SIAM J. NUMER. ANAL.
Vol. 62, No. 3, pp. 1212–1225

POINTWISE GRADIENT ESTIMATE OF THE RITZ PROJECTION*

LARS DIENING†, JULIAN ROLFES†, AND ABNER J. SALGADO‡

Abstract. Let $\Omega \subset \mathbb{R}^n$ be a convex polytope ($n \le 3$). The Ritz projection is the best approximation, in the $W_0^{1,2}$-norm, to a given function in a finite element space. When such finite element spaces are constructed on the basis of quasiuniform triangulations, we show a pointwise estimate on the Ritz projection. Namely, the gradient at any point in Ω is controlled by the Hardy–Littlewood maximal function of the gradient of the original function at the same point. From this estimate, the stability of the Ritz projection on a wide range of spaces that are of interest in the analysis of PDEs immediately follows. Among those are weighted spaces, Orlicz spaces, and Lorentz spaces.

Key words. maximal function, gradient estimates, Muckenhoupt weights, Ritz projection

MSC codes. 65N30, 65N80, 65N12

DOI. 10.1137/23M1571800

1. Introduction.

To approximate solutions of partial differential equations (PDEs), in particular, those that are second order and elliptic, the finite element method has emerged as the method of choice. A finite element scheme is nothing but a Galerkin approximation with a particular choice of finite dimensional subspace (piecewise polynomials subject to a triangulation of the domain) and a particular basis. It is fair to say that the study of the properties of finite element schemes for second order linear elliptic second order equations in an energy setting has reached a state of maturity. In short, the Ritz projection, which is the best approximation in the $W_0^{1,2}$-norm (see section 2 for notation), possesses optimal approximation properties when these are measured in the energy norm, which usually is a norm equivalent to the $W^{1,2}$-norm. This reduces the numerical analysis of a finite element scheme to a question of approximation theory, and this is usually resolved by constructing a suitable interpolant.

On the other hand, the study of the properties of the Ritz projection in nonenergy norms has been the subject of intensive study with many classical results, recent progress, and still some open questions. We refer the reader to the introductions of [15] and [7] for some historical accounts. It is fair to say that the development of this subject is obscured by technicalities, and it is far from settled. Nevertheless, apart from the intrinsic interest such estimates may present, these become important when dealing, for instance, with nonlinear or coupled problems, or even when in a linear problem the data is sufficiently rough that the functional setting that provides well-posedness is no longer the energy one (see, for instance, [10]), or when the energy norm is not equivalent to the usual $W^{1,2}$-norm (see [22]).

*Received by the editors May 10, 2023; accepted for publication (in revised form) October 25, 2023; published electronically May 21, 2024.
https://doi.org/10.1137/23M1571800
Funding: The work of the authors was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant SFB 1283/2 2021 - 317210226. The main results of this paper were obtained in February 2022 when the third author was visiting Bielefeld University for the conference "*Nonlocal Equations: Analysis and Numerics.*" The work of the third author was also partially supported by the National Science Foundation grant DMS-2111228.
†Department of Mathematics, University of Bielefeld, Postfach 10 01 31, 33501 Bielefeld, Germany (lars.diening@uni-bielefeld.de, julian.rolfes@uni-bielefeld.de, http://www.bi-discrete.com).
‡Department of Mathematics, University of Tennessee, Knoxville, TN 37996 USA (asalgad1@utk.edu, https://math.utk.edu/people/abner-j-salgado/).

The purpose of this work is to make a contribution in this direction. In this work we concentrate on the Ritz projection subject to homogeneous Dirichlet boundary values. Homogeneity is merely a matter of convenience. It might be possible to consider Neumann boundary conditions, see [31] for results in this direction. In fact one would only need suitable Green's function estimates. Treating mixed boundary values, however, is difficult in this context since the regularity corresponds to the one of a slit domain. For such domains the needed estimates of the Green's functions are not valid. We show that, over quasiuniform meshes, the gradient of the Ritz projection at any point in the domain is controlled by the Hardy–Littlewood maximal operator of the gradient of the original function at the same point. This pointwise estimate not only immediately implies stability of the Ritz projection in any function space where the maximal operator is bounded butalso elucidates the action of the Ritz projection, i.e., finite element approximation. It is a sort of averaging procedure.

Our presentation is organized as follows. In section 2 we introduce notation. The statement of our main result, Theorem 3.1, is presented in section 3. Here we also collect a list of corollaries. Some of these recover known results, whereas others are truly new and may find application in the finite element approximation of, for instance, nonlinear elliptic problems with nonstandard growth conditions [8]. The proof of our main result is the content of section 4. For clarity, this proof is split into several steps that comprise the bulk of this section.

2. Notation and preliminaries. We begin by introducing some notation and specifying the framework under which we shall operate. The relation $A \lesssim B$ means that there is a constant c for which $A \leq cB$. The value of this constant may change at each occurrence. More importantly, this constant does not depend on A, B, or discretization parameters. $A \approx B$ means that $A \lesssim B$ and $B \lesssim A$.

Throughout our work, $\Omega \subset \mathbb{R}^n$, $n \leq 3$, is a bounded convex polytope. While convexity is essential for our arguments, the dimensional restriction is merely an artifact of our methods. Given $x \in \mathbb{R}^n$, we denote its Euclidean norm by $|x|$. By $B(x,r)$ we denote the open ball with center $x \in \mathbb{R}^n$ and radius $r > 0$. For a measurable set $E \subset \mathbb{R}^n$ we denote by $|E|$ its Lebesgue measure. $L^0(\Omega)$ denotes the collection of functions $\Omega \to \mathbb{R}$ that are measurable. For $p \in [1,\infty]$ and $k \in \mathbb{N}$ we denote by $L^p(\Omega)$ and $W^{k,p}(\Omega)$, respectively, the usual Lebesgue and Sobolev spaces. The subspace of $W^{k,p}(\Omega)$ that consists of functions vanishing on the boundary is denoted by $W^{k,p}_0(\Omega)$. We immediately notice that, whenever $w \in W^{k,p}_0(\Omega)$, its extension to $\mathbb{R}^n \setminus \Omega$ by zero, denoted by $\tilde{w}$, is such that $\tilde{w} \in W^{k,p}(\mathbb{R}^n)$. For this reason, whenever necessary, we shall make this extension by zero without explicit mention or change of notation. By $L^1_{\mathrm{loc}}(\mathbb{R}^n)$ we denote the space of locally integrable functions. For $f \in L^0(\mathbb{R}^n)$ the (centered) Hardy–Littlewood maximal operator M of f is

$$M[f](x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)| \, \mathrm{d}y \tag{2.1}$$

for all $x \in \mathbb{R}^n$. With this notation $M[f]$ readily extends to vector valued functions. If X is a normed space, we shall denote by $\| \cdot \|_X$ its norm. If this norm comes from an inner product, this will be denoted by $\langle \cdot, \cdot \rangle_X$. We shall make no distinction between scalar and vector valued functions or their spaces, as this will be clear from the context. For $\alpha \in (0,1]$ we let $C^{0,\alpha}(\overline{\Omega})$ denote the space of Hölder continuous functions with seminorm

$$|f|_{C^{0,\alpha}(\overline{\Omega})} = \sup_{x,y \in \overline{\Omega}} \frac{|f(x) - f(y)|}{|x - y|^\alpha} \tag{2.2}$$

and norm $\|f\|_{C^{0,\alpha}(\overline{\Omega})} = \|f\|_{L^\infty(\Omega)} + |f|_{C^{0,\alpha}(\overline{\Omega})}$.

Let $\mathbb{T} = \{\mathcal{T}_h\}_{h>0}$ be a quasiuniform family of conforming triangulations of Ω in the sense of Ciarlet [3, p. 124] where, for $h > 0$, the triangulation $\mathcal{T}_h$ has mesh size h. For $k \in \mathbb{N}$ we denote by

$$\mathcal{L}^1_k(\mathcal{T}_h) = \left\{ w_h \in C(\overline{\Omega}) : w_{h|T} \in \mathbb{P}_k \ \forall T \in \mathcal{T}_h \right\}$$

the Lagrange space of degree k, where $\mathbb{P}_k$ is the space of polynomials of degree at most k. We set $V_h = \mathcal{L}^1_k(\mathcal{T}_h) \cap W^{1,1}_0(\Omega)$ and immediately observe that $V_h \subset W^{1,\infty}_0(\Omega)$. The Ritz projection $R_h : W^{1,1}_0(\Omega) \to V_h$ is defined by

$$\langle \nabla R_h u, \nabla \phi_h \rangle_{L^2(\Omega)} = \langle \nabla u, \nabla \phi_h \rangle_{L^2(\Omega)} \qquad \forall \phi_h \in V_h. \tag{2.3}$$

We comment that this mapping is the orthogonal projections onto V_h with respect to the $W^{1,2}_0(\Omega)$-seminorm. The following local error estimate for R_h can be found in [7, Theorem 1]. In fact, it holds for more general families of triangulations than quasiuniform ones.

Proposition 2.1 (local error estimate). *Let $w \in W^{1,\infty}_0(\Omega)$ and $\mathbb{T}$ be a quasiuniform family of triangulations of a polytype Ω. Let $z \in \Omega$ and $h > 0$. Define $D = \Omega \cap B_d(z)$ with $d \geq k_0 h$, where k_0 is sufficiently large. We have, for every $w_h \in V_h$,*

$$\begin{aligned} |\nabla(w - R_h w)(z)| &\lesssim \|\nabla(w - w_h)\|_{L^\infty(D)} + d^{-1}\|w - w_h\|_{L^\infty(D)} \\ &\quad + d^{-\frac{n}{2}-1}\|w - R_h w\|_{L^2(D)}, \end{aligned}$$

where the implicit constant is independent of w, h, and z.

Proof. As mentioned before, this is essentially [7, Theorem 1]. However, in that result, as stated, the point z is where $\|\nabla(w - R_h w)\|_{L^\infty(\Omega)}$ is attained. One merely needs to examine the proof to see that this point may be arbitrary. ∎

3. Statement of the main result and corollaries.

We are now in position to state the main result of our work.

Theorem 3.1 (pointwise estimate). *Let $\Omega \subset \mathbb{R}^n$, for $n \in \{2,3\}$, be a convex polytope and $\mathbb{T} = \{\mathcal{T}_h\}_{h>0}$ be a family of conforming and quasiuniform triangulations of Ω. For every $u \in W^{1,1}_0(\Omega)$ and almost every $z \in \Omega$ we have*

$$|\nabla R_h u(z)| \lesssim M[\nabla u](z), \tag{3.1}$$

where the implicit constant is independent of z, u, and h and depends on $\mathbb{T}$ only through its shape regularity constants.

Before we embark on the proof of this result, we immediately mention that it implies the stability of the Ritz projection in any space where the Hardy–Littlewood maximal operator is bounded. For the sake of completeness we present a far from exhaustive list of examples: (weighted) L^p spaces (see section 3.3), Lorentz spaces (see sections 3.1 and 3.4); Orlicz spaces (see section 3.2) and (weighted) variable exponent spaces (see section 3.5).

3.1. Lorentz spaces.

Let μ be a measure on Ω, $p \in [1, \infty)$, and $q \in [1, \infty]$. The Lorentz spaces are defined as

$$L^{p,q}(\mu, \Omega) = \left\{ f \in L^0(\mu, \Omega) : \|f\|_{L^{p,q}(\mu,\Omega)} < \infty \right\},$$

where

$$\|f\|_{L^{p,q}(\mu,\Omega)} = \begin{cases} \left(q\int_0^\infty t^q \mu_f(t)^{q/p} \frac{\mathrm{d}t}{t}\right)^{1/q}, & q<\infty, \\ \sup_{t>0} t\mu_f(t)^{1/p}, & q=\infty, \end{cases} \tag{3.2}$$

and

$$\mu_f(t) = \mu\left(\{x\in\Omega : |f(x)|>t\}\right)$$

is the distribution function of f. We recall that, for $p\in[1,\infty)$, $L^{p,p}(\mu,\Omega)=L^p(\mu,\Omega)$ with equivalence of norms [12, Proposition 1.4.5]. Finally, if μ is the Lebesgue measure, we simply denote these spaces by $L^{p,q}(\Omega)$.

COROLLARY 3.2 (Lorentz stability). *In the setting of Theorem* 3.1 *assume, in addition, that* $p\in(1,\infty)$ *and* $q\in(1,\infty]$*, or that* $p=1$ *and* $q=\infty$*. Then we have*

$$\|\nabla R_h u\|_{L^{p,q}(\Omega)} \lesssim \|\nabla u\|_{L^{p,q}(\Omega)},$$

where the implicit constant is independent of u *and* h*. In particular, for* $p\in(1,\infty]$*, we have*

$$\|\nabla R_h u\|_{L^{p}(\Omega)} \lesssim \|\nabla u\|_{L^{p}(\Omega)}.$$

Proof. Consider first the case $p=1$ and $q=\infty$. Owing to, for instance, [12, Theorem 2.1.6] we have $M: L^{1,\infty}\to L^{1,\infty}$ boundedly.

For $p>1$ it suffices to invoke [17, Theorems A, section 5.2]. ∎

We comment that the boundedness of the Ritz projection in $W^{1,p}$ spaces has already been presented in [23, 15, 7]. Thus, the case $p\in(1,\infty)$ of Corollary 3.2 can also be obtained by the Marcinkiewicz interpolation theorem as presented in [2, Theorem 5.3.2].

3.2. Orlicz spaces. Another new result is stability in Orlicz spaces. We say that $\varphi:(0,\infty)\to(0,\infty)$ is an Orlicz function if it is nonnegative and increasing and

$$\varphi(0+) = \lim_{t\downarrow 0}\varphi(t) = 0, \qquad \varphi(\infty) = \lim_{t\to\infty}\varphi(t)=\infty.$$

If φ is an Orlicz function and, in addition, it is convex and satisfies

$$\lim_{t\downarrow 0}\frac{\varphi(t)}{t} = \lim_{t\to\infty}\frac{t}{\varphi(t)} = 0,$$

then we say that it is an N-function.

For an N-function φ, we define its corresponding Orlicz space as

$$L^\varphi(\Omega) = \left\{ f\in L^0(\Omega) : \|f\|_{L^\varphi(\Omega)}<\infty\right\},$$
$$\|f\|_{L^\varphi(\Omega)} = \inf_{\lambda>0}\left\{ \int_\Omega \varphi\left(\frac{1}{\lambda}|f(x)|\right)\mathrm{d}x \le 1\right\}.$$

We refer the reader to [18] for further properties of such spaces.

Given an N-function φ, we say that $\varphi\in\nabla_2$ if there exists $a>1$ such that

$$\varphi(t)\le \frac{1}{2a}\varphi(at) \qquad \forall t\ge 0.$$

Corollary 3.3 (Orlicz stability). *In the setting of Theorem* 3.1 *let* $\varphi \in \nabla_2$. *Then*

$$\|\nabla R_h u\|_{L^\varphi(\Omega)} \lesssim \|\nabla u\|_{L^\varphi(\Omega)},$$

where the implicit constant is independent of u *and* h.

Proof. According to [17, Theorem 1.2.1(v)], if $\varphi \in \nabla_2$, then the maximal function is bounded on $L^\varphi(\Omega)$. Apply Theorem 3.1 to conclude. ∎

Remark 3.4 (Simonenko indices). Given an N-function φ define

$$h_\varphi(\lambda) = \sup_{t>0} \frac{\varphi(\lambda t)}{\varphi(t)}, \qquad \lambda > 0.$$

The upper and lower Simonenko indices of φ are, respectively,

$$p_\varphi^- = \lim_{\lambda \downarrow 0} \frac{\log h_\varphi(\lambda)}{\log \lambda}, \qquad p_\varphi^+ = \lim_{\lambda \to \infty} \frac{\log h_\varphi(\lambda)}{\log \lambda}.$$

We comment that $\varphi \in \nabla_2$ implies $p_\varphi^- > 1$ so that the condition in Corollary 3.3 is consistent with the results of Corollary 3.2.

On the other hand, we say that an N-function is power-like if $p_\varphi^+ < \infty$. According to [26] (see also [16]), the space $L^\varphi(\Omega)$ is an intermediate space between $L^p(\Omega)$ and $L^q(\Omega)$ provided the Simonenko indices satisfy

$$1 \le p \le p_\varphi^- \le p_\varphi^+ \le q \le \infty.$$

Thus, in the case of $\varphi \in \nabla_2$ and power-like, the results of Corollary 3.3 could be obtained by interpolation. Since, however, we are not assuming $p_\varphi^+ < \infty$, this is truly a new result.

3.3. Muckenhoupt weighted spaces. Next we extend the results of [10] to the optimal range of indices. We recall that a function $0 \le \omega \in L^1_{\mathrm{loc}}(\Omega)$ is called a weight. For $p \in [1,\infty)$ we say that a weight ω belongs to the Muckenhoupt class $\mathcal{A}_p$ if

$$[\omega]_{\mathcal{A}_p} = \begin{cases} \sup_Q \left(\frac{1}{|Q|} \int_Q \omega(x) \,\mathrm{d}x \right) \left(\frac{1}{|Q|} \int_Q \omega^{-\frac{1}{p-1}}(x) \,\mathrm{d}x \right)^{p-1}, & p > 1, \\ \sup_Q \left(\frac{1}{|Q|} \int_Q \omega(x) \,\mathrm{d}x \right) \|\omega^{-1}\|_{L^\infty(Q)}, & p = 1, \end{cases}$$

where the supremum is over all cubes $Q \subseteq \Omega$ with sides parallel to the coordinate axes. Weighted Lebesgue spaces are defined, for $p \in (1,\infty)$ and $\omega \in \mathcal{A}_p$, as

$$L^p(\omega,\Omega) = \left\{ f \in L^0(\Omega) : \|f\|_{L^p(\omega,\Omega)} < \infty \right\},$$
$$\|f\|_{L^p(\omega,\Omega)} = \left[\int_\Omega |f(x)|^p \omega(x) \,\mathrm{d}x \right]^{1/p}.$$

Corollary 3.5 (weighted stability). *Under the assumptions of Theorem* 3.1 *let* $p \in (1,\infty)$ *and* $\omega \in \mathcal{A}_p$. *Then,*

$$\|\nabla R_h u\|_{L^p(\omega,\Omega)} \lesssim \|\nabla u\|_{L^p(\omega,\Omega)},$$

where the implicit constant is independent of u *and* h.

Proof. It suffices to recall that, provided $\omega \in \mathcal{A}_p$, the Hardy–Littlewood maximal operator is bounded on weighted spaces; see [12, Theorem 7.1.9(b)]. ∎

As we mentioned above, this result generalizes [10, Corollary 3.3], where such an estimate is obtained, but with $\omega \in \mathcal{A}_{p/2}$, a strictly smaller class.

3.4. Weighted Lorentz spaces. Let ω be a weight. Here we are concerned with weighted Lorentz spaces $L^{p,q}(\omega,\Omega)$; i.e., the measure in (3.2) is $\mu = \omega\,\mathrm{d}x$.

COROLLARY 3.6 (weighted stability). *In the setting of Theorem* 3.1 *let* $p \in (1,\infty)$*,* $q \in (1,\infty]$*, and* $\omega \in \mathcal{A}_p$*. Then*

$$\|\nabla R_h u\|_{L^{p,q}(\omega,\Omega)} \lesssim \|\nabla u\|_{L^{p,q}(\omega,\Omega)},$$

where the implicit constant is independent of u *and* h*.*

Proof. According to [17, Theorem 5.2.1], given the range of exponents, we have that $M : L^{p,q}(\omega,\Omega) \to L^{p,q}(\omega,\Omega)$ boundedly if $\omega \in \mathcal{A}_p$. ∎

3.5. Weighted variable exponent spaces. As a final application we mention weighted variable exponent spaces. A variable exponent is $p \in L^0(\Omega)$ such that $p(\Omega) \subset [1,\infty]$. Given a variable exponent and a weight $0 \le \omega \in L^1_{\mathrm{loc}}(\Omega)$ we define weighted variable exponent Lebesgue spaces as

$$L^{p(\cdot)}_\omega(\Omega) = \left\{ f \in L^0(\Omega) : \|f\|_{L^{p(\cdot)}_\omega(\Omega)} < \infty \right\},$$
$$\|f\|_{L^{p(\cdot)}_\omega(\Omega)} = \inf_{\lambda>0} \left\{ \int_\Omega \left| \frac{1}{\lambda} f(x)\omega(x) \right|^{p(x)} \mathrm{d}x \le 1 \right\}.$$

We refer the reader to [6, 9] for an extensive treatise on these spaces.

Given a variable exponent p, we say that $p \in \mathcal{P}^{\log}(\Omega)$ if

$$\left| \frac{1}{p(x)} - \frac{1}{p(y)} \right| \lesssim \frac{1}{\log(\mathrm{e} + 1/|x-y|)} \quad \forall x, y \in \Omega,$$

and, there is $p_\infty \ge 1$ such that

$$\left| \frac{1}{p(x)} - \frac{1}{p_\infty} \right| \lesssim \frac{1}{\log(\mathrm{e} + 1/|x|)} \quad \forall x \in \Omega.$$

If p is a variable exponent, then p' is its Hölder conjugate, that is, the variable exponent that satisfies

$$\frac{1}{p(x)} + \frac{1}{p'(x)} = 1$$

for almost every $x \in \Omega$. We say that the weight ω satisfies the generalized Muckenhoupt condition, denoted by $\omega \in \mathcal{A}$, if

$$\|\chi_Q\|_{L^{p(\cdot)}_\omega(\Omega)} \|\chi_Q\|_{L^{p'(\cdot)}_{\omega^{-1}}(\Omega)} \approx |Q|$$

for every cube Q with sides parallel to the coordinate axes. Here χ_Q is the characteristic function of Q.

Remark 3.7 ($\mathcal{A}$ versus $\mathcal{A}_p$). If $p(x) = p \in (1,\infty)$ for all $x \in \Omega$, then it is known that

$$\|f\|^p_{L^{p(\cdot)}_\omega(\Omega)} = \int_\Omega |f(x)\omega(x)|^p \,\mathrm{d}x = \|f\omega\|_{L^p(\Omega)} = \|f\|_{L^p(\omega^p,\Omega)}.$$

Thus, we see that $\omega \in \mathcal{A}$ is equivalent to $\mu = \omega^p \in \mathcal{A}_p$.

COROLLARY 3.8 (variable exponent stability). *Under the assumptions of Theorem* 3.1 *let* $p \in \mathcal{P}^{\log}(\Omega)$ *with* $\operatorname{ess\,inf}_{x\in\Omega} p(x) > 1$, *and* $\omega \in \mathcal{A}$. *Then*

$$\|\nabla R_h u\|_{L^{p(\cdot)}_\omega(\Omega)} \lesssim \|\nabla u\|_{L^{p(\cdot)}_\omega(\Omega)},$$

where the implicit constant is independent of u *and* h.

Proof. Under the given assumptions the Hardy–Littlewood maximal operator is bounded on $L^{p(\cdot)}_\omega(\Omega)$; see [4], [9, Theorem 4.3.8], and [9, Theorem 5.8.6]. ∎

3.6. Other extensions and variations. As we mentioned after Theorem 3.1, the list we have provided is not exhaustive. For instance, under certain conditions, one can also assert the boundedness in Orlicz–Musielak spaces [5].

On the other hand, there are some spaces where the stability remains open. Notable examples are $\mathcal{H}^1(\Omega)$, the atomic Hardy space, and $L^1(\Omega)$.

4. Proof of the main result. We now focus on the proof of Theorem 3.1. The technique that we shall follow will be a combination of weighted norm inequalities, as in [23], and local estimates, as presented in [7]. We shall also rely on some estimates on the Green's function that hold, for $n \in \{2,3\}$, in convex polytopes.

PROPOSITION 4.1 (Green's function estimates). *Let* $\Omega \subset \mathbb{R}^n$, *with* $n \in \{2,3\}$, *be a convex polytope and* $G : \overline{\Omega} \times \overline{\Omega} \to \mathbb{R}$ *be the Green's function associated to this domain. Then, for every* $i \in \{1,\dots,n\}$,

$$|\partial_{x_i} G(x,\xi)| \lesssim \frac{1}{|x-\xi|^{n-1}} \qquad \forall x,\xi \in \overline{\Omega}.$$

In addition, there is $\alpha \in (0,1]$, *depending only on the inner angles of* Ω, *such that, for every* $i,j \in \{1,\dots,n\}$ *and all* $x,y,\xi \in \Omega$, *we have*

$$\frac{|\partial_{x_i} G(x,\xi) - \partial_{y_i} G(y,\xi)|}{|x-y|^\alpha} \lesssim |x-\xi|^{-n-\alpha+1} + |y-\xi|^{-n-\alpha+1},$$
$$\frac{\left|\partial_{x_i}\partial_{\xi_j} G(x,\xi) - \partial_{y_i}\partial_{\xi_j} G(y,\xi)\right|}{|x-y|^\alpha} \lesssim |x-\xi|^{-n-\alpha} + |y-\xi|^{-n-\alpha}.$$

Proof. The first bound can be found in [14, Theorem 3.3(iv)] for $n \geq 3$ and [11, Proposition 1 (9)] for $n = 2$.

In the case $n = 3$, the Hölder estimates on the first and mixed derivatives can be found in [15, formula (1.4)]. When $n = 2$ [10, Lemma 2.1] presents a proof of the estimate for the mixed derivative. The estimate on the first derivative follows the same proof presented in [15]. ∎

Notice that the Hölder estimates on derivatives of G are the only instances where our dimensional restriction plays a role. As soon as the estimates in Proposition 4.1 are valid for more dimensions, the proof of Theorem 3.1 follows verbatim.

4.1. Approximation of identity. The technique of weighted norms [20, 21] relies on the construction of a regularized distance function and its properties. Here we rephrase some of the properties of such a function that may help elucidate the reason for its use. For $K, \gamma > 0$ to be chosen we define $\varphi_1 : \mathbb{R}^n \to \mathbb{R}$ by

$$\varphi_1(x) = c_1 \left(|x|^2 + K^2\right)^{-\frac{n+\gamma}{2}}, \tag{4.1}$$

where c_1 is such that $\int_{\mathbb{R}^n} \varphi_1(x)\,\mathrm{d}x = 1$. Now, for $\epsilon > 0$ and $z \in \Omega$, we define

$$\varphi_\epsilon(x) = \epsilon^{-n}\varphi_1(x/\epsilon) = c_1\epsilon^{-n}\left(|x/\epsilon|^2 + K^2\right)^{-\frac{n+\gamma}{2}} = c_1\epsilon^{\gamma}\left(|x|^2 + K^2\epsilon^2\right)^{-\frac{n+\gamma}{2}},$$
$$\varphi_{\epsilon,z}(x) = \varphi_\epsilon(z - x).$$

Notice that the family $\{\varphi_\epsilon\}_{\epsilon>0}$ is an approximation of the identity.

LEMMA 4.2 (convolution estimate). *For every $\epsilon > 0$, $z \in \Omega$, and $f \in L^0(\Omega)$ we have*

$$\|\varphi_{\epsilon,z} f\|_{L^1(\Omega)} = (\varphi_\epsilon * |f|)(z) \lesssim M[f](z),$$

where the constant is independent of ϵ, z, and f. Here, f is extended by zero outside of Ω.

Proof. Since the function φ_1 is radial and decreasing, it suffices to invoke Theorem 2.2 of section 2.2 in [27]; see also [12, Theorem 2.1.10]. ∎

4.2. Regularized Green's function. To establish our main estimate we shall rely on a pointwise representation. We fix $h > 0$ and let $z \in \Omega$ be such that $z \in \mathring{T}$ for some $T \in \mathcal{T}_h$. Owing to shape regularity, see [19, formula (12)], [28, Lemma 2.2] and [25]; there is a function $\delta_z \in C_0^\infty(T)$ such that

$$\int_T \delta_z(x)P(x)\,\mathrm{d}x = P(z) \quad \forall P \in \mathbb{P}_k, \qquad \|D^m\delta_z\|_{L^\infty(\Omega)} \lesssim h^{-n-m}, \quad m \in \mathbb{N}_0.$$

Fix $l \in \{1, \dots, n\}$. The regularized Green's function is $g_z \in W_0^{1,2}(\Omega)$ such that

$$\langle \nabla g_z, \nabla v\rangle_{L^2(\Omega)} = \langle \delta_z, \partial_l v\rangle_{L^2(\Omega)} \quad \forall v \in W_0^{1,2}(\Omega). \tag{4.2}$$

Owing to the fact that the right-hand side in (4.2) is compactly supported in Ω, we can, using Proposition 4.1, obtain some Hölder regularity for g_z. This is the content of the following result.

PROPOSITION 4.3 (estimates on g_z). *Let $z \in \mathring{T} \in \mathcal{T}_h$ and g_z solve (4.2). Then, for every $i \in \{1, \dots, n\}$ and all $x, y \notin T$, $x \neq y$, we have*

$$\frac{|\partial_i g_z(x) - \partial_i g_z(y)|}{|x - y|^\alpha} \lesssim \max_{\xi\in T}\left(|x - \xi|^{-n-\alpha} + |y - \xi|^{-n-\alpha}\right),$$

where the exponent $\alpha \in (0, 1]$ is the same as in Proposition 4.1. Moreover,

$$\|\nabla g_z\|_{L^\infty(\Omega)} \lesssim h^{-n}.$$

Proof. We begin by using the pointwise representation of g_z in terms of the Green's function G, and the fact that δ_z is supported on T to obtain

$$\partial_i g_z(x) - \partial_i g_z(y) = -\int_T \left(\partial_{x_i} G(x, \xi) - \partial_{y_i} G(y, \xi)\right)\partial_l \delta_z(\xi)\,\mathrm{d}\xi.$$

We now invoke Proposition 4.1 to obtain

$$\begin{aligned}\frac{|\partial_i g_z(x) - \partial_i g_z(y)|}{|x - y|^\alpha} &\le \sup_{\xi\in T} \frac{|\partial_{\xi_l}\partial_{x_i} G(x, \xi) - \partial_{\xi_l}\partial_{y_i} G(y, \xi)|}{|x - y|^\alpha}\|\delta_z\|_{L^1(T)} \\ &\lesssim \max_{\xi\in T}\left(|x - \xi|^{-n-\alpha} + |y - \xi|^{-n-\alpha}\right),\end{aligned}$$

as claimed.

To obtain the second estimate we observe that δ_z is supported on T and use its scaling properties to assert that, for any $i \in \{1,\ldots,n\}$, we have

$$|\partial_i g_z(x)| = \left| \int_T \partial_{x_i} G(x,\xi) \partial_l \delta_z(\xi) \,\mathrm{d}\xi \right| \lesssim \int_T |x-\xi|^{1-n} h^{-n-1} \,\mathrm{d}\xi \lesssim h^{-n}.$$

All estimates have been proved. ∎

4.3. Step 1: Pointwise representation. We now begin with the proof of Theorem 3.1 per se. Owing to the properties of δ_z we have that

$$\begin{aligned}\partial_l R_h u(z) &= \langle \delta_z, \partial_l R_h u \rangle_{L^2(\Omega)} = \langle \nabla g_z, \nabla R_h u \rangle_{L^2(\Omega)} = \langle \nabla R_h g_z, \nabla u \rangle_{L^2(\Omega)} \\ &= \langle \delta_z, \partial_l u \rangle_{L^2(\Omega)} + \langle \nabla (R_h g_z - g_z), \nabla u \rangle_{L^2(\Omega)},\end{aligned}$$

where we used (4.2) and the definition of the Ritz projection (2.3). From the definition of δ_z it follows immediately that

$$\left| \langle \delta_z, \partial_l u \rangle_{L^2(\Omega)} \right| \lesssim M[\nabla u](z).$$

On the other hand, we estimate the second term as

$$\left| \langle \nabla (R_h g_z - g_z), \nabla u \rangle_{L^2(\Omega)} \right| \le \| \varphi_{h,z} \nabla u \|_{L^1(\Omega)} \left\| \varphi_{h,z}^{-1} \nabla (R_h g_z - g_z) \right\|_{L^\infty(\Omega)}.$$

Owing to Lemma 4.2,

$$\| \varphi_{h,z} \nabla u \|_{L^1(\Omega)} \lesssim M[\nabla u](z).$$

Thus, if we define

$$\mathcal{G}_h = \sup_{z\in\Omega} \left\| \varphi_{h,z}^{-1} \nabla (R_h g_z - g_z) \right\|_{L^\infty(\Omega)}, \tag{4.3}$$

we see that the heart of the matter is to provide a uniform, in h, estimate for this quantity.

In summary, the rest of the proof consists in showing the following result.

PROPOSITION 4.4 (uniform estimate). *In the setting of Theorem* 3.1 *there are* $K > k_0$ *and* $\gamma \in (0,\alpha)$ *such that, if* φ_1 *is defined as in* (4.1)*, we have*

$$\mathcal{G}_h \lesssim 1,$$

where the constant is independent of $h > 0$*, and* $\mathcal{G}_h$ *was defined in* (4.3)*. Here* k_0 *is as in Proposition* 2.1*, and* α *is as in Proposition* 4.1*.*

4.4. Step 2: Dyadic decomposition. Fix $z \in \Omega$. Define, for $j \in \mathbb{N}_0$, $d_j = 2^j K h$. We decompose the domain Ω into the following annuli:

$$
\begin{aligned}
B_j &= \{x \in \Omega : |x - z| < d_j\}, \qquad & A_j &= B_j \setminus B_{j-1}, \\
A_j^{+} &= B_{j+1} \setminus B_{j-2}, & A_j^{++} &= B_{j+2} \setminus B_{j-3},
\end{aligned} \tag{4.4}
$$

with the convention that, for $j < 0$, $B_j = \emptyset$. For $S \subset \overline{\Omega}$ we also define

$$
\mathcal{N}_h(S) = \bigcup \{T \in \mathcal{T}_h : S \cap T \neq \emptyset\}.
$$

The use of this dyadic decomposition lies in the fact that on each annulus the regularized distance function $\varphi_{h,z}$ is almost constant.

LEMMA 4.5 (distance estimates). *Assume that $K > 2$. For all $j \geq 0$ we have*

$$
\varphi_{h,z}(x) \approx h^{\gamma} d_j^{-n-\gamma} \qquad \forall x \in A_j,
$$

and

$$
\operatorname{dist}\left(A_j^{+}, \mathcal{N}_h(\Omega \setminus A_j^{++})\right) \approx d_j,
$$

where the implicit constants are independent of h. As a consequence, for $j \geq 3$ we have

$$
|\nabla g_z|_{C^{0,\alpha}(A_j^{++})} \lesssim d_j^{-n-\alpha},
$$

where $\alpha \in (0,1]$ is as in Proposition 4.1.

Proof. The estimate on $\varphi_{h,z}$ follows by definition. The estimate on the distance between A_j^{+} and $\mathcal{N}_h(\Omega \setminus A_j^{++})$ does so as well.

On the other hand, if $j \geq 3$, $x, y \in A_j^{++}$, and $\xi \in T$, then $|x - \xi|, |y - \xi| \approx d_j$. We can then refine the estimate of Proposition 4.3 to conclude

$$
\frac{|\partial_i g_z(x) - \partial_i g_z(y)|}{|x - y|^{\alpha}} \lesssim d_j^{-n-\alpha}. \qquad \square
$$

4.5. Step 3: Reduction to interpolation and duality. Now fix some $z \in \Omega$ and define

$$
\mathcal{G}_{h,z} = \left\| \varphi_{h,z}^{-1} \nabla (R_h g_z - g_z) \right\|_{L^{\infty}(\Omega)}.
$$

Let $\mathrm{j} \in \mathbb{N}_0$ now be such that

$$
\mathcal{G}_{h,z} = \left\| \varphi_{h,z}^{-1} \nabla (g_z - R_h g_z) \right\|_{L^{\infty}(A_\mathrm{j})}.
$$

Using the distance estimates of Lemma 4.5, we can also assert that

$$
\mathcal{G}_{h,z} \lesssim h^{-\gamma} d_\mathrm{j}^{n+\gamma} \|\nabla (g_z - R_h g_z)\|_{L^{\infty}(A_\mathrm{j})}.
$$

Now choose $K \geq k_0$, where k_0 was introduced in Proposition 2.1. Then we have $\operatorname{diam} A_\mathrm{j} \approx d_j \geq k_0 h$, so that with a simple covering argument we may obtain that

$$
\begin{aligned}
\mathcal{G}_{h,z} &\lesssim h^{-\gamma} d_\mathrm{j}^{n+\gamma} \|\nabla (g_z - R_h g_z)\|_{L^{\infty}(A_\mathrm{j})} \\
&\lesssim h^{-\gamma} d_\mathrm{j}^{n+\gamma} \Big(\|\nabla (g_z - \Pi_h g_z)\|_{L^{\infty}(A_\mathrm{j}^{+})} + d_\mathrm{j}^{-1} \|g_z - \Pi_h g_z\|_{L^{\infty}(A_\mathrm{j}^{+})} \\
&\qquad + d_\mathrm{j}^{-\frac{n}{2}-1} \|g_z - R_h g_z\|_{L^{2}(A_\mathrm{j}^{+})} \Big) = \mathrm{I} + \mathrm{II} + \mathrm{III},
\end{aligned} \tag{4.5}
$$

where Π_h is, for instance, the so-called Scott–Zhang interpolant [24], or any other interpolant satisfying local stability and approximation properties. The first two terms will be handled using interpolation estimates, whereas the last one is controlled by duality.

4.6. Step 4: Bound of $\mathrm{I}+\mathrm{II}$ via interpolation estimates. It is our goal now to bound $\mathrm{I}+\mathrm{II}$ using the approximation properties of Π_h and the regularity of g_z. This regularity, however, depends on the distance between A_{j}^+ and z. If $\mathrm{j}\geq 3$, then we can invoke the estimate in Lemma 4.5 to see that

$$\begin{aligned}\|g_z-\Pi_h g_z\|_{L^\infty(A_{\mathrm{j}}^+)}+h\|\nabla(g_z-\Pi_h g_z)\|_{L^\infty(A_{\mathrm{j}}^+)} &\lesssim h^{1+\alpha}|\nabla g_z|_{C^{0,\alpha}(A_{\mathrm{j}}^{++})}\\ &\lesssim h^{1+\alpha}d_{\mathrm{j}}^{-n-\alpha}.\end{aligned}$$

As a consequence, since $0<\gamma<\alpha$,

$$\begin{aligned}\mathrm{I}+\mathrm{II} &\lesssim h^{-\gamma}d_{\mathrm{j}}^{n+\gamma}\left(h^\alpha d_{\mathrm{j}}^{-n-\alpha}+h^{1+\alpha}d_{\mathrm{j}}^{-n-\alpha-1}\right)\lesssim\left(\frac{h}{d_{\mathrm{j}}}\right)^{\alpha-\gamma}+\left(\frac{h}{d_{\mathrm{j}}}\right)^{1+\alpha-\gamma}\\ &\leq\frac{1}{K^{\alpha-\gamma}}+\frac{1}{K^{1+\alpha-\gamma}}.\end{aligned}\tag{4.6}$$

If, on the other hand, $\mathrm{j}<3$, we use the second bound of Proposition 4.3 to obtain

$$\|g_z-\Pi_h g_z\|_{L^\infty(A_{\mathrm{j}}^+)}+h\|\nabla(g_z-\Pi_h g_z)\|_{L^\infty(A_{\mathrm{j}}^+)}\lesssim h\|\nabla g_z\|_{L^\infty(\Omega)}\lesssim h^{1-n}.$$

In this case then we get

$$\begin{aligned}\mathrm{I}+\mathrm{II} &\lesssim h^{-\gamma}d_{\mathrm{j}}^{n+\gamma}\left(h^{-n}+d_{\mathrm{j}}^{-1}h^{1-n}\right)\lesssim h^{-n-\gamma}d_3^{n+\gamma}+h^{1-n-\gamma}d_3^{n+\gamma-1}\\ &\lesssim K^{n+\gamma}+K^{n+\gamma-1}.\end{aligned}\tag{4.7}$$

Gathering (4.6) and (4.7), we arrive at

$$\mathrm{I}+\mathrm{II}\lesssim\max\left\{\frac{1}{K^{\alpha-\gamma}}+\frac{1}{K^{1+\alpha-\gamma}},K^{n+\gamma}+K^{n+\gamma-1}\right\}.\tag{4.8}$$

4.7. Step 5: Bound of III by duality. We bound III by duality. Define

$$\mathcal{S}_{\mathrm{j}}=\left\{v\in C_0^\infty(\Omega):\|v\|_{L^2(\Omega)}\leq 1,\ \operatorname{supp}(v)\subset A_j^+\right\}\tag{4.9}$$

so that

$$\|g_z-R_h g_z\|_{L^2(A_{\mathrm{j}}^+)}=\sup_{0\neq v\in\mathcal{S}_{\mathrm{j}}}\langle g_z-R_h g_z,v\rangle_{L^2(\Omega)}.$$

Fix $v\in\mathcal{S}_{\mathrm{j}}$ and let $w_v\in W_0^{1,2}(\Omega)$ solve

$$-\Delta w_v=v,\ \text{in }\Omega,\qquad w_v=0,\ \text{on }\partial\Omega.\tag{4.10}$$

Then, by Galerkin orthogonality,

$$\begin{aligned}\langle g_z-R_h g_z,v\rangle_{L^2(\Omega)} &=\langle\nabla(g_z-R_h g_z),\nabla w_v\rangle_{L^2(\Omega)}\\ &=\langle\nabla(g_z-R_h g_z),\nabla(w_v-\Pi_h w_v)\rangle_{L^2(\Omega)}\\ &=\langle\varphi_{h,z}^{-1}\nabla(g_z-R_h g_z),\varphi_{h,z}\nabla(w_v-\Pi_h w_v)\rangle_{L^2(\Omega)}.\end{aligned}$$

An application of Hölder's inequality then allows us to conclude that

$$\mathrm{III}\leq h^{-\gamma}d_{\mathrm{j}}^{\frac{n}{2}+\gamma-1}\sup_{v\in\mathcal{S}_j}\|\varphi_{h,z}\nabla(w_v-\Pi_h w_v)\|_{L^1(\Omega)}\mathcal{G}_h.$$

Notice that if, in this last estimate, the term that is multiplying $\mathcal{G}_h$ is sufficiently small, then it could be absorbed on the left-hand side in (4.5). This possibility is explored in the following result.

LEMMA 4.6 (duality bound). *Let $\mathcal{S}_{\mathfrak{j}}$ be defined as in* (4.9) *and* $\gamma \in (0,\alpha)$. *There is a constant, independent of* $\mathfrak{j}$, z, *and* h, *such that*

$$h^{-\gamma} d_{\mathfrak{j}}^{\frac{n}{2}+\gamma-1} \sup_{v\in\mathcal{S}_{\mathfrak{j}}} \|\varphi_{h,z}\nabla(w_v - \Pi_h w_v)\|_{L^1(\Omega)} \le C\left(\frac{1}{K} + \frac{1}{K^{\alpha-\gamma}}\right),$$

where $w_v \in W_0^{1,2}(\Omega)$ *is the solution to* (4.10) *and* α *is as in Proposition* 4.1.

Proof. Let $v \in \mathcal{S}_{\mathfrak{j}}$ be arbitrary. Using Lemma 4.5 and scaling, we have

$$\begin{aligned} h^{-\gamma} d_{\mathfrak{j}}^{\frac{n}{2}+\gamma-1} \|\varphi_{h,z}\nabla(w_v - \Pi_h w_v)\|_{L^1(A_{\mathfrak{j}}^{++})} &\lesssim d_{\mathfrak{j}}^{-\frac{n}{2}-1} \|\nabla(w_v - \Pi_h w_v)\|_{L^1(A_{\mathfrak{j}}^{++})} \\ &\le d_{\mathfrak{j}}^{-1} \|\nabla(w_v - \Pi_h w_v)\|_{L^2(A_{\mathfrak{j}}^{++})} \\ &\lesssim \frac{h}{d_{\mathfrak{j}}} |w_v|_{W^{2,2}(\Omega)} \lesssim \frac{1}{K}\|v\|_{L^2(\Omega)} \le \frac{1}{K}, \end{aligned}$$

where, since Ω is convex, we used a regularity estimate on w_v.

To control the norm in $\Omega \setminus A_{\mathfrak{j}}^{++}$ observe that, owing to the estimates of Proposition 4.1, for every $i \in \{1,\dots,n\}$ we have for every $x, y \in \Omega \setminus A_{\mathfrak{j}}^{++}$ that

$$\begin{aligned} \frac{|\partial_i w_v(x) - \partial_i w_v(y)|}{|x-y|^\alpha} &\le \int_{A_{\mathfrak{j}}^{+}} \frac{|\partial_{x_i} G(x,\xi) - \partial_{y_i} G(y,\xi)|}{|x-y|^\alpha} |v(\xi)|\,\mathrm{d}\xi \\ &\lesssim \max_{\xi\in A_{\mathfrak{j}}^{+}} \left(|x-\xi|^{-n-\alpha+1} + |y-\xi|^{-n-\alpha+1}\right) \int_{A_{\mathfrak{j}}^{+}} |v(\xi)|\,\mathrm{d}\xi \\ &\lesssim d_{\mathfrak{j}}^{-n-\alpha+1} d_{\mathfrak{j}}^{\frac{n}{2}} \|v\|_{L^2(A_{\mathfrak{j}}^{+})} \le d_{\mathfrak{j}}^{1-\alpha-\frac{n}{2}}, \end{aligned}$$

where we used the second distance estimate of Lemma 4.5. This shows that

$$|\nabla w_v|_{C^{0,\alpha}(\mathcal{N}_h(\Omega\setminus A_{\mathfrak{j}}^{++}))} \lesssim d_{\mathfrak{j}}^{1-\alpha-\frac{n}{2}}.$$

To shorten notation let $e_w = w_v - \Pi_h w_v$. We use that $\|\varphi_{h,z}\|_{L^1(\Omega)} = 1$ and the recently obtained regularity estimate to proceed as follows:

$$\begin{aligned} h^{-\gamma} d_{\mathfrak{j}}^{\frac{n}{2}+\gamma-1} \|\varphi_{h,z}\nabla e_w\|_{L^1(\Omega\setminus A_{\mathfrak{j}}^{++})} &\le h^{-\gamma} d_{\mathfrak{j}}^{\frac{n}{2}+\gamma-1} \|\nabla e_w\|_{L^\infty(\Omega\setminus A_{\mathfrak{j}}^{++})} \\ &\le h^{-\gamma} d_{\mathfrak{j}}^{\frac{n}{2}+\gamma-1} h^\alpha d_{\mathfrak{j}}^{1-\alpha-\frac{n}{2}} = \left(\frac{h}{d_{\mathfrak{j}}}\right)^{\alpha-\gamma} \le \frac{1}{K^{\alpha-\gamma}}. \end{aligned}$$

We combine both bounds to conclude. ∎

With Lemma 4.6 at hand we conclude that

$$\mathrm{III} \lesssim \left(\frac{1}{K} + \frac{1}{K^{\alpha-\gamma}}\right)\mathcal{G}_h. \tag{4.11}$$

4.8. Step 6: Final step. Gathering all the estimates. With the aid of (4.8) and (4.11), estimate (4.5) reduces to

$$\mathcal{G}_{h,z} \lesssim \max\left\{\frac{1}{K^{\alpha-\gamma}} + \frac{1}{K^{1+\alpha-\gamma}}, K^{n+\gamma} + K^{n+\gamma-1}\right\} + \left(\frac{1}{K} + \frac{1}{K^{\alpha-\gamma}}\right)\mathcal{G}_{h,z},$$

provided $K \ge k_0$, where k_0 is defined as in Proposition 2.1; and $\gamma \in (0,\alpha)$, with α as in Proposition 4.1. Notice the hidden constant here is independent of z. Since $\mathcal{G}_h = \sup_{z\in\Omega} \mathcal{G}_{h,z}$, we can now, if necessary, choose an even bigger K to conclude the proof of Proposition 4.4 and, as a consequence, that of Theorem 3.1.

REFERENCES

[1] C. Bennett, R. A. DeVore, and R. Sharpley, *Weak-L^∞ and BMO*, Ann. of Math. (2), 113 (1981), pp. 601–611, https://doi.org/10.2307/2006999.

[2] J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction*, Grundlehren Math. Wiss. 223, Springer-Verlag, Berlin, New York, 1976.

[3] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics in Applied Mathematics 40, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002, http://dx.doi.org/10.1137/1.9780898719208. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174].

[4] D. Cruz-Uribe, L. Diening, and P. Hästö, *The maximal operator on weighted variable Lebesgue spaces*, Fract. Calc. Appl. Anal., 14 (2011), pp. 361–374, https://doi.org/10.2478/s13540-011-0023-7.

[5] D. Cruz-Uribe and P. Hästö, *Extrapolation and interpolation in generalized Orlicz spaces*, Trans. Amer. Math. Soc., 370 (2018), pp. 4323–4349, https://doi.org/10.1090/tran/7155.

[6] D. V. Cruz-Uribe and A. Fiorenza, *Variable Lebesgue Spaces: Foundations and Harmonic Analysis*, Appl. Numer. Harmon. Anal., Birkhäuser/ Springer, Heidelberg, 2013, https://doi.org/10.1007/978-3-0348-0548-3.

[7] A. Demlow, D. Leykekhman, A. H. Schatz, and L. B. Wahlbin, *Best approximation property in the W^1_∞ norm for finite element methods on graded meshes*, Math. Comp., 81 (2012), pp. 743–764, https://doi.org/10.1090/S0025-5718-2011-02546-9.

[8] L. Diening and F. Ettwein, *Fractional estimates for non-differentiable elliptic systems with general growth*, Forum Math., 20 (2008), pp. 523–556, https://doi.org/10.1515/FORUM.2008.027.

[9] L. Diening, P. Harjulehto, P. Hästö, and M. Růžička, *Lebesgue and Sobolev Spaces with Variable Exponents*, Lecture Notes in Math. 2017, Springer, Berlin, Heidelberg, 2011, https://doi.org/10.1007/978-3-642-18363-8.

[10] I. Drelichman, R. G. Durán, and I. Ojea, *A weighted setting for the numerical approximation of the Poisson problem with singular sources*, SIAM J. Numer. Anal., 58 (2020), pp. 590–606, https://doi.org/10.1137/18M1213105.

[11] S. J. Fromm, *Potential space estimates for Green potentials in convex domains*, Proc. Amer. Math. Soc., 119 (1993), pp. 225–233, https://doi.org/10.2307/2159846.

[12] L. Grafakos, *Classical Fourier Analysis*, 3rd ed., Grad. Texts in Math. 249, Springer, New York, 2014, https://doi.org/10.1007/978-1-4939-1194-3.

[13] L. Grafakos, *Modern Fourier Analysis*, 3rd ed., Grad. Texts in Math. 250, Springer, New York, 2014, https://doi.org/10.1007/978-1-4939-1230-8.

[14] M. Grüter and K.-O. Widman, *The Green function for uniformly elliptic equations*, Manuscripta Math., 37 (1982), pp. 303–342, https://doi.org/10.1007/BF01166225.

[15] J. Guzmán, D. Leykekhman, J. Rossmann, and A. H. Schatz, *Hölder estimates for Green's functions on convex polyhedral domains and their applications to finite element methods*, Numer. Math., 112 (2009), pp. 221–243, https://doi.org/10.1007/s00211-009-0213-y.

[16] A. Y. Karlovich and L. Maligranda, *On the interpolation constant for Orlicz spaces*, Proc. Amer. Math. Soc., 129 (2001), pp. 2727–2739, https://doi.org/10.1090/S0002-9939-01-06162-7.

[17] V. Kokilashvili and M. Krbec, *Weighted Inequalities in Lorentz and Orlicz Spaces*, World Scientific Publishing, River Edge, NJ, 1991, https://doi.org/10.1142/9789814360302.

[18] M. A. Krasnoselskii and J. B. Rutickii, *Convex Functions and Orlicz Spaces*, P. Noordhoff, Groningen, 1961; translated from the first Russian edition by Leo F. Boron.

[19] D. Leykekhman and B. Vexler, *Finite element pointwise results on convex polyhedral-domains*, SIAM J. Numer. Anal., 54 (2016), pp. 561–587, http://dx.doi.org/10.1137/15M1013912.

[20] F. Natterer, *Über die punktweise Konvergenz finiter Elemente*, Numer. Math., 76 (1975), pp. 67–77, https://doi.org/10.1007/BF01419529.

[21] J. A. Nitsche, *L_∞-convergence of finite element approximation*, in Journées "Éléments Finis" (Rennes, 1975), Univ. Rennes, Rennes, France, 1975, 18.

[22] 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), pp. 733–791, https://doi.org/10.1007/s10208-014-9208-x.

[23] R. Rannacher and R. Scott, *Some optimal error estimates for piecewise linear finite element approximations*, Math. Comp., 38 (1982), pp. 437–445, https://doi.org/10.2307/2007280.

[24] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493, https://doi.org/10.2307/2008497.

[25] R. Scott, *Finite element convergence for singular data*, Numer. Math., 21 (1973/74), pp. 317–327, http://dx.doi.org/10.1007/BF01436386.
[26] I. B. Simonenko, *Interpolation and extrapolation of linear operators in Orlicz spaces*, Mat. Sb. (N.S.), 63 (1964), pp. 536–553.
[27] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton Math. Ser. 30, Princeton University Press, Princeton, NJ, 1970.
[28] V. Thomée and L. B. Wahlbin, *Stability and analyticity in maximum-norm for simplicial Lagrange finite element semidiscretizations of parabolic equations with Dirichlet-boundary conditions*, Numer. Math., 87 (2000), pp. 373–389, http://dx.doi.org/10.1007/s002110000184.
[29] A. Torchinsky, *Real-Variable Methods in Harmonic Analysis*, reprint of the 1986 original, Dover, Mineola, NY, 2004.