.

[46] B. Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.*, 165:207–226, 1972.

[47] F. Müller, D. Schötzau, and Ch. Schwab. Symmetric interior penalty discontinuous Galerkin methods for elliptic problems in polygons. *SIAM J. Numer. Anal.*, 55(5):2490–2521, 2017.

[48] 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.

[49] R.H. Nochetto, E. Otárola, and A.J. Salgado. Piecewise polynomial interpolation in Muckenhoupt weighted Sobolev spaces and applications. *Numer. Math.*, 132(1):85–130, 2016.

[50] R.H. Nochetto and A. Veeser. Primer of adaptive finite element methods. In *Multiscale and adaptivity: modeling, numerics and applications*, volume 2040 of *Lecture Notes in Math.*, pages 125–225. Springer, Heidelberg, 2012.

[51] M.A. Olshanskii and A. Reusken. On the convergence of a multigrid method for linear reaction-diffusion problems. *Computing*, 65(3):193–202, 2000.

[52] E. Otárola. *A PDE approach to numerical fractional diffusion*. PhD thesis, University of Maryland, College Park, 2014.

[53] H.-G. Roos, M. Stynes, and L. Tobiska. *Numerical methods for singularly perturbed differential equations*, volume 24 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1996. Convection-diffusion and flow problems.

[54] S.A. Sauter and Ch. Schwab. *Boundary element methods*, volume 39 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2011. Translated and expanded

from the 2004 German original.

[55] R. Schneider. *Multiskalen- und Wavelet-Matrixkompression*. Advances in Numerical Mathematics. B. G. Teubner, Stuttgart, 1998. Analysisbasierte Methoden zur effizienten Lösung großer vollbesetzter Gleichungssysteme. [Analysis-based methods for the efficient solution of large nonsparse systems of equations].

[56] R. Schneider, O. Reichmann, and Ch. Schwab. Wavelet solution of variable order pseudodifferential equations. *Calcolo*, 47(2):65–101, 2010.

[57] D. Schötzau and Ch. Schwab. Exponential convergence for *hp*-version and spectral finite element methods for elliptic problems in polyhedra. *M3AS*, 25(9):1617–1661, 2015.

[58] D. Schötzau and Ch. Schwab. Exponential convergence of hp-fem for elliptic problems in polyhedra: Mixed boundary conditions and anisotropic polynomial degrees. *Journ. Found. Comput. Math.*, 18(3):595–660, 2018.

[59] Ch. Schwab. *p- and hp-finite element methods*. Numerical Mathematics and Scientific Computation. The Clarendon Press, Oxford University Press, New York, 1998. Theory and applications in solid and fluid mechanics.

[60] Ch. Schwab and M. Suri. The *p* and *hp* versions of the finite element method for problems with boundary layers. *Math. Comp.*, 65(216):1403–1429, 1996.

[61] Ch. Schwab, M. Suri, and C. Xenophontos. The *hp* finite element method for problems in mechanics with boundary layers. *Comput. Methods Appl. Mech. Engrg.*, 157(3-4):311–333, 1998. Seventh Conference on Numerical Methods and Computational Mechanics in Science and Engineering (NMCM 96) (Miskolc).

[62] L.R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[63] P.R. Stinga and J.L. Torrea. Extension problem and Harnack's inequality for some fractional operators. *Comm. Partial Differential Equations*, 35(11):2092–2122, 2010.

[64] L. Tartar. *An introduction to Sobolev spaces and interpolation spaces*, volume 3 of *Lecture Notes of the Unione Matematica Italiana*. Springer, Berlin, 2007.

[65] B.O. Turesson. *Nonlinear potential theory and weighted Sobolev spaces*, volume 1736 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000.