ASC Report No. 38/2014

Local high-order regularization and applications to hp-methods (extended version)

M. Karkulik, J.M. Melenk

Institute for Analysis and Scientific Computing — Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-05-6

Most recent ASC Reports

37/2014 *P. Amodio, T. Levitina, G. Settanni, E. Weinmüller*
Whispering gallery modes in oblate spheroidal cavities: calculations with a variable stepsize

36/2014 *J. Burkotova, I. Rachunkova, S. Stanek and E. Weinmüller*
Analytical and numerical treatment of singular linear BVPs with unsmooth inhomogeneity

35/2014 *E. Weinmüller*
Collocation - a powerful tool for solving singular ODEs and DAEs

34/2014 *M. Feischl, T. Führer, G. Gantner, A. Haberl, D. Praetorius*
Adaptive boundary element methods for optimal convergence of point errors

33/2014 *M. Halla, L. Nannen*
Hardy space infinite elements for time-harmonic two-dimensional elastic waveguide problems

32/2014 *A. Feichtinger and E. Weinmüller*
Numerical treatment of models from applications using BVPSUITE

31/2014 *C. Abert, M. Ruggeri, F. Bruckner, C. Vogler, G. Hrkac, D. Praetorius, and D. Suess*
Self-consistent micromagnetic simulations including spin-diffusion effects

30/2014 *J. Schöberl*
C++11 Implementation of Finite Elements in NGSolve

29/2014 *A. Arnold and J. Erb*
Sharp entropy decay for hypocoercive and non-symmetric Fokker-Planck equations with linear drift

28/2014 *G. Kitzler and J. Schöberl*
A high order space momentum discontinuous Galerkin method for the Boltzmann equation

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: admin@asc.tuwien.ac.at
WWW: http://www.asc.tuwien.ac.at
FAX: +43-1-58801-10196

ISBN 978-3-902627-05-6

Local high-order regularization and applications to hp-methods (extended version) *

M. Karkulik[†] J.M. Melenk[‡]

November 20, 2014

Abstract

We develop a regularization operator based on smoothing on a locally defined length scale. This operator is defined on L^1 and has approximation properties that are given by the local regularity of the function it is applied to and the local length scale. Additionally, the regularized function satisfies inverse estimates commensurate with the approximation orders. By combining this operator with a classical hp-interpolation operator, we obtain an hp-Clément type quasi-interpolation operator, i.e., an operator that requires minimal smoothness of the function to be approximated but has the expected approximation properties in terms of the local mesh size and polynomial degree. As a second application, we consider residual error estimates in hp-boundary element methods that are explicit in the local mesh size and the local approximation order.

Key words: Clément interpolant, quasi-interpolation, hp-FEM, hp-BEM
AMS Subject Classification: 65N30, 65N35, 65N50

1 Introduction

The regularization (or mollification or smoothing) of a function is a basic tool in analysis and the theory of functions, cf., e.g., [Bur98]. In its simplest form on the full space $\mathbb{R}^d$ one chooses a compactly supported smooth mollifier ρ with $\|\rho\|_{L^1(\mathbb{R}^d)} = 1$, introduces for $\varepsilon \in (0,1)$ the scaled mollifier $\rho_\varepsilon(\mathbf{x}) = \varepsilon^{-d}\rho(\mathbf{x}/\varepsilon)$ and defines the regularization u_ε of a function $u \in L^1(\mathbb{R}^d)$ as the convolution of u with the mollifier ρ_ε, i.e., $u_\varepsilon := \rho_\varepsilon \star u$. It is well known that this regularized function satisfies certain "inverse estimates" and has certain approximation properties if the function u has some Sobolev regularity. That is, if one denotes by $\omega_\varepsilon := \cup_{\mathbf{x}\in\omega} B_\varepsilon(\mathbf{x})$ the "ε-neighborhood" of a domain ω, then one has with the usual Sobolev spaces H^s,

$$\text{inverse estimate:} \qquad \|u_\varepsilon\|_{H^k(\omega)} \lesssim \varepsilon^{m-k}\|u\|_{H^m(\omega_\varepsilon)}, \quad k \geq m;$$
$$\text{simultaneous approximation property:} \qquad \|u - u_\varepsilon\|_{H^k(\omega)} \lesssim \varepsilon^{m-k}\|u\|_{H^m(\omega_\varepsilon)}, \quad 0 \leq k \leq m.$$

*Supported by CONICYT through FONDECYT project 3140614.
†Pontificia Universidad Católica de Chile, Facultad de Matemáticas, Vicuna Mackenna 4860, Santiago, Chile. mkarkulik@mat.puc.cl
‡Vienna University of Technology, Institute for Analysis and Scientific Computing. Wiedner Hauptstrasse 8-10, A-1040 Vienna. melenk@tuwien.ac.at

The regularized function u_ε is obtained from u by an averaging of u on a *fixed* length scale ε. It is the purpose of the present paper to derive analogous estimates for operators that are based on averaging on a *spatially varying* length scale (see Theorem 2.2). Let us mention that averaging with a spatially varying length scale has been used in [Sha85] to obtain an inverse trace theorem.

For many purposes of numerical analysis, the tool corresponding to the regularization technique in analysis is quasi-interpolation. In the finite element community, such operators are often associated with the names of Clément [Clé75] or Scott & Zhang [SZ90]. Many variants exist, but they all rely, in one way or another, on averaging on a length scale that is given by the local mesh size. The basic results for the space $S^{1,1}(\mathcal{T})$ of continuous, piecewise linear functions on a triangulation $\mathcal{T}$ of a domain Ω take the following form:

$$\begin{aligned} &\text{inverse estimate:} && \|I^{Cl}u\|_{H^k(K)} \lesssim h_K^{-k}\|I^{Cl}u\|_{L^2(K)} \lesssim h_K^{-k}\|u\|_{L^2(\omega_K)}, \qquad k \in \{0,1\},\\ &\text{approximation property:} && \|u - I^{Cl}u\|_{L^2(K)} \lesssim h_K\|u\|_{H^1(\omega_K)}; \end{aligned}$$

here, h_K stands for the diameter of the element $K \in \mathcal{T}$, and $\omega_K = \cup\{\overline{K'} \,|\, \overline{K'} \cap \overline{K} \neq \emptyset\}$ is the patch of neighboring elements of K. Quasi-interpolation operators of the above type are not restricted to piecewise linears on affine triangulations. The literature includes many extensions and refinements of the original construction of [Clé75] to account for boundary conditions, isoparametric elements, Hermite elements, anisotropic meshes, or hanging nodes, see [SZ90, BG98, GS02, Car99, CH09, Ape99, Aco01, Ran12]. Explicit constants for stability or approximation estimates for quasi-interpolation operators are given in [Ver99]. It is worth stressing that the typical h-version quasi-interpolation operators have *simultaneous* approximation properties in a scale of Sobolev spaces including fractional order Sobolev spaces.

In the hp-version of the finite element method (or the closely related spectral element method) the quasi-interpolation operator maps into the space $S^{p,1}(\mathcal{T})$ of continuous, piecewise polynomials of arbitrarily high degree p on a mesh $\mathcal{T}$. There the situation concerning quasi-interpolation with p-explicit approximation properties in scales of Sobolev spaces and corresponding inverse estimates is much less developed. In particular, for inverse estimates it is well known that in contrast to the h-version, elementwise polynomial inverse estimates do not match the approximation properties of polynomials so that some appropriate substitute needs to be found.

In the hp-version finite element method, the standard approach to the construction of piecewise polynomial approximants on unstructured meshes is to proceed in two steps: In a first step, polynomial approximations are constructed for every element separately; in a second step, the continuity requirements are enforced by using lifting operators. The first step thus falls into the realm of classical approximation theory and a plethora of results are available there, see, e.g., [DL93]. Polynomial approximation results developed in the hp-FEM/spectral element literature focused mostly (but not exclusively) on L^2-based spaces and include [AK99, Guo06, CHQZ06, BS87, DB05, CD05] and [CQ82, BM92, BM97, BDM99, BDM07, Qua84]. The second step is concerned with removing the interelement jumps. In the L^2-based setting, appropriate liftings can be found in [BCMP91, MS97, BDM07, BDM92] although the key lifting goes back at least to [Gag57]. While optimal (in p) convergence rates can be obtained with this approach, the function to be approximated is required to have some regularity since traces on the element boundary need to be defined. In conclusion, this route does not appear to lead to approximation operators for functions with minimal regularity (i.e., L^2 or even L^1). It is possible, however, to construct quasi-interpolation operators in an hp-context as done in [Mel05]. There, the construction is performed patchwise instead of elementwise and thus circumvents the need for lifting operators.

The present work takes a new approach to the construction of quasi-interpolation operators suitable for an *hp*-setting. These quasi-interpolation operators are constructed as the concatenation of two operators, namely, a smoothing operator and a classical polynomial interpolation operator. The smoothing operator turns an L^1-function into a C^∞-function by a local averaging procedure just as in the case of constant ε mentioned at the beginning of the introduction. The novel aspect is that the length scale on which the averaging is done may be linked to the *local* mesh size h and the *local* approximation order p; essentially, we select the local length scale $\varepsilon \sim h/p$. The resulting function $\mathcal{I}_\varepsilon u$ is smooth, and one can quantify $u - \mathcal{I}_\varepsilon u$ locally in terms of the local regularity of u and the local length scale h/p. Additionally, the averaged function $\mathcal{I}_\varepsilon u$ satisfies appropriate inverse estimates. The smooth function $\mathcal{I}_\varepsilon u$ can be approximated by piecewise polynomials using classical interpolation operators, whose approximation properties are well understood. In total, one arrives at a quasi-interpolation operator.

Our two-step construction that is based on first smoothing and then employing a classical interpolation operator has several advantages. The smoothing operator $\mathcal{I}_\varepsilon$ is defined merely in terms of a length scale function ε and not explicitly in terms of a mesh. Properties of the mesh are only required for the second step, the interpolation step. Hence, quasi-interpolation operators for a variety of meshes including those with hanging nodes can be constructed; the requirement is that a classical interpolation operator for smooth functions be available with the appropriate approximation properties. Also in H^1-conforming settings of regular meshes (i.e., no hanging nodes), the two-step construction can lead to improved results: In [MPS13], an *hp*-interpolation operator is constructed that leads to optimal H^1-conforming approximation in the broken H^2-norm under significant smoothness assumptions. The present technique allows us to reduce this regularity requirement to the minimal H^2-regularity. Finally, we mention that on a technical side, the present construction leads to a tighter domain of dependence for the quasi-interpolant than the construction in [Mel05].

Another feature of our construction is that it naturally leads to simultaneous approximation results in scales of (positive order) Sobolev spaces. Such simultaneous approximations have many applications, for example in connection with singular perturbation problems, [MW14]. The simultaneous approximation properties in a scale of Sobolev spaces makes *hp*-quasi-interpolation operators available for (positive order) fractional order Sobolev spaces, which are useful in *hp*-BEM. As an application, we employ our *hp*-quasi-interpolation operator for the *a posteriori* error estimation in *hp*-BEM (on shape regular meshes) involving the hypersingular operator, following [CMPS04] for the *h*-BEM.

Above, we stressed the importance of inverse estimates satisfied by the classical low order quasi-interpolation operators. The smoothed function $\mathcal{I}_\varepsilon u$ satisfies inverse estimates as well. This can be used as a substitute for the lack of a direct inverse estimate for the *hp*-quasi-interpolant. We illustrate how this inverse estimate property of $\mathcal{I}_\varepsilon u$ can be exploited in conjunction with (local) approximation properties of $\mathcal{I}_\varepsilon$ for *a posteriori* error estimation in *hp*-BEM. Specifically, we generalize the reliable *h*-BEM *a posteriori* error estimator of [Car97, CMS01] for the single layer BEM operator to the *hp*-BEM setting.

We should mention a restriction innate to our approach. Our averaging operator $\mathcal{I}_\varepsilon$ is based on volume averaging. In this way, the operator can be defined on L^1. However, this very approach limits the ability to incorporate boundary conditions. We note that the classical *h*-FEM Scott-Zhang operator [SZ90] successfully deals with boundary conditions by using averaging on boundary faces instead of volume elements. While such a technique could be employed here as well, it is beyond the scope of the present paper. Nevertheless, we illustrate in Theorem 2.4 and Corollary 3.6 what is possible within our framework of pure volume averaging.

This work is organized as follows: In Section 2 we present the main result of the paper, that is, the averaging operator $\mathcal{I}_\varepsilon$. This operator is defined in terms of a spatially varying length scale, which we formally introduce in Definition 2.1. The stability and approximation properties of $\mathcal{I}_\varepsilon$ are studied locally and collected in Theorem 2.2. The following Section 3 is devoted to applications of the operator $\mathcal{I}_\varepsilon$. In Section 3.1 (Theorem 3.3) we show how to generate a quasi-interpolation operator from $\mathcal{I}_\varepsilon$ and a classical interpolation operator. In this construction, one has to define a length scale function from the local mesh size and the local approximation order. This is done in Lemma 3.1, which may be of independent interest. Section 3.2 is devoted to *a posteriori* error estimation in *hp*-BEM: Corollary 3.9 addresses the single layer operator and Corollary 3.12 deals with the hypersingular operator. The remainder of the paper is devoted to the proof of Theorem 2.2. Since the averaging is performed on a spatially varying length scale, we will require variations of the Faà Di Bruno formula in Lemmas 4.4, 4.5. We conclude the paper with an appendix in which we show that for domains that are star-shaped with respect to a ball, the constants in the Sobolev embedding theorems (with the exception of certain limiting cases) can be controlled solely in terms of the diameter and the "chunkiness" of the domain. This result is obtained by a careful tracking of constants in the proof of the Sobolev embedding theorem. However, since this statement does not seem to be explicitly available in the literature, we include its proof in the appendix.

2 Notation and main result

Points in physical space $\mathbb{R}^d$ are denoted by small boldface letters, e.g., $\mathbf{x} = (x_1, \ldots, x_d)$. Multiindices in $\mathbb{N}_0^d$ are also denoted by small boldface letters, e.g., $\mathbf{r}$, and are used for partial derivatives, e.g., $D^{\mathbf{r}}u$, which have order $r = |\mathbf{r}| = \sum_{i=1}^d \mathbf{r}_i$. We also use the notation $\mathbf{x}^{\mathbf{r}} = \prod_{i=1}^d x_i^{r_i}$. A ball with radius r centered at $\mathbf{x} \in \mathbb{R}^d$ is denoted by $B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^d \mid |\mathbf{y}-\mathbf{x}| < r\}$, and we abbreviate $B_r := B_r(0)$. For open Lipschitz domains $\Omega \subset \mathbb{R}^d$, $C^\infty(\Omega)$ is the space of functions with derivatives of all order, and $C_0^\infty(\Omega)$ is the space of functions with derivatives of all orders and compact support in Ω. By $W^{r,p}(\Omega)$ for $r \in \mathbb{N}_0$ and $p \in [1, \infty]$ we denote the standard Sobolev space of functions with distributional derivatives of order r being in $L^p(\Omega)$, equipped with norm $\|u\|^p_{W^{r,p}(\Omega)} = \|u\|^p_{r,p,\Omega} = \sum_{|\mathbf{r}|\le r} \|D^{\mathbf{r}}u\|^p_{L^p(\Omega)}$ and seminorm $|u|^p_{W^{r,p}(\Omega)} = |u|^p_{r,p,\Omega} = \sum_{|\mathbf{r}|=r} \|D^{\mathbf{r}}u\|^p_{L^p(\Omega)}$. We will also work with fractional order spaces: for $s \in (0,1)$ and $p \in [1,\infty)$ we define Aronstein-Slobodeckij seminorms

$$|u|_{W^{s,p}(\Omega)} = |u|_{s,p,\Omega} = \left(\int_\Omega \int_\Omega \frac{|u(\mathbf{x}) - u(\mathbf{y})|^p}{|\mathbf{x}-\mathbf{y}|^{sp+d}}\, d\mathbf{y}\, d\mathbf{x}\right)^{1/p};$$

for $s = \lfloor s \rfloor + \sigma$ with $\lfloor s \rfloor = \sup\{n \in \mathbb{N}_0 \,|\, n \le s\}$ and $\sigma \in (0,1)$ we set $|u|^p_{W^{s,p}(\Omega)} = |u|^p_{s,p,\Omega} = \sum_{|\mathbf{s}|=\lfloor s \rfloor} |D^{\mathbf{s}}u|^p_{W^{\sigma,p}(\Omega)}$. The full norm $\|u\|_{W^{s,p}(\Omega)}$ is given by $\|u\|^p_{W^{s,p}(\Omega)} = \|u\|^p_{s,p,\Omega} = \|u\|^p_{W^{\lfloor s \rfloor,p}(\Omega)} + |u|^p_{W^{s,p}(\Omega)}$. By $A \lesssim B$ we mean that there is a constant $C > 0$ that is independent of relevant parameters such as the mesh size, polynomial degree and the like with $A \le C \cdot B$. In order to state our main result, we need the following definition.

Definition 2.1 (Λ-admissible length scale function)**.** *Let* $\Lambda := (\Lambda_{\mathbf{r}})_{\mathbf{r}\in\mathbb{N}_0^d}$ *be a sequence of positive numbers. A function* $\varepsilon : \Omega \to \mathbb{R}$ *is called a* Λ-admissible length scale function*, if*

(i) $\varepsilon \in C^\infty(\Omega)$,

(ii) $0 < \varepsilon \le 1$,

(iii) $|D^{\mathbf{r}}\varepsilon| \leq \Lambda_{\mathbf{r}} |\varepsilon|^{1-|\mathbf{r}|}$ *pointwise in* Ω *for all* $\mathbf{r} \in \mathbb{N}_0^d$.

The following theorem is the main result of the present work. Its proof will be given in Section 5.4 below.

Theorem 2.2. *Let* $\Omega \subset \mathbb{R}^d$ *be a Lipschitz domain. Let* $k_{\max} \in \mathbb{N}_0$ *and* $\Lambda := (\Lambda_{\mathbf{r}})_{\mathbf{r}\in\mathbb{N}_0^d}$ *be a sequence of positive numbers. Then, there exists a constant* $0 < \beta < 1$*, such that for every* Λ*-admissible length scale function* $\varepsilon \in C^\infty(\Omega)$ *there exists an operator* $\mathcal{I}_\varepsilon : L^1_{loc}(\Omega) \to C^\infty(\Omega)$ *with the following properties (i)–(iii). In the estimates below,* $\omega \subset \Omega$ *is an arbitrary open set and* $\omega_\varepsilon \subset \Omega$ *denotes its "neighborhood" given by*

$$\omega_\varepsilon := \Omega \cap \bigcup_{\mathbf{x}\in\omega} B_{\beta\varepsilon(\mathbf{x})}(\mathbf{x}). \tag{2.1}$$

(i) Suppose that the pair $(s,p) \in \mathbb{N}_0 \times [1,\infty]$ *satisfies* $s \leq k_{\max}+1$*. Assume that the pair* $(r,q) \in \mathbb{R}\times[1,\infty]$ *satisfies* $(0 \leq s \leq r \in \mathbb{R}$ *and* $q \in [1,\infty))$ *or* $(0 \leq s \leq r \in \mathbb{N}_0$ *and* $q \in [1,\infty])$*. Then*

$$|\mathcal{I}_\varepsilon u|_{r,q,\omega} \leq C_{r,q,s,p,\Lambda,\Omega} \sum_{|\mathbf{s}|\leq s} \left\| \varepsilon^{s-r+d(1/q-1/p)} D^{\mathbf{s}} u \right\|_{0,p,\omega_\varepsilon}, \tag{2.2}$$

where $C_{r,q,s,p,\Lambda,\Omega}$ *depends only on* r, q, s, p, $(\Lambda_{\mathbf{r}'})_{|\mathbf{r}'|\leq\lceil r\rceil}$*, and* Ω.

(ii) Suppose $0 \leq s \in \mathbb{R}$, $r \in \mathbb{N}_0$ *with* $s \leq r \leq k_{\max}+1$*, and* $1 \leq p \leq q < \infty$*. Define* $\mu := d(p^{-1}-q^{-1})$*. Assume that either* $(r = s+\mu$ *and* $p > 1)$ *or* $(r > s+\mu)$*. Then*

$$|u - \mathcal{I}_\varepsilon u|_{s,q,\omega} \leq C_{s,q,r,p,\Lambda,\Omega} \sum_{|\mathbf{r}|\leq r} \left\| \varepsilon^{r-s+d(1/q-1/p)} D^{\mathbf{r}} u \right\|_{0,p,\omega_\varepsilon}, \tag{2.3}$$

where $C_{s,q,r,p,\Lambda,\Omega}$ *depends only on* s, q, r, p, $(\Lambda_{\mathbf{s}'})_{|\mathbf{s}'|\leq\lceil s\rceil}$*, and* Ω.

(iii) Suppose s, $r \in \mathbb{N}_0$ *with* $s \leq r \leq k_{\max}+1$*, and* $1 \leq p < \infty$*. Define* $\mu := d/p$*. Assume that either* $(r = s+\mu$ *and* $p = 1)$ *or* $(r > s+\mu$ *and* $p > 1)$*. Then*

$$|u - \mathcal{I}_\varepsilon u|_{s,\infty,\omega} \leq C_{s,r,p,\Lambda,\Omega} \sum_{|\mathbf{r}|\leq r} \left\| \varepsilon^{r-s-d/p} D^{\mathbf{r}} u \right\|_{0,p,\omega_\varepsilon}, \tag{2.4}$$

where $C_{s,r,p,\Lambda,\Omega}$ *depends only on* s, r, p, $(\Lambda_{\mathbf{s}'})_{|\mathbf{s}'|\leq s}$*, and* Ω.

A few comments concerning Theorem 2.2 are in order:

Remark 2.3.

1. *The stability properties (part (i)) and the approximation properties (parts (ii), (iii)) involve* unweighted *(possibly fractional) Sobolev norms on the left-hand side and weighted integer order norms on the right-hand side. Our reason for admitting fractional Sobolev spaces only on one side of the estimate (here: the left-hand side) is that in this case the local length scale can be incorporated fairly easily into the estimate.*

2. *The pairs* (s,q) *for the left-hand side and* (r,p) *for the right-hand side in part (ii) are linked to each other. Essentially, the parameter combination of* (s,q) *and* (r,p) *in part (ii) is the one known from the classical Sobolev embedding theorems; the only possible exception are certain cases related to the limiting case* $p = 1$*. This connection to the Sobolev embedding theorems arises from the proof of Theorem 2.2, which employs Sobolev embedding theorems and scaling arguments.*

3. *In the classical Sobolev embedding theorems, the embedding into L^∞-based spaces is special, since the embedding is actually into a space of continuous functions. Part (ii) therefore excludes the case $q = \infty$, and some results for the special case $q = \infty$ are collected in part (iii).* ■

The following variant of Theorem 2.2 allows for the preservation of homogeneous boundary conditions:

Theorem 2.4. *The operator $\mathcal{I}_\varepsilon$ in Theorem 2.2 can be modified such that the following is true for all $u \in L^1_{loc}(\Omega)$:*

(i) The statements (i)—(iii) of Theorem 2.2 are valid, if one replaces ω_ε of the right-hand sides with $\widetilde{\omega}_\varepsilon := \cup_{\mathbf{x}\in\omega} B_{\beta\varepsilon(\mathbf{x})}(\mathbf{x})$ and simultaneously replaces u on the right-hand sides with $\widetilde{u} := u\chi_\Omega$ (i.e., u is extended by zero outside Ω). This assumes that $\widetilde{u}$ is as regular on $\widetilde{\omega}$ as the right hand-sides of (i)—(iii) dictate.

(ii) The function $\mathcal{I}_\varepsilon u$ vanishes near $\partial\Omega$. More precisely, assuming that $\varepsilon \in C(\overline{\Omega})$, then there is $\lambda > 0$ such that

$$\mathcal{I}_\varepsilon u = 0 \qquad \text{on } \cup_{\mathbf{x}\in\partial\Omega} B_{\lambda\varepsilon(\mathbf{x})}(\mathbf{x}).$$

Proof. The proof follows by a modification of the proof of Theorem 2.2. In Theorem 2.2, the value $(\mathcal{I}_\varepsilon v)(\mathbf{x})$ for an $\mathbf{x}$ near $\partial\Omega$ is obtained by an averaging of v on a ball $\mathbf{b} + B_{\delta\varepsilon(\mathbf{x})}(\mathbf{x})$ where $\mathbf{b}$ is chosen (in dependence on $\mathbf{x}$ and $\varepsilon(\mathbf{x})$) in a such a way that $\mathbf{b} + B_{\delta\varepsilon(\mathbf{x})}(\mathbf{x}) \subset \Omega$. In order to ensure that $\mathcal{I}_\varepsilon v$ vanishes near $\partial\Omega$, one can modify this procedure: one extends v by zero outside Ω and selects the translation $\mathbf{b}$ so that the averaging region $\mathbf{b} + B_{\delta\varepsilon(\mathbf{x})}(\mathbf{x}) \subset \mathbb{R}^d \setminus \Omega$. In this way, the desired condition (ii) is ensured. The statement (i) follows in exactly the same way as in the proof of Theorem 2.2. □

3 Applications to *hp*-methods

Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain. A partition $\mathcal{T} = \{K\}_{K\in\mathcal{T}}$ of Ω is a collection of open and mutually disjoint sets such that $\overline{\Omega} = \cup_{K\in\mathcal{T}}\overline{K}$. We set

$$\omega_K := \text{interior}\left(\cup\{\overline{K'}\,|\,\overline{K'}\cap\overline{K} \neq \emptyset\}\right).$$

In order to simplify the notation, we will write also "$K' \in \omega_K$" to mean $K' \in \mathcal{T}$ such that $K' \subset \omega_K$. A partition $\mathcal{T}$ of Ω is called a mesh, if every element K is the image of the reference simplex $\widehat{K} \subset \mathbb{R}^d$ under an affine map $F_K : \widehat{K} \to K$. A mesh $\mathcal{T}$ is assumed to be regular in the sense of Ciarlet [Cia76], i.e., it is not allowed to have hanging nodes (this restriction is not essential and only made for convenience of presentation–see Remark 3.5 below). To every element K we associate the mesh width $h_K := \text{diam}(K)$, and we define $h \in L^\infty(\Omega)$ as $h(\mathbf{x}) = h_K$ for $\mathbf{x} \in K$. We call a mesh γ-shape regular if

$$h_K^{-1}\left\|F_K'\right\| + h_K\left\|(F_K')^{-1}\right\| \leq \gamma \qquad \text{for all } K \in \mathcal{T}.$$

A γ-shape regular mesh is locally quasi-uniform, i.e., there is a constant C_γ which depends only on γ such that

$$C_\gamma^{-1}h_K \leq h_{K'} \leq C_\gamma h_K \qquad \text{for } \overline{K}\cap\overline{K'} \neq \emptyset. \tag{3.1}$$

A polynomial degree distribution $\mathbf{p}$ on a partition $\mathcal{T}$ is a multiindex $\mathbf{p} = (p_K)_{K\in\mathcal{T}}$ with $p_K \in \mathbb{N}_0$. A polynomial degree distribution is said to be γ_p-shape regular if

$$\gamma_p^{-1}(p_K + 1) \leq p_{K'} + 1 \leq \gamma_p p_K + 1 \qquad \text{for } \overline{K} \cap \overline{K'} \neq \emptyset. \tag{3.2}$$

We define a function $p \in L^\infty(\Omega)$ by $p(\mathbf{x}) = p_K$ for $\mathbf{x} \in K$. For $r \in \{0, 1\}$, a mesh $\mathcal{T}$ and a polynomial degree distribution $\mathbf{p}$ we introduce

$$\mathcal{S}^{\mathbf{p},r}(\mathcal{T}) = \left\{u \in H^r(\Omega) \mid \forall K \in \mathcal{T} : u|_K \circ F_K \in \mathcal{P}_{p_K}(\widehat{K})\right\}, \text{ where } \mathcal{P}_p(\widehat{K}) = \operatorname{span}\left\{\mathbf{x}^{\mathbf{k}} \mid |\mathbf{k}| \leq p\right\}. \tag{3.3}$$

The next lemma shows that shape-regular meshes and polynomial degree distributions allow for the construction of length scale functions that are essentially given by h/p and for which the sequence $(\Lambda_{\mathbf{r}})_{\mathbf{r}\in\mathbb{N}_0^r}$ depends solely only on the mesh parameters γ and γ_p.

Lemma 3.1. *Let $\Omega \subset \mathbb{R}^d$ be a domain and $\mathcal{T}$ be a partition of Ω. Suppose that for all $K \in \mathcal{T}$ it holds*

$$\begin{aligned}
C_{\text{reg}}^{-1} h_K \leq h_{K'} \leq C_{\text{reg}} h_K \qquad &\textit{for all } \overline{K'} \cap \overline{K} \neq \emptyset,\\
\left|B_{C_{\text{reg}}^{-1} h_K}(\mathbf{x}) \cap K\right| \geq C_{\text{geo}} h_K^d \qquad &\textit{for all } \mathbf{x} \in K,\\
\operatorname{card}\{K \in \mathcal{T} \,|\, K \subset \omega_K\} \leq C_{\text{patch}},&
\end{aligned}$$

with constants $C_{\text{reg}}, C_{\text{geo}}, C_{\text{patch}} > 0$ independent of K. Let $\mathbf{p}$ be a γ_p-shape regular polynomial degree distribution on $\mathcal{T}$, i.e., (3.2) holds. Then, there exists a Λ-admissible length scale function $\varepsilon \in C^\infty(\Omega)$ such that for all $K \in \mathcal{T}$ it holds

$$\varepsilon|_K \sim \frac{h_K}{p_K + 1} \qquad \textit{and} \qquad \varepsilon|_K \leq h_{K'} \qquad \textit{for all } K' \in \omega_K. \tag{3.4}$$

The hidden constants in (3.4) as well as the sequence $\Lambda = (\Lambda_{\mathbf{r}})_{\mathbf{r}\in\mathbb{N}_0^d}$ can be controlled in terms of the parameters $C_{\text{reg}}, C_{\text{geo}}, C_{\text{patch}}$, and γ_p only.

Proof. Denote by χ_K the characteristic function of the element $K \in \mathcal{T}$ and define

$$\varepsilon := \min\{1, \operatorname{diam}(\Omega)^{-1}\} C_{\text{patch}}^{-1} C_{\text{reg}}^{-C_{\text{patch}}} \sum_{K\in\mathcal{T}} \frac{h_K}{p_K + 1} \chi_K \star \rho_{C_{\text{reg}}^{-1} h_K}.$$

Here, ρ is the standard mollifier given by $\rho(\mathbf{x}) = C_\rho \exp(-1/(1 - |\mathbf{x}|^2))$ for $|\mathbf{x}| < 1$. The constant $C_\rho > 0$ is chosen such that $\int_{\mathbb{R}^d} \rho(\mathbf{x})\, d\mathbf{x} = 1$, and $\rho_\delta(\mathbf{x}) := \delta^{-d}\rho(\mathbf{x}/\delta)$. Clearly, $\varepsilon \in C^\infty(\Omega)$. Note that ρ is positive, hence $\varepsilon > 0$, and all summands in the definition of ε are positive. Furthermore, it holds $C_{\text{geo}} \lesssim (\chi_K \star \rho_{C_{\text{reg}}^{-1} h_K})|_K$. We conclude that

$$\begin{aligned}
\varepsilon|_K &\geq \min\{1, \operatorname{diam}(\Omega)^{-1}\} C_{\text{patch}}^{-1} C_{\text{reg}}^{-C_{\text{patch}}} \frac{h_K}{p_K + 1} (\chi_K \star \rho_{C_{\text{reg}}^{-1} h_K})|_K\\
&\gtrsim \min\{1, \operatorname{diam}(\Omega)^{-1}\} C_{\text{patch}}^{-1} C_{\text{reg}}^{-C_{\text{patch}}} C_{\text{geo}} \frac{h_K}{p_K + 1}.
\end{aligned}$$

Now choose a $\widetilde{K} \in \omega_K$ with $h_{\widetilde{K}} = \min_{K' \in \omega_K} h_{K'}$ and note that $\operatorname{supp}(\chi_K \star \rho_{h_{\widetilde{K}}}) \subseteq \omega_K$. Then, it holds also that $\operatorname{supp}(\chi_K \star \rho_{C_{\text{reg}}^{-1} h_K}) \subseteq \operatorname{supp}(\chi_K \star \rho_{h_{\widetilde{K}}}) \subseteq \omega_K$. Furthermore, due to the properties of ρ it holds $\chi_{K'} \star \rho_{C_{\text{reg}}^{-1} h_{K'}} \leq 1$ for all $K' \in \mathcal{T}$. We conclude that

$$
\begin{aligned}
\varepsilon|_K &= \min\{1, \operatorname{diam}(\Omega)^{-1}\} C_{\text{patch}}^{-1} C_{\text{reg}}^{-C_{\text{patch}}} \sum_{K' \in \omega_K} \frac{h_{K'}}{p_{K'} + 1} \left(\chi_{K'} \star \rho_{C_{\text{reg}}^{-1} h_{K'}} \right) |_K \\
&\leq \min\{1, \operatorname{diam}(\Omega)^{-1}\} C_{\text{reg}}^{-C_{\text{patch}}} \frac{\max_{K' \in \omega_K} h_{K'}}{\min_{K' \in \omega_K} p_{K'} + 1}.
\end{aligned}
$$

This shows first that $\varepsilon|_K \lesssim h_K/(p_K + 1)$. Furthermore, since

$$
\max_{K' \in \omega_K} h_{K'} \leq C_{\text{reg}}^{C_{\text{patch}}} \min_{K' \in \omega_K} h_{K'},
$$

we conclude that $\varepsilon|_K \leq 1$, and finally $\varepsilon|_K \leq h_{K'}$ for $K' \in \omega_K$. The same arguments show that for $\mathbf{x} \in K$, it holds that

$$
|D^{\mathbf{r}} \varepsilon(\mathbf{x})| \lesssim \left| C_{\text{reg}}^{|\mathbf{r}|} \sum_{K' \in \omega_K} \frac{h_{K'}^{1-|\mathbf{r}|}}{p_{K'} + 1} \chi_{K'} \star (D^{\mathbf{r}} \rho)_{C_{\text{reg}}^{-1} h_{K'}}(\mathbf{x}) \right| \lesssim \frac{h_K^{1-|\mathbf{r}|}}{p_K + 1} \lesssim \left(\frac{h_K}{p_K + 1} \right)^{1-|\mathbf{r}|} \lesssim \varepsilon(\mathbf{x})^{1-|\mathbf{r}|}.
$$

Hence, we have shown (i)—(iii) of Definition 2.1. $\square$

Remark 3.2. *Lemma 3.1 is formulated for general partitions which fulfill the given assumptions. We emphasize that customary γ-shape regular, affine partitions into simplices or quadrilaterals are admissible in Lemma 3.1.* ■

3.1 Quasi-interpolation operators

The following theorem shows how to construct a quasi-interpolation operator by combining the smoothing operator $\mathcal{I}_\varepsilon$ from Theorem 2.2 with a classical interpolation operator, which is allowed to require significant smoothness (e.g., point evaluations).

Theorem 3.3. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. Let $\mathcal{T}$ be a γ-shape regular mesh and $\mathbf{p}$ a γ_p-shape regular polynomial degree distribution on $\mathcal{T}$ with $p_K \geq 1$ for all $K \in \mathcal{T}$.*

Suppose that there are s, $r' \in \mathbb{N}_0$, $p \in [1, \infty)$ and a linear operator $\Pi^{hp} : W^{r',p}(\Omega) \to \mathcal{S}^{\mathbf{p},1}(\mathcal{T})$ with

$$
\left| u - \Pi^{hp} u \right|_{s,p,K}^p \leq C_\Pi \left(\frac{h_K}{p_K} \right)^{p(r'-s)} \|u\|_{r',p,\omega_K}^p \qquad \forall K \in \mathcal{T}. \tag{3.5}
$$

Then, there exists a linear operator $\mathcal{I}^{hp} : L^1_{\text{loc}}(\Omega) \to \mathcal{S}^{\mathbf{p},1}(\mathcal{T})$ with

$$
\left| u - \mathcal{I}^{hp} u \right|_{s,p,K}^p \leq C C_\Pi \left(\frac{h_K}{p_K} \right)^{p(r-s)} \|u\|_{r,p,\cup_{K' \subset \omega_K} \overline{\omega_{K'}}}^p \qquad \forall K \in \mathcal{T}
$$

for all $r \in \mathbb{N}_0$ with $s \leq r \leq r'$. The constant C depends only on the mesh parameters γ, γ_p, and s, r, r', and Ω.

Proof. Choose ε from Lemma 3.1 and $k_{\max} = r' - 1$. We use the operator $\mathcal{I}_\varepsilon$ of Theorem 2.2 and define $\mathcal{I}^{hp} := \Pi^{hp} \circ \mathcal{I}_\varepsilon$. Then, according to (2.3) and (3.5), it holds

$$\begin{aligned}\left|u - \mathcal{I}^{hp} u\right|_{s,p,K} &\leq |u - \mathcal{I}_\varepsilon u|_{s,p,K} + \left|\mathcal{I}_\varepsilon u - \Pi^{hp}\mathcal{I}_\varepsilon u\right|_{s,p,K} \\ &\lesssim \left(\frac{h_K}{p_K}\right)^{r-s} \|u\|_{r,p,\omega_K} + v \left(\frac{h_K}{p_K}\right)^{r'-s} \|\mathcal{I}_\varepsilon u\|_{r',p,\omega_K}.\end{aligned}$$

Theorem 2.2 implies

$$\begin{aligned}\|\mathcal{I}_\varepsilon u\|^p_{r',p,\omega_K} &= \sum_{K' \in \omega_K} \left(\sum_{j=0}^{r} |\mathcal{I}_\varepsilon u|^p_{j,p,K'} + \sum_{j=r}^{r'} |\mathcal{I}_\varepsilon u|^p_{j,p,K'} \right) \\ &\lesssim \sum_{K' \in \omega_K} \left(\sum_{j=0}^{r} |u|^p_{j,p,\omega_{K'}} + \sum_{j=r}^{r'} \left(\frac{h_K}{p_K}\right)^{p(r-j)} |u|^p_{r,p,\omega_{K'}} \right) \\ &\lesssim \sum_{K' \in \omega_K} \left(\frac{h_K}{p_K}\right)^{p(r-r')} \|u\|^p_{r,p,\omega_{K'}},\end{aligned}$$

where we additionally used that $K_\varepsilon \subset \omega_K$ due to (3.4). This shows the result. □

As an application of Theorem 3.3, we construct an interpolation operator with simultaneous approximation properties on regular meshes. The novel feature of the operator of Corollary 3.4 is that it provides the optimal rate of convergence in the broken H^2-norm, which can be of interest in the analysis of *hp*-Discontinuous Galerkin methods.

Corollary 3.4. *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$ be a polygonal/polyhedral domain. Fix $r_{max} \in \mathbb{N}_0$. Let $\mathcal{T}$ be a γ-shape regular mesh and $\mathbf{p}$ be a γ_p-shape regular polynomial degree distribution with $p_K \geq 1$ for all $K \in \mathcal{T}$. Define $\widehat{p}_K := \min\{p_{K'} \,|\, K' \subset \omega_K\}$. Then, there is an operator $\mathcal{I}^{hp} : L^1_{\text{loc}}(\Omega) \to \mathcal{S}^{\mathbf{p},1}(\mathcal{T})$ such that for every $r \in \{0, 1, \ldots, r_{max}\}$*

$$\left|u - \mathcal{I}^{hp} u\right|_{\ell,2,K} \leq C h_K^{\min\{\widehat{p}_K+1,r\}-\ell} p_K^{-(r-\ell)} \|u\|_{r,2,\cup_{K' \in \omega_K} \overline{\omega'_K}} \qquad \text{for } \ell = 0, 1, \ldots, \min\{2, r\}. \quad (3.6)$$

The constant C depends only on the shape regularity constants γ, γ_p and on r_{max} and Ω.

Proof. See Appendix B for details. □

Remark 3.5. *(non-regular meshes/hanging nodes) The meshes in Theorem 3.3 and Corollary 3.4 are assumed to be regular, i.e., no hanging nodes are allowed. Furthermore, the meshes are assumed to be affine and simplicial. The proof of Theorem 3.3 shows that these restrictions are not essential: It relies on the smoothing operator $\mathcal{I}_\varepsilon$ (which is essentially independent of the meshes) and some suitable polynomial approximation operator on the mesh $\mathcal{T}$. If a polynomial approximation operator is available on meshes with hanging nodes, or on meshes that have other types of elements (e.g., quadrilaterals) or non-affine elements, then similar arguments as in the proof of Theorem 3.3 can be applied.* ■

So far, we have only made use of Theorem 2.2. Its modification, Theorem 2.4, allows for the incorporation of boundary conditions. It is worth pointing out that regularity of the zero extension $\chi_\Omega u$ of u is required, which limits the useful parameter range. Nevertheless, Theorem 2.4 allows us to develop an *hp*-Clément interpolant that preserves homogeneous boundary conditions:

Corollary 3.6. *Let the hypotheses on the mesh and polynomial degree distribution be as in Theorem 3.3. Then there exists a linear operator* $\mathcal{I}^{hp} : L^1_{loc}(\Omega) \to \mathcal{S}^{\mathbf{p},1}(\mathcal{T}) \cap H^1_0(\Omega)$ *such that for all* $K \in \mathcal{T}$ *and all* $u \in H^1_0(\Omega)$*:*

$$\begin{aligned} \|\mathcal{I}^{hp}u\|_{L^2(K)} &\leq C\|u\|_{L^2(\omega_K)}, \\ |u - \mathcal{I}^{hp}u|_{H^\ell(K)} &\leq C\left(\frac{h_K}{p_K}\right)^{1-\ell}\|u\|_{H^1(\omega_K)}, \qquad \ell \in \{0,1\}. \end{aligned}$$

The constant C *depends only on the mesh parameters* γ, γ_p *and* Ω.

3.2 Residual error estimation in hp-boundary element methods

Let $\Gamma := \partial\Omega$ be the boundary of a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$. If $d = 2$, we assume $\operatorname{diam}(\Omega) < 1$. Two basic problems in boundary element methods (BEM) involve the *single layer operator* $\mathrm{V} : H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$ and the *hypersingular operator* $\mathrm{D} : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$. We refer to [Cos88, HW08, McL00, Néd88] for a detailed discussion of these operators and to the monographs [Ste08, SS11] for boundary element methods in general. In the simplest BEM settings, one studies the following two problems:

$$\text{Find } \varphi \in H^{-1/2}(\Gamma) \text{ s.t. } \mathrm{V}\,\varphi = f; \tag{3.7}$$

$$\text{Find } u \in H^{1/2}(\Gamma) \text{ s.t. } \mathrm{D}\,u = g; \tag{3.8}$$

here, the right-hand sides are given data with $f \in H^{1/2}(\Gamma)$ and $g \in H^{-1/2}(\Gamma)$ such that $\langle g, 1\rangle_\Gamma = 0$. In a conforming Galerkin setting, one takes finite-dimensional subspaces $V_N \subset H^{-1/2}(\Gamma)$ and $W_N \subset H^{1/2}(\Gamma)$ and defines the Galerkin approximations $\varphi_N \in V_N$ and $u_N \in W_N$ by

$$\text{Find } \varphi_N \in V_N \text{ s.t. } \langle \mathrm{V}\,\varphi_N, v\rangle_\Gamma = \langle f, v\rangle_\Gamma \qquad \forall v \in V_N, \tag{3.9}$$

$$\text{Find } u_N \in W_N \text{ s.t. } \langle \mathrm{D}\,u_N, v\rangle_\Gamma = \langle g, v\rangle_\Gamma \qquad \forall v \in W_N. \tag{3.10}$$

Residual *a posteriori* error estimation for these Galerkin approximations is based on bounding

$$\|f - \mathrm{V}\,\varphi_N\|_{H^{1/2}(\Gamma)} \qquad \text{and} \qquad \|g - \mathrm{D}\,u_N\|_{H^{-1/2}(\Gamma)}. \tag{3.11}$$

These norms are non-local, which results in two difficulties. First, it makes them hard to evaluate in a computational environment. Second, they cannot be used as indicators for local mesh refinement. The ultimate goal of residual error estimation is to obtain a fully localized, computable error estimator based on the equation's residual. Several residual error estimators have been presented in an h-version BEM context, for example, [Ran86, Ran89, Ran93, CS95, CS96, Car97, CMS01, CMPS04]. The localization of norms in the hp-BEM context is a more delicate question, and to our knowledge there are no results on fully localized residual error estimates. The only result that we are aware of is [CFS96], were it is shown that the error can be bounded reliably by the product of two localized residual error estimators. In the following, we generalize the local residual estimators of [CMS01] (for the single layer operator) and [CMPS04] (for the hypersingular operator), which were developed in an h-BEM setting, to the hp-BEM in Corollaries 3.9 and 3.12.

3.2.1 Spaces and meshes on surfaces

Sobolev spaces and local parametrizations by charts

The Sobolev spaces $H^s(\Gamma)$, $s \in [0,1]$, can be defined using an open cover $\mathcal{U} := \{U_j\}_{j=1}^n$ of Γ, local bi-lipschitz parametrizations $\left\{\chi_j : B_2^{d-1} \to U_j\right\}_{j=1}^n$ and a partition of unity $\{\beta_j\}_{j=1}^n$ subordinate to $\mathcal{U}$, cf. [SS11]. Here, B_2^{d-1} is the open ball of radius 2 in $\mathbb{R}^{d-1}$. A norm is given by $\|u\|_{H^s(\Gamma)}^2 := \sum_{j=1}^n \|(\beta_j u)\circ\chi_j\|_{H^s(B_2^{d-1})}^2$. The space $L^2(\Gamma)$ is defined equivalently with respect to the surface measure $d\sigma$, and the space $H^1(\Gamma)$ is defined equivalently via the norm $\|u\|_{H^1(\Gamma)}^2 := \|u\|_{L^2(\Gamma)}^2 + \|\nabla_\Gamma u\|_{L^2(\Gamma)}^2$, where ∇_Γ denotes the surface gradient. The space $L^2(\Gamma)$ is equipped with the scalar product $\int_\Gamma u \cdot v \; d\sigma$. The space $H^{-1/2}(\Gamma)$ is the dual space of $H^{1/2}(\Gamma)$ with respect to the extended $L^2(\Gamma)$-scalar product. Additionally, spaces of fractional order can be defined via interpolation between $L^2(\Gamma)$ and $H^1(\Gamma)$.

Meshes and piecewise polynomial spaces

As in the case of volume discretizations discussed above, we restrict our attention to affine triangulations of Γ. A triangulation $\mathcal{T}$ of Γ is a partition of Γ into (relatively) open disjoint elements K. Every element K is the image of the reference simplex $\widehat{K} \subset \mathbb{R}^{d-1}$ under an affine element map $F_K : \widehat{K} \to K \subset \Gamma \subset \mathbb{R}^d$. The mesh width h_K is given by $h_K := \operatorname{diam}(K)$. Since the (affine) element maps F_K map from $\mathbb{R}^{d-1}$ to $\mathbb{R}^d$, the shape-regularity requirement takes the following form:

$$h_K^{-1}\|F_K'\| + h_K^2\| \left((F_K')^\top F_K'\right)^{-1} \| \leq \gamma \quad \text{for all } K \in \mathcal{T}. \tag{3.12}$$

The local comparability (3.1) is ensured by (3.12). The spaces $\mathcal{S}^{\mathbf{p},0}(\mathcal{T})$ and $\mathcal{S}^{\mathbf{p},1}(\mathcal{T})$ are defined as in (3.3), but with Γ instead of Ω. We will also require the local comparability of the polynomial degree spelled out in (3.2). As at the outset of Section 3, we denote by h and p the piecewise constant functions given by $h|_K = h_K$ and $p|_K = p_K$.

3.2.2 Single layer operator

Lemma 3.7. *Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain and let $\Gamma = \partial\Omega$ be its boundary. Let $\mathcal{T}$ be a γ-shape regular mesh on Γ, and let $\mathbf{p}$ be a γ_p-shape regular polynomial degree distribution on $\mathcal{T}$. Then it holds*

$$\|u\|_{H^{1/2}(\Gamma)} \lesssim \left\|h^{-1/2}\hat{p}^{1/2}u\right\|_{L^2(\Gamma)} + \left\|h^{1/2}\hat{p}^{-1/2}\nabla_\Gamma u\right\|_{L^2(\Gamma)} \qquad \text{for } u \in H^1(\Gamma),$$

where $\hat{p} := \max(1,p)$ and the hidden constant depends only on γ and γ_p.

Proof. Define $\mathcal{T}_j := \{\chi_j^{-1}(K \cap U_j) \mid K \in \mathcal{T}\}$. As the χ_j are bi-lipschitz, $\{\mathcal{T}_j\}_{j=1}^n$ are partitions of B_2^{d-1} which fulfill the assumptions of Lemma 3.1. Hence, we obtain a family $\{\varepsilon_j\}_{j=1}^n$ of Λ-admissible length scale functions on B_2^{d-1} with $\varepsilon_j|_{\chi_j^{-1}(K)} \sim h_K/p_K$. Define $\nu := \min_{j=1}^n \operatorname{dist}(\operatorname{supp}(\beta_j \circ \chi_j), \partial B_2^{d-1})/2$. As the support of $\beta_j \circ \chi_j$ is compact in B_2^{d-1}, it holds $0 < \nu < 1$ and $\{\nu\varepsilon_j\}_{j=1}^n$ is a family of Λ-admissible length scale functions on B_2^{d-1}. With Theorem 2.2 we construct operators $\mathcal{I}_{\nu\varepsilon_j}$:

$L^1_{\mathrm{loc}}(B_2^{d-1}) \to C^\infty(B_2^{d-1})$, and the choice of ν shows $\mathrm{supp}(\mathcal{I}_{\nu\varepsilon_j}(\beta_j u \circ \chi_j)) \subset B_2^{d-1}$. Theorem 2.2 then shows

$$\begin{aligned}
\|\beta_j u \circ \chi_j\|_{H^{1/2}(B_2^{d-1})} &\le \|\mathcal{I}_{\nu\varepsilon_j}(\beta_j u \circ \chi_j)\|_{H^{1/2}(B_2^{d-1})} + \|\beta_j u \circ \chi_j - \mathcal{I}_{\nu\varepsilon_j}(\beta_j u \circ \chi_j)\|_{H^{1/2}(B_2^{d-1})} \\
&\lesssim \left\|\varepsilon_j^{-1/2} \beta_j u \circ \chi_j\right\|_{L^2(B_2^{d-1})} + \left\|\varepsilon_j^{1/2} \nabla(\beta_j u \circ \chi_j)\right\|_{L^2(B_2^{d-1})} \\
&\lesssim \left\|\varepsilon_j^{-1/2} u \circ \chi_j\right\|_{L^2(B_2^{d-1})} + \left\|\varepsilon_j^{1/2} \nabla(u \circ \chi_j)\right\|_{L^2(B_2^{d-1})} \\
&\lesssim \left\|h^{-1/2} \hat{p}^{1/2} u\right\|_{L^2(U_j \cap \Gamma)} + \left\|h^{1/2} \hat{p}^{-1/2} \nabla_\Gamma u\right\|_{L^2(U_j \cap \Gamma)},
\end{aligned}$$

where we used $\|D\chi_j\|_{L_\infty(B_2^{d-1})} \simeq 1$ in the last step. Taking the sum over j concludes the result. □

The following can be seen as a generalization of the results of [Car97, CMS01] to obtain a residual *a posteriori* error estimator for *hp*-boundary elements for weakly singular integral equations.

Theorem 3.8. *Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain and let $\Gamma = \partial\Omega$ be its boundary. Let $\mathcal{T}$ be a γ-shape regular mesh on Γ, and let $\mathbf{p}$ be a γ_p-shape regular polynomial degree distribution on $\mathcal{T}$. Suppose that $u \in H^1(\Gamma)$ satisfies*

$$\int_\Gamma u \cdot \phi_{hp} \, d\sigma = 0 \quad \text{for all } \phi_{hp} \in \mathcal{S}^{\mathbf{p},0}(\mathcal{T}). \tag{3.13}$$

Then, with $\hat{p} := \max(1, p)$,

$$\|u\|_{H^{1/2}(\Gamma)} \le C_{\gamma,\gamma_p} \left\|h^{1/2} \hat{p}^{-1/2} \nabla_\Gamma u\right\|_{L^2(\Gamma)},$$

where the constant C_{γ,γ_p} depends only on the shape-regularity constants γ, γ_p, and on Γ.

Proof. Denote by Π the $L^2(\Gamma)$-orthogonal projection onto $\mathcal{S}^{\mathbf{p},0}(\mathcal{T})$. Then, due to the orthogonality (3.13), it holds

$$\left\|h^{-1/2} \hat{p}^{1/2} u\right\|_{L^2(\Gamma)} = \left\|h^{-1/2} \hat{p}^{1/2} (1 - \Pi) u\right\|_{L^2(\Gamma)} \lesssim \left\|h^{1/2} \hat{p}^{-1/2} \nabla_\Gamma u\right\|_{L^2(\Gamma)},$$

where the last estimate follows from well-known approximation results. Finally, Lemma (3.7) concludes the result. □

We explicitly formulate the residual error estimate that results from Theorem 3.8:

Corollary 3.9 (*hp*-*a posteriori* error estimation for single layer operator)**.** *Let $f \in H^1(\Gamma)$ and let $\varphi \in H^{-1/2}(\Gamma)$ solve (3.7). Suppose that $\mathcal{T}$ is a γ-shape regular mesh on Γ and $\mathbf{p}$ is a γ_p-shape regular polynomial degree distribution. Let $V_N = \mathcal{S}^{\mathbf{p},0}(\mathcal{T})$ in (3.9) and let φ_N be the solution of (3.9). Then with residual $R_N := f - \mathrm{V}\,\varphi_N$ the Galerkin error $\varphi - \varphi_N$ satisfies*

$$\|\varphi - \varphi_N\|_{H^{-1/2}(\Gamma)} \le C\|R_N\|_{H^{1/2}(\Gamma)} \le C\|\,(h/\hat{p})^{1/2}\, \nabla_\Gamma R_N\|_{L^2(\Gamma)}.$$

Proof. The first estimate expresses the boundedness of the operator V^{-1}. The second estimate follows from Theorem 3.8 and the Galerkin orthogonalities. □

3.2.3 Hypersingular operator

Lemma 3.10. *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a Lipschitz domain, and let $\Gamma = \partial\Omega$. Let $\mathcal{T}$ a γ-shape regular mesh on Γ, and let $\mathbf{p}$ be a γ_p-shape regular polynomial degree distribution on $\mathcal{T}$ with $p_K \geq 1$ for all $K \in \mathcal{T}$. Then, there exists a linear operator $\mathcal{I}_{\mathcal{T}}^{\mathbf{p}} : H^{1/2}(\Gamma) \to \mathcal{S}^{\mathbf{p},1}(\mathcal{T})$ such that*

$$\left\| h^{-1/2} p^{1/2} (u - \mathcal{I}_{\mathcal{T}}^{\mathbf{p}} u) \right\|_{L^2(\Gamma)} \lesssim \|u\|_{H^{1/2}(\Gamma)} .$$

Proof. Define the linear smoothing operator $\mathcal{I}_{\mathcal{T}}$ by $\mathcal{I}_{\mathcal{T}} u := \sum_{j=1}^n \mathcal{I}_{\nu\varepsilon_j}(\beta_j u \circ \chi_j) \circ \chi_j^{-1}$, where the $\mathcal{I}_{\nu\varepsilon_j}$ are as in the proof of Lemma 3.7. We have $\mathcal{I}_{\mathcal{T}} u \in H^1(\Gamma)$ and $\mathcal{I}_{\mathcal{T}} u \in C^\infty(\overline{\Gamma_k})$ on every planar side Γ_k of Γ. From $\mathcal{I}_{\mathcal{T}} u$, we can construct in an elementwise fashion as in Corollary 3.4 a function $\mathcal{I}_{\mathcal{T}}^{\mathbf{p}} u \in \mathcal{S}^{\mathbf{p},1}(\mathcal{T})$ such that

$$\left\| \mathcal{I}_{\mathcal{T}} u - \mathcal{I}_{\mathcal{T}}^{\mathbf{p}} u \right\|_{L^2(K)} \lesssim h_K^2 p_K^{-2} \left| \mathcal{I}_{\mathcal{T}} u \right|_{H^2(K)} .$$

Theorem 2.2 shows

$$\begin{aligned}
\left\| h^{-1} p \left(\mathcal{I}_{\mathcal{T}} u - \mathcal{I}_{\mathcal{T}}^{\mathbf{p}} u \right) \right\|_{L^2(\Gamma)}^2 &= \sum_{K \in \mathcal{T}} \left\| h^{-1} p \left(\mathcal{I}_{\mathcal{T}} u - \mathcal{I}_{\mathcal{T}}^{\mathbf{p}} u \right) \right\|_{L^2(K)}^2 \lesssim \sum_{K \in \mathcal{T}} h_K^2 p_K^{-2} \left| \mathcal{I}_{\mathcal{T}} u \right|_{H^2(K)}^2 \\
&\lesssim \sum_{j=1}^n \sum_{K \in \mathcal{T}} h_K^2 p_K^{-2} \| \mathcal{I}_{\nu\varepsilon_j}(\beta_j u \circ \chi_j) \circ \chi_j^{-1} \|_{H^2(K \cap U_j)}^2 \\
&\lesssim \sum_{j=1}^n \sum_{K \in \mathcal{T}} h_K^2 p_K^{-2} \| \mathcal{I}_{\nu\varepsilon_j}(\beta_j u \circ \chi_j) \|_{H^2(B_2^{d-1} \cap \chi_j^{-1}(K))}^2 \\
&\lesssim \sum_{j=1}^n \sum_{K \in \mathcal{T}} \| \beta_j u \circ \chi_j \|_{H^1(B_2^{d-1} \cap \chi_j^{-1}(\omega_K))}^2 \\
&\lesssim \sum_{j=1}^n \| \beta_j u \circ \chi_j \|_{H^1(B_2^{d-1})}^2 = \|u\|_{H^1(\Gamma)}^2 .
\end{aligned}$$

The same argument shows $\left\| \mathcal{I}_{\mathcal{T}} u - \mathcal{I}_{\mathcal{T}}^{\mathbf{p}} u \right\|_{L^2(\Gamma)} \lesssim \|u\|_{L^2(\Gamma)}$. An interpolation argument shows that $\left\| h^{-1/2} p^{1/2} \left(\mathcal{I}_{\mathcal{T}} u - \mathcal{I}_{\mathcal{T}}^{\mathbf{p}} u \right) \right\|_{L^2(\Gamma)} \lesssim \|u\|_{H^{1/2}(\Gamma)}$. Analogously, we obtain

$$\begin{aligned}
\left\| h^{-1} p (u - \mathcal{I}_{\mathcal{T}} u) \right\|_{L^2(\Gamma)}^2 &= \left\| h^{-1} p \sum_{j=1}^n \left(\beta_j u - \mathcal{I}_{\nu\varepsilon_j}(\beta_j u \circ \chi_j) \circ \chi_j^{-1} \right) \right\|_{L^2(\Gamma)}^2 \\
&\lesssim \sum_{j=1}^n \left\| \varepsilon_j^{-1} \left(\beta_j u \circ \chi_j - \mathcal{I}_{\nu\varepsilon_j}(\beta_j u \circ \chi_j) \right) \right\|_{L^2(B_2^{d-1})}^2 \\
&\lesssim \sum_{j=1}^n \| \beta_j u \circ \chi_j \|_{H^1(B_2^{d-1})}^2 = \|u\|_{H^1(\Gamma)}^2 .
\end{aligned}$$

Here, the second estimate follows as we can bound the approximation error of $\mathcal{I}_{\nu\varepsilon_j}$ locally on every element $\chi_j^{-1}(K \cap U_j)$. Likewise, we obtain $\|u - \mathcal{I}_{\mathcal{T}} u\|_{L^2(\Gamma)} \lesssim \|u\|_{L^2(\Gamma)}$. An interpolation argument again shows that $\left\| h^{-1/2} p^{1/2} \left(u - \mathcal{I}_{\mathcal{T}} u \right) \right\|_{L^2(\Gamma)} \lesssim \|u\|_{H^{1/2}(\Gamma)}$. The triangle inequality finally shows the result. $\square$

The following can be seen as a generalization of the results of [CMPS04] to obtain a residual *a posteriori* error estimator for *hp*-boundary elements for hypersingular integral equations.

Theorem 3.11. *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a Lipschitz domain and let $\Gamma = \partial\Omega$ be its boundary. Let $\mathcal{T}$ be a γ-shape regular mesh on Γ, and let $\mathbf{p}$ a γ_p-shape regular polynomial degree distribution on $\mathcal{T}$ with $p_K \geq 1$ for all $K \in \mathcal{T}$. Suppose that $u \in L^2(\Gamma)$ satisfies*

$$\int_\Gamma u \cdot \phi_{hp}\, d\sigma = 0 \qquad \text{for all } \phi_{hp} \in \mathcal{S}^{\mathbf{p},1}(\mathcal{T}). \tag{3.14}$$

Then, for a constant C_{γ,γ_p} that depends only on the shape-regularity constants γ, γ_p of the mesh and on Γ,

$$\|u\|_{H^{-1/2}(\Gamma)} \leq C_{\gamma,\gamma_p} \left\| h^{1/2} p^{-1/2} u \right\|_{L^2(\Gamma)}.$$

Proof. The orthogonality (3.14), Lemma 3.10, and Cauchy-Schwarz show for any $v \in H^{1/2}(\Gamma)$

$$\begin{aligned}\int_\Gamma u \cdot v\, d\sigma = \int_\Gamma u \cdot (v - \mathcal{I}v)\, d\sigma &\lesssim \left\| h^{1/2} p^{-1/2} u \right\|_{L^2(\Gamma)} \left\| h^{-1/2} p^{1/2} (v - \mathcal{I}v) \right\|_{L^2(\Gamma)} \\ &\lesssim \left\| h^{1/2} p^{-1/2} u \right\|_{L^2(\Gamma)} \|v\|_{H^{1/2}(\Gamma)}.\end{aligned}$$

The definition of $H^{-1/2}(\Gamma)$ as dual space of $H^{1/2}(\Gamma)$ shows the result. $\square$

Corollary 3.12 (*hp*-*a posteriori* error estimation for hypersingular operator)**.** *Let $g \in L^2(\Gamma)$ with $\langle g, 1\rangle_\Gamma = 0$ and let $u \in H^{1/2}(\Gamma)$ be defined by (3.8). Suppose that $\mathcal{T}$ is a γ-shape regular mesh on Γ and $\mathbf{p}$ is a γ_p-shape regular polynomial degree distribution with $p_K \geq 1$ for all $K \in \mathcal{T}$. Let $W_N = \mathcal{S}^{\mathbf{p},1}(\mathcal{T})$ and $u_N \in W_N$ be given by (3.10). Then with the residual $R_N := g - \mathrm{D}\, u_N$*

$$|u - u_N|_{H^{1/2}(\Gamma)} \leq C \|R_N\|_{H^{-1/2}(\Gamma)} \leq C \|(h/p)^{1/2} R_N\|_{L^2(\Gamma)}.$$

Proof. The first estimate follows from the (semi-)ellipticity of the hypersingular operator $\mathrm{D} : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$. The second estimate follows from Theorem 3.11 and Galerkin orthogonality. $\square$

4 Technical results for the proof of Theorem 2.2

This section is about tools and technical results that will be used in the remainder. The letter ρ will always denote a *mollifier*, i.e., a nonnegative function $\rho \in C_0^\infty(\mathbb{R}^d)$ with *(i)* $\rho(\mathbf{x}) = 0$ for $|\mathbf{x}| \geq 1$, and *(ii)* $\int_{\mathbb{R}^d} \rho(\mathbf{x}) d\mathbf{x} = 1$. For $\delta > 0$, we write $\rho_\delta(\mathbf{x}) := \rho(\mathbf{x}/\delta)\, \delta^{-d}$, so that $\rho_\delta(\mathbf{x}) = 0$ for $|\mathbf{x}| \geq \delta$ and $\int_{\mathbb{R}^d} \rho_\delta(\mathbf{x}) d\mathbf{x} = 1$. A mollifier ρ is said to be of order $k_{\max} \in \mathbb{N}_0$ if

$$\int_{\mathbb{R}^d} \mathbf{y}^{\mathbf{s}} \rho(\mathbf{y}) d\mathbf{y} = 0 \qquad \text{for every multi-index } \mathbf{s} \in \mathbb{N}_0^d \text{ with } 1 \leq |\mathbf{s}| \leq k_{max}. \tag{4.1}$$

(Note that this condition is void if $k_{max} = 0$.) The condition (4.1) implies that a convolution with a mollifier of order $k_{\max}$ reproduces polynomials of degree up to k_{max}.

A lot of results will be proved in a local fashion. In order to transform these local results into global ones, we will use Besicovitch's covering theorem, see [EG92a]. It is recalled here for the reader's convenience:

Proposition 4.1 (Besicovitch covering theorem)**.** *There is a constant N_d (depending only on the spatial dimension d) such that the following holds: For any collection $\mathcal{F}$ of non-empty, closed balls in $\mathbb{R}^d$ with $\sup\{\operatorname{diam} B \mid B \in \mathcal{F}\} < \infty$, and the set A of the mid-points of the balls $B \in \mathcal{F}$, there are subsets $\mathcal{G}_1, \ldots, G_{N_d} \subset \mathcal{F}$ such that for each $i = 1, \ldots, N_d$, the family $\mathcal{G}_i$ is a countable set of pairwise disjoint balls and*

$$A \subset \bigcup_{i=1}^{N_d} \bigcup_{B \in \mathcal{G}_i} B.$$

□

An open set $S \subset \mathbb{R}^d$ is said to be *star-shaped with respect to a ball B*, if the closed convex hull of $\{\mathbf{x}\} \cup B$ is a subset of S for every $\mathbf{x} \in S$. The *chunkiness parameter* of S is defined as $\eta(S) := \operatorname{diam}(S)/\rho_{\max}$, where

$$\rho_{\max} := \sup\{\rho \mid S \text{ is star-shaped with respect to a ball of radius } \rho\},$$

cf. [BS08, Def. 4.2.16]. We will frequently employ Sobolev embedding theorems, and it will be necessary to control the constants in terms of the chunkiness of the underlying domain. Results of this type are well known for integer order spaces, cf. [Ada75, Ch. V]. In Appendix A, we give a self-contained proof that also for fractional order spaces the constants of the Sobolev embedding theorem for star-shaped domains can be controlled in terms of the chunkiness parameter and the diameter of the domain. This results in the following embedding theorem.

Theorem 4.2 (embedding theorem)**.** *Let $\eta > 0$, and let $\Omega \subset \mathbb{R}^d$ be a bounded domain with chunkiness parameter $\eta(\Omega) \leq \eta$. Let $s, r, p, q \in \mathbb{R}$ with $0 \leq s \leq r < \infty$ and $1 \leq p \leq q < \infty$ and set $\mu := d(p^{-1} - q^{-1})$. Assume that ($r = s + \mu$ and $p > 1$) or ($r > s + \mu$). Then there exists a constant $C_{s,q,r,p,\eta,d}$ (depending only on the quantities indicated) such that*

$$|u|_{s,q,\Omega} \leq C_{s,q,r,p,\eta,d} \operatorname{diam}(\Omega)^{-\mu} \left(\operatorname{diam}(\Omega)^{r-s} |u|_{r,p,\Omega} + \sum_{\substack{r' \in \mathbb{N}_0: \\ \lceil s \rceil \leq r' \leq r}} \operatorname{diam}(\Omega)^{r'-s} |u|_{r',p,\Omega} \right).$$

Furthermore, if s, $r \in \mathbb{N}_0$, $s \leq r$, set $\mu' := d/p$. Assume that ($r = s + \mu'$ and $p = 1$) or ($r > s + \mu'$ and $p > 1$). Then there exists a constant $C_{s,r,p,\eta,d}$ (depending only on the quantities indicated) such that

$$|u|_{s,\infty,\Omega} \leq C_{s,r,p,\eta,d} \operatorname{diam}(\Omega)^{-\mu'} \sum_{r'=s}^{r} \operatorname{diam}(\Omega)^{r'-s} |u|_{r',p,\Omega}.$$

Proof. As Ω is star-shaped with respect to a ball of radius $\operatorname{diam}(\Omega)/(2\eta)$, the scaled domain $\widehat{\Omega} := \operatorname{diam}(\Omega)^{-1}\Omega$ is star-shaped with respect to a ball of radius $1/(2\eta)$. For the first result, we employ Theorem A.1 and scaling arguments to obtain the stated right-hand side with the sum extending over $r' \in \{0, 1, \ldots, \lfloor r \rfloor\}$ instead of $\{\lceil s \rceil, \ldots, \lfloor r \rfloor\}$. The restriction of the summation to $r' \in \{\lceil s \rceil, \ldots, \lfloor r \rfloor\}$ follows from the observation that the left-hand side vanishes for polynomials of degree $\lceil s \rceil - 1$. Hence, one can use of polynomial approximation result of [DS80] in the usual way by replacing u with $u - \pi$, where π is the polynomial approximation given in [DS80].

The L^∞-estimate follows from [BS08, Lem. 4.3.4] and scaling arguments. □

Another tool we will use is the classical Bramble-Hilbert Lemma. The following version can be found in [BS08, Lemma 4.3.8].

Lemma 4.3 (Bramble-Hilbert)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with chunkiness parameter $\eta(\Omega) \leq \eta < \infty$. Then, for all $u \in W^{m,p}(S)$ with $p \geq 1$, there is a polynomial $\pi \in \mathcal{P}^{m-1}(S)$ such that*

$$|u - \pi|_{k,p,S} \leq C_{m,d,\eta} \operatorname{diam}(S)^{m-k} |u|_{m,p,S}, \qquad \text{for all } k = 0, \ldots, m.$$

The constant $C_{m,d,\eta}$ depends only on m, d, and η. □

The aim of the next lemma is to formulate a version of the Faà di Bruno formula, which is a formula for computing higher derivatives of composite functions. For $s, \ell \in \mathbb{N}_0$ we denote by $\mathcal{M}_{s,\ell}$ a set of multi-indices $\mathcal{M}_{s,\ell} = \{\mathbf{t}_i\}_{i=1}^{\ell} \subset \mathbb{N}_0^d$ such that $|\mathbf{t}_i| \geq 1$ and $\sum_{i=1}^{\ell} (|\mathbf{t}_i| - 1) = s$.

Lemma 4.4 (Faà di Bruno)**.** *For every $\mathbf{s} \in \mathbb{N}_0^d$ with $|\mathbf{s}| \geq 1$ and every $\mathbf{r} \in \mathbb{N}_0^d$ with $|\mathbf{r}| \leq |\mathbf{s}|$ and every set $\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}$ with $1 \leq \ell \leq |\mathbf{r}|$ there is a polynomial $P_{\mathbf{s},\mathbf{r},\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}} : \mathbb{R}^d \to \mathbb{R}$ of degree $|\mathbf{r}|$ such that the following is true: For any $\varepsilon \in C^\infty(\mathbb{R}^d)$, $\mathbf{z} \in \mathbb{R}^d$, and $u \in C^\infty(\mathbb{R}^d)$ the derivative $D^{\mathbf{s}}_{\mathbf{x}} u(\mathbf{x}')$, $\mathbf{x}' := \mathbf{x} + \mathbf{z}\varepsilon(\mathbf{x})$, can be written in the form*

$$D^{\mathbf{s}}_{\mathbf{x}} u(\mathbf{x}') = (D^{\mathbf{s}} u)(\mathbf{x}') + \sum_{|\mathbf{r}| \leq |\mathbf{s}|} (D^{\mathbf{r}} u)(\mathbf{x}') \sum_{\substack{\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell} \\ 1 \leq \ell \leq |\mathbf{r}|}} P_{\mathbf{s},\mathbf{r},\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}}(\mathbf{z}) \prod_{\mathbf{t} \in \mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}} D^{\mathbf{t}} \varepsilon(\mathbf{x}). \tag{4.2}$$

We employ the convention that empty sums take the value zero and empty products the value 1. *Furthermore, if $P_{\mathbf{s},\mathbf{r},\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}}$ is constant, then $P_{\mathbf{s},\mathbf{r},\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}} \equiv 0$.*

Proof of Lemma 4.4. Introduce the function $\widetilde{u}$ as

$$\widetilde{u}(\mathbf{x}) := u(\mathbf{x} + \mathbf{z}\varepsilon(\mathbf{x})). \tag{4.3}$$

We use the shorthand $\partial_i = \partial \cdot / \partial x_i$ and compute for $|\mathbf{s}| = 1$

$$\begin{aligned} \partial_i \widetilde{u}(\mathbf{x}) &= (\partial_i u)(\mathbf{x}')(1 + \mathbf{z}_i \partial_i \varepsilon(\mathbf{x})) + \sum_{j \neq i} (\partial_j u)(\mathbf{x}') \mathbf{z}_j \partial_i \varepsilon(\mathbf{x}) \\ &= (\partial_i u)(\mathbf{x}') + \sum_{j=1}^{d} (\partial_j u)(\mathbf{x}') \mathbf{z}_j \partial_i \varepsilon(\mathbf{x}), \end{aligned}$$

which we recognize to be of the form (4.2). We now proceed by induction. To that end, we assume that formula (4.2) is true for all multiindices $\mathbf{s}' \in \mathbb{N}_0^d$ with $|\mathbf{s}'| \leq n$. Then for $\mathbf{s} = (\mathbf{s}'_1, \ldots, \mathbf{s}'_{i-1}, \mathbf{s}'_i + 1, \mathbf{s}'_{i+1}, \ldots, \mathbf{s}'_d)$ we compute with the induction hypothesis:

$$\begin{aligned} D^{\mathbf{s}} \widetilde{u}(\mathbf{x}') &= \partial_i D^{\mathbf{s}'} \widetilde{u}(\mathbf{x}') = (D^{\mathbf{s}} u)(\mathbf{x}') + \sum_{j=1}^{d} (D^{\mathbf{s}'} \partial_j u)(\mathbf{x}') \mathbf{z}_j \partial_i \varepsilon(\mathbf{x}) \\ &\quad + \partial_i \left(\sum_{|\mathbf{r}| \leq |\mathbf{s}'|} (D^{\mathbf{r}} u)(\mathbf{x}') \sum_{\substack{\mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell} \\ 1 \leq \ell \leq |\mathbf{r}|}} P_{\mathbf{s},\mathbf{r},\mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell}}(\mathbf{z}) \prod_{\mathbf{t} \in \mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell}} D^{\mathbf{t}} \varepsilon(\mathbf{x}) \right) =: T_1 + T_2 + T_3. \end{aligned}$$

Hence, $T_1 + T_2$ consists of terms of the desired form. For the term T_3, we compute

$$
\begin{aligned}
\partial_i (D^{\mathbf{r}} u)(\mathbf{x}') &= (D^{\mathbf{r}} \partial_i u)(\mathbf{x}') + \sum_{j'=1}^{d} (D^{\mathbf{r}} \partial_{j'} u)(\mathbf{x}') \mathbf{z}_{j'} \partial_i \varepsilon(\mathbf{x}) \\
\partial_i \prod_{\mathbf{t} \in \mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell}} D^{\mathbf{t}} \varepsilon(\mathbf{x}) &= \sum_{\mathbf{t} \in \mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell}} (D^{\mathbf{t}} \partial_i \varepsilon(\mathbf{x})) \prod_{\substack{\mathbf{t}' \in \mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell} \\ \mathbf{t}' \neq \mathbf{t}}} D^{\mathbf{t}'} \varepsilon(\mathbf{x}).
\end{aligned}
$$

Hence, T_3 has the form

$$
\begin{aligned}
T_3 = {} & \sum_{|\mathbf{r}| \leq |\mathbf{s}'|} D^{\mathbf{r}} \partial_i u(\mathbf{x}') \sum_{\substack{\mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell} \\ 1 \leq \ell \leq |\mathbf{r}|}} P_{\mathbf{s}',\mathbf{r},\mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell}}(\mathbf{z}) \prod_{\mathbf{t} \in \mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell}} D^{\mathbf{t}} \varepsilon(\mathbf{x}) \\
& + \sum_{|\mathbf{r}| \leq |\mathbf{s}'|} \sum_{j'=1}^{d} D^{\mathbf{r}} \partial_{j'} u(\mathbf{x}') \sum_{\substack{\mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell} \\ 1 \leq \ell \leq |\mathbf{r}|}} \mathbf{z}_{j'} \partial_i \varepsilon(\mathbf{x}) P_{\mathbf{s}',\mathbf{r},\mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell}}(\mathbf{z}) \prod_{\mathbf{t} \in \mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell}} D^{\mathbf{t}} \varepsilon(\mathbf{x}) \\
& + \sum_{|\mathbf{r}| \leq |\mathbf{s}'|} D^{\mathbf{r}} u(\mathbf{x}') \sum_{\substack{\mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell} \\ 1 \leq \ell \leq |\mathbf{r}|}} P_{\mathbf{s}',\mathbf{r},\mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell}}(\mathbf{z}) \sum_{\mathbf{t} \in \mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell}} (D^{\mathbf{t}} \partial_i \varepsilon(\mathbf{x})) \prod_{\substack{\mathbf{t}' \in \mathcal{M}_{|\mathbf{s}'|-|\mathbf{r}|,\ell} \\ \mathbf{t}' \neq \mathbf{t}}} D^{\mathbf{t}'} \varepsilon(\mathbf{x}).
\end{aligned}
$$

Since $|\mathbf{s}| = |\mathbf{s}'| + 1$, we see that each of the three sums has the stipulated form. This concludes the induction argument. $\square$

For the function $\widetilde{u}$ given by (4.3), the next lemma quantifies the difference $u - \widetilde{u}$ if ε is a Λ-admissible length scale function.

Lemma 4.5. *Let $\Omega \subset \mathbb{R}^d$ be a domain, and let $\varepsilon \in C^\infty(\Omega)$ be a Λ-admissible length scale function. Let $u \in C^\infty(\mathbb{R}^d)$. Then, for a multiindex $\mathbf{s} \in \mathbb{N}_0^d$, the derivative $D^{\mathbf{s}}_{\mathbf{x}} u(\mathbf{x}')$, $\mathbf{x}' := \mathbf{x} + \mathbf{z}\varepsilon(\mathbf{x})$, can be written in the form*

$$
D^{\mathbf{s}}_{\mathbf{x}} u(\mathbf{x}') = (D^{\mathbf{s}} u)(\mathbf{x}') + \sum_{|\mathbf{r}| \leq |\mathbf{s}|} (D^{\mathbf{r}} u)(\mathbf{x}') E_{\mathbf{s},\mathbf{r}}(\mathbf{z}, \mathbf{x}), \tag{4.4}
$$

where the smooth functions $E_{\mathbf{s},\mathbf{r}}$ are polynomials of degree $|\mathbf{r}|$ in the first component and satisfy

$$
\sup_{|\mathbf{z}| \leq R} |D^{\mathbf{t}}_{\mathbf{z}} E_{\mathbf{s},\mathbf{r}}(\mathbf{z}, \mathbf{x})| \leq C_{\Lambda,\mathbf{s},R} |\varepsilon(\mathbf{x})|^{|\mathbf{r}|-|\mathbf{s}|} \qquad \forall \mathbf{x} \in \Omega. \tag{4.5}
$$

The constants $C_{\Lambda,\mathbf{s},R}$ depend only on $\mathbf{s}$, R, and $(\Lambda_{\mathbf{s}'})_{|\mathbf{s}'| \leq |\mathbf{s}|}$.

Proof. Apply Lemma 4.4 and define

$$
E_{\mathbf{r}}(\mathbf{z}, \mathbf{x}) = \sum_{\substack{\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell} \\ 1 \leq \ell \leq |\mathbf{r}|}} P_{\mathbf{s},\mathbf{r},\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}}(\mathbf{z}) \prod_{\mathbf{t} \in \mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}} D^{\mathbf{t}} \varepsilon(\mathbf{x}).
$$

Clearly, $E_{\mathbf{r}}$ is smooth and is a polynomial of degree $\mathbf{r}$ in the first component. Let $\mathbf{x} \in \Omega$. According to the properties of a Λ-admissible length scale function and the definition of the set $\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}$ for

$1 \le \ell \le |\mathbf{r}|$, there holds

$$\prod_{\mathbf{t}\in\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}} \left|D^{\mathbf{t}}\varepsilon(\mathbf{x})\right| \le \prod_{\mathbf{t}\in\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}} \Lambda_{\mathbf{t}} \left|\varepsilon(\mathbf{x})\right|^{1-|\mathbf{t}|} \\ \le \max_{|\mathbf{s}'|\le|\mathbf{s}|} (\Lambda_{\mathbf{s}'}) \prod_{\mathbf{t}\in\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}} \left|\varepsilon(\mathbf{x})\right|^{1-|\mathbf{t}|} = \max_{|\mathbf{s}'|\le|\mathbf{s}|} (\Lambda_{\mathbf{s}'}) \left|\varepsilon(\mathbf{x})\right|^{|\mathbf{r}|-|\mathbf{s}|}. \tag{4.6}$$

We can conclude the proof by setting

$$C_{\Lambda,\mathbf{s},R} := \max_{|\mathbf{s}'|\le|\mathbf{s}|} (\Lambda_{\mathbf{s}'}) \sup_{|\mathbf{z}|\le R} \sum_{\substack{\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell} \\ 1\le\ell\le|\mathbf{r}|}} \left|D^{\mathbf{t}} P_{\mathbf{s},\mathbf{r},\mathcal{M}_{|\mathbf{s}|-|\mathbf{r}|,\ell}}(\mathbf{z})\right|.$$

□

5 Higher-order volume regularization

5.1 Regularization on a reference ball

Throughout this section, ε denotes a Λ-admissible length scale function with Lipschitz constant L_ε. From Definition 2.1 (iii) it follows that $L_\varepsilon \le \max_{|\mathbf{r}|=1} \Lambda_{\mathbf{r}}$, and the right-hand side depends only on $(\Lambda_{\mathbf{r}})_{|\mathbf{r}|=1}$. Our goal is to construct operators that, given $\mathbf{z} \in \mathbb{R}^d$, employ regularization on a length scale $\varepsilon(\mathbf{z})$, which therefore determines the quality of the approximation at $\mathbf{z}$. We will analyze such operators on balls $B_r(\mathbf{x})$ that have a radius which is comparable to the value of $\varepsilon(\mathbf{x})$. Hence, it will be convenient to use reference configurations and scaling arguments. For fixed $\mathbf{x}$, we define the scaling map

$$T_{\mathbf{x}} : \mathbf{z} \mapsto \mathbf{x} + \varepsilon(\mathbf{x})\mathbf{z}.$$

In classical finite element approximation theory, the *pull-back* $u \circ T_{\mathbf{x}}$ of a function u is approximated on a reference configuration. This approximation is also analyzed on the reference configuration, and scaling arguments provide the current quality of the approximation to u (given by powers of the underlying length scale). As stated above, the construction that is carried out in this work defines the approximation of the pull-back also in terms of the local length scale. In order to obtain a fixed length scale on our reference configuration, i.e., to make the approximation properties on the reference configuration independent of a specific length scale, it will be convenient to define the function $\varepsilon_{\mathbf{x}}$ by

$$\mathbf{z} \mapsto \varepsilon_{\mathbf{x}}(\mathbf{z}) := \frac{\varepsilon(T_{\mathbf{x}}(\mathbf{z}))}{\varepsilon(\mathbf{x})}.$$

The next lemma shows that $\varepsilon_{\mathbf{x}}$ does, in essence, only depend on Λ, but not on $\mathbf{x}$. We construct parameters $\alpha, \beta, \delta \in \mathbb{R}$ where δ will be used to define the regularization operator and α, β will be used to define balls on which the regularization error will be analyzed. In subsequent sections, these parameters need to be adjusted also according to properties of the domain of interest Ω, more precisely, its Lipschitz character and in particular the Lipschitz constant $L_{\partial\Omega}$ of $\partial\Omega$. Hence, the parameters α, β, δ will be chosen with in dependence on $L_{\partial\Omega}$ and an additional parameter L.

Lemma 5.1. *Let $\Omega \subset \mathbb{R}^d$ be an (arbitrary) domain and $\varepsilon \in C^\infty(\Omega)$ be a Λ-admissible length scale function with Lipschitz constant L_ε. Clearly, L_ε depends solely on $(\Lambda_{\mathbf{r}})_{|\mathbf{r}|=1}$. Then:*

(i) For fixed $\alpha \in \left(0, \min\left(1, L_\varepsilon^{-1}/2\right)\right)$ let $\mathbf{x} \in \Omega$ be such that $T_{\mathbf{x}}(B_\alpha) \subset \Omega$. Then

$$2^{-1} \leq \varepsilon_{\mathbf{x}}(\mathbf{z}) \leq 2 \qquad \text{for all } \mathbf{z} \in B_\alpha. \tag{5.1}$$

(ii) One may choose $0 < \alpha,\ \delta,\ \beta < \min\left(1, L_\varepsilon^{-1}/2\right)$ with

$$2\delta + \alpha < \beta < \min\left(1, L_\varepsilon^{-1}/2\right). \tag{5.2}$$

The parameters α, β, and δ depend only on $(\Lambda_{\mathbf{r}})_{|\mathbf{r}|=1}$.

(iii) Given (arbitrary) parameters $L_{\partial\Omega}$, $L > 0$ one may choose $\alpha,\ \beta,\ \delta > 0$ such that

$$\max\left((L_{\partial\Omega}+1)(\delta+\alpha)+\delta, \alpha+\delta(L+1)(1+L_\varepsilon\alpha)\right) < \beta < \min\left(1, L_\varepsilon^{-1}/2\right). \tag{5.3}$$

The parameters depend only on $(\Lambda_{\mathbf{r}})_{|\mathbf{r}|=1}$ and on L, $L_{\partial\Omega}$.

(iv) Let $K \subset \Omega$ be compact. Then the parameters α, β, δ can be chosen such that in addition to (5.3) the following additional property holds:

$$\mathbf{x} \in \Omega \qquad \Longrightarrow \left(\text{ either } \overline{B_{2\beta}(\mathbf{x})} \subset \Omega \qquad \text{or } \overline{B_{2\beta}(\mathbf{x})} \cap K = \emptyset \right). \tag{5.4}$$

Proof. Fix $0 < \alpha < \min(1, L_\varepsilon^{-1}/2)$ and $\mathbf{x} \in \Omega$ such that $T_{\mathbf{x}}(B_\alpha) \subset \Omega$. For $\mathbf{z} \in B_\alpha$ we conclude with the reverse triangle inequality

$$\pm\left(\varepsilon(T_{\mathbf{x}}(\mathbf{z})) - \varepsilon(\mathbf{x})\right) \leq |\varepsilon(T_{\mathbf{x}}(\mathbf{z})) - \varepsilon(\mathbf{x})| \leq L_\varepsilon |\mathbf{x} - T_{\mathbf{x}}(\mathbf{z})| \leq L_\varepsilon \alpha \varepsilon(\mathbf{x}).$$

This yields

$$\frac{\varepsilon(T_{\mathbf{x}}\mathbf{z})}{\varepsilon(\mathbf{x})} \leq L_\varepsilon\alpha + 1 \qquad \text{and} \qquad \frac{\varepsilon(T_{\mathbf{x}}\mathbf{z})}{\varepsilon(\mathbf{x})} \geq 1 - L_\varepsilon\alpha,$$

from which (5.1) follows. The additional features (5.3) can be achieved by adjusting α and δ.

Finally, (iv) follows from compactness of K. $\square$

Next, we will prove an approximation result on the reference configuration.

Lemma 5.2. *Let $\Omega \subset \mathbb{R}^d$ be an (arbitrary) domain. Let ρ be a mollifier of order $k_{\max}$, and let $\varepsilon \in C^\infty(\Omega)$ be a Λ-admissible length scale function. Choose α, β, δ such that (5.2) holds. For $\mathbf{x} \in \Omega$ such that $T_{\mathbf{x}}(B_\alpha) \subset \Omega$ and a function $v \in L^1_{loc}(B_\beta)$ define*

$$\mathbf{z} \mapsto \mathcal{E}_{\mathbf{x}} v(\mathbf{z}) := \int_{\mathbf{y} \in B_\beta} v(\mathbf{y}) \rho_{\delta\varepsilon_{\mathbf{x}}(\mathbf{z})}(\mathbf{z} - \mathbf{y}). \tag{5.5}$$

(i) Let $(s,p) \in \mathbb{N}_0 \times [1,\infty]$ satisfy $s \leq k_{\max} + 1$. Assume ($s \leq r \in \mathbb{R}$ and $q \in [1,\infty)$) or ($s \leq r \in \mathbb{N}_0$ and $q \in [1,\infty]$). Then it holds

$$|\mathcal{E}_{\mathbf{x}} v|_{r,q,B_\alpha} \leq C_{r,q,s,p,\Lambda} |v|_{s,p,B_\beta}, \tag{5.6a}$$

where $C_{r,q,s,p,\Lambda}$ depends only on r, q, s, p and $(\Lambda_{\mathbf{r}'})_{|\mathbf{r}'| \leq \lceil r \rceil}$ as well as ρ, k_{max}, and α, β.

(ii) Suppose $0 \le s \in \mathbb{R}$, $r \in \mathbb{N}_0$ *with* $s \le r \le k_{\max}+1$, *and* $1 \le p \le q < \infty$. *Define* $\mu := d(p^{-1} - q^{-1})$. *Assume that* ($r = s + \mu$ *and* $p > 1$) *or* ($r > s + \mu$). *Then it holds that*

$$|v - \mathcal{E}_{\mathbf{x}} v|_{s,q,B_\alpha} \le C_{s,q,r,p,\Lambda} |v|_{r,p,B_\beta}, \tag{5.6b}$$

where $C_{s,q,r,p,\Lambda}$ *depends only on* s, q, r, p, *and* $(\Lambda_{\mathbf{s}'})_{|\mathbf{s}'| \le \lceil s \rceil}$ *as well as* ρ, k_{max}, *and* α, β.

(iii) Suppose s, $r \in \mathbb{N}_0$ *with* $s \le r \le k_{\max}+1$, *and* $1 \le p < \infty$. *Define* $\mu := d/p$. *Assume that* ($r = s + \mu$ *and* $p = 1$) *or* ($r > s + \mu$ *and* $p > 1$). *Then it holds that*

$$|v - \mathcal{E}_{\mathbf{x}} v|_{s,\infty,B_\alpha} \le C_{s,r,p,\Lambda} |v|_{r,p,B_\beta}, \tag{5.6c}$$

where $C_{s,r,p,\Lambda}$ *depends only on* s, r, p, *and* $(\Lambda_{\mathbf{s}'})_{|\mathbf{s}'| \le s}$ *as well as* ρ, k_{max}, *and* α, β.

Proof. For a multi-index $\mathbf{r}$ we have

$$|D^{\mathbf{r}}_{\mathbf{z}} \varepsilon_{\mathbf{x}}(\mathbf{z})| = \varepsilon(\mathbf{x})^{-1} |D^{\mathbf{r}}_{\mathbf{z}} \varepsilon(\mathbf{x} + \varepsilon(\mathbf{x})\mathbf{z})| = \varepsilon(\mathbf{x})^{-1} |(D^{\mathbf{r}}\varepsilon)(\mathbf{x} + \varepsilon(\mathbf{x})\mathbf{z})| \, |\varepsilon(\mathbf{x})|^{|\mathbf{r}|} \le \Lambda_{\mathbf{r}} |\varepsilon_{\mathbf{x}}(\mathbf{z})|^{1-|\mathbf{r}|}, \tag{5.7}$$

from which we conclude that $\varepsilon_{\mathbf{x}} \in C^\infty(B_\alpha)$ is also a Λ-admissible length scale function. In view of density, we may assume that $v \in C^\infty(\overline{B_\beta})$, so that we can interchange differentiation and integration. Setting $\mathbf{z}' := \mathbf{z} - \delta\varepsilon(\mathbf{z})\mathbf{y}$, the Faà di Bruno formula from Lemma 4.5 shows

$$\begin{aligned} D^{\mathbf{r}}_{\mathbf{z}} v(\mathbf{z}') &= (D^{\mathbf{r}} v)(\mathbf{z}') + \sum_{|\mathbf{t}| \le |\mathbf{r}|} (D^{\mathbf{t}} v)(\mathbf{z}') \, E_{\mathbf{r},\mathbf{t}}(\delta\mathbf{y}, \mathbf{z}) \\ &= (\delta\varepsilon_{\mathbf{x}}(\mathbf{z}))^{-|\mathbf{r}|} D^{\mathbf{r}}_{\mathbf{y}} v(\mathbf{z}') + \sum_{|\mathbf{t}| \le |\mathbf{r}|} (\delta\varepsilon_{\mathbf{x}}(\mathbf{z}))^{-|\mathbf{t}|} D^{\mathbf{t}}_{\mathbf{y}} v(\mathbf{z}') \, E_{\mathbf{r},\mathbf{t}}(\delta\mathbf{y}, \mathbf{z}). \end{aligned} \tag{5.8}$$

We obtain with integration by parts, the product rule, and the support properties of ρ

$$\begin{aligned} D^{\mathbf{r}}_{\mathbf{z}} \mathcal{E}_{\mathbf{x}} v(\mathbf{z}) = (-1)^{|\mathbf{r}|} &\int_{\mathbf{y} \in B_1(0)} v(\mathbf{z}') \, D^{\mathbf{r}}\rho(\mathbf{y}) \, (\delta\varepsilon_{\mathbf{x}}(\mathbf{z}))^{-|\mathbf{r}|} \\ &+ \sum_{|\mathbf{t}| \le |\mathbf{r}|} (-1)^{|\mathbf{t}|} \sum_{\mathbf{s} \le \mathbf{t}} \binom{\mathbf{s}}{\mathbf{t}} \int_{\mathbf{y} \in B_1(0)} v(\mathbf{z}') D^{\mathbf{s}}_{\mathbf{y}} E_{\mathbf{r},\mathbf{t}}(\delta\mathbf{y}, \mathbf{z}) D^{\mathbf{t}-\mathbf{s}}\rho(\mathbf{y}) \, (\delta\varepsilon_{\mathbf{x}}(\mathbf{z}))^{-|\mathbf{t}|}. \end{aligned}$$

Taking into account Lemmas 4.5 and 5.1, we obtain for $r \in \mathbb{N}_0$ the estimate

$$|\mathcal{E}_{\mathbf{x}} v|_{r,\infty,B_\alpha} \le C_{\Lambda,r} \|v\|_{0,1,B_\beta}. \tag{5.9}$$

Hölder's inequality then shows (5.6a) for the case $s = 0$ and integer r. For $r \notin \mathbb{N}_0$ and $q \in [1, \infty)$, we use the Embedding Theorem 4.2 (although for the present case of the ball B_α, a simpler argument could be used) in the following way: observing $0 < \lceil r \rceil - r < 1$, we select $1 < p_\star < q$ such that

$$\left(\frac{\lceil r \rceil - r}{d} > \frac{1}{p_\star} - \frac{1}{q} \right).$$

Then the Embedding Theorem 4.2 and estimate (5.9) show

$$|\mathcal{E}_{\mathbf{x}} v|_{r,q,B_\alpha} \lesssim \|\mathcal{E}_{\mathbf{x}} v\|_{\lceil r \rceil, p_\star, B_\alpha} \lesssim C_{\Lambda,\lceil r \rceil} \|v\|_{0,1,B_\beta}. \tag{5.10}$$

Again, Hölder's inequality shows (5.6a) for the case $s = 0$. Next, let $1 \le s \le k_{\max} + 1$ and $s \le r$. Then for any polynomial $\pi \in \mathcal{P}_{s-1}(B_\beta)$ we have $|\pi|_{r,q,B_\beta} = 0$ and $\mathcal{E}_{\mathbf{x}}\pi = \pi$. Estimates (5.9) or (5.10) and the Bramble-Hilbert lemma 4.3 then show

$$|\mathcal{E}_{\mathbf{x}} v|_{r,q,B_\alpha} \le C_{\Lambda,\lceil r\rceil} \inf_{\pi\in\mathcal{P}_{s-1}(B_\beta)} \|v-\pi\|_{0,1,B_\beta} \le C_{\Lambda,r,s} |v|_{s,1,B_\beta},$$

from which (5.6a) follows for $s \ge 1$ again by application of Hölder's inequality.

Proof of (ii), (iii): Since $\mathcal{E}_{\mathbf{x}}\pi = \pi$ for any polynomial $\pi \in \mathcal{P}_{r-1}(B_\beta)$, the triangle inequality yields

$$\begin{aligned}|v - \mathcal{E}_{\mathbf{x}} v|_{s,q,B_\alpha} &\le |v-\pi|_{s,q,B_\alpha} + |\mathcal{E}_{\mathbf{x}}(v-\pi)|_{s,q,B_\alpha} \\ &\lesssim \|v-\pi\|_{r,p,B_\alpha} + |v-\pi|_{\lfloor s\rfloor,p,B_\beta} \le \|v-\pi\|_{r,p,B_\beta},\end{aligned}$$

where we used the embedding results of Theorem 4.2 as well as (5.6a) in the second step. The estimates (5.6b) and (5.6c) now again follow by application of the Bramble-Hilbert lemma 4.3. $\square$

5.2 Regularization in the interior of Ω

With the results of the last section, we can define a regularization operator that will determine $\mathcal{I}_\varepsilon$ of Theorem 2.2 in the interior of Ω; near $\partial\Omega$, we will need a modification introduced in Section 5.3 below. The properties of $\mathcal{I}_\varepsilon$ in the interior will be analyzed using scaling arguments. We show that the regularization operator $\mathcal{E}$ satisfies inverse estimates (Lemma 5.3, (i)) in addition to having approximation properties (Lemma 5.3, (ii) and (iii)).

Lemma 5.3. *Let ρ be a mollifier of order k_{max}, $p \in [1,\infty)$, Let $\Omega \subset \mathbb{R}^d$ be an (arbitrary) domain and let $\varepsilon \in C^\infty(\Omega)$ be a Λ-admissible length scale function. Choose α, β, δ according to Lemma 5.1. Define*

$$\Omega_\varepsilon := \{\mathbf{x} \in \Omega \,|\, T_{\mathbf{x}}(B_\beta) \subset \Omega\}.$$

For a function $u \in L^1_{loc}(\mathbb{R}^d)$ define

$$\mathcal{E}u(\mathbf{z}) := \int_{\mathbf{y}} u(\mathbf{y})\rho_{\delta\varepsilon(\mathbf{z})}(\mathbf{z}-\mathbf{y}), \qquad \text{for } \mathbf{z} \in \Omega.$$

Then, for $\mathbf{x} \in \Omega_\varepsilon$:

(i) Suppose $(s,p) \in \mathbb{N}_0 \times [1,\infty]$ satisfies $s \le k_{\max}+1$. Assume ($s \le r \in \mathbb{R}$ and $q \in [1,\infty)$) or ($s \le r \in \mathbb{N}_0$ and $q \in [1,\infty]$). Then it holds

$$|\mathcal{E}u|_{r,q,B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})} \le C_{r,q,s,p,\Lambda}\varepsilon(\mathbf{x})^{s-r+d(1/q-1/p)} |u|_{s,p,B_{\beta\varepsilon(\mathbf{x})}(\mathbf{x})}, \tag{5.11a}$$

where $C_{r,q,s,p,\Lambda}$ depends only on r, q, s, p and $(\Lambda_{\mathbf{r}'})_{|\mathbf{r}'|\le\lceil r\rceil}$ as well as ρ, $k_{\max}$, and α, β, δ.

(ii) Suppose $0 \le s \in \mathbb{R}$, $r \in \mathbb{N}_0$ with $s \le r \le k_{\max}+1$, and $1 \le p \le q < \infty$. Define $\mu := d(p^{-1} - q^{-1})$. Assume that ($r = s + \mu$ and $p > 1$) or ($r > s + \mu$). Then it holds that

$$|u - \mathcal{E}u|_{s,q,B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})} \le C_{s,q,r,p,\Lambda}\varepsilon(\mathbf{x})^{r-s+d(1/q-1/p)} |u|_{r,p,B_{\beta\varepsilon(\mathbf{x})}(\mathbf{x})}, \tag{5.11b}$$

where $C_{s,q,r,p,\Lambda}$ depends only on s, q, r, p, and $(\Lambda_{\mathbf{s}'})_{|\mathbf{s}'|\le\lceil s\rceil}$ as well as ρ, $k_{\max}$, and α, β, δ.

(iii) Suppose s, $r \in \mathbb{N}_0$ with $s \leq r \leq k_{\max}+1$, and $1 \leq p < \infty$. Define $\mu := d/p$. Assume that ($r = s+\mu$ and $p = 1$) or ($r > s+\mu$ and $p > 1$). Then it holds that

$$|v - \mathcal{E}_{\mathbf{x}} v|_{s,\infty,B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})} \leq C_{s,r,p,\Lambda}\varepsilon(\mathbf{x})^{r-s-d/p} |v|_{r,p,B_{\beta\varepsilon(\mathbf{x})}(\mathbf{x})} . \tag{5.11c}$$

where $C_{s,r,p,\Lambda}$ depends only on s, r, p, and $(\Lambda_{\mathbf{s}'})_{|\mathbf{s}'|\leq s}$ as well as ρ, $k_{\max}$, and α, β, δ.

Proof. First, we check that $(\mathcal{E}u) \circ T_{\mathbf{x}} = \mathcal{E}_{\mathbf{x}}(u \circ T_{\mathbf{x}})$. To that end, let $\mathbf{x} \in \Omega_\varepsilon$ and $\mathbf{z} \in B_\beta$. We employ the substitution $\mathbf{y} = T_{\mathbf{x}}(\mathbf{w})$ and note that $T_{\mathbf{x}}(\mathbf{z}) - T_{\mathbf{x}}(\mathbf{w}) = \varepsilon(\mathbf{x})(\mathbf{z}-\mathbf{w})$:

$$\begin{aligned}
(\mathcal{E}u)(T_{\mathbf{x}}(\mathbf{z})) &= \int_{\mathbf{y}} u(\mathbf{y})\rho\left(\frac{T_{\mathbf{x}}(\mathbf{z})-\mathbf{y}}{\delta\varepsilon(T_{\mathbf{x}}(\mathbf{z}))}\right)\varepsilon(T_{\mathbf{x}}(\mathbf{z}))^{-d} \\
&= \int_{\mathbf{w}} u(T_{\mathbf{x}}\mathbf{w})\rho\left(\frac{T_{\mathbf{x}}(\mathbf{z})-T_{\mathbf{x}}(\mathbf{w})}{\delta\varepsilon(T_{\mathbf{x}}(\mathbf{z}))}\right)\varepsilon(T_{\mathbf{x}}(\mathbf{z}))^{-d}\varepsilon(\mathbf{x})^{d} \\
&= \int_{\mathbf{w}} (u\circ T_{\mathbf{x}})(\mathbf{w})\rho\left(\frac{\mathbf{z}-\mathbf{w}}{\delta\varepsilon_{\mathbf{x}}(\mathbf{z})}\right)\varepsilon_{\mathbf{x}}(\mathbf{z})^{-d} = \mathcal{E}_{\mathbf{x}}(u\circ T_{\mathbf{x}})(\mathbf{z}).
\end{aligned}$$

Now, the estimates (5.11) follow from Lemma 5.2 using scaling arguments, cf. [Heu14]. □

5.3 Regularization on a half-space with a Lipschitz-boundary

Lemma 5.3 focusses on regularizing a given function v in the *interior* of Ω. If we want to take full advantage of v's regularity, this approach does not

extend up to the boundary. The reason lies in the construction: the value $(\mathcal{E}v)(\mathbf{x})$ is defined by an averaging process in a ball of radius $\delta\varepsilon(\mathbf{x})$ around $\mathbf{x}$ and thus requires v to be defined in the ball $B_{\delta\varepsilon(\mathbf{x})}(\mathbf{x})$. Hence, it cannot be defined in the point $\mathbf{x}$ if $\delta\varepsilon(\mathbf{x})$ is bigger than $\mathbf{x}$'s distance to the boundary $\partial\Omega$. In the present section, we therefore propose a modification of the averaging operator that is based on averaging not on the ball $B_{\delta\varepsilon(\mathbf{x})}(\mathbf{x})$ but on the ball $\mathbf{b}+B_{\delta\varepsilon(\mathbf{x})}(\mathbf{x})$, where the vector $\mathbf{b}$ (which depends on $\mathbf{x}$ and $\varepsilon(\mathbf{x})$) is such that $\mathbf{b}+B_{\delta\varepsilon(\mathbf{x})}(\mathbf{x}) \subset \Omega$ (see (5.15) for the precise definition), and an application of Taylor's formula.

Following Stein, [Ste70], we call Ω a special Lipschitz domain, if it has the form

$$\Omega := \left\{\mathbf{x} \in \mathbb{R}^d \mid x_d > f_{\partial\Omega}(\mathbf{x}_{d-1})\right\},$$

where $f_{\partial\Omega} : \mathbb{R}^{d-1} \to \mathbb{R}$ is a Lipschitz continuous function with Lipschitz constant $L_{\partial\Omega}$. In this coordinate system, every point $\mathbf{x} \in \mathbb{R}^d$ has the form $\mathbf{x} = (\mathbf{x}_{d-1}, x_d) \in \mathbb{R}^d$. For each $\mathbf{x} = (\mathbf{x}_{d-1}, x_d) \in \mathbb{R}^d$, we can then define the set

$$\mathfrak{C}_{\mathbf{x}} := \{(\mathbf{y}_{d-1}, y_d) \in \mathbb{R}^d \colon (y_d - x_d) > L_{\partial\Omega}|\mathbf{y}_{d-1} - \mathbf{x}_{d-1}|\}.$$

It is easy to see that $\mathfrak{C}_{\mathbf{x}}$ is a convex cone with apex at $\mathbf{x}$. It will also be convenient to introduce the vector $\mathbf{e}_d = (0, 0, \dots, 1) \in \mathbb{R}^d$. We will again analyze the regularization operator near the boundary on balls $B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})$. Due to the construction mentioned above, the regularization operator takes the values of the underlying function on sets which are not balls anymore. We call these sets $\veebar_{\mathbf{x}}$ (cf. (5.12)). The next two lemmas analyze their properties.

Lemma 5.4. *Let Ω be a special Lipschitz domain with Lipschitz constant $L_{\partial\Omega}$. The following two statements hold:*

(i) *If $\mu > 0$ and $\tau > (L_{\partial\Omega} + 1)(\mu + 1)$, then for all $r > 0$ and $\mathbf{x} \in \Omega$ the set*

$$\mathfrak{C}'_{\mathbf{x}} := \bigcup_{\mathbf{z} \in B_r(\mathbf{x}) \cap \Omega} \mathfrak{C}_{\mathbf{z}}$$

is star-shaped with respect to the ball $B_{\mu r}(\mathbf{x} + \tau r \mathbf{e}_d)$.

(ii) *There is a constant $L > 1$, which depends only on $L_{\partial\Omega}$, such that for all $r > 0$ and $\mathbf{x} \in \Omega$, all $\mathbf{z} \in B_r(\mathbf{x})$, and all $r_0 > 0$ it holds that*

$$B_{r_0}(\mathbf{z} + L r_0 \mathbf{e}_d) \subset \mathfrak{C}'_{\mathbf{x}}.$$

Proof. First we show (i): For $\mu > 0$, suppose that $\tau > (L_{\partial\Omega} + 1)(\mu + 1)$. Now let $r > 0$ and $\mathbf{x} \in \Omega$ be arbitrary. It suffices to show that $B_{\mu r}(\mathbf{x} + \tau r \mathbf{e}_d) \subset \mathfrak{C}_{\mathbf{z}}$ for all $\mathbf{z} \in B_r(\mathbf{x})$. Therefore, let $\mathbf{y} \in B_{\mu r}(\mathbf{x} + \tau r \mathbf{e}_d)$ and note that

$$\begin{aligned} y_d - z_d = y_d - (x_d + \tau r) + (x_d + \tau r - x_d) + x_d - z_d \geq -\mu r + \tau r - r = (\tau - \mu - 1) r \\ |\mathbf{y}_{d-1} - \mathbf{z}_{d-1}| \leq |\mathbf{y}_{d-1} - \mathbf{x}_{d-1}| + |\mathbf{x}_{d-1} - \mathbf{z}_{d-1}| \leq \mu r + r = (\mu + 1) r. \end{aligned}$$

Due to the choice of τ, we conclude the statement. To conclude *(ii)*, we choose $L > L_{\partial\Omega} + 1$. For $\mathbf{y} \in B_{r_0}(\mathbf{z} + L r_0 \mathbf{e}_d)$, we compute

$$y_d - z_d \geq L r_0 - r_0 = (L - 1) r_0 \quad \text{and} \quad |\mathbf{y}_{d-1} - \mathbf{z}_{d-1}| \leq r_0,$$

and conclude $\mathbf{y} \in \mathfrak{C}_{\mathbf{z}}$, from which (ii) follows. $\square$

Lemma 5.5. *Let Ω be a special Lipschitz domain with Lipschitz constant $L_{\partial\Omega}$ and let $\varepsilon \in C^\infty(\Omega)$ be a Λ-admissible length scale function. Fix a compact set $K \subset \Omega$. Then there are α, β, δ, L, $\tau > 0$ such that* (5.1)–(5.4) *hold, and such that with*

$$\veebar_{\mathbf{x}} := B_{\beta\varepsilon(\mathbf{x})}(\mathbf{x}) \cap \bigcup_{\mathbf{z} \in B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x}) \cap \Omega} \mathfrak{C}_{\mathbf{z}} \quad \text{for } \mathbf{x} \in \Omega, \tag{5.12}$$

the following statements (i)–(iii) are true:

(i) *For $\mathbf{x}_0 \in \Omega$, the point $\mathbf{x}_0 + \delta\varepsilon(\mathbf{x}_0) L \mathbf{e}_d$ satisfies*

$$\operatorname{dist}(\mathbf{x}_0 + \delta\varepsilon(\mathbf{x}_0) L \mathbf{e}_d, \partial\Omega) > \delta\varepsilon(\mathbf{x}_0).$$

(ii) *For every $\mathbf{x} \in \Omega$ and every $\mathbf{x}_0 \in B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x}) \cap \Omega$ there holds $B_{\delta\varepsilon(\mathbf{x}_0)}(\mathbf{x}_0 + L\delta\varepsilon(\mathbf{x}_0)\mathbf{e}_d) \subset \veebar_{\mathbf{x}}$.*

(iii) *The set $\veebar_{\mathbf{x}}$ is star-shaped with respect to $B_{\delta\varepsilon(\mathbf{x})}(\mathbf{x} + \tau\alpha\varepsilon(\mathbf{x})\mathbf{e}_d)$.*

Proof. First, choose L from Lemma (5.4), (ii). Then, choose α, β, $\delta > 0$ according to Lemma 5.1. As $\beta > (L_{\partial\Omega} + 1)(\delta/\alpha + 1)\alpha + \delta$, we can choose τ from Lemma 5.4, (i), where we set $\mu = \delta/\alpha$, in a way such that $\beta > \tau\alpha + \delta$. This shows

$$B_{\delta\varepsilon(\mathbf{x})}(\mathbf{x} + \tau\alpha\varepsilon(\mathbf{x})\mathbf{e}_d) \subset B_{\beta\varepsilon(\mathbf{x})}(\mathbf{x}). \tag{5.13}$$

Furthermore, due to the choice of α in Lemma 5.1 we have $\varepsilon(\mathbf{x}_0) \le (1+L_\varepsilon\alpha)\varepsilon(\mathbf{x})$ for all $\mathbf{x}_0 \in B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})$. Hence, in view of (5.3), we infer

$$\beta\varepsilon(\mathbf{x}) \ge [\alpha + \delta(L+1)(1+L_\varepsilon\alpha)]\,\varepsilon(\mathbf{x}) \ge \alpha\varepsilon(\mathbf{x}) + (L+1)\delta\varepsilon(\mathbf{x}_0),$$

which implies

$$B_{\delta\varepsilon(\mathbf{x}_0)}(\mathbf{x}_0 + L\delta\varepsilon(\mathbf{x}_0)\mathbf{e}_d) \subset B_{\beta\varepsilon(\mathbf{x})}(\mathbf{x}) \quad \text{for all } \mathbf{x}_0 \in B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})\,. \tag{5.14}$$

The statement *(i)* follows from *(ii)*. Statements *(ii)* and *(iii)* are seen as follows: For $\mathbf{x}_0 \in B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})$ we choose $r = \alpha\varepsilon(\mathbf{x})$ and $r_0 = \delta\varepsilon(\mathbf{x}_0)$ in Lemma 5.4, (ii) and obtain

$$B_{\delta\varepsilon(\mathbf{x}_0)}(\mathbf{x}_0 + L\delta\varepsilon(\mathbf{x}_0)\mathbf{e}_d) \subset \bigcup_{\mathbf{z}\in B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})\cap\Omega} \mathfrak{C}_{\mathbf{z}}.$$

Together with (5.14), this shows (ii). Furthermore, choosing $r = \alpha\varepsilon(\mathbf{x})$ in Lemma 5.4, (i), we see that

$$\bigcup_{\mathbf{z}\in B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})\cap\Omega} \mathfrak{C}_{\mathbf{z}} \qquad \text{is star-shaped w.r.t. } B_{\delta\varepsilon(\mathbf{x})}(\mathbf{x} + \tau\alpha\varepsilon(\mathbf{x})\mathbf{e}_d).$$

Together with (5.13), this shows (iii). $\square$

Lemma 5.6. *Let ρ be a mollifier of order k_{max}, and let Ω be a special Lipschitz domain with Lipschitz constant $L_{\partial\Omega}$. Assume that $\varepsilon \in C^\infty(\Omega)$ is a Λ-admissible length scale function. Choose α, β, δ, L, τ according to Lemma 5.5. Then define the operator*

$$\mathcal{G}u(\mathbf{x}_0) := \sum_{|\mathbf{k}|\le k_{max}} \frac{(-\delta\varepsilon(\mathbf{x}_0)L\mathbf{e}_d)^{\mathbf{k}}}{\mathbf{k}!}\left(D^{\mathbf{k}}(u\star\rho_{\delta\varepsilon(\mathbf{x}_0)})\right)(\mathbf{x}_0 + \delta\varepsilon(\mathbf{x}_0)L\mathbf{e}_d)\,, \qquad \mathbf{x}_0\in\Omega. \tag{5.15}$$

(i) Suppose $(s,p)\in\mathbb{N}_0\times[1,\infty]$ satisfies $s \le k_{\max}+1$. Assume and ($s \le r \in \mathbb{R}$ and $q\in[1,\infty)$) or ($s \le r \in \mathbb{N}_0$ and $q \in [1,\infty]$). Then it holds

$$|\mathcal{G}v|_{r,q,B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})\cap\Omega} \le C_{r,q,s,p,\Lambda,L_{\partial\Omega}}\varepsilon(\mathbf{x})^{s-r+d(1/q-1/p)}\,|v|_{s,p,\veebar_{\mathbf{x}}}\,, \tag{5.16a}$$

where $C_{r,q,s,p,\Lambda,L_{\partial\Omega}}$ depends only on r, q, s, p, $(\Lambda_{\mathbf{r}'})_{|\mathbf{r}'|\le\lceil r\rceil}$, $L_{\partial\Omega}$, as well as on ρ, k_{max}, α, β, δ.

(ii) Suppose $0 \le s \in \mathbb{R}$, $r\in\mathbb{N}_0$ with $s \le r \le k_{\max}+1$, and $1 \le p \le q < \infty$. Define $\mu := d(p^{-1}-q^{-1})$. Assume that ($r = s+\mu$ and $p>1$) or ($r > s+\mu$). Then it holds that

$$|v-\mathcal{G}v|_{s,q,B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})\cap\Omega} \le C_{s,q,r,p,\Lambda,L_{\partial\Omega}}\varepsilon(\mathbf{x})^{r-s+d(1/q-1/p)}\,|v|_{r,p,\veebar_{\mathbf{x}}}\,, \tag{5.16b}$$

where $C_{s,q,r,p,\Lambda,L_{\partial\Omega}}$ depends only on s, q, r, p, $(\Lambda_{\mathbf{s}'})_{|\mathbf{s}'|\le\lceil s\rceil}$, $L_{\partial\Omega}$ as well as on ρ, k_{max}, α, β, δ.

(iii) Suppose s, $r\in\mathbb{N}_0$ with $s\le r\le k_{\max}+1$, and $1\le p<\infty$. Define $\mu := d/p$. Assume that ($r = s+\mu$ and $p=1$) or ($r > s+\mu$ and $p>1$). Then it holds that

$$|v-\mathcal{G}v|_{s,\infty,B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})\cap\Omega} \le C_{s,r,p,\Lambda,L_{\partial\Omega}}\varepsilon(\mathbf{x})^{r-s-d/p}\,|v|_{r,p,\veebar_{\mathbf{x}}}\,, \tag{5.16c}$$

where $C_{s,r,p,\Lambda,L_{\partial\Omega}}$ depends only on s, r, p, $(\Lambda_{\mathbf{s}'})_{|\mathbf{s}'|\le s}$, $L_{\partial\Omega}$ as well as on ρ, k_{max}, α, β, δ.

Proof of Lemma 5.6. Since $u \mapsto u \star \rho_{\delta\varepsilon(\mathbf{x}_0)}$ is a classical convolution operator with fixed length scale, we may write

$$D^{\mathbf{k}}(u \star \rho_{\delta\varepsilon(\mathbf{x}_0)}) = u \star (D^{\mathbf{k}}\rho_{\delta\varepsilon(\mathbf{x}_0)}) = u \star (D^{\mathbf{k}}\rho)_{\delta\varepsilon(\mathbf{x}_0)}(\delta\varepsilon(\mathbf{x}_0))^{-|\mathbf{k}|},$$

and a change of variables gives (assuming u is smooth)

$$D^{\mathbf{r}}_{\mathbf{x}_0}\mathcal{G}u(\mathbf{x}_0) = \sum_{|\mathbf{k}|\le k_{max}} \frac{(-L\mathbf{e}_d)^{\mathbf{k}}}{\mathbf{k}!} \int_{\mathbf{y}\in B_1(0)} D^{\mathbf{r}}_{\mathbf{x}_0} u\left(\mathbf{x}_0'\right) D^{\mathbf{k}}\rho(\mathbf{y}),$$

where $\mathbf{x}_0' := \mathbf{x}_0 + \varepsilon(\mathbf{x}_0)\,(\delta\mathbf{e}_d L + \delta\mathbf{y})$. The Faà di Bruno formula from Lemma 4.4 shows

$$\begin{aligned} D^{\mathbf{r}}_{\mathbf{x}_0} u\left(\mathbf{x}_0'\right) &= (D^{\mathbf{r}}u)\left(\mathbf{x}_0'\right) + \sum_{|\mathbf{t}|\le|\mathbf{r}|} \left(D^{\mathbf{t}}u\right)\left(\mathbf{x}_0'\right) E_{\mathbf{r},\mathbf{t}}\left(\delta\mathbf{e}_d L + \delta\mathbf{y}, \mathbf{x}_0\right) \\ &= (\delta\varepsilon(\mathbf{x}_0))^{-|\mathbf{r}|} D^{\mathbf{r}}_{\mathbf{y}} u\left(\mathbf{x}_0'\right) + \sum_{|\mathbf{t}|\le|\mathbf{r}|} (\delta\varepsilon(\mathbf{x}_0))^{-|\mathbf{t}|} D^{\mathbf{t}}_{\mathbf{y}} u\left(\mathbf{x}_0'\right) E_{\mathbf{r},\mathbf{t}}\left(\delta\mathbf{e}_d L + \delta\mathbf{y}, \mathbf{x}_0\right). \end{aligned}$$

We obtain with integration by parts and the product rule

$$\begin{aligned} D^{\mathbf{r}}_{\mathbf{x}_0}\mathcal{G}u(\mathbf{x}_0) = &\sum_{|\mathbf{k}|\le k_{max}} \frac{(-L\mathbf{e}_d)^{\mathbf{k}}}{\mathbf{k}!}(-1)^{|\mathbf{r}|} \int_{\mathbf{y}\in B_1(0)} u\left(\mathbf{x}_0'\right) D^{\mathbf{k}+\mathbf{r}}\rho(\mathbf{y})\,(\delta\varepsilon(\mathbf{x}_0))^{-|\mathbf{r}|} \\ &+ \sum_{|\mathbf{k}|\le k_{max}} \frac{(-L\mathbf{e}_d)^{\mathbf{k}}}{\mathbf{k}!} \sum_{|\mathbf{t}|\le|\mathbf{r}|} (-1)^{|\mathbf{t}|} \sum_{\mathbf{s}\le\mathbf{t}} \binom{\mathbf{s}}{\mathbf{t}} \\ &\quad \int_{\mathbf{y}\in B_1(0)} u(\mathbf{x}_0') D^{\mathbf{s}}_{\mathbf{y}} E_{\mathbf{r},\mathbf{t}}(\delta\mathbf{e}_d L + \delta\mathbf{y}, \mathbf{x}_0) D^{\mathbf{k}+\mathbf{t}-\mathbf{s}}\rho(\mathbf{y})\,(\delta\varepsilon(\mathbf{x}_0))^{-|\mathbf{t}|}. \end{aligned} \tag{5.17}$$

The estimates now follow as in Lemma 5.2. We apply Lemma 4.5, Lemma 5.5, (ii), and integration by parts to obtain from identity (5.17) the estimate

$$|\mathcal{G}v|_{r,\infty,B_{\alpha\varepsilon(\mathbf{x})}(\mathbf{x})\cap\Omega} \le C_{r,\Lambda,L_{\partial\Omega}}\varepsilon(\mathbf{x})^{-r}\,\|v\|_{0,1,\veebar_{\mathbf{x}}}.$$

This is the analogue of (5.9), such that the remainder of the proof follows as in Lemma 5.2. Note that we employ the embedding results of Theorem 4.2 as well as the Bramble-Hilbert Lemma 4.3 on the domain $\veebar_{\mathbf{x}}$, which we may since Lemma 5.5 shows that $\eta(\veebar_{\mathbf{x}}) \lesssim 1$ uniformly in $\mathbf{x}$. $\square$

5.4 Proof of Theorem 2.2

In the last section, we derived results on the stability and approximation properties of the two operators $\mathcal{E}$ and $\mathcal{G}$. These results, however, are strongly localized. Now we show how they can be globalized. The main ingredients are a partition of unity, the local properties of the smoothing operators, and Besicovitch's Covering theorem.

We start with a lemma that follows from Besicovitch's Covering theorem (Proposition 4.1):

Lemma 5.7. *Let $\Omega \subset \mathbb{R}^d$ be an arbitrary domain and $\varepsilon \in C^\infty(\Omega)$ be a Λ-admissible length scale function with Lipschitz constant L_ε. Let α, $\beta > 0$ satisfy (5.2). Let N_d be given by Proposition 4.1.*

Let $\omega \subset \Omega$ be arbitrary. Then there exist points $\mathbf{x}_{ij} \in \omega$, $i = 1, \ldots, N_d$, $j \in \mathbb{N}$, such that for the closed balls

$$B_{ij} := \overline{B_{\frac{\alpha}{2}\varepsilon(\mathbf{x}_{ij})}(\mathbf{x}_{ij})} \subset \widetilde{B}_{ij} := \overline{B_{\alpha\varepsilon(\mathbf{x}_{ij})}(\mathbf{x}_{ij})} \subset \widehat{B}_{ij} := \overline{B_{\beta\varepsilon(\mathbf{x}_{ij})}(\mathbf{x}_{ij})}, \tag{5.18}$$

the following is true:

(i) $\omega \subset \cup_{i=1}^{N_d} \cup_{j\in\mathbb{N}} B_{ij}$.

(ii) *For each* $i \in \{1,\dots,N_d\}$, *the balls* $\{B_{ij} \mid j \in \mathbb{N}\}$ *are pairwise disjoint.*

(iii) *For each* $i \in \{1,\dots,N_d\}$, *the balls* $\widehat{B}_{ij}$, $j \in \mathbb{N}$ *satisfy an overlap property: There exists* $C_{overlap} > 0$, *which depends solely on* d, α, β, L_ε, *such that*

$$\operatorname{card}\{j' \mid \widehat{B}_{ij} \cap \widehat{B}_{ij'} \neq \emptyset\} \leq C_{overlap} \qquad \forall j \in \mathbb{N}.$$

(iv) *Let* $q \in [1,\infty)$ *and* $\sigma \in (0,1)$. *Then for a constant* C *that depends solely on* σ, q, d, *and* α:

$$\sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}} \int_{\mathbf{x}\in B_{ij}} \int_{\mathbf{y}\in\omega\setminus\widetilde{B}_{ij}} \frac{|v(\mathbf{x})|^q}{|\mathbf{x}-\mathbf{y}|^{\sigma q+d}}\, d\mathbf{y}\, d\mathbf{x} \leq C \sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}} \varepsilon(\mathbf{x}_{ij})^{-\sigma q}\|v\|^q_{0,q,\widetilde{B}_{ij}} \tag{5.19}$$

$$\sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}} \int_{\mathbf{x}\in B_{ij}} \int_{\mathbf{y}\in\omega\setminus\widetilde{B}_{ij}} \frac{|v(\mathbf{y})|^q}{|\mathbf{x}-\mathbf{y}|^{\sigma q+d}}\, d\mathbf{y}\, d\mathbf{x} \leq C \sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}} \varepsilon(\mathbf{x}_{ij})^{-\sigma q}\|v\|^q_{0,q,\widetilde{B}_{ij}}. \tag{5.20}$$

Proof. Proof of (i), (ii): The set $\mathcal{F} := \left\{\overline{B_{\alpha\varepsilon(\mathbf{x})/2}(\mathbf{x})} \mid \mathbf{x} \in \omega\right\}$ is a closed cover of ω. According to Besicovitch's Covering Theorem (Proposition 4.1), there are countable subsets $\mathcal{G}_1,\dots,\mathcal{G}_{N_d}$ of $\mathcal{F}$, the elements of every subset being pairwise disjoint, and

$$\omega \subset \bigcup_{i=1}^{N_d}\bigcup_{j\in\mathbb{N}} \overline{B_{\alpha\varepsilon(\mathbf{x}_{ij})/2}(\mathbf{x}_{ij})},$$

where $B_{ij} = \overline{B_{\alpha\varepsilon(\mathbf{x}_{ij})/2}(\mathbf{x}_{ij})} \in \mathcal{G}_i$ for $j \in \mathbb{N}$.

Proof of (iii): Suppose $\widehat{B}_{ij} \cap \widehat{B}_{ij'} \neq \emptyset$. Then, the Lipschitz continuity of ε gives

$$\left|\varepsilon(\mathbf{x}_{ij}) - \varepsilon(\mathbf{x}_{ij'})\right| \leq L_\varepsilon \left\|\mathbf{x}_{ij} - \mathbf{x}_{ij'}\right\| \leq L_\varepsilon\beta\left(\varepsilon(\mathbf{x}_{ij}) + \varepsilon(\mathbf{x}_{ij'})\right).$$

Since $\beta L_\varepsilon < 1$ due to our assumption (5.2), we conclude

$$\widehat{B}_{ij} \cap \widehat{B}_{ij'} \neq \emptyset \implies \varepsilon(\mathbf{x}_{ij}) \leq \varepsilon(\mathbf{x}_{ij'})\frac{1+L_\varepsilon\beta}{1-L_\varepsilon\beta}. \tag{5.21}$$

For $\mathbf{z} \in B_{ij'}$ with $\widehat{B}_{ij} \cap \widehat{B}_{ij'} \neq \emptyset$ we use (5.21) to estimate

$$\begin{aligned} \|\mathbf{z} - \mathbf{x}_{ij}\| &\leq \left\|\mathbf{z} - \mathbf{x}_{ij'}\right\| + \left\|\mathbf{x}_{ij'} - \mathbf{x}_{ij}\right\| \leq \alpha\varepsilon(\mathbf{x}_{ij'})/2 + \beta\left(\varepsilon(\mathbf{x}_{ij}) + \varepsilon(\mathbf{x}_{ij'})\right) \\ &\leq \varepsilon(\mathbf{x}_{ij})\left(\frac{\alpha}{2}\frac{1+L_\varepsilon\beta}{1-L_\varepsilon\beta} + \beta + \beta\frac{1+L_\varepsilon\beta}{1-L_\varepsilon\beta}\right) =: \varepsilon(\mathbf{x}_{ij})C_{\text{big}}. \end{aligned} \tag{5.22}$$

Note that all $B_{ij'}$ with $j' \in \mathbb{N}$ are disjoint for a fixed i. Therefore, with C_d denoting the volume of the unit sphere in $\mathbb{R}^d$

$$\sum_{\substack{j'\in\mathbb{N} \\ \widehat{B}_{ij}\cap\widehat{B}_{ij'}\neq\emptyset}} \left|B_{ij'}\right| \leq \left|B_{\varepsilon(\mathbf{x}_{ij})C_{\text{big}}}(\mathbf{x}_{ij})\right| = C_d\left(\varepsilon(\mathbf{x}_{ij})C_{\text{big}}\right)^d, \tag{5.23}$$

Next, we bound the number of balls that intersect a given one. We use (5.21) and (5.23) to see

$$\begin{aligned}\operatorname{card}\left\{j' \mid \widehat{B}_{ij}\cap\widehat{B}_{ij'}\neq\emptyset\right\} &= \sum_{\substack{j'\in\mathbb{N}\\ \widehat{B}_{ij}\cap\widehat{B}_{ij'}\neq\emptyset}} \frac{\left|B_{ij'}\right|}{\left|B_{ij'}\right|} = \frac{2^d}{\alpha^d C_d}\sum_{\substack{j'\in\mathbb{N}\\ \widehat{B}_{ij}\cap\widehat{B}_{ij'}\neq\emptyset}}\frac{\left|B_{ij'}\right|}{\varepsilon(\mathbf{x}_{ij'})^d}\\ &\leq \frac{2^d}{\alpha^d C_d}\left(\frac{1+L_\varepsilon\beta}{1-L_\varepsilon\beta}\right)^d \varepsilon(\mathbf{x}_{ij})^{-d}\sum_{\substack{j'\in\mathbb{N}\\ \widehat{B}_{ij}\cap\widehat{B}_{ij'}\neq\emptyset}}\left|B_{ij'}\right| \qquad (5.24)\\ &\leq \left(\frac{2C_{\text{big}}}{\alpha}\frac{1+L_\varepsilon\beta}{1-L_\varepsilon\beta}\right)^d =: C_{\text{overlap}}.\end{aligned}$$

Proof of (iv): We start with the simpler estimate, (5.19). It follows directly from the observation that for $\mathbf{x}\in B_{ij}$ we have $B_{\frac{\alpha}{2}\varepsilon(\mathbf{x}_{ij})}(\mathbf{x})\subset B_{\alpha\varepsilon(\mathbf{x}_{ij})}(\mathbf{x}_{ij})=\widetilde{B}_{ij}$ so that

$$\int_{\mathbf{y}\in\omega\setminus\widetilde{B}_{ij}}\frac{1}{|\mathbf{x}-\mathbf{y}|^{\sigma q+d}}\,d\mathbf{y}\leq\int_{\mathbf{y}\in\mathbb{R}^d\setminus B_{\frac{\alpha}{2}\varepsilon(\mathbf{x}_{ij})}(0)}\frac{1}{|\mathbf{y}|^{\sigma q+d}}\,d\mathbf{y}=C_{\alpha,d,\sigma,q}\varepsilon(\mathbf{x}_{ij})^{-\sigma q},$$

where the constant $C_{\alpha,d,\sigma,q}$ depends on the quantities indicated.

We turn to the estimate (5.20). We essentially repeat the arguments of [Fae02, Lemma 3.1]. Using $\omega\subset\cup_{ij}B_{ij}$ and the notation χ_A for the characteristic function of a set A, we have to estimate

$$\begin{aligned}&\sum_{i=1}^{N_d}\sum_{i'=1}^{N_d}\sum_{j\in\mathbb{N}}\sum_{j'\in\mathbb{N}}\int_{\mathbf{x}\in B_{ij}}\int_{\mathbf{y}\in B_{i'j'}\setminus\widetilde{B}_{ij}}\frac{|v(\mathbf{y})|^q}{|\mathbf{x}-\mathbf{y}|^{\sigma q+d}}\,d\mathbf{y}\,d\mathbf{x}\\ &=\sum_{i'=1}^{N_d}\sum_{j'\in\mathbb{N}}\int_{\mathbf{y}\in B_{i'j'}}|v(\mathbf{y})|^q\int_{\mathbf{x}\in\mathbb{R}^d}\sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}}\chi_{B_{ij}}(\mathbf{x})\chi_{B_{i'j'}\setminus\widetilde{B}_{ij}}(\mathbf{y})\frac{1}{|\mathbf{x}-\mathbf{y}|^{\sigma q+d}}\,d\mathbf{x}\,d\mathbf{y}.\end{aligned}\qquad(5.25)$$

To that end, we analyze $\sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}}\chi_{B_{ij}}(\mathbf{x})\chi_{B_{i'j'}\setminus\widetilde{B}_{ij}}(\mathbf{y})$ in more detail. Pick $\lambda>0$ such that

$$1-\frac{\alpha}{2}L_\varepsilon-\lambda L_\varepsilon<1\qquad\text{and}\qquad\frac{\alpha}{2}\frac{1-\frac{\alpha}{2}L_\varepsilon-\lambda L_\varepsilon}{1+\alpha/2}>\lambda>0;\qquad(5.26)$$

this is possible due to (5.2). We claim that the following is true:

$$\mathbf{y}\in B_{i'j'}\qquad\Longrightarrow\qquad C(\mathbf{x},\mathbf{y}):=\sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}}\chi_{B_{ij}}(\mathbf{x})\chi_{B_{i'j'}\setminus\widetilde{B}_{ij}}(\mathbf{y})\leq\begin{cases}N_d & \text{if } |\mathbf{x}-\mathbf{y}|\geq\lambda\varepsilon(\mathbf{x}_{i'j'})\\ 0 & \text{if } |\mathbf{x}-\mathbf{y}|<\lambda\varepsilon(\mathbf{x}_{i'j'}).\end{cases}\qquad(5.27)$$

The desired final estimate (5.20) then follows from inserting (5.27) into (5.25) and the introduction of polar coordinates to evaluate the integral in $\mathbf{x}$. In order to see (5.27), fix $(i',j')\in\{1,\ldots,N_d\}\times\mathbb{N}$. Since the sets $\{B_{ij}\,|\,j\in\mathbb{N}\}$ are pairwise disjoint for each i, it is clear that $C(\mathbf{x},\mathbf{y})\leq N_d$. We have therefore to ascertain that $|\mathbf{x}-\mathbf{y}|<\lambda\varepsilon(\mathbf{x}_{i'j'})$ implies $C(\mathbf{x},\mathbf{y})=0$. Let $|\mathbf{x}-\mathbf{y}|<\lambda\varepsilon(\mathbf{x}_{i'j'})$. We proceed by contradiction. Suppose $C(\mathbf{x},\mathbf{y})\neq 0$. Then, we must have $\mathbf{x}\in B_{ij}$, $\mathbf{y}\in B_{i'j'}$ and $\mathbf{y}\notin\widetilde{B}_{ij}$. Hence, we conclude $|\mathbf{y}-\mathbf{x}|\geq\frac{\alpha}{2}\varepsilon(\mathbf{x}_{ij})$. Thus,

$$\frac{\alpha}{2}\varepsilon(\mathbf{x}_{ij})\leq|\mathbf{x}-\mathbf{y}|<\lambda\varepsilon(\mathbf{x}_{i'j'}).\qquad(5.28)$$

Next, we use the Lipschitz continuity of ε:

$$\begin{aligned}\varepsilon(\mathbf{x}_{i'j'}) \le \varepsilon(\mathbf{x}_{ij}) + L_\varepsilon|\mathbf{x}_{ij} - \mathbf{x}_{i'j'}| &\le \varepsilon(\mathbf{x}_{ij}) + L_\varepsilon|\mathbf{x}_{ij} - \mathbf{x}| + L_\varepsilon|\mathbf{x} - \mathbf{y}| + L_\varepsilon|\mathbf{y} - \mathbf{x}_{i'j'}| \\ &\le \varepsilon(\mathbf{x}_{ij}) + \frac{\alpha}{2}L_\varepsilon\varepsilon(\mathbf{x}_{ij}) + \lambda L_\varepsilon\varepsilon(\mathbf{x}_{i'j'}) + \frac{\alpha}{2}L_\varepsilon\varepsilon(\mathbf{x}_{i'j'}).\end{aligned}$$

Rearranging the terms yields

$$\left(1 - \frac{\alpha}{2}L_\varepsilon - \lambda L_\varepsilon\right)\varepsilon(\mathbf{x}_{i'j'}) \le \left(1 + \frac{\alpha}{2}\right)\varepsilon(\mathbf{x}_{ij}).$$

Inserting this in (5.28) yields

$$\frac{\alpha}{2}\frac{1 - \frac{\alpha}{2}L_\varepsilon - \lambda L_\varepsilon}{1 + \alpha/2}\varepsilon(\mathbf{x}_{i'j'}) \le |\mathbf{x} - \mathbf{y}| < \lambda\varepsilon(\mathbf{x}_{i'j'}),$$

which contradicts (5.26). $\square$

Proof of Theorem 2.2. As it is standard in the treatment of Lipschitz domains, we proceed with a localization procedure. Let $\{U^j\}_{j=0}^m$ be an open cover of Ω such that $\{U^j\}_{j=1}^m$ is an open cover of $\partial\Omega$. Let $\{\eta^j\}_{j=0}^m$ form a partition of unity subordinate to this open cover (see [Ada75, Thm. 3.14]), i.e., $\sum_{j=0}^m \eta^j = 1$ on $\overline{\Omega}$, and $0 \le \eta^j \in C_0^\infty(U^j)$. We furthermore assume that $\operatorname{supp}(\eta^0) \subset \Omega$. Next, we choose α, β, δ, τ, and L according to Lemma 5.5 where we set $K = \operatorname{supp}(\eta^0)$. Since $\varepsilon \le 1$, this implies in particular that

$$\mathbf{x} \in \Omega \qquad \Longrightarrow \qquad \text{either} \quad \overline{B_{2\beta\varepsilon(\mathbf{x})}(\mathbf{x})} \subset \Omega \quad \text{or} \quad \overline{B_{2\beta\varepsilon(\mathbf{x})}(\mathbf{x})} \cap \operatorname{supp}(\eta^0) = \emptyset. \tag{5.29}$$

With these choices, we may apply Lemmas 5.3 and 5.6.

For $u \in L^1_{loc}(\Omega)$ we write $u = u\eta^0 + \sum_{j=1}^m u\eta^j$, and we extend $u\eta^j$ to $\mathbb{R}^d$ by zero. The operator $\mathcal{E}$ of Lemma 5.3 will be applied to $u\eta^0$, while the operator $\mathcal{G}$ of Lemma 5.6 will be applied to $u\eta^j$ for $j \ge 1$. We can assume that $\partial\Omega \cap U^j$ can (after translation) be extended to $\{\mathbf{x} \in \mathbb{R}^d \mid \mathbf{x}_d = g^j(\mathbf{x}_1, \dots, \mathbf{x}_{d-1})\}$ and g^j has the same Lipschitz constant as $\partial\Omega$, i.e., we consider $\partial\Omega \cap U^j$ as a special Lipschitz domain. Thus, we see that the operator

$$\mathcal{I}u := \mathcal{E}(u\eta^0) + \sum_{j=1}^m \mathcal{G} \circ Q^j(u\eta^j),$$

Q^j being an appropriate coordinate transformation, is well defined.

Let $\omega \subset \Omega$ and let the closed balls B_{ij}, $\widetilde{B}_{ij}$, $\widehat{B}_{ij}$ be given by (5.18) of Lemma 5.7. We first show the stability (2.2) of $\mathcal{I}_\varepsilon$ in the case that $r \in \mathbb{N}_0$ and $s \in \mathbb{N}_0$. Applying the triangle inequality to the definition gives

$$|\mathcal{I}_\varepsilon u|_{r,q,\omega} \le \left|\mathcal{E}(u\eta^0)\right|_{r,q,\omega} + \sum_{\ell=1}^m \left|\mathcal{G}(u\eta^\ell)\right|_{r,q,\omega}. \tag{5.30}$$

We start by focussing on the contribution $\mathcal{E}(u\eta^0)$ in (5.30). First, we remark for $\mathcal{E}(u\eta^0)$ that (5.29) implies

$$\mathbf{x} \in \widetilde{K} := \operatorname{supp}(\mathcal{E}(u\eta^0)) \qquad \Longrightarrow \qquad B_{\beta\varepsilon(\mathbf{x})}(\mathbf{x}) \subset \Omega. \tag{5.31}$$

We note the covering property stated in Lemma 5.7, (i). With the estimate (5.11a) of Lemma 5.3 we get

$$\begin{aligned}\left|\mathcal{E}(u\eta^0)\right|^q_{r,q,\omega} = \left|\mathcal{E}(u\eta^0)\right|^q_{r,q,\omega\cap\widetilde{K}} &\leq \sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}} \left|\mathcal{E}(u\eta^0)\right|^q_{r,q,B_{ij}\cap\widetilde{K}} \\ &\overset{(5.31),(5.11a)}{\lesssim} \sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}} \varepsilon(\mathbf{x}_{ij})^{q(s-r)+d(1-q/p)} \left|u\eta^0\right|^q_{s,p,\widehat{B}_{ij}}.\end{aligned}$$

Noting that $\varepsilon(\mathbf{x}_{ij}) \simeq \varepsilon(\mathbf{z})$ for all $\mathbf{z} \in \widehat{B}_{ij} \cap \Omega$ we obtain

$$\varepsilon(\mathbf{x}_{ij})^{q(s-r)+d(1-q/p)} \left|u\eta^0\right|^q_{s,p,\widehat{B}_{ij}} \lesssim \varepsilon(\mathbf{x}_{ij})^{q(s-r)+d(1-q/p)} \left\|u\right\|^q_{s,p,\Omega\cap\widehat{B}_{ij}} \lesssim \sum_{|\mathbf{s}|\leq s} \left\|\varepsilon^{s-r+d(1/q-1/p)} D^{\mathbf{s}} u\right\|^q_{0,p,\Omega\cap\widehat{B}_{ij}}.$$

We combine the last two estimates and (5.24) to arrive at

$$\left|\mathcal{E}(u\eta^0)\right|^q_{r,q,\omega} \lesssim \sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}}\sum_{|\mathbf{s}|\leq s} \left\|\varepsilon^{s-r+d(1/q-1/p)} D^{\mathbf{s}} u\right\|^q_{0,p,\Omega\cap\widehat{B}_{ij}} \lesssim N_d C_{\text{overlap}} \sum_{|\mathbf{s}|\leq s} \left\|\varepsilon^{s-r+d(1/q-1/p)} D^{\mathbf{s}} u\right\|^q_{0,p,\omega_\varepsilon}.$$

This concludes the proof of the stability bound (2.2) for $r,\ s \in \mathbb{N}_0$. We now turn to the case $r \in \mathbb{R}\setminus\mathbb{N}_0$ and $q \in [1,\infty)$ together with $s \in \mathbb{N}_0$. For $|\mathbf{r}| = \lfloor r \rfloor$ and $\sigma = r - \lfloor r \rfloor \in (0,1)$, we recall the definition of the sets B_{ij} and $\widetilde{B}_{ij}$ and write

$$\begin{aligned}\left|D^{\mathbf{r}}\mathcal{E}(u\eta^0)\right|^q_{\sigma,q,\omega} = \left|D^{\mathbf{r}}\mathcal{E}(u\eta^0)\right|^q_{\sigma,q,\omega\cap\widetilde{K}} &\leq \sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}} \int_{B_{ij}\cap\widetilde{K}}\int_{\omega\cap\widetilde{K}} \frac{\left|D^{\mathbf{r}}\mathcal{E}(u\eta^0)(\mathbf{x}) - D^{\mathbf{r}}\mathcal{E}(u\eta^0)(\mathbf{y})\right|^q}{|\mathbf{x}-\mathbf{y}|^{\sigma q+d}}\, d\mathbf{y}\, d\mathbf{x} \\ &\leq \sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}} \left|D^{\mathbf{r}}\mathcal{E}(u\eta^0)\right|_{\sigma,q,\widetilde{B}_{ij}\cap\widetilde{K}} + \int_{B_{ij}\cap\widetilde{K}}\int_{(\omega\cap\widetilde{K})\setminus\widetilde{B}_{ij}} \frac{\left|D^{\mathbf{r}}\mathcal{E}(u\eta^0)(\mathbf{x}) - D^{\mathbf{r}}\mathcal{E}(u\eta^0)(\mathbf{y})\right|^q}{|\mathbf{x}-\mathbf{y}|^{\sigma q+d}}\, d\mathbf{y}\, d\mathbf{x}.\end{aligned}$$

Lemma 5.7, (iv) leads to

$$\left|D^{\mathbf{r}}\mathcal{E}(u\eta^0)\right|^q_{\sigma,q,\omega} \lesssim \sum_{i=1}^{N_d}\sum_{j\in\mathbb{N}} \left|D^{\mathbf{r}}\mathcal{E}(u\eta^0)\right|^q_{\sigma,q,\widetilde{B}_{ij}\cap\widetilde{K}} + \varepsilon(\mathbf{x}_{ij})^{-\sigma q} \left\|D^{\mathbf{r}}\mathcal{E}(u\eta^0)\right\|^q_{0,q,\widetilde{B}_{ij}}.$$

Recalling that (5.31) ensures that nontrivial terms in this sum correspond to pairs (i,j) with $\widehat{B}_{ij} \subset \Omega$, we may use once more estimate (5.11a) of Lemma 5.3. The finite overlap property of the sets $\widehat{B}_{ij}$ then shows the stability estimate (2.2) for $r \in \mathbb{R}\setminus\mathbb{N}_0$ and $q \in [1,\infty)$. The same arguments as above apply for the parts with $\mathcal{G}$ in (5.30) if we use estimate (5.16a) of Lemma 5.6.

Finally, the estimates (2.3) and (2.4) can be shown exactly as (2.2) if we use estimates (5.11b) or (5.11c) of Lemma 5.3 and estimates (5.16b) or (5.16c) of Lemma 5.6. $\square$

A Sobolev embedding theorems (Proof of Theorem 4.2)

The purpose of the appendix is the proof of the core of Theorem 4.2, that is, we show for domains that are star shaped with respect to a ball, that the constants in some Sobolev embedding theorems depend solely on the "chunkiness parameter" and the diameter of the domain. This can be seen by tracking the domain dependence in the proof given in [Mur68, Mur67]. For the reader's convenience, we present below the essential steps of this proof with an emphasis on the dependence on the geometry.

Theorem A.1. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with* $\operatorname{diam}(\Omega) = 1$. *Assume Ω is star-shaped with respect to the ball $B_\rho := B_\rho(0)$ of radius $\rho > 0$. Let $0 \leq s \leq r < \infty$ and $1 \leq p \leq q < \infty$. Set*

$$\mu := d\left(\frac{1}{p} - \frac{1}{q}\right). \tag{A.1}$$

Assume that one of the following two possibilities takes place:

(a) $r = s + \mu$ and $p > 1$;

(b) $r > s + \mu$ and $p \geq 1$.

Then there exists $C = C(s, q, r, p, \rho, d)$ depending only on the constants indicated such that

$$|u|_{W^{s,q}(\Omega)} \leq C(s, q, r, p, \rho, d)\|u\|_{W^{r,p}(\Omega)}.$$

Proof. The case $s = 0$ is handled in Lemmas A.8 and A.12, while the case $s \in (0, 1)$ can be found in Lemmas A.11 and A.13. For $|\mathbf{t}| = \lfloor s \rfloor$, these two cases imply for any derivative $D^{\mathbf{t}}u$ the estimate $|D^{\mathbf{t}}u|_{W^{s-\lfloor s \rfloor,q}(\Omega)} \leq C\|u\|_{W^{r,p}(\Omega)}$. □

Remark A.2. 1. *The case $p = 1$ in conjunction with $r = s+\mu$ is excluded in Theorem A.1. This is due to our method of proof. The Sobolev embedding theorem in the form given in [AF03, Thm. 7.38] suggests that Theorem A.1 also holds in this case.*

2. *The star-shapedness in Theorem A.1 is not the essential ingredient for our control of the constants of the embedding theorems. It suffices that Ω satisfies the interior cone condition (A.2) with explicit control of the Lipschitz constant of the function ψ and the parameter T. That is, the impact of the geometry on the final estimates is captured by the Lipschitz constant* $\operatorname{Lip}(\psi)$ *of ψ, the parameter T, and d.*

Lemma A.3. *Let Ω be as in Theorem A.1. Then there exists a Lipschitz continuous function*[1] *$\psi : \mathbb{R}^d \to \mathbb{R}^d$ and constants C, $\widetilde{C} > 0$, and $T \in (0, 1]$, which depend solely on the chunkiness parameter ρ, such that the following is true:*

(i) For every $\mathbf{x} \in \Omega$ it holds

$$C_{\mathbf{x},t} := \{\mathbf{x} + t(\psi(\mathbf{x}) + \mathbf{z}) \,|\, \mathbf{z} \in B_1, \quad 0 < t < T\} \subset \Omega. \tag{A.2}$$

(ii) $\|\psi\|_{L^\infty(\mathbb{R}^d)} + \|\nabla\psi\|_{L^\infty(\mathbb{R}^d)} \leq C$.

(iii) For every $t \in [0, T]$, the map $\Psi_t : \mathbb{R}^d \to \mathbb{R}^d$ given by with $\Psi_t(\mathbf{x}) := \mathbf{x} + t\psi(\mathbf{x})$ is invertible and bilipschitz, i.e., Ψ and its inverse $\Psi_t^{-1} : \mathbb{R}^d \to \mathbb{R}^d$ are Lipschitz continuous. Furthermore, $\|\nabla\Psi_t\|_{L^\infty(\mathbb{R}^d)} \leq 1 + t\widetilde{C}$ as well as $\|\nabla\Psi_t^{-1}\|_{L^\infty(\mathbb{R}^d)} \leq 1 + t\widetilde{C}$. Additionally, $\Psi_t(\Omega) + B_1 \subset \Omega$.

[1] In fact, the mapping is smooth.

Proof. Recall that Ω is star-shaped with respect to the ball B_ρ, whose center is the origin. Let χ be a smooth cut-off function supported by $B_{\rho/2}$ with $\chi \equiv 1$ on $B_{\rho/4}$ and $0 \leq \chi \leq 1$. Let $\psi(\mathbf{x}) := L\frac{\mathbf{x}}{|\mathbf{x}|}(\chi(\mathbf{x}) - 1)$, where the parameter $L > 0$ will be chosen sufficiently large below. Then $\|\psi\|_{L^\infty(\mathbb{R}^d)} \leq L$ (if the space $\mathbb{R}^d$ is endowed with the Euclidean norm) and $\|\nabla\psi\|_{L^\infty(\mathbb{R}^d)} \leq CL$ for a constant $C > 0$ that depends solely on the choice of χ. For $\mathbf{x} \in \Omega \setminus B_\rho$, geometric considerations and diam $= 1$ show that $\{\mathbf{x} + t(\psi(\mathbf{x}) + \mathbf{z}) \,|\, \mathbf{z} \in B_1\}$ is contained in the infinite cone with apex $\mathbf{x}$ that contains the ball B_ρ provided that L is sufficiently large, specifically, $L \sim 1/\rho$. Hence, by taking T sufficiently small (essentially, $T \sim 1/L$) we can ensure the condition (A.2). This shows (i) and (ii). The assertion (iii) follows from suitably reducing T. We note that for t with $t\|\nabla\psi\|_{L^\infty(\mathbb{R}^d)} < 1$, the map $\Psi(\mathbf{x}) = \mathbf{x} + t\psi(\mathbf{x})$ is invertible as a map $\mathbb{R}^d \to \mathbb{R}^d$ by the Banach Fixed Point Theorem. To see that Ψ_t^{-1} is Lipschitz continuous, we let $\mathbf{y}$, $\mathbf{y}' \in \mathbb{R}^d$ and let $\mathbf{x}$, $\mathbf{x}'$ satisfy $\Psi_t(\mathbf{x}) = \mathbf{y}$, $\Psi_t(\mathbf{x}') = \mathbf{y}'$. Then $\mathbf{x} - \mathbf{x}' = \mathbf{y} - \mathbf{y}' - t(\psi(\mathbf{x}) - \psi(\mathbf{x}'))$ so that $|\mathbf{x} - \mathbf{x}'| \leq |\mathbf{y} - \mathbf{y}'| + t\|\nabla\psi\|_{L^\infty(\mathbb{R}^d)}|\mathbf{x} - \mathbf{x}'|$. The assumption $t\|\nabla\psi\|_{L^\infty(\mathbb{R}^d)} < 1$ then implies the result. The argument also shows that $\nabla\Psi_t$ has the stated t-dependence. □

The method of proof of Theorem A.1 relies on appropriate smoothing. The following lemma provides two different representations of a function u in terms of an averaged version $M(u)$. These two presentations will be needed to treat both the case of fractional and integer order Sobolev regularity. Let $\omega \in C_0^\infty(\mathbb{R}^d)$ with $\operatorname{supp}(\omega) \subset B_1$ and $\int_{B_1} \omega(\mathbf{z})\, d\mathbf{z} = 1$. Then we have the following representation formula:

Lemma A.4. *Let ψ and T be as in Lemma A.3. For $u \in C^\infty(\Omega)$ and $t \in [0, T]$ define*

$$M(u)(t, \mathbf{x}) := \int_{\mathbf{z}\in B_1} \omega(\mathbf{z})u(\mathbf{x} + t(\mathbf{z} + \psi(\mathbf{x})))\, d\mathbf{z}. \tag{A.3}$$

Then we have the two representation formulas

$$u(\mathbf{x}) - M(u)(t, \mathbf{x}) = M_R(u)(t, \mathbf{x}) = M_S(u)(t, \mathbf{x})$$

for any $t \in [0, T]$, where

$$M_R(u)(t, \mathbf{x}) := -\int_{\tau=0}^{t} \int_{\mathbf{z}\in B_1} \underbrace{\omega(\mathbf{z})(\mathbf{z} + \psi(\mathbf{x}))}_{=:\omega_1(\mathbf{x},\mathbf{z})} \cdot \nabla u(\mathbf{x} + \tau(\mathbf{z} + \psi(\mathbf{x})))\, d\mathbf{z}\, d\tau, \tag{A.4}$$

$$\begin{aligned} M_S(u)(t, \mathbf{x}) := -\int_{\tau=0}^{t} \int_{\mathbf{z}\in B_1} \int_{\mathbf{z}'\in B_1} \omega(\mathbf{z}') \frac{\omega_2(\mathbf{x}, \mathbf{z})}{\tau} \\ \left[u(\mathbf{x} + \tau(\mathbf{z} + \psi(\mathbf{x}))) - u(\mathbf{x} + \tau(\mathbf{z}' + \psi(\mathbf{x})))\right] d\mathbf{z}'\, d\mathbf{z}\, d\tau \end{aligned} \tag{A.5}$$

where the function ω_2 is given by

$$\omega_2(\mathbf{x}, \mathbf{z}) := -\left(d\omega(\mathbf{z}) + \nabla\omega(\mathbf{z}) \cdot (\mathbf{z} + \psi(\mathbf{x}))\right). \tag{A.6}$$

Proof. Since $M(u)(0, \mathbf{x}) = u(\mathbf{x})$, we have

$$u(\mathbf{x}) = M(u)(t, \mathbf{x}) - \int_{\tau=0}^{t} \partial_\tau M(u)(\tau, \mathbf{x})\, d\tau. \tag{A.7}$$

Interchanging differentiation and integration yields $\partial_\tau M(u)(\tau,\mathbf{x}) = \int_{\mathbf{z}\in B_1} \omega(\mathbf{z})(\mathbf{z}+\psi(\mathbf{x}))\cdot\nabla u(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})))\,d\mathbf{z}$, which is formula (A.4). In order to see (A.5), we start again with (A.7). A change of variables shows (u is implicitly extended by zero outside Ω)

$$M(u)(\tau,\mathbf{x}) = \int_{\mathbf{y}\in\mathbb{R}^d} \underbrace{\tau^{-d}\omega((\mathbf{y}-\mathbf{x})/\tau - \psi(\mathbf{x}))}_{=:\widetilde{\omega}(\mathbf{y},\mathbf{x},\tau)} u(\mathbf{y})\,d\mathbf{y}.$$

Hence, $\partial_\tau M(u)(\tau,\mathbf{x}) = \int_{\mathbf{y}\in\mathbb{R}^d} \partial_\tau\widetilde{\omega}(\mathbf{y},\mathbf{x},\tau)u(\mathbf{y})\,d\mathbf{y}$, where $\partial_\tau\widetilde{\omega}$ is given explicitly by

$$\begin{aligned}\partial_\tau\widetilde{\omega}(\mathbf{y},\mathbf{x},\tau) &= -\frac{1}{\tau^{d+1}}\left\{d\omega((\mathbf{y}-\mathbf{z})/\tau-\psi(\mathbf{x})) + \nabla\omega((\mathbf{y}-\mathbf{x})/\tau-\psi(\mathbf{x}))\cdot(\mathbf{y}-\mathbf{x})/\tau\right\}\\ &= \tau^{-(d+1)}\omega_2(\mathbf{y}-\mathbf{z})/\tau-\psi(\mathbf{x})).\end{aligned}$$

As $M(1)\equiv 1$, we get $\partial_\tau M(1)\equiv 0$, i.e., $\displaystyle\int_{\mathbf{y}\in\mathbb{R}^d} \partial_\tau\widetilde{\omega}(\mathbf{y},\mathbf{x},\tau)\,d\mathbf{y}\equiv 0$. Therefore, for arbitrary $\mathbf{y}'$

$$\partial_\tau M(u)(\tau,\mathbf{x}) = \int_{\mathbf{y}\in\mathbb{R}^d} \partial_\tau\widetilde{\omega}(\mathbf{x},\mathbf{y},\tau)(u(\mathbf{y})-u(\mathbf{y}'))\,d\mathbf{y}.$$

Multiplication with $\omega((\mathbf{y}'-\mathbf{x})/\tau-\psi(\mathbf{x}))$, integration over $\mathbf{y}'$, and a change of variables yield

$$\partial_\tau M(u)(\tau,\mathbf{x}) = \int_{\mathbf{z}'\in B_1}\int_{\mathbf{z}\in B_1} \omega(\mathbf{z}')\frac{\omega_2(\mathbf{x},\mathbf{z})}{\tau}(u(\mathbf{x}+\tau(\mathbf{z}+\psi(\mathbf{x}))) - u(\mathbf{x}+\tau(\mathbf{z}'+\psi(\mathbf{x})))\,d\mathbf{z}\,d\mathbf{z}'.$$

Inserting this in (A.7) then gives the representation (A.5). □

Remark A.5. *Higher order representation formulas are possible, see, e.g., [Mur68].*

A.1 The case of integer s in Theorem A.1

The limiting case $p=1$ is special. In the results below, it will appear often in conjunction with a parameter $\sigma\geq 0$. In the interest of brevity, we formulate a condition that we will require repeatedly in the sequel:

$$(\tilde{\sigma}=0 \quad\text{and}\quad \tilde{p}>1) \qquad\text{or}\qquad (\tilde{\sigma}>0 \quad\text{and}\quad \tilde{p}\geq 1)\,. \tag{A.8}$$

For a ball $B_t\subset\mathbb{R}^d$ of radius $t>0$, we write $|B_t|\sim t^d$ for its (Lebesgue) measure.

Lemma A.6. *Let $1\leq\tilde{p}\leq\tilde{q}<\infty$ and $u\in L^{\tilde{p}}(\Omega)$. Let Ω, T, and ψ be as in Lemma A.3. Set $\tilde{\mu}=d(\tilde{p}^{-1}-\tilde{q}^{-1})$. Define, for $t\in(0,T)$, the function*

$$U(\mathbf{x},t) := \int_{\mathbf{z}\in B_1} |u(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})))|\,d\mathbf{z}. \tag{A.9}$$

Then the following two estimates hold:

(i) There exists $C_1=C_1(\tilde{p},\tilde{q},\mathrm{Lip}(\psi),d)>0$ depending only on the quantities indicated such that

$$\|U(t,\cdot)\|_{L^{\tilde{q}}(\Omega)} \;\leq\; C_1 t^{-\tilde{\mu}}\|u\|_{L^{\tilde{p}}(\Omega)}. \tag{A.10}$$

(ii) If the pair $(\tilde\sigma,\tilde p)$ satisfies (A.8), then there exists $C_2 = C_2(\tilde p,\tilde q,\operatorname{Lip}(\psi),\tilde\sigma,d) > 0$ depending only on the quantities indicated such that

$$\left\| \int_{t=0}^{T} t^{-1+\tilde\mu+\tilde\sigma} U(t,\cdot)\,dt \right\|_{L^{\tilde q}(\Omega)} \quad\leq\quad C_2 T^{\tilde\sigma}\left[T^{\tilde\mu-d}\|u\|_{L^1(\Omega)} + \|u\|_{L^{\tilde p}(\Omega)}\right]. \tag{A.11}$$

Proof. Proof of (A.10): We will use the elementary estimate

$$\|v\|_{L^{\tilde q}(\Omega)} \leq \|v\|^{\tilde p/\tilde q}_{L^{\tilde p}(\Omega)}\|v\|^{1-\tilde p/\tilde q}_{L^\infty(\Omega)} \qquad \text{for } \tilde p\leq\tilde q<\infty \text{ and } v\in L^\infty(\Omega). \tag{A.12}$$

We start with $\|U(t,\cdot)\|_{L^\infty(\Omega)}$. Letting $\tilde p' = \tilde p/(\tilde p-1)$ be the conjugate exponent of $\tilde p$, we compute:

$$\begin{aligned} |U(t,\mathbf{x})| &= \int_{\mathbf{z}\in B_1} |u(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x}))|\,d\mathbf{z} \leq |B_1|^{1/\tilde p'}\left(\int_{\mathbf{z}\in B_1}|u(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x}))|^{\tilde p}\,d\mathbf{z}\right)^{1/\tilde p} \\ &= |B_1|^{1/\tilde p'}t^{-d/\tilde p}\left(\int_{\mathbf{z}\in tB_1}|u(\mathbf{x}+\mathbf{z}+t\psi(\mathbf{x}))|^{\tilde p}\,d\mathbf{z}\right)^{1/\tilde p} \leq |B_1|^{1/\tilde p'}t^{-d/\tilde p}\|u\|_{L^{\tilde p}(\Omega)}. \end{aligned} \tag{A.13}$$

For $\|U(t,\cdot)\|_{L^{\tilde p}(\Omega)}$, we use the Minkowski inequality (A.31) and the change of variables formula (cf., e.g., [EG92b, Thm. 2, Sec. 3.3.3] for the case of bi-Lipschitz changes of variables although in the present case, the mapping is smooth!) to compute

$$\begin{aligned} \|U(t,\cdot)\|_{L^{\tilde p}(\Omega)} &= \left(\int_{\mathbf{x}\in\Omega}\left|\int_{\mathbf{z}\in B_1}|u(\Psi_t(\mathbf{x})+t\mathbf{z})|\,d\mathbf{z}\right|^{\tilde p}\,d\mathbf{x}\right)^{1/\tilde p} \\ &\overset{(A.31)}{\leq} \int_{\mathbf{z}\in B_1}\left(\int_{\mathbf{x}\in\Omega}|u(\Psi_t(\mathbf{x})+t\mathbf{z})|^{\tilde p}\,d\mathbf{x}\right)^{1/\tilde p}\,d\mathbf{z} \\ &\leq \int_{\mathbf{z}\in B_1}\left(\int_{\mathbf{x}\in\Omega}|u(\mathbf{x})|^{\tilde p}\|\det D\Psi_t^{-1}\|_{L^\infty(\mathbb{R}^d)}\,d\mathbf{x}\right)^{1/\tilde p}\,d\mathbf{z} \leq C\|u\|_{L^{\tilde p}(\Omega)}, \end{aligned} \tag{A.14}$$

where the constant C depends only on T and $\operatorname{Lip}(\psi)$ (cf. Lemma A.3 for the t-dependence of Ψ_t). Inserting (A.13) and (A.14) in (A.12) yields (A.10).

Proof of (A.11) for the case $\tilde\sigma>0$ together with $\tilde p\geq 1$: This is a simple consequence of (A.10).

Proof of (A.11) for the case $\tilde\sigma=0$ together with $\tilde p>1$: From Lemma A.3 we have $\|\psi\|_{L^\infty(\mathbb{R}^d)}\leq C$. Hence, $B_1+\psi(\mathbf{x})\subset B_{1+C}$ uniformly in $\mathbf{x}$. If we implicitly assume that u is extended by zero outside of Ω, we can estimate

$$|U(t,\mathbf{x})| = \int_{\mathbf{z}\in B_1}|u(\mathbf{x}+t\mathbf{z}+t\psi(\mathbf{x})|\,d\mathbf{z} = \int_{\mathbf{z}\in B_1+\psi(\mathbf{x})}|u(\mathbf{x}+t\mathbf{z})|\,d\mathbf{z} \leq \int_{\mathbf{z}\in B_{1+C}}|u(\mathbf{x}+t\mathbf{z})|\,d\mathbf{z}.$$

Our goal is to show that the function

$$\mathbf{x}\mapsto\int_{t=0}^{T} t^{-1+\tilde\mu}\int_{\mathbf{z}\in B_1}u(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})))\,d\mathbf{z}$$

is in $L^{\tilde q}(\Omega)$ provided that $u\in L^{\tilde p}(\Omega)$. This is shown with Lemma A.14 below. We assume that u is extended by zero outside of Ω and bound the $L^{\tilde q}(\mathbb{R}^d)$-norm of

$$\mathbf{x}\mapsto\int_{t=0}^{T} t^{-1+\tilde\mu}\int_{\mathbf{z}\in B_{1+C}}|u(\mathbf{x}+t\mathbf{z})|\,d\mathbf{z}.$$

From Lemma A.15 (and a density argument to be able to work with $|u|$ instead of u) we get

$$\begin{aligned}&\int_{t=0}^{T} t^{-1+\tilde\mu}\int_{\mathbf{z}\in B_{1+C}} |u(\mathbf{x}+t\mathbf{z})|\,d\mathbf{z} = \\ &\frac{1}{d-\tilde\mu}\left[T^{\tilde\mu-d}\int_{\mathbf{z}\in B_{(1+C)R}}|u(\mathbf{x}+\mathbf{z})|\,d\mathbf{z} - \int_{\mathbf{z}\in B_{(1+C)R}}|\mathbf{z}|^{\tilde\mu-d}|u(\mathbf{x}+\mathbf{z})|\,d\mathbf{z}\right].\end{aligned} \tag{A.15}$$

The first contribution in (A.15) is estimated directly. For the second contribution, we obtain from Lemma A.14 (with $\lambda = d-\tilde\mu$, $r^{-1} = 1+\frac{\tilde\mu}{d}-\tilde p^{-1}$ so that $1-r^{-1} = \tilde p^{-1}-\frac{\tilde\mu}{d} = \tilde q^{-1}$)

$$\left\|\int_{\mathbf{z}\in B_{(1+C)R}}|\mathbf{z}|^{\tilde\mu-d}|u(\cdot+\mathbf{z})|\,d\mathbf{z}\right\|_{L^{\tilde q}(\mathbb{R}^d)} \lesssim \|u\|_{L^{\tilde p}(\mathbb{R}^d)},$$

which is the desired estimate. $\square$

Lemma A.7. *Let Ω, T, and ψ be as in Lemma A.3. Let $1\le\tilde p\le\tilde q<\infty$ and define $\tilde\mu := d(\tilde p^{-1}-\tilde q^{-1})$ as in (A.1). Let $\tilde s\in(0,1)$ and $u\in W^{\tilde s,\tilde p}(\Omega)$. Define, for $t\in(0,T)$, the function*

$$V(t,\mathbf{x}) := \int_{\mathbf{z}\in B_1}\int_{\mathbf{z}'\in B_1}|u(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x}))) - u(\mathbf{x}+t(\mathbf{z}'+\psi(\mathbf{x})))|\,d\mathbf{z}'\,d\mathbf{z}. \tag{A.16}$$

Then, the following two assertions hold true:

(i) There exists $C_1 = C_1(\tilde p,\tilde q,\tilde s,d,\mathrm{Lip}(\psi),T)$ such that

$$\|V(t,\cdot)\|_{L^{\tilde q}(\Omega)} \le C_1 t^{\tilde s-\tilde\mu}|u|_{W^{\tilde s,\tilde p}(\Omega)} \qquad \forall t\in(0,T). \tag{A.17}$$

(ii) If the pair $(\tilde\sigma,\tilde p)$ satisfies (A.8), then there is a constant $C_2 = C_2(\tilde p,\tilde q,\tilde s,d,\mathrm{Lip}(\psi),\sigma,T)$ such that

$$\left\|\int_{t=0}^{T} t^{-1+\tilde\mu-\tilde s+\tilde\sigma}V(t,\cdot)\,dt\right\|_{L^{\tilde q}(\Omega)} \le C_2|u|_{W^{\tilde s,\tilde p}(\Omega)}. \tag{A.18}$$

Proof. The proof is structurally similar to that of Lemma A.6. Define

$$v(\mathbf{y},\mathbf{y}') = \frac{|u(\mathbf{y})-u(\mathbf{y}')|}{|\mathbf{y}-\mathbf{y}'|^{s+d/p}}.$$

Letting $\tilde p' = \tilde p/(\tilde p-1)$ be the conjugate exponent of $\tilde p$, we compute

$$\begin{aligned}V(t,\mathbf{x}) &= \int_{\mathbf{z}}\int_{\mathbf{z}'}\frac{|u(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x}))) - u(\mathbf{x}+t(\mathbf{z}'+\psi(\mathbf{x})))|}{|\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})) - (\mathbf{x}+t(\mathbf{z}'+\psi(\mathbf{x})))|^{\tilde s+d/\tilde p}}(t|\mathbf{z}-\mathbf{z}'|)^{\tilde s+d/\tilde p}\,d\mathbf{z}'\,d\mathbf{z} \\ &\le (2t)^{\tilde s+d/\tilde p}\int_{\mathbf{z}}\int_{\mathbf{z}'} v(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})),\mathbf{x}+t(\mathbf{z}'+\psi(\mathbf{x}))\,d\mathbf{z}'\,d\mathbf{z} \\ &\le (2t)^{\tilde s+d/\tilde p}t^{-d}|B_t|^{1/\tilde p'}\int_{\mathbf{z}\in B_1}\|v(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})),\cdot)\|_{L^{tildep}(\Omega)}\,d\mathbf{z}, \\ &\le Ct^{\tilde s}\int_{\mathbf{z}\in B_1}\|v(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})),\cdot)\|_{L^{\tilde p}(\Omega)}\,d\mathbf{z}.\end{aligned} \tag{A.19}$$

We recognize that, up to the factor $t^{\tilde s}$, the right-hand side is a function of the form studied in Lemma A.6. The bounds (A.17) and (A.18) therefore follow from Lemma A.6. $\square$

Lemma A.8. *Let Ω be as in Theorem A.1. Assume $r \geq 0$, $1 \leq p \leq q < \infty$ and set $\mu := d(p^{-1} - q^{-1})$ as in (A.1). Assume that one of the following two cases occurs:*

(a) $r = \mu$ in conjunction with $p > 1$;

(b) $r > \mu \notin \mathbb{N}$ in conjunction with $p \geq 1$.

Then there is a constant $C = C(p, q, r, \mathrm{Lip}(\psi), T, d)$, which depends solely on the quantities indicated (and the assumption that $\operatorname{diam}\Omega \leq 1$), such that

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{W^{r,p}(\Omega)}.$$

Proof. We can assume $\mu > 0$, which implies $r > 0$ and $q > 1$. First we start with some preliminaries. For $L \in \mathbb{N}$ with

$$\frac{L}{d} \leq 1 - \frac{1}{q}. \tag{A.20}$$

define recursively the values $p_0, p_1, \ldots, p_L \in [1, \infty)$ by

$$\frac{1}{p_0} := \frac{1}{q}, \qquad \frac{1}{p_i} =: \frac{1}{p_{i-1}} + \frac{1}{d}, \qquad i = 1, \ldots, L. \tag{A.21}$$

Note that indeed $p_i \geq 1$ since (A.20) implies $p_L \geq 1$ in view of $p_L^{-1} = q^{-1} + Ld^{-1}$. Furthermore, we have the stronger assertion

$$1 < p_i, \qquad i = 0, \ldots, L - 1. \tag{A.22}$$

We claim that

$$\begin{cases} \|u\|_{L^{p_{i-1}}(\Omega)} \leq C\|u\|_{W^{1,p_i}(\Omega)}, & i = 1, \ldots, L-1, \\ \|u\|_{L^{p_{L-1}}(\Omega)} \leq C\|u\|_{W^{1,p_L}(\Omega)}, & \text{if } p_L > 1. \end{cases} \tag{A.23}$$

which implies in particular

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{W^{L,p_L}(\Omega)} \qquad \text{if } p_L > 1. \tag{A.24}$$

In order to see (A.23) we use the representation $u(\mathbf{x}) = M(u)(T, \mathbf{x}) + M_R(u)(T, \mathbf{x})$ from Lemma A.4 and the fact that ω is fixed and bounded and that we have control over ψ and $\nabla\psi$ by Lemma A.3. Hence we arrive at the following bound:

$$|u(\mathbf{x})| \leq C \left[\int_{\mathbf{z} \in B_1} |u(\mathbf{x} + T(\mathbf{z} + \psi(\mathbf{x})))| \, d\mathbf{z} + \int_{t=0}^{T} \int_{\mathbf{z} \in B_1} |\nabla u(\mathbf{x} + t(\mathbf{z} + \psi(\mathbf{x})))| \, d\mathbf{z} \, dt \right]. \tag{A.25}$$

For $i = 1, \ldots, L-1$, the first term in (A.25) is bounded by (A.10). The second term in (A.25) is bounded with (A.11) where we choose $\tilde{\mu} = d(p_i^{-1} - p_{i-1}^{-1}) = 1$ and $\tilde{\sigma} = 0$, which we may due to (A.22). This gives

$$\|u\|_{L^{p_{i-1}}(\Omega)} \leq C \left[T^{-1}\|u\|_{L^{p_i}(\Omega)} + \|\nabla u\|_{L^{p_i}(\Omega)} \right] \leq C\|u\|_{W^{1,p_i}(\Omega)},$$

which is the first part of (A.23). The case $i = L$ in (A.23) follows in the same way if $p_L > 1$. Using (A.24) we can show the desired result as follows.

(i) The case $r = \mu \in \mathbb{N}$.
Note that the choice $L = r$ satisfies (A.20) and that $p_L = p > 1$. Therefore, (A.24) is the desired estimate.

(ii) The case $\mu \notin \mathbb{N}$ and $\mu \le r < \lceil \mu \rceil$.
As $p_{\lfloor \mu \rfloor} > p \ge 1$, we obtain with (A.24) the bound $\|u\|_{q,\Omega} \lesssim \|u\|_{\lfloor \mu \rfloor, p_{\lfloor \mu \rfloor},\Omega}$.

For $|\mathbf{t}| = \lfloor \mu \rfloor$ we write $D^{\mathbf{t}} u(\mathbf{x}) = M(D^{\mathbf{t}} u)(T,\mathbf{x}) + M_S(D^{\mathbf{t}} u)(T,\mathbf{x})$ and use Lemma A.7, where we set $\tilde{q} = p_{\lfloor \mu \rfloor}$, $\tilde{p} = p$, and $\tilde{s} = r - \lfloor \mu \rfloor$. Since $\lfloor \mu \rfloor = d(p_{\lfloor \mu \rfloor}^{-1} - q^{-1})$, we see $\tilde{\mu} - \tilde{s} = \mu - r$, such that for $\mu = r$ we choose $\tilde{\sigma} = 0$ and hence require $p > 1$, while for $\mu < r$ we can choose $\tilde{\sigma} > 0$ and hence $p \ge 1$. This shows $|u|_{\lfloor \mu \rfloor, p_{\lfloor \mu \rfloor},\Omega} \lesssim |u|_{r,p,\Omega}$.

For $|\mathbf{t}| \le \lfloor \mu \rfloor - 1$ (given $\mu > 1$) we write $D^{\mathbf{t}} u(\mathbf{x}) = M(D^{\mathbf{t}} u)(T,\mathbf{x}) + M_R(D^{\mathbf{t}} u)(T,\mathbf{x})$ and use Lemma A.6, where we set $\tilde{q} = p_{\lfloor \mu \rfloor}$, $\tilde{p} = p$. Since then $\tilde{\mu} = \mu - \lfloor \mu \rfloor < 1$, we obtain for any $p \ge 1$ that $|u|_{|\mathbf{t}|, p_{\lfloor \mu \rfloor},\Omega} \lesssim |u|_{|\mathbf{t}|+1,p,\Omega}$.

(iii) The case $\mu \notin \mathbb{N}$ and $\lceil \mu \rceil \le r$.
It suffices to consider $r = \lceil \mu \rceil$. As $p_{\lfloor \mu \rfloor} > p \ge 1$, we obtain with (A.24) the bound $\|u\|_{q,\Omega} \lesssim \|u\|_{\lfloor \mu \rfloor, p_{\lfloor \mu \rfloor},\Omega}$.

For $|\mathbf{t}| \le \lfloor \mu \rfloor$ we write $D^{\mathbf{t}} u(\mathbf{x}) = M(D^{\mathbf{t}} u)(T,\mathbf{x}) + M_R(D^{\mathbf{t}} u)(T,\mathbf{x})$ and use Lemma A.6, where we set $\tilde{q} = p_{\lfloor \mu \rfloor}$, $\tilde{p} = p$. Since then $\tilde{\mu} = \mu - \lfloor \mu \rfloor < 1$, we obtain for any $p \ge 1$ that $|u|_{|\mathbf{t}|, p_{\lfloor \mu \rfloor},\Omega} \lesssim |u|_{|\mathbf{t}|+1,p,\Omega}$. In total, this yields $\|u\|_{q,\Omega} \le \|u\|_{\lceil \mu \rceil,p,\Omega}$.

□

A.2 The case of fractional s in Theorem A.1

The analog of Lemma A.6 is the following result.

Lemma A.9. *Let Ω, T, ψ be as in Theorem A.1. Let $1 \le \tilde{p} \le \tilde{q} < \infty$. Set $\tilde{\mu} = d(\tilde{p}^{-1} - \tilde{q}^{-1})$. Let $K = K(\mathbf{x},\mathbf{z})$ be defined on $\Omega \times \mathbb{R}^d$ with $\operatorname{supp} K(\mathbf{x},\cdot) \subset B_1$ for every $\mathbf{x} \in \Omega$. Let K be bounded (bound $\|K\|_{L^\infty}$) and Lipschitz continuous with Lipschitz constant $\mathrm{Lip}(K)$. Define the function*

$$V(t,\mathbf{x}) := \int_{\mathbf{z} \in B_1} K(\mathbf{x},\mathbf{z}) u(\mathbf{x} + t(\mathbf{z} + \psi(\mathbf{x})))\, d\mathbf{z}.$$

(i) Let $\tilde{s} \in (0,1)$. Then there exists $C_1 = C_1(\tilde{p},\tilde{q},\tilde{s},d,\mathrm{Lip}(\psi),T,\|K\|_{L^\infty},\mathrm{Lip}(K))$ such that

$$|V(t,\cdot)|_{W^{\tilde{s},\tilde{q}}(\Omega)} \le C_1 t^{-\tilde{\mu}-\tilde{s}} \|u\|_{L^{\tilde{p}}(\Omega)}.$$

(ii) Let $\tilde{s} \in (0,1)$ and assume that the pair $(\tilde{\sigma},\tilde{p})$ satisfies (A.8). Then there exists a constant $C_2 = C_2(\tilde{p},\tilde{q},\tilde{s},d,\mathrm{Lip}(\psi),T,\|K\|_{L^\infty},\mathrm{Lip}(K),\sigma)$ such that

$$\left| \int_{t=0}^{T} t^{-1+\tilde{s}+\tilde{\mu}+\tilde{\sigma}} V(t,\cdot) \right|_{W^{\tilde{s},\tilde{q}}(\Omega)} \le C_2 \|u\|_{L^{\tilde{p}}(\Omega)}.$$

Proof. Proof of (i): Let $\mathbf{x}$, $\mathbf{y} \in \Omega$ and $t \in (0,T)$. Define the translation $\mathbf{t} := \frac{\mathbf{x}-\mathbf{y}}{t} + \psi(\mathbf{x}) - \psi(\mathbf{y})$ and denote by χ_A the characteristic function of a set A. An affine change of variables gives for $\mathbf{x}$, $\mathbf{y} \in \Omega$ and $t \in (0,T)$

$$\begin{aligned} V(t,\mathbf{x}) &= \int_{\mathbf{z}\in B_1} K(\mathbf{x},\mathbf{z})u(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})))\,d\mathbf{z} = \int_{\mathbf{z}\in\mathbb{R}^d} K(\mathbf{x},\mathbf{z})u(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})))\chi_{B_1}(\mathbf{z})\,d\mathbf{z} \\ &= \int_{\mathbf{z}'\in\mathbb{R}^d} K(\mathbf{x},\mathbf{z}'-\mathbf{t})u(\mathbf{y}+t(\mathbf{z}'+\psi(\mathbf{y})))\chi_{B_1+\mathbf{t}}(\mathbf{z}')\,d\mathbf{z}' \\ &= V(t,\mathbf{y}) + \int_{\mathbf{z}} \underbrace{[K(\mathbf{x},\mathbf{z}-\mathbf{t})\chi_{B_1+\mathbf{t}}(\mathbf{z}) - K(\mathbf{y},\mathbf{z})\chi_{B_1}(\mathbf{z})]}_{=:B(\mathbf{x},\mathbf{y},\mathbf{z})} u(\mathbf{y}+t(\mathbf{z}+\psi(\mathbf{y}))\,d\mathbf{z}. \end{aligned}$$

We estimate the function B. We have the obvious estimate $|B(\mathbf{x},\mathbf{y},\cdot)| \le \|K\|_{L^\infty}[\chi_{B_1}+\chi_{B_1+\mathbf{t}}]$. For further estimates, we start by noting

$$|\mathbf{t}| \le |\mathbf{x}-\mathbf{y}|/t + \|\nabla\psi\|_{L^\infty}|\mathbf{x}-\mathbf{y}| \le C|\mathbf{x}-\mathbf{y}|/t, \tag{A.26}$$

where C depends on the Lipschitz constant of ψ and on T. We also note

$$\begin{aligned} &\mathbf{z}\in B_1\setminus(B_1+\mathbf{t}) &&\Longrightarrow (|\mathbf{z}|\le 1 \quad\wedge\quad |\mathbf{z}-\mathbf{t}|\ge 1) &&\Longrightarrow 1-|\mathbf{t}| \le |\mathbf{z}| \le 1, \\ &\mathbf{z}\in(\mathbf{t}+B_1)\setminus B_1 &&\Longrightarrow (|\mathbf{z}|\ge 1 \quad\wedge\quad |\mathbf{z}-\mathbf{t}|\le 1) &&\Longrightarrow 1-|\mathbf{t}| \le |\mathbf{z}-\mathbf{t}| \le 1. \end{aligned}$$

Since K is Lipschitz continuous on $\Omega\times\mathbb{R}^d$ and $\operatorname{supp} K(\mathbf{x},\cdot)\subset B_1$ we get $|K(\mathbf{x},\mathbf{z})|\le C\mathbf{t}$ for every $\mathbf{z}\in R := (B_1\setminus(B_1+\mathbf{t}))\cup((B_1+\mathbf{t})\setminus B_1)$, where the constant C depends only on the Lipschitz constant of K. We therefore get $|B(\mathbf{x},\mathbf{y},\mathbf{z})|\le C\mathbf{t}$ for $\mathbf{z}\in R$. For the case $\mathbf{z}\in B_1\cap(B_1+\mathbf{t})$, we get from the Lipschitz continuity of K that $|B(\mathbf{x},\mathbf{y},\mathbf{z})| \le C[|\mathbf{x}-\mathbf{y}|+|\mathbf{x}-\mathbf{y}|/t] \le C|\mathbf{x}-\mathbf{y}|/t$. Putting together the above estimates for B, we arrive at

$$|B(\mathbf{x},\mathbf{y},\mathbf{z})| \le C\min\{1,|\mathbf{x}-\mathbf{y}|/t\}\,[\chi_{B_1}(\mathbf{z})+\chi_{\mathbf{t}+B_1}(\mathbf{z})]\,.$$

In total, we get

$$\begin{aligned} |V(t,\mathbf{x})-V(t,\mathbf{y})| &\le C\min\{1,|\mathbf{x}-\mathbf{y}|/t\}\int_{B_1\cup\mathbf{t}+B_1} |u(\mathbf{y}+t(\mathbf{z}+\psi(\mathbf{y})))|\,d\mathbf{z} \\ &= C\min\{1,|\mathbf{x}-\mathbf{y}|/t\}\left[\int_{\mathbf{z}\in B_1}|u(\mathbf{x}+t(\mathbf{x}+\psi(\mathbf{x})))|\,d\mathbf{z} + \int_{\mathbf{z}\in B_1}|u(\mathbf{y}+t(\mathbf{y}+\psi(\mathbf{y})))|\,d\mathbf{z}\right] \\ &\le C\min\{1,|\mathbf{x}-\mathbf{y}|/t\}\,[U(t,\mathbf{x})+U(t,\mathbf{y})]\,, \end{aligned}$$

where, in the last step we have inserted the definition of the function U from (A.9). Therefore,

$$\frac{|V(t,\mathbf{x})-V(t,\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|^{\tilde{s}+d/\tilde{q}}} \le C\frac{\min\{1,|\mathbf{x}-\mathbf{y}|/t\}}{|\mathbf{x}-\mathbf{y}|^{\tilde{s}+d/\tilde{q}}}\,[U(t,\mathbf{x})+U(t,\mathbf{y})]\,. \tag{A.27}$$

In view of the symmetry in the variables $\mathbf{x}$ and $\mathbf{y}$, we will only consider one type of integral. We compute

$$\begin{aligned} &\int_{\mathbf{y}\in\Omega}|U(t,\mathbf{x})|^{\tilde{q}}\int_{\mathbf{y}\in\Omega}\frac{(\min\{1,|\mathbf{x}-\mathbf{y}|/t\})^{\tilde{q}}}{|\mathbf{x}-\mathbf{y}|^{\tilde{s}\tilde{q}+d}}\,d\mathbf{y} \\ &\qquad\lesssim |U(t,\mathbf{x})|^{\tilde{q}}\left[\int_{r=0}^{t}\frac{(r/t)^{\tilde{q}}}{r^{\tilde{s}\tilde{q}+d}}r^{d-1}\,dr + \int_{r=t}^{\infty} r^{-\tilde{s}\tilde{q}-d}r^{d-1}\,dr\right] \lesssim t^{-\tilde{s}\tilde{q}}|U(t,\mathbf{x})|^{\tilde{q}}, \end{aligned} \tag{A.28}$$

where the hidden constants depend only on $\tilde{s}$ and $\tilde{q}$. We conclude

$$|V(t,\cdot)|^{\tilde{q}}_{W^{\tilde{s},\tilde{q}}(\Omega)} \leq Ct^{-\tilde{s}\tilde{q}} \|U(t,\cdot)\|^{\tilde{q}}_{L^{\tilde{q}}(\Omega)} \leq Ct^{-\tilde{s}\tilde{q}-\tilde{\mu}\tilde{q}} \|u\|^{\tilde{q}}_{L^{\tilde{p}}(\Omega)},$$

where the last step follows from (A.10).

Proof of (ii): Starting from (A.27) we have to estimate

$$I := \int_{\mathbf{x}\in\Omega} \int_{\mathbf{y}\in\Omega} \frac{1}{|\mathbf{x}-\mathbf{y}|^{\tilde{s}\tilde{q}+d}} \left| \int_{t=0}^{T} t^{-1+\tilde{\mu}+\tilde{\sigma}} \min\{1, |\mathbf{x}-\mathbf{y}|/t\} U(t,\mathbf{x})\, dt \right|^{\tilde{q}} d\mathbf{y}\, d\mathbf{x}.$$

Applying the Minkowski inequality (A.31), we obtain (recalling the calculation performed in (A.28))

$$\begin{aligned} I &\leq \int_{\mathbf{x}\in\Omega} \left\{ \int_{t=0}^{T} \left(\int_{\mathbf{y}\in\Omega} \frac{1}{|\mathbf{x}-\mathbf{y}|^{\tilde{s}\tilde{q}+d}} t^{(-1+\tilde{\mu}+\tilde{\sigma})\tilde{q}} \left(\min\{1, |\mathbf{x}-\mathbf{y}|/t\}\right)^{\tilde{q}} |U(t,\mathbf{x})|^{\tilde{q}}\, d\mathbf{y} \right)^{1/\tilde{q}} dt \right\}^{\tilde{q}} d\mathbf{x} \\ &\lesssim \int_{\mathbf{x}\in\Omega} \left\{ \int_{t=0}^{T} |U(t,\mathbf{x})| t^{-1+\tilde{\mu}+\tilde{\sigma}}\, dt \right\}^{\tilde{q}} d\mathbf{x} \lesssim \|u\|^{\tilde{q}}_{L^{\tilde{p}}(\Omega)}, \end{aligned}$$

where, in the last step, we used (A.11). □

The analog of Lemma A.7 is as follows:

Lemma A.10. *Let Ω, T, ψ be as in Theorem A.1. Let $K = K(\mathbf{x},\mathbf{z},\mathbf{z}')$ be defined on $\Omega \times \mathbb{R}^d \times \mathbb{R}^d$ with $\operatorname{supp} K(\mathbf{x},\cdot,\cdot) \subset B_1 \times B_1$ for every $\mathbf{x} \in \Omega$. Let K be bounded (bound $\|K\|_{L^\infty}$) and Lipschitz continuous with Lipschitz constant $\operatorname{Lip}(K)$.*

Let $\tilde{s}$, $\tilde{r} \in (0,1)$. Let $1 \leq \tilde{p} \leq \tilde{q} < \infty$. Define $\tilde{\mu} = d(\tilde{p}^{-1} - \tilde{q}^{-1})$. For $u \in W^{\tilde{r},\tilde{p}}(\Omega)$ define for $t \in (0,T)$, the function

$$V(t,\mathbf{x}) := \int_{\mathbf{z}\in B_1} \int_{\mathbf{z}'\in B_1} K(\mathbf{x},\mathbf{z},\mathbf{z}') \left[u(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x}))) - u(\mathbf{x}+t(\mathbf{z}'+\psi(\mathbf{x}))) \right] d\mathbf{z}'\, d\mathbf{z}. \tag{A.29}$$

Then:

(i) There exists $C_1 = C_1(\tilde{p},\tilde{q},\tilde{r},\tilde{s},T,\operatorname{Lip}(\psi),d,\|K\|_{L^\infty},\operatorname{Lip}(K))$ such that

$$|V(t,\cdot)|_{W^{\tilde{s},\tilde{p}}(\Omega)} \leq C_1 t^{-\tilde{s}-\tilde{\mu}+\tilde{r}} |u|_{W^{\tilde{r},\tilde{p}}(\Omega)} \qquad \text{for all } t \in (0,T).$$

(ii) If the pair $(\tilde{\sigma},\tilde{p})$ satisfies (A.8), then there exists $C_2 = C_2(\tilde{p},\tilde{q},\tilde{r},\tilde{s},T,\operatorname{Lip}(\psi),d,\|K\|_{L^\infty},\operatorname{Lip}(K),\tilde{\sigma})$ such that

$$\left\| \int_{t=0}^{T} t^{-1+\tilde{s}-\tilde{r}+\tilde{\mu}+\tilde{\sigma}} V(t,\cdot)\, dt \right\|_{W^{\tilde{s},\tilde{q}}(\Omega)} \leq C_2 |u|_{W^{\tilde{r},\tilde{p}}(\Omega)}.$$

Proof. We proceed as in the proof of Lemma A.9. With the translation vector $\mathbf{t}$ there and the analogous change of variables in $\mathbf{z}$ and $\mathbf{z}'$ we obtain

$$V(t,\mathbf{x}) - V(t,\mathbf{y}) = \int_{\mathbf{z}\in B_1} \int_{\mathbf{z}'\in B_1} B(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{z}') \left[u(\mathbf{y}+t(\mathbf{z}+\psi(\mathbf{y}))) - u(\mathbf{y}+t(\mathbf{z}'+\psi(\mathbf{y}))) \right] d\mathbf{z}'\, d\mathbf{z},$$

where

$$B(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{z}') := K(\mathbf{x},\mathbf{z}-\mathbf{t},\mathbf{z}'-\mathbf{t})\chi_{B_1+\mathbf{t}}(\mathbf{z})\chi_{B_1+\mathbf{t}}(\mathbf{z}') - K(\mathbf{y},\mathbf{z},\mathbf{z}')\chi_{B_1}(\mathbf{z})\chi_{B_1}(\mathbf{z}').$$

As in the proof of Lemma A.9, we get with $Z := B_1 \cup (B_1+\mathbf{t})$:

$$|B(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{z}')| \le C\min\{1,|\mathbf{x}-\mathbf{y}|/t\}\chi_Z(\mathbf{z})\chi_Z(\mathbf{z}'). \tag{A.30}$$

Upon setting

$$v(\mathbf{y},\mathbf{y}') = \frac{|u(\mathbf{y})-u(\mathbf{y}')|}{|\mathbf{y}-\mathbf{y}'|^{\tilde r + d/\tilde p}}$$

we get in analogy to the procedure in (A.19)

$$\begin{aligned}|V(t,\mathbf{x})-V(t,\mathbf{y})| &\lesssim \min\{1,|\mathbf{x}-\mathbf{y}|/t\}t^{\tilde r}\int_{\mathbf{z}\in Z}\|v(\mathbf{y}+t(\mathbf{z}+\psi(\mathbf{y})),\cdot)\|_{L^{\tilde p}(\Omega)}\,d\mathbf{z}\\ &= \min\{1,|\mathbf{x}-\mathbf{y}|/t\}t^{\tilde r}\int_{\mathbf{z}\in B_1}\|v(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})),\cdot\|_{L^{\tilde p}(\Omega)} + \|v(\mathbf{y}+t(\mathbf{z}+\psi(\mathbf{y})),\cdot)\|_{L^{\tilde p}(\Omega)}\,d\mathbf{z}\end{aligned}$$

We recognize the similarity with the situation in Lemma A.7. We set $U(t,\mathbf{x}) := \int_{\mathbf{z}\in B_1}\|v(\mathbf{x}+t(\mathbf{z}+\psi(\mathbf{x})),\cdot)\|_{L^{\tilde p}(\Omega)}$ and arrive at

$$\frac{|V(t,\mathbf{x})-V(t,\mathbf{y})|}{|\mathbf{x}-\mathbf{x}|^{\tilde s+d/\tilde q}} \lesssim \frac{\min\{1,|\mathbf{x}-\mathbf{y}|/t\}}{|\mathbf{x}-\mathbf{y}|^{\tilde s+d/\tilde q}}t^{\tilde r}\left[U(t,\mathbf{x})+U(t,\mathbf{y})\right].$$

Again, given the symmetry in the variables $\mathbf{x}$ and $\mathbf{y}$, we get as in (A.28)

$$|V(t,\cdot)|_{W^{\tilde s,\tilde q}(\Omega)} \lesssim t^{-\tilde s-\tilde\mu+\tilde r}|u|_{W^{\tilde r,\tilde p}(\Omega)},$$

which is the assertion of part (i) of the lemma. For part (ii) of the lemma we proceed analogously to the proof of Lemma A.9, (ii). □

We now come to the analog of Lemma A.8.

Lemma A.11. *Let Ω be as in Theorem A.1. Let $1 \le p \le q < \infty$. Define $\mu = d(p^{-1}-q^{-1})$ as in (A.1). Let $s \in (0,1)$ with $s+\mu \le 1$ and assume that one of the following cases occurs:*

(a) $r = s+\mu$ and $p > 1$.

(b) $r > s+\mu$ and $p \ge 1$.

Then, there is a constant $C = C(p,q,r,s,\mathrm{Lip}(\psi),T,d)$ such that

$$|u|_{W^{s,q}(\Omega)} \le C\|u\|_{W^{r,p}(\Omega)}.$$

Proof. We do not lose generality if we assume $\mu > 0$. The proof is divided into several steps.

(i) The case $s+\mu \le r < 1$.
We use Lemma A.4 and write $u(\mathbf{x}) = M(u)(T,\mathbf{x}) + M_S(u)(T,\mathbf{x})$. From Lemma A.9, (i) we get $|M(u)(T,\cdot)|_{W^{s,q}(\Omega)} \le C\|u\|_{L^p(\Omega)}$. Lemma A.10, (ii) implies $|M_S(u)(T,\cdot)|_{W^{s,q}(\Omega)} \le C|u|_{W^{r,p}(\Omega)}$ if either $r-s=\mu$ together with $p>1$ or $r-s>\mu$ together with $p \ge 1$.

(ii) The case $s + \mu < 1 \leq r$.
It suffices to consider the case $r = 1$. We use Lemma A.4 and write $u(\mathbf{x}) = M(u)(T, \mathbf{x}) + M_R(u)(T, \mathbf{x})$. In Lemma A.9 we choose $\tilde{s} = s$, $\tilde{q} = q$, $\tilde{p} = p$ and obtain due to (i) the bound $|M(u)(T, \cdot)|_{W^{s,q}(\Omega)} \lesssim \|u\|_{L^p(\Omega)}$. Furthermore, since $\tilde{s} + \tilde{\mu} < 1$, we can choose $\tilde{\sigma} > 0$ and obtain $|M_R(u)(T, \cdot)|_{s,q,\Omega} \lesssim \|u\|_{1,p,\Omega}$.

(iii) The case $s + \mu = 1 \leq r$.
We use again the representation $u(\mathbf{x}) = M(u)(T, \mathbf{x}) + M_R(T, \mathbf{x})$. The contribution $M(u)(T, \cdot)$ is treated again with Lemma A.9, (i). If $r = 1$, the contribution $M_R(u)(T, \cdot)$ is handled by Lemma A.9, (ii), where we choose $\tilde{\sigma} = 0$ and hence require $p > 1$. If, on the other hand, $r > 1$, choose an arbitrary $\mu_\star$ with $0 < \mu_\star < \mu$ and $r > 1 + \mu - \mu_\star$ and define $p_\star$ via $\mu_\star = d(p_\star^{-1} - q^{-1})$. Note that $p_\star > p \geq 1$. As $s + \mu_\star < 1$, we obtain from step (ii) that $|u|_{s,q,\Omega} \lesssim \|u\|_{1,p_\star,\Omega}$. As $r - 1 > \mu - \mu_\star = d(p^{-1} - p_\star^{-1}) \notin \mathbb{N}$, we obtain from Lemma A.8, (b), that $\|u\|_{1,p_\star,\Omega} \lesssim \|u\|_{r,p,\Omega}$.

□

The following result complements Lemma A.8 with the cases $r > \mu \in \mathbb{N}$.

Lemma A.12. *Let Ω be as in Theorem A.1. Assume $r \geq 0$, $1 \leq p \leq q < \infty$ and set $\mu := d(p^{-1} - q^{-1})$ as in (A.1). Assume that $r > \mu \in \mathbb{N}$. Then there is a constant $C = C(p, q, r, \mathrm{Lip}(\psi), T, d)$, which depends solely on the quantities indicated (and the assumption that* $\operatorname{diam} \Omega \leq 1$*), such that*

$$\|u\|_{L^q(\Omega)} \leq C \|u\|_{W^{r,p}(\Omega)}.$$

Proof. Choose a $\mu_\star \notin \mathbb{N}$ with $\mu - 1 < \mu_\star < \mu$ and define $p_\star$ via $\mu_\star = d(p_\star^{-1} - q^{-1})$. Note that $p_\star > p \geq 1$. Lemma A.8 shows $\|u\|_{q,\Omega} \lesssim \|u\|_{\mu_\star, p_\star, \Omega}$. In a second step, observe that

$$1 = \underbrace{\mu_\star - (\mu - 1)}_{<1} + \underbrace{\mu - \mu_\star}_{d(p^{-1} - p_\star^{-1})} < \underbrace{r - (\mu - 1)}_{>1}$$

Hence we can apply Lemma A.11 for $|\mathbf{t}| = \mu - 1$ and obtain $\left|D^{\mathbf{t}} u\right|_{\mu_\star - (\mu-1), p_\star, \Omega} \lesssim \left\|D^{\mathbf{t}} u\right\|_{r - (\mu - 1), p, \Omega}$. Furthermore, since $\mu - \mu_\star < 1$, we can use Lemma A.8 for $|\mathbf{t}| \leq \mu - 1$ to obtain $\left|D^{\mathbf{t}} u\right|_{p_\star, \Omega} \lesssim \left\|D^{\mathbf{t}} u\right\|_{1,p,\Omega}$. As we took all terms of $\|u\|_{\mu_\star, p_\star, \Omega}$ into account, the result follows. □

The following results complements Lemma A.11 with the case $s + \mu > 1$.

Lemma A.13. *Let Ω be as in Theorem A.1. Let $1 \leq p \leq q < \infty$. Define $\mu = d(p^{-1} - q^{-1})$ as in (A.1). Let $s \in (0, 1)$ with $s + \mu > 1$ and assume that one of the following cases occurs:*

(a) $r = s + \mu$ and $p > 1$.

(b) $r > s + \mu$ and $p \geq 1$.

Then, there is a constant $C = C(p, q, r, s, \mathrm{Lip}(\psi), T, d)$ such that

$$|u|_{W^{s,q}(\Omega)} \leq C \|u\|_{W^{r,p}(\Omega)}.$$

Proof. Define $p_\star$ by $1 - s = d(p_\star^{-1} - q^{-1})$. As $s + \mu > 1$, it holds $p_\star > p \geq 1$. By Lemma A.11, we get $|u|_{W^{s,q}(\Omega)} \lesssim \|u\|_{W^{1,p_\star}(\Omega)}$. In a second step, observe that $r - 1 \geq s - 1 + \mu = d\left(p^{-1} - p_\star^{-1}\right)$, and hence Lemmas A.8 and A.12 imply the estimate $\|u\|_{W^{1,p_\star}(\Omega)} \leq \|u\|_{W^{r,p}(\Omega)}$. This concludes the proof. □

A.3 Auxiliary results

We need the Minkowski inequality (cf. [Ste70, Appendix A.1], [DL93, Chap. 2, eqn. (1.6)]

$$\left(\int_{\mathbf{y}\in\mathcal{Y}}\left(\int_{\mathbf{x}\in\mathcal{X}}|F(\mathbf{x},\mathbf{y})|\,d\mathbf{x}\right)^p d\mathbf{y}\right)^{1/p} \le \int_{\mathbf{x}\in\mathcal{X}}\left(\int_{\mathbf{y}\in\mathcal{Y}}|F(\mathbf{x},\mathbf{y})|^p\,d\mathbf{y}\right)^{1/p} d\mathbf{x}, \qquad 1\le p<\infty. \quad \text{(A.31)}$$

We will also need an application of the Marcinkiewicz interpolation theorem which is worked out in [RS75, Example 4, Sec. IX.4]:

Lemma A.14. *Let $1<p,q<\infty$ and assume $0<\lambda<d$. Let $p^{-1}+r^{-1}+\lambda d^{-1}=2$. Then*

$$\int_{\mathbf{x}\in\mathbb{R}^d}\int_{\mathbf{y}\in\mathbb{R}^d}\frac{|f(\mathbf{x})||g(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|^\lambda}\,d\mathbf{x},\,d\mathbf{y}\le C\|f\|_{L^p(\mathbb{R}^d)}\|g\|_{L^r(\mathbb{R}^d)}$$

for all $f\in L^p(\mathbb{R}^d)$, $g\in L^r(\mathbb{R}^d)$. The constant C depends only on p, r, λ, and d. That is, the map $f\mapsto\int_{\mathbb{R}^d}f(\mathbf{y})|\mathbf{x}-\mathbf{y}|^{-\lambda}\,d\mathbf{y}$ is a bounded linear map from $L^p(\mathbb{R}^d)$ to $L^{r/(r-1)}(\mathbb{R}^d)$.

Furthermore, we need the following result.

Lemma A.15. *Let $u\in C_0^\infty(\mathbb{R}^d)$ and B_R be a ball of radius R centered at the origin. Let $s\in[0,1)$ for $d\ge2$ and $s\in(0,1)$ for $d=1$. Then for $T>0$*

$$\int_{t=0}^T t^{-s}\int_{\mathbf{z}\in B_R}u(t\mathbf{z})\,d\mathbf{z}\,dt=\frac{1}{d-1+s}\left[\int_{\mathbf{z}\in B_{RT}}|\mathbf{z}|^{-s-d+1}u(\mathbf{z})\,d\mathbf{z}-T^{-s-d+1}\int_{B_{RT}}u(\mathbf{z})\,d\mathbf{z}\right].$$

Proof. The proof follows from the introduction of polar coordinates, Fubini, and an integration by parts. We use the polar coordinate (r,ω) with $\omega\in\partial B_1$. Then

$$\begin{aligned}\int_{t=0}^T t^{-s}\int_{\mathbf{z}\in B_R}u(t\mathbf{z}) &= \int_{\omega\in\partial B_1}\int_{t=0}^T t^{-s}\int_{r=0}^R u(tr\omega)r^{d-1} = \int_{\omega\in\partial B_1}\int_{t=0}^T t^{-s}t^{-(d-1)-1}\int_{\rho=0}^{Rt}u(\rho\omega)\rho^{d-1}\\ &= \int_{\omega\in\partial B_1}\frac{1}{-s-d+1}\left[t^{-s-d+1}\int_{\rho=0}^{Rt}u(\rho\omega)\rho^{d-1}\Big|_{t=0}^{t=T}-\int_{t=0}^T t^{-s-d+1}u(t\omega)t^{d-1}\right]\\ &= \frac{1}{-s-d+1}\left[T^{-s-d+1}\int_{B_{RT}}u(\mathbf{z})\,d\mathbf{z}-\int_{\mathbf{z}\in B_{RT}}|\mathbf{z}|^{-s-d+1}u(\mathbf{z})\,d\mathbf{z}\right].\end{aligned}$$

□

B Element-by-element approximation for variable polynomial degree in 3D

In this section, we generalize the operator Π_p^{MPS} of [MPS13, Thm. B.3] to variable (γ_p-shape regular) polynomial degree distributions. Structurally, we proceed as in [MPS13, Appendix B]: we define polynomial approximation operators that permit an "element by element" construction of a global piecewise polynomial approximation operator. To that end, we will fix the operator in vertices, edges, faces, and elements separately. We need to define approximation operators defined on edges, face, elements. Since we need to distinguish between "low" and "high" polynomial degree, we introduce **two** operators for edges, faces, and elements.

Definition B.1. *Let $\widehat{K} \subset \mathbb{R}^3$ be the reference tetrahedron. Let $\mathcal{V}$ be the set of the 4 vertices, $\mathcal{E}$ be the set of the 6 edges, and $\mathcal{F}$ be the set of the 4 faces of $\widehat{K}$. Let p_e, p_f, p_K be polynomial degrees associated with an edge e, a face f, and the element $\widehat{K}$. For an edge e we denote by $\mathcal{V}(e)$ the set of endpoints. For a face f, we denote by $\mathcal{V}(f)$ the set of vertices of f and by $\mathcal{E}(f)$ the set of edges of f.*

(i) (edge operators) For an edge e with associated polynomial degree p_e define the operators $\pi_e^h : C^\infty(\overline{e}) \to \mathcal{P}_{p_e}(e)$ and $\pi_e^p : C^\infty(\overline{e}) \to \mathcal{P}_{\lfloor p_e/4 \rfloor}(e)$ by

- *$(\pi_e^h u) \in \mathcal{P}_{p_e}$ is the unique minimizer of*

$$v \mapsto \|u - v\|_{L^2(e)}$$

  under the constraint $(\pi_e^h u)(V) = u(V)$ for all $V \in \mathcal{V}(e)$.

- *$(\pi_e^p u) \in \mathcal{P}_{\lfloor p_e/4 \rfloor}$ is the unique minimizer of*

$$v \mapsto p_e^4 \sum_{j=0}^{4} p_e^{-j} |u - v|_{H^j(e)}$$

  under the constraint that $\partial_e^j v(V) = \partial_e^j u(V)$ for $j \in \{0, 1, 2, 3\}$ and all $V \in \mathcal{V}(e)$. Here, ∂_e denote the tangential derivative along e. For π_e^p to be meaningful, we have to require $p_e \geq 28$.

(ii) (face operators) For a face f with associated polynomial degree p_f define the operators $\pi_f^h : C^\infty(\overline{f}) \to \mathcal{P}_{p_f}(f)$ and $\pi_f^p : C^\infty(\overline{f}) \to \mathcal{P}_{\lfloor p_f/2 \rfloor}(f)$ as follows, assuming *that a continuous, piecewise polynomial (of degree $\leq p_f$) approximation $\pi_{\partial f} u$ on the boundary of f is given:*

- *$(\pi_f^h u) \in \mathcal{P}_{p_f}$ is the unique minimizer of*

$$v \mapsto \|u - v\|_{L^2(f)}$$

  under the constraint $(\pi_f^h u)|_e = (\pi_{\partial f} u)|_e$ for all $e \in \mathcal{E}(f)$.

- *Assume that $\pi_{\partial f} u$ is given by $\pi_e^p u$ for all three edges $e \in \mathcal{E}(f)$. Then, $(\pi_f^p u) \in \mathcal{P}_{2\lfloor p_f/4 \rfloor}$ is the unique minimizer of*

$$v \mapsto p_f^4 \sum_{j=0}^{4} p_f^{-j} |u - v|_{H^j(f)}$$

under the constraint that $(\pi_f^p u)|_e = \pi_e^p u$ for all $e \in \mathcal{E}(f)$ and additionally the mixed derivatives at the 3 vertices $V \in \mathcal{V}(f)$ satisfy $(\partial_{e_1,V}\partial_{e_2,V}\pi_f u)(V) = (\partial_{e_1,V}\partial_{e_2,V} u)(V)$; here, at a vertex $V \in \mathcal{V}(f)$, e_1 and e_2 are the two edges of f meeting at V and $\partial_{e_1,V}$, $\partial_{e_2,V}$ represent derivatives along these directions (taking V as the common origin). Note that this requires $2\lfloor p_f/4\rfloor \geq \max_{e\in\mathcal{E}(f)}\lfloor p_e/4\rfloor$ and thus $p_f \geq 16$.

(iii) (element operators) Assume that $\pi_{\partial\widehat{K}} u$ is a continuous, piecewise polynomial (of degree $\leq p_K$) approximation on $\partial\widehat{K}$. Define $(\pi_K^p u) \in \mathcal{P}_{p_K}$ as the unique minimizer of

$$v \mapsto p_f^4 \sum_{j=0}^{4} p_f^{-j} |u - v|_{H^j(\widehat{K})}$$

under the constraint that $(\pi_K^p u)|_{\partial\widehat{K}} = \pi_{\partial\widehat{K}} u$.

Theorem B.2. *Consider the reference tetrahedron $\widehat{K}$. Fix $p_{ref} \in \mathbb{N}$. Let the element degree p_K, the face degrees p_f and the edge degrees p_e satisfy the "minimium rule":*

$$\begin{aligned} &1 \leq p_e \leq p_f && \forall f \in \mathcal{F}, \quad \forall e \in \mathcal{E}(f), \\ &1 \leq p_f \leq p_K && \forall f \in \mathcal{F}. \end{aligned}$$

Assume the polynomial degrees are comparable*, i.e., there is C_{comp} such that*

$$\begin{aligned} &1 \leq p_e \leq p_f \leq C_{comp} p_e && \forall f \in \mathcal{F}, \quad \forall e \in \mathcal{E}(f), \\ &1 \leq p_f \leq p_K \leq C_{comp} p_f && \forall f \in \mathcal{F}. \end{aligned}$$

Define

$$p := \min\{p_e \,|\, e \in \mathcal{E}\}.$$

(Note that the minimum rule implies additionally $p \leq p_f$ for all $f \in \mathcal{F}$ and a fortiori *$p \leq p_K$.)*

Define the approximation operator $\pi : C^\infty(\overline{\widehat{K}}) \to \mathcal{P}_{p_K}$ as follows:

1. *(vertices) Require $(\pi u)(V) = u(V)$ for all vertices $V \in \mathcal{V}$.*
2. *(edge) Let $e \in \mathcal{E}$. If $\lfloor p_e/4\rfloor \geq p_{ref}$, then set $(\pi u)|_e := \pi_e^p u$. Else, set $(\pi u)|_e = \pi_e^h u$.*
3. *(face) Let $f \in \mathcal{F}$. By step 2, $(\pi u)|_{\partial f}$ is fixed. If $(\pi u)|_{\partial f}$ is given by $\pi_e^p u$ for all three edges $e \in \mathcal{E}(f)$, then set $(\pi u)|_f := \pi_f^p u$. Else, set $(\pi u)|_f = \pi_f^h u$.*
4. *(element) The last three steps have fixed $(\pi u)|_{\partial\widehat{K}}$. Set $\pi u := \pi_K^p u$.*

Then:

(i) (approximation property) For each $s > 5$ there is $C_s > 0$ such that

$$\sum_{j=0}^{2} p^{2-j} \|u - \pi u\|_{H^j(\widehat{K})} \leq C_s p^{-(s-2)} \|u\|_{H^s(\widehat{K})} \qquad \forall u \in H^s(\widehat{K}). \tag{B.1}$$

(ii) (polynomial reproduction)

$$\pi u = u \qquad \forall u \in \mathcal{P}_{\lfloor p/4 \rfloor} \qquad\qquad \textit{if } p \geq 4,$$
$$\pi u = u \qquad \forall u \in \mathcal{P}_{p} \qquad\qquad \textit{if } p_e < 4p_{ref} \textit{ for } \text{all } \textit{edges } e \in \mathcal{E}.$$

(iii) (locality)

- *in each vertex $V \in \mathcal{V}$, πu is completely determined by $u|_V$.*
- *on each edge $e \in \mathcal{E}$, $(\pi u)|_e$ is completely determined by $u|_e$, p_e, and p_{ref}.*
- *on each face $f \in \mathcal{F}$, $(\pi u)|_f$ is completely determined by $u|_f$, p_f, the polynomial degrees p_e, $e \in \mathcal{E}(f)$, and p_{ref}.*

Proof. Proof of (iii): This follows by construction.

Proof of (ii): If $p \geq 4$, then inspection of the construction shows that $\pi u = u$ for all polynomials of degree $\lfloor p/4 \rfloor$. If $p_e < 4p_{ref}$ for *all* edges $e \in \mathcal{E}$, then $(\pi u)|_e = \pi_e^h u = u|_e$ for all polynomials u of degree p. Since πu is of the form $\pi_e^h u$ on all edges, the face values $(\pi u)|_f$ are also given by $\pi_f^h u$. Hence, polynomials of degree p are reproduced on all faces and thus also on the element.

Proof of (i): As a first step, we reduce the question to the case that

$$\lfloor p_e/4 \rfloor \geq p_{ref} \qquad \forall e \in \mathcal{E}, \qquad \lfloor p_f/2 \rfloor \geq p_{ref} \qquad \forall f \in \mathcal{F}.$$

In the converse case, one of the edge polynomials $p_{e'}$ satisfies $p_{e'} \leq 4(p_{ref}+1)$ or one face f' satisfies $p_{f'} \leq 2(p_{pref}+1)$. In view of the comparability of the degrees, this implies that

$$\begin{aligned}
p_e &\leq 4C_{comp}(p_{ref}+1) \qquad \forall e \in \mathcal{E}, \\
p_f &\leq 4C_{comp}(p_{ref}+1) \qquad \forall f \in \mathcal{F}, \\
p_K &\leq 4C_{comp}^2(p_{ref}+1).
\end{aligned}$$

In other words: the polynomial degrees are bounded and therefore only finitely many cases for the operator π can arise. By norm equivalence on finite dimensional space, the bound (B.1) holds.

We may now assume $\lfloor p_e/4 \rfloor \geq p_{ref}$ for all edges $e \in \mathcal{E}$ and $\lfloor p_f/2 \rfloor \geq p_{ref}$ for all faces. Recall that $p = \min\{p_e \,|\, e \in \mathcal{E}\}$ and that by our assumption of the "minimum rule" we therefore have $p \leq p_f \leq p_K$. Define $\widetilde{p} := \lfloor p/4 \rfloor$. We may assume $\widetilde{p} \geq 1$. Then we can proceed as in the proof of [MPS13, Thm. B.3] with $\widetilde{p}$ taking the role of p in the proof of [MPS13, Thm. B.3], from where the condition $s > 5$ arises. This leads to (B.1). $\square$

We are now in position to prove Corollary 3.4.

Proof. (of Corollary 3.4) We only consider the case $d = 3$. Let $s > \max\{5, r_{max}\}$, with r_{max} of the statement of corollary 3.4. Let ε be defined by Lemma 3.1. Then, the smoothed function $\mathcal{I}_\varepsilon u$ satisfies by the same reasoning as in the proof of Theorem 3.3 for $0 \leq r \leq q \leq s$ the bound

$$\|\mathcal{I}_\varepsilon u\|_{q,p,K} \leq C \left(\frac{h_K}{p_K}\right)^{r-q} \|u\|_{r,p,\omega_K} \qquad \forall K \in \mathcal{T}. \tag{B.2}$$

Next, we wish to employ the operator of Theorem B.2. To that end, we associate with each edge e and each face f of the triangulation $\mathcal{T}$ a polynomial degree by the "minimum rule", i.e.,

$$p_e := \min\{p_K \,|\, e \text{ is an edge of } K\}, \qquad p_f := \min\{p_K \,|\, f \text{ is a face of } K\}.$$

The definition of the edge and face polynomial degrees implies

$$p_f \geq p_e \qquad \forall \text{ edges } e \text{ of a face } f,$$
$$p_K \geq p_f \qquad \forall \text{ faces } f \text{ of an element } K,$$

which is the "minimum rule" required in Theorem B.2. Fix an element K and let $p := \min\{p_e \,|\, e \text{ is an edge of } K\}$. The γ-shape regularity of the mesh and the polynomial degree distribution implies the existence of C_{comp} such that (independent of K)

$$p_{max} := \max\{p_e \,|\, e \text{ is an edge of } K\} \leq C_{comp} p = C_{comp} \min\{p_e \,|\, e \text{ is an edge of } K\}. \tag{B.3}$$

We recognize that the definition of $\widehat{p}_K$ in the statement of Corollary 3.4 coincides with p:

$$\widehat{p}_K = p. \tag{B.4}$$

Fix $p_{ref} \in \mathbb{N}$ so that

$$p_{ref} \geq C_{comp} r_{max}. \tag{B.5}$$

The approximation operator π of Theorem B.2 satisfies on the reference element $\widehat{K}$

$$\sum_{j=0}^{2} p^{2-j} |v - \pi v|_{H^j(\widehat{K})} \leq C p^{-(s-2)} \|v\|_{H^s(\widehat{K})}. \tag{B.6}$$

In order to get the correct powers of h_K, we exploit that π reproduces polynomials by Theorem B.2. We consider two cases:
1. Case: $p_{max} < 4p_{ref}$. In this case, Theorem B.2 implies immediately that $\pi v = v$ for all $v \in \mathcal{P}_p$.
2. Case: $p_{max} \geq 4p_{ref}$. In this case,

$$C_{comp} p \geq p_{max} \geq 4p_{ref} \geq 4C_{comp} r_{max}$$

so that $p/4 \geq r_{max}$. Hence, Theorem B.2 implies $\pi v = v$ for all $v \in \mathcal{P}_{r_{max}}$.

Combining the two cases yields

$$\pi v = v \qquad \forall v \in \mathcal{P}_{\min\{p, r_{max}\}}. \tag{B.7}$$

Hence combining (B.2), (B.6), and (B.7) together with the usual scaling arguments and the Bramble-Hilbert lemma 4.3 leads to

$$\begin{aligned}
&\sum_{j=0}^{2} h_K^j p^{-j} |\mathcal{I}_\varepsilon u - \pi \mathcal{I}_\varepsilon u|_{H^j(K)} \leq C p^{-s} \sum_{q=1+\min\{p,r_{max}\}}^{s} h_K^q |\mathcal{I}_\varepsilon u|_{H^q(K)} \\
&\leq C p^{-s} \left[\sum_{q=1+\min\{p,r_{max}\}}^{r} h_K^q |\mathcal{I}_\varepsilon u|_{H^q(K)} + \sum_{q=r+1}^{s} h_K^q |\mathcal{I}_\varepsilon u|_{H^q(K)} \right] \\
&\leq C p^{-s} h_K^{1+\min\{p,r_{max}\}} \|\mathcal{I}_\varepsilon u\|_{H^r(K)} + C p^{-s} \sum_{q=r+1}^{s} h_K^q |\mathcal{I}_\varepsilon u|_{H^q(K)} \\
&\leq C p^{-s} h_K^{1+\min\{p,r_{max}\}} \|\mathcal{I}_\varepsilon u\|_{H^r(K)} + C p^{-s} \sum_{q=r+1}^{s} h_K^q \left(\frac{h_K}{p_K} \right)^{r-q} \|u\|_{r,2,\omega_K} \\
&\leq C \left[\frac{h_K^r}{p_K^r} + p_K^{-s} h_K^{1+\min\{p,r\}} \right] \|u\|_{r,2,\omega_K} \overset{r \leq r_{max} \leq s}{\leq} C p_K^{-r} \left[h_K^r + h_K^{1+\min\{p_K,r\}} \right] \|u\|_{r,2,\omega_K},
\end{aligned}$$

which proves the desired result. $\square$

References

[Aco01] Gabriel Acosta. Lagrange and average interpolation over 3D anisotropic elements. *J. Comput. Appl. Math.*, 135(1):91–109, 2001.

[Ada75] Robert A. Adams. *Sobolev spaces.* Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[AF03] Robert A. Adams and John J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam).* Elsevier/Academic Press, Amsterdam, second edition, 2003.

[AK99] Mark Ainsworth and David Kay. The approximation theory for the p-version finite element method and application to non-linear elliptic PDEs. *Numer. Math.*, 82(3):351–388, 1999.

[Ape99] Thomas Apel. Interpolation of non-smooth functions on anisotropic finite element meshes. *M2AN*, 33(6):1149–1185, 1999.

[BCMP91] I. Babuška, A. Craig, J. Mandel, and J. Pitkäranta. Efficient preconditioning for the p-version finite element method in two dimensions. *SIAM J. Numer. Anal.*, 28(3):624–661, 1991.

[BDM92] C. Bernardi, M. Dauge, and Y. Maday. Trace liftings which preserve polynomials. *C.R. Acad. Sci. Paris, Série I*, 315:333–338, 1992.

[BDM99] C. Bernardi, M. Dauge, and Y. Maday. *Spectral Methods for Axisymmetric domains.* North Holland/Gauthier-Villars, 1999.

[BDM07] C. Bernardi, M. Dauge, and Y. Maday. Polynomials in the Sobolev world (version 2). Technical Report 14, IRMAR, 2007.

[BG98] C. Bernardi and V. Girault. A local regularization operator for triangular and quadrilateral finite elements. *SIAM J. Numer. Anal.*, 35(5):1893–1916 (electronic), 1998.

[BM92] C. Bernardi and Y. Maday. *Approximations spectrales de problèmes aux limites elliptiques.* Mathématiques & Applications. Springer Verlag, 1992.

[BM97] C. Bernardi and Y. Maday. Spectral methods. In P.G. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, Vol. 5.* North Holland, Amsterdam, 1997.

[BS87] I. Babuška and Manil Suri. The optimal convergence rate of the p-version of the finite element method. *SIAM J. Numer. Anal.*, 24(4):750–776, 1987.

[BS08] Susanne C. Brenner and L. Ridgway Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics.* Springer, New York, third edition, 2008.

[Bur98] Victor I. Burenkov. *Sobolev spaces on domains*, volume 137 of *Teubner-Texte zur Mathematik [Teubner Texts in Mathematics].* B. G. Teubner Verlagsgesellschaft mbH, Stuttgart, 1998.

[Car97] Carsten Carstensen. An a posteriori error estimate for a first-kind integral equation. *Math. Comp.*, 66(217):139–155, 1997.

[Car99] Carsten Carstensen. Quasi-interpolation and a posteriori error analysis in finite element methods. *M2AN Math. Model. Numer. Anal.*, 33(6):1187–1202, 1999.

[CD05] W. Cao and L. Demkowicz. Optimal error estimate of a projection based interpolation for the p-version approximation in three dimensions. *Comput. Math. Appl.*, 50(3-4):359–366, 2005.

[CFS96] Carsten Carstensen, Stefan A. Funken, and Ernst P. Stephan. A posteriori error estimates for hp-boundary element methods. *Appl. Anal.*, 61(3-4):233–253, 1996.

[CH09] C. Carstensen and Jun Hu. Hanging nodes in the unifying theory of a posteriori finite element error control. *J. Comput. Math.*, 27(2-3):215–236, 2009.

[CHQZ06] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral methods*. Scientific Computation. Springer-Verlag, Berlin, 2006. Fundamentals in single domains.

[Cia76] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland Publishing Company, 1976.

[Clé75] Ph. Clément. Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique*, 9(R-2):77–84, 1975.

[CMPS04] Carsten Carstensen, M. Maischak, D. Praetorius, and E. P. Stephan. Residual-based a posteriori error estimate for hypersingular equation on surfaces. *Numer. Math.*, 97(3):397–425, 2004.

[CMS01] C. Carstensen, M. Maischak, and E. P. Stephan. A posteriori error estimate and h-adaptive algorithm on surfaces for Symm's integral equation. *Numer. Math.*, 90(2):197–213, 2001.

[Cos88] Martin Costabel. Boundary integral operators on Lipschitz domains: Elementary results. *SIAM J. Math. Anal.*, 19:613–626, 1988.

[CQ82] C. Canuto and A. Quarteroni. Approximation results for orthogonal polynomials in Sobolev spaces. *Math. Comp.*, 38(257):67–86, 1982.

[CS95] Carsten Carstensen and Ernst P. Stephan. A posteriori error estimates for boundary element methods. *Math. Comp.*, 64(210):483–500, 1995.

[CS96] Carsten Carstensen and Ernst P. Stephan. Adaptive boundary element methods for some first kind integral equations. *SIAM J. Numer. Anal.*, 33(6):2166–2183, 1996.

[DB05] L. Demkowicz and A. Buffa. H^1, $H(\mathrm{curl})$ and $H(\mathrm{div})$-conforming projection-based interpolation in three dimensions. Quasi-optimal p-interpolation estimates. *Comput. Methods Appl. Mech. Engrg.*, 194(2-5):267–296, 2005.

[DL93] Ronald A. DeVore and George G. Lorentz. *Constructive approximation*, volume 303 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993.

[DS80] T. Dupont and L.R. Scott. Polynomial approximation of functions in Sobolev spaces. *Math. Comp.*, 34:441–463, 1980.

[EG92a] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[EG92b] L.C. Evans and R.F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.

[Fae02] Birgit Faermann. Localization of the Aronszajn-Slobodeckij norm and application to adaptive boundary element methods. II. The three-dimensional case. *Numer. Math.*, 92(3):467–499, 2002.

[Gag57] E. Gagliardo. Caratterizzazione delle tracce sulla frontiera relative ad alcune classi di funzioni in n variabili. *Rend. Sem. Mat. Univ. Padova*, 27:284–305, 1957.

[GS02] V. Girault and L. R. Scott. Hermite interpolation of nonsmooth functions preserving boundary conditions. *Math. Comp.*, 71(239):1043–1074 (electronic), 2002.

[Guo06] Benqi Guo. Approximation theory for the p-version of the finite element method in three dimensions. I. Approximabilities of singular functions in the framework of the Jacobi-weighted Besov and Sobolev spaces. *SIAM J. Numer. Anal.*, 44(1):246–269 (electronic), 2006.

[Heu14] Norbert Heuer. On the equivalence of fractional-order Sobolev semi-norms. *J. Math. Anal. Appl.*, 417(2):505–518, 2014.

[HW08] George C. Hsiao and Wolfgang L. Wendland. *Boundary integral equations*, volume 164 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 2008.

[McL00] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, 2000.

[Mel05] J. M. Melenk. hp-interpolation of nonsmooth functions and an application to hp-a posteriori error estimation. *SIAM J. Numer. Anal.*, 43(1):127–155, 2005.

[MPS13] J. M. Melenk, A. Parsania, and S. Sauter. General DG-methods for highly indefinite Helmholtz problems. *J. Sci. Comput.*, 57(3):536–581, 2013.

[MS97] Rafael Muñoz-Sola. Polynomial liftings on a tetrahedron and applications to the h-p version of the finite element method in three dimensions. *SIAM J. Numer. Anal.*, 34(1):282–314, 1997.

[Mur67] Tosinobu Muramatu. A proof for the imbedding theorems for Sobolev spaces. *Proc. Japan Acad.*, 43:87–92, 1967.

[Mur68] Tosinobu Muramatu. On imbedding theorems for Sobolev spaces and some of their generalizations. *Publ. Res. Inst. Math. Sci. Ser. A*, 3:393–416, 1967/1968.

[MW14] J.M. Melenk and T. Wihler. A posteriori error analysis of hp-fem for singularly perturbed problems reaction-diffusion equations. Technical report, 2014.

[Néd88] J.-C. Nédélec. Approximation of integral equations with nonintegrable kernels. In *The mathematics of finite elements and applications, VI (Uxbridge, 1987)*, pages 343–352. Academic Press, London, 1988.

[Qua84] Alfio Quarteroni. Some results of Bernstein and Jackson type for polynomial approximation in L^p-spaces. *Japan J. Appl. Math.*, 1(1):173–181, 1984.

[Ran86] E. Rank. Adaptivity and accuracy estimation for finite element and boundary integral element methods. In I. Babuška, O. C. Zienkiewicz, J. Gago, and E. R. de A. Oliveira, editors, *Accuracy Estimates and Adaptive Refinements in Finite Element Computations*. John Wiley & Sons, 1986.

[Ran89] Ernst Rank. Adaptive h-, p- and hp-versions for boundary integral element methods. *Internat. J. Numer. Methods Engrg.*, 28(6):1335–1349, 1989.

[Ran93] E. Rank. Adaptivity and accuracy estimation for FEM and BIEM. In *Adaptive finite and boundary element methods*, Internat. Ser. Comput. Engrg., pages 1–46. Comput. Mech., Southampton, 1993.

[Ran12] Alexander Rand. Average interpolation under the maximum angle condition. *SIAM J. Numer. Anal.*, 50(5):2538–2559, 2012.

[RS75] Michael Reed and Barry Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[Sha85] V. V. Shan′kov. The averaging operator with variable radius, and the inverse trace theorem. *Sibirsk. Mat. Zh.*, 26(6):141–152, 191, 1985.

[SS11] Stefan A. Sauter and Christoph Schwab. *Boundary element methods*, volume 39 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2011. Translated and expanded from the 2004 German original.

[Ste70] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[Ste08] Olaf Steinbach. *Numerical approximation methods for elliptic boundary value problems*. Springer, New York, 2008. Finite and boundary elements, Translated from the 2003 German original.

[SZ90] L. Ridgway Scott and Shangyou Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[Ver99] Rüdiger Verfürth. Error estimates for some quasi-interpolation operators. *M2AN Math. Model. Numer. Anal.*, 33(4):695–713, 1999.