MATHEMATICS OF COMPUTATION
http://dx.doi.org/10.1090/mcom/3089
Article electronically published on March 3, 2016

MULTILEVEL METHODS FOR NONUNIFORMLY ELLIPTIC OPERATORS AND FRACTIONAL DIFFUSION

LONG CHEN, RICARDO H. NOCHETTO, ENRIQUE OTÁROLA, AND ABNER J. SALGADO

Abstract. We develop and analyze multilevel methods for nonuniformly elliptic operators whose ellipticity holds in a weighted Sobolev space with an A_2–Muckenhoupt weight. Using the so-called Xu-Zikatanov (XZ) identity, we derive a nearly uniform convergence result under the assumption that the underlying mesh is quasi-uniform. As an application we also consider the so-called α-harmonic extension to localize fractional powers of elliptic operators. Motivated by the scheme proposed by the second, third and fourth authors, we present a multilevel method with line smoothers and obtain a nearly uniform convergence result on anisotropic meshes. Numerical experiments illustrate the performance of our method.

1. Introduction

In this work we are interested in the development and analysis of efficient and fast solvers for equations that arise from finite element discretizations of elliptic boundary value problems with the most general class of coefficients that allow for a regularity theory [30]. As an application, we consider the equations that appear in the discretization of the fractional Laplacian based on the scheme proposed and analyzed in [44].

Fractional and nonlocal operators can be found in a number of applications such as biophysics [15], finance [19, 56], turbulence [22], image processing [31, 32], materials science [3], optimization [27], porous media flow [25], peridynamics [26, 46], nonlocal continuum field theories [28] and others. From this it is evident that the particular type of operator appearing in applications can vary widely and that a unified analysis of their discretizations might be well beyond our reach. A more modest, but nevertheless quite ambitious, goal is to develop an analysis and approximation of a model operator that is representative of a particular class. This is the purpose of our recent research program, in which we deal with an important

Received by the editor March 17, 2014 and, in revised form, April 25, 2015.

2010 *Mathematics Subject Classification.* Primary 65N55, 65F10, 65N22, 65N30, 35S15, 65N12.

Key words and phrases. Finite elements, weighted Sobolev spaces, Muckenhoupt weights, anisotropic estimates, multilevel methods.

The first author has been supported by NSF grants DMS-1115961, DMS-1418934, and DOE prime award # DE-SC0006903.

The second and fourth authors have been supported in part by NSF grants DMS-1109325 and DMS-1411808.

The third author was supported in part by the NSF grants DMS-1109325 and DMS-1411808 and by CONICYT through a CONICYT-FULBRIGHT Fellowship.

The fourth author was supported by NSF grant DMS-1418784.

nonlocal operator: fractional powers of the Dirichlet Laplace operator $(-\Delta)^s$, with $s \in (0,1)$, which for convenience we will simply call the fractional Laplacian.

In previous work [44] we proposed a discretization technique for this operator and provided an *a priori* error analysis for it. In this paper, we shall be interested in fast multilevel methods for the approximate solution of the discrete problems that arise from the discretization of the fractional Laplacian. In other words, we shall be concerned with efficient solution techniques for discretizations of the following problem. Let Ω be an open and bounded subset of $\mathbb{R}^n$ $(n \geq 1)$, with boundary $\partial\Omega$. Given $s \in (0,1)$ and a smooth enough function f, find u such that

$$(1.1) \qquad (-\Delta)^s u = f, \text{ in } \Omega, \quad u = 0, \text{ on } \partial\Omega.$$

The fractional Laplacian is a nonlocal operator (see [16, 17, 42]), which is one of the main difficulties in studying and solving problem (1.1). To localize it, Caffarelli and Silvestre showed in [17] that any power of the fractional Laplacian in $\mathbb{R}^n$ can be determined as a Dirichlet-to-Neumann operator via an extension problem on the upper half-space $\mathbb{R}^{n+1}_+$. For a bounded domain Ω, this result has been adapted in [18, 50], thus obtaining an extension problem which is now posed on the semi-infinite cylinder $\mathcal{C} = \Omega \times (0,\infty)$. This extension is the following mixed boundary value problem:

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

where $\partial_L\mathcal{C} = \partial\Omega \times [0,\infty)$ denotes the lateral boundary of $\mathcal{C}$, and

$$(1.3) \qquad \frac{\partial \mathscr{U}}{\partial \nu^\alpha} = -\lim_{y \to 0^+} y^\alpha \partial_y \mathscr{U}$$

is the the so-called conormal exterior derivative of $\mathscr{U}$ with ν being the unit outer normal to $\mathcal{C}$ at $\Omega \times \{0\}$. The parameter α is defined as

$$(1.4) \qquad \alpha = 1 - 2s \in (-1,1).$$

Finally, d_s is a positive normalization constant which depends only on s; see [17] for details. We will call y the *extended variable* and the dimension $n+1$ in $\mathbb{R}^{n+1}_+$ the *extended dimension* of problem (1.2).

The following simple strategy to find the solution of (1.1) has been proposed and analyzed in [44]: given a sufficiently smooth function f we solve (1.2), thus obtaining a function $\mathscr{U} = \mathscr{U}(x',y)$, and set $u : x' \in \Omega \mapsto u(x') = \mathscr{U}(x',0) \in \mathbb{R}$ to obtain the solution of (1.1).

For an overview of the existing numerical techniques used to solve problems involving fractional diffusion such as the matrix transference technique and the contour integral method, we refer to [44]. In addition to [44], the recent work of Bonito and Pasciak [6] discretizes fractional powers of elliptic operators via the integral formulation for self-adjoint operators discussed, for instance, in [5, Chapter 10.4].

The main advantage of the algorithm proposed in [44] is that we are solving the local problem (1.2) instead of dealing with the nonlocal operator $(-\Delta)^s$ of problem (1.1). However, this comes at the expense of incorporating one more dimension into the problem, thus raising the question of how computationally efficient this approach is. A quest for the answer motivates the study of multilevel methods, since it is known that they are the most efficient techniques for the solution of discretizations of partial differential equations; see [12, 13, 37, 53]. Multigrid methods for equations of the type (1.2), however, are not very well understood.

The purpose of this work is twofold and hinges on the multilevel framework developed in [8, 9, 53] and the Xu-Zikatanov identity [55]. First, we show nearly uniform convergence of a multilevel method for a class of general nonuniformly elliptic equations of the form

$$-\mathrm{div}(\mathcal{A}(x)\nabla u) = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega,$$

where the matrix $\mathcal{A}$ is elliptic in the following sense:

$$\omega(x)|\xi|^2 \lesssim \xi^\intercal \mathcal{A}(x)\xi \lesssim \omega(x)|\xi|^2,$$

for all ξ and almost every x. Here the relation $a \lesssim b$ indicates that $a \leq Cb$, with a constant C and ω belongs to the so-called Muckenhoupt class A_2; see Definition 2.1 for details. The multilevel meshes are quasi-uniform and point-wise smoothers (e.g., Gauss-Seidel smoother) are used. Second, we derive an almost uniform convergence of a multilevel method for the local problem that arises from our PDE approach to the fractional Laplacian (1.2) on anisotropic meshes [44, 45]. For the fractional Laplacian, [44] shows that a quasi-uniform mesh cannot yield quasi-optimal error estimates and, consequently, the mesh in the extended dimension must be graded towards the bottom of the cylinder, thus becoming anisotropic. We apply line smoothers over vertical lines in the extended domain and prove that the corresponding multigrid $\mathcal{V}$-cycle converges nearly uniformly. For both problems under consideration, by nearly uniformly we mean that the contraction factor of our multigrid methods depends on the number of levels and thus logarithmically on the problem size. With this level of generality, this seems unavoidable without further assumptions.

At this point we feel compelled to emphasize that although our results are not optimal, we consider the largest possible class of weights. Moreover:

- Muckenhoupt weights are the largest class of weights for which there is a well-established regularity theory for elliptic PDEs [30].
- The Muckenhoupt condition is not only sufficient but also necessary for the continuity of the Hardy-Littlewood maximal function and singular integral operators of Calderón-Zygmund type [41, 43]. This is fundamental not only in the aforementioned regularity theory but also in the study of the structural properties of the underlying weighted function spaces.
- Nonuniformly elliptic or weighted problems have been considered in the literature before. For instance, [11] proves optimal convergence for a nonuniformly elliptic equation. However, [11] hinges on a collection of rather *ad hoc* assumptions (two dimensions, particular geometry, specific properties of the weight ω), and there is no hint as to how to develop a general theory. In contrast, our results rely solely on membership of ω in the class A_2 and are valid in any dimension.
- A related work is [36], where the authors show a uniform norm equivalence for a multilevel space decomposition under the assumption that the weight belongs to the smaller class A_1. Their results and techniques, however, do not apply to our setting since, simply put, an A_1-weight is "almost bounded", which is too restrictive; see Remark 2.2 for details. We make no regularity assumption on the weight ω and show that our estimates solely depend on the A_2-constant $C_{2,\omega}$.

- In light of the previous two observations, our multilevel theory presents the most general framework where point smoothing (such as Gauss-Seidel) could work.

We propose an algorithm with complexity $\mathcal{O}(M^{n+1}\log M)$ for computing a nearly optimal approximation of the fractional Laplacian problem (1.1) in $\Omega \subset \mathbb{R}^n$, where M denotes the number of degrees of freedom in each direction. Notice that using the intrinsic integral formulation of the fractional Laplacian [16, 17], a discretization would result in a dense matrix with $\mathcal{O}(M^{2n})$. Special techniques such as fast multipole methods [35], the $\mathcal{H}$-matrix methods [39] or wavelet methods [40, 49] might be applied to reduce the complexity of storage and manipulation of the dense matrix as well as the complexity of solvers.

The outline of this paper is as follows. In Section 2, we introduce the notation and functional framework we shall work with. Section 3 contains the salient results about the finite element approximation of nonuniformly elliptic equations including the fractional Laplacian on anisotropic meshes. Here we also collect the relevant properties of a quasi-interpolant which are crucial to obtaining the convergence analysis of our multilevel methods. In Section 4, we recall the theory of subspace corrections [53] and the Xu-Zikatanov identity [55]. We present multigrid algorithms for nonuniformly elliptic equations discretized on quasi-uniform meshes in Section 5 and prove their nearly uniform convergence. We adapt the algorithms and analysis of Section 5 to the fractional Laplacian discretized on anisotropic meshes in Section 6. This requires a line smoother along the extended direction. Finally, to illustrate the performance of our methods and the sharpness of our results, we present a series of numerical experiments in Section 7.

2. Notation and preliminaries

2.1. **Notation.** Throughout this work, Ω is an open, bounded and connected subset of $\mathbb{R}^n$, with $n \geq 1$. The boundary of Ω is denoted by $\partial\Omega$. Unless specified otherwise, we will assume that $\partial\Omega$ is Lipschitz. We define the semi-infinite cylinder by $\mathcal{C} = \Omega \times (0,\infty)$, and its lateral boundary by $\partial_L \mathcal{C} = \partial\Omega \times [0,\infty)$. Given $\mathcal{Y} > 0$, we define the truncated cylinder by $\mathcal{C}_{\mathcal{Y}} = \Omega \times (0,\mathcal{Y})$ and $\partial_L \mathcal{C}_{\mathcal{Y}}$ accordingly.

Throughout our discussion, when dealing with elements defined in $\mathbb{R}^{n+1}$, we shall need to distinguish the extended dimension. A vector $x \in \mathbb{R}^{n+1}$ is denoted by

$$x = (x^1, \dots, x^n, x^{n+1}) = (x', x^{n+1}) = (x', y),$$

with $x^i \in \mathbb{R}$ for $i = 1, \dots, n+1$, $x' \in \mathbb{R}^n$ and $y \in \mathbb{R}$. The upper half-space in $\mathbb{R}^{n+1}$ will be denoted by

$$\mathbb{R}^{n+1}_+ = \left\{ x = (x', y) : x' \in \mathbb{R}^n,\ y \in \mathbb{R},\ y > 0 \right\}.$$

The relation $a \lesssim b$ indicates that $a \leq Cb$, with a constant C that does not depend on a, b and the important multilevel discretization parameters J and h_J (see Section 4 for their definitions), but it might depend on s and Ω. The value of C might change at each occurrence.

If X and Y are topological vector spaces, we write $X \hookrightarrow Y$ to denote that X is continuously embedded in Y. We denote by X' the dual of X. If X is normed, we denote by $\|\cdot\|_X$ its norm.

2.2. **Weighted Sobolev spaces.** In the Caffarelli-Silvestre extension (1.2), the parameter $\alpha = 1 - 2s \in (-1, 1)$. Consequently, the weight y^α is degenerate ($\alpha > 0$) or singular ($\alpha < 0$), thereby making problem (1.2) nonuniformly elliptic. The natural space for problem (1.2) is no longer the standard space H^1 but rather the weighted Sobolev space $H^1(y^\alpha, \mathcal{C})$, where the weight $|y|^\alpha$ belongs to the so-called Muckenhoupt class $A_2(\mathbb{R}^{n+1})$; see [30, 43, 51]. For completeness, we recall the definition of Muckenhoupt classes.

Definition 2.1 (Muckenhoupt class A_p)**.** Let $n \geq 1$ and $\omega \in L^1_{loc}(\mathbb{R}^n)$ be such that $\omega(x) > 0$ for a.e. $x \in \mathbb{R}^n$. We say that $\omega \in A_p(\mathbb{R}^n)$, $1 < p < \infty$, if there exists a positive constant $C_{p,\omega}$ such that

$$\sup_B \left(\frac{1}{|B|} \int_B \omega \,\mathrm{d}x \right) \left(\frac{1}{|B|} \int_B \omega^{1/(1-p)} \,\mathrm{d}x \right)^{p-1} = C_{p,\omega} < \infty, \tag{2.1}$$

where the supremum is taken over all balls B in $\mathbb{R}^n$ and $|B|$ denotes the Lebesgue measure of B. In addition, we define

$$A_\infty(\mathbb{R}^n) = \bigcup_{p>1} A_p(\mathbb{R}^n) \quad \text{and} \quad A_1(\mathbb{R}^n) = \bigcap_{p>1} A_p(\mathbb{R}^n).$$

If ω belongs to the Muckenhoupt class $A_p(\mathbb{R}^n)$, we say that ω is an A_p-weight, and we call the constant $C_{p,\omega}$ in (2.1) the A_p-constant of ω.

Remark 2.2 (Characterization of the A_1-class). A useful characterization of the A_1-Muckenhoupt class is given in [47]: $\omega \in A_1(\mathbb{R}^n)$ if and only if

$$\sup_B \frac{\|\omega^{-1}\|_{L^\infty(B)}}{|B|} \int_B \omega \,\mathrm{d}x = C_{1,\omega} < \infty. \tag{2.2}$$

Since $\alpha \in (-1, 1)$, it is immediate that $|y|^\alpha \in A_2(\mathbb{R}^{n+1})$ but $|y|^\alpha \notin A_1(\mathbb{R}^{n+1})$.

From the A_p-condition and Hölder's inequality it follows that an A_p-weight satisfies the so-called *strong doubling property*, which is essential in the analysis of the multigrid methods introduced in §4. The proof of this property is standard; see [51, Proposition 1.2.7] or [45, Proposition 2.2] for more details.

Proposition 2.1 (Strong doubling property)**.** *Let $\omega \in A_p(\mathbb{R}^n)$ with $1 < p < \infty$ and let $E \subset \mathbb{R}^n$ be a measurable subset of a ball $B \subset \mathbb{R}^N$. Then*

$$\omega(B) \leq C_{p,\omega} \left(\frac{|B|}{|E|} \right)^p \omega(E). \tag{2.3}$$

For an A_p-weight we define weighted L^p spaces as follows.

Definition 2.3 (Weighted Lebesgue spaces)**.** Let $\omega \in A_p(\mathbb{R}^n)$, and let $D \subset \mathbb{R}^n$ be an open and bounded domain. For $1 < p < \infty$, we define the weighted Lebesgue space $L^p(\omega, D)$ as the set of measurable functions u on D for which

$$\|u\|_{L^p(\omega,D)} = \left(\int_D |u|^p \omega \,\mathrm{d}x \right)^{1/p} < \infty.$$

Based on the fact that $L^p(\omega, D) \hookrightarrow L^1_{loc}(D)$ (cf. [45, Proposition 2.3]), it makes sense to talk about weak derivatives of functions in $L^p(\omega, D)$. We define weighted Sobolev spaces as follows.

Definition 2.4 (Weighted Sobolev spaces)**.** Let $D \subset \mathbb{R}^n$ be an open and bounded domain, $\omega \in A_p(\mathbb{R}^n)$ with $1 < p < \infty$ and $m \in \mathbb{N}$. The weighted Sobolev space $W^m_p(\omega, D)$ is the space of functions $u \in L^p(\omega, D)$ such that for any multiindex κ

with $|\kappa| \leq m$, the weak derivatives $D^\kappa u \in L^p(\omega, D)$. We endow $W_p^m(\omega, D)$ with the following seminorm and norm:

$$|u|_{W_p^m(\omega,D)} = \left(\sum_{|\kappa|=m} \|D^\kappa u\|^p_{L^p(\omega,D)} \right)^{1/p}, \quad \|u\|_{W_p^m(\omega,D)} = \left(\sum_{j\leq m} |u|^p_{W_p^j(\omega,D)} \right)^{1/p},$$

respectively. We also define $\mathring{W}_p^m(\omega, D)$ as the closure of $\mathcal{C}_0^\infty(D)$ in $W_p^m(\omega, D)$.

Owing to the fact that $\omega \in A_p$, most of the properties of classical Sobolev spaces have a weighted counterpart; see [30, 33, 51]. In particular we have the following result (cf. [51, Proposition 2.1.2, Corollary 2.1.6] and [33, Theorem 1]).

Proposition 2.2 (Properties of weighted Sobolev spaces)**.** *Let $D \subset \mathbb{R}^n$ be an open and bounded domain, $1 < p < \infty$, $\omega \in A_p(\mathbb{R}^n)$ and $m \in \mathbb{N}$. The spaces*

$$W_p^m(\omega, D) \quad \text{and} \quad \mathring{W}_p^m(\omega, D)$$

are complete, and $W_p^m(\omega, D) \cap \mathcal{C}^\infty(D)$ is dense in $W_p^m(\omega, D)$.

2.3. **The Caffarelli-Silvestre extension problem.** Here we explore problem (1.2); we refer the reader to [16, 17, 44] for details. Since problem (1.2) is posed on the unbounded domain $\mathcal{C}$, it cannot be directly approximated with finite element-like techniques. However, as [44, Proposition 3.1] shows, the solution $\mathscr{U}$ decays exponentially in y so that, by truncating the cylinder $\mathcal{C}$ to $\mathcal{C}_\mathcal{Y}$ and setting a vanishing Dirichlet boundary condition on the upper boundary $y = \mathcal{Y}$, we incur only an exponentially small error in terms of $\mathcal{Y}$ [44, Theorem 3.5].

For $\alpha = 1 - 2s$ we have $|y|^\alpha \in A_2(\mathbb{R}^n)$ and we define

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

Proposition 2.2 states that $\mathring{H}_L^1(y^\alpha, \mathcal{C}_\mathcal{Y})$ is a Hilbert space. We also define

$$\mathbb{H}^s(\Omega) = [H_0^1(\Omega), L^2(\Omega)]_{1-s}, \tag{2.4}$$

for $0 < s < 1$, which is the natural space for the solution u of problem (1.1); let $\mathbb{H}^{-s}(\Omega)$ be the dual of $\mathbb{H}^s(\Omega)$. As [44, Proposition 2.5] shows, the trace operator

$$\mathring{H}_L^1(y^\alpha, \mathcal{C}_\mathcal{Y}) \ni w \mapsto \operatorname{tr}_\Omega w \in \mathbb{H}^s(\Omega)$$

is well defined. We approximate problem (1.2) by: find $v \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_\mathcal{Y})$ such that

$$\int_{\mathcal{C}_\mathcal{Y}} y^\alpha \nabla v \cdot \nabla \phi = d_s \langle f, \operatorname{tr}_\Omega \phi \rangle_{\mathbb{H}^{-s}(\Omega) \times \mathbb{H}^s(\Omega)}, \quad \forall \phi \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_\mathcal{Y}), \tag{2.5}$$

where $\langle \cdot, \cdot \rangle_{\mathbb{H}^{-s}(\Omega) \times \mathbb{H}^s(\Omega)}$ denotes the duality pairing between $\mathbb{H}^s(\Omega)$ and $\mathbb{H}^{-s}(\Omega)$. Finally, we recall the exponential convergence result [44, Theorem 3.5]:

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

where λ_1 denotes the first eigenvalue of the Dirichlet Laplace operator and $\mathcal{Y}$ is the truncation parameter.

3. Finite element discretization of nonuniformly elliptic equations

Let Ω be an open and bounded subset of $\mathbb{R}^n$ $(n \geq 1)$ with boundary $\partial\Omega$ and let $f \in L^2(\omega^{-1}, \Omega)$. We now focus on a finite element method for the following nonuniformly elliptic boundary value problem: find $u \in H_0^1(\omega, \Omega)$ that solves

$$-\mathrm{div}(\mathcal{A}(x)\nabla u) = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega, \tag{3.1}$$

where $\mathcal{A} : \Omega \to \mathbb{R}^{n\times n}$, $\mathcal{A} = \mathcal{A}^\intercal$ and satisfies the nonuniform ellipticity condition:

$$\omega(x)|\xi|^2 \lesssim \xi^\intercal \mathcal{A}(x)\xi \lesssim \omega(x)|\xi|^2, \qquad \forall \xi \in \mathbb{R}^n, \quad \text{a.e. } x \in \Omega. \tag{3.2}$$

The function ω belongs to the Muckenhoupt class A_2, which is defined by (2.1). Examples of nonuniformly elliptic equations are the harmonic extension problem related with the fractional Laplace operator [16, 17], elliptic PDEs in an axisymmetric three dimensional domain with axisymmetric data [4, 34], and equations modeling the motion of particles in a central potential field in quantum mechanics [2].

Nonuniformly elliptic equations of the type (3.1)-(3.2) have been studied in [30]. Given $f \in L^2(\omega^{-1}, \Omega)$, there exists a unique solution $u \in H_0^1(\omega, \Omega)$ [30, Theorem 2.2]. Notice that by taking the weight ω to be y^α we see that the underlying differential operator in (2.5) is a particular instance of $-\mathrm{div}(\mathcal{A}(x)\nabla u)$ in (3.1).

We define the bilinear form

$$a(u, v) = \int_\Omega \mathcal{A}\nabla u \cdot \nabla v \,\mathrm{d}x, \tag{3.3}$$

which is clearly continuous and coercive in $H_0^1(\omega, \Omega)$. Then, a weak formulation of problem (3.1) reads: find $u \in H_0^1(\omega, \Omega)$ such that

$$a(u, v) = \int_\Omega f v \,\mathrm{d}x, \quad \forall v \in H_0^1(\omega, \Omega). \tag{3.4}$$

3.1. Finite element approximation on quasi-uniform meshes.

To avoid technical difficulties, we assume Ω to be a polyhedral domain. Let $\mathscr{T} = \{T\}$ be a mesh of Ω into elements T (simplices or cubes) such that

$$\bar{\Omega} = \bigcup_{T\in\mathscr{T}} T, \qquad |\Omega| = \sum_{T\in\mathscr{T}} |T|.$$

The partition $\mathscr{T}$ is assumed to be conforming or compatible. We denote by $\mathbb{T}$ a collection of conforming meshes. We say that $\mathbb{T}$ is *shape regular* if there exists a constant $\sigma > 1$ such that, for all $\mathscr{T} \in \mathbb{T}$, we have

$$\max\{\sigma_T : T \in \mathscr{T}\} \leq \sigma, \tag{3.5}$$

where $\sigma_T := h_T/\rho_T$ is the shape coefficient of T. For simplicial elements, $h_T = \mathrm{diam}(T)$ and ρ_T is the diameter of the largest sphere inscribed in T [14, 24]. For the definition of h_T and ρ_T in the case of n-rectangles, we refer to [24].

We assume that the collection of meshes $\mathbb{T}$ is conforming and satisfies the regularity assumption (3.5), which says that the element shape does not degenerate with refinement. A refinement method generating meshes satisfying the shape regular condition (3.5) will be called *isotropic refinement*. A particular instance of an isotropic refinement is the so-called quasi-uniform refinement. We recall that $\mathbb{T}$ is quasi-uniform if it is shape regular and for all $\mathscr{T} \in \mathbb{T}$ we have

$$\max\{h_T : T \in \mathscr{T}\} \lesssim \min\{h_T : T \in \mathscr{T}\},$$

where the hidden constant is independent of $\mathscr{T}$. In this case, all the elements on the same refinement level are of comparable size. We define $h_{\mathscr{T}} = \max_{T\in\mathscr{T}} h_T$.

Given a mesh $\mathscr{T} \in \mathbb{T}$, we define the finite element space of continuous piecewise polynomials of degree one:

$$\mathbb{V}(\mathscr{T}) = \left\{ W \in \mathcal{C}^0(\bar{\Omega}) : \; W|_T \in \mathcal{P}(T) \; \forall T \in \mathscr{T}, \; W|_{\partial\Omega} = 0 \right\}, \tag{3.6}$$

where for a simplicial element T, $\mathcal{P}(T)$ corresponds to the space of polynomials of total degree at most one, i.e., $\mathbb{P}_1(T)$, and for n-rectangles, $\mathcal{P}(T)$ stands for the space of polynomials of degree at most one in each variable, i.e., $\mathbb{Q}_1(T)$.

The finite element approximation of u, solution of problem (3.1), is defined as the unique discrete function $U_{\mathscr{T}} \in \mathbb{V}(\mathscr{T})$ satisfying

$$a(U_{\mathscr{T}}, W) = \int_\Omega fW \,\mathrm{d}x, \quad \forall W \in \mathbb{V}(\mathscr{T}). \tag{3.7}$$

3.2. **Quasi-interpolation operator.** Let us recall the main properties of the quasi-interpolation operator $\Pi_{\mathscr{T}}$ introduced and analyzed in [45]. This operator is based on local averages over stars, and then it is well defined for functions in $L^p(\omega, \Omega)$. We summarize its construction and its approximation properties as follows; see [45] for details.

Given a mesh $\mathscr{T} \in \mathbb{T}$ and $T \in \mathscr{T}$, we denote by $\mathcal{N}(T)$ the set of nodes of T. We set $\mathcal{N}(\mathscr{T}) := \bigcup_{T\in\mathscr{T}} \mathcal{N}(T)$ and $\mathring{\mathcal{N}}(\mathscr{T}) := \mathcal{N}(\mathscr{T}) \cap \Omega$. Then, any discrete function $W \in \mathbb{V}(\mathscr{T})$ is characterized by its nodal values on the set $\mathring{\mathcal{N}}(\mathscr{T})$. Moreover, the functions $\phi_{\mathtt{v}} \in \mathbb{V}(\mathscr{T})$, $\mathtt{v} \in \mathring{\mathcal{N}}(\mathscr{T})$, such that $\phi_{\mathtt{v}}(\mathtt{w}) = \delta_{\mathtt{vw}}$ for all $\mathtt{w} \in \mathcal{N}(\mathscr{T})$ are the canonical basis of $\mathbb{V}(\mathscr{T})$, and $W = \sum_{\mathtt{v}\in\mathring{\mathcal{N}}(\mathscr{T})} W(\mathtt{v})\phi_{\mathtt{v}}$.

Given a vertex $\mathtt{v} \in \mathcal{N}(\mathscr{T})$, we define the star or patch around $\mathtt{v}$ as $S_{\mathtt{v}} = \bigcup_{T\ni\mathtt{v}} T$, and for $T \in \mathscr{T}$ we define its patch as $S_T = \bigcup_{\mathtt{v}\in T} S_{\mathtt{v}}$. For each vertex $\mathtt{v} \in \mathcal{N}(\mathscr{T})$, we define $h_{\mathtt{v}} = \min\{h_T : \mathtt{v} \in T\}$.

Let $\psi \in \mathcal{C}^\infty(\mathbb{R}^n)$ be such that $\int \psi = 1$ and $\operatorname{supp} \psi \subset B$, where B denotes the ball in $\mathbb{R}^n$ of radius r centered at zero with $r \leq 1/\sigma$, with σ defined by (3.5). For $\mathtt{v} \in \mathring{\mathcal{N}}(\mathscr{T})$, we define the rescaled smooth function

$$\psi_{\mathtt{v}}(x) = \frac{1}{h_{\mathtt{v}}^n} \psi\left(\frac{\mathtt{v} - x}{h_{\mathtt{v}}}\right).$$

Given a smooth function v, we denote by $P^1 v(x, z)$ the Taylor polynomial of degree one of the function v in the variable z about the point x, i.e.,

$$P^1 v(x, z) = v(x) + \nabla v(x) \cdot (z - x).$$

Then, given $\mathtt{v} \in \mathring{\mathcal{N}}(\mathscr{T})$ and a function $v \in W^1_p(\omega, \Omega)$, we define the corresponding averaged Taylor polynomial of first degree of v about the vertex $\mathtt{v}$ as

$$Q^1_{\mathtt{v}} v(z) = \int P^1 v(x, z) \psi_{\mathtt{v}}(x) \,\mathrm{d}x. \tag{3.8}$$

Since $\operatorname{supp} \psi_{\mathtt{v}} \subset S_{\mathtt{v}}$, the integral appearing in (3.8) can be written over $S_{\mathtt{v}}$. Moreover, integration by parts shows that $Q^1_{\mathtt{v}}$ is well defined for functions in $L^1(\Omega)$; see [5, Proposition 4.1.12]. Consequently, [44, Proposition 2.3] implies that $Q^1_{\mathtt{v}}$ is also well defined for functions in $L^p(\omega, \Omega)$ with $\omega \in A_p(\mathbb{R}^n)$.

Given $\omega \in A_p(\mathbb{R}^n)$ and $v \in L^p(\omega, \Omega)$, we define the quasi-interpolant $\Pi_{\mathscr{T}} v$ as the unique function $\Pi_{\mathscr{T}} v \in \mathbb{V}(\mathscr{T})$ that satisfies $\Pi_{\mathscr{T}} v(\mathtt{v}) = Q^1_{\mathtt{v}}(\mathtt{v})$ if $\mathtt{v} \in \mathring{\mathcal{N}}(\mathscr{T})$, and $\Pi_{\mathscr{T}} v(\mathtt{v}) = 0$ if $\mathtt{v} \in \mathcal{N}(\mathscr{T}) \cap \partial\Omega$, i.e., $\Pi_{\mathscr{T}} v(\mathtt{v}) = \sum_{\mathtt{v}\in\mathring{\mathcal{N}}(\mathscr{T})} Q^1_{\mathtt{v}}(\mathtt{v})\phi_{\mathtt{v}}$. For this operator, [45, Section 5] proves stability and interpolation error estimates in the weighted L^p-norm and W^1_p-seminorm. We recall these results for completeness.

Proposition 3.1 (Weighted stability and local error estimate I)**.** *Let $T \in \mathscr{T}$, $\omega \in A_p(\mathbb{R}^n)$ and $v \in L^p(\omega, S_T)$. Then, we have the following local stability bound:*

$$\|\Pi_{\mathscr{T}} v\|_{L^p(\omega,T)} \lesssim \|v\|_{L^p(\omega,S_T)}. \tag{3.9}$$

If, in addition, $v \in W^1_p(\omega, S_T)$, then we have the local interpolation error estimate

$$\|v - \Pi_{\mathscr{T}} v\|_{L^p(\omega,T)} \lesssim h_{\mathtt{v}} \|\nabla v\|_{L^p(\omega,S_T)}. \tag{3.10}$$

The hidden constants in both inequalities depend only on $C_{p,\omega}$, ψ and σ.

Proposition 3.2 (Weighted stability and local error estimate II)**.** *Let $T \in \mathscr{T}$, $\omega \in A_p(\mathbb{R}^n)$ and $v \in W^1_p(\omega, S_T)$. Then, we have the following local stability bound:*

$$\|\nabla \Pi_{\mathscr{T}} v\|_{L^p(\omega,T)} \lesssim \|\nabla v\|_{L^p(\omega,S_T)}. \tag{3.11}$$

If, in addition, $v \in W^2_p(\omega, S_T)$, then

$$\|\nabla(v - \Pi_{\mathscr{T}} v)\|_{L^p(\omega,T)} \lesssim h_{\mathtt{v}} \|D^2 v\|_{L^p(\omega,S_T)}. \tag{3.12}$$

The hidden constants in both inequalities depend only on $C_{p,\omega}$, ψ and σ.

3.3. **Finite element approximation on anisotropic meshes.** Let us now focus our attention on the finite element discretization of problem (2.5). To do so, we must first study the regularity of its solution $\mathscr{U}$. An error estimate for v, solution of (2.5), depends on the regularity of $\mathscr{U}$ as well [44, §4.1]. The second order regularity of $\mathscr{U}$ is much worse in the extended direction, as the following estimates from [44, Theorem 2.7] reveal: for $\beta > 2\alpha + 1$,

$$\|\Delta_{x'} \mathscr{U}\|_{L^2(y^\alpha,\mathcal{C})} + \|\partial_y \nabla_{x'} \mathscr{U}\|_{L^2(y^\alpha,\mathcal{C})} \lesssim \|f\|_{\mathbb{H}^{1-s}(\Omega)}, \tag{3.13}$$

$$\|\mathscr{U}_{yy}\|_{L^2(y^\beta,\mathcal{C})} \lesssim \|f\|_{L^2(\Omega)}. \tag{3.14}$$

Therefore, *graded* meshes in the extended variable y play a fundamental role.

Estimates (3.13)–(3.14) motivate the construction of a mesh over $\mathcal{C}_\mathcal{Y}$ with cells of the form $T = K \times I$, where $K \subset \mathbb{R}^n$ is an element that is isoparametrically equivalent either to the unit cube $[0,1]^n$ or the unit simplex in $\mathbb{R}^n$ and $I \subset \mathbb{R}$ is an interval. Let $\mathscr{T}_\Omega = \{T\}$ be a conforming and shape regular mesh of Ω. To obtain a global regularity assumption for $\mathscr{T}_\mathcal{Y}$ we assume that there is a constant $\sigma_\mathcal{Y}$ such that if $T_1 = K_1 \times I_1$ and $T_2 = K_2 \times I_2 \in \mathscr{T}_\mathcal{Y}$ have nonempty intersection, then

$$h_{I_1} h_{I_2}^{-1} \le \sigma_\mathcal{Y}, \tag{3.15}$$

where $h_I = |I|$. Exploiting the Cartesian structure of the mesh it is possible to handle anisotropy in the extended variable and obtain estimates of the form

$$\|v - \Pi_{\mathscr{T}_\mathcal{Y}} v\|_{L^2(y^\alpha,T)} \lesssim h_{\mathtt{v}'} \|\nabla_{x'} v\|_{L^2(y^\alpha,S_T)} + h_{\mathtt{v}''} \|\partial_y v\|_{L^2(y^\alpha,S_T)},$$
$$\|\partial_{x_j}(v - \Pi_{\mathscr{T}_\mathcal{Y}} v)\|_{L^2(y^\alpha,T)} \lesssim h_{\mathtt{v}'} \|\nabla_{x'} \partial_{x_j} v\|_{L^2(y^\alpha,S_T)} + h_{\mathtt{v}''} \|\partial_y \partial_{x_j} v\|_{L^2(y^\alpha,S_T)},$$

with $j = 1, \dots, n+1$, where $h_{\mathtt{v}'} = \min\{h_K : \mathtt{v}' \in K\}$, $h_{\mathtt{v}''} = \min\{h_I : \mathtt{v}'' \in I\}$ and v is the solution of problem (2.5); see [44, §4.2.3 and §4.2.4] for details. However, since $\mathscr{U}_{yy} \approx y^{-\alpha-1}$ as $y \approx 0$, we realize that $\mathscr{U} \notin H^2(y^\alpha, \mathcal{C})$ and the second estimate is not meaningful for $j = n+1$. In view of the regularity estimate (3.14) it is necessary to measure the regularity of $\mathscr{U}_{yy}$ with a different weight and thus compensate with a graded mesh in the extended dimension. This makes anisotropic estimates essential.

To simplify the analysis and implementation of multilevel techniques, we consider a sequence of nested discretizations constructed as follows: We introduce a sequence of nested uniform partitions of the unit interval $\{\mathcal{T}_k\}$, with mesh points $\widehat{y}_{l,k}$, for $l = 0, \dots, M_k$ and $k = 0, \dots, J$. We obtain a family of meshes of the interval $[0, \mathscr{Y}]$ given by the mesh points

$$y_{l,k} = \mathscr{Y} \widehat{y}_{l,k}^{\gamma}, \quad l = 0, \dots, M_k, \tag{3.16}$$

where $\gamma > 3/(1-\alpha)$. Then, for $k = 0, \dots, J$, we consider a quasi-uniform triangulation $\mathscr{T}_{\Omega,k}$ of the domain Ω and construct the mesh $\mathscr{T}_{\mathscr{Y},k}$ as the tensor product of $\mathscr{T}_{\Omega,k}$ and the partition given in (3.16); hence $\#\mathscr{T}_{\mathscr{Y},k} = M_k \, \#\mathscr{T}_{\Omega,k}$. Assuming that $\#\mathscr{T}_{\Omega,k} \approx M_k^n$ we have $\#\mathscr{T}_{\mathscr{Y},k} \approx M_k^{n+1}$. Finally, since $\mathscr{T}_{\Omega,k}$ is shape regular and quasi-uniform, $h_{\mathscr{T}_{\Omega,k}} \approx (\#\mathscr{T}_{\Omega,k})^{-1/n}$. All these considerations allow us to obtain the following result [44, Theorem 5.4 and Remark 5.5].

Theorem 3.1 (Error estimate)**.** *Denote by $V_{\mathscr{T}_{\mathscr{Y},k}} \in \mathbb{V}(\mathscr{T}_{\mathscr{Y},k})$ the Galerkin approximation of problem* (2.5)*. Then,*

$$\|\nabla(\mathscr{U} - V_{\mathscr{T}_{\mathscr{Y},k}})\|_{L^2(y^\alpha, \mathcal{C})} \lesssim |\log(\#\mathscr{T}_{\mathscr{Y},k})|^s (\#\mathscr{T}_{\mathscr{Y},k})^{-1/(n+1)} \|f\|_{\mathbb{H}^{1-s}(\Omega)},$$

where $\mathscr{Y} \approx \log(\#\mathscr{T}_{\mathscr{Y},k})$.

We notice that the anisotropic meshes of the cylinder $\mathcal{C}_{\mathscr{Y}}$ considered above are semi-structured by construction. They are generated as the tensor product of an unstructured grid $\mathscr{T}_\Omega$ together with the structured mesh $\mathcal{T}_k$.

Notice that the approximation estimates (3.9)–(3.12) are local and thus valid under the weak shape regularity condition (3.15). Owing to the tensor product structure of the mesh, we have the following anisotropic error estimate.

Lemma 3.2 (Weighted L^2 anisotropic error estimate)**.** *Let $v \in \mathring{H}_L^1(y^\alpha, \mathcal{C}_{\mathscr{Y}})$ be the solution of problem* (2.5)*. Then, the quasi-interpolation operator $\Pi_{\mathscr{T}_{\mathscr{Y}}}$ satisfies the following error estimate:*

$$\|v - \Pi_{\mathscr{T}_{\mathscr{Y}}} v\|_{L^2(y^\alpha, \mathcal{C}_{\mathscr{Y}})} \lesssim \#\mathscr{T}_{\mathscr{Y}}^{-1/(n+1)} \left(\|\nabla_{x'} v\|_{L^2(y^\alpha, \mathcal{C}_{\mathscr{Y}})} + \|\partial_y v\|_{L^2(y^\alpha, \mathcal{C}_{\mathscr{Y}})} \right).$$

Proof. This is a direct consequence of the results from [45, §5] together with the Cartesian structure of the mesh $\mathscr{T}_{\mathscr{Y}}$. □

A simple application of the mean value theorem yields

$$y_{l+1,k} - y_{l,k} = \frac{\mathscr{Y}}{M_k^\gamma}\left((l+1)^\gamma - l^\gamma\right) \le \gamma \frac{\mathscr{Y}}{M_k} \left(\frac{l+1}{M_k} \right)^{\gamma-1} \le \gamma \frac{\mathscr{Y}}{M_k}, \tag{3.17}$$

for every $l = 0, \dots, M_k - 1$, where $\gamma > 3/(1-\alpha) = 3/(2s)$ according to (3.16). In other words, since the mesh size of the quasi-uniform mesh $\mathscr{T}_{\Omega,k}$ is $\mathcal{O}(M_k^{-1})$, the size of the partitions in the extended variable y can be uniformly controlled by $h_{\mathscr{T}_{\Omega,k}}$ for $k = 0, \dots, J$. However, γ blows up as $s \downarrow 0$.

4. Multilevel space decomposition and multigrid methods

In this section, we present a $\mathcal{V}$-cycle multigrid algorithm based on the method of subspace corrections [9, 53], and we present the key identity of Xu and Zikatanov [55] in order to analyze the convergence of the proposed multigrid algorithm.

4.1. **Multilevel decomposition.** We follow [7, 8] to present a multilevel decomposition of the space $\mathbb{V}(\mathscr{T})$. Assume that we have an initial conforming mesh $\mathscr{T}_0$ made of simplices or cubes and a nested sequence of discretizations $\{\mathscr{T}_k\}_{k=0}^J$ where, for $k>0$, $\mathscr{T}_k$ is obtained by uniform refinement of $\mathscr{T}_{k-1}$. We then obtain a nested sequence, in the sense of trees, of quasi-uniform meshes $\mathscr{T}_0 \le \mathscr{T}_1 \le \cdots \le \mathscr{T}_J = \mathscr{T}$. Denoting by $h_k := h_{\mathscr{T}_k}$ the mesh size of the mesh $\mathscr{T}_k$, we have that $h_k \approx \rho^k$ for some $\rho \in (0,1)$, and then $J \approx |\log h_J|$. Let $\mathbb{V}_k := \mathbb{V}(\mathscr{T}_k)$ denote the corresponding finite element space over $\mathscr{T}_k$ defined by (3.6). We thus get a sequence of nested spaces $\mathbb{V}_0 \subset \mathbb{V}_1 \subset \cdots \subset \mathbb{V}_J = \mathbb{V}$ and a macro space decomposition $\mathbb{V} = \sum_{k=0}^J \mathbb{V}_k$. Note the redundant overlapping of the multilevel decomposition above; in particular, the sum is not direct. We now introduce a space micro-decomposition. We start by defining $\mathcal{N}_k := \mathcal{N}(\mathscr{T}_k) = \dim \mathbb{V}_k$, i.e., the number of interior vertices of the mesh $\mathscr{T}_k$. In order to deal with point and line Gauss-Seidel smoothers, we introduce the following sets of indices: For $j = 1,\dots,\mathcal{M}_k$ we denote by $\mathcal{I}_{k,j}$ a subset of the index set $\{1,2,\dots,\mathcal{N}_k\}$ and assume $\mathcal{I}_{k,j}$ satisfies

$$\bigcup_{j=1}^{\mathcal{M}_k} \mathcal{I}_{k,j} = \{1,2,\dots,\mathcal{N}_k\}.$$

The sets $\mathcal{I}_{k,j}$ may overlap: given $0 < j_1, j_2 \le \mathcal{M}_k$ such that $j_1 \neq j_2$, we may have $\mathcal{I}_{k,j_1} \cap \mathcal{I}_{k,j_2} \neq \emptyset$. This overlap, however, is finite and independent of J and $\mathcal{N}_k$.

Upon denoting the standard nodal basis of $\mathbb{V}_k$ by $\phi_{k,i}$, $i = 1,\dots,\mathcal{N}_k$, we define $\mathbb{V}_{k,j} = \operatorname{span}\{\phi_{k,i} : i \in \mathcal{I}_{k,j}\}$ and we have the space decomposition

$$\mathbb{V} = \sum_{k=0}^{J} \sum_{j=1}^{\mathcal{M}_k} \mathbb{V}_{k,j}. \tag{4.1}$$

4.2. **Multigrid algorithm.** We now describe the multigrid algorithm for the nonuniformly elliptic problem (3.1). We start by introducing several auxiliary operators. For $k = 0,\dots,J$, we define the operator $A_k : \mathbb{V}_k \to \mathbb{V}_k$ by

$$(A_k v_k, w_k)_{L^2(\omega,\Omega)} = a(v_k, w_k), \quad \forall v_k, w_k \in \mathbb{V}_k,$$

where the bilinear form a is defined in (3.3). Notice that this operator is symmetric and positive definite with respect to the weighted L^2-inner product. The projection operator $P_k : \mathbb{V}_J \to \mathbb{V}_k$ in the a-inner product is defined by

$$a(P_k v, w_k) = a(v, w_k), \quad \forall w_k \in \mathbb{V}_k, \tag{4.2}$$

and the weighted L^2-projection $Q_k : \mathbb{V}_J \to \mathbb{V}_k$ is defined by

$$(Q_k v, w_k)_{L^2(\omega,\Omega)} = (v, w_k)_{L^2(\omega,\Omega)}, \quad w_k \in \mathbb{V}_k.$$

We define, analogously, the operators $A_{k,j} : \mathbb{V}_{k,j} \to \mathbb{V}_{k,j}$, $P_{k,j} : \mathbb{V}_k \to \mathbb{V}_{k,j}$ and $Q_{k,j} : \mathbb{V}_k \to \mathbb{V}_{k,j}$. The operator $A_{k,j}$ can be regarded as the restriction of A_k to the subspace $\mathbb{V}_{k,j}$, and its matrix representation, which is the sub-matrix of A_k obtained by deleting the indices $i \notin \mathcal{I}_{k,j}$, is symmetric and positive definite. On the other hand, the operators $P_{k,j}$ and $Q_{k,j}$ denote the projections with respect to the a- and the weighted L^2-inner products into $\mathbb{V}_{k,j}$, respectively. We also remark that the matrix representation of the operator $Q_{k,j}$ is the so-called restriction operator, and the prolongation operator $Q_{k,j}^T$ corresponds to the natural embedding $\mathbb{V}_{k,j} \hookrightarrow \mathbb{V}_k$. The following property, which is of fundamental importance, will be used frequently:

$$A_{k,j} P_{k,j} = Q_{k,j} A_k. \tag{4.3}$$

With this notation we define a symmetric $\mathcal{V}$-cycle multigrid method as in [10, Algorithm 3.1], with $m \geq 1$ pre- and post-smoothing steps of the form

$$v \leftarrow v + A_{k,j}^{-1} Q_{k,j}(r_k - A_k v),$$

where r_k is the residual in $\mathbb{V}_k$ and $1 \leq j \leq \mathcal{M}_k$. When $m = 1$, it is equivalent to the application of successive subspace corrections (SSC) to the decomposition (4.1) with exact sub-solvers $A_{k,j}^{-1}$ so that the $\mathcal{V}$-cycle multigrid method has a smoother at each level of block Gauss-Seidel type [8, 53]. In particular, if we consider a nodal decomposition $\mathcal{I}_{k,j} = \{j\}$ we obtain a point-wise Gauss-Seidel smoother. On the other hand, if the indices in $\mathcal{I}_{k,j}$ are such that the corresponding vertices lie on a straight line, we obtain the so-called line Gauss-Seidel smoother, which will be essential to efficiently solve problem (1.2) with anisotropic elements.

4.3. **Analysis of the multigrid method.** In order to prove the nearly uniform convergence of the symmetric $\mathcal{V}$-cycle multigrid method without any assumptions, we rely on the following fundamental identity developed by Xu and Zikatanov [55]; see also [21, 23] for alternative proofs.

Theorem 4.1 (XZ identity)**.** *Let $\mathbb{V}$ be a Hilbert space with inner product $(\cdot,\cdot)_A$ and norm $\|\cdot\|_A$. For $j = 0, \ldots, J$ let $\mathbb{V}_j \subset \mathbb{V}$ be closed subspaces of $\mathbb{V}$ that satisfy $\mathbb{V} = \sum_{j=0}^{J} \mathbb{V}_j$. Denote by $P_j : \mathbb{V} \to \mathbb{V}_j$ the orthogonal projection in the a-inner product onto $\mathbb{V}_j$ defined in* (4.2)*. Then, the following identity holds:*

$$\left\| \prod_{j=0}^{J} (I - P_j) \right\|_A^2 = 1 - \frac{1}{1 + c_0},$$

where

$$c_0 = \sup_{\|\nu\|_A = 1} \inf_{\sum_{i=0}^{J} \nu_i = \nu} \sum_{i=0}^{J} \left\| P_i \sum_{j=i+1}^{J} \nu_j \right\|_A^2 . \tag{4.4}$$

The XZ identity given by Theorem 4.1, the properties of the interpolation operator $\Pi_{\mathscr{T}}$ defined in §3.2, the stability of the nodal decomposition stated in Lemma 5.1 below, and the weighted inverse inequality proved in Lemma 5.2 below will allow us to obtain nearly uniform convergence of the symmetric $\mathcal{V}$-cycle without resorting to any regularity assumptions on the solution (see [7, 8, 10, 53] for details) or the weight ω.

5. Analysis of multigrid methods on quasi-uniform grids

In this section we consider the $\mathcal{V}$-cycle multigrid method applied to solve the weighted discrete problem (3.7) on quasi-uniform meshes. We consider standard point-wise Gauss-Seidel smoothers and prove convergence with a nearly optimal rate up to a factor $J \eqsim |\log h_J|$. Our main contribution is the extension of the standard multigrid analysis [12, 13, 37, 53] to include weights belonging to the Muckenhoupt class $A_2(\mathbb{R}^N)$. An optimal result for weights in the $A_1(\mathbb{R}^N)$-class is derived in [36]. Nevertheless, since our main motivation is the fractional Laplacian and the weight $y^\alpha \in A_2(\mathbb{R}^N) \setminus A_1(\mathbb{R}^N)$, we need to consider the larger class $A_2(\mathbb{R}^N)$.

5.1. **Stability of the nodal decomposition in the weighted L^2-norm.** Here we show that the nodal decomposition is stable in the weighted L^2-norm. Equivalently, the mass matrix for this inner product is spectrally equivalent to its diagonal.

Lemma 5.1 (Stability of the nodal decomposition)**.** *Let $\mathscr{T} \in \mathbb{T}$ be a quasi-uniform mesh, and let $v \in \mathbb{V}(\mathscr{T})$. Then, we have the following norm equivalence:*

$$\sum_{i=1}^{\mathcal{N}(\mathscr{T})} \|v_i\|^2_{L^2(\omega,\Omega)} \lesssim \|v\|^2_{L^2(\omega,\Omega)} \lesssim \sum_{i=1}^{\mathcal{N}(\mathscr{T})} \|v_i\|^2_{L^2(\omega,\Omega)}, \tag{5.1}$$

where $v = \sum_{i=1}^{\mathcal{N}(\mathscr{T})} v_i$ denotes the nodal decomposition for v, and the hidden constants in each inequality above only depend on the dimension n and $C_{2,\omega}$.

Proof. Let $\widehat{T} \subset \mathbb{R}^n$ be a reference element and $\{\widehat{\phi}_1, \dots, \widehat{\phi}_{\mathcal{N}_{\widehat{T}}}\}$ be its local shape functions, where $\mathcal{N}_{\widehat{T}}$ is the number of vertices of $\widehat{T}$. A standard argument shows that

$$\widehat{c}_1 \left(\int_{\widehat{T}} \widehat{\omega} \right) \sum_{i=1}^{\mathcal{N}_{\widehat{T}}} \widehat{V}_i^2 \le \|\widehat{v}\|^2_{L^2(\widehat{\omega},\widehat{T})} \le \widehat{c}_2 \left(\int_{\widehat{T}} \widehat{\omega} \right) \sum_{i=1}^{\mathcal{N}_{\widehat{T}}} \widehat{V}_i^2,$$

where $0 < \widehat{c}_1 \le \widehat{c}_2$, $\widehat{v} = \sum_{i=1}^{\mathcal{N}_{\widehat{T}}} \widehat{V}_i \widehat{\phi}_i$ and $\widehat{\omega}$ is a weight; see [29, Lemma 9.7]. Now, given $T \in \mathscr{T}$, we denote by $F_T : \widehat{T} \to T$ the mapping such that $\hat{v} = v \circ F_T$. Since the A_2-class is invariant under isotropic dilations [45, Proposition 2.1], a scaling argument shows that

$$\left(\int_T \omega \right) \sum_{i=1}^{\mathcal{N}_T} V_i^2 \lesssim \|v\|^2_{L^2(\omega,T)} \lesssim \left(\int_T \omega \right) \sum_{i=1}^{\mathcal{N}_T} V_i^2.$$

It remains to show that $\int_T \omega \approx \int_T \omega \phi_i^2$. The fact that $0 \le \phi_i \le 1$ immediately yields

$$\int_T \omega \phi_i^2 \le \int_T \omega.$$

The converse inequality follows from the *strong doubling property* of ω given in Proposition 2.1. In fact, setting $E = \{x \in T : \phi_i^2 \ge \frac{1}{2}\} \subset T$, we have

$$\int_T \omega \phi_i^2 \ge \int_E \omega \phi_i^2 \ge \frac{1}{2} \int_E \omega \ge \frac{1}{2C_{2,\omega}} \left(\frac{|E|}{|T|} \right)^2 \int_T \omega.$$

Finally, notice that the supports of the nodal basis functions $\{\phi_i\}_{i=1}^{\mathcal{N}(\mathscr{T})}$ have a finite overlap which is independent of the refinement level; i.e., for every $i = 1, \dots, \mathcal{N}(\mathscr{T})$, the number $n(i) = \#\{j : \operatorname{supp} \phi_i \cap \operatorname{supp} \phi_j \neq \emptyset\}$ is uniformly bounded. We arrive at (5.1) summing over all the elements $T \in \mathscr{T}$. $\square$

Let us now show a weighted inverse inequality.

Lemma 5.2 (Weighted inverse inequality)**.** *Let $\mathscr{T} \in \mathbb{T}$ be a quasi-uniform mesh, and let $T \in \mathscr{T}$ and $v \in \mathbb{V}(\mathscr{T})$. Then, we have the following inverse inequality:*

$$\|\nabla v\|_{L^2(\omega,T)} \lesssim h_{\mathscr{T}}^{-1} \|v\|_{L^2(\omega,T)}. \tag{5.2}$$

Proof. Since $\mathscr{T}$ is quasi-uniform with mesh size $h_{\mathscr{T}}$, we have $|\nabla \phi_i| \lesssim h_{\mathscr{T}}^{-1}$ and

$$\int_T \omega |\nabla v|^2 \lesssim h_{\mathscr{T}}^{-2} \sum_{i=1}^{\mathcal{N}_T} V_i^2 \int_T \omega,$$

where we have used the nodal decomposition of $v = \sum_{i=1}^{\mathcal{N}_T} V_i \phi_i$. As in the proof of Lemma 5.1, Proposition 2.1 yields

$$\int_T \omega \lesssim C_{2,\omega} \int_T \omega \phi_i^2,$$

so that we obtain

$$\int_T \omega |\nabla v|^2 \lesssim C_{2,\omega} h_{\mathscr{T}}^{-2} \sum_{i=1}^{\mathcal{N}_T} V_i^2 \int_T \omega \phi_i^2 \lesssim C_{2,\omega} h_{\mathscr{T}}^{-2} \int_T \omega v^2,$$

where, in the last step, we have used (5.1). This concludes the proof. □

5.2. **Convergence analysis.** We now present a convergence analysis of the multigrid $\mathcal{V}$-cycle applied to solve the weighted discrete problem (3.7) over quasi-uniform meshes and with standard point-wise Gauss-Seidel smoothers: $\mathcal{M}_k = \mathcal{N}_k$ and $\mathcal{I}_{k,j} = \{j\}$ for $j = 1, \dots \mathcal{N}_k$. The main ingredients are the stability of the nodal decomposition obtained in Lemma 5.1, the weighted inverse inequality of Lemma 5.2, and the properties of the quasi-interpolant introduced in Section 3. We follow [52, 54].

Theorem 5.3 (Convergence of symmetric $\mathcal{V}$-cycle multigrid)**.** *The multigrid $\mathcal{V}$-cycle with point-wise Gauss-Seidel smoother over quasi-uniform meshes is convergent with a contraction rate*

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

where C is independent of the mesh size, and it depends on the weight ω only through the constant $C_{2,\omega}$ defined in (2.1).

Proof. By the XZ identity stated in Theorem 4.1, we only need to estimate

$$c_0 = \sup_{\|v\|_{H_0^1(\omega,\Omega)}=1} \inf_{\sum_{k=0}^J \sum_{i=1}^{\mathcal{N}_k} v_{k,i} = v} \sum_{k=0}^J \sum_{i=1}^{\mathcal{N}_k} \left\| \nabla \left(P_{k,i} \sum_{(l,j)\succ(k,i)} v_{l,j} \right) \right\|_{L^2(\omega,\Omega)}^2, \tag{5.3}$$

where $\succ$ stands for the so-called *lexicographic ordering*, i.e.,

$$(l,j) \succ (k,i) \Leftrightarrow \begin{cases} l > k, \\ l = k \quad \text{and} \quad j > i. \end{cases}$$

We recall that $k = 0, \dots, J$, $j = 1, \dots, \mathcal{N}_k$ and the operator $P_{k,i} : \mathbb{V}_k \to \mathbb{V}_{k,i}$ is the projection with respect to the bilinear form a. For $k = 0, \dots, J$ we denote by $\Pi_{\mathscr{T}_k}$ the quasi-interpolation operator defined in §3.2 over the mesh $\mathscr{T}_k$. Next, we introduce the telescopic multilevel decomposition

$$v = \sum_{k=0}^J v_k, \qquad v_k = (\Pi_{\mathscr{T}_k} - \Pi_{\mathscr{T}_{k-1}})v, \qquad \Pi_{\mathscr{T}_{-1}} v := 0, \tag{5.4}$$

along with the nodal decomposition $v_k = \sum_{i=1}^{\mathcal{N}_k} v_{k,i}$, for each level k. Consequently, the right hand side of (5.3) can be rewritten by using the telescopic multilevel decomposition (5.4) as follows:

$$V_{k,i} := \sum_{(l,j)\succ(k,i)} v_{l,j} = \sum_{l=k+1}^{J}\sum_{j=1}^{\mathcal{N}_k} v_{l,j} + \sum_{j=i+1}^{\mathcal{N}_k} v_{k,j} = \sum_{l=k+1}^{J} v_l + \sum_{j=i+1}^{\mathcal{N}_k} v_{k,j}$$
$$= v - \Pi_{\mathscr{T}_k} v + \sum_{j=i+1}^{\mathcal{N}_k} v_{k,j},$$

because $\Pi_{\mathscr{T}_k}$ is invariant over $\mathbb{V}_k$. Therefore, we have

$$\|\nabla P_{k,i} V_{k,i}\|^2_{L^2(\omega,\Omega)} \lesssim \|\nabla P_{k,i}(v - \Pi_{\mathscr{T}_k} v)\|^2_{L^2(\omega,\Omega)} + \left\| \nabla P_{k,i} \sum_{j=i+1}^{\mathcal{N}_k} v_{k,j} \right\|^2_{L^2(\omega,\Omega)}$$
$$\lesssim \|\nabla(v - \Pi_{\mathscr{T}_k} v)\|^2_{L^2(\omega,D_{k,i})} + \sum_{\substack{j=i+1\\ D_{k,i}\cap D_{k,j}\neq\emptyset}}^{\mathcal{N}_k} \|\nabla v_{k,j}\|^2_{L^2(\omega,\Omega)},$$

where $D_{k,i} = \operatorname{supp}\phi_{k,i}$. Adding over $i = 1, \dots, \mathcal{N}_k$ and using the finite overlapping property of the sets $D_{k,i}$ yield

$$\sum_{i=1}^{\mathcal{N}_k} \sum_{\substack{j=i+1\\ D_{k,i}\cap D_{k,j}\neq\emptyset}}^{\mathcal{N}_k} \|\nabla v_{k,j}\|^2_{L^2(\omega,\Omega)} \lesssim \sum_{i=1}^{\mathcal{N}_k} \|\nabla v_{k,i}\|^2_{L^2(\omega,\Omega)},$$

whence, the weighted inverse inequality (5.2) gives

$$\sum_{i=1}^{\mathcal{N}_k} \|\nabla P_{k,i} V_{k,i}\|^2_{L^2(\omega,\Omega)} \lesssim \|\nabla(v - \Pi_{\mathscr{T}_k} v)\|^2_{L^2(\omega,\Omega)} + \sum_{i=1}^{\mathcal{N}_k} h_k^{-2} \|v_{k,i}\|^2_{L^2(\omega,\Omega)}.$$

We resort to (stability of the operator $\Pi_{\mathscr{T}_k}$) Proposition 3.2 and (stability of the micro-decomposition) Lemma 5.1 to arrive at

$$\sum_{i=1}^{\mathcal{N}_k} \|\nabla P_{k,i} V_{k,i}\|^2_{L^2(\omega,\Omega)} \lesssim \|\nabla v\|^2_{L^2(\omega,\Omega)} + h_k^{-2} \|v_k\|^2_{L^2(\omega,\Omega)}.$$

Since $v_k = (\Pi_{\mathscr{T}_k} - \Pi_{\mathscr{T}_{k-1}})v$, we utilize the approximation properties of $\Pi_{\mathscr{T}_k}$, given in Proposition 3.1, to deduce

$$\|v_k\|_{L^2(\omega,\Omega)} \le \|v - \Pi_{\mathscr{T}_k} v\|_{L^2(\omega,\Omega)} + \|v - \Pi_{\mathscr{T}_{k-1}} v\|_{L^2(\omega,\Omega)} \lesssim h_k \|\nabla v\|_{L^2(\omega,\Omega)}.$$

This implies $\sum_{i=1}^{\mathcal{N}_k} \|\nabla P_{k,i} V_{k,i}\|^2_{L^2(\omega,\Omega)} \lesssim \|\nabla v\|^2_{L^2(\omega,\Omega)}$, and adding over k from 0 to J yields $c_0 \lesssim J$, which completes the proof. □

6. A multigrid method for the fractional Laplace operator on tensor product anisotropic meshes

As we explained in §3.3, the regularity estimate (3.14) implies the necessity of graded meshes in the extended variable y. This allows us to recover an almost-optimal error estimate for the finite element approximation of problem (1.2) [44, Theorem 5.4]. In fact, finite elements on quasi-uniform meshes have *poor* approximation properties for small values of the parameter s. The isotropic error estimates of [44, Theorem 5.1] are not optimal, which makes anisotropic estimates essential. For this reason, in this section we develop a multilevel theory for problem (1.2) having in mind anisotropic partitions in the extended variable y and the multilevel setting described in Section 4 for the nonuniformly elliptic equation (3.1). We shall obtain nearly uniform convergence of a $\mathcal{V}$-cycle multilevel method for the problem (1.2) without any regularity assumptions on the solution or weight. We consider line Gauss-Seidel smoothers. The analysis is an adaptation of the results presented in [52] for anisotropic elliptic equations, and it is again based on the XZ identity [55].

6.1. **A multigrid algorithm with line smoothers.** The success of multigrid methods for uniformly elliptic operators is due to the fact that the smoothers are effective in reducing the nonsmooth (high frequency) components of the error and the coarse grid corrections are effective in reducing the smooth (low frequency) components. However, the effectiveness of both strategies depends crucially on several factors such as the anisotropy of the mesh. A key ingredient in the design and analysis of a multigrid method on anisotropic meshes is the use of the so-called line smoothers; see [1, 11, 38, 48].

Intuitively, when solving the α-harmonic extension (1.2) on graded meshes, the approximation from the coarse grid is dominated by the larger mesh size in the x'-direction, and thus the coarse grid correction cannot capture the smaller scale in the y-direction. One possible solution is the use of semi-coarsening, i.e., coarsening only the y-direction until the mesh sizes in both directions are comparable. Another solution is the use of line smoothing, i.e., solving sub-problems restricted to one vertical line. We shall use the latter approach, which is relatively easy to implement for tensor-product meshes.

Let us describe the decomposition of $\mathbb{V}_J = \mathbb{V}(\mathscr{T}_{\mathcal{Y}_J})$ that we shall use. To do so, we follow the notation of §4.1. We set $\mathcal{M}_k$ to be the number of interior nodes of $\mathscr{T}_{\Omega,k}$ and define, for $j = 1, \ldots, \mathcal{M}_k$, the set $\mathcal{I}_{k,j}$ as the collection of indices for the vertices that lie on the line $\{\mathrm{v}'_j\} \times [0, \mathcal{Y})$ at the level k. The decomposition is then given by (4.1). This decomposition is also stable, which allows us to obtain the appropriate anisotropic inverse inequalities; see Lemma 6.1 below.

Owing to the nature of the decomposition, the smoother requires the evaluation of $A_{k,j}^{-1}$ which corresponds to the action of the operator over a vertical line. This can be efficiently realized since the corresponding matrix is tri-diagonal.

Lemma 6.1 (Nodal stability and anisotropic inverse inequalities)**.** *Let $\mathscr{T}_{\mathcal{Y}}$ be a tensor product graded grid, which is quasi-uniform in Ω and graded in the extended variable so that* (3.16) *holds. If $v = \sum_{j=1}^{\mathcal{M}_J} v_j \in \mathbb{V}(\mathscr{T}_{\mathcal{Y}})$, then*

$$\sum_{j=1}^{\mathcal{M}_J} \|v_j\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \lesssim \|v\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \lesssim \sum_{j=1}^{\mathcal{M}_J} \|v_j\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})}. \tag{6.1}$$

Moreover, we have the following inverse inequalities:

$$\|\nabla_{x'} v\|_{L^2(y^\alpha,T)} \lesssim h_K^{-1} \|v\|_{L^2(y^\alpha,T)}, \qquad \|\partial_y v\|_{L^2(y^\alpha,T)} \lesssim h_I^{-1} \|v\|_{L^2(y^\alpha,T)}, \tag{6.2}$$

where $T = K \times I$ is a generic element of $\mathscr{T}_{\mathcal{Y}}$.

Proof. The nodal stability (6.1) follows along the same lines of Lemma 5.1 upon realizing that the functions $v_j = v_j(x', y)$ are defined on the vertical lines (v'_j, y) with $y \in [0, \mathcal{Y})$ and the index j corresponds to a nodal decomposition in Ω. Using that $|\nabla_{x'} \phi_i| \lesssim h_K^{-1}$ and $|\partial_y \phi_i| \lesssim h_I^{-1}$, we derive (6.2) inspired by Lemma 5.2. □

We now examine the $\mathcal{V}$-cycle multigrid method applied to the decomposition (4.1) with exact sub-solvers on $\mathbb{V}_{k,j}$, i.e., with line smoothers; see [10, §III.12] and [52]. A key observation in favor of subspaces $\{\mathbb{V}_{k,j}\}_{j=1}^{\mathcal{M}_k}$ follows.

Lemma 6.2 (Nodal stability of y-derivatives)**.** *Under the same assumptions of Lemma* 6.1 *we have*

$$\sum_{j=1}^{\mathcal{M}_J} \|\partial_y v_j\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} \lesssim \|\partial_y v\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} \lesssim \sum_{j=1}^{\mathcal{M}_J} \|\partial_y v_j\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})}. \tag{6.3}$$

Proof. We just proceed as in Lemma 5.1 with v replaced by $\partial_y v = \sum_{j=1}^{\mathcal{M}_J} \partial_y v_j$. □

Exploiting Theorem 4.1, the properties of the quasi-interpolation operator $\Pi_{\mathscr{T}_k}$ defined in §3.2, and Lemmas 6.1 and 6.2, we obtain the nearly uniform convergence of the symmetric $\mathcal{V}$-cycle multigrid method. We follow [52, 54].

Theorem 6.3 (Convergence of multigrid methods with line smoothers)**.** *The symmetric $\mathcal{V}$-cycle multigrid method with line smoothing converges with a contraction rate*

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

where C is independent of the number of degrees of freedom. The constant C depends on the weight y^α only through the constant C_{2,y^α} and on s like $C \approx \gamma$, where γ is the parameter that defines the graded mesh (3.16).

Proof. We again use the XZ identity (4.1) and modify the arguments in the proof of Theorem 5.3. We introduce the telescopic multilevel decomposition

$$v = \sum_{k=0}^{J} v_k, \qquad v_k = (\Pi_{\mathscr{T}_{\mathscr{Y},k}} - \Pi_{\mathscr{T}_{\mathscr{Y},k-1}})v, \qquad \Pi_{\mathscr{T}_{\mathscr{Y},-1}} v := 0 \tag{6.4}$$

and the line decomposition $v_k = \sum_{j=1}^{\mathcal{M}_k} v_{k,j}$. With the same arguments as in the proof of Theorem 5.3 and denoting $V_{k,i} = \sum_{(l,j)\succ(k,i)} v_{l,j}$, we arrive at the inequality

$$\sum_{i=1}^{\mathcal{M}_k} \|\nabla P_{k,i} V_{k,i}\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} \lesssim \|\nabla(v - \Pi_{\mathscr{T}_{\mathscr{Y},k}} v)\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} + \sum_{j=1}^{\mathcal{M}_k} \|\nabla v_{k,j}\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})}, \tag{6.5}$$

where we used the finite overlapping property of the sets $\mathcal{I}_{k,j}$; see §4.1. It remains to estimate both terms in (6.5). The stability of the quasi-interpolant $\Pi_{\mathscr{T}_{\mathscr{Y},k}}$ stated in (3.11) ([44, Theorems 4.7 and 4.8] and [45, Lemma 5.1]) yields

$$\|\nabla(v - \Pi_{\mathscr{T}_{\mathscr{Y},k}} v)\|_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} \lesssim \|\nabla v\|_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})}. \tag{6.6}$$

To estimate the second term in (6.5) we begin by noticing that

$$\sum_{j=1}^{\mathcal{M}_k} \|\nabla v_{k,j}\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} = \sum_{j=1}^{\mathcal{M}_k} \|\nabla_{x'} v_{k,j}\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} + \sum_{j=1}^{\mathcal{M}_k} \|\partial_y v_{k,j}\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})}. \tag{6.7}$$

The first term is estimated via the first weighted inverse inequality (6.2) and the stability of the nodal decomposition (6.1), that is,

$$\sum_{j=1}^{\mathcal{M}_k} \|\nabla_{x'} v_{k,j}\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} \lesssim \sum_{j=1}^{\mathcal{M}_k} {h'_k}^{-2} \|v_{k,j}\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})} \lesssim {h'_k}^{-2} \|v_k\|^2_{L^2(y^\alpha,\mathcal{C}_\mathscr{Y})}, \tag{6.8}$$

where h_k' is the mesh size in the x' direction at level k. The approximation property of $\Pi_{\mathscr{T}_{\mathcal{Y},k}}$ stated in Lemma 3.2 ([45, Theorem 5.7]) and the definition of v_k yield

$$\begin{aligned}\|v_k\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} &\le \|v-\Pi_{\mathscr{T}_{\mathcal{Y},k}}v\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} + \|v-\Pi_{\mathscr{T}_{\mathcal{Y},k-1}}v\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})}\\ &\lesssim h_k'\|\nabla_{x'}v\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} + h_k''\|\partial_y v\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})}\end{aligned}$$

where h_k'' denotes the *maximal* mesh size in the y direction at level k. Using (3.17) we see that $h_k'' \lesssim \gamma h_k'$, and replacing the estimate above in (6.8), we obtain

$$\sum_{j=1}^{\mathcal{M}_k} \|\nabla_{x'}v_{k,j}\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \lesssim \|\nabla v\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})}, \tag{6.9}$$

which bounds the first term in (6.7). To estimate the second term, we resort to Lemma 6.2, namely

$$\sum_{j=1}^{\mathcal{M}_k} \|\partial_y v_{k,j}\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \lesssim \|\partial_y v_k\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})}. \tag{6.10}$$

Finally, inequalities (6.9) and (6.10) allow us to conclude that

$$\sum_{j=1}^{\mathcal{M}_k} \|\nabla v_{k,j}\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \lesssim \|\nabla v\|^2_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})},$$

which together with (6.6) yields the desired result after summing over k. □

Remark 6.4 (Dependence on s). We point out the use of (3.17), which in turn implies $h_k'' \lesssim \gamma h_k'$, to derive (6.9). This translates into $C \approx \gamma$ in Theorem 6.3 and, since $\gamma > 3/(1-\alpha) = 3/(2s)$, in deterioration of the contraction factor as $s \downarrow 0$. We explore a remedy in §7.3.

7. Numerical illustrations for fractional diffusion

Here we present numerical experiments to support our theory. We consider

$$n = 1, \quad \Omega = (0,1), \quad u = \sin(3\pi x), \tag{7.1}$$
$$n = 2, \quad \Omega = (0,1)^2, \quad u = \sin(2\pi x_1)\sin(2\pi x_2), \tag{7.2}$$

and $\mathcal{Y} = 1$. The length $\mathcal{Y}$ of the cylinder in the extended direction is fixed, as discussed in [44], so that it captures the exponential decay of the solution. All of our algorithms are implemented based on the MATLAB© software package *i*FEM [20].

7.1. **Multigrid with line smoothers on graded meshes.** We partition Ω into a uniform grid of size $h_{\mathscr{T}_\Omega}$, and we construct a graded mesh in the extended direction using (3.16) with $\gamma = 3/2s + 0.1$ and $M = 1/h_{\mathscr{T}_\Omega}$. The points are ordered columnwise so that the indices associated to vertical lines are easily accessible. Starting from $h_{\mathscr{T}_\Omega,0} = 0.25$ we obtain a sequence of meshes by halving the mesh size of Ω and applying (3.16) in the extended direction with double number of mesh points.

We assemble the matrix corresponding to the finite element discretization of (3.7) on each level. The natural embedding $\mathbb{V}(\mathscr{T}_k) \to \mathbb{V}(\mathscr{T}_{k+1})$ for $k = 0,\dots,J-1$ gives us the prolongation matrix between two consecutive levels. Notice that the prolongation in the x'-direction is obtained by standard averaging, while in the extended direction the weights must be modified to take into account the grading of the mesh. The restriction matrix is the transpose of the prolongation matrix.

As discussed in Section 6 we must use vertical line smoothers to attain efficiency. The tri-diagonal sub-matrix corresponding to one vertical line is inverted exactly by the built-in direct solver in MATLAB©. Red-black ordering of the indices in the x'-direction is used to further improve the efficiency of the line smoothers. We perform three pre- and post-smoothing steps. Starting from a zero initial guess we use as exit criterion that the ℓ^2-norm of the relative residual is smaller than 10^{-7}.

Tables 1 and 2 show the number of iterations for the implemented multigrid method for the one and two dimensional problems, respectively. As we see, the method converges almost uniformly with respect to the number of degrees of freedom. Notice that the number of iterations for $s = 0.15$ is significantly larger than those for the remaining tested cases. This can be explained by the fact that, as Theorem 6.3 states, the contraction factor depends on $\gamma \approx 1/s$, and thus we observe a preasymptotic regime where the number of iterations grows. This is exactly the case for the one dimensional problem, and we would expect a similar behavior in the two dimensional case. However, since the extended problem is now in three dimensions, the size of the problems grows rather quickly, and thus computational resources were not sufficient to deal with the cases $h_{\mathscr{T}_\Omega} = \frac{1}{256}, \frac{1}{512}$. In §7.3 we propose a modification of the graded mesh to address this issue.

Table 1. Number of iterations for a multigrid method for the one dimensional fractional Laplacian using a line smoother in the extended direction. The mesh in Ω is uniform of size $h_{\mathscr{T}_\Omega}$. The mesh in the extended direction is graded according to (3.16).

$h_{\mathscr{T}_\Omega}$	DOFs	$s = 0.15$	$s = 0.3$	$s = 0.6$	$s = 0.8$
$\frac{1}{16}$	289	7	6	5	5
$\frac{1}{32}$	1,089	13	9	6	6
$\frac{1}{64}$	4,225	25	10	6	6
$\frac{1}{128}$	16,641	33	11	6	6
$\frac{1}{256}$	66,049	37	10	6	6
$\frac{1}{512}$	263,169	38	10	6	7

Table 2. Number of iterations for a multigrid method for the two dimensional fractional Laplacian using a line smoother in the extended direction. The mesh in Ω is uniform of size $h_{\mathscr{T}_\Omega}$. The mesh in the extended direction is graded according to (3.16).

$h_{\mathscr{T}_\Omega}$	DOFs	$s = 0.15$	$s = 0.3$	$s = 0.6$	$s = 0.8$
$\frac{1}{16}$	4,913	10	7	6	5
$\frac{1}{32}$	35,937	19	8	6	6
$\frac{1}{64}$	274,625	34	9	6	6
$\frac{1}{128}$	2,146,689	47	9	6	6

We also tested a point Gauss-Seidel smoother for the one dimensional case $\Omega = (0,1)$. Except for the trivial case $h_{\mathscr{T}_\Omega} = 1/16$, the corresponding $\mathcal{V}$-cycle is not able to achieve the desired accuracy in 200 iterations.

7.2. **Multigrid methods on quasi-uniform meshes.** Even though the approximation of the Caffarelli-Silvestre extension of the fractional Laplace operator on quasi-uniform meshes in the extended direction is suboptimal, let us use this problem to illustrate the convergence properties of the multilevel method, developed in Section 5, for general A_2 weights. The setting is the same as in §7.1, but we use quasi-uniform meshes and a point-wise Gauss-Seidel smoother. Tables 3 and 4 show the number of iterations with respect to the number of degrees of freedom and s. The convergence is almost uniform with respect to the number of unknowns as well as the parameter $s \in (0,1)$.

TABLE 3. Number of iterations for a multigrid method with point-wise Gauss-Seidel smoothers on uniform meshes for the one dimensional fractional Laplacian.

$h_{\mathscr{T}_\Omega}$	DOFs	$s = 0.15$	$s = 0.3$	$s = 0.6$	$s = 0.8$
$\frac{1}{16}$	289	12	13	13	14
$\frac{1}{32}$	1,089	15	15	15	17
$\frac{1}{64}$	4,225	15	16	16	17
$\frac{1}{128}$	16,641	15	16	16	18
$\frac{1}{256}$	66,049	15	15	16	18
$\frac{1}{512}$	263,169	15	15	16	18

TABLE 4. Number of iterations for a multigrid method with point-wise Gauss-Seidel smoothers on uniform meshes for the two dimensional fractional Laplacian.

$h_{\mathscr{T}_\Omega}$	DOFs	$s = 0.15$	$s = 0.3$	$s = 0.6$	$s = 0.8$
$\frac{1}{16}$	4,913	13	12	13	15
$\frac{1}{32}$	35,937	15	15	15	17
$\frac{1}{64}$	274,625	15	16	16	18
$\frac{1}{128}$	2,146,689	15	16	16	19

7.3. **Modified mesh grading.** Examining the proof of Theorem 6.3, we realize that the critical step (6.9) consists in the application of (3.17), namely $h_k'' \lesssim \gamma h_k'$, which deteriorates as s becomes small because $\gamma > 3/(1-\alpha) = 3/(2s)$. Numerically, this effect can be seen in Tables 1 and 2, where, for instance, the number of iterations needed for $s = 0.15$ is significantly larger than those for all the other tested values. As a result, the contraction rate of Theorem 6.3 becomes $1 - 1/(1 + C\gamma J)$. Here we explore computationally how to overcome this issue. We construct a mesh such that the maximum mesh size in the extended direction is uniformly bounded, with respect to s, by the uniform mesh size in the x'-direction without changing the ratio of degrees of freedom in Ω and the extended direction by more than a constant.

Let us begin with some heuristics. To control the aspect ratio h''_k/h'_k uniformly on $s \in (0,1)$, we may apply some extra refinements to the largest elements in the y direction, increasing the number of degrees of freedom of $\mathscr{T}_\mathcal{Y}$ just by a constant factor. We denote by $\tilde{\mathscr{T}}_\mathcal{Y}$ the resulting mesh and we notice that $\mathbb{V}(\mathscr{T}_\mathcal{Y}) \subset \mathbb{V}(\tilde{\mathscr{T}}_\mathcal{Y})$. Thus, Galerkin orthogonality implies

$$\|\nabla(v - V_{\tilde{\mathscr{T}}_\mathcal{Y}})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \le \|\nabla(v - V_{\mathscr{T}_\mathcal{Y}})\|_{L^2(y^\alpha,\mathcal{C}_\mathcal{Y})} \lesssim (\#\mathscr{T}_\mathcal{Y})^{-\frac{1}{n+1}} \approx (\#\tilde{\mathscr{T}}_\mathcal{Y})^{-\frac{1}{n+1}}.$$

We build on this idea through a modification of the mapping function below.

Let $F : (0,1) \to (0,\mathcal{Y})$ be an increasing and differentiable function such that $F(0) = 0$ and $F(1) = \mathcal{Y}$. By mapping a uniform grid of $(0,1)$ via the function F, we can construct a graded mesh with mesh points given by $y_l = F(l/M)$ for $l = 1,\dots,M$: $F(\xi) = \mathcal{Y}\xi^\gamma$ yields (3.16) and a linear F gives a uniform grid. The mean value theorem implies

$$y_{l+1} - y_l = \frac{F'(c_l)}{M} \le \frac{1}{M}\max\left\{|F'(\xi)| : \xi \in \left[\frac{l}{M}, \frac{l+1}{M}\right]\right\},$$

which shows that the map of (3.16) is not uniformly bounded with respect to s.

For this reason, we instead consider the following construction: Let $(\xi_\star, y_\star) \in (0,1)^2$, which we will call the *transition point*, and define the mapping

$$F(\xi) = y_\star\mathcal{Y}\left(\frac{\xi}{\xi_\star}\right)^\gamma,\ 0 < \xi \le \xi_\star, \quad F(\xi) = \mathcal{Y}\left(\frac{1-y_\star}{1-\xi_\star}(\xi - \xi_\star) + y_\star\right),\ \xi_\star < \xi < 1.$$

Over the interval $(0,\xi_\star)$ the mapping F defines the same type of graded mesh, but over $(\xi_\star,1)$ it defines a uniform mesh. Let us now choose the transition point to obtain a bound on the derivative of F. We have

$$F'(\xi) = \gamma\mathcal{Y}\frac{y_\star}{\xi_\star}\left(\frac{\xi}{\xi_\star}\right)^{\gamma-1},\ 0 < \xi \le \xi_\star, \quad F'(\xi) = \mathcal{Y}\frac{1-y_\star}{1-\xi_\star},\ \xi_\star < \xi < 1, \tag{7.3}$$

so that $\mathcal{S} := \max_{\xi\in[0,1]}|F'(\xi)| = \mathcal{Y}\max\left\{\gamma\frac{y_\star}{\xi_\star}, \frac{1-y_\star}{1-\xi_\star}\right\}$. Given $\xi_\star$ we choose $y_\star$ to have $\gamma\frac{y_\star}{\xi_\star} = \frac{1-y_\star}{1-\xi_\star}$, i.e., $y_\star = \left(1 + \gamma\frac{1-\xi_\star}{\xi_\star}\right)^{-1}$. This yields $F \in \mathcal{C}^1([0,1])$ and, more importantly,

$$\mathcal{S} = \gamma\mathcal{Y}\frac{y_\star}{\xi_\star} = \mathcal{Y}\frac{\gamma}{\xi_\star + (1-\xi_\star)\gamma} \le \mathcal{Y}\frac{1}{1-\xi_\star}.$$

We can now choose $\xi_\star$ to gain control of $\mathcal{S}$. For instance, $\xi_\star = 0.5$ gives us that $\mathcal{S} \le 2\mathcal{Y}$, and $\xi_\star = 0.75$ gives us that $\mathcal{S} \le 4\mathcal{Y}$. In the experiments presented below we choose $\xi_\star = 0.75$. The theory presented in Section 6 still applies.

The modified graded meshes have asymptotically the same distribution of points near the bottom part of the cylinder, and so they are also capable of capturing the singular behavior of the solution $\mathscr{U}$. However, near the top part, the aspect ratio is uniformly controlled by a factor 4. The modified mesh is only applied for $\gamma > 4$. Therefore for $s = 0.3$, 0.6 and 0.8, no modification is needed in the original mesh.

Upon constructing a mesh with this modification, we can develop a $\mathcal{V}$-cycle multigrid solver with vertical line smoothers. Comparisons of this approach with the setting of §7.1 are shown in Tables 5 and 6. From them we can conclude that the strong anisotropic behavior of the mesh grading (3.16) affects the performance of the $\mathcal{V}$-cycle multigrid with vertical line smoothers. For the original graded meshes, there is a preasymptotic regime where the number of iterations increases faster than $\log J$. The modification of the mesh proposed in (7.3) allows us to obtain an almost uniform number of iterations for all problem sizes without sacrificing the near optimal order of convergence of the method.

TABLE 5. Comparison of the multilevel solver with vertical line smoother over two graded meshes for the one dimensional fractional Laplacian, $s = 0.15$. *Legends*: The original mesh, given by (3.16), is denoted by o, whereas the modification proposed in (7.3) is denoted by m; I – iterations, E – error in the energy norm.

$h_{\mathscr{T}_\Omega}$	DOFs	I(o)	I(m)	E(o)	E(m)
$\frac{1}{16}$	289	7	7	0.1556	0.1739
$\frac{1}{32}$	1,089	13	9	0.0828	0.0937
$\frac{1}{64}$	4,225	25	10	0.0426	0.0485
$\frac{1}{128}$	16,641	33	10	0.0216	0.0246
$\frac{1}{256}$	66,049	37	11	0.0109	0.0124
$\frac{1}{512}$	263,169	38	11	0.0055	0.0062

TABLE 6. Comparison of the multilevel solver with vertical line smoother over two graded meshes for the two dimensional fractional Laplacian, $s = 0.15$. *Legends*: The original mesh, given by (3.16), is denoted by o, whereas the modification proposed in (7.3) is denoted by m; I – iterations, E – error in the energy norm.

$h_{\mathscr{T}_\Omega}$	DOFs	I(o)	I(m)	E(o)	E(m)
$\frac{1}{16}$	4,913	10	8	0.1070	0.1198
$\frac{1}{32}$	35,937	19	11	0.0570	0.0646
$\frac{1}{64}$	274,625	34	12	0.0294	0.0334
$\frac{1}{128}$	2,146,689	47	13	0.0149	0.0170

REFERENCES

[1] T. Apel and J. Schöberl, *Multigrid methods for anisotropic edge refinement*, SIAM J. Numer. Anal. **40** (2002), no. 5, 1993–2006 (electronic), DOI 10.1137/S0036142900375414. MR1950630 (2003m:65225)

[2] D. Arroyo, A. Bespalov, and N. Heuer, *On the finite element method for elliptic problems with degenerate and singular coefficients*, Math. Comp. **76** (2007), no. 258, 509–537, DOI 10.1090/S0025-5718-06-01910-7. MR2291826 (2008e:65336)

[3] P. W. Bates, *On some nonlocal evolution equations arising in materials science*, Nonlinear dynamics and evolution equations, Fields Inst. Commun., vol. 48, Amer. Math. Soc., Providence, RI, 2006, pp. 13–52. MR2223347 (2007g:35097)

[4] Z. Belhachmi, C. Bernardi, and S. Deparis, *Weighted Clément operator and application to the finite element discretization of the axisymmetric Stokes problem*, Numer. Math. **105** (2006), no. 2, 217–247, DOI 10.1007/s00211-006-0039-9. MR2262757 (2008c:65310)

[5] M. Š. Birman and M. Z. Solomjak, *Spektralnaya teoriya samosopryazhennykh operatorov v gilbertovom prostranstve* (Russian), Leningrad. Univ., Leningrad, 1980. MR609148 (82k:47001)

[6] A. Bonito and J. E. Pasciak, *Numerical approximation of fractional powers of elliptic operators*, Math. Comp. **84** (2015), no. 295, 2083–2110, DOI 10.1090/S0025-5718-2015-02937-8. MR3356020

[7] J. H. Bramble and J. E. Pasciak, *New convergence estimates for multigrid algorithms*, Math. Comp. **49** (1987), no. 180, 311–329, DOI 10.2307/2008314. MR906174 (89b:65234)

[8] J. H. Bramble, J. E. Pasciak, J. P. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp. **57** (1991), no. 195, 23–45, DOI 10.2307/2938661. MR1079008 (91m:65158)

[9] J. H. Bramble, J. E. Pasciak, J. P. Wang, and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition*, Math. Comp. **57** (1991), no. 195, 1–21, DOI 10.2307/2938660. MR1090464 (92d:65094)

[10] J. H. Bramble and X. Zhang, *The analysis of multigrid methods*, Handbook of numerical analysis, Vol. VII, Handb. Numer. Anal., VII, North-Holland, Amsterdam, 2000, pp. 173–415. MR1804746 (2001m:65183)

[11] J. H. Bramble and X. Zhang, *Uniform convergence of the multigrid V-cycle for an anisotropic problem*, Math. Comp. **70** (2001), no. 234, 453–470, DOI 10.1090/S0025-5718-00-01222-9. MR1709148 (2001g:65134)

[12] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp. **31** (1977), no. 138, 333–390. MR0431719 (55 #4714)

[13] A. Brandt, *Multigrid Techniques: 1984 Guide with Applications to Fluid Dynamics*, GMD-Studien [GMD Studies], vol. 85, Gesellschaft für Mathematik und Datenverarbeitung mbH, St. Augustin, 1984. MR772748 (87c:65139b)

[14] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Applied Mathematics, vol. 15, Springer, New York, 2008. MR2373954 (2008m:65001)

[15] A. Bueno-Orovio, D. Kay, V. Grau, B. Rodriguez, and K. Burrage, *Fractional diffusion models of cardiac electrical propagation: role of structural heterogeneity in dispersion of repolarization*, J. R. Soc. Interface, 11(97), 2014.

[16] X. Cabré and Y. Sire, *Nonlinear equations for fractional Laplacians II: Existence, uniqueness, and qualitative properties of solutions*, Trans. Amer. Math. Soc. **367** (2015), no. 2, 911–941, DOI 10.1090/S0002-9947-2014-05906-0. MR3280032

[17] L. Caffarelli and L. Silvestre, *An extension problem related to the fractional Laplacian*, Comm. Partial Differential Equations **32** (2007), no. 7-9, 1245–1260, DOI 10.1080/03605300600987306. MR2354493 (2009k:35096)

[18] 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** (2011), no. 8, 1353–1384, DOI 10.1080/03605302.2011.562954. MR2825595 (2012h:35361)

[19] P. Carr, H. Geman, D. B. Madan, and M. Yor, *The fine structure of asset returns: An empirical investigation*, Journal of Business, **75** (2002), 305–33.

[20] L. Chen, *i*FEM: An integrated finite element methods package in matlab, Technical report, University of California at Irvine, 2009.

[21] L. Chen, *Deriving the X-Z identity from auxiliary space method*, Domain decomposition methods in science and engineering XIX, Lect. Notes Comput. Sci. Eng., vol. 78, Springer, Heidelberg, 2011, pp. 309–316, DOI 10.1007/978-3-642-11304-8_35. MR2867674

[22] W. Chen, *A speculative study of* 2/3*-order fractional laplacian modeling of turbulence: Some thoughts and conjectures*, Chaos **16** (2006), no. 2, 1–11.

[23] D. Cho, J. Xu, and L. Zikatanov, *New estimates for the rate of convergence of the method of subspace corrections*, Numer. Math. Theory Methods Appl. **1** (2008), no. 1, 44–56. MR2401666 (2009b:65077)

[24] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics in Applied Mathematics, vol. 40, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)]. MR1930132

[25] J. Cushman and T. Glinn, *Nonlocal dispersion in media with continuously evolving scales of heterogeneity*, Trans. Porous Media **13** (1993), 123–138.

[26] Q. Du, M. Gunzburger, R. B. Lehoucq, and K. Zhou, *Analysis of the volume-constrained peridynamic Navier equation of linear elasticity*, J. Elasticity **113** (2013), no. 2, 193–217, DOI 10.1007/s10659-012-9418-x. MR3102595

[27] G. Duvaut and J.-L. Lions, *Inequalities in Mechanics and Physics,* translated from the French by C. W. John, Grundlehren der Mathematischen Wissenschaften, 219, Springer-Verlag, Berlin-New York, 1976. MR0521262 (58 #25191)

[28] A. C. Eringen, *Nonlocal Continuum Field Theories*, Springer-Verlag, New York, 2002. MR1918950 (2003m:74003)

[29] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Applied Mathematical Sciences, vol. 159, Springer-Verlag, New York, 2004. MR2050138 (2005d:65002)

[30] E. B. Fabes, C. E. Kenig, and R. P. Serapioni, *The local regularity of solutions of degenerate elliptic equations*, Comm. Partial Differential Equations **7** (1982), no. 1, 77–116, DOI 10.1080/03605308208820218. MR643158 (84i:35070)

[31] P. Gatto and J. S. Hesthaven, *Numerical approximation of the fractional Laplacian via hp-finite elements, with an application to image denoising*, J. Sci. Comput. **65** (2015), no. 1, 249–270, DOI 10.1007/s10915-014-9959-1. MR3394445

[32] G. Gilboa and S. Osher, *Nonlocal operators with applications to image processing*, Multiscale Model. Simul. **7** (2008), no. 3, 1005–1028, DOI 10.1137/070698592. MR2480109 (2010b:94006)

[33] V. Gol′dshtein and A. Ukhlov, *Weighted Sobolev spaces and embedding theorems*, Trans. Amer. Math. Soc. **361** (2009), no. 7, 3829–3850, DOI 10.1090/S0002-9947-09-04615-7. MR2491902 (2010b:46068)

[34] J. Gopalakrishnan and J. E. Pasciak, *The convergence of V-cycle multigrid algorithms for axisymmetric Laplace and Maxwell equations*, Math. Comp. **75** (2006), no. 256, 1697–1719 (electronic), DOI 10.1090/S0025-5718-06-01884-9. MR2240631 (2007g:65116)

[35] L. Greengard and V. Rokhlin, *A fast algorithm for particle simulations*, J. Comput. Phys. **73** (1987), no. 2, 325–348, DOI 10.1016/0021-9991(87)90140-9. MR918448 (88k:82007)

[36] M. Griebel, K. Scherer, and A. Schweitzer, *Robust norm equivalencies for diffusion problems*, Math. Comp. **76** (2007), no. 259, 1141–1161 (electronic), DOI 10.1090/S0025-5718-07-01973-4. MR2299769 (2008d:65149)

[37] W. Hackbusch, *Multigrid Methods and Applications*, Springer Series in Computational Mathematics, vol. 4, Springer-Verlag, Berlin, 1985. MR814495 (87e:65082)

[38] W. Hackbusch, *The frequency decomposition multi-grid method. I. Application to anisotropic equations*, Numer. Math. **56** (1989), no. 2-3, 229–245, DOI 10.1007/BF01409786. MR1018302 (90i:65212)

[39] W. Hackbusch, *A sparse matrix arithmetic based on $\mathcal{H}$-matrices. I. Introduction to $\mathcal{H}$-matrices*, Computing **62** (1999), no. 2, 89–108, DOI 10.1007/s006070050015. MR1694265 (2000c:65039)

[40] H. Harbrecht and R. Schneider, *Rapid solution of boundary integral equations by wavelet Galerkin schemes*, Multiscale, nonlinear and adaptive approximation, Springer, Berlin, 2009, pp. 249–294, DOI 10.1007/978-3-642-03413-8_8. MR2648376 (2011k:65170)

[41] T. P. Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) **175** (2012), no. 3, 1473–1506, DOI 10.4007/annals.2012.175.3.9. MR2912709

[42] N. S. Landkof, *Foundations of Modern Potential Theory,* translated from the Russian by A. P. Doohovskoy, Die Grundlehren der mathematischen Wissenschaften, Band 180, Springer-Verlag, New York-Heidelberg, 1972. MR0350027 (50 #2520)

[43] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. **165** (1972), 207–226. MR0293384 (45 #2461)

[44] 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** (2015), no. 3, 733–791, DOI 10.1007/s10208-014-9208-x. MR3348172

[45] R. H. Nochetto, E. Otárola, and A. J. Salgado, *Piecewise polynomial interpolation in Muckenhoupt weighted Sobolev spaces and applications*, Numer. Math. **132** (2016), no. 1, 85–130, DOI 10.1007/s00211-015-0709-6. MR3439216

[46] S. A. Silling, *Reformulation of elasticity theory for discontinuities and long-range forces*, J. Mech. Phys. Solids **48** (2000), no. 1, 175–209, DOI 10.1016/S0022-5096(99)00029-0. MR1727557 (2000i:74008)

[47] E. M. Stein, *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals,* with the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993. MR1232192 (95c:42002)

[48] R. Stevenson, *Robustness of multi-grid applied to anisotropic equations on convex domains and on domains with re-entrant corners*, Numer. Math. **66** (1993), no. 3, 373–398, DOI 10.1007/BF01385703. MR1246963 (94i:65047)

[49] R. Stevenson, *Adaptive wavelet methods for solving operator equations: an overview*, Multiscale, nonlinear and adaptive approximation, Springer, Berlin, 2009, pp. 543–597, DOI 10.1007/978-3-642-03413-8_13. MR2648381 (2011k:65196)

[50] P. R. Stinga and J. L. Torrea, *Extension problem and Harnack's inequality for some fractional operators*, Comm. Partial Differential Equations **35** (2010), no. 11, 2092–2122, DOI 10.1080/03605301003735680. MR2754080 (2012c:35456)

[51] B. O. Turesson, *Nonlinear Potential Theory and Weighted Sobolev Spaces*, Lecture Notes in Mathematics, vol. 1736, Springer-Verlag, Berlin, 2000. MR1774162 (2002f:31027)

[52] Y. Wu, L. Chen, X. Xie, and J. Xu, *Convergence analysis of V-cycle multigrid methods for anisotropic elliptic equations*, IMA J. Numer. Anal. **32** (2012), no. 4, 1329–1347, DOI 10.1093/imanum/drr043. MR2991830

[53] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev. **34** (1992), no. 4, 581–613, DOI 10.1137/1034116. MR1193013 (93k:65029)

[54] J. Xu, L. Chen, and R. H. Nochetto, *Optimal multilevel methods for* $H(\mathrm{grad})$*,* $H(\mathrm{curl})$*, and* $H(\mathrm{div})$ *systems on graded and unstructured grids*, Multiscale, nonlinear and adaptive approximation, Springer, Berlin, 2009, pp. 599–659, DOI 10.1007/978-3-642-03413-8_14. MR2648382 (2011k:65178)

[55] J. Xu and L. Zikatanov, *The method of alternating projections and the method of subspace corrections in Hilbert space*, J. Amer. Math. Soc. **15** (2002), no. 3, 573–597, DOI 10.1090/S0894-0347-02-00398-3. MR1896233 (2003f:65095)

[56] C.-S. Zhang, *Adaptive Finite Element Methods for Variational Inequalities: Theory and Applications In Finance*, ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)–University of Maryland, College Park, 2007. MR2711028

Department of Mathematics, University of California at Irvine, Irvine, California 92697

E-mail address: chenlong@math.uci.edu

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

E-mail address: rhn@math.umd.edu

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

E-mail address: enrique.otarola@usm.cl

Department of Mathematics, University of Tennessee, Knoxville, Tennessee 37996

E-mail address: asalgad1@utk.edu