OPTIMAL CONVERGENCE RATES FOR THE FINITE ELEMENT APPROXIMATION OF THE SOBOLEV CONSTANT

LIVIU I. IGNAT AND ENRIQUE ZUAZUA

ABSTRACT. We establish optimal convergence rates for the $P1$ finite element approximation of the Sobolev constant in arbitrary dimensions $N \geq 2$ and for Lebesgue exponents $1 < p < N$. Our analysis relies on a refined study of the Sobolev deficit in suitable quasi-norms, which have been introduced and utilized in the context of finite element approximations of the p-Laplacian. The proof further involves sharp estimates for the finite element approximation of Sobolev minimizers.

1. INTRODUCTION

This paper is devoted to the study of the $P1$ finite element approximation of the Sobolev constant

$$S(p,N) = \inf_{u \in \dot{W}^{1,p}(\mathbb{R}^N)} \frac{\|Du\|_{L^p(\mathbb{R}^N)}}{\|u\|_{L^{p^*}(\mathbb{R}^N)}}, \tag{1.1}$$

which ensures the validity of the optimal Sobolev inequality

$$\|u\|_{L^{p^*}(\mathbb{R}^N)} \leq S(p,N)\|Du\|_{L^p(\mathbb{R}^N)}, \ \forall u \in \dot{W}^{1,p}(\mathbb{R}^N), \tag{1.2}$$

in dimensions $N \geq 2$ and for exponents $1 \leq p < N$, where

$$p^* = \frac{Np}{N-p} \tag{1.3}$$

denotes the Sobolev conjugate of p.

When $1 < p < N$, the minimum constant $S(p,N)$ is attained on an $(N+2)$-dimensional manifold $\mathcal{M}$, see [2, 20]. The case $p = 1$, which involves distinct features and will not be addressed in this work, was treated in [13, 18], where the sharp constant was determined.

The Sobolev constant plays an important role in various areas including, of course, existence and regularity of solutions for nonlinear PDE.

For the purposes of numerical analysis, we restrict our attention to bounded domains. To simplify the presentation, we consider the unit ball $B \subset \mathbb{R}^N$ with $N \geq 2$, equipped with a finite element mesh of characteristic size h. Let V_h denote the corresponding finite-dimensional subspace of $W_0^{1,p}(B)$, consisting of P^1 finite element functions on B that are continuous, piecewise linear, and vanish on the boundary. These functions are extended by zero outside of B. A precise construction of V_h is provided in Section 2.

Sobolev's inequality still holds in $W_0^{1,p}(B)$, with the same sharp constant (see for example [1]), but, in this case, the infimum is well-known not to be achieved:

$$S(p,N) = \inf_{u \in W_0^{1,p}(B)} \frac{\|Du\|_{L^p(B)}}{\|u\|_{L^{p^*}(\mathbb{R}^N)}}. \tag{1.4}$$

1991 *Mathematics Subject Classification.* 46E35, 65N30.

Key words and phrases. Sobolev inequalities, Approximation and Stability, Finite element method, Quasi-Norms.

Obviously, the same inequality holds in the finite-element subspace V_h as well, with the following, slightly larger, minimal constant

$$S_h(p,N) := \min_{u_h\in V_h} \frac{\|Du_h\|_{L^p(\mathbb{R}^N)}}{\|u_h\|_{L^{p^*}(\mathbb{R}^N)}}. \tag{1.5}$$

Obviously

$$S_h(p,N) \geq S(p,N). \tag{1.6}$$

In view of the convergence properties of finite element methods one can also prove that

$$S_h(p,N) \to S(p,N), \text{ as } h \to 0. \tag{1.7}$$

Our goal here is to provide sharp convergence rates. This is relevant, of course, for the obtention of sharp finite element convergence rates for PDE solutions in which the continuous analysis relies on the fine use of the Sobolev inequality. This problem was already addressed by Antonietti and Pratelli [1] for $p=2$, $N=3$, who proved suboptimal convergence rates.

This kind of problems has been previously considered in a number of related contexts. In particular, the answer is well-known for Poincaré's inequality, related to the first eigenvalue of the Laplacian in a bounded domain Ω of $\mathbb{R}^N$, $N\geq 1$, defined as

$$\lambda_1(\Omega) = \inf_{u\in H_0^1(\Omega)} \frac{\int_\Omega |Du|^2dx}{\int_\Omega u^2dx}.$$

The corresponding finite element approximation is defined as

$$\lambda_{1h}(V_h) = \inf_{u_h\in V_h} \frac{\int_\Omega |Du_h|^2dx}{\int_\Omega u_h^2dx},$$

and it is well known to satisfy [7, Prop. 6.30, p. 315], [3, Section 8, p. 700]:

$$\lambda_{1h}(V_h) - \lambda_1(\Omega) \simeq d(\phi, V_h)^2 = \inf_{u_h\in V_h} \|\phi - u_h\|^2_{H^1(\Omega)} \sim h^2,$$

where $\phi \in H_0^1(\Omega)$ is the first eigenfunction associated with $\lambda_1(\Omega)$. In this case the optimal convergence rate is given by the H^1-distance $d(\phi, V_h)$, the distance form the first continuous eigenfunction (normalised to one in $H_0^1(\Omega)$). This distance turns out to be of order h given that ϕ belongs to $H_0^1(\Omega)$ for Ω smooth or convex.

Poincaré's inequality can be seen as the simplest instance of the type of problems addressed in this work: determining convergence rates for the finite element approximation of key constants in the functional analysis of PDEs. In the Poincaré setting, the minimum is attained by the first eigenfunction of the Laplacian, which is unique up to normalization in $H_0^1(\Omega)$, and the problem is of linear-quadratic nature.

In contrast, the Sobolev inequality presents a genuinely nonlinear character. In bounded domains, the infimum is not attained, although explicit families of minimizing sequences can be constructed via scaling from the exact minimizers in $\mathbb{R}^N$. Our main objective is to develop and combine the necessary tools to fully address this more intricate setting, while building a broader methodology that could also be applied to related problems, such as those involving Hardy constants (see [11]).

The main result of this paper is the following optimal converge rate:.

Theorem 1.1. *Let $N\geq 2$ and $1<p<N$ and V_h the space of $P1$ finite elements space in the unit ball B. Then*

$$S_h(p,N) - S(p,N) \simeq h^{\alpha(p,N)}, \tag{1.8}$$

where

$$\alpha(p,N) = \frac{2(N-p)}{N+p-2}. \tag{1.9}$$

We compare our results with the earlier work of Antonietti and Pratelli [1]. In the specific case $p = 2$ and $N = 3$, they established the estimate

$$h^\gamma \lesssim S_h(2,N) - S(2,N) \lesssim h^{1/3}, \tag{1.10}$$

for some exponent $\gamma > 2 \cdot (26)^2/3$. In contrast, we prove the sharp convergence rate $h^{2/3}$, which improves upon the upper bound $h^{1/3}$ obtained in [1]. Moreover, our result is shown to be optimal, as we also establish a matching lower bound.

Our proof, which also lays the foundation for a systematic methodology to address similar problems, relies on several novel ingredients:

- The first key ingredient in our proof is the Sobolev deficit
$$\delta(u) = \frac{\|Du\|_{L^p(\mathbb{R}^N)}}{\|u\|_{L^{p*}(\mathbb{R}^N)}} - S(p,N), \;\; u \in \dot{W}^{1,p}(\mathbb{R}^N).$$
Building on the analysis developed in [15], we adapt this concept to the $P1$ finite element setting. However, as we will show, this adaptation alone is insufficient, as it yields suboptimal convergence rates.
- Achieving sharp convergence rates requires the use of quasi-norms introduced in the context of finite element approximations of the p-Laplacian [4, 12]. These quasi-norms play a crucial role in capturing the nonlinear structure of the problem with the necessary accuracy.

Let us recall that in the finite element setting, when approximating solutions u of the continuous p-Laplacian problem, the error is not measured in the classical $W^{1,p}$-norm. This norm would require $W^{2,p}$ regularity of the solution, which is generally not available. Instead, the error is measured in quasi-norms specifically adapted to the function u, namely:

$$|u-v|_{p,2} = \left(\int (|Du| + |D(u-v)|)^{p-2} |D(u-v)|^2 dx \right)^{1/2}.$$

This tool, introduced in [4, 12], allows to handle the degeneracy of the p-Laplacian, in order to obtain sharp error bounds.

The key idea underlying the proof of Theorem 1.1 is to derive optimal convergence rates by exploiting a form of Taylor expansion for the Sobolev deficit:

$$\delta(u) \simeq f(d(u,\mathcal{M})), \quad \forall\, u \text{ with } \delta(u) << 1,\ \|Du\|_{L^p(\mathbb{R}^N)} = 1, \tag{1.11}$$

where f is a suitable function and d denotes a suitable distance to the manifold $\mathcal{M}$.

The classical approach (see [15]) in studying the deficit is to obtain a lower bound of the type

$$\delta(u) \gtrsim f(d(u,\mathcal{M})), \tag{1.12}$$

for $d(u,v) = \|D(u-v)\|_{L^p(\mathbb{R}^N)}$ and $f(s) = s^{\max\{2,p\}}$. However, as we shall see below, when proceeding this way, the obtained upper counterpart is of order $\min\{2,p\}$. The obtention of sharp bilateral bounds of the same order requires to replace the Sobolev distance by the quasi-norms above. In fact, in [14], for the particular case $p \geq 2$, a variant of the quasi-norm above is also employed.

One of the main contributions of the present paper is to carefully adapt these methodologies to the distance defined by the quasi-norm: $d(u,v) = |u-v|_{p,2}$ (see Lemma 4.1). This together with the upper bound in Lemma 4.3 gives us an estimate of the type (1.11).

As mentioned above, the results in [1] are suboptimal due to two key limitations: they rely on a version of (1.12) where the distance is defined by the weaker norm $d(u,v) = |u-v|_{L^{p^*}(\mathbb{R}^N)}$, and they employ only a one-sided estimate for $\delta(u)$, in contrast to the two-sided expansion used in (1.11).

The proof of Theorem 1.1 is structured as follows:

- We begin by establishing the upper bound. To this end, we consider $u_{\lambda,h}$, the projection of a minimizer $U_\lambda \in \mathcal{M}$ (with $\lambda > 0$) onto the finite element space V_h. Using the estimate $\delta(u_{\lambda,h}) \lesssim f\big(d(u_{\lambda,h}, U_\lambda)\big)$, we then evaluate the distance $d(u_{\lambda,h}, U_\lambda)$ in terms of the mesh size h. Optimizing with respect to the parameter λ yields the desired upper bound for the difference $S_h(p,N) - S(p,N)$.
- To establish the lower bound, we consider the discrete minimizer $u_h \in V_h$, for which $\delta(u_h) = S_h(p,N) - S(p,N)$ is known to be small when h is small. A concentration-compactness argument ensures the existence of a minimizer $U_h \in \mathcal{M}$ such that $\delta(u_h) \gtrsim f\big(d(u_h, U_h)\big)$. The final step consists in deriving a sharp lower estimate for the distance $d(u_h, U_h)$.

When the estimates for the Sobolev deficit are used in the classical framework – i.e., in terms of the $W^{1,p}(\mathbb{R}^N)$-norm, as in [15] – rather than via the quasi-norms introduced above, it is still possible to obtain lower and upper bounds. However, these bounds are not sharp. Specifically, this approach yields:

Proposition 1.2. *In the setting of Theorem 1.1,*

$$h^{\gamma(p,N)\max\{2,p\}} \lesssim S_h(p,N) - S(p,N) \lesssim h^{\gamma(p,N)\min\{2,p\}}, \tag{1.13}$$

where

$$\gamma(p,N) = \frac{N-p}{N+p(p-2)}. \tag{1.14}$$

When $p = 2$ and $N \geq 3$, Proposition 1.2 above yields the same sharp result as Theorem 1.1, namely

$$S_h(2,N) - S(2,N) \simeq h^{\frac{2(N-2)}{N}}.$$

Our results lead also to some interesting and challenging questions:

- *Sharper asymptotics.* A natural question that arises is whether the lower and upper bounds can be refined to identify the exact limit
$$\lim_{h\to 0} h^{-\alpha(p,N)}(S_h(p,N) - S(p,N)).$$
In a weaker form, one could instead aim to derive sharper, explicit estimates for the multiplicative constants in the asymptotic unilateral bounds, namely:
$$\underline{\lim}_{h\to 0}\, h^{-\alpha(p,N)}(S_h(p,N) - S(p,N))$$
and
$$\overline{\lim}_{h\to 0}\, h^{-\alpha(p,N)}(S_h(p,N) - S(p,N)).$$
It is worth noting, however, that achieving such refinements may require significant further developments and new technical tools. Moreover, the precise behavior of the asymptotics could depend sensitively on the geometric properties of the finite element mesh.

- *General domains.* Our analysis is limited to the case of the ball B to simplify the presentation. But it can be extended to any bounded, Lipschitz domain which is star-shaped with respect to an internal ball. Also, the same problem could be formulated and analysed for more general finite-element methods. We refer to [1, Remark 3, 4] for a review on the possible extensions one could pursue. One could also consider the problem of the finite element approximation in the whole space $\mathbb{R}^N$.
- *Case* $p = 1$. The case $p = 1$ is even more interesting since the set of minimizers is totally different in this case, $\lambda\chi_B$, with $\lambda \in \mathbb{R}$ and for some ball B. They do not belong to $W^{1,1}(\mathbb{R}^N)$ but $BV(\mathbb{R}^N)$, see [9, 16].

The paper is organized as follows:

- In Section 2, we present in detail the finite element method under consideration and establish some preliminary results on the finite element approximation.
- Section 3 reviews existing work and methodology related to eigenvalue approximation. This classical approach is sufficient to derive upper bounds for the numerical gap $S_h(2, N) - S(2, N)$ without resorting to the Sobolev deficit.
- In Section 4, we recall known results on the Sobolev deficit and demonstrate how the framework developed in [15] can be adapted to our setting to prove Theorem 1.1. We also derive an upper bound on the Sobolev deficit in terms of quasi-norms (see Lemma 4.3). Furthermore, we analyze the class $\mathcal{M}$ of minimizers in Sobolev's inequality (see [2, 20]) to obtain estimates of their Sobolev norms inside and outside the unit ball, and to quantify their distance to the finite element subspace V_h.
- Section 5 contains the complete proofs of the main results. We begin with the proof of Proposition 1.2, which addresses the core difficulties of the problem, and proceed to establish the sharp convergence result stated in Theorem 1.1.
- Finally, the Appendix collects the proofs of several technical results that play a crucial role in the main arguments of the paper.

2. Preliminaries on finite elements

2.1. **The finite element basis.** Let us consider Ω a polyhedral domain in $\mathbb{R}^N$. Of course, depending on the dimension, we should refer to intervals, in dimension $N = 1$, polygons, in dimension $N = 2$, polyhedra, in dimension $N = 3$ and polytopes in arbitrary dimension N. But to simplify the presentation we will generically use the term "polyhedron" without distinguishing the dimension. For each positive h we construct a partition $\mathcal{T}_h$ of the domain Ω into a finite set of polyhedra T satisfying the following properties ([10, p. 38, p. 51], see also [17, Appedix B and C] for a concise presentation):

(1) $\cup_{T\in\mathcal{T}_h} T = \overline{\Omega}$,
(2) Each $T \in \mathcal{T}_h$ is closed and its interior non-empty,
(3) For distinct T and T' their interior are disjoint,
(4) If $T, T' \in \mathcal{T}_h$, $T \neq T'$, then either $T \cap T' = \emptyset$, $T \cap T'$ is a common m-face, $m \in \{0, \dots, N-1\}$.

For each $T \in \mathcal{T}_h$ we denote by ρ_T and h_T the diameter of the largest ball contained in T and the diameter of T respectively. We set

$$h = h(\mathcal{T}_h) = \max_{T\in\mathcal{T}_h} h_T.$$

We will consider a set of regular meshes $(\mathcal{T}_h)_{h>0}$: there exists $\sigma > 0$, independent of h, such that

$$\frac{h_T}{\rho_T} \leq \sigma, \ \forall\, T \in \mathcal{T}_h, \ \forall\, h > 0. \tag{2.1}$$

Also the mesh is assumed to be quasi-uniform, i.e.

$$\inf_{h>0} \frac{\min_{T\in\mathcal{T}_h} h_T}{\max_{T\in\mathcal{T}_h} h_T} > 0.$$

Each element of the mesh $\mathcal{T}_h$ is the image of a reference simplex (interval $(N=1)$/triangle$(N=2)$/tetrahedron$(N=3)$/N-simplex, in general) through an affine mapping $F_T : \mathbb{R}^N \to \mathbb{R}^N$,

$$F_T(\widehat{x}) = B_T\widehat{x} + b_T,$$

B_T being an invertible $N \times N$ matrix, $b \in \mathbb{R}^N$, i.e.

$$F_T(\widehat{T}) = T, \quad \forall\, T \in \mathcal{T}_h.$$

Matrix $B_T = \nabla F_T$ satisfies (cf. [17, Lemma C. 12, p. 735]):

$$\|B_T\| \leq \frac{h_T}{\rho_{\widehat{T}}}, \quad \|B_T^{-1}\| \leq \frac{h_{\widehat{T}}}{\rho_T}$$

and

$$|\det B| = |JF_T| = \frac{|T|}{|\widehat{T}|}.$$

For a fixed partition or grid $\mathcal{T}_h$, we define the finite element space V_h as

$$V_h = \{f \in C(\overline{\Omega}); f \circ F_T \in \mathbb{P}^1(\widehat{T}), \ \forall T \in \mathcal{T}_h, \ f = 0 \ \text{ on } \partial\Omega\},$$

where $\mathbb{P}^1(\widehat{T})$ is the space of linear polynomials on $\widehat{T}$.

Our analysis is carried out in the unit ball, $\Omega = B$. This requirers a first approximation of B by means of polyhedral domains $B_h \subset B$ (as in [1]) such that all the nodes of ∂B_h are on ∂B. The finite element subspace V_h employed to define the numerical approximation of the Sobolev constant will be the one corresponding to the polyhedral domain B_h, so that $V_h \subset H_0^1(B_h) \cap C(\overline{B}_h)$. The elements of V_h can be extended by zero outside B_h so that V_h can also be viewed as a finite-dimensional subspace of $H_0^1(B)$.

2.2. **Approximation by finite elements.** We first recall some classical finite element approximation results in Sobolev spaces [6, Th. 4.4.20]: For any polyhedral domain Ω in $\mathbb{R}^N$, and partition $\mathcal{T}_h$ as above (see [6, Th. 4.4.4] for the complete set of restrictions), the global piecewise linear interpolant I^h, mapping $W^{2,p}(\Omega)$ into V_h satisfies

$$\Big(\sum_{T\in\mathcal{T}_h} \|u - I^h u\|^p_{W^{s,p}(T)}\Big)^{1/p} \leq Ch^{2-s}\|u\|_{W^{2,p}(\Omega)}, \ \forall u \in W^{2,p}(\Omega), \tag{2.2}$$

for all $s \in \{0, 1\}$ and $p > N/2$.

Of course the multiplicative constant C in (2.2) depends on N, p, s and the domain Ω, but it is independent of u. All along the paper the dependence of the constant C on each of the parameters of the problem (in particular the dimension N and the exponent p) will not be made explicit in the notation. Constants in our estimates are normally independent of the function u under consideration (or its discrete counterparts), unless otherwise stated.

The above restriction on $p > N/2$ is necessary when dealing with general functions in $W^{2,p}(\Omega)$. The fact that $p > N/2$ ensures the continuous embedding of $W^{2,p}(\Omega)$ into $C^0(\bar{\Omega})$, needed to define the interpolation I^h by taking the pointwise values over vertices of the mesh.

We now adapt this result to the particular case of $W^{2,\infty}(\Omega)$ functions.

Lemma 2.1. *Let $\mathcal{T}_h$ be a regular mesh on a polyhedral domain $\Omega \in \mathbb{R}^N$, $N \geq 1$, and $s \in \{0,1\}$. There, for all $1 \leq p < \infty$ and $|\alpha| = s$, there exists a positive constant $C = C(N,p,s,\sigma)$ such that:*

$$\Big(\sum_{T\in\mathcal{T}_h} \|D^\alpha(u - I^h u)\|^p_{L^p(T)}\Big)^{1/p} \leq Ch^{2-s}\Big(\sum_{T\in\mathcal{T}_h} |T|\|D^2 u\|^p_{L^\infty(T)}\Big)^{1/p}, \ \forall u \in W^{2,\infty}(\Omega), \tag{2.3}$$

and

$$\max_{T\in\mathcal{T}_h} \|D^\alpha(u - I^h u)\|_{L^\infty(T)} \leq Ch^{2-s} \max_{T\in\mathcal{T}_h} \|D^2 u\|_{L^\infty(T)}, \ \forall u \in W^{2,\infty}(\Omega). \tag{2.4}$$

Moreover, for any $T \in \mathcal{T}_h$ and $p \geq 1$:

$$\|D^\alpha I^h u\|_{L^p(T)} \lesssim h_T^{-|\alpha|+\frac{N}{p}} \|u\|_{L^\infty(T)}, p \geq 1, \ \forall\ u \in C^1(T). \tag{2.5}$$

This result can be proved by slightly modifying the arguments in [6, Th. 4.4.4, Chapter 4] and it can be extended to more general finite elements and $u \in C^m(\overline{\Omega})$ but this is out of the scope of this paper. The proof will be postponed to the Appendix.

The above upper bounds are optimal when considering functions that are strictly convex in one direction.

We recall first the following result, proved in [8, Lemma 2.5], will be instrumental: for any $1 \leq p < \infty$, any strictly convex function $u \in C^2(\Omega)$ and any simplex $T \in \mathcal{T}_h$ with vertices $\{a_k\}_{k=1}^{N+1}$, the interpolation error of u by its linear interpolant $I^h u$ admits the following lower bound:

$$\|u - I^h u\|^p_{L^p(T)} \geq C_p \min_{x\in T} |\lambda_1(D^2 u)(x)|^p |T| \sum_{1\leq i\neq j\leq N+1} |a_i - a_j|^{2p}.$$

In particular when $\mathcal{T}_h$ is quasi-regular we have

$$\|u - I^h u\|^p_{L^p(T)} \geq C_p h^{N+2p} \min_{x\in T} |\lambda_1(D^2 u)(x)|^p.$$

We establish a similar estimate for the gradients.

Lemma 2.2. *For any $p \in (1,\infty)$ there exists a positive constant $C(p)$ such that for any $T \in \mathcal{T}_h$ and any $u \in C^2(T)$*

$$\min_{A\in\mathbb{R}^N} \int_T |Du - A|^p dx \geq C(p)\rho_T^{N+p} \max_{\xi\in\mathbf{S}^{N-1}} \min_{x\in\overline{\Omega}} |\xi^T D^2 u(x)\xi|^p.$$

The proof will be given in the Appendix.

3. Eigenvalue approximation and the case $p = 2$

Let us first recall the classical estimates for the eigenvalues of the Laplacian, [3]. We focus on the simplest case since our aim is to present those tools that will be useful to handle the Sobolev constant for $p = 2$. We refer to [3] for other extensions.

We adopt the notations in [3, Section 8, p. 697]. Let V a real Hilbert space and $a(\cdot,\cdot)$ be a symmetric, continuous and coercive bilinear form on V. Let H be another Hilbert space such that $V \subset H$, with compact embedding, b a continuous symmetric bilinear form on H, satisfying $b(u,u) > 0$, for all $u \in V$, $u \neq 0$. Let $V_h \subset V$ be a family of finite-dimensional spaces of V.

Let λ_1 be the first eigenvalue of *the form a, relative to the form b*. Then, there exists a nonzero $\phi_1 \in V$ such that

$$a(\phi_1, v) = \lambda_1(\phi, v), \forall\ v \in V.$$

In a similar way we define λ_{1h} in the space V_{1h}:

$$a(\phi_{1h}, v) = \lambda_{1h}(\phi_{1h}, v), \forall\, v_h \in V_h.$$

An important result in the eigenvalue approximation theory (Babuška-Osborn [3, Section 8, p. 700], which goes back to Chatelin [7, Prop. 6.30, p. 315]) asserts that:

$$C_1\varepsilon_h^2 \leq \lambda_{1h} - \lambda_1 \leq C_2\varepsilon_h^2 \tag{3.1}$$

where

$$\varepsilon_h = d_V(\phi_1, V_h) = \inf_{u_h \in V_h} \|\phi_1 - u_h\|_V.$$

The upper bound together with the trivial estimate $0 \leq \lambda_{1h} - \lambda_1$, can be also found in [19, Th. 6.4-2]. If the family $(V_h)_{h>0}$ satisfies

$$\forall\, u \in V,\ \lim_{h\to 0} \inf_{v_h \in V_h} \|u - u_h\|_V = 0$$

one obtains that $\varepsilon_h \to 0$. But the obtention of the convergence order of λ_{1h} towards λ_1 requires a finer analysis.

Let us now explain how, inspired on these classical results, one can prove the upper bound statements in Theorem 1.1 and Proposition 1.2 for the Sobolev constant when $p = 2$, $N \geq 3$.

For notational simplicity we write S instead of $S(2, N)$. Let V_h the space of functions constructed in Section 2 extended with zero outside the unit ball. We define also the finite element counterpart:

$$S_h := \min_{u_h \in V_h} \frac{\|Du_h\|_{L^2(\mathbb{R}^N)}}{\|u_h\|_{L^{2^*}(\mathbb{R}^N)}}.$$

Let us consider a minimizer $U \in \mathcal{M}$ with $\|DU\|_{L^2(\mathbb{R}^N)} = 1$, $\|U\|_{L^{2^*}(\mathbb{R}^N)} = 1/S$ and the operator $\mathcal{L} = -U^{2-2^*}\Delta$ on $H = L^2(U^{2^*-2})$ as in [5]. It follows that $\lambda_1 = S^{2^*}$ is the first eigenvalue of $\mathcal{L}$ with the corresponding eigenfunction U. Let be $V = \dot{H}^1(\mathbb{R}^N)$, and the bilinear form $a : V \times V \to \mathbb{R}$ defined as follows:

$$a(u, v) = -\int_{\mathbb{R}^N} (\mathcal{L}u)v\, U^{2^*-2}\, dx = \int_{\mathbb{R}^N} Du \cdot Dv\, dx.$$

Applying the previous classical results on eigenvalue approximation, [3, 7], in a finite dimensional subspace $V_h \subset V$, we obtain that

$$\lambda_{1h} = \min_{u_h \in V_h} \frac{\int_{\mathbb{R}} |Du_h|^2 dx}{\int_{\mathbb{R}^N} u_h^2\, U^{2^*-2} dx}$$

satisfies

$$\lambda_{1h} - S^{2^*} \simeq d_V^2(U, V_h)$$

where

$$d_V(U, V_h) = \inf_{u_h \in V_h} \|D(U - u_h)\|_{L^2(\mathbb{R}^N)}.$$

Let $u_h \in V_h$ be the minimizer corresponding to λ_{1h} above. Using that

$$\int_{\mathbb{R}^N} u_h^2 U^{2^*-2} dx \leq \|u_h\|^2_{L^{2^*}(\mathbb{R}^N)} \|U\|^{2^*-2}_{L^{2^*}(\mathbb{R}^N)} = \|u_h\|^2_{L^{2^*}(\mathbb{R}^N)} S^{2-2^*}$$

we obtain that

$$S^{2^*-2} S_h^2 = S^{2^*-2} \frac{\int_{\mathbb{R}} |Du_h|^2 dx}{\|u_h\|^2_{L^{2^*}(\mathbb{R}^N)}} \leq \frac{\int_{\mathbb{R}} |Du_h|^2 dx}{\int_{\mathbb{R}^N} u_h^2\, U^{2^*-2} dx} = \lambda_{1h}.$$

Combining the two estimates for λ_{1h} we get

$$S_h^2 - S^2 \lesssim d_V^2(U, V_h).$$

Tacking the infimum over $U \in \mathcal{M}_1 := \{U \in \mathcal{M}, \|DU\|_{L^2(\mathbb{R}^N)} = 1\}$ we obtain

$$S_h - S \lesssim \inf_{U \in \mathcal{M}_1} d_V^2(U, V_h).$$

It remains to estimate the above right hand side to obtain the upper bound in Theorem 1.1. However, for the lower bound one has to use estimates for the Sobolev deficit as described in the introduction. Let us briefly explain how this can be obtained in the above setting.

Let us now take $u_h \in V_h$ with $\|Du_h\|_{L^2(\mathbb{R}^N)} = 1$, the function for which S_h above is attained. The inequalities for the Sobolev deficit in [5] show that

$$S_h - S \gtrsim d_V^2(u_h, \mathcal{M}).$$

When $S_h - S$ is small (something that can be assured as soon as $h > 0$ is small enough) there exists $U_{\lambda_h} \in \mathcal{M}$ such that $d_V(u_h, \mathcal{M}) = d_V(u_h, U_{\lambda_h})$ and then $\|DU_{\lambda_h}\|_{L^2(\mathbb{R}^N)} \geq 1/2$.

Hence

$$S_h - S \gtrsim d_V^2(U_{\lambda_h}, u_h) \geq \inf_{U \in \mathcal{M}_1} d_V^2(U, V_h).$$

We finally obtain that

$$S_h - S \simeq \inf_{U \in \mathcal{M}_1} d_V^2(U, V_h). \tag{3.2}$$

The results in the next sections will show that for piecewise finite element method

$$\inf_{U \in \mathcal{M}_1} d_V^2(U, V_h) \simeq h^{2N/(N-2)}.$$

Our goal in the next section is to extend this analysis to the general case $1 < p < N$, and to establish the two-sided bounds stated in Theorem 1.1.

4. The Sobolev constant and finite element approximation of minimizers

This Section is devoted to gather a number of technical results whose proofs are postponed to the Appendix at the end of the paper.

Let us recall some classical facts about the Sobolev inequality: for any $1 < p < N$, $N \geq 2$, it holds

$$\|Du\|_{L^p(\mathbb{R}^N)} \geq S(p, N)\|u\|_{L^{p^*}(\mathbb{R}^N)}, \; \forall\, u \in \dot{W}^{1,p}(\mathbb{R}^N), \tag{4.1}$$

where $p^* = \frac{pN}{N-p}$. The optimal constant $S(p, N)$ was obtained by Aubin [2] and Talenti [20] and shown to be attained for the $(n+2)$-dimensional manifold $\mathcal{M}$ of functions of the form

$$U_{c,\lambda,x_0}(x) = cU_{\lambda,x_0}(x) = c\lambda U(\lambda^{p/(N-p)}(x - x_0)), \tag{4.2}$$

with

$$U(x) = u_0(|x|);\; u_0(r) = \frac{k_0}{(1 + |r|^{\frac{p}{p-1}})^{\frac{N-p}{p}}}. \tag{4.3}$$

where k_0 is chosen such that $\|DU\|_{L^p(\mathbb{R}^N)} = 1$. In particular, for $c = 1$, $\|DU_{\lambda,x_0}\|_{L^p(\mathbb{R}^N)}$=1.

The $n+2$ free parameters of the manifold $\mathcal{M}$ are the center of gravity $x_0 \in \mathbb{R}^N$, the positive parameter $\lambda > 0$ and the constant $c \in \mathbb{R}$.

4.1. **The p - Sobolev deficit.** Let us now consider the so called p-Sobolev deficit

$$\delta(u)=\frac{\|Du\|_{L^p(\mathbb{R}^N)}}{\|u\|_{L^{p*}(\mathbb{R}^N)}}-S(p,N),\ \forall\ u\in \dot W^{1,p}(\mathbb{R}^N).$$

It is well known from the works of Bianchi and Egnel [5] and more recently of Figalli and Zhang [15], that for every $1<p<N$, there exists a positive constant $c(p,N)$ such that

$$\delta(u)\geq c(p,N)\inf_{v\in\mathcal{M}}\Big(\frac{\|Du-Dv\|_{L^p(\mathbb{R}^N)}}{\|Du\|_{L^p(\mathbb{R}^N)}}\Big)^{\max\{2,p\}},\ \forall\ u\in\ W^{1,p}(\mathbb{R}^N). \tag{4.4}$$

In fact, the results in [15] include a refined estimate in the range $2\leq p<N$, which is particularly useful for our analysis. This estimate involves a quasi-norm, rather than the classical Sobolev seminorm $\|Du-Dv\|_{L^p(\mathbb{R}^N)}$ and is reminiscent of the approach introduced in [4, 12] to obtain sharper convergence rates for the p-Laplacian. The following lemma presents a refined version of these estimates, valid for the full range $1<p<N$, expressed in terms of the quasi-norms used in finite element theory.

Lemma 4.1. *Let $1<p<N$. There is a positive constant $c(p,N)$ such that for any $u\in \dot W^{1,p}(\mathbb{R}^N)$ there exists a function $v=v_u\in\mathcal{M}$ such that*

$$\frac{\delta(u)}{c(p,N)}\geq\frac{\int_{\mathbb{R}^n}(|Du-Dv|+|Dv|)^{p-2}|Du-Dv|^2\,dx}{\|Du\|^p_{L^p(\mathbb{R}^N)}}+\Big(\frac{\|Du-Dv\|_{L^p(\mathbb{R}^N)}}{\|Du\|_{L^p(\mathbb{R}^N)}}\Big)^{\max\{2,p\}}. \tag{4.5}$$

Remark 4.2. In particular we obtain

$$\delta(u)\geq c\inf_{v\in\mathcal{M}}\frac{\int_{\mathbb{R}^n}(|Du-Dv|+|Dv|)^{p-2}|Du-Dv|^2\,dx}{\|Du\|^p_{L^p(\mathbb{R}^N)}},\ \forall\ u\in\dot W^{1,p}(\mathbb{R}^N). \tag{4.6}$$

Proof. The proof follows the lines of Theorem 1.1 in [15].

By homogeneity we can assume that $\|Du\|_{L^p(\mathbb{R}^N)}=1$. Also it is sufficient to consider the case where $\delta(u)<<1$. Lemma 4.1 in [15] shows that, when the deficit is small, there exists $v\in\mathcal{M}$ such that $\varepsilon:=\|Du-Dv\|_{L^p(\mathbb{R}^N)}$ is small and $u-v$ is orthonormal to $T_v\mathcal{M}$. For full details, we refer the reader to the proof of Theorem 1.1 and Lemma 4.1 in [15].

For $1<p\leq 2N/(N+2)$, [15, formulas (4.4), (4.5)] give

$$\begin{aligned}\delta(u)&\geq c_0(p,N)\int_{\mathbb{R}^N}\min\{|D(u-v)|^p,|Dv|^{p-2}|D(u-v)|^2\}dx\\&=c_0(p,N)\int_{\mathbb{R}^N}\min\{|D(u-v)|^{p-2},|Dv|^{p-2}\}|D(u-v)|^2dx.\\&\geq c_1(p,N)\|D(u-v)\|^2_{L^p(\mathbb{R}^N)}.\end{aligned} \tag{4.7}$$

When $p<2$ we also get

$$\begin{aligned}\delta(u)&\geq c_0(p,N)\int_{\mathbb{R}^N}\min\{|D(u-v)|^{p-2},|Dv|^{p-2}\}|D(u-v)|^2dx\\&\geq c_0(p,N)\int_{\mathbb{R}^N}(|D(u-v)|+|Dv|)^{p-2}|D(u-v)|^2dx.\end{aligned} \tag{4.8}$$

Both inequalities yield to (4.1) when $1<p\leq 2N/(N+2)$.

When $p \in (2N/(N+2), 2)$ repeating the arguments in [15] we have

$$\begin{aligned}\delta(u) &\geq c_0(p,N)\int_{\mathbb{R}^N}\min\{|D(u-v)|^{p-2},|Dv|^{p-2}\}|D(u-v)|^2dx - C_1\int_{\mathbb{R}^N}|u-v|^{p^*}dx\\ &\geq \frac{c_0(p,N)}{2}\int_{\mathbb{R}^N}\min\{|D(u-v)|^{p-2},|Dv|^{p-2}\}|D(u-v)|^2dx + \frac{c_1(p,N)}{4}\|D(u-v)\|^2_{L^p(\mathbb{R}^N)}\\ &\quad + \frac{c_1(p,N)}{4}\|D(u-v)\|^2_{L^p(\mathbb{R}^N)} - C_1\|u-v\|^{p^*}_{L^{p^*}(\mathbb{R}^N)}.\end{aligned}$$

Since $p^* > 2$ the last term in the right hand side is positive by the Sobolev inequality if $\|D(u-v)\|_{L^p(\mathbb{R}^N)}$ is small.

For $p \geq 2$ the proof of [15, Th. 1.1] gives

$$\begin{aligned}\delta(u) &\geq c_0(p,N)\Big(\int_{\mathbb{R}^N}|Dv|^{p-2}|D(u-v)|^2\,dx + \int_{\mathbb{R}^N}|D(u-v)|^p\,dx\Big)\\ &\geq c_1(p,N)\|Du-Dv\|^p_{L^p(\mathbb{R}^N)}\end{aligned}$$

and the proof is finished.

□

We now turn our attention to the upper bounds of the deficit.

Lemma 4.3. *Let $1<p<N$. There exists a positive constant $C(p,N)$ and a positive number ε_0 such that*

$$\frac{\delta(u)}{C(p,N)} \leq \frac{\int_{\mathbb{R}^n}(|Du|+|D(u-v)|)^{p-2}|Du-Dv|^2}{\|Du\|^p_{L^p(\mathbb{R}^N)}}\,dx, \quad p\geq 2, \tag{4.9}$$

and

$$\frac{\delta(u)}{C(p,N)} \leq \frac{\int_{\mathbb{R}^n}(|Du|+|D(u-v)|)^{p-2}|Du-Dv|^2}{\|Du\|^p_{L^p(\mathbb{R}^N)}} + \Big(\frac{\int_{\mathbb{R}^N}|Dv|^{p-1}|D(u-v)|\,dx}{\|Du\|^p_{L^p(\mathbb{R}^N)}}\Big)^2, \quad 1<p<2, \tag{4.10}$$

hold for all $u \in W^{1,p}(\mathbb{R}^N)$ with $\delta(u) < S(p,N)$ and $v \in \mathcal{M}$ such that $\|D(u-v)\|_{L^p(\mathbb{R}^N)} \leq \varepsilon_0\|Du\|_{L^p(\mathbb{R}^N)}$.

Remark 4.4. As we shall see, we shall prove that, for any $1<p<N$, the following inequality holds:

$$\delta(u) \leq \frac{C(p,N)}{\|Du\|^p_{L^p(\mathbb{R}^N)}}\Big(\int_{\mathbb{R}^N}(|Dv|+|D(u-v)|)^{p-2}|D(u-v)|^2dx + \int_{\mathbb{R}^n}|Dv|^{p-2}|Du-Dv|^2dx\Big). \tag{4.11}$$

This inequality immediately leads to estimate (4.9).

However, when $1<p<2$, as we shall below (see remark 7.3 in the proof of Lemma 4.10, Appendix 7.7), the last term on the right hand side, namely

$$\int_{\mathbb{R}^n}|Dv|^{p-2}|Du-Dv|^2dx$$

is singular. Thus, it has to be replaced by

$$\Big(\int_{\mathbb{R}^N}|Dv|^{p-1}|D(u-v)|\,dx\Big)^2,$$

which is well defined also in the range $1<p<2$ when $u \in V_h$ and $v \in \mathcal{M}$.

As a consequence of this lemma, we can deduce an upper counterpart to the lower bound in (4.4). Although one might expect that an upper bound for the deficit would be easier to obtain than the lower bound, we have not found such a result in the literature. This estimate will play an important role in the proof of Proposition 1.2.

Corollary 4.5. *There is a constant $C = C(p, N)$ and a positive number ε_0 such that*

$$\delta(u) \le C \inf_{v\in\mathcal{M}} \Big(\frac{\|Du - Dv\|_{L^p(\mathbb{R}^N)}}{\|Du\|_{L^p(\mathbb{R}^N)}}\Big)^{\min\{2,p\}}. \tag{4.12}$$

holds for all $u \in W^{1,p}(\mathbb{R}^N)$ satisfying

$$d(u, \mathcal{M}) = \inf_{v\in\mathcal{M}} \|Du - Dv\|_{L^p(\mathbb{R}^N)} \le \varepsilon_0 \|Du\|_{L^p(\mathbb{R}^N)}.$$

Remark 4.6. This estimate is sharp. In fact, the same argument as in [15, Remark 1.2] shows that if

$$\delta(u) \le C \inf_{v\in\mathcal{M}} \Big(\frac{\|Du - Dv\|_{L^p(\mathbb{R}^N)}}{\|Du\|_{L^p(\mathbb{R}^N)}}\Big)^{\beta} \tag{4.13}$$

holds, then, necessarily, $\beta \le \min\{2, p\}$.

Note that there is a gap in the exponents, $\max\{2, p\}$ for the lower bound in (4.4), while the upper bound in (4.12) holds with $\min\{2, p\}$.

4.2. **Estimates on the minimizers $\mathcal{M}$.** In this section we establish various estimates for the minimizers $U_{\lambda,x_0} \in \mathcal{M}$ that, as indicated above, are defined as in (4.2) and (4.3). Their proofs are given in the Appendix.

Lemma 4.7. *Let $1 < p < N$ and $N \ge 2$. There exist two positive constants $c_1 = c_1(N, p)$ and $c_2 = c_2(N, p)$ such that for any $|x_0| < 1$ the minimizer U_{λ,x_0} satisfies*

$$c_1\lambda^{\frac{p}{N-p}(N+\frac{p}{p-1})} \le \int_{|x|<1} |DU_{\lambda,x_0}|^p dx \le c_2\lambda^{\frac{p}{N-p}(N+\frac{p}{p-1})},\ \forall\ 0 < \lambda < 1, \tag{4.14}$$

$$1 - \frac{c_2\lambda^{-\frac{p}{p-1}}}{(1 - |x_0|)^{\frac{N-p}{p-1}}} \le \int_{|x|<1} |DU_{\lambda,x_0}|^2 dx \le 1 - c_1\lambda^{-\frac{p}{p-1}},\ \forall\ \lambda > 1. \tag{4.15}$$

Remark 4.8. Observe that, when $|x_0| \ge 1$, the exterior of the unit ball contains a half space passing through x_0. Thus,

$$\int_{|x|\ge1} |DU_{\lambda,x_0}|^2 dx \ge \frac{1}{2}\int_{\mathbb{R}^N} |DU_{\lambda,x_0}|^2 dx = \frac{1}{2}$$

for all $\lambda > 0$. This illustrates behavior opposite to the energy concentration phenomenon described by (4.15) in the case $|x_0| < 1$.

Lemma 4.9. *Let $1 < p < N$. The second order derivatives of the minimizer U_{λ,x_0} satisfy the following estimates:*

$$|D^2U_{\lambda,x_0}(x)| \le C_{N,p}\lambda^{\frac{N+p}{N-p}} a(\lambda^{\frac{p}{N-p}}|x - x_0|),\ x \in \mathbb{R}^N, \tag{4.16}$$

where

$$a(r) = \frac{u_0'(r)}{r} = r^{\frac{2-p}{p-1}}\Big(1 + r^{\frac{p}{p-1}}\Big)^{-\frac{N}{p}}$$

and

$$\int_{|x|<1} |D^2 U_{\lambda,x_0}(x)|^p dx \le C_{N,p}\lambda^{\frac{p^2}{N-p}}. \tag{4.17}$$

Moreover, there exists a positive constant $A_{N,p}$, a finite covering of the plane with open sets $\mathbb{R}^N \subset \cup_{k\in F}\Gamma_k$ and a set $\{\xi_k\}_{k\in F} \in \mathbf{S}^1$ such that

$$|\xi_k^T D^2 U_{\lambda,x_0}(x)\xi_k| \ge A_{N,p}\lambda^{\frac{N+p}{N-p}} a(\lambda^{\frac{p}{N-p}}|x-x_0|), \quad \forall \lambda > 0, x \in \Gamma_k,\ k \in F. \tag{4.18}$$

The proofs of these results are presented in the Appendix.

4.3. **Approximation of the minimizers.** As discussed in Section 3, it is necessary to estimate the distance between the minimizers in $\mathcal{M}$ and the space V_h, the $P1$ finite element space introduced in Section 2.

To obtain the upper bounds in Proposition 1.2 and Theorem 1.1, we do not need to measure how far all the minimizers in $\mathcal{M}$ are from the space V_h; it suffices to sample just one. In the next lemma, we estimate the distance between the minimizer $U_{\lambda,0}$ and its projection onto V_h, both in the classical Sobolev norm and in the quasi-norms. Since the upper bound for the Sobolev deficit obtained in Lemma 4.3 includes an additional term when $1 < p < 2$, we also need to estimate this remainder term.

Lemma 4.10. *Let $N \ge 2$ and $1 < p < N$. For any $\lambda > 1$ and $h < 1/2$ such that $h\lambda^{\frac{p}{N-p}} \le 1$ the following holds*

$$\int_{\mathbb{R}^N} |D(U_{\lambda,0} - u_h)|^p dx \lesssim \lambda^{-\frac{p}{p-1}} + (h\lambda^{\frac{p}{N-p}})^p, \tag{4.19}$$

$$\int_{\mathbb{R}^N} (|DU_{\lambda,0}| + |D(U_{\lambda,0} - u_h)|)^{p-2} |D(U_{\lambda,0} - u_h)|^2 dx \lesssim \lambda^{-\frac{p}{p-1}} + (h\lambda^{\frac{p}{N-p}})^2, \tag{4.20}$$

$$\int_{\mathbb{R}^N} |DU_{\lambda,0}|^{p-1} |D(U_{\lambda,0} - u_h)| dx \lesssim \lambda^{-\frac{p}{p-1}} + h\lambda^{\frac{p}{N-p}}, 1 < p < 2, \tag{4.21}$$

where $u_h \in V_h$ is given by $u_h = I^h(U_{\lambda,0} - (U_{\lambda,0})_{|\partial B_h})$ extended with zero outside B_h.

Remark 4.11. The above lemma implies

$$\inf_{u_h\in V_h} \|D(U_{\lambda,0} - u_h)\|_{L^p(\mathbb{R}^N)} \lesssim \lambda^{-\frac{1}{p-1}} + h\lambda^{\frac{p}{N-p}}.$$

Optimising this estimate with $\lambda^{-\frac{1}{p-1}} = h^{-\frac{(N-p)}{N+p^2-2p}}$ we get

$$\inf_{u_h\in V_h} \|D(U_{\lambda,0} - u_h)\|_{L^p(\mathbb{R}^N)} \lesssim h^{\gamma(p,N)}$$

where

$$\gamma(p, N) = \frac{N-p}{N + p(p-2)}.$$

A similar argument works for the quasi-norms

$$\inf_{u_h\in V_h} \int_{\mathbb{R}^N} (|DU_{\lambda,0}| + |D(U_{\lambda,0} - u_h)|)^{p-2} |D(U_{\lambda,0} - u_h)|^2 dx \lesssim h^{\alpha(p,N)}$$

with

$$\alpha(p, N) = \frac{2(N-p)}{N+p-2}.$$

5. Proof of the main results

In this section we first prove Proposition 1.2

$$h^{\gamma(p,N)\max\{2,p\}} \lesssim S_h(p,N) - S(p,N) \lesssim h^{\gamma(p,N)\min\{2,p\}}, \tag{5.1}$$

and secondly we consider the improved result in Theorem 1.1

$$h^{\alpha(p,N)} \lesssim S_h(p,N) - S(p,N) \lesssim h^{\alpha(p,N)}. \tag{5.2}$$

The proof of Theorem 1.1 follows the same overall strategy as that of Proposition 1.2, with the key difference that the Sobolev distances used in Proposition 1.2 are replaced by quasi-norms of the finite element analysis the p-Laplacian. This allows to obtain sharp lower and upper bounds of the same asymptotic order in h.

The proof of Proposition 1.2 (and therefore that of Theorem 1.1 as well) is conducted in two main steps:

- Step 1. In the first step we prove the upper bound in (5.1) (or in (5.2)). For that we employ the finite element projection of suitable continuous minimizer U_λ by choosing optimally the values of $\lambda >> 1$ and $h << 1$.
- Step 2. In the second step we prove the lower bound. Employing the general lower bound for the deficit in (4.1), this is done by a suitable approximation in $\mathcal{M}$ of the finite-element minimizer.

We consider the unit ball B. The Sobolev constant in the unit ball B coincides with that in the whole space, although it is not attained in B; see [1].

We construct the space V_h as in Section 2 and use that

$$S_h(p,N) = \min_{u_h\in V_h} \frac{\|Du_h\|_{L^p(B)}}{\|u_h\|_{L^{p^*}(B)}} = \min_{u_h\in V_h} \frac{\|Du_h\|_{L^p(\mathbb{R}^N)}}{\|u_h\|_{L^{p^*}(\mathbb{R}^N)}},$$

identifying V_h with the space of functions defined everywhere by extending them by zero outside their support.

For notational simplicity, throughout this section we will write S_h and S in place of the full expressions for the two constants introduced above.

Proof of Proposition 1.2. As described above we proceed in two main steps.
Step I. The upper bound. To obtain un upper bound we choose a suitable U_λ and the finite element approximation $u_{h,\lambda} = I^h(U_\lambda - U_{\lambda|\partial B_h}) \in V_h$ extended by zero outside B_h, with a suitable $\lambda >> 1$ and $h << 1$.

Lemma 4.10 shows that choosing

$$\lambda^{\frac{1}{p-1}} = h^{-\frac{(N-p)}{N+p^2-2p}},$$

we have

$$\|Du_{h,\lambda} - DU_\lambda\|_{L^p(\mathbb{R}^N)} \sim h^{\gamma(p,N)}.$$

This allows us to use Corollary 4.5 to obtain

$$S_h \le \frac{\|Du_{h,\lambda}\|_{L^p(\mathbb{R}^N)}}{\|u_{h,\lambda}\|_{L^{p^*}(\mathbb{R}^N)}} \le S + C(p,N)\Big(\frac{\|Du_{h,\lambda} - DU_\lambda\|_{L^p(\mathbb{R}^N)}}{\|Du_{h,\lambda}\|_{L^p(\mathbb{R}^N)}}\Big)^{\min\{2,p\}}.$$

Since $\|DU_\lambda\|_{L^p(\mathbb{R}^N)} = 1$ we also get $\|Du_{h,\lambda}\|_{L^p(\mathbb{R}^N)} \simeq 1$. Thus, we get the desired upper bound

$$S_h - S \lesssim h^{\gamma(p,N)\min\{2,p\}}.$$

Step II. The lower bound. Let $u_h \in V_h$ be a minimizer of S_h with, for simplicity, $\|Du_h\|_{L^p(\mathbb{R}^N)} = 1$.

From the previous upper bound and the estimate of the Sobolev deficit (4.4) we have that there exists $U_{c_h,\lambda_h,x_h} = c_h U_{\lambda_h,x_h} \in \mathcal{M}$ such that

$$h^{\gamma(p,N)\min\{2,p\}} \gtrsim S_h - S = \delta(u_h) \gtrsim \Big(\frac{\|Du_h - c_h DU_{\lambda_h,x_h}\|_{L^p(\mathbb{R}^N)}}{\|Du_h\|_{L^p(\mathbb{R}^N)}}\Big)^{\max\{2,p\}}.$$

We now need to get a lower bound on the right hand side of this inequality in terms of h. For this we need to gather further information on c_h, x_h and λ_h. We therefore proceed in two steps.

- In the following we prove that $c_h \to 1$, $x_h \in B_h$, $\lambda_h \geq \lambda_0$ and $h\lambda_h^{\frac{p}{N-p}} \leq \Lambda$ for some positive constants λ_0, Λ as $h \to 0$.

  Since $\|Du_h\|_{L^p(\mathbb{R}^N)} = \|DU_{\lambda_h,x_h}\|_{L^p(\mathbb{R}^N)} = 1$ it follows that $|c_h - 1| \leq h^{\gamma(p,N)\min\{2,p\}}$. Choosing h small we trivially have $|c_h|^p > 1/2$ and then

$$\begin{aligned} o(1) =& \delta(u_h) = \|Dw_h - c_h DU_{\lambda_h,x_h}\|^p_{L^p(\mathbb{R}^N)} = |c_h|^p \int_{B_h^c} |DU_{\lambda,x_h}|^p dx + \int_{B_h} |c_h DU_{\lambda,x_h} - Dw_h|^p dx \\ \gtrsim& \int_{B^c} |DU_{\lambda_h,x_h}|^p dx + \int_{B_h} |DU_{\lambda_h,x_h} - c_h^{-1} Du_h|^p dx \\ =& \int_{B^c} |DU_{\lambda_h,x_h}|^p dx + I_h. \end{aligned}$$

  Assuming $|x_h| \geq 1$, Remark 4.8 shows that for all $\lambda > 0$ the exterior of the unit ball B^c contains always a half space passing through x_h and then

$$o(1) \gtrsim \int_{B^c} |DU_{\lambda_h,x_h}|^p dx \geq \frac{1}{2}.$$

  Hence $|x_h| < 1$. A similar argument shows that $x_h \notin B \setminus B_h$.

  Let us assume that, up to a subsequence, $\lambda_h \to 0$. By estimate (4.14) we obtain

$$o(1) \gtrsim \int_{B^c} |DU_{\lambda_h,x_h}|^p dx = 1 - \int_B |DU_{\lambda_h,x_h}|^p dx \geq 1 - c_1 \lambda_h^{\frac{p}{N-p}(N+\frac{p}{p-1})},$$

  which gives a contraction. Hence, there exists a constant λ_0 such that $\lambda_h \geq \lambda_0$ for all sufficiently small h.

  Let us now prove that the sequence $(\lambda_h^{\frac{p}{N-p}} h)_{h>0}$ is uniformly bounded. Suppose, by contradiction and up to a subsequence, that $\lambda_h^{p/(N-p)} h \to \infty$. We will show in this case that

$$I_h \gtrsim 1.$$

  Let $x_h \in B_h$ be the point where the maximum of the integrand I_h is essentially concentrated. This point lies in some triangle $T_{0h} \in \mathcal{T}_h$. Without loss of generality, we may assume that the ball $B_{\lambda^{-p/(N-p)}}(x_h) \subset T_{0h}$. If this inclusion does not hold, we can instead consider the intersection $T_{0h} \cap B_{\lambda^{-p/(N-p)}}(x_h)$, which still contains a fixed proportion of the ball. Due to the mesh regularity condition (2.1), this proportion is independent of h.

Since Dw_h is constant inside triangle T_{0h}, and satisfies $c_h^{-1}Dw_h = A_h$ we can apply estimate (4.18) from Lemma 4.9 to obtain a contradiction:

$$o(1) \gtrsim I_h \geq \int_{|x-x_h|<\lambda_h^{-\frac{p}{N-p}}} |\lambda^{\frac{N}{N-p}} DU(\lambda_h^{\frac{p}{N-p}}(x-x_h)) - A_h|^p dx$$
$$= \int_{|x|<1} |DU(x) - A_h\lambda_h^{-\frac{N}{N-p}}|^p dx \geq \inf_{A\in\mathbb{R}^N} \int_{|x|<1} |DU(x) - A|^p dx \geq C_0 > 0.$$

- With the above considerations on x_h, λ_h and $h\lambda_h^{\frac{p}{N-p}}$ we proceed to prove the desired lower bound. Choosing, if needed, a smaller value of h, we can assume

$$h\lambda_h^{\frac{p}{N-p}} \leq \Lambda < \frac{\lambda_0^p}{10} \leq \frac{\lambda_h^{\frac{p}{N-p}}}{10}.$$

We use estimate (4.15) to obtain

$$\int_{B^c} |DU_{\lambda_h,x_h}|^p \, dx \geq c_1 \lambda_h^{-\frac{p}{p-1}}, \tag{5.3}$$

for some positive constant c_1.

Let us now estimate the term I_h. Using that u_h is piecewise constant in each triangle $T \in \mathcal{T}_h$, by Lemma 2.2 and estimate (4.18), we obtain

$$I_h = \sum_{T\in\mathcal{T}_h} \int_T |DU_{\lambda_h,x_h} - Du_h|^p \, dx \geq \sum_{T\in\mathcal{T}_h} \min_{A_T\in\mathbb{R}^N} \int_T |DU_{\lambda_h,x_h} - A_T|^p \, dx$$
$$\gtrsim \sum_{T\in\mathcal{T}_h} h^{N+p} \lambda^{\frac{p(N+p)}{N-p}} \min_{x\in T} a^p(\lambda_h^{\frac{p}{N-p}}|x-x_h|).$$

We denote by $\mathcal{T}_h^1$ the elements $T \in \mathcal{T}_h$ which have a nonempty intersection with the ball $\{|x - x_h| \geq 2h\}$. Any element $T \in \mathcal{T}_h^1$ is included in the exterior of the ball $\{|x - x_h| < h\}$ and for any $x \in T \subset \mathcal{T}_h^1$ we have

$$\max_{\overline{x}\in T} |\overline{x} - x_h| \leq |x - x_h| + h_T \leq |x - x_h| + h \leq 2|x - x_h|$$

and

$$|x - x_h| \leq \max_{\overline{x}\in T} |\overline{x} - x_h| \leq 2 \min_{\underline{x}\in T} |\underline{x} - x_h| \leq 2|x - x_h|.$$

This implies that

$$I_h \geq \sum_{T\in\mathcal{T}_h^1} h^{N+p}\lambda_h^{\frac{p(N+p)}{N-p}} \min_{x\in T} a^p(\lambda_h^{\frac{p}{N-p}}|x-x_h|) \gtrsim h^p \lambda^{\frac{p(N+p)}{N-p}} \sum_{T\in\mathcal{T}_h^1} \int_T a^p(\lambda_h^{\frac{p}{N-p}}|x-x_h|)\, dx$$
$$\gtrsim h^p \lambda_h^{\frac{p(N+p)}{N-p}} \int_{h<|x-x_h|,|x|<1-h} a^p(\lambda_h^{\frac{p}{N-p}}|x-x_h|)\, dx.$$

If $|x_h| < 1/2$ then the set $\{h < |x-x_h| < 1/4\}$ is included in $\omega_h = \{h < |x-x_h|, |x| < 1-h\} \subset B_h$ since $|x| \leq |x_h| + |x-x_h| < 1/2 + 1/4 < 1-h$. This implies

$$I_h \gtrsim h^p \lambda_h^{\frac{p(N+p)}{N-p}} \int_{h<|x-x_h|<1/4} a^p(\lambda_h^{\frac{p}{N-p}}|x-x_h|)\, dx \simeq h^p \lambda_h^{\frac{p^2}{N-p}} \int_{h\lambda_h^{\frac{p}{N-p}}<|y|<\lambda_h^{\frac{p}{N-p}}/4} a^p(|y|)\, dy$$
$$\gtrsim h^p \lambda_h^{\frac{p^2}{N-p}} \int_{\Lambda<|x|<\lambda_0^p/4} a^p(|x|)\, dx.$$

When $|x_h| > 1/2$, it may happen that x_h lies close to the boundary of B_h. Nevertheless, even in the worst-case scenario, we can always construct a cone centered at x_h that is entirely contained within the set ω_h. Assume that $x_h = (x_{1h}, 0')$ with $x_{1h} \in [1/2, 1]$. Thus

$$\{x = (x_1, x'), |x' - 0'| \le \alpha|x_{1h} - x_1|, 2h \le x_{1h} - x_1 \le \frac{1}{2}\} \subset \omega_h$$

since, in this case, $|x| \le |x_1| + \alpha|x_1 - x_{1h}| \le \alpha x_{1h} + (1-\alpha)x_{1h} = x_{1h} - 2(1-\alpha)h < 1-h$ for $\alpha < 1/2$. Denoting

$$\omega' = \{y = (y_1, y'), |y'| \le \alpha|y_1|, 2h \le y_1 \le \frac{1}{2}\}$$

we obtain, after a change of variables, that

$$I_1 \gtrsim h^p \lambda_h^{\frac{p^2}{N-p}} \int_{y \in \lambda_h^{\frac{p}{N-p}} \omega'} a^p(|y|)dy \gtrsim h^p \lambda_h^{\frac{p^2}{N-p}},$$

where we have used that the set $\lambda_h^{\frac{p}{N-p}}\omega'$ contains a subset with positive measure, independent of the small parameter h,

$$\begin{aligned}\{y = (y_1, y'), |y'| \le &\alpha|y_1|, 2\Lambda \le y_1 \le \frac{\lambda_0^p}{2}\} \\ &\subset \{y = (y_1, y'), |y'| \le \alpha|y_1|, 2h\lambda_h^{\frac{p}{N-p}} \le y_1 \le \frac{\lambda_h^{\frac{p}{N-p}}}{2}\} \subset \lambda_h^{\frac{p}{N-p}}\omega'.\end{aligned}$$

Putting together the estimates for I_h and (5.3) we find that

$$S_h - S \gtrsim (\lambda_h^{-\frac{1}{p-1}} + h\lambda_h^{\frac{p}{N-p}})^{\max\{2,p\}} \gtrsim h^{\gamma(p,N)\max\{2,p\}}.$$

which finishes the proof.

□

Proof of Theorem 1.1. We proceed as above, but now considering the quasi-norms.

Step I. The upper bound. The proof follows the same ideas as in Proposition 1.2 by taking $u_{h,\lambda} = I^h(U_\lambda - U_{\lambda|\partial B_h}) \in V_h$. In view of estimate (4.19), $\|Du_{h,\lambda} - DU_\lambda\|_p$ is small enough when λ is large and $h\lambda^{\frac{p}{N-p}}$ is small. This allows us to apply the upper bounds for the Sobolev deficit in Lemma 4.3.

For $p \ge 2$ the deficit estimates in Lemma 4.3 give

$$\delta(u_{h,\lambda}) \lesssim \int_{\mathbb{R}^N} (|Du_{h,\lambda}| + |DU_\lambda - Du_{h,\lambda}|)^{p-2}|DU_\lambda - Du_{h,\lambda}|^2 dx$$

and estimate (4.20) leads to

$$\delta(u_{h,\lambda}) \lesssim \lambda^{-\frac{p}{p-1}} + (h\lambda^{\frac{p}{N-p}})^2.$$

Taking λ such that $\lambda^{\frac{p}{p-1}} = (h\lambda^{\frac{p}{N-p}})^2$ we find the desired estimate $\delta(u_h) \lesssim h^{\alpha(p,N)}$ with

$$\alpha(p, N) = \frac{2(N-p)}{N+p-2}.$$

For $1<p<2$ we apply Lemma 4.3 to get

$$\delta(u_{h,\lambda})\leq \int_{\mathbb{R}^N}(|Du_{h,\lambda}|+|D(U_\lambda-u_{h,\lambda})|)^{p-2}|D(U_\lambda-u_{h,\lambda})|^2dx + \Big(\int_{\mathbb{R}^N}|DU_\lambda|^{p-1}|D(U_\lambda-u_{h,\lambda}|\,dx\Big)^2.$$

Using estimates (4.20) and (4.21) we obtain

$$\delta(u_{h,\lambda})\lesssim \lambda^{-\frac{p}{p-1}}+(h\lambda^{\frac{p}{N-p}})^2.$$

Taking λ conveniently, the proof of the upper bound is obtained as in the proof of Proposition 1.2.

Step II. The lower bound. As in the proof of Proposition 1.2 let us take $u_h\in V_h$ with $\|Du_h\|_{L^p(\mathbb{R}^N)}=1$ and $\|u_h\|_{L^{p^*}(\mathbb{R}^N)}=1/S_h$. From the upper bound estimate in Step I we have

$$h^{\alpha(p,N)}\gtrsim S_h-S=\delta(u_h).$$

By Lemma 4.1 there exists $U^h=U_{c_h,\lambda_h,x_h}$ such that and (4.5) holds:

$$\delta(u_h)\gtrsim \int_{\mathbb{R}^N}(|Du_h|+|D(U^h-u_h)|)^{p-2}|D(U^h-u_h)|^2dx+\int_{\mathbb{R}^N}|D(U^h-u_h)|^pdx.$$

The same analysis in Step II of the proof of Proposition 1.2 shows that $c_h\simeq 1$, $h\lambda_h^{\frac{p}{N-p}}\leq\Lambda$ and $\lambda_h\geq\lambda_0>0$. Also $\|D(U^h-u_h)\|_{L^p(\mathbb{R}^N)}\leq 1$ for h small and $\|DU^h\|_{L^p(\mathbb{R}^N)}\simeq 1$.

For $p\geq 2$ the same analysis as in the proof of Proposition 1.2 shows that there exists a set ω independent of h such that

$$\begin{aligned}\delta(u_h)&\gtrsim \int_{\mathbb{R}^N}(|Du_h|+|D(U^h-u_h)|)^{p-2}|D(U^h-u_h)|^2dx\\&\gtrsim\int_{\mathbb{R}^N}|DU^h|^{p-2}|D(U^h-u_h)|^2dx\\&\gtrsim \lambda_h^{-\frac{p}{p-1}}+(h\lambda_h^{\frac{p}{N-p}})^2\int_\omega |u'(|y|)|^{p-2}a^2(|y|)dy\gtrsim h^{\alpha(p,N)}.\end{aligned}$$

For $1<p<2$ let us first recall the following inequality for quasi-norms in [4, Lemma 2.2]

$$\int_\Omega(|Du|+|Dv|)^{p-2}dx\geq\frac{\|D(u-v)\|^2_{L^p(\Omega)}}{(\|Du\|_{L^p(\Omega)}+\|Dv\|_{L^p(\Omega)})^{2-p}}.$$

Using that $\|DU^h\|_{L^p(\mathbb{R}^N)}\simeq 1$ we get

$$\begin{aligned}\delta(u_h)&\gtrsim\int_{\mathbb{R}^N}(|Du_h|+|D(U^h-u_h)|)^{p-2}|D(U_h-u_h)|^2dx\\&\gtrsim\int_{B_h^c}|DU^h|^pdx+\int_{B_h}(|Du_h|+|D(U^h-u_h)|)^{p-2}|D(U^h-u_h)|^2dx\\&\gtrsim\lambda_h^{-\frac{p}{p-1}}+\frac{\|D(U^h-u_h)\|^2_{L^p(B_h)}}{(\|DU^h\|_{L^p(B_h)}+\|D(U^h-u_h)\|_{L^p(B_h)})^{2-p}}\\&\gtrsim\lambda_h^{-\frac{p}{p-1}}+\|D(U^h-u_h)\|^2_{L^p(B_h)}.\end{aligned}$$

From Step II of the proof of Proposition 1.2 we know that $\|D(U^h-u_h)\|_{L^p(B_h)}\gtrsim h\lambda_h^{\frac{p}{N-p}}$, hence

$$\delta(u_h)\gtrsim\lambda_h^{-\frac{p}{p-1}}+(h\lambda_h^{\frac{p}{N-p}})^2\gtrsim h^{\alpha(p,N)}$$

and the proof is completed. □

6. Conclusions and perspectives

In this paper, we have established sharp convergence rates for the $P1$ finite element approximations of the Sobolev constant. Our approach is inspired by existing techniques for analyzing the deficit function in Sobolev inequalities in the continuous setting, but extends them through significant new developments. These advances are particularly crucial given that the Sobolev constant is not attained in bounded domains, necessitating a careful analysis of approximation rates for explicitly known minimizing sequences.

Moreover, achieving sharp convergence rates requires working with quasi-norms commonly used in the finite element analysis of the p-Laplacian, rather than with the standard $W^{1,p}(\mathbb{R}^N)$ norms. Our methodology builds on prior results concerning finite element approximation rates for Laplacian eigenvalues, adapting them to the context of the Sobolev constant.

The techniques developed here can also be adapted to analyze $P1$ finite element approximations of other fundamental constants in the theory of PDEs and mathematical physics, such as the Hardy constant, which is preliminarily discussed in [11]. This direction will be pursued in future work, as it requires a substantial refinement of the methodology introduced in this paper.

Finally, the questions explored here naturally extend to other settings, including the approximation of the Sobolev constant using different classes of finite element methods. However, addressing these extensions will require considerable additional work.

Acknowledgements. The authors thank Endre Süli for his valuable suggestion of employing the quasi-norms in [4] to sharpen the preliminary convergence results in Proposition 1.2 to Theorem 1.1, at the occasion of the Workshop "Modelling, PDE analysis and computational mathematics in materials science", Prague, September 2024.

L. I. Ignat was supported by a grant of the Ministry of Research, Innovation and Digitization, CCCDI - UEFISCDI, project number ROSUA-2024-0001, within PNCDI IV.

E. Zuazua was funded by the Alexander von Humboldt-Professorship program, the ModConFlex Marie Curie Action, HORIZON-MSCA-2021-dN-01, the COST Action MAT-DYN-NET, the Transregio 154 Project "Mathematical Modelling, Simulation and Optimization Using the Example of Gas Networks" of the DFG, grant "Hybrid Control and Estimation of Semi-Dissipative Systems: Analysis, Computation, and Machine Learning", AFOSR Proposal 24IOE027 and grants PID2020-112617GB-C22/AEI/10.13039/501100011033 and TED2021-131390B-I00/ AEI/10.13039/501100011033 of MINECO (Spain), and Madrid Government - UAM Agreement for the Excellence of the University Research Staff in the context of the V PRICIT (Regional Programme of Research and Technological Innovation).

References

[1] P. F. Antonietti and A. Pratelli. Finite element approximation of the Sobolev constant. *Numer. Math.*, 117(1):37–64, 2011.
[2] T. Aubin. Problèmes isopérimétriques et espaces de Sobolev. *J. Differential Geometry*, 11(4):573–598, 1976.
[3] I. Babuška and J. Osborn. Eigenvalue problems. In *Handbook of numerical analysis, Vol. II*, volume II of *Handb. Numer. Anal.*, pages 641–787. North-Holland, Amsterdam, 1991.
[4] J. W. Barrett and W. B. Liu. Finite element approximation of the p-Laplacian. *Math. Comp.*, 61(204):523–537, 1993.
[5] G. Bianchi and H. Egnell. A note on the Sobolev inequality. *J. Funct. Anal.*, 100(1):18–24, 1991.
[6] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[7] F. Chatelin. *Spectral approximation of linear operators*, volume 65 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.
[8] L. Chen, P. Sun, and J. Xu. Optimal anisotropic meshes for minimizing interpolation errors in L^p-norm. *Math. Comp.*, 76(257):179–204, 2007.
[9] A. Cianchi. A quantitative Sobolev inequality in BV. *J. Funct. Anal.*, 237(2):466–481, 2006.
[10] P. G. Ciarlet. *The finite element method for elliptic problems*, volume Vol. 4 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978.
[11] F. Della Pietra, G. Fantuzzi, L. I. Ignat, A. L. Masiello, G. Paoli, and E. Zuazua. Finite element approximation of the Hardy constant. *J. Convex Anal.*, 31(2):497–523, 2024.
[12] C. Ebmeyer and W. Liu. Quasi-norm interpolation error estimates for the piecewise linear finite element approximation of p-Laplacian problems. *Numer. Math.*, 100(2):233–258, 2005.
[13] H. Federer and W. H. Fleming. Normal and integral currents. *Ann. of Math. (2)*, 72:458–520, 1960.
[14] A. Figalli and R. Neumayer. Gradient stability for the Sobolev inequality: the case $p \geq 2$. *J. Eur. Math. Soc. (JEMS)*, 21(2):319–354, 2019.
[15] A. Figalli and Y. R.-Y. Zhang. Sharp gradient stability for the Sobolev inequality. *Duke Math. J.*, 171(12):2407–2459, 2022.
[16] N. Fusco, F. Maggi, and A. Pratelli. The sharp quantitative Sobolev inequality for functions of bounded variation. *J. Funct. Anal.*, 244(1):315–341, 2007.
[17] V. John. *Finite element methods for incompressible flow problems*, volume 51 of *Springer Series in Computational Mathematics*. Springer, Cham, 2016.
[18] V. G. Maz'ja. Classes of domains and imbedding theorems for function spaces. *Soviet Math. Dokl.*, 1:882–885, 1960.
[19] P.-A. Raviart and J.-M. Thomas. *Introduction à l'analyse numérique des équations aux dérivées partielles*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983.
[20] G. Talenti. Best constant in Sobolev inequality. *Ann. Mat. Pura Appl. (4)*, 110:353–372, 1976.

7. Appendix

7.1. **Piecewise linear approximations.** In this section we prove the two Lemmas in Section 2.

Proof of Lemma 2.1. We proceed as in [6, Th. 4.4.4, Chapter 4]. We first prove that for any element T and any $u \in W^{2,\infty}(\Omega)$ we have

$$\|D^\alpha(u - \mathcal{I}_T u)\|_{L^p(T)} \leq C h_T^{2-|\alpha|+\frac{N}{p}} \|D^2 u\|_{L^\infty(T)} \tag{7.1}$$

where $\mathcal{I}_T$ is the local linear interpolation operator on T. This gives immediately the case $p = \infty$. Summing over all the elements of $\mathcal{T}_h$ gives the desired estimate for $1 \leq p < \infty$.

To prove (7.1), we consider the case when $T = \hat{T}$ is the reference element, $h_T = 1$ and d its diameter. The general case then follows by a homogeneity argument and the quasi-uniform assumption on the mesh $\mathcal{T}_h$.

Writing $\mathcal{I}_{\hat{T}}$ in explicit form as in [6, Lemma 4.4.1, p. 105] we find that

$$\|D^\alpha \mathcal{I}_{\hat{T}} u\|_{C(\hat{T})} \leq C\|u\|_{C(\hat{T})}, |\alpha| \leq 1. \tag{7.2}$$

Note that here we refer to the reference finite element $\hat{T}$. So, the inequality is not affected by any power of the mesh size h. This will enter later, by scaling, when dealing with meshes of characteristic size h, as in the statement of the Lemma.

Consider $x_0 \in \hat{T}$ and the linear Taylor approximation

$$T^2_{x_0} u(x) = \sum_{|\alpha| \leq 1} \frac{1}{\alpha!} D^\alpha u(x_0)(x - x_0)^\alpha.$$

We emphasize that since we are working with regular functions $W^{2,\infty}(\Omega)$ (and therefore in $C^1(\bar{\Omega})$) instead of $W^{2,p}$, we do not need to use the averaged Taylor polynomial as in the classical framework [6, Section 4.1]. Since $T^2_{x_0} \in \mathbb{P}^1(\widehat{T})$ it follows that $\mathcal{I}_{\hat{T}} T^2_{x_0} = T^2_{x_0}$ and the reminder

$$(R^2u)(x) = u(x) - T^2_{x_0}u(x) = 2\sum_{|\alpha|=2}(x-x_0)^\alpha \int_0^1 \frac{1}{\alpha!} s D^\alpha u(x+s(x_0-x))ds$$

satisfies (see also [6, (4.3.5), p. 101])

$$|(R^2u)(x)| \le C(\text{diameter}(\hat{T}))^2 \|D^2u\|_{L^\infty(\hat{T})}. \tag{7.3}$$

It follows that, for any $|\alpha| \le 1$,

$$D^\alpha(u - \mathcal{I}_{\hat{T}}u) = D^\alpha(u - T^2_{x_0}u) + D^\alpha(T^2_{x_0} - \mathcal{I}_{\hat{T}}u) = D^\alpha(u - T^2_{x_0}u) + D^\alpha(\mathcal{I}_{\hat{T}}(T^2_{x_0}u - u)).$$

In view of estimates (7.2) and (7.3), for any $u \in W^{2,\infty}(\hat{T})$,

$$\begin{aligned} |D^\alpha(u - \mathcal{I}_{\hat{T}}u)(x)| &\le \|D^\alpha(u - T^2_{x_0}u)\|_{C(\hat{T})} + \|D^\alpha(\mathcal{I}_{\hat{T}}(T^2_{x_0}u - u))\|_{L^\infty(\hat{T})} \\ &\le \|R^2u\|_{L^\infty(\hat{T})} + \|T^2_{x_0}u - u\|_{L^\infty(\hat{T})} \le 2\|R^2u\|_{L^\infty(\hat{T})} \le C\|D^2u\|_{L^\infty(\hat{T})}. \end{aligned}$$

Integrating over $\hat{T}$ we get the desired result.

Let us now consider the case of a general element T and the affine transform B_T mapping $\hat{T}$ to T, and $\hat{u}$ in u. Using the previous estimate in $\hat{T}$ and the fact that all the elements are quasi-regular we obtain:

$$\begin{aligned} \int_T |D^\alpha_x(u - \mathcal{I}_T u)|^p dx &\le C h_T^{-|\alpha|p} |\det B_T| \int_{\hat{T}} |D^\alpha_{\hat{x}}(\hat{u} - \mathcal{I}_{\hat{T}}\hat{u})|^p d\hat{x} \\ &\le C h_T^{-|\alpha|p+N} \int_{\hat{T}} |D^\alpha_{\hat{x}}(\hat{u} - \mathcal{I}_{\hat{T}}\hat{u})|^p d\hat{x} \\ &\le C h_T^{-|\alpha|p+N} \|D^2_{\hat{x}}\hat{u}\|^p_{L^\infty(\hat{T})} \le C h_T^{-|\alpha|p+N} (h_T^2 \|D^2_x u\|_{L^\infty(T)})^p \\ &= C h_T^{(2-|\alpha|)p+N} \|D^2_x u\|^p_{L^\infty(T)}. \end{aligned}$$

In the case of estimate (2.5) we use (7.2) to obtain

$$\begin{aligned} \int_T |D^\alpha_x(\mathcal{I}_T u)|^p dx &\le C h_T^{-|\alpha|p} |\det B_T| \int_{\hat{T}} |D^\alpha_{\hat{x}}(\mathcal{I}_{\hat{T}}\hat{u})|^p d\hat{x} \le C h_T^{-|\alpha|p+N} \int_{\hat{T}} |D^\alpha_{\hat{x}}(\mathcal{I}_{\hat{T}}\hat{u})|^p d\hat{x} \\ &\le C h_T^{-|\alpha|p+N} \|\hat{u}\|^p_{C(\hat{T})} \le C h_T^{-|\alpha|p+N} \|u\|^p_{C(T)}. \end{aligned}$$

This shows that

$$\|D^\alpha I^h u\|_{L^p(T)} \lesssim h_T^{-|\alpha|+\frac{N}{p}} \|u\|_{L^\infty(T)},$$

and finishes the proof. □

Before proving Lemma 2.2 let us first consider the one-dimensional case.

Lemma 7.1. *Let $-\infty < a < b < \infty$ and $u \in C^2([a,b])$. For any $p \in (1,\infty)$ there exists a positive constant $C(p)$ such that*

$$\int_a^b |u'(r) - A|^p dr \ge C(p)(b-a)^{p+1} \inf_{r\in[a,b]} |u''(r)|^p, \ \forall\ A \in \mathbb{R}.$$

Proof. Consider the function $f(A) = \int_a^b |u'(r) - A|^p dr$. It is a convex function satisfying $\lim_{A\to\pm\infty} = \infty$. It attains its minimum at a point A_0 such that

$$\int_a^b |u'(r) - A_0|^{p-1}\,\mathrm{sgn}((u'(r) - A_0))dr = 0.$$

It follows that there exists $r_0 \in (a,b)$ such that $u'(r_0) = A_0$. Using that

$$|u'(r) - u'(r_0)| \geq |r - r_0| \inf_{[a,b]} |u''(r)|$$

we get the desired estimate

$$f(A) \geq \inf_{[a,b]} |u''(r)|^p \int_a^b |r - r_0|^p dr = C(p)(b-a)^{p+1} \inf_{r\in[a,b]} |u''(r)|^p.$$

This finishes the proof. □

Proof of Lemma 2.2. To fix the ideas we prove it for the direction x_1. For any $A_1 \in \mathbb{R}$ we have:

$$\int_T |\partial_{x_1} u - A_1|^p dx \geq C(p)\rho_T^{N+p} \min_{x\in T} |\partial_{x_1x_1} u(x)|^p.$$

Take a cube $C_{\rho_T} \subset B_{\rho_T} \subset T$. To simplify the presentation let us assume $C_{\rho_T} = [0,L]^N$. We write $x = (x_1, x')$ and apply the one-dimensional result in Lemma 7.1:

$$\begin{aligned}\int_T |\partial_{x_1} u - A_i|^p dx \geq \int_{C_{\rho_T}} |\partial_{x_1} u - A_i|^p dx &= \int_{[0,L]^{N-1}} \int_0^L |\partial_{x_1} u - A_i|^p dx_1 dx' \\ &\geq C(p)L^{N+p} \min_{x\in T} |\partial_{x_1x_1} u(x)|^p \geq C(p)\rho_T^{N+p} \min_{x\in T} |\partial_{x_1x_1} u(x)|^p.\end{aligned}$$

Let us now consider the general case. It is enough to prove that, for any $A \in \mathbb{R}^N$ and $\xi \in \mathbf{S}^{N-1}$, it holds

$$\int_{B_{\rho_T}} |Du - A|^p dx \geq C(p)\rho_T^{N+p} \min_{x\in\overline{\Omega}} |\xi^T D^2 u(x)\xi|^p.$$

Consider a rotation R of the unit sphere $\mathbf{S}^{N-1}$ such that $Re_1 = \xi$ and $v(y) = u(Ry)$. Using the previous step for v and the fact that, $Dv(y) = R^T Du(Ry)$, $D^2 v(y) = R^T D^2 u(Ry) R$ we get for any $A \in \mathbb{R}^N$ that

$$\begin{aligned}\min_{x\in B_{\rho_T}} |\xi^T D^2 u(x)\xi|^p &= \min_{x\in B_{\rho_T}} |e_1^T R^T D^2 u(x) Re_1|^p = \min_{x\in B_{\rho_T}} |e_1^T D^2 v(x) e_1|^p \\ &\leq \frac{1}{C(p)\rho_T^{N+p}} \int_{B_{\rho_T}} |Dv - A|^p dx = \frac{1}{C(p)\rho_T^{N+p}} \int_{B_{\rho_T}} |R^T(Du) - A|^p dx \\ &= \frac{1}{C(p)\rho_T^{N+p}} \int_{B_{\rho_T}} |Du - RA|^p dx.\end{aligned}$$

Replacing A by $R^T A$ gives the desired result. □

7.2. **An elementary inequality.** The following holds:

Lemma 7.2. *Let $q \in (1,\infty)$. Then, there exists positive constants A_q and B_q such that*

$$\big||a+b|^q - |a|^q - q|a|^{q-2}ab\big| \leq A_q|a|^{q-2}|b|^2 + B_q|b|^q,\ \forall a,b \in \mathbb{R}. \tag{7.4}$$

Furthermore $A_q = 0$ if $q \in (1,2]$.

On the other hand, for suitable A_q, B_q and C_q the following inequalities hold:

$$|x+y|^q \leq |x|^q + q|x|^{q-2}x\cdot y + A_q|x|^{q-2}|y|^2 + B_q|y|^q,\ \forall x,y \in \mathbb{R}^N \tag{7.5}$$

with $A_q = 0$ *when* $q \in (1,2]$, *and*

$$|x+y|^q \le |x|^q + q|x|^{q-2}x\cdot y + C_q \frac{(|x|+|y|)^q}{|x|^2+|y|^2}|y|^2,\ \forall x,y\in\mathbb{R}^N. \tag{7.6}$$

Proof. Some cases are treated in [14, Lemma 3.2]. For completeness we give here a short proof.

Let us begin with (7.4). By homogeneity we can reduce the proof to the following one:

$$\big||1+z|^q - 1 - qz\big| \le A_q|z|^2 + B_q|z|^q, \forall z\in\mathbb{R}.$$

It is clear that for $|z|>1$ the inequality is true with $A_q = 0$ independent of the value of q. When $|z|<1$ we have to consider the case when $z\simeq 0$. Taylor's expansion yields

$$|1+z|^q - 1 - qz = O(|z|^2) \le O(|z|^{\min\{2,q\}}),\ \ z\sim 0.$$

So, for $q<2$, one can choose $A_q = 0$.

Inequality (7.5) holds trivially when $|x| = 0$. Also, one can consider only the case when $|x|=1$ and after a rotation we can also assume $x=(1,0,\dots,0)$. It is then sufficient to prove that that

$$\Big((1+y_1)^2 + |y'|^2\Big)^{q/2} \le 1 + qy_1 + A_q|y|^2 + B_q|y|^q$$

for $y=(y_1,y')\in\mathbb{R}\times\mathbb{R}^{N-1}$.

When $q\le 2$ we easily have

$$\Big((1+y_1)^2 + |y'|^2\Big)^{q/2} \le |1+y_1|^q + |y'|^q \le 1+qy_1 + B_q|y_1|^q + |y'|^q \le 1 + qy_1 + \tilde{B}_q|y|^q.$$

For $q>2$ we distinguish cases $q/2\le 2$ and $q/2>2$. When $q/2\le 2$, in view of inequality (7.4),

$$\begin{aligned}\Big((1+y_1)^2 + |y'|^2\Big)^{q/2} &\le |1+y_1|^q + \frac{q}{2}|1+y_1|^{q-2}|y'|^2 + B_{q/2}|y'|^q\\ &\le |1+y_1|^q + C_q(|y|^2+|y|^q)\\ &\le 1+qy_1 + A_q|y_1|^2 + B_q|y_1|^q + C_q(|y|^2+|y|^q)\\ &\le 1+qy_1 + C_q'(|y|^2+|y|^q).\end{aligned}$$

When $q/2>2$, using again inequality (7.4), we have

$$\begin{aligned}\Big((1+y_1)^2 + |y'|^2\Big)^{q/2} &\le |1+y_1|^q + \frac{q}{2}|1+y_1|^{q-2}|y'|^2 + A_{q/2}|1+y_1|^{q-4}|y'|^4 + B_{q/2}|y'|^q\\ &\le |1+y_1|^q + C_q(|y|^2+|y|^q+|y|^4)\\ &\le 1+qy_1 + A_q|y_1|^2 + B_q|y_1|^q + C_q(|y|^2+|y|^q)\\ &\le 1+qy_1 + C_q'(|y|^2+|y|^q).\end{aligned}$$

Let us now consider the last inequality (7.6). For $q\ge 2$ it is an immediate consequence of the second one. When $q\le 2$ it remains to prove that

$$\Big((1+y_1)^2 + |y'|^2\Big)^{q/2} \le 1+qy_1 + C_q\frac{(1+|y|)^q}{1+|y|^2}|y|^2.$$

In this case we easily have

$$\Big((1+y_1)^2 + |y'|^2\Big)^{q/2} \le |1+y_1|^q + |y'|^q \le 1+qy_1 + B_q|y_1|^q + |y'|^q \le 1 + qy_1 + \tilde{B}_q|y|^q.$$

For $|y| \geq 1$ we get the desired estimate since $|y|^q \lesssim \frac{(1+|y|)^q}{1+|y|^2}|y|^2$. When $|y| \leq 1$ we use that $(1+x)^\alpha \leq 1+\alpha x$ for all $0<\alpha<1$ and $x>-1$:

$$\Big((1+y_1)^2+|y'|^2\Big)^{q/2} = (1+2y_1+y_1^2+|y'|^2)^{q/2} \leq 1+qy_1+\frac{q}{2}(y_1^2+|y'|^2) \leq 1+qy_1+\frac{q}{2}|y|^2$$
$$\lesssim \frac{(1+|y|)^q}{1+|y|^2}|y|^2.$$

The proof is now completed. □

7.3. **Proof of Lemma 4.3.** Let u be such that $\|u\|_{L^{p^*}(\mathbb{R}^N)}=1$ and

$$\delta(u)=\|Du\|_{L^p(\mathbb{R}^N)}-S(p,N)<S(p,N).$$

Then $S(p,N)\leq \|Du\|_{L^p(\mathbb{R}^N)}\leq 2S(p,N)$ and, using

$$|y^p-x^p|=|(y-x)\xi^{p-1}|\geq |x|^{p-1}|y-x|, \quad \forall |x|<|y|,$$

we get

$$\delta(u)=\|Du\|_{L^p(\mathbb{R}^N)}-S(p,N)\leq \frac{1}{(S(p,N))^{p-1}}(\|Du\|^p_{L^p(\mathbb{R}^N)}-(S(p,N))^p). \tag{7.7}$$

It suffices, therefore, to estimate the right hand side of the above inequality.

Let $v\in\mathcal{M}$ and set $\varepsilon=\|D(u-v)\|_{L^p(\mathbb{R}^N)}$. We write

$$u=v+\|Du-Dv\|_{L^p(\mathbb{R}^N)}\frac{u-v}{\|Du-Dv\|_{L^p(\mathbb{R}^N)}}:=v+\varepsilon\varphi,$$

where $\|D\varphi\|_{L^p(\mathbb{R}^N)}=1$. Also, $\|Dv\|_{L^p(\mathbb{R}^N)}\simeq \|Du\|_{L^p(\mathbb{R}^N)}\simeq S(p,N)\|v\|_{L^{p^*}(\mathbb{R}^N)}\simeq S(p,N)$. The two cases $1<p<2$ and $2\leq p<N$ are treated differently.

Case I. $2\leq p<N$**.** Inequality (7.6) shows that for any $q>1$ and $x,y\in\mathbb{R}^N$ the following holds

$$|x+y|^q\leq |x|^q+q|x|^{q-2}x\cdot y+C_q(|x|+|y|)^{q-2}|y|^2. \tag{7.8}$$

Thus we obtain

$$\|Du\|^p_{L^p(\mathbb{R}^N)}=\int_{\mathbb{R}^N}|D(v+\varepsilon\varphi)|^pdx$$
$$\leq \int_{\mathbb{R}^N}|Dv|^pdx+\varepsilon p|Dv|^{p-2}Dv\cdot D\varphi+C_p\varepsilon^2(|Dv|+\varepsilon|D\varphi|)^{p-2}|D\varphi|^2dx.$$

We use that for $p^*>1$ inequality $|1+x|^{p^*}\geq 1+p^*x$ holds or all $x\in\mathbb{R}$ to get

$$\|u\|^p_{L^{p^*}(\mathbb{R}^N)}\geq \Big(\int_{\mathbb{R}^N}(|v|^{p^*}+\varepsilon p^*|v|^{p^*-2}v\varphi)\,dx\Big)^{p/p^*}=\|v\|^p_{L^{p^*}}\Big(1+\varepsilon p^*\frac{\int_{\mathbb{R}^N}|v|^{p^*-2}v\varphi\,dx}{\|v\|^{p^*}_{L^{p^*}}}\Big)^{p/p^*}.$$

Since

$$\Big|\int_{\mathbb{R}^N}|v|^{p^*-2}v\varphi\,dx\Big|\leq \|v\|^{p^*-1}_{L^{p^*}(\mathbb{R}^N)}\|\varphi\|_{L^{p^*}(\mathbb{R}^N)}\lesssim 1$$

we obtain, by Taylor's expansion,

$$\|u\|^p_{L^{p^*}}\geq \|v\|^p_{L^{p^*}}\Big(1+\varepsilon p\frac{\int_{\mathbb{R}^N}|v|^{p^*-2}v\varphi\,dx}{\|v\|^p_{L^{p^*}}}-C\varepsilon^2\Big(\frac{\int_{\mathbb{R}^N}|v|^{p^*-2}v\varphi\,dx}{\|v\|^{p^*}_{L^{p^*}}}\Big)^2\Big).$$

Putting together the above two estimates and using that

$$\int_{\mathbb{R}^N}|Dv|^pdx=(S(p,N))^p\|v\|^p_{L^{p^*}(\mathbb{R}^N)}, \tag{7.9}$$

$$\int_{\mathbb{R}^N} |Dv|^{p-2}Dv \cdot D\varphi dx = (S(p,N))^p \|v\|^{p-p^*}_{L^{p^*}(\mathbb{R}^N)} \int_{\mathbb{R}^N} |v|^{p^*-2} v\varphi dx, \tag{7.10}$$

we get

$$\begin{aligned}\delta(u) &\lesssim \|Du\|^p_{L^p(\mathbb{R}^N)} - (S(p,N))^p \|u\|^p_{L^{p*}} \\ &= \varepsilon^2 \int_{\mathbb{R}^N} (|Dv| + \varepsilon|D\varphi|)^{p-2} |D\varphi|^2 dx + \varepsilon^2 \Big(\int_{\mathbb{R}^N} |v|^{p^*-2} v\varphi \, dx\Big)^2 \|v\|^{p-2p^*}_{L^{p^*}(\mathbb{R}^N)}.\end{aligned}$$

Hölder's inequality and [14, Lemma 3.3] give us

$$\Big(\int_{\mathbb{R}^N} |v|^{p^*-2} v\varphi dx\Big)^2 \le \Big(\int_{\mathbb{R}^N} |v|^{p^*-2} \varphi^2 \, dx\Big)\Big(\int_{\mathbb{R}^N} |v|^{p^*} \, dx\Big) \lesssim \|v\|^{p^*}_{L^{p^*}(\mathbb{R}^N)} \int_{\mathbb{R}^N} |Dv|^{p-2} |D\varphi|^2 \, dx.$$

This shows that

$$\begin{aligned}\delta(u) &\lesssim \varepsilon^2 \int_{\mathbb{R}^N} (|Dv| + \varepsilon|D\varphi|)^{p-2} |D\varphi|^2 dx + \varepsilon^2 \int_{\mathbb{R}^N} |Dv|^{p-2} |D\varphi|^2 \, dx \\ &= \int_{\mathbb{R}^N} (|Dv| + |D(u-v)|)^{p-2} |D(u-v)|^2 dx + \int_{\mathbb{R}^N} |Dv|^{p-2} |D(u-v)|^2 \, dx.\end{aligned}$$

This is exactly the estimate in Remark 4.4. Since $p \ge 2$ we immediately obtain (4.9).

Case II. $1 < p < 2$. In this setting,

$$\|u\|^{p^*}_{L^{p*}} = \int_{\mathbb{R}^N} |v + \varepsilon\varphi|^{p^*} \, dx \ge \int_{\mathbb{R}^N} |v|^{p^*} \, dx + \varepsilon p^* \int_{\mathbb{R}^N} |v|^{p^*-2} v\varphi \, dx$$

and, by inequality (7.6),

$$\begin{aligned}\|Du\|^{p^*}_{L^p(\mathbb{R}^N)} &= \Big(\int_{\mathbb{R}^N} |D(v+\varepsilon\varphi)|^p \, dx\Big)^{p^*/p} \\ &\le \Big(\int_{\mathbb{R}^N} |Dv|^p + \varepsilon p |Dv|^{p-2} Dv D\varphi + C_p \varepsilon^2 (|Dv| + \varepsilon|D\varphi|)^{p-2} |D\varphi|^2 \, dx\Big)^{p^*/p} \\ &= \|Dv\|^{p^*}_{L^p(\mathbb{R}^N)} \Big(1 + \frac{\varepsilon p \int_{\mathbb{R}^N} |Dv|^{p-2} Dv D\varphi + C_p \int_{\mathbb{R}^N} \varepsilon^2 (|Dv| + \varepsilon|D\varphi|)^{p-2} |D\varphi|^2}{\|Dv\|^p_{L^p(\mathbb{R}^N)}} \, dx\Big)^{p^*/p}.\end{aligned}$$

For small x we have

$$|1+x|^{p^*/p} \le 1 + \frac{p^*}{p} x + Cx^2$$

and this gives us

$$\begin{aligned}\|Du\|^{p^*}_{L^p(\mathbb{R}^N)} \le& \|Dv\|^{p^*}_{L^p(\mathbb{R}^N)} + \varepsilon p^* \|Dv\|^{p^*-p}_{L^p(\mathbb{R}^N)} \int_{\mathbb{R}^N} |Dv|^{p-2} Dv D\varphi \, dx \\ &+ \varepsilon^2 \frac{C_p p^*}{p} \|Dv\|^{p^*-p}_{L^p(\mathbb{R}^N)} \int_{\mathbb{R}^N} (|Dv| + \varepsilon|D\varphi|)^{p-2} |D\varphi|^2 dx \\ &+ C \|Dv\|^{p^*-2p}_{L^p(\mathbb{R}^N)} \Big(\varepsilon p \int_{\mathbb{R}^N} |Dv|^{p-2} Dv D\varphi \, dx + C_p \int_{\mathbb{R}^N} \varepsilon^2 (|Dv| + \varepsilon|D\varphi|)^{p-2} |D\varphi|^2 dx\Big)^2.\end{aligned}$$

Since for $p < 2$ we trivially have

$$\varepsilon^2 \int_{\mathbb{R}^N} (|Dv| + \varepsilon|D\varphi|)^{p-2} |D\varphi|^2 dx \le \varepsilon^p \int_{\mathbb{R}^N} |D\varphi|^p dx = \varepsilon^p.$$

Identities (7.9) and (7.10) imply that

$$\begin{aligned}\delta(u) &\lesssim \|Du\|^{p^*}_{L^p(\mathbb{R}^N)} - (S(p,N))^p\|u\|^{p^*}_{L^{p^*}}\\ &\lesssim \varepsilon^2\int_{\mathbb{R}^N}(|Dv|+\varepsilon|D\varphi|)^{p-2}|D\varphi|^2dx + \Big(\varepsilon\int_{\mathbb{R}^N}|Dv|^{p-2}DvD\varphi\, dx\Big)^2\\ &= \int_{\mathbb{R}^N}(|Dv|+|D(u-v)|)^{p-2}|D(u-v)|^2dx + \Big(\int_{\mathbb{R}^N}|Dv|^{p-1}|D(u-v)|\, dx\Big)^2\end{aligned}$$

which finishes the proof.

7.4. **Proof of Corollary 4.5.** Let us consider $u\in \dot W^{1,p}$ with $\|Du\|_{L^p(\mathbb{R}^N)}=1$ such that $d(u,\mathcal{M})=\varepsilon$ with ε a small positive parameter, so that there exists $v=v_u\in\mathcal{M}$ such that $d(u,\mathcal{M})=\|D(u-v)\|_{L^p(\mathbb{R}^N)}=\varepsilon$. Since $v\in\mathcal{M}$, $\|u\|_{L^{p^*}(\mathbb{R}^N)}\simeq\|v\|_{L^{p^*}(\mathbb{R}^N)}$ and $\delta(u)$ is that Lemma 4.3 applies. For $p\geq 2$ and $\varepsilon<1$ we have

$$\begin{aligned}\delta(u) &\lesssim \int_{\mathbb{R}^n}(|Du|+|Dv|)^{p-2}|D(u-v)|^2 \lesssim \int_{\mathbb{R}^n}(|D(u-v)|+|Du|)^{p-2}|D(u-v)|^2\\ &\lesssim \|D(u-v)\|^p_{L^p(\mathbb{R}^N)} + \|Du\|^{p-2}_{L^p(\mathbb{R}^N)}\|D(u-v)\|^2_{L^p(\mathbb{R}^N)}\\ &\lesssim \|D(u-v)\|^2_{L^p(\mathbb{R}^N)}.\end{aligned}$$

A similar argument works for $1<p<2$ since $\|Dv\|_{L^p(\mathbb{R}^N)}\leq 2$ for small ε:

$$\begin{aligned}\delta(u) &\lesssim \int_{\mathbb{R}^N}(|Dv|+|D(u-v)|)^{p-2}|D(u-v)|^2dx + \Big(\int_{\mathbb{R}^N}|Dv|^{p-1}|D(u-v)|dx\Big)^2\\ &\lesssim \int_{\mathbb{R}^N}|D(u-v)|^pdx + \Big(\int_{\mathbb{R}^N}|Dv|^pdx\Big)^{2/p'}\Big(\int_{\mathbb{R}^N}|D(u-v)|^pdx\Big)^{2/p}\\ &\lesssim \|D(u-v)\|^p_{L^p(\mathbb{R}^N)} + \|D(u-v)\|^2_{L^p(\mathbb{R}^N)} \lesssim \|D(u-v)\|^p_{L^p(\mathbb{R}^N)}.\end{aligned}$$

This finishes the proof.

7.5. **Proof of Lemma 4.7.** Let us assume, without loss of generality, that $x_0=(x_{01},0,\dots,0)$ with $|x_{01}|<1$. Using the explicit form of $U_{\lambda,x_0}(x)$, and denoting

$$A_\lambda=\{(y_1,y')\in\mathbb{R}\times\mathbb{R}^{N-1},\ |(y_1+\lambda^{\frac{p}{N-p}}x_{01},y')|\leq\lambda^{\frac{p}{N-p}}\}$$

we obtain, by an elementary change of variables,

$$\int_{|x|<1}|DU_{\lambda,x_0}|^pdx=\int_{A_\lambda}|DU(y)|^pdy=C_N\int_{A_\lambda}\frac{|y|^{\frac{p}{p-1}}dy}{(1+|y|^{\frac{p}{p-1}})^N}.$$

Observe that, since $|x_{01}|<1$, $A_\lambda\subset\{y\in\mathbb{R}^N,\ |y|\leq 2\lambda^{\frac{p}{N-p}}\}$. Then the upper bound in (4.14) holds for all $\lambda>0$:

$$\int_{|x|<1}|DU_{\lambda,x_0}|^pdx\lesssim\int_{|y|\leq 2\lambda^{\frac{p}{N-p}}}\frac{|y|^{\frac{p}{p-1}}dy}{(1+|y|^{\frac{p}{p-1}})^N}\lesssim\int_0^{2\lambda^{\frac{p}{N-p}}}\frac{r^{N-1+\frac{p}{p-1}}dr}{(1+r^{\frac{p}{p-1}})^N}.$$

For the lower bound in (4.14) observe that for $|x_{01}|\leq 1/2$, there is always a ball centered at the origin included in A_λ, $\{|y|<\lambda^{\frac{p}{N-p}}/2\}\subset A_\lambda$. Then

$$\int_{|x|<1}|DU_{\lambda,x_0}|^pdx\gtrsim\int_{|y|<\lambda^{\frac{p}{N-p}}/2}\frac{|y|^{\frac{p}{p-1}}dy}{(1+|y|^{\frac{p}{p-1}})^N}\gtrsim\lambda^{\frac{p}{N-p}(N+\frac{p}{p-1})},\ \forall\,\lambda<1.$$

Let us now consider the case when $1/2 < |x_{01}| < 1$ and choose the particular case $1/2 < x_{01} < 1$. Observe that

$$y \in A_\lambda \iff |y'|^2 + y_1^2 + 2y_1\lambda^{\frac{p}{N-p}}x_{01} \le \lambda^{\frac{2p}{N-p}}(1 - x_{01}^2)$$

and then, for λ small enough,

$$B_\lambda := \{(y_1, y') \in \mathbb{R} \times \mathbb{R}^{N-1}, |y'|^2 + y_1^2 + 4y_1\lambda^{\frac{p}{N-p}} \le 0,\ y_1 \le 0\} \subset A_\lambda \subset \{|y| < 1\}.$$

Moreover, $B_\lambda = \lambda^{\frac{p}{N-p}} B_1$ and then

$$\int_{B_\lambda} \frac{|y|^{\frac{p}{p-1}}dy}{(1 + |y|^{\frac{p}{p-1}})^N} \ge \frac{1}{2^N} \int_{B_\lambda} |y|^{\frac{p}{p-1}}dy \simeq \lambda^{\frac{p}{N-p}(N+\frac{p}{p-1})} \int_{B_1} |y|^{\frac{p}{p-1}}dy.$$

The proof of estimate (4.14) is finished.

Let us now consider the inequalities in (4.15). They are reduced to estimate the minimizers outside the unit ball

$$\int_{|x|>1} |DU_{\lambda,x_0}|^p dx = C_N \int_{A_\lambda^c} \frac{|y|^{\frac{p}{p-1}}dy}{(1 + |y|^{\frac{p}{p-1}})^N}.$$

Using that for $\lambda > 1$, $\{y \in \mathbb{R}^N, |y| \ge 2\lambda^{\frac{p}{N-p}}\} \subset A_\lambda^c$ we have

$$\int_{A_\lambda^c} \frac{|y|^{\frac{p}{p-1}}dy}{(1 + |y|^{\frac{p}{p-1}})^N} \ge \int_{|y|\ge 2\lambda^{\frac{p}{N-p}}} \frac{|y|^{\frac{p}{p-1}}dy}{(1 + |y|^{\frac{p}{p-1}})^N} \ge \int_{|y|\ge 2\lambda^{\frac{p}{N-p}}} |y|^{-\frac{p(N-1)}{p-1}}dy \simeq \lambda^{-\frac{p}{p-1}}.$$

Also, observe that $A_\lambda = \lambda^{\frac{p}{N-p}} A_1$ and any $y \in A_1^c$, i.e. $|y + (x_{01}, 0)| \ge 1$, satisfies $|y| \ge 1 - |x_{01}|$. Hence

$$\begin{aligned}
\int_{A_\lambda^c} \frac{|y|^{\frac{p}{p-1}}dy}{(1 + |y|^{\frac{p}{p-1}})^N} &= \lambda^{\frac{p}{N-p}(N+\frac{p}{p-1})} \int_{A_1^c} \frac{|y|^{\frac{p}{p-1}}dy}{(1 + \lambda^{\frac{p^2}{(N-p)(p-1)}}|y|^{\frac{p}{p-1}})^N} \\
&\le \lambda^{\frac{p}{N-p}(N+\frac{p}{p-1})} \int_{|y|>1-|x_{01}|} \frac{|y|^{\frac{p}{p-1}}dy}{(1 + \lambda^{\frac{p^2}{(N-p)(p-1)}}|y|^{\frac{p}{p-1}})^N} \\
&\le \lambda^{-\frac{p}{p-1}} \int_{|y|\ge 1-|x_{01}|} |y|^{-\frac{p}{p-1}(N-1)}dy = \frac{\lambda^{-\frac{p}{p-1}}}{(1 - |x_{01}|)^{\frac{N-p}{p-1}}}.
\end{aligned}$$

This proves (4.15) and finishes the proof.

7.6. **Proof of Lemma 4.9.** Using a translation and a scaling argument it is sufficient to consider the case $x_0 = 0$ and $\lambda = 1$. For the radially symmetric function $U(x) = u(|x|)$ we have

$$\partial_{x_i x_j} U(x) = u''(r)\frac{x_i x_j}{r^2} + u'(r)\frac{\delta_{ij}r^2 - x_i x_j}{r^3}$$

and the eigenvalues of the Hessian matrix are $\lambda_k(D^2(x)) \in \{u''(r), \frac{u'(r)}{r}\}$.

Explicit computation show that

$$u_0''(r) = \frac{c_{N,p}}{p-1}(1 + r^{\frac{p}{p-1}})^{-\frac{N}{p}-1}(1 - (N-1)r^{\frac{p}{p-1}})r^{\frac{2-p}{p-1}},$$

$$u_0'(r) = c_{N,p} r^{\frac{1}{p-1}}(1 + r^{\frac{p}{p-1}})^{-\frac{N}{p}}.$$

Thus

$$\Delta U(x) = \frac{c_{N,p}}{p-1} r^{\frac{2-p}{p-1}} \Big(1 + r^{\frac{p}{p-1}}\Big)^{-\frac{N}{p}-1} \Big[(p-2)(N-1)(r^{\frac{p}{p-1}} + 1) + N - 2\Big].$$

and for any $1 \leq i, j \leq n$ we have

$$|\partial_{x_i x_j} U(x)| \lesssim |u''(r)| + \frac{|u'(r)|}{r} \lesssim r^{\frac{2-p}{p-1}} \Big(1 + r^{\frac{p}{p-1}}\Big)^{-\frac{N}{p}}.$$

It implies that under the assumption $1 < p < N$, $D^2 U \in L^p(\mathbb{R}^N)$. Indeed,

$$\int_{\mathbb{R}^N} |D^2 U(x)|^p dx \lesssim \int_0^\infty r^{N-1} r^{\frac{(2-p)p}{p-1}} \Big(1 + r^{\frac{p}{p-1}}\Big)^{-N} dr < \infty$$

since

$$N + \frac{p(2-p)}{p-1} - \frac{Np}{p-1} < 0 \Leftrightarrow p > 1$$

and

$$N + \frac{p(2-p)}{p-1} > 0 \Leftrightarrow p \in \Big(\frac{N+2-\sqrt{N^2+4}}{2}, \frac{N+2+\sqrt{N^2+4}}{2}\Big) \supseteq (1, N).$$

A scaling argument proves (4.16) and

$$\int_{|x|<1} |D^2 U_{\lambda, x_0}(x)|^p dx \lesssim \lambda^{\frac{p^2}{N-p}}.$$

Let us now consider the lower bounds. Explicit computations show that

$$\Delta U(x) = \frac{c_{N,p}}{p-1} r^{\frac{2-p}{p-1}} \Big(1 + r^{\frac{p}{p-1}}\Big)^{-\frac{N}{p}-1} \Big[(p-2)(N-1) r^{\frac{p}{p-1}} + 1 + (p-1)(N-1)\Big].$$

For $p \geq 2$ we have

$$|\Delta U(x)| \geq \frac{c_{N,p}(p-2)(N-1)}{p-1} r^{\frac{2-p}{p-1}} \Big(1 + r^{\frac{p}{p-1}}\Big)^{-\frac{N}{p}-1} = A_{N,p} r^{\frac{2-p}{p-1}} \Big(1 + r^{\frac{p}{p-1}}\Big)^{-\frac{N}{p}-1}.$$

The sets

$$\Gamma_k = \Big\{x \in \mathbb{R}^N : |\partial_{x_k x_k} u(x)| > \frac{A_{N,p}}{2N} r^{\frac{2-p}{p-1}} \Big(1 + r^{\frac{p}{p-1}}\Big)^{-\frac{N}{p}-1},\ k = 1, \ldots, N\Big\}$$

cover the space $\mathbb{R}^N$ and then the conclusion holds.

When $1 < p < 2$ the proof is more delicate since there exists a set $\{x : |x|^{\frac{p}{p-1}} = 1/(N-1)\}$ where $u''(r) = 0$ and then $\lambda_1(D^2 U(x)) = 0$. Also, there is a point $x = 0$ where $u'(r)/r = \lambda_1(D^2 U(x)) = 0$. However, since the Hessian is non-degenerate, we can cover these sets with cylinders in which there exist directions on $\mathbf{S}^1$ along which the associated quadratic form admits a uniform lower bound.

Let us take a point $x \in \{x : |x|^{\frac{p}{p-1}} = 1/(N-1)\}$. For simplicity assume $x = (x_0, 0') \in \mathbb{R} \times \mathbb{R}^{N-1}$. Choosing $\varepsilon = 1/100$, there exists δ_0 such that

$$|\lambda_1(H(x))| = |u''(|x||) < \varepsilon a(|x|)$$

and

$$|\lambda_2(H(x))| = |u'(|x|)/|x|| > a(|x|)$$

for all $x = (x_1, x')$ in the cylinder $C_{\delta_0} = \{|x_1 - x_0| < \delta_0, |x' - 0'| < \delta_0\}$. Using that

$$|(\xi_1, 0')^T H(x)(\xi_1, 0')| = |\xi_1^2||u''(|x|)| < \varepsilon a(|x|),$$

and

$$|(0, \xi')^T H(x)(0, \xi')| = |\xi'|^2 \frac{|u'(|x|)|}{|x|} > a(|x|),$$

there exists α_ε such that for all $\xi \in \{\xi \in \mathbf{S}^1, |\xi_1| < \alpha_\varepsilon |\xi'|\}$ it holds

$$|\xi^T H(x)\xi| > \frac{1}{2} a(|x|), \forall x \in C_{\delta_0}.$$

Covering the set $x \in \{x : |x|^{\frac{p}{p-1}} = 1/(N-1)\}$ with a finite number of cylinders we obtain the desired property. A similar argument works when $|x| = 0$.

7.7. **Proof of Lemma 4.10.** For simplicity we write U_λ instead of $U_{\lambda,0}$. We split all the integrals over B_h and B_h^c. Since functions in V_h vanish outside B_h, there is no contribution from this region. In view of the way the polyhedral domain B_h is constructed, $B_h^c \subset \{|x| > 1-h\}$. The same arguments as in the proof of (4.15) lead to

$$\int_{B_h^c} |DU_\lambda|^p dx \leq \int_{|x|>1-h} |DU_\lambda|^p dx \leq \int_{|x|>1/2} |DU_\lambda|^p dx \lesssim \lambda^{-\frac{p}{p-1}}.$$

It remains to estimate the integrals in B_h. We proceed in several steps.

Step I. Proof of (4.19). Let us denote

$$I_h = \int_{B_h} |DU_\lambda - Du_h|^p dx.$$

When $p > N/2$ classical estimates for the linear interpolator [6, Theorem 4.4.20, p. 108] and the result in (4.17) show that

$$I_h \leq h^p \|D^2 U_\lambda\|^p_{L^p(B_h)} \lesssim (h\lambda^{\frac{p}{N-p}})^p,$$

which finishes the proof of (4.19).

Let us now consider the case $1 < p \leq N/2$, in which we will make use of the estimate provided in Lemma 2.1.We estimate differently the terms in the triangles T which are outside the ball $\{|x| < 2h\}$. If $x \in T \in \mathcal{T}_h$ with $T \subset \{|x| > 2h\}$, we have $|x| > 2h > 2h_T$, $|x| \leq \inf_{x\in T} |x| + h_T < \inf_{x \in T} |x| + |x|/2$ and then $|x| \leq \sup_{x\in T} |x| \leq 2 \inf_{x\in T} |x| \leq 2|x|$. In this case estimate (4.16) gives us

$$|T| \|D^2 U_\lambda\|^p_{L^\infty(T)} \lesssim |T| \sup_{x\in T} \lambda^{\frac{p(N+p)}{N-p}} a(\lambda^{\frac{p}{N-p}}|x|) \lesssim \lambda^{\frac{p(N+p)}{N-p}} \int_T a^p(\lambda^{\frac{p}{N-p}}|x|) dx.$$

Using Lemma 2.1 it follows that

$$\begin{aligned} I_{1h} := \sum_{T\in\mathcal{T}_h, T\subset\{|x|>2h\}} \int_T |DU_\lambda - Du_h|^p dx &\leq h^p \sum_{T\in\mathcal{T}_h, T\subset\{|x|>2h\}} |T| \|D^2 U_\lambda\|^p_{L^\infty(T)} \\ &\lesssim h^p \lambda^{\frac{p(N+p)}{N-p}} \int_{2h<|x|<1} a^p(\lambda^{\frac{p}{N-p}}|x|) dx = h^p \lambda^{\frac{p^2}{N-p}} \int_0^\infty a^p(r) r^{N-1} dr \lesssim h^p \lambda^{\frac{p^2}{N-p}}. \end{aligned}$$

The triangles $T \in \mathcal{T}_h$ that are not entirely contained outside the ball $\{|x| < 2h\}$ are included within the larger ball $\{|x| < 3h\}$. We denote by I_{2h} the part of the integral corresponding to these triangles. We recall the following estimates for the linear interpolator (2.5)

$$\|DI^h u\|_{L^p(T)} \lesssim h_T^{-1+\frac{N}{p}} \|u\|_{L^\infty(T)}, \ p \geq 1.$$

From the construction of B_h we have $U_{\lambda|\partial B_h} = U_{\lambda|\partial B} = C_\lambda$. It follows that $Du_h = DI^h(U_\lambda - (U_\lambda)_{|\partial B_h}) = DI^h(U_\lambda)$ and thus

$$\begin{aligned} I_{2h} &\leq \sum_{T\in\mathcal{T}_h, T\subset\{|x|<3h\}} \int_T (|DU_\lambda|^p + |Du_h|^p)dx \lesssim \int_{|x|<3h} |DU_\lambda|^p dx + h^{N-p}\|U_\lambda\|^p_{L^\infty(|x|<3h)} \\ &\lesssim \int_0^{3h\lambda^{p/(N-p)}} |u_0'(s)|^p s^{N-1} ds + h^{N-p}\lambda^p \|U\|^p_{L^\infty(\mathbb{R}^N)} \\ &\lesssim (h\lambda^{\frac{p}{N-p}})^{N+\frac{p}{p-1}} + (h\lambda^{\frac{p}{N-p}})^{N-p} \lesssim (h\lambda^{\frac{p}{N-p}})^{N-p}. \end{aligned}$$

Summing with the estimate for I_{1h} and using that $1 < p \leq N/2$ we get the desired result

$$I_h \lesssim (h\lambda^{\frac{p}{N-p}})^p + (h\lambda^{\frac{p}{N-p}})^{N-p} \lesssim (h\lambda^{\frac{p}{N-p}})^p.$$

Step II. Proof of (4.20)**.** Let us first consider the case $p \geq 2$. For $p > N/2$ we use the error estimates for the quasi-norms in [12, Th. 3.1] to get

$$\begin{aligned} I_h &:= \int_{B_h} (|DU_\lambda| + |DU_\lambda - Du_h|)^{p-2}|DU_\lambda - Du_h|^2 dx \\ &\lesssim h^2 \int_{B_h} (|DU_\lambda| + h|D^2U_\lambda|)^{p-2}|D^2U_\lambda|^2 dx \\ &\lesssim h^2 \int_{B_h} |DU_\lambda|^{p-2}|D^2U_\lambda|^2 dx + h^p \int_{B_h} |D^2U_\lambda|^p dx \\ &\leq (h\lambda^{\frac{p}{N-p}})^2 \int_0^\infty |u'(r)|^{p-2}a^2(r)r^{N-1}dr + h^p\lambda^{\frac{p^2}{N-p}} \int_0^\infty a^p(r)r^{N-1}dr \\ &\lesssim (h\lambda^{\frac{p}{N-p}})^2 + (h\lambda^{\frac{p}{N-p}})^p \lesssim (h\lambda^{\frac{p}{N-p}})^2. \end{aligned}$$

Let us consider the case $2 \leq p \leq N/2$. We split $I_h = I_{1h} + I_{2h}$ as in Step I. For the first term we use similar arguments as in Step I:

$$\begin{aligned} I_{1h} &:= \sum_{T\in\mathcal{T}_h, T\subset\{|x|>2h\}} \int_T (|DU_\lambda| + |DU_\lambda - Du_h|)^{p-2}|DU_\lambda - Du_h|^2 dx \\ &\leq \sum_{T\in\mathcal{T}_h, T\subset\{|x|>2h\}} \int_T (|DU_\lambda|^{p-2}|DU_\lambda - Du_h|^2 + |DU_\lambda - Du_h|^p) dx \\ &\leq h^2 \sum_{T\in\mathcal{T}_h, T\subset\{|x|>2h\}} \|DU_\lambda\|^{p-2}_{L^\infty(T)} \int_T |DU_\lambda - Du_h|^2 dx + h^p\lambda^{\frac{p^2}{N-p}} \\ &\lesssim h^p\lambda^{\frac{p^2}{N-p}} + h^2 \int_{|x|>2h} |DU_\lambda|^{p-2}|D^2U_\lambda|^2 dx \\ &\leq (h\lambda^{\frac{p}{N-p}})^p + (h\lambda^{\frac{p}{N-p}})^2 \int_0^\infty |u'(r)|^{p-2}a^2(r)r^{N-1}dr \lesssim (h\lambda^{\frac{p}{N-p}})^2. \end{aligned}$$

For the term I_{2h} we proceed as in the Step I since $N - p \geq p \geq 2$:

$$\begin{aligned} I_{2h} &\leq \sum_{T\in\mathcal{T}_h, T\subset\{|x|<3h\}} \int_T (|DU_\lambda| + |DU_\lambda - Du_h|)^{p-2}|DU_\lambda - Du_h|^2 dx \\ &\leq \sum_{T\in\mathcal{T}_h, T\subset\{|x|<3h\}} \int_T (|DU_\lambda|^p + |Du_h|^p) dx \lesssim (h\lambda^{\frac{p}{N-p}})^{N-p} \leq (\lambda^{\frac{p}{N-p}})^2. \end{aligned}$$

Let us now consider the case $1 < p < 2$:

$$\begin{aligned}
I_{1h} &= \sum_{T\in\mathcal{T}_h, T\subset\{|x|>2h\}} \int_T (|DU_\lambda| + |DU_\lambda - Du_h|)^{p-2}|DU_\lambda - Du_h|^2 dx \\
&\le \sum_{T\in\mathcal{T}_h, T\subset\{|x|>2h\}} h^2 \|D^2U_\lambda\|^2_{L^\infty(T)} \int_T |DU_\lambda|^{p-2}dx \\
&\le h^2 \int_{|x|>2h} (\lambda^{\frac{N}{N-p}})^{p-2}|u'(\lambda^{\frac{p}{N-p}}|x|)|^{p-2}\lambda^{\frac{2(N+p)}{N-p}}|a(\lambda^{\frac{p}{N-p}}|x|)|^2 dx \\
&\le (h\lambda^{\frac{p}{N-p}})^2 \int_0^\infty |u'(r)|^{p-2}a^2(r)r^{N-1}dr \lesssim (h\lambda^{\frac{p}{N-p}})^2.
\end{aligned}$$

We use that for $p \le 2$ function a is uniformly bounded near the origin to get

$$\begin{aligned}
I_{2h} &\le \sum_{T\in\mathcal{T}_h, T\subset\{|x|<3h\}} \int_T (|DU_\lambda| + |DU_\lambda - Du_h|)^{p-2}|DU_\lambda - Du_h|^2 dx \\
&\le \sum_{T\in\mathcal{T}_h, T\subset\{|x|<3h\}} \int_T |DU_\lambda - Du_h|^p dx \le h^p \sum_{T\in\mathcal{T}_h, T\subset\{|x|<3h\}} |T|\|D^2U_\lambda\|^p_{L^\infty(T)} \\
&\le h^{p+N}\lambda^{\frac{(N+p)p}{N-p}} a^p(\lambda^{\frac{p}{N-p}}h) \lesssim (h\lambda^{\frac{p}{N-p}})^{N+p}.
\end{aligned}$$

Remark 7.3. At this point it is important top observe that, even if the following inequality holds

$$\begin{aligned}
I_{2h} &\le \sum_{T\in\mathcal{T}_h, T\subset\{|x|<3h\}} \int_T (|DU_\lambda| + |DU_\lambda - Du_h|)^{p-2}|DU_\lambda - Du_h|^2 dx \\
&\le \sum_{T\in\mathcal{T}_h, T\subset\{|x|<3h\}} \int_T |DU_\lambda|^{p-2}|DU_\lambda - Du_h|^2 dx,
\end{aligned}$$

it is useless in practice since the last integral diverges over the triangle T containing the origin. This is so as $|DU_\lambda|^{p-2}$ fails to be integrable in a neighborhood of the origin, in contrast to $|DU_\lambda|^p$, which remains integrable. This is why the estimate in (4.11) had to be improved to the one in (4.9).

Step III. Proof of (4.21). With similar notations as in the previous steps

$$\begin{aligned}
I_{1h} &:= \sum_{T\in\mathcal{T}_h, T\subset\{|x|>2h\}} \int_T |DU_\lambda|^{p-1}|DU_\lambda - Du_h| dx \\
&\le h \sum_{T\in\mathcal{T}_h, T\subset\{|x|>2h\}} \|D^2U_\lambda\|_{L^\infty(T)} \int_T |DU_\lambda|^{p-1}dx \\
&\lesssim h(\lambda^{1+\frac{p}{N-p}})^{p-1}\lambda^{\frac{(N+p)}{N-p}} \int_{2h<|x|<1} |u'(\lambda^{\frac{p}{N-p}}|x|)|^{p-1}a(\lambda^{\frac{p}{N-p}}|x|)dx \\
&\le h\lambda^{\frac{N(p-1)}{N-p}}\lambda^{\frac{(N+p-pN)}{N-p}} \int_0^\infty |u'(r)|^{p-1}a(r)r^{N-1}dr \lesssim h\lambda^{\frac{p}{N-p}}.
\end{aligned}$$

For the second term I_{2h} we proceed as before but we cannot compare $\|D^2u_\lambda\|^p_{L^\infty(T)}$ with its integral. Instead, since we are in the ball $\{|x| < 3h\}$ and $p < 2$ we can compare it with the values at $|x| = h$ and use that $a(r) \lesssim r^{\frac{2-p}{p-1}}$ for $0 < r < 1$:

$$\begin{aligned}
I_{2h} &\le h \sum_{T\in\mathcal{T}_h, T\subset\{|x|<3h\}} \int_T |DU_\lambda|^{p-1} \|D^2 U_\lambda\|_{L^\infty(T)} \\
&\le h\lambda^{\frac{(N+p)}{N-p}} a(\lambda^{\frac{p}{N-p}} h)\lambda^{\frac{N(p-1)}{N-p}} \int_{|x|<3h} |u'(\lambda^{\frac{p}{N-p}}|x|)|^{p-1} dx \\
&\le h\lambda^{\frac{(p+pN)}{N-p}} (\lambda^{\frac{p}{N-p}} h)^{\frac{2-p}{p-1}} \lambda^{\frac{-pN}{N-p}} \int_0^\infty |u'(r)|^{p-1} r^{N-1} dr \\
&\le (h\lambda^{\frac{p}{N-p}})^{\frac{1}{p-1}} \lesssim h\lambda^{\frac{p}{N-p}}.
\end{aligned}$$

The proof is now complete.

[1] Institute of Mathematics "Simion Stoilow" of the Romanian Academy, 21 Calea Grivitei Street, 010702 Bucharest, Romania.
[2] National University of Science and Technology Politehnica Bucharest, 313 Splaiul Independenţei, 060042 Bucharest, Romania.
[3] Academy of Romanian Scientists, Ilfov Street, no. 3, Bucharest, Romania.
E-mail address: liviu.ignat@gmail.com

[1] Friedrich-Alexander-Universität Erlangen-Nürnberg, Department of Mathematics, Chair for Dynamics, Control, Machine Learning and Numerics (Alexander von Humboldt Professorship), Cauerstr. 11, 91058 Erlangen, Germany.
[2] Chair of Computational Mathematics, Deusto University, 48007 Bilbao, Basque Country, Spain.
[3] Universidad Autónoma de Madrid, Departamento de Matemáticas, Ciudad Universitaria de Cantoblanco, 28049 Madrid, Spain.
E-mail address, Corresponding author: enrique.zuazua@fau.de