Energy norm error estimates for averaged discontinuous Galerkin methods: multidimensional case

Ferenc Izsák

Department of Applied Analysis and Computational Mathematics & MTA-ELTE NumNet Research Group, Eötvös Loránd University, Pázmány P. stny. 1C, 1117 - Budapest, Hungary

Abstract

A mathematical analysis is presented for a class of interior penalty (IP) discontinuous Galerkin approximations of elliptic boundary value problems. In the framework of the present theory one can derive some overpenalized IP bilinear forms avoiding the notion of fluxes and artificial penalty terms. The main idea is to start from bilinear forms for the local average of discontinuous approximations which are rewritten using the theory of distributions. It is pointed out that a class of overpenalized IP bilinear forms can be obtained using a perturbation of these. Also, error estimations can be derived between the local averages of the discontinuous approximations and the analytic solution in the H^1-seminorm. Using the local averages, the analysis is performed in a conforming framework without any assumption on extra smoothness for the solution of the original boundary value problem.

Keywords: discontinuous Galerkin methods, elliptic boundary value problems, error estimation

2010 MSC: 65N12, 65N15, 65N30

1. Introduction

Discontinuous Galerkin (dG) methods have been introduced and used from the end of the seventies, first for linear transport problems. Later similar methods were constructed for elliptic boundary value problems [1] and nowadays it is available for the numerical solution of almost all kind of problems based on PDE's.

These methods have proved their usefulness in several simulations of real-life phenomena [2], [3], [4]. The most favorable property of the corresponding numerical methods is that the local mesh refinement can be easily performed giving rise to efficient adaptive strategies.

An important milestone in the systematic analysis of dG methods for elliptic boundary value problems was the

Email address: izsakf@cs.elte.hu (Ferenc Izsák)

paper [5]. This pioneering work served as a basis of the consecutive works concerning a priori and a posteriori error estimates [6], *hp*-adaptive methods [7], time dependent problems [8]. For an up-to-date summary of the theoretical achievements for dG methods we refer to the recent monograph [9] and for implementation issues to the monographs [10] and [11].

At the same time, the above analysis should be improved in some aspects. First, which can be considered as a didactic issue, the choice of the corresponding bilinear forms would deserve more motivation. After recasting the elliptic problem in a mixed form, numerical fluxes and penalty terms are defined which lead to different bilinear forms. No a priori suggestion or motivation is mentioned to propose an appropriate choice of the fluxes. Additionally, penalty terms contain user-defined constants, which should be tuned to ensure the ellipticity of the corresponding bilinear form. At the same time, too large constant would grow the condition number of the stiffness matrix in the computations. On the optimal choice for elliptic problems we refer to [12] and for Maxwell equations to [13]. A possible remedy of these problems is to use overpenalization [14], [15], which results inheritly elliptic bilinear forms. Also, as pointed out in [15] and [16], one can use appropriate preconditioners to deal with large condition numbers.

The second issue to be addressed is the assumption on extra-regularity of the analytic solution in the original analysis. This problem was solved in the meantime: in [17] the author developed an analysis based on a Strang type lemma [18], which could successfully deal with the non-conformity of the dG type approximation.

The most important issue is the norm for the convergence. The choice of the bilinear form implies a mesh-dependent norm, which is a real mathematical artifact. The convergence is proved with respect to this norm or in a weaker, *e.g.*, in the L_2-norm. At the same time, in the corresponding real-life problems the natural norm is usually the H^1-norm (or seminorm). Note that there are some achievements which point out the usefulness of the interior penalty (IP) methods. For these methods, one can obtain convergence in the so-called BV norm which does not depend on the actual mesh [19], [20] and can be related to broken Sobolev norms.

Based on the above issues, the aim of the present work is to contribute to the *error analysis* of some existing dG methods for elliptic boundary value problems by proposing an alternative of the commonly used theoretical basis in [5]. In particular, we derive overpenalized interior penalty bilinear forms avoiding the notion of numerical fluxes or recasting them into a mixed form. The new idea is to use the local average of the discontinuous approximation from the beginning. The only heuristic detail in our analysis is the choice of the local average. The main benefit of the analysis is that it can be done in an H^1-conforming framework such that one can prove the quasi optimal convergence of the local average with respect to the natural H^1-seminorm for Dirichlet problems. This work is a generalization of the paper [21] concerning the one-dimensional case. For numerical experiments demonstrating the performance and usefulness of overpenalized dG methods we refer to [14], [15] and [16].

The idea to use *post*processing (or smoothing or filtering) for dG approximations has already appeared in the literature [22]. In the last years, many related results have been achieved: involved algorithms were developed for linear hyperbolic problems in [23] and their accuracy-increasing property was verified also for advection-diffusion problems with respect to negative Sobolev norms [24]. The accurate computation of the corresponding convolutions is challenging, see the recent developments in [25] and [26].

The setup of the article is as follows. We first introduce the tools for our analysis. Then the main results are presented along with the main consequences: first, the derivation of the overpenalized IP methods as an alternative of the conventional one in [5] and second, the practice of the proposed method is discussed examining its computational costs. Finally, the detailed proofs are given without the most technical details, which are placed in the Appendix.

2. Mathematical preliminaries

We investigate the finite element solution of the elliptic boundary value problem

$$\begin{cases} -\Delta u(\mathbf{x}) = g(\mathbf{x}) \quad \mathbf{x} \in \Omega \subset \mathbb{R}^d \\ u(\mathbf{x}) = 0 \quad \mathbf{x} \in \partial\Omega, \end{cases} \tag{1}$$

where Ω is a polyhedral Lipschitz-domain and $g \in L_2(\Omega)$ is given.

Discontinuous finte element spaces. The finite element approximation is computed on a simplicial mesh $\mathcal{T}_{\mathtt{h}}$. The symbol $\mathcal{F}$ denotes the set of interelement faces. Using the notation $\rho(K)$ for the radius of the maximal inscribed ball into K, we introduce the mesh parameter $h \leq 1$ with

$$h = \min\{\rho(K)/2 : K \in \mathcal{T}_{\mathtt{h}}\}.$$

We also use the notations h_K, h_f and h_{Ω_j} for the diameter of the corresponding subdomain or face. The error analysis can be applied in consecutive refinement steps (also in an hp-adaptive procedure) if the corresponding family $\{\mathcal{T}_{\mathtt{h}}\}_{\mathtt{h}\in\mathcal{H}}$ of meshes is non-degenerated, i.e. there exists a parameter N_* such that for all $K \in \{\mathcal{T}_{\mathtt{h}}\}_{\mathtt{h}\in\mathcal{H}}$ we have

$$\operatorname{diam} K \leq N_* \rho(K). \tag{2}$$

Also, hanging nodes are allowed provided that their number is bounded in the refinement steps. At the same time, the constants in the estimates may depend on the maximal polynomial degree and the number of hanging nodes on the faces. To reduce the complexity of the forthcoming analysis, we do not track this dependence. For the numerical solution we use the finite element space

$$\mathbb{P}_{h,\mathbf{k}} = \{u \in L_2(\Omega) : u|_K \in P_{k_j}(\Omega_j) \text{ for all } \Omega_j \in \mathcal{T}_{\mathtt{h}}\},$$

where $\mathbf{k} = (k_1, k_2, \dots)$ and $P_{k_j}(\Omega_j)$ denotes the linear space of polynomials of total degree k_j on the subdomain Ω_j. This notation will also be used for interelement faces and for balls instead of Ω_j.

Jumps and averages. We introduce the notation $\boldsymbol{\nu}_j$ for the outward normal of Ω_j with any symbol j. We also make use of the conventional notation $\{\!\{\cdot\}\!\} : \mathbb{P}_{h,\mathbf{k}} \to L_2(\mathcal{F})$ and $[\![\cdot]\!] : \mathbb{P}_{h,\mathbf{k}} \to L_2(\mathcal{F})$ for the average and jump operators which are given on each interelement face $f_\Omega = \bar{\Omega}_+ \cap \bar{\Omega}_-$ with

$$\{\!\{v\}\!\}_{f_\Omega}(\mathbf{x}) = \frac{1}{2}(v(\mathbf{x}_+) + v(\mathbf{x}_-)) \quad \text{and} \quad [\![v]\!]_{f_\Omega}(\mathbf{x}) = \boldsymbol{\nu}_+ v(\mathbf{x}_+) + \boldsymbol{\nu}_- v(\mathbf{x}_-),$$

where $v(\mathbf{x}_\pm) = \lim_{\Omega_\pm \supset \mathbf{x}_n \to \mathbf{x}} v(\mathbf{x}_n)$. On each boundary face $f \subset \partial\Omega$ we simply define

$$\{\!\{v\}\!\}_f(\mathbf{x}) = v(\mathbf{x}) \quad \text{and} \quad [\![v]\!]_f(\mathbf{x}) = \boldsymbol{\nu}(\mathbf{x}) v(\mathbf{x}).$$

Norms, scalar products and bilinear forms. The $L_2(\Omega^*)$-norm on a generic domain Ω^* will be denoted with $\|\cdot\|_{\Omega^*}$ and the corresponding scalar product with $(\cdot,\cdot)_{\Omega^*}$. In case of $\Omega^* = \Omega$ or if the support of the terms is given, we omit the subscript. Similar notation is applied for the scalar product and the corresponding L_2 norm on $\mathcal{F}$ and on a single interelement face f.

With these, the most popular dG approximation of u in (1) is the so-called symmetric interior penalty dG method which is given with the bilinear form $a_{\mathrm{IP}} : \mathbb{P}_{\mathbf{h},k} \times \mathbb{P}_{\mathbf{h},k} \to \mathbb{R}$ as follows:

$$a_{\mathrm{IP}}(u,v) = (\nabla_h u, \nabla_h v) - \sum_{f\in\mathcal{F}} (\{\!\{\nabla_h u\}\!\}, [\![v]\!])_f + (\{\!\{\nabla_h v\}\!\}, [\![u]\!])_f + \sum_{f\in\mathcal{F}} \sigma_h([\![u]\!], [\![v]\!])_f, \tag{3}$$

where ∇_h denotes the piecewise gradient on the subdomains in $\mathcal{T}_h$ and $\sigma_h \in \mathbb{R}$ denotes a penalty parameter, which is proportional with $(\operatorname{diam} f)^{-1}$ in the conventional setting and with $(\operatorname{diam} f)^{-r}$ in case of overpenalized methods, see [14] and [15]. The notation $\nabla_f [\![u]\!]$ will be used for the gradient of the jump functions defined on the interelement face f: i.e. for each $\boldsymbol{\tau}_1$ unit (tangential) vector $\boldsymbol{\tau}_1$ contained in f and for any $\mathbf{x} \in f$

$$\boldsymbol{\tau}_1 \cdot \nabla_f [\![u]\!](\mathbf{x}) := \lim_{\delta\to 0} \frac{[\![u]\!](\mathbf{x} + \delta\boldsymbol{\tau}_1) - [\![u]\!](\mathbf{x})}{\delta}.$$

The notation $\lambda_d(\cdot)$ will be used to the d-dimensional Lebesque measure.

Averaging, conforming finite element space. For the local average we use the piecewise constant function $\eta_h : \mathbb{R}^n \to \mathbb{R}$ depending also on the parameter $s > 1$ with

$$\eta_h(\mathbf{x}) = \begin{cases} \frac{1}{B_{h^s,d}} & |\mathbf{x}| \le h^s \\ 0 & |\mathbf{x}| > h^s, \end{cases}$$

where $B(\mathbf{x}, r)$ denotes the closed ball with radius r centered at $\mathbf{x}$ and $B_{h^s,d} = \lambda_d(B(\mathbf{0}, h^s))$. The analysis makes use only two properties of η_h: this is symmetric with respect to the origin and $\int_{\mathbb{R}^d} \eta_h = 1$ such that $\eta_h * u$ is the local average of the function $u : \mathbb{R}^d \to \mathbb{R}$. Also, a straightforward computation gives that $\operatorname{supp} \eta_h * \eta_h = B(\mathbf{0}, 2h^s)$ and $\int_{B(\mathbf{0},2h^s)} \eta_h * \eta_h = 1$. These facts will be used without further reference.

The analysis of the conforming approach will be carried out in the space

$$\mathbb{P}_{h,\mathbf{k},s} = \{\eta_h * u_0|_{\Omega_h} : u_0 \text{ is the zero extension of } u \in \mathbb{P}_{h,\mathbf{k}}\},$$

where $\Omega_h = \{\mathbf{x} \in \mathbb{R}^d : d(\mathbf{x}, \Omega) < h^s\}$. Obviously, $\mathbb{P}_{h,\mathbf{k},s} \subset H_0^1(\Omega_h)$. We use the notation $\Omega_{j,h}$ in a similar sense and $\tilde{\Omega}_j = \operatorname{int} \{\bar{\Omega}_k \in \mathcal{T}_h : \bar{\Omega}_j \cap \bar{\Omega}_k \neq \emptyset\}$ for the patch of Ω_j, where $\operatorname{int}(\mathcal{S})$ denotes the interior of a set $\mathcal{S}$.

Reference subdomain pairs. To extend the standard scaling arguments we first define a reference set $\mathcal{K}$ of neighboring simplex pairs (K_+, K_-) having the interelement face $f = \bar{K}_+ \cap \bar{K}_-$ such that the following conditions hold:

- $f \subset 0 \times \mathbb{R}^{d-1}$ and one vertex of f is $\mathbf{0} \in \mathbb{R}^d$
- the maximum edge-length of f is *one*
- K_+ and K_- satisfy the condition on non-degeneracy, see (2).

Then for any neighboring subdomains $\Omega_+, \Omega_- \in \mathcal{T}_\mathbf{h}$ there is a pair $(K_+, K_-) \in \mathcal{K}$ and an affine linear map $A_\Omega : K_+ \cup K_- \to \Omega_+ \cup \Omega_-$ with $A_\Omega(K_+) = \Omega_+$ and $A_\Omega(K_-) = \Omega_-$, moreover

$$A_\Omega(\mathbf{x}) = A_{\Omega,0}(h_\Omega \mathbf{x}), \tag{4}$$

where h_Ω denotes the maximum edge length of f_Ω and $A_{\Omega,0}$ is an isometry; see also Fig. 1.

Accordingly, for any $v \in \mathbb{P}_{h,\mathbf{k}}(\Omega_+ \cup \Omega_-)$ we have $v_0 := v \circ A_\Omega \in \mathbb{P}_{h,\mathbf{k}}(K_+ \cup K_-)$, moreover, using (4) the following equalities are valid:

$$[\![v_0]\!](\mathbf{x}) = [\![v]\!](A_\Omega(\mathbf{x})) \quad \text{and} \quad \eta_{h_0} * v_0(\mathbf{x}) = \eta_{h_0 h^{\frac{1}{s}}} * v(A_\Omega \mathbf{x}), \tag{5}$$

whenever the operation $\eta_{h_0} *$ makes sense.

We also use the notation $h_\Omega \cdot K_\pm = \{h_\Omega \mathbf{x} : \mathbf{x} \in K_\pm\}$ and similarly $h_\Omega \cdot f$ and introduce the interior domain $\Omega_{j0} = \{\mathbf{x} \in \Omega_j : B(\mathbf{x}, h^s) \subset \Omega_j\}$ and the interior face $f_0 \subset f$ similarly. The definition of h ensures that these are non-empty.

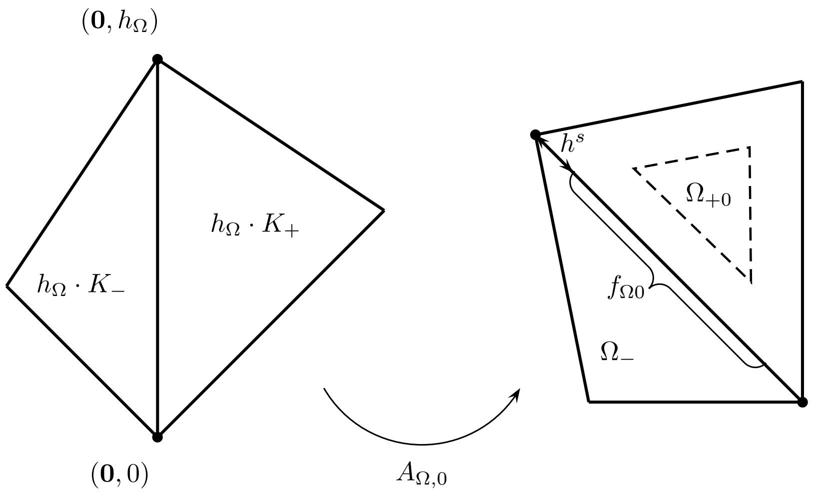

Figure 1: The transformation of the reference subdomain pair, the interior domain Ω_{+0} and the interior face $f_{\Omega 0}$ in the 2-dimensional case.

BV spaces and distributions. The space $\mathrm{BV}(\Omega)$ of real valued functions on Ω with bounded variations is defined with

$$\mathrm{BV}(\Omega) = \left\{ u : \Omega \to \mathbb{R} : \sup_{\substack{\boldsymbol{\phi} \in [C_c^1(\Omega)]^d \\ \|\boldsymbol{\phi}\|_\infty = 1}} \int_\Omega u \,\mathrm{div}\, \boldsymbol{\phi} := |u|_{\mathrm{BV}} < \infty \right\}$$

and is equipped with the seminorm $|\cdot|_{\mathrm{BV}}$, where $\|\cdot\|_\infty$ denotes the maximum norm on $C_c^1(\Omega)$. This seminorm can also be given as

$$|\cdot|_{\mathrm{BV}} = \int_\Omega \mathrm{d}|\partial u|,$$

where $|\partial u|$ is the Radon measure generated by the distributional derivative of u. The dual pairing between a distribution S and a test function ϕ is denoted using angle brackets: $\langle S, \phi \rangle$.

In the estimates, the notation $g_1 \lesssim g_2$ means the existence of a constant C - which does not depend on the mesh parameter but possibly on the local polynomial degree - such that $g_1 \leq C \cdot g_2$. We also use the notation $g_1 \sim g_2$ provided that both $g_1 \lesssim g_2$ and $g_2 \lesssim g_1$ are satisfied.

3. Results and discussion

The basic idea of the present *analysis* is to look for a smoothed dG approximation immediately in the bilinear form. In this case, in the background we can compute with discontinuous basis functions in $\mathbb{P}_{h,\mathbf{k}}$ while still having the freedom to choose them independently on the neighboring subdomains. On the other hand, as we compute conforming approximations, we can use the entire armory of the classical finite element analysis.

The smoothed (or averaged) dG approximation consists of finding $\eta_h * u_h \in \mathbb{P}_{h,\mathbf{k},s}$ such that for all $\eta_h * v_h \in \mathbb{P}_{h,\mathbf{k},s}$ we have

$$a_\eta(u_h, v_h) := a_\eta^+(\eta_h * u_h, \eta_h * v_h) := (\nabla(\eta_h * u_h), \nabla(\eta_h * v_h)) = (g_0, \eta_h * v_h), \tag{6}$$

where the bilinear forms $a_\eta : \mathbb{P}_{h,\mathbf{k}} \times \mathbb{P}_{h,\mathbf{k}} \to \mathbb{R}$ and $a_\eta^+ : \mathbb{P}_{h,\mathbf{k},s} \times \mathbb{P}_{h,\mathbf{k},s} \to \mathbb{R}$ are defined by (6) and g_0 denotes the zero extension of g to Ω_h. Whenever the spaces $\mathbb{P}_{h,\mathbf{k},s} \not\subset H_0^1(\Omega)$ we call the method H^1-conforming since each space is in $H_0^1(\Omega_h)$.

To prove the first main result we introduce the modified IP bilinear form $a_{\mathrm{IP},s} : \mathbb{P}_{h,\mathbf{k}} \times \mathbb{P}_{h,\mathbf{k}} \to \mathbb{R}$ with

$$a_{\mathrm{IP,s}}(u, v) = (\nabla_h u, \nabla_h v) - \sum_{f\in\mathcal{F}} (\{\!\{\nabla_h u\}\!\}, [\![v]\!])_f + (\{\!\{\nabla_h v\}\!\}, [\![u]\!])_f + \sum_{f\in\mathcal{F}} \sigma_{s,h}([\![u]\!], [\![v]\!])_f, \tag{7}$$

where

$$\sigma_{s,h}([\![u]\!], [\![v]\!])_f = \begin{cases} \frac{16}{3\pi^2} h^{-s}([\![u]\!], [\![v]\!])_f & \text{for } d = 2 \\ \frac{3}{5} h^{-s}([\![u]\!], [\![v]\!])_f & \text{for } d = 3 \end{cases} \tag{8}$$

and the corresponding finite element approximation $u_{\mathrm{IP},s}$ for which

$$a_{\mathrm{IP},s}(u_{\mathrm{IP},s}, v) = (g, v) \quad \forall v \in \mathbb{P}_{h,\mathbf{k}}. \tag{9}$$

Theorem 1. *Assume that $3s > d + 2$. Then the IP bilinear form in (7) is a perturbation of a_η in the sense that*

$$|a_\eta(u, v) - a_{IP,s}(u, v)| \lesssim h^{s-1}(1 + h^{3s-d-2})\|\nabla(\eta_h * u)\|\|\nabla(\eta_h * v)\|.$$

Since the bilinear form a_η is a slight modification of $a_{\mathrm{IP},s}$, we expect that the local average of the approximations of u_h and $u_{\mathrm{IP},s}$ are also close to each other. In precise terms, we have the following.

Theorem 2. *Assume that $3s > d + 2$. Then for the finite element approximations u_h and $u_{IP,s}$ we have*

$$\|\nabla(\eta_h * u_{IP,s} - \eta_h * u_h)\| \lesssim h^{s-1}\|\nabla(\eta_h * u_h)\| + \max_j h_{\Omega_j}^d \|\eta_h * g - g_0\|.$$

The main result of this work is the error estimation for the averaged IP approximation in the $H^1(\Omega)$-seminorm.

Theorem 3. *The averaged interior penalty approximation is quasi optimal with respect to the real energy seminorm in the following sense:*

$$\|\nabla(u - \eta_h * u_{IP,s})\| \lesssim \inf_{v_h\in\mathbb{P}_{h,\mathbf{k}}} \|u - \eta_h * v_h\|_1 + \mathcal{O}(h^{s-\frac{1}{2}}) + \max_j h_{\Omega_j}^d \|\eta_h * g - g_0\|.$$

The new derivation of IP bilinear forms. Based on the results of the paper, we propose the following introduction of IP methods for the numerical solution of (1).

- Introduce the H^1-conforming finite element discretization (6).
- Since the $a_{\mathrm{IP},s}$ bilinear form is a perturbation of a_η and given more explicitly, one should compute $u_{\mathrm{IP},s}$ in the practice.

Note that this approach contains one heuristic detail, namely the choice of the smoothing operator in (6). At the same time, we can completely avoid the problem of choosing some numerical flux or penalty parameter and the corresponding bilinear form is obviously elliptic.

The new computational proccedure and its costs. Using Theorem 3, we propose the following approximation of (1).

(i) Compute the solution $u_{\mathrm{IP},s}$ of the problem in (9).

(ii) Compute the local average $\eta_h * u_{\mathrm{IP},s}$.

Note that in (9) an overpenalized IP method is applied. In the classical approach [15], the penalty term is defined by

$$\eta \sum_{f\in\mathcal{F}} \frac{s}{|\mathrm{diam}\, f|^{2s+1}} \int_f \Pi_f^{s-1} [\![u]\!]\, \Pi_f^{s-1} [\![v]\!] \,,$$

where s is a positive integer and Π_f^{s-1} denotes the L_2-orthogonal projection to $P_{s-1}(f)$. In our approach,

- it is not necessary to compute orthogonal projections
- we can use any parameter s with $3s > d+2$, which results in less ill-conditioned linear systems
- the constant η is given in (8).

The only extra cost compared with a conventional method is the computation of the local average. The most straightforward way to compute it is to apply Gauss quadrature on a ball. Another possibility is to take a convolution operator with the characteristic function of a ball. To approximate it one can make of use of the `convn.m` subroutine of MATLAB. Note that the efficient computation of more involved smoothing procedures have been already developed, see [25] and [26].

4. Analysis and proofs

We make use of the following inequalities, which can be proved using simple scaling arguments and the assumption that the family $\{\mathcal{T}_{\mathrm{h}}\}_{\mathrm{h}\in\mathcal{H}}$ of meshes is non-degenerated.

Proposition 1. *For all faces and domains in the family $\{\mathcal{T}_{\boldsymbol{h}}\}_{\boldsymbol{h}\in\mathcal{H}}$ of meshes we have the following inequalities:*

$$\max_{B(\mathbf{0},h^s)} |u| \sim h^{-\frac{sd}{2}} \|u\|_{B(\mathbf{0},h^s)} \qquad \forall u \in \mathbb{P}_k(B(\mathbf{0},h^s)), \tag{10}$$

$$\max_f [\![u]\!] \lesssim h^{1-d} \int_f |\,[\![u]\!]\,| \qquad \forall\, [\![u]\!] \in \mathbb{P}_k(f), \tag{11}$$

$$\max_K |\nabla^2 u| \lesssim h_K^{-\frac{d}{2}-2} \|u\|_K \leq h^{-\frac{d}{2}-2} \|u\|_K \qquad \forall u \in \mathbb{P}_k(K), \tag{12}$$

$$\max_f \nabla_f^2 [\![u]\!] \lesssim h_f^{-d-1} \int_f |\,[\![u]\!]\,| \lesssim h^{-d-1} \int_f |\,[\![u]\!]\,| \qquad \forall\, [\![u]\!] \in \mathbb{P}_k(f), \tag{13}$$

$$\|\nabla u\|_{B(\mathbf{0},h^s)} \lesssim h^{\frac{(s-1)d}{2}} \|\nabla u\|_{B(\mathbf{0},h)} \lesssim h^{-1} h^{\frac{(s-1)d}{2}} \|u\|_{B(\mathbf{0},h)} \qquad \forall u \in \mathbb{P}_k(B(\mathbf{0},h)). \qquad \square \tag{14}$$

We also need an estimate between the discontinuous function $\nabla_h u$ and its local average $\eta_h * \nabla_h u$ with a convergence rate depending on h. For this a Taylor expansion is developed about all $\mathbf{x} \in \Omega_{j0}$ giving for an arbitrary $\mathbf{y} \in \Omega_j$ that

$$u(\mathbf{y}) = u(\mathbf{x}) + \nabla u(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) + \frac{1}{2} \nabla^2 u(\boldsymbol{\xi}_{\mathbf{y}})(\mathbf{y} - \mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) \tag{15}$$

for some $\boldsymbol{\xi}_{\mathbf{y}}$ in the section $(\mathbf{x}, \mathbf{y})$. Integrating both sides over $B(\mathbf{x}, h^s)$ yields

$$B_{h^s,d} \cdot (\eta_h * u(\mathbf{x})) = B_{h^s,d} \cdot u(\mathbf{x}) + \int_{B(\mathbf{x},h^s)} \frac{1}{2} \nabla^2 u(\boldsymbol{\xi}_{\mathbf{y}})(\mathbf{y} - \mathbf{x}) \cdot (\mathbf{y} - \mathbf{x})$$

and therefore

$$\eta_h * u(\mathbf{x}) - u(\mathbf{x}) = \frac{1}{2 \cdot B_{h^s,d}} \int_{B(\mathbf{0},h^s)} \nabla^2 u(\boldsymbol{\xi}_{\mathbf{y}}) |\mathbf{y}|^2 \,\mathrm{d}\mathbf{y}. \tag{16}$$

Proposition 2. *For all $u \in \mathbb{P}_{h,\boldsymbol{k}}$ and subdomain Ω_j we have*

$$\|\nabla_h u - \eta_h * \nabla_h u\|_{\Omega_j} \lesssim h^{\frac{s-1}{2}} \|\nabla_h u\|_{\tilde{\Omega}_j}.$$

Proof: We first use the triangle inequality

$$\begin{aligned}\|\nabla_h u - \eta_h * \nabla_h u\|_{\Omega_j} &\le \|\nabla_h u - \eta_h * \nabla_h u\|_{\Omega_{j0}} + \|\nabla_h u - \eta_h * \nabla_h u\|_{\Omega_j \setminus \Omega_{j0}} \\ &\le \|\nabla_h u - \eta_h * \nabla_h u\|_{\Omega_{j0}} + \|\nabla_h u\|_{\Omega_j \setminus \Omega_{j0}} + \|\eta_h * \nabla_h u\|_{\Omega_j \setminus \Omega_{j0}},\end{aligned} \tag{17}$$

where the contributions are estimated separately. We obviously have the estimate $\lambda_d(\Omega_j \setminus \Omega_{j0}) \lesssim h^s h_{\Omega_j}^{d-1}$ such that a simple scaling argument gives

$$\|\nabla_h u\|_{\Omega_j \setminus \Omega_{j0}} \lesssim h^{\frac{s-1}{2}} \|\nabla_h u\|_{\Omega_j}. \tag{18}$$

This also implies, using (14) in the second line with $s = 1$ that

$$\begin{aligned}&\|\eta_h * \nabla_h u\|^2_{\Omega_j \setminus \Omega_{j0}} \le \lambda_d(\Omega_j \setminus \Omega_{j0}) \max_{\tilde{\Omega}_j} |\nabla_h u|^2 \\ &\le h_{\Omega_j}^{d-1} h^s \max_{\tilde{\Omega}_j} |\nabla_h u|^2 \le h_{\Omega_j}^{d-1} h^s h_{\Omega_j}^{-d} \|\nabla_h u\|^2_{\tilde{\Omega}_j} = h^{s-1} \|\nabla_h u\|^2_{\tilde{\Omega}_j}.\end{aligned} \tag{19}$$

Finally, combining the inequalities in (16) and (12) we arrive at the estimate

$$\begin{aligned}|\nabla_h u - \eta_h * \nabla_h u|_{\Omega_{j0}}| &\le \frac{1}{2 \cdot \lambda_d(B(\mathbf{x}, h^s))} \max_{\mathbf{y} \in \Omega_j} |\nabla^3 u(\mathbf{y})| \int_{B(\mathbf{0}, h^s)} |\mathbf{y}|^2 \, \mathrm{d}\mathbf{y} \\ &\lesssim h^{-sd} \max_{\mathbf{y} \in \Omega_j} |\nabla^3 u(\mathbf{y})| h^{s(d+2)} \lesssim h^{2s} h^{-\frac{d}{2}-2} \|\nabla u\|_{\Omega_j}.\end{aligned}$$

Therefore, using (10) we obtain

$$\|\nabla_h u - \eta_h * \nabla_h u\|_{\Omega_{j0}} \le h^{2s-2} \|\nabla_h u\|_{\Omega_j}. \tag{20}$$

The estimates (18), (19) and (20) with (17) imply then the inequality in the proposition. $\square$

Remark: For functions $v \in C^2(\mathbb{R}^d)$ one can easily estimate the difference in Proposition 2. Moreover, it turns out that the convergence rate of the difference $\int_{\mathbb{R}^d} |\eta_h * v|^2 - |v|^2$ characterizes the Sobolev space $H^1(\mathbb{R}^n)$, see [27].

The chief problem in the estimations with convolution terms is that the scaling arguments can not be applied in a straightforward way. Whenever we use polynomial spaces the function space $\{\eta_h * v : v \in \mathbb{P}_{h,\mathbf{k}}, 0 < h < h_0\}$ is infinite dimensional, which makes the following proofs non-trivial.

Proposition 3. *There exists $h_0 > 0$ such that for all h with $h^{1-\frac{1}{s}} < h_0$ and $v \in \mathbb{P}_{h,\boldsymbol{k}}(\Omega_+ \cup \Omega_-)$ we have*

$$\int_{f\Omega} | [\![v]\!] | \lesssim \int_{\Omega_+ \cup \Omega_-} |\nabla(\eta_h * v)| \tag{21}$$

and for $s \ge \frac{3}{2}$

$$\int_{f\Omega} | [\![v]\!] | \lesssim h^{\frac{d}{2}} h_\Omega^{\frac{s-1}{2}} \sqrt{\int_{\Omega_+ \cup \Omega_-} |\nabla(\eta_h * v)|^2}. \tag{22}$$

The corresponding proof is postponed to the Appendix.

4.1. The bilinear form

To give the bilinear form (6) in a more explicit form, we first need some identities for distributional derivatives. We first decompose the gradient of a function $u \in \mathbb{P}_{h,\mathbf{k}}(\Omega)$ as follows.

Lemma 1. *For all $u \in \mathbb{P}_{h,\boldsymbol{k}}(\Omega)$ we have*

$$\nabla u = \nabla_h u + [\![u]\!]_{\mathcal{D}} = \nabla_h u + \sum_{f \in \mathcal{F}} [\![u]\!]_f \tag{23}$$

in the sense of distributions, i.e. $[\![u]\!]_{\mathcal{D}} \in [\mathcal{D}^d(\Omega)]^*$ is a distribution with

$$\langle [\![u]\!]_{\mathcal{D}}, \boldsymbol{\phi}\rangle = -\sum_{f\in\mathcal{F}} \int_f [\![u]\!]_f \cdot \boldsymbol{\phi} := -\int_{\mathcal{F}} [\![u]\!] \cdot \boldsymbol{\phi} = -([\![u]\!], \boldsymbol{\phi})_{\mathcal{F}},$$

where $[\![u]\!]_f = [\![u]\!]_{\mathcal{D}}|_f$ and we also have supp $[\![u]\!]_{\mathcal{D}} = \mathcal{F}$.

Proof: Obviously, for all $\boldsymbol{\phi} \in [\mathcal{D}(\Omega)]^d$ we have

$$\begin{aligned}\langle \nabla u, \boldsymbol{\phi}\rangle &= -\langle u, \operatorname{div} \boldsymbol{\phi}\rangle = -\sum_{\Omega_j \in \mathcal{T}_h} \int_{\Omega_j} u \operatorname{div} \boldsymbol{\phi} = \sum_{\Omega_j \in \mathcal{T}_h} \int_{\Omega_j} \nabla u \cdot \boldsymbol{\phi} - \sum_{\Omega_j \in \mathcal{T}_h} \int_{\partial\Omega_j} u|_{\Omega_j} \boldsymbol{\nu}_j \cdot \boldsymbol{\phi} \\ &= \sum_{\Omega_j \in \mathcal{T}_h} \int_{\Omega_j} \nabla u \cdot \boldsymbol{\phi} - \sum_{f\in\mathcal{F}} \int_f [\![u]\!]_f \cdot \boldsymbol{\phi}|_f,\end{aligned}$$

which proves the statement. □

Remarks: The decomposition in Lemma 1 is indeed a Lebesgue decomposition [28], [29] of the Radon measure corresponding to the distributional derivative ∇u. The role of the jump terms in this context is analyzed in [30], Section 10.

The symbol $[\![\cdot]\!]_{\mathcal{D}}$ can be understood both as a distribution supported on the interelement faces and the singular measure in the corresponding Lebesgue decomposition. The connection between $[\![u]\!]_{\mathcal{D}}$ and the classical function $[\![u]\!]$ is highlighted in Lemma 1.

The negative sign is a weakness of the conventional notation. This is already transparent in the one-dimensional case: whenever the Heaviside step function $H : \mathbb{R} \to \mathbb{R}$ is increasing, by definition we have $[\![H]\!](0) = -1$.

For the consecutive derivations we need also an identity regarding the convolution of distributions.

Lemma 2. *For all $u \in \mathbb{P}_{h,\boldsymbol{k}}$ the convolution $\eta_h * [\![u]\!]_{\mathcal{D}}$ is regular, which will be identified with the corresponding locally integrable function. With this, for all bounded function $\mathbf{w} : \Omega \to \mathbb{R}^d$ we have*

$$\langle \eta_h * [\![u]\!]_{\mathcal{D}}, \mathbf{w}\rangle = ([\![u]\!], \eta_h * \mathbf{w})_{\mathcal{F}}.$$

Similar statement holds for $[\![u]\!]_f$ with the following identity:

$$\langle \eta_h * [\![u]\!]_f \,, \mathbf{w}\rangle = ([\![u]\!] \,, \eta_h * \mathbf{w})_f.$$

Proof: Since both η_h and $[\![u]\!]_{\mathcal{D}}$ are compactly supported, we get by definition (see [31], Definition 2.1) and by Lemma 1 that for each $\boldsymbol{\phi} \in [C_0^\infty(\Omega)]^d$ the following equality is valid:

$$\begin{aligned}
&\langle \eta_h * [\![u]\!]_{\mathcal{D}} \,, \boldsymbol{\phi}\rangle = \langle [\![u]\!]_{\mathcal{D}} \,, \mathbf{y} \to \eta_h(\mathbf{x} \to \boldsymbol{\phi}(\mathbf{x}+\mathbf{y}))\rangle = \langle [\![u]\!]_{\mathcal{D}} \,, \mathbf{y} \to \int_{\mathbb{R}^d} \eta_h(\mathbf{x})\boldsymbol{\phi}(\mathbf{x}+\mathbf{y})\,\mathrm{d}\mathbf{x}\rangle \\
&= -\int_{\mathcal{F}} [\![u]\!](\mathbf{y}) \int_{\mathbb{R}^d} \eta_h(\mathbf{x})\boldsymbol{\phi}(\mathbf{x}+\mathbf{y})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} = -\int_{\mathcal{F}} [\![u]\!](\mathbf{y}) \int_{\mathbb{R}^d} \eta_h(\mathbf{z}-\mathbf{y})\boldsymbol{\phi}(\mathbf{z})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{y} \\
&= -\int_{\mathcal{F}} [\![u]\!](\mathbf{y}) \int_{\mathbb{R}^d} \eta_h(\mathbf{y}-\mathbf{z})\boldsymbol{\phi}(\mathbf{z})\,\mathrm{d}\mathbf{z}\,\mathrm{d}\mathbf{y} = -\int_{\mathcal{F}} [\![u]\!](\mathbf{y})\, \eta_h * \boldsymbol{\phi}(\mathbf{y})\,\mathrm{d}\mathbf{y} = -([\![u]\!] \,, \eta_h * \boldsymbol{\phi})_{\mathcal{F}}.
\end{aligned} \tag{24}$$

On the other hand, according to [31], page 337, Exercise 10, $\eta_h * [\![u]\!]_{\mathcal{D}}$ is locally integrable such that the statement of the lemma is valid for all bounded functions $\mathbf{w}$ as it was stated. The statement can also be applied for $\eta_h * [\![u]\!]_f$ and the derivation in (24) can be modified in an obvious way changing $[\![u]\!]_{\mathcal{D}}$ to $[\![u]\!]_f$ and the symbol $\mathcal{F}$ to f. $\square$

As an obvious consequence of Lemma 1 and Lemma 2 we get the following.

Corollary 1. *The left hand side $a_\eta(u, v)$ of (6) can be rewritten as*

$$\begin{aligned}
&(\eta_h * \nabla_h u, \eta_h * \nabla_h v) + \langle \eta_h * \nabla_h u, \eta_h * [\![v]\!]_{\mathcal{D}}\rangle + \langle \eta_h * \nabla_h v, \eta_h * [\![u]\!]_{\mathcal{D}}\rangle + \langle \eta_h * [\![u]\!]_{\mathcal{D}} \,, \eta_h * [\![v]\!]_{\mathcal{D}}\rangle \\
&= (\eta_h * \nabla_h u, \eta_h * \nabla_h v) + (\eta_h * \nabla_h u, \eta_h * [\![v]\!]) + (\eta_h * \nabla_h v, \eta_h * [\![u]\!]) + (\eta_h * [\![u]\!]_{\mathcal{D}} \,, \eta_h * [\![v]\!]_{\mathcal{D}}) \\
&= (\eta_h * \nabla_h u, \eta_h * \nabla_h v) - (\eta_h * \eta_h * \nabla_h u, [\![v]\!])_{\mathcal{F}} - (\eta_h * \eta_h * \nabla_h v, [\![u]\!])_{\mathcal{F}} \\
&+ \left(\sum_{f\in\mathcal{F}} \eta_h * [\![u]\!]_f \,, \sum_{f\in\mathcal{F}} \eta_h * [\![v]\!]_f\right).
\end{aligned} \tag{25}$$

Remarks: The second line is given directly using the decomposition in (23), in the third line we have identified $\eta * [\![u]\!]$ and $\eta * [\![v]\!]$ with the corresponding locally integrable functions, which is related to the lifted forms of the dG methods as each scalar product corresponds to a volume integral. On the other hand, the second and third terms in the fourth line are integrals which can be computed on faces according to (24).
The locally integrable function $\eta_h * [\![u]\!]_f$ will be given explicitly in Lemma 5.

5. Comparison with the IP bilinear form

We compare our bilinear form (25) with the IP bilinear form (3) componentwise.

Comparison of the first terms.

Lemma 3. *For all $u, v \in \mathbb{P}_{h,\boldsymbol{k}}(\Omega)$ we have*

$$|(\nabla_h u, \nabla_h v) - (\eta_h * \nabla_h u, \eta_h * \nabla_h v)| \le h^{s-1} \|\eta_h * \nabla_h u\| \|\eta_h * \nabla_h v\|.$$

Proof: We obviously have

$$\begin{aligned} &|(\nabla_h u, \nabla_h v) - (\eta_h * \nabla_h u, \eta_h * \nabla_h v)| \\ &\le |(\nabla_h u - \eta_h * \nabla_h u, \nabla_h v) + (\eta_h * \nabla_h u, \nabla_h v - \eta_h * \nabla_h v)| \\ &\le \|\nabla_h u - \eta_h * \nabla_h u\| \|\nabla_h v\| + \|\eta_h * \nabla_h u\| \|\nabla_h v - \eta_h * \nabla_h v\|. \end{aligned} \tag{26}$$

Also, application of the estimate in Proposition 2 and a simple scaling argument implies for each subdomain Ω_j that

$$\begin{aligned} &\|\nabla_h u|_{\Omega_{j0}}\| \le \|\nabla_h u - \eta_h * \nabla_h u\|_{\Omega_{j0}} + \|\eta_h * \nabla_h u\|_{\Omega_{j0}} \\ &\le h^{\frac{s-1}{2}} \|\nabla_h u\|_{\Omega_j} + \|\eta_h * \nabla_h u\|_{\Omega_{j0}} \lesssim h^{\frac{s-1}{2}} \|\nabla_h u\|_{\Omega_{j0}} + \|\eta_h * \nabla_h u\|_{\Omega_{j0}} \end{aligned}$$

and therefore,

$$\|\nabla_h u\|_{\Omega_{j0}} \lesssim \|\eta_h * \nabla_h u\|_{\Omega_{j0}},$$

which can be used to obtain the following inequality:

$$\|\nabla_h u\|_{\Omega_j} \lesssim \|\nabla_h u\|_{\Omega_{j0}} \lesssim \|\eta_h * \nabla_h u\|_{\Omega_{j0}} \le \|\eta_h * \nabla u\|_{\Omega_j}. \tag{27}$$

Therefore, using again Proposition 2 we also have

$$\|\nabla_h u - \eta_h * \nabla_h u\|_{\Omega_j} \lesssim h^{\frac{s-1}{2}} \|\eta_h * \nabla u\|_{\Omega_j}. \tag{28}$$

Taking the square of (27) and (28) for each index j and summing them we have

$$\|\nabla_h u\| \lesssim \|\eta_h * \nabla u\| \quad \text{and} \quad \|\nabla_h u - \eta_h * \nabla_h u\| \lesssim h^{\frac{s-1}{2}} \|\eta_h * \nabla u\|$$

which can be used in (26) to obtain

$$\begin{aligned} &|(\nabla_h u, \nabla_h v) - (\eta_h * \nabla_h u, \eta_h * \nabla_h v)| \\ &\lesssim h^{\frac{s-1}{2}} \|\eta_h * \nabla u\| \|\eta_h * \nabla v\| + h^{\frac{s-1}{2}} \|\eta_h * \nabla v\| \|\eta_h * \nabla u\| \end{aligned}$$

as stated in the lemma. $\square$

Comparison of the second and third terms. To compare the second and third terms in (25) and (3) we use the notation in Fig. 1 and the corresponding explanation.

To analyze the average of the approximations we use the following statement on integral means.

Proposition 4. *For each $u \in \mathbb{P}_{h,\boldsymbol{k}}(\Omega_- \cup \Omega_+)$ and $\mathbf{x} \in f_\Omega$ with $B(\mathbf{x}, 2h^s) \subset \Omega_- \cup \Omega_+$ there exist $\bar{\mathbf{x}}_- \in B_-(\mathbf{x}, 2h^s)$ and $\bar{\mathbf{x}}_+ \in B_+(\mathbf{x}, 2h^s)$ such that*

$$u(\bar{\mathbf{x}}_-) = 2 \int_{B_-(\mathbf{0},2h^s)} u(\mathbf{x}-\mathbf{z}) \cdot \eta_h * \eta_h(\mathbf{z})\, \mathrm{d}\mathbf{z}$$

and similarly,

$$u(\bar{\mathbf{x}}_+) = 2 \int_{B_+(\mathbf{0},2h^s)} u(\mathbf{x}-\mathbf{z}) \cdot \eta_h * \eta_h(\mathbf{z})\, \mathrm{d}\mathbf{z},$$

where $B_-(\mathbf{0}, 2h^s)$ and $B_+(\mathbf{0}, 2h^s)$ denote the half-ball with non-positive and non-negative first coordinates, respectively.

The proof is postponed to the Appendix.

We also note here that the set $\{\mathbf{x} \in f_\Omega : B(\mathbf{x}, 2h^s) \subset \Omega_- \cup \Omega_+\}$ is non-empty by the definition of h and the relation $h^s < h$. This fact will be used in the following two statements.

Proposition 5. *For all $u \in \mathbb{P}_{h,\boldsymbol{k}}(\Omega_- \cup \Omega_+)$ we have the following inequality:*

$$\max_{\substack{\mathbf{x}\in f_\Omega \\ B(\mathbf{x},2h^s)\subset\Omega_+\cup\Omega_-}} |\eta_h * \eta_h * \nabla_h u(\mathbf{x}) - \{\!\!\{\nabla_h u\}\!\!\}(\mathbf{x})| \lesssim h^{s-\frac{d}{2}-1} \|\nabla_h u\|_{B(\Omega_+\cup\Omega_-)}. \tag{29}$$

Proof: Using the result of Proposition 4 we rewrite the difference on the left hand side of (29) as follows:

$$\begin{aligned}
&\eta_h * \eta_h * \nabla_h u(\mathbf{x}) - \{\!\!\{\nabla_h u\}\!\!\}(\mathbf{x}) \\
&= \frac{1}{2}\Bigg(2\int_{B_-(\mathbf{0},2h^s)} \nabla_h u(\mathbf{x}-\mathbf{z}) \cdot \eta_h * \eta_h(\mathbf{z})\, \mathrm{d}\mathbf{z} + 2\int_{B_+(\mathbf{0},2h^s)} \nabla_h u(\mathbf{x}-\mathbf{z}) \cdot \eta_h * \eta_h(\mathbf{z})\, \mathrm{d}\mathbf{z} \\
&\qquad - \nabla_h u(\mathbf{x}_-) - \nabla_h u(\mathbf{x}_+)\Bigg) \\
&= \frac{1}{2}\left(\nabla_h u(\bar{\mathbf{x}}_-) - \nabla_h u(\mathbf{x}_-) + \nabla_h u(\bar{\mathbf{x}}_+) - \nabla_h u(\mathbf{x}_+)\right).
\end{aligned} \tag{30}$$

We use then the estimate

$$|\nabla_h u(\bar{\mathbf{x}}_-) - \nabla_h u(\mathbf{x}_-)| \leq 2h^s \cdot \sup_{\mathbf{z}\in(\mathbf{x}_-,(0-,\mathbf{y}))} \|\nabla_h^2 u(\mathbf{z})\|$$

in (30) to see that

$$\begin{aligned}
&\max_{\substack{\mathbf{x}\in f_\Omega \\ B(\mathbf{x},2h^s)\subset\Omega_+\cup\bar{\Omega}_-}} |\eta_h * \eta_h * \nabla_h u(\mathbf{x}) - \{\!\!\{\nabla_h u\}\!\!\}(\mathbf{x})| \\
&\leq \frac{1}{2} \cdot 2h^s \cdot \left(\max_{\mathbf{z}\in f_\Omega \hat{+} B(\mathbf{0},2h^s)} \|\nabla_h^2 u(\mathbf{z})\| + \max_{\mathbf{z}\in f_\Omega \hat{+} B(\mathbf{0},2h^s)} \|\nabla_h^2 u(\mathbf{z})\|\right) = 2h^s \max_{\mathbf{z}\in f_\Omega \hat{+} B(\mathbf{0},2h^s)} \|\nabla_h^2 u(\mathbf{z})\|.
\end{aligned}$$

The last term here can be estimated using scaling arguments as

$$
\begin{aligned}
&2h^s \max_{\mathbf{z}\in f_\Omega \dot{+} B(\mathbf{0},2h^s)} \|\nabla_h^2 u(\mathbf{z})\| \lesssim h^s \sqrt{h^{-s} h_\Omega^{1-d}} \|\nabla_h^2 u\|_{f_\Omega \dot{+} B(\mathbf{0},2h^s)} \\
&\lesssim h^{\frac{1+s-d}{2}} \|\nabla_h^2 u\|_{f_\Omega \dot{+} B(\mathbf{0},2h^s)} \lesssim h^{\frac{1+s-d}{2}} h^{\frac{s-1}{2}} \|\nabla_h^2 u\|_{\Omega_+\cup\Omega_-} \lesssim h^{s-\frac{d}{2}-1} \|\nabla_h u\|_{\Omega_+\cup\Omega_-},
\end{aligned}
$$

which proves the statement of the proposition. $\square$

We can now relate the third and second terms in the proposed bilinear form (25) and the IP bilinear form.

Lemma 4. *For arbitrary $u, v \in \mathbb{P}_{h,\boldsymbol{k}}$ we have the following inequality:*

$$
|(\eta_h * \eta_h * \nabla_h u, [\![v]\!])_{\mathcal{F}} - (\{\!\{\nabla_h u\}\!\}, [\![v]\!])_{\mathcal{F}}| \lesssim h^{s-1} \|\nabla(\eta_h * u)\| \|\nabla(\eta_h * v)\|.
$$

Proof: Using the result of Proposition 5 and Proposition 3 we obtain the following estimation on the interelement face f_Ω between Ω_+ and Ω_-:

$$
\begin{aligned}
&|(\eta_h * \eta_h * \nabla_h u, [\![v]\!])_{f_\Omega} - (\{\!\{\nabla_h u\}\!\}, [\![v]\!])_{f_\Omega}| = |(\eta_h * \eta_h * \nabla_h u - \{\!\{\nabla_h u\}\!\}, [\![v]\!])_{f_\Omega}| \\
&\leq \max_{\mathbf{x}\in f_\Omega} |(\eta_h * \eta_h * \nabla_h u - \{\!\{\nabla_h u\}\!\}(\mathbf{x})| \int_{f_\Omega} |[\![v]\!]| \\
&\lesssim \max_{\substack{\mathbf{x}\in f_\Omega \\ B(\mathbf{x},2h^s)\subset \Omega_+\cup\bar{\Omega}_-}} |(\eta_h * \eta_h * \nabla_h u - \{\!\{\nabla_h u\}\!\}(\mathbf{x})| \int_{f_\Omega} |[\![v]\!]| \\
&\leq h^{s-\frac{d}{2}-1} \left(\frac{h}{h_\Omega}\right)^{\frac{d}{2}} \|\nabla_h u\|_{\Omega_+\cup\Omega_-} \int_f |[\![v]\!]| \leq h^{s-\frac{d}{2}-1} \left(\frac{h}{h_\Omega}\right)^{\frac{d}{2}} \|\nabla_h u\|_{\Omega_+\cup\Omega_-} \int_{\Omega_+\cup\Omega_-} |\nabla(\eta_h * v)| \\
&\lesssim h^{s-\frac{d}{2}-1} \left(\frac{h}{h_\Omega}\right)^{\frac{d}{2}} \|\nabla_h(\eta_h * u)\|_{\Omega_+\cup\Omega_-} h_\Omega^{\frac{d}{2}} \|\nabla(\eta_h * v)\|_{\Omega_+\cup\Omega_-} \\
&\lesssim h^{s-1} \|\nabla_h(\eta_h * u)\|_{\Omega_+\cup\Omega_-} \|\nabla(\eta_h * v)\|_{\Omega_+\cup\Omega_-}.
\end{aligned}
$$

Summing up these inequalities for each interelement face f_Ω and using the discrete Cauchy–Schwarz inequality result in the estimate

$$
\begin{aligned}
&|\sum_{f_\Omega\in\mathcal{F}} (\eta_h * \eta_h * \nabla_h u, [\![v]\!])_{f_\Omega} - \sum_{f_\Omega\in\mathcal{F}} (\{\!\{\nabla_h u\}\!\}, [\![v]\!])_{f_\Omega}| \leq \sum_{f_\Omega\in\mathcal{F}} |(\eta_h * \eta_h * \nabla_h u, [\![v]\!])_{f_\Omega} - (\{\!\{\nabla_h u\}\!\}, [\![v]\!])_{f_\Omega}| \\
&\leq \sum_{f_\Omega\in\mathcal{F}} h^{s-1} \|\nabla_h(\eta_h * u)\|_{\Omega_+\cup\Omega_-} \|\nabla_h(\eta_h * v)\|_{\Omega_+\cup\Omega_-} \\
&\lesssim h^{s-1} \sqrt{\sum_{f_\Omega\in\mathcal{F}} \|\nabla(\eta_h * u)\|^2_{\Omega_+\cup\Omega_-}} \sqrt{\sum_{f_\Omega\in\mathcal{F}} \|\nabla(\eta_h * v)\|^2_{\Omega_+\cup\Omega_-}} \lesssim h^{s-1} \|\nabla(\eta_h * u)\| \|\nabla(\eta_h * v)\|
\end{aligned}
$$

as stated in the lemma. $\square$

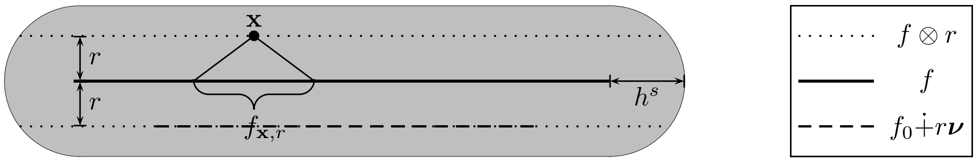

Figure 2: Interelement face f with the support of $\eta_h * [\![u]\!]_f$ (shaded) and the sets $f_{\mathbf{x},r}$, $f \otimes r$ and $f_0 \dot{+} r\boldsymbol{\nu}_+$.

Comparison of the fourth terms. To relate the last term in (25) with the penalty term in the IP bilinear form, we rewrite the locally integrable function $\eta_h * [\![v]\!]_{\mathcal{D}} = \eta_h * \sum_{f\in\mathcal{F}} [\![v]\!]_f$ (see Lemma 1 and Lemma 2) in a more explicit form.

Lemma 5. *For each $v \in \mathbb{P}_{\mathbf{h},k}$ and $f \in \mathcal{F}$ the locally integrable function corresponding to $\eta_h * [\![v]\!]_f$ can be given as*

$$\eta_h * [\![v]\!]_{\mathcal{D}} = \sum_{f\in\mathcal{F}} \eta_h * [\![v]\!]_f (\mathbf{x}) = \sum_{f\in\mathcal{F}} \int_f \eta_h(\mathbf{x}-\mathbf{y}) [\![v]\!]_f (\mathbf{y}) \,\mathrm{d}\mathbf{y}. \tag{31}$$

This result can also serve as a good argument for why we applied the same notation for the convolution corresponding to the jump of v and the jump function. Since the proof is a bit technical it is postponed to the Appendix.

To analyze the right hand side of (31), we introduce the following sets which are depicted in Figure 2.

$$f \otimes r = \{\mathbf{x} \in \langle f\rangle \dot{+} r\boldsymbol{\nu}_1 \cup \langle f\rangle \dot{+} r\boldsymbol{\nu}_2 : d(\mathbf{x}, f) \le h^s\}$$

and

$$f_0 \otimes r = (f_0 \dot{+} r\boldsymbol{\nu}_1) \cup (f_0 \dot{+} r\boldsymbol{\nu}_2),$$

where $\langle f\rangle$ denotes the affine subspace generated by $f = \Omega_+ \cap \Omega_-$.

Observe that $\eta_h * [\![u]\!]_f (\mathbf{x})$ can be nonzero if $\mathbf{x} \in f \otimes r$ for some $r < h^s$ and then we use the notation $f_{\mathbf{x},r} = B(\mathbf{x}, h^s) \cap f$, which is a ball in $\langle f\rangle$ centered at the projection of $\mathbf{x}$ on f with the radius $\sqrt{h^{2s} - r^2}$ such that $\lambda(f_{\mathbf{x},r}) = B_{\sqrt{h^{2s}-r^2},d-1}$.

With these, we can rewrite (31) as

$$\eta_h * [\![u]\!]_f (\mathbf{x}) = \frac{1}{B_{h^s,d}} \int_{f_{\mathbf{x},r}} [\![u]\!] .$$

In this way, using Lemma 5 the integral in the last term of (25) *on a face f* can be rewritten as

$$(\eta_h * [\![u]\!]_f , \eta_h * [\![v]\!]_f) = \frac{1}{[B_{h^s,d}]^2} \int_{-h^s}^{h^s} \int_{f\otimes r} \int_{f_{\mathbf{x},r}} [\![u]\!] (\mathbf{s}) \,\mathrm{d}\mathbf{s} \int_{f_{\mathbf{x},r}} [\![v]\!] (\mathbf{s}) \,\mathrm{d}\mathbf{s} \,\mathrm{d}\mathbf{x} \,\mathrm{d}r. \tag{32}$$

We intend to relate this term with the following:

$$\frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2 \,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r. \tag{33}$$

To work with smooth functions, both in (32) and (33) we have to restrict the integrals on $f_0\otimes r$ and to f_0, respectively. Since $\lambda(f)\sim h^{d-1}$ and $\lambda(f\setminus f_0)\sim h^{d-2}h^s$, a scaling argument implies the following estimates:

$$\begin{aligned}
&I_1(r) := \\
&\left|\int_{f\otimes r}\int_{f_{\mathbf{x},r}}[\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\int_{f_{\mathbf{x},r}}[\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{x} - \int_{f_0}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}[\![u]\!]\,(s)\,\mathrm{d}\mathbf{s}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}[\![v]\!]\,(s)\,\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{x}\right| \\
&= \left|\int_{f\otimes r}\int_{f_{\mathbf{x},r}}[\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\int_{f_{\mathbf{x},r}}[\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{x} - \int_{f_0\otimes r}\int_{f_{\mathbf{x},r}}[\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\int_{f_{\mathbf{x},r}}[\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{x}\right| \\
&\lesssim \frac{h^{d-2}h^s}{h^{d-1}}\int_{f\otimes r}\left|\int_{f_{\mathbf{x},r}}[\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\right|\left|\int_{f_{\mathbf{x},r}}[\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\right|\,\mathrm{d}\mathbf{x} = h^{s-1}\int_{f\otimes r}\left|\int_{f_{\mathbf{x},r}}[\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\right|\left|\int_{f_{\mathbf{x},r}}[\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\right|\,\mathrm{d}\mathbf{x}
\end{aligned} \tag{34}$$

and

$$\begin{aligned}
I_2(r) &:= \left|\int_{f_0}[B_{\sqrt{h^{2s}-r^2},d-1}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x} - \int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\right| \\
&\lesssim \frac{h^{d-2}h^s}{h^{d-1}}\int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2|\,[\![u]\!]\,[\![v]\!]\,|(\mathbf{x})\,\mathrm{d}\mathbf{x} \lesssim h^{s-1}(h^{2s}-r^2)^{d-1}\int_f |\,[\![u]\!]\,[\![v]\!]\,|(\mathbf{x})\,\mathrm{d}\mathbf{x}.
\end{aligned} \tag{35}$$

Remark: The estimation of I_1 is still valid if we use $f_{00}\subset f$ with $\lambda(f\setminus f_{00})\sim h^{d-2}h^s$.

For the forthcoming computations, we also give the magnitude of the following integrals:

$$\int_{-h^s}^{h^s}(h^{2s}-r^2)^{d-1}\,\mathrm{d}r = \mathcal{O}(h^{2sd-s}) \tag{36}$$

$$\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}|\mathbf{s}-\mathbf{x}|^2\,\mathrm{d}\mathbf{s} = \mathcal{O}(\sqrt{h^{2s}-r^2}), \tag{37}$$

which can be verified with a straightforward computation.

Lemma 6. *For all $u,v\in\mathbb{P}_{h,\boldsymbol{k}}$ and $\Omega_+,\Omega_-\in\mathcal{T}_h$ we have the following inequality*

$$\begin{aligned}
&\left|(\eta_h * [\![u]\!]_f\,, \eta_h * [\![v]\!]_f) - \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r\right| \\
&\le h^{s-1}(1+h^{3s-d-2})\|\nabla(\eta_h * u)\|_{\Omega_+\cup\Omega_-}\|\nabla(\eta_h * v)\|_{\Omega_+\cup\Omega_-}.
\end{aligned} \tag{38}$$

Proof: Using (32) and a triangle inequality with (34) and (35) we have

$$\begin{aligned}
&\left|(\eta_h * [\![u]\!]_f\,, \eta_h * [\![v]\!]_f) - \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r\right| \\
&\le \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s} I_1(r)\,\mathrm{d}r + \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s} I_2(r)\,\mathrm{d}r \\
&+ \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\left|\int_{f_0}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}[\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}[\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{x}\right. \\
&\qquad \left. - \int_{f_0}[B_{\sqrt{h^{2s}-r^2},d-1}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|\,\mathrm{d}r \\
&\le \frac{h^{s-1}}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_{f\otimes r}\left|\int_{f_{\mathbf{x},r}}[\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\right|\left|\int_{f_{\mathbf{x},r}}[\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\right|\,\mathrm{d}\mathbf{x}\,\mathrm{d}r \\
&+ \frac{h^{s-1}}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2|\,[\![u]\!]\,[\![v]\!]\,|(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r \\
&+ \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\left|\int_{f_0}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}[\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}[\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\right. \\
&\left. -[B_{\sqrt{h^{2s}-r^2},d-1}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|\,\mathrm{d}r.
\end{aligned} \tag{39}$$

The error terms here are estimated separately.

We first use (11) and (36) to obtain

$$\begin{aligned}
&\frac{h^{s-1}}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_{f\otimes r}\left|\int_{f_{\mathbf{x},r}}[\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\right|\left|\int_{f_{\mathbf{x},r}}[\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\right|\,\mathrm{d}\mathbf{x}\,\mathrm{d}r \\
&\lesssim h^{s-2sd-1}\int_{-h^s}^{h^s}\lambda_{d-1}(f\otimes r)[B_{\sqrt{h^{2s}-r^2},d-1}]^2\max_f\,[\![u]\!]\max_f\,[\![v]\!]\;\mathrm{d}r \\
&\lesssim h^{s-2sd-1}\int_{-h^s}^{h^s}h_\Omega^{d-1}(h^{2s}-r^2)^{d-1}\cdot h_\Omega^{1-d}\int_f|\,[\![u]\!]\,|h_\Omega^{1-d}\int_f|\,[\![v]\!]\,|\,\mathrm{d}r \\
&\le h^{s-2sd-d}\int_{-h^s}^{h^s}(h^{2s}-r^2)^{d-1}\,\mathrm{d}r\int_f|\,[\![u]\!]\,|\int_f|\,[\![v]\!]\,| \\
&= h^{s-2sd-d}h^{2sd-s}\int_f|\,[\![u]\!]\,|\int_f|\,[\![v]\!]\,| = h^{-d}\int_f|\,[\![u]\!]\,|\int_f|\,[\![v]\!]\,|.
\end{aligned} \tag{40}$$

We proceed similarly for the second term in (39):

$$\begin{aligned}
&\frac{h^{s-1}}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2\,|[\![u]\!]\,(\mathbf{x})\,[\![v]\!]\,(\mathbf{x})|\;\mathrm{d}\mathbf{x}\,\mathrm{d}r \\
&\lesssim h^{s-2sd-1}h^{2sd-s}\int_f|\,[\![u]\!]\,[\![v]\!]\,| \lesssim h^{-1}h_\Omega^{1-d}\int_f|\,[\![u]\!]\,|\int_f|\,[\![v]\!]\,| \le h^{-d}\int_f|\,[\![u]\!]\,|\int_f|\,[\![v]\!]\,|.
\end{aligned} \tag{41}$$

We finally estimate the third term in (39). Using the expansion in (15) on f_0 with the surface gradient $\nabla_f [\![u]\!] := \nabla [\![u]\!]$ and integrating both sides on the ball $B(\mathbf{x}, \sqrt{h^{2s}-r^2})$ implies

$$\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})} [\![u]\!](\mathbf{s})\,\mathrm{d}\mathbf{s} = B_{\sqrt{h^{2s}-r^2},d-1} [\![u]\!](\mathbf{x}) + \frac{1}{2}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})} \nabla^2 [\![u]\!](\boldsymbol{\xi}_{\mathbf{s}})(\mathbf{s}-\mathbf{x})\cdot(\mathbf{s}-\mathbf{x})\,\mathrm{d}\mathbf{s}. \tag{42}$$

Taking the product of (42) for $[\![u]\!]$ and $[\![v]\!]$ and using (37) and (13) we obtain

$$\begin{aligned}
&\left|\int_{f_0}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})} [\![u]\!](\mathbf{s})\,\mathrm{d}\mathbf{s}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})} [\![v]\!](\mathbf{s})\,\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{x} - [B_{\sqrt{h^{2s}-r^2},d-1}]^2 [\![u]\!][\![v]\!](\mathbf{x})\,\mathrm{d}\mathbf{x}\right|\\
&\le \int_{f_0}\left|B_{\sqrt{h^{2s}-r^2},d-1}[\![u]\!](\mathbf{x})\frac{1}{2}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}\nabla^2[\![v]\!](\boldsymbol{\xi}_{\mathbf{s}})|\mathbf{s}-\mathbf{x}|^2\,\mathrm{d}\mathbf{s}\right|\\
&+\left|B_{\sqrt{h^{2s}-r^2},d-1}[\![v]\!](\mathbf{x})\frac{1}{2}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}\nabla^2[\![u]\!](\boldsymbol{\xi}_{\mathbf{s}})|\mathbf{s}-\mathbf{x}|^2\,\mathrm{d}\mathbf{s}\right|\\
&+\left|\frac{1}{4}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}\nabla^2[\![u]\!](\boldsymbol{\xi}_{\mathbf{s}})|\mathbf{s}-\mathbf{x}|^2\,\mathrm{d}\mathbf{s}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}\nabla^2[\![v]\!](\boldsymbol{\xi}_{\mathbf{s}})|\mathbf{s}-\mathbf{x}|^2\,\mathrm{d}\mathbf{s}\right|\,\mathrm{d}\mathbf{x}\\
&\lesssim \max_f|\nabla^2[\![v]\!]|B_{\sqrt{h^{2s}-r^2},d-1}\int_f|[\![u]\!]|(\mathbf{x})\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}|\mathbf{s}-\mathbf{x}|^2\,\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{x}\\
&+\max_f|\nabla^2[\![u]\!]|B_{\sqrt{h^{2s}-r^2},d-1}\int_f|[\![v]\!]|(\mathbf{x})\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}|\mathbf{s}-\mathbf{x}|^2\,\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{x}\\
&+\max_f|\nabla^2[\![v]\!]|\max_f|\nabla^2[\![v]\!]|\int_{f_0}(h^{2s}-r^2)^{d+1}\,\mathrm{d}\mathbf{x}\\
&\lesssim h^{-d-1}\int_f|[\![v]\!]|B_{\sqrt{h^{2s}-r^2},d-1}\int_f|[\![u]\!]|(h^{2s}-r^2)^{\frac{d+1}{2}} + h^{-2d-2}\int_f|[\![v]\!]|\int_f|[\![u]\!]|(h^{2s}-r^2)^{d+1}\\
&\lesssim h^{-d-1}(h^{2s}-r^2)^d\int_f|[\![v]\!]|\int_f|[\![u]\!]| + h^{-2d-2}(h^{2s}-r^2)^{d+1}\int_f|[\![v]\!]|\int_f|[\![u]\!]|\\
&= (h^{-d-1}(h^{2s}-r^2)^d + h^{-2d-2}(h^{2s}-r^2)^{d+1})\int_f|[\![v]\!]|\int_f|[\![u]\!]|.
\end{aligned}$$

In this way, we can estimate the last term in (39) as

$$\begin{aligned}
&\frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\left|\int_{f_0}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}[\![u]\!](\mathbf{s})\,\mathrm{d}\mathbf{s}\int_{B(\mathbf{x},\sqrt{h^{2s}-r^2})}[\![v]\!](\mathbf{s})\,\mathrm{d}\mathbf{s} - [B_{\sqrt{h^{2s}-r^2},d-1}]^2[\![u]\!][\![v]\!](\mathbf{x})\,\mathrm{d}\mathbf{x}\right|\,\mathrm{d}r\\
&\lesssim \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}(h^{-d-1}(h^{2s}-r^2)^d + h^{-2d-2}(h^{2s}-r^2)^{d+1})\int_f|[\![v]\!]|\int_f|[\![u]\!]|\,\mathrm{d}r\\
&\lesssim h^{-2sd}\int_f|[\![v]\!]|\int_f|[\![u]\!]|\int_{-h^s}^{h^s}h^{-d-1}(h^{2s}-r^2)^d + h^{-2d-2}(h^{2s}-r^2)^{d+1})\,\mathrm{d}r\\
&\lesssim h^{-2sd}\int_f|[\![v]\!]|\int_f|[\![u]\!]|\cdot(h^{-d-1}h^{2sd+s} + h^{-2d-2}h^{2sd+3s})\\
&= (h^{-d-1+s} + h^{-2d-2+3s})\int_f|[\![v]\!]|\int_f|[\![u]\!]|
\end{aligned}$$

and therefore, using (39), (40), (41) and the estimate (22) in Proposition 3 we finally obtain

$$
\begin{aligned}
&\left|(\eta_h * [\![u]\!]_f \,, \eta_h * [\![v]\!]_f) - \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2\, [\![u]\!]\, [\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r\right| \\
&\lesssim (h^{-d} + h^{-d-1+s} + h^{-2d-2+3s})\int_f |\,[\![v]\!]\,| \int_f |\,[\![u]\!]\,| \\
&\lesssim (h^{-d} + h^{-d-1+s} + h^{-2d-2+3s})h^d h^{s-1}\|\nabla(\eta_h * u)\|_{\Omega_-\cup\Omega_+}\|\nabla(\eta_h * v)\|_{\Omega_-\cup\Omega_+} \\
&\lesssim h^{s-1}(1 + h^{-d+3s-2})\|\nabla(\eta_h * u)\|_{\Omega_-\cup\Omega_+}\|\nabla(\eta_h * v)\|_{\Omega_-\cup\Omega_+}
\end{aligned}
$$

as we have stated. $\square$

Lemma 7. *For each different faces f_j and f_k we have the estimate*

$$
\left|(\eta_h * [\![u]\!]_{f_j} \,, \eta_h * [\![v]\!]_{f_k})\right| \lesssim h^{s-1}\|\nabla(\eta_h * u)\|_{\Omega_{j+}\cup\Omega_{j-}}\|\nabla(\eta_h * v)\|_{\Omega_{k+}\cup\Omega_{k-}}. \tag{43}
$$

Proof: Note that the scalar product on the left hand side has to be computed on $S_{j,k} := \operatorname{supp}\eta_h * [\![u]\!]_{f_j} \cap \operatorname{supp}\eta_h * [\![u]\!]_{f_j}$. We first focus to the two-dimensional case. We use the transformation $\mathcal{S}_{j,k} : S_{j,k} \to [-h^s, h^s] \times [-h^s, h^s]$, which maps a point in conv $\{f_j, f_k\}$ with distances s_1 and s_2 measured from f_j and f_k, respectively to the point (s_1, s_2) and acts similarly for the three remaining points with these distances. This transformation is depicted in Figure 3. A straightforward computation shows that the determinant corresponding to this transformation is bounded as long as the condition of non-degeneracy is satisfied. Using the principle of the estimation in (40) we obtain the following estimate:

$$
\begin{aligned}
&\left|(\eta_h * [\![u]\!]_{f_j} \,, \eta_h * [\![v]\!]_{f_k})\right| = \frac{1}{[B_{h^s,d}]^2}\left|\int_{S_{j,k}}\int_{f_{j,\mathbf{x},r}} [\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\int_{f_{k,\mathbf{x},r}} [\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{x}\right| \\
&\lesssim \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_{-h^s}^{h^s}\left|\int_{f_{j,\mathbf{x},r}} [\![u]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\int_{f_{k,\mathbf{x},r}} [\![v]\!]\,(\mathbf{s})\,\mathrm{d}\mathbf{s}\right|\,\mathrm{d}\mathbf{x} \\
&\lesssim \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\max_{f_j}|\,[\![u]\!]\,|B_{\sqrt{h^{2s}-s_1^2}}\,\mathrm{d}s_1\int_{-h^s}^{h^s}\max_{f_k}|\,[\![v]\!]\,|B_{\sqrt{h^{2s}-s_2^2}}\,\mathrm{d}s_2 \\
&\lesssim \frac{1}{[B_{h^s,d}]^2}\max_{f_j}|\,[\![u]\!]\,|\max_{f_k}|\,[\![v]\!]\,|\left(\int_0^{h^s}(h^{2s}-s_1^2)^{\frac{d-1}{2}}\,\mathrm{d}s_1\right)^2 \\
&\lesssim \frac{1}{[B_{h^s,d}]^2}h^{2-2d}\int_{f_j}|\,[\![u]\!]\,|\int_{f_k}|\,[\![v]\!]\,|h^{2sd} \lesssim h^{2-2d}\int_{f_j}|\,[\![u]\!]\,|\int_{f_k}|\,[\![v]\!]\,| \\
&\lesssim h^{2-2d}h^s h_\Omega^{d-1}\|\nabla(\eta_h * u)\|_{\Omega_{j+}\cup\Omega_{j-}}\|\nabla(\eta_h * v)\|_{\Omega_{k+}\cup\Omega_{k-}} \\
&\lesssim h^{1+s-d}\|\nabla(\eta_h * u)\|_{\Omega_{j+}\cup\Omega_{j-}}\|\nabla(\eta_h * v)\|_{\Omega_{k+}\cup\Omega_{k-}}.
\end{aligned}
$$

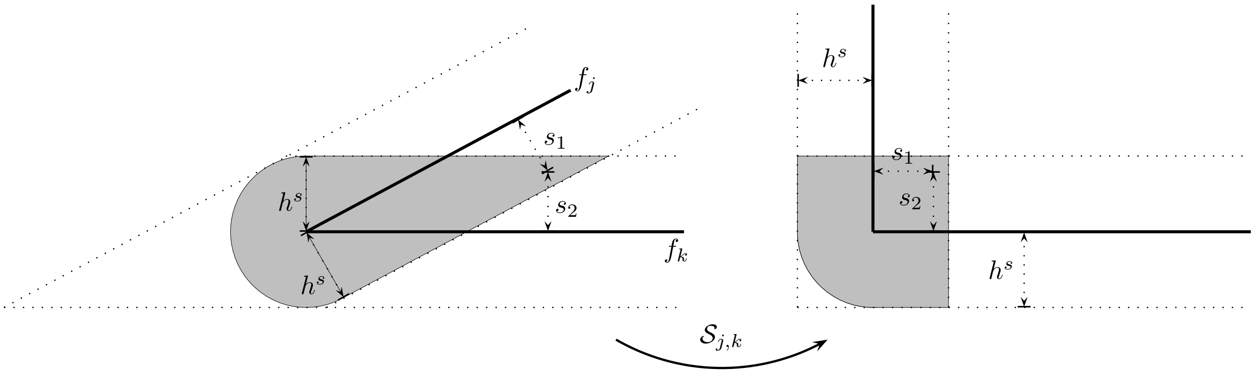

Figure 3: The support $S_{j,k} = \operatorname{supp} \eta_h * [\![u]\!]_{f_j} \cap \operatorname{supp} \eta_h * [\![u]\!]_{f_j}$ (left, shaded) and its transformation $\mathcal{S}_{j,k}$ into $[-h^s, h^s] \times [-h^s, h^s]$ in a 2-dimensional setup.

In the three-dimensional case, we also have to integrate along the common edge $e_{j,k} = \bar{f}_j \cap f_k$, which similarly to the above estimate gives

$$
\begin{aligned}
&\left|\left(\eta_h * [\![u]\!]_{f_j}, \eta_h * [\![v]\!]_{f_k}\right)\right| = \frac{1}{[B_{h^s,d}]^2} \left| \int_{S_{j,k}} \int_{f_{j,\mathbf{x},r}} [\![u]\!](\mathbf{s})\, \mathrm{d}\mathbf{s} \int_{f_{k,\mathbf{x},r}} [\![v]\!](\mathbf{s})\, \mathrm{d}\mathbf{s}\, \mathrm{d}\mathbf{x} \right| \\
&\lesssim \frac{1}{[B_{h^s,d}]^2} \int_{e_{j,k}} \int_{-h^s}^{h^s} \int_{-h^s}^{h^s} \left| \int_{f_{j,\mathbf{x},r}} [\![u]\!](\mathbf{s})\, \mathrm{d}\mathbf{s} \int_{f_{k,\mathbf{x},r}} [\![v]\!](\mathbf{s})\, \mathrm{d}\mathbf{s} \right| \mathrm{d}\mathbf{x} \\
&\lesssim h^{2-2d} h^s h_\Omega h_\Omega^{d-1} \|\nabla(\eta_h * u)\|_{\Omega_{j+} \cup \Omega_{j-}} \|\nabla(\eta_h * v)\|_{\Omega_{k+} \cup \Omega_{k-}} \\
&\lesssim h^{2+s-d} \|\nabla(\eta_h * u)\|_{\Omega_{j+} \cup \Omega_{j-}} \|\nabla(\eta_h * v)\|_{\Omega_{k+} \cup \Omega_{k-}}.
\end{aligned}
$$

Summarized, in both cases $d = 2$ and $d = 3$ we finaly arrive at

$$
\left|\left(\eta_h * [\![u]\!]_{f_j}, \eta_h * [\![v]\!]_{f_k}\right)\right| \lesssim h^{s-1} \|\nabla(\eta_h * u)\|_{\Omega_{j+} \cup \Omega_{j-}} \|\nabla(\eta_h * v)\|_{\Omega_{k+} \cup \Omega_{k-}}
$$

as stated in the lemma. $\square$

Corollary 2. *For all $u, v \in \mathbb{P}_{h,\boldsymbol{k}}$ we have*

$$
\begin{aligned}
&\left| \left(\sum_{f \in \mathcal{F}} \eta_h * [\![u]\!]_f, \sum_{f \in \mathcal{F}} \eta_h * [\![u]\!]_f \right) - \sum_{f \in \mathcal{F}} \frac{1}{[B_{h^s,d}]^2} \int_{-h^s}^{h^s} \int_f [B_{\sqrt{h^{2s} - r^2}, d-1}]^2 [\![u]\!] [\![v]\!](\mathbf{x})\, \mathrm{d}\mathbf{x}\, \mathrm{d}r \right| \\
&\leq h^{s-1}(1 + h^{3s-d-2}) \|\nabla(\eta_h * u)\| \|\nabla(\eta_h * v)\|.
\end{aligned}
$$

Proof: Using the triangle inequality with (7) and taking the sum of the inequalities in (38) and applying the discrete Cauchy–Schwarz inequality $|\sum_{j\in J} a_j b_j| \leq \sqrt{\sum_{j\in J} a_j^2}\sqrt{\sum_{j\in J} b_j^2}$ we obtain

$$\begin{aligned}
&\left|\left(\sum_{f\in\mathcal{F}} \eta_h * [\![u]\!]_f\,, \sum_{f\in\mathcal{F}} \eta_h * [\![u]\!]_f\right) - \sum_{f\in\mathcal{F}} \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r\right| \\
&\leq 2\left|\sum_{j\neq k}(\eta_h * [\![u]\!]_{f_j}\,, \eta_h * [\![u]\!]_{f_k})\right| \\
&+\left|\sum_{f\in\mathcal{F}}(\eta_h * [\![u]\!]_f\,, \eta_h * [\![v]\!]_f) - \sum_{f\in\mathcal{F}} \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r\right| \\
&\leq h^{s-1}\sum_{f\in\mathcal{F}} \|\nabla(\eta_h * u)\|_{\Omega_+\cup\Omega_-}\|\nabla(\eta_h * v)\|_{\Omega_+\cup\Omega_-} \\
&+\sum_{f\in\mathcal{F}}\left|(\eta_h * [\![u]\!]_f\,, \eta_h * [\![v]\!]_f) - \frac{1}{[B_{h^s,d}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r\right| \\
&\leq h^{s-1}(2 + h^{3s-d-2})\sum_{f\in\mathcal{F}} \|\nabla(\eta_h * u)\|_{\Omega_+\cup\Omega_-}\|\nabla(\eta_h * v)\|_{\Omega_+\cup\Omega_-} \\
&\lesssim h^{s-1}(1 + h^{3s-d-2})\|\nabla(\eta_h * u)\|\|\nabla(\eta_h * v)\|
\end{aligned}$$

as stated in the corollary. $\square$

Remark: The difference in the corollary is a higher-order term compared to $\|\nabla(\eta_h * u)\|\|\nabla(\eta_h * v)\|$ provided that $4s - d - 3 > 0$.

Finally, we compute the approximation of the penalty term in (33), which appears in Lemma 6.

- For $d = 2$ we have

$$\begin{aligned}
&\frac{1}{[B_{h^s,2}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},1}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r \\
&= \frac{1}{h^{4s}\pi^2}\int_{-h^s}^{h^s} 4(h^{2s} - r^2)\,\mathrm{d}r \int_f [\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x} = \frac{16}{3\pi^2}h^{-s}\int_f [\![u]\!]\,[\![v]\!]\,.
\end{aligned}$$

- For $d = 3$ we have

$$\begin{aligned}
&\frac{1}{[B_{h^s,3}]^2}\int_{-h^s}^{h^s}\int_f [B_{\sqrt{h^{2s}-r^2},2}]^2\,[\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r \\
&= \frac{9}{16h^{6s}\pi^2}\int_{-h^s}^{h^s} \pi^2(h^{2s} - r^2)^2\,\mathrm{d}r \int_f [\![u]\!]\,[\![v]\!]\,(\mathbf{x})\,\mathrm{d}\mathbf{x} = \frac{3}{5}h^{-s}\int_f [\![u]\!]\,[\![v]\!]\,.
\end{aligned}$$

Proof of Theorem 1: Using Lemma 3, Lemma 4 and Corollary 2 we obtain

$$
\begin{aligned}
&|a_\eta(u,v) - a_{\mathrm{IP},s}(u,v)| \leq |(\eta_h * \nabla_h u, \eta_h * \nabla_h v) - (\nabla_h u, \nabla_h v)| \\
&+ \left| \sum_{f\in\mathcal{F}} (\eta_h * \eta_h * \nabla_h u, [\![v]\!])_f + (\eta_h * \eta_h * \nabla_h v, [\![u]\!])_f - (\{\!\{\nabla_h u\}\!\}, [\![v]\!])_f - (\{\!\{\nabla_h v\}\!\}, [\![u]\!]_f) \right| \\
&+ \left| \sum_{f\in\mathcal{F}} (\eta_h * [\![u]\!], \eta_h * [\![v]\!])_f - \sum_{f\in\mathcal{F}} \frac{1}{[B_{h^s,d}]^2} \int_{-h^s}^{h^s} \int_f [B_{\sqrt{h^{2s}-r^2},d-1}]^2 [\![u]\!] [\![v]\!] (\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}r \right| \\
&\lesssim (h^{s-1} + h^{s-1}(1 + h^{3s-d-2})) \|\nabla(\eta_h * u)\| \|\nabla(\eta_h * v)\|
\end{aligned}
$$

as stated in the theorem. $\square$

Proof of Theorem 2: Since u_h solves (6) and $u_{\mathrm{IP},s} \in \mathbb{P}_{h,\mathbf{k}}$ we have

$$
(\nabla(\eta_h * u_h), \nabla(\eta_h * (u_h - u_{\mathrm{IP},s})) = (g_0, \eta_h * (u_h - u_{\mathrm{IP},s}))
$$

such that using the equality

$$
(\eta_h * w_1, w_2) = (w_1, \eta_h * w_2)
$$

for compactly supported functions $w_1, w_2 \in L_1(\mathbb{R}^d)$ and the definition of $u_{\mathrm{IP},s}$ in (9) we obtain

$$
\begin{aligned}
&(\nabla(\eta_h * (u_h - u_{\mathrm{IP},s})), \nabla(\eta_h * (u_h - u_{\mathrm{IP},s}))) \\
&= (g_0, \eta_h * (u_h - u_{\mathrm{IP},s})) - (\nabla(\eta_h * u_{\mathrm{IP},s}), \nabla(\eta_h * (u_h - u_{\mathrm{IP},s}))) \\
&= (g_0, \eta_h * (u_h - u_{\mathrm{IP},s})) - a_{\mathrm{IP},s}(u_{\mathrm{IP},s}, u_h - u_{\mathrm{IP},s}) - (\nabla(\eta_h * u_{\mathrm{IP},s}), \nabla(\eta_h * (u_h - u_{\mathrm{IP},s}))) \\
&+ a_{\mathrm{IP},s}(u_{\mathrm{IP},s}, u_h - u_{\mathrm{IP},s}) \\
&= (\eta_h * g, u_h - u_{\mathrm{IP},s}) - (g, u_h - u_{\mathrm{IP},s}) - (\nabla(\eta_h * (u_{\mathrm{IP},s} - u_h)), \nabla(\eta_h * (u_h - u_{\mathrm{IP},s}))) \\
&+ a_{\mathrm{IP},s}(u_{\mathrm{IP},s} - u_h, u_h - u_{\mathrm{IP},s}) - (\nabla(\eta_h * u_h), \nabla(\eta_h * (u_h - u_{\mathrm{IP}}))) + a_{\mathrm{IP},s}(u_h, u_h - u_{\mathrm{IP},s}).
\end{aligned}
\tag{44}
$$

We note that the application of (27) to $u_h - u_{\mathrm{IP},s}$ (instead of $\nabla_h u$) and the Friedrichs's inequality imply

$$
\|u_h - u_{\mathrm{IP}}\| \lesssim \|\eta_h * (u_h - u_{\mathrm{IP}})\| \lesssim \max_j h_{\Omega_j}^d \|\nabla(\eta_h * (u_h - u_{\mathrm{IP}}))\|
$$

and therefore, using Theorem 1 for the last two pair of terms in (44), we obtain that

$$
\begin{aligned}
&\|\nabla(\eta_h * (u_h - u_{\mathrm{IP},s}))\|^2 \\
&\lesssim \|\eta_h * g - g\| \|u_h - u_{\mathrm{IP},s}\| + h^{s-1}(\|\nabla(\eta_h * (u_{\mathrm{IP},s} - u_h))\|^2 + \|\nabla(\eta_h * u_h)\| \|\nabla(\eta_h * (u_h - u_{\mathrm{IP},s}))\|) \\
&\lesssim \max_j h_{\Omega_j}^d \|\eta_h * g - g\| \|\nabla(\eta_h * (u_h - u_{\mathrm{IP},s}))\| + h^{s-1}(1 + h^{3s-d-2}) \|\nabla(\eta_h * (u_{\mathrm{IP},s} - u_h))\|^2 \\
&+ h^{s-1}(1 + h^{3s-d-2}) \|\nabla(\eta_h * u_h)\| \|\nabla(\eta_h * (u_h - u_{\mathrm{IP},s}))\|
\end{aligned}
$$

such that we finally get

$$(1-h^{s-1})\|\nabla(\eta_h*(u_h-u_{\mathrm{IP},s}))\| \lesssim \max_j h_{\Omega_j}^d\|\eta_h*g-g\| + h^{s-1}(1+h^{3s-d-2})\|\nabla(\eta_h*u_h)\|,$$

which implies the estimate in the theorem. □

To state quasi optimal convergence we observe that for each h we have $\eta_h * u_h \in \mathbb{P}_{h,\mathbf{k},s} \subset H_0^1(\Omega_h)$. This means that the method in (6) is not conforming since the approximation in general is not in $H_0^1(\Omega)$. Also, the bilinear form a_η^+ is non-consistent in the sense that the zero extension u_0 of u is not necessarily the solution of (6) for any h. In this way, we need to apply the Strang lemma [18], Section 2.3.2. For this we first note that for some constants C_1 and C_2 we have for all $0 \neq w_1, w_2 \in H_0^1(\Omega_h)$ that

$$|a_\eta^+(w_1,w_2)| \le C_1\|w_1\|_{H_0^1(\Omega_h)}\|w_2\|_{H_0^1(\Omega_h)}$$

and

$$C_2 \le \frac{a_\eta^+(w_1,w_2)}{\|w_1\|_{H_0^1(\Omega_h)}\|w_2\|_{H_0^1(\Omega_h)}}.$$

Lemma 8. *The numerical solution $\eta_h * u_h$ of (6) approximates u in quasi optimal way in the sense that*

$$\|\nabla(u-\eta_h*u_h)\|_{\Omega_h} \lesssim \inf_{v_h\in\mathbb{P}_{h,\boldsymbol{k}}}\|\nabla(u-\eta_h*v_h)\| + h^{s-\frac{1}{2}}\|\nabla u\|.$$

Proof: Since in this proof it is essential whether a scalar product is defined on Ω or on Ω_h, we indicate it in the subscript. A direct application of Lemma 2.25 in [18] gives that

$$\begin{aligned}&\|\nabla(u-\eta_h*u_h)\|_{\Omega_h}\\ &\le (1+\frac{C_1}{C_2})\inf_{v_h\in\mathbb{P}_{h,\mathbf{k}}}\|\nabla(u-\eta_h*v_h)\| + \sup_{\eta_h*v_h\in\mathbb{P}_{h,\mathbf{k},s}}\frac{(\nabla u_0,\nabla(\eta_h*v_h))_{\Omega_h}-(g_0,\eta_h*v_h)_{\Omega_h}}{\|\eta_h*v_h\|_{1,\Omega_h}},\end{aligned} \tag{45}$$

where the lower indices denote zero extensions. Using these, the second term in (45) can be rewritten as

$$\begin{aligned}&(\nabla u_0,\nabla(\eta_h*v_h))_{\Omega_h}-(g_0,\eta_h*v_h)_{\Omega_h} = (\nabla u,\nabla(\eta_h*v_h))_\Omega-(g,\eta_h*v_h)_\Omega\\ &= (-\Delta u,\eta_h*v_h)_\Omega + \langle\boldsymbol{\nu}\cdot\nabla u,\eta_h*v_h\rangle_{\partial\Omega}-(g,\eta_h*v_h)_\Omega = \langle\boldsymbol{\nu}\cdot\nabla u,\eta_h*v_h\rangle_{\partial\Omega}.\end{aligned}$$

Therefore, we can estimate (45) to obtain

$$\|\nabla(u-\eta_h*u_h)\|_\Omega \lesssim \inf_{v_h\in\mathbb{P}_{h,\mathbf{k}}}\|\nabla(u-\eta_h*v_h)\| + \sup_{\eta_h*v_h\in\mathbb{P}_{h,\mathbf{k},s}}\frac{\langle\boldsymbol{\nu}\cdot\nabla u,\eta_h*v_h\rangle_{\partial\Omega}}{\|\eta_h*v_h\|_{1,\Omega_h}}. \tag{46}$$

For the rest, it is sufficient to estimate the second term here. We apply a classical trace inequality in Ω and in $\Omega_h \setminus \Omega$ which implies

$$\begin{aligned}
&\sup_{\eta_h * v_h \in \mathbb{P}_{h,\mathbf{k},s}} \frac{\langle \boldsymbol{\nu} \cdot \nabla u, \eta_h * v_h \rangle_{\partial\Omega}}{\|\eta_h * v_h\|_{1,\Omega_h}} \leq \sup_{\eta_h * v_h \in \mathbb{P}_{h,\mathbf{k},s}} \frac{\|\boldsymbol{\nu} \cdot \nabla u\|_{-\frac{1}{2},\partial\Omega} \|\eta_h * v_h\|_{\frac{1}{2},\partial\Omega}}{\|\eta_h * v_h\|_{1,\Omega_h}} \\
&\lesssim \|u\|_{1,\Omega} \sup_{\eta_h * v_h \in \mathbb{P}_{h,\mathbf{k},s}} \frac{\|\eta_h * v_h\|_{\frac{1}{2},\partial\Omega}}{\|\eta_h * v_h\|_{1,\Omega_h}} \lesssim \|u\|_{1,\Omega} \sup_{\eta_h * v_h \in \mathbb{P}_{h,\mathbf{k},s}} \frac{\|\eta_h * v_h\|_{1,\Omega_h \setminus \Omega}}{\|\eta_h * v_h\|_{1,\Omega_h}} \\
&\lesssim \|u\|_{1,\Omega} \sup_{\eta_h * v_h \in \mathbb{P}_{h,\mathbf{k},s}} \frac{\|\nabla(\eta_h * v_h)\|_{\Omega_h \setminus \Omega}}{\|\nabla(\eta_h * v_h)\|_{\Omega_h}}.
\end{aligned} \tag{47}$$

Observe that the numerator can be rewritten, using Lemma 1, as

$$\|\nabla(\eta_h * v_h)\|_{\Omega_h \setminus \Omega} = \|\eta_h * \nabla_h v_h + \eta_h * [\![v_h]\!]\,\|_{\Omega_h \setminus \Omega} \leq \|\eta_h * \nabla_h v_h\|_{\Omega_h \setminus \Omega} + \|\eta_h * [\![v_h]\!]\,\|_{\Omega_h \setminus \Omega}.$$

To analyze the term $\|\eta_h * [\![v_h]\!]\,\|_{\Omega_h \setminus \Omega}$ we use the notations corresponding to Lemma 5 and Fig. 2. Furthermore, we define

$$f_{00} = \{\mathbf{x} \in f : d(\mathbf{x}, \partial\Omega) > h^s\}$$

such that

$$\operatorname{supp} \eta_h * [\![v_h]\!]_f\,|_{\Omega_h \setminus \Omega} \subset \operatorname{supp} \eta_h * [\![v_h]\!]_f \setminus \{f_0 \otimes r : r \in [0, h^s]\},$$

see also Fig. 4. According to (34) and the remark after (35) the first term $\frac{1}{B^2_{h^s,d}} \int_{-h^s}^{h^s} I_1(r)\,\mathrm{d}r$ in the second line of (39) provides an upper bound for $\|\eta_h * [\![v_h]\!]_f\,\|_{\Omega_h \setminus \Omega}$. Therefore, the estimate in (40) implies

$$\|\eta_h * [\![v_h]\!]_f\,\|^2_{\Omega_h \setminus \Omega} \leq h^{-d} \left[\int_f [\![v_h]\!]\right]^2 \lesssim h^{-d} h^d h^{s-1} \|\nabla(\eta_h * v_h)\|^2_{\Omega_+ \cup \Omega_-}.$$

Taking their sum for all interelement faces gives then

$$\|\eta_h * [\![v_h]\!]\,\|_{\Omega_h \setminus \Omega} \lesssim h^{\frac{s-1}{2}} \|\nabla(\eta_h * v_h)\|_{\Omega}. \tag{48}$$

Note that the condition on non-degeneracy implies that for all subdomains $\Omega_k \subset \tilde{\Omega}_j$ we have

$$\lambda(\Omega_k) \sim h_\Omega^d \tag{49}$$

and

$$\lambda(\Omega_{j,h}) \sim h^s h_\Omega^{d-1}. \tag{50}$$

Using then (10), (49), (50) and (27) we obtain that

$$\|\eta_h * \nabla_h v_h\|^2_{\Omega_{j,h}} \leq \lambda(\Omega_{j,h}) \max_{\tilde{\Omega}_j} |\nabla_h v_h|^2 \lesssim \lambda(\Omega_{j,h}) h_\Omega^{-d} \|\nabla_h v_h\|^2_{\tilde{\Omega}_j} \lesssim h_\Omega^{s-1} \|\nabla_h v_h\|^2_{\tilde{\Omega}_j} \lesssim h_\Omega^{s-1} \|\eta_h * \nabla v_h\|^2_{\tilde{\Omega}_j},$$

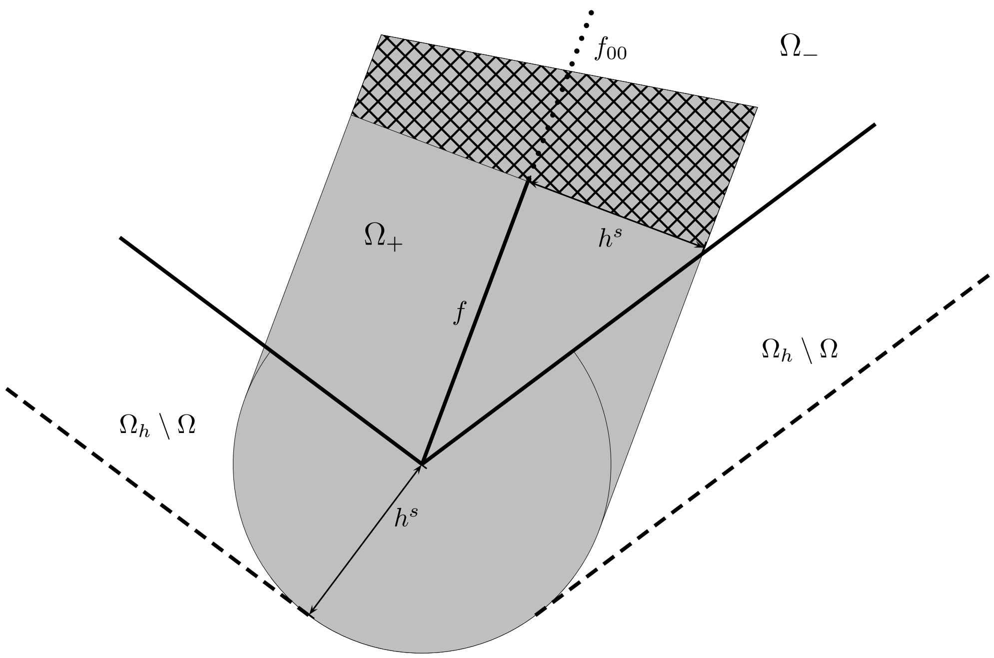

Figure 4: The support of $\eta_h * [\![v_h]\!]_f$(shaded) and the set $\{f_{00} \otimes r : r \in [0, h^s]\}$ (with crosshatch) in a 2-dimensional setup.

which can be summed for all subdomain-patches to arrive at

$$\|\eta_h * \nabla_h v_h\|^2_{\Omega_h \setminus \Omega} \lesssim \sum_{\Omega_j \in \mathcal{T}_h} \|\eta_h * \nabla_h v_h\|^2_{\Omega_{j,h}} \lesssim h^{s-1} \sum_{\Omega_j \in \mathcal{T}_h} \|\eta_h * \nabla v_h\|^2_{\tilde{\Omega}_j} \lesssim h^{s-1} \|\eta_h * \nabla v_h\|^2. \tag{51}$$

We can use (48) and (51) to complete the estimation in (47) as

$$\sup_{\eta_h * v_h \in \mathbb{P}_{h,\mathbf{k},s}} \frac{\langle \boldsymbol{\nu} \cdot \nabla u, \eta_h * v_h \rangle_{\partial\Omega}}{\|\eta_h * v_h\|_{1,\Omega_h}} \le \|u\|_{1,\Omega} \sup_{\eta_h * v_h \in \mathbb{P}_{h,\mathbf{k},s}} \frac{h^{s-\frac{1}{2}} \|\eta_h * \nabla v\|^2}{\|\nabla(\eta_h * v_h)\|_{\Omega_h}} \lesssim \|u\|_{1,\Omega} h^{s-\frac{1}{2}},$$

which together with (46) gives the estimate in the lemma. $\square$

Proof of Theorem 3: A triangle inequality and the estimates in Theorem 2 and Lemma 8 imply that

$$\begin{aligned} \|\nabla(u - \eta_h * u_{\mathrm{IP},s})\| &\lesssim \|\nabla(u - \eta_h * u_h)\| + \|\nabla(\eta_h * u_{\mathrm{IP},s} - \eta_h * u_h)\| \\ &\lesssim \inf_{v_h \in \mathbb{P}_{h,\mathbf{k}}} \|u - \eta_h * v_h\|_1 + \mathcal{O}(h^{s-\frac{1}{2}}) + \max_j h^d_{\Omega_j} \|\eta_h * g - g_0\|, \end{aligned}$$

as stated in the theorem. $\square$

Appendix

Following the notations in [32] we introduce the smooth function $\Phi : \mathbb{R}^d \to \mathbb{R}$ with

$$\Phi(\mathbf{x}) = \begin{cases} Ce^{\frac{1}{|\mathbf{x}|^2-1}} & \text{if } |\mathbf{x}| < 1 \\ 0 \quad \text{if } |\mathbf{x}| \geq 1 \end{cases} \qquad \text{and} \qquad \int_{B(\mathbf{0},1)} \Phi = 1$$

and define $\Phi_\delta : \mathbb{R}^d \to \mathbb{R}$ by $\Phi_\delta := \left(\frac{1}{\delta}\right)^d \Phi(\frac{\mathbf{x}}{\delta})$. Additionally, for $\mathbf{x} \in \mathbb{R}^d$ we use the notation $\Phi_{\delta,\mathbf{x}}$ for the function given simply by the

$$\Phi_{\delta,\mathbf{x}}(\mathbf{x} + \mathbf{y}) := \Phi_\delta(\mathbf{y}).$$

We use the following proposition; for the proof we refer to [32], pages 713–716.

Proposition 6. *For an arbitrary bounded Lipschitz domain U and parameter $p \in [1, \infty)$ the following statements are valid.*

(i) For $f \in C(U)$ we have $\lim\limits_{\delta \to 0} \Phi_\delta * f \to f$ uniformly.

(ii) For any $f \in L_{p,loc}(U)$ we have $\lim\limits_{\delta \to 0} \Phi_\delta * f \to f$ in $L_{p,loc}(U)$.

We also need the following statements.

Lemma 9. *If for all $\mathbf{x} \in K_1 \cup K_2$ we have the limit*

$$\lim_{\delta \to 0} \langle \eta_h * [\![u]\!]_f \, , \Phi_{\delta,\mathbf{x}} \rangle = \tilde{f}(\mathbf{x}) \tag{52}$$

then $\eta_h * [\![u]\!]_f$ can be identified with $\tilde{f}$. Also, for the function $\eta_h *_f \Phi_{\delta,\mathbf{x}} : \mathcal{F} \to \mathbb{R}$ given by

$$f \ni \mathbf{y} \to \int_{\mathbb{R}^d} \eta_h(\mathbf{y} - \mathbf{z}) \Phi_{\delta,\mathbf{x}}(\mathbf{z}) \, \mathrm{d}\mathbf{z}$$

we have the convergence

$$\lim_{\delta \to 0} \eta_h *_f \Phi_{\delta,\mathbf{x}} = \eta_h(-\mathbf{x} + \cdot) \quad \text{in } L_1(f). \tag{53}$$

Proof: We first note that

$$\langle \eta_h * [\![u]\!]_f \, , \Phi_{\delta,\mathbf{x}} \rangle = \int_{\mathbb{R}^d} \eta_h * [\![u]\!]_f \, (\mathbf{x} + \mathbf{y}) \Phi_{\delta,\mathbf{x}}(\mathbf{x} + \mathbf{y}) \, \mathrm{d}\mathbf{y} = \int_{\mathbb{R}^d} \eta_h * [\![u]\!]_f \, (\mathbf{x} + \mathbf{y}) \Phi_\delta(\mathbf{y}) \, \mathrm{d}\mathbf{y}$$
$$= \int_{\mathbb{R}^d} \eta_h * [\![u]\!]_f \, (\mathbf{x} + \mathbf{y}) \Phi_\delta(-\mathbf{y}) \, \mathrm{d}\mathbf{y} = \Phi_\delta * \eta_h * [\![u]\!]_f \, (\mathbf{x}).$$

In this way, according to property (ii) we can rewrite the condition in (52) as

$$\lim_{\delta\to 0} \Phi_\delta * \eta_h * [\![u]\!]_f \to \tilde{f} \quad \text{in } L_{1,\text{loc}}(\mathbb{R}^d). \tag{54}$$

Using the property (i) above, the fact that $\eta_h * [\![u]\!]_f$ is locally integrable and the limit in (54), we have that for each function $g \in C_0^\infty(\Omega)$ the following equality is valid:

$$\langle \eta_h * [\![u]\!]_f, g\rangle = (\eta_h * [\![u]\!]_f, g) = \lim_{\delta\to 0}(\eta_h * [\![u]\!]_f, \Phi_\delta * g) = \lim_{\delta\to 0}(\Phi_\delta * \eta_h * [\![u]\!]_f, g) = (\tilde{f}, g),$$

which proves the first statement of the lemma.

To prove the second statement we rewrite $\eta_h *_f \Phi_{\delta,\mathbf{x}}$ as

$$\eta_h *_f \Phi_{\delta,\mathbf{x}}(\mathbf{y}) = \int_{\mathbb{R}^d} \eta_h(\mathbf{y}-\mathbf{z})\Phi_\delta(-\mathbf{x}+\mathbf{z})\,\mathrm{d}\mathbf{z} = \int_{\mathbb{R}^d} \eta_h(\mathbf{y}-\mathbf{x}-\mathbf{z})\Phi_\delta(\mathbf{z})\,\mathrm{d}\mathbf{z}.$$

Accordingly, we have the pointwise convergence

$$\eta_h *_f \Phi_{\delta,\mathbf{x}}(\mathbf{y}) \to \eta_h(\mathbf{y}-\mathbf{x}).$$

On the other hand,

$$\int_{\mathbb{R}^d} \eta_h(\mathbf{y}-\mathbf{x}-\mathbf{z})\Phi_\delta(\mathbf{z}) \le \max_{\mathbb{R}^d} |\eta_h|$$

so that the function $\max_{\mathbb{R}^d} |\eta_h| \cdot \mathbb{1} \in L_1(\mathcal{F})$ delivers an upper bound for each function $\eta_h *_f \Phi_{\delta,\mathbf{x}}$. The statement is therefore an obvious consequence of the Lebesgue dominant convergence theorem. □

Proof of Lemma 5: We compute $\eta_h * [\![u]\!]_f$ based on the first statement in Lemma 9. For this, we use Lemma 1 and (53), which implies for each $f \in \mathcal{F}$ the following:

$$\begin{aligned}\lim_{\delta\to 0}\langle \eta_h * [\![u]\!]_f, \Phi_{\delta,\mathbf{x}}\rangle &= \lim_{\delta\to 0}\langle [\![u]\!]_f, \eta_h * \Phi_{\delta,\mathbf{x}}\rangle = \lim_{\delta\to 0}\int_f [\![u]\!]_f(\mathbf{y})\eta_h *_f \Phi_{\delta,x}(\mathbf{y})\,\mathrm{d}\mathbf{y} \\ &= \int_f [\![u]\!]_f(\mathbf{y}) \lim_{\delta\to 0}\eta_h *_f \Phi_{\delta,x}(\mathbf{y})\,\mathrm{d}\mathbf{y} = \int_f [\![u]\!]_f(\mathbf{y})\eta_h(\mathbf{y}-\mathbf{x})\,\mathrm{d}\mathbf{y},\end{aligned}$$

the sum of which gives the statement in the lemma. □

Proof of Proposition 3 We first prove (21). According to Lemma 1 and the consecutive remark, we have obviously that

$$\int_{f_0} |\,[\![v]\!]\,| \le |v|_{\text{BV}}, \tag{55}$$

where the BV seminorm is taken on $\overline{K_+ \cup K_-}$. For the next step we use a scaling argument and introduce the function space

$$\bar{\mathbb{P}}_K := \mathbb{P}_{h,\mathbf{k}}|_{K_+\cup K_-} / \langle \mathbb{1}\rangle,$$

which is the restriction of $\mathbb{P}_{h,\mathbf{k}}$ to $K_+ \cup K_-$ factorized with the constant functions. The BV seminorm on this function space becomes a norm, and accordingly, we use the notation $\|\cdot\|_{\mathrm{BV}}$. We next prove that for all $\epsilon > 0$ there is $h_0 > 0$ such that for all $h < h_0$ and $v \in \bar{\mathbb{P}}_K$ we have

$$\|\eta_h * v - v\|_{\mathrm{BV}} < \epsilon \|v\|_{\mathrm{BV}}. \tag{56}$$

For this we consider a normed basis $\{v_1, v_2, \ldots, v_D\}$ of $v \in \bar{\mathbb{P}}_K$ with respect to the BV norm and define the Euclidean norm $\|\cdot\|_E$ generated by this basis such that

$$\left| \sum_{j=1}^{D} a_j v_j \right|_E^2 = \sum_{j=1}^{D} a_j^2. \tag{57}$$

This norm should be equivalent with the BV norm, i.e. there is a constant c_0 with

$$\|v\|_E \leq c_0 \|v\|_{\mathrm{BV}} \quad \forall\, v \in \bar{\mathbb{P}}_{K_0}. \tag{58}$$

Note that this constant should be not the same for all pair of neighboring subdomains, but it is a continuous function of the position of the vertices. In particular, if we fix the edge f_0 of length *one*, then f_0 is fixed for $d = 2$ and for $d = 3$ the remaining vertex should be in a compact set if the condition of non-degeneracy holds true. Therefore, the constant c_0 has a finite maximum. Similarly, if we fix now an arbitrary interelement face chosen above the remaining node of K_{0-} and K_{0+} can lie in a compact set. In this way, for each pair of neighboring subdomains with at least one interelement edge of length *one* there is a uniform constant c_0 in (58). Also, since we have a finite basis, and η_h is a Dirac series, there is h_0 such that for all $h < h_0$ we have

$$\|\eta_h * v_j - v_j\|_{\mathrm{BV}} \leq \frac{\epsilon}{c_0 D} \|v_j\|_{\mathrm{BV}} = \frac{\epsilon}{c_0 \sqrt{D}} \quad \forall\, j \in \{1, 2, \ldots, D\}. \tag{59}$$

We obtain also here that (59) is valid for all pair of neighboring subdomains with at least one interelement edge of length *one* with a uniform parameter h_0. Then using (59), (57) and (58) we have that for any $0 < h < h_0$ and $v = \sum_{j=1}^{D} a_j v_j \in \bar{\mathbb{P}}_K$ the following inequality is valid:

$$\begin{aligned} \|\eta_h * v - v\|_{\mathrm{BV}} &= \| \sum_{j=1}^{D} \eta_h * v_j - v_j \|_{\mathrm{BV}} \leq \sum_{j=1}^{D} \|\eta_h * a_j v_j - a_j v_j\|_{\mathrm{BV}} \\ &\leq \frac{\epsilon}{c_0 \sqrt{D}} \sum_{j=1}^{D} |a_j| \leq \frac{\epsilon}{c_0} \sqrt{\sum_{j=1}^{D} |a_j|^2} = \frac{\epsilon}{c_0} \|v\|_E \leq \epsilon \|v\|_{\mathrm{BV}}, \end{aligned}$$

which proves the inequality in (56). Consequently, we also have

$$(1 - \epsilon)\|v\|_{\mathrm{BV}} \leq \|v - \eta_h * v\|_{\mathrm{BV}} + \|\eta_h * v\|_{\mathrm{BV}} - \epsilon \|v\|_{\mathrm{BV}} \leq \|\eta_h * v\|_{\mathrm{BV}}.$$

In the last step we relate (55) and (56) and use that $\eta_h * v$ is differentiable to obtain

$$\int_{f_0} [\![v]\!] \leq \|v\|_{\mathrm{BV}} \leq \frac{1}{1-\epsilon}\|\eta_h * v\|_{\mathrm{BV}} = \frac{1}{1-\epsilon}\int_{K_+\cup K_-} |\nabla(\eta_h * v)|. \tag{60}$$

To prove the statement of the lemma for two arbitrary neighboring subdomains Ω_+ and Ω_- we use (60) and the equalities in (5) which give

$$\begin{aligned}\int_{f_\Omega} [\![v]\!] &= h_\Omega^{d-1}\int_{f_0} [\![v_0]\!] \lesssim h_\Omega^{d-1}\int_{K_+\cup K_-} |\nabla(\eta_{h_0} * v_0)| \\ &= h_\Omega^{-d} h_\Omega^{d-1}\int_{\Omega_+\cup\Omega_-} h_\Omega^1 |\nabla(\eta_{h_0 h_\Omega^{\frac{1}{s}}} * v)| = \int_{\Omega_+\cup\Omega_-} |\nabla(\eta_{h_0 h_\Omega^{\frac{1}{s}}} * v)|.\end{aligned} \tag{61}$$

The inequality remains true if the lower index $h_0 h_\Omega^{\frac{1}{s}}$ is changed to a smaller one since this is equivalent with the choice of a smaller index h_0. Obviously the condition $h^{1-\frac{1}{s}} < h_0$ implies $h_0 h_\Omega^{\frac{1}{s}} > h$ and using (61) with the previous remark gives that

$$\int_{f_\Omega} [\![v]\!] \leq \int_{\Omega_+\cup\Omega_-} |\nabla(\eta_h * v)|$$

as stated in the inequality (21).

To prove (22), we again use the geometric setup shown in Figure 1. With this we obtain

$$\begin{aligned}&\|\eta_h * \nabla u\|_{h_\Omega\cdot K_+\cup h_\Omega\cdot K_-}\left(\int_{h_\Omega\cdot f_0 \pm h^s} 1\right)^{\frac{1}{2}} \geq \|\eta_h * \nabla u\|_{h_\Omega\cdot f_0\pm h^s}\left(\int_{h_\Omega\cdot f_0\pm h^s} 1\right)^{\frac{1}{2}} \\ &= \|\nabla(\eta_h * u)\|_{h_\Omega\cdot f_0\pm h^s}\left(\int_{h_\Omega\cdot f_0\pm h^s} 1\right)^{\frac{1}{2}} \geq \int_{h_\Omega\cdot f_0\pm h^s} |\nabla(\eta_h * u)| \\ &\gtrsim \int_{h_\Omega\cdot f_0}\left|\int_{-h^s}^{h^s} \nabla(\eta_h * u)(x,\mathbf{y})\,\mathrm{d}x\right|\,\mathrm{d}\mathbf{y} = \int_{h_\Omega\cdot f_0} |(\eta_h * u)(h^s,\mathbf{y}) - (\eta_h * u)(-h^s,\mathbf{y})|\,\mathrm{d}\mathbf{y} \\ &\geq \int_{h_\Omega\cdot f_0} |u(h^s,\mathbf{y}) - u(-h^s,\mathbf{y})| - |(\eta_h * u)(h^s,\mathbf{y}) - u(h^s,\mathbf{y})| - |u(-h^s,\mathbf{y}) - (\eta_h * u)(-h^s,\mathbf{y})|\,\mathrm{d}\mathbf{y}.\end{aligned} \tag{62}$$

To continue with the estimate we note that u is differentiable twice in $B((-h^s,\mathbf{y}),h^s)$ and according to (42) we have

$$\begin{aligned}&|(\eta_h * u)(h^s,\mathbf{y}) - u(h^s,\mathbf{y})\,\mathrm{d}\mathbf{y}| \leq \frac{1}{B_{h^s,d}}\cdot\frac{1}{2}\max_{h_\Omega\cdot K_-}|\nabla^2 u|\int_{B(0,h^s)} |\mathbf{s}|^2\,\mathrm{d}\mathbf{s} \\ &\lesssim \max_{h_\Omega\cdot K_-}|\nabla^2 u| h^{-sd} h^{s(d+2)} = \max_{h_\Omega\cdot K_-}|\nabla^2 u| h^{2s}.\end{aligned}$$

Therefore, using (62) we have

$$\|\eta_h * \nabla u\|_{h_\Omega\cdot K_+\cup h_\Omega\cdot K_-}\left(\int_{h_\Omega\cdot f_0\pm h^s} 1\right)^{\frac{1}{2}} \gtrsim \int_{h_\Omega\cdot f_0} |[\![u]\!]| - h^{2s}\lambda(h_\Omega\cdot f_0)\max_{h_\Omega\cdot K_+\cup h_\Omega\cdot K_-}|\nabla^2 u|,$$

which can be rewritten with the aid of (12) and the condition $s \geq \frac{3}{2}$ as

$$\int_{h \cdot f_0} |[\![u]\!]| \lesssim h^{2s} h_\Omega^{d-1} \max_{h \cdot K_+ \cup h \cdot K_-} |\nabla^2 u| + h^{\frac{s}{2}} h_\Omega^{\frac{d}{2}-\frac{1}{2}} \|\eta_h * \nabla u\|_{h_\Omega \cdot K_+ \cup h_\Omega \cdot K_-}$$
$$\lesssim h^{2s} h_\Omega^{d-1} h_\Omega^{-\frac{d}{2}-1} \|\nabla u\|_{h_\Omega \cdot K_+ \cup h_\Omega \cdot K_-} + h^{\frac{s}{2}} h_\Omega^{\frac{d}{2}-\frac{1}{2}} \|\eta_h * \nabla u\|_{h_\Omega \cdot K_+ \cup h_\Omega \cdot K_-}$$
$$\lesssim h_\Omega^{\frac{d}{2}} (h^{\frac{s}{2}} h_\Omega^{-\frac{1}{2}} + h^{2s} h_\Omega^2) \|\eta_h * \nabla u\|_{h \cdot K_+ \cup h \cdot K_-} \lesssim h^{\frac{s}{2}} h_\Omega^{\frac{d}{2}-\frac{1}{2}} \|\eta_h * \nabla u\|_{h \cdot K_+ \cup h \cdot K_-}$$

A summation with respect to the faces gives then the desired inequality. □

Acknowledgments

The author acknowledges the support of the Hungarian Research Fund OTKA (grants PD10441 and K104666). The author is also supported by the János Bolyai Research Fellowship of the Hungarian Academy of Sciences.

References

[1] G. A. Baker, Finite element methods for elliptic equations using nonconforming elements, Math. Comp. 31 (137) (1977) 45–59.

[2] C. Dawson, E. J. Kubatko, J. J. Westerink, C. Trahan, C. Mirabito, C. Michoski, N. Panda, Discontinuous Galerkin methods for modeling hurricane storm surge, Adv. Water Res. 34 (9) (2011) 1165–1176.

[3] Discontinuous galerkin methods for computational aerodynamics. 3D adaptive flow simulation with the DLR PADGE code, Aerosp. Sci. Technol. 14 (7) (2010) 512 – 519. doi:http://dx.doi.org/10.1016/j.ast.2010.04.002.

[4] J. Tago, V. M. Cruz-Atienza, J. Virieux, V. Etienne, F. J. Sánchez-Sesma, A 3D *hp*-adaptive discontinuous Galerkin method for modeling earthquake dynamics, J. GEOPHYS. RES.-SOL. EA 117 (B9).

[5] D. N. Arnold, F. Brezzi, B. Cockburn, L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal. 39 (5) (2001/02) 1749–1779 (electronic).

[6] O. A. Karakashian, F. Pascal, A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems, SIAM J. Numer. Anal. 41 (6) (2003) 2374–2399.

[7] P. Houston, D. Schötzau, T. P. Wihler, Energy norm a posteriori error estimation of the hp-adaptive discontinuous Galerkin methods for elliptic problems, Math. Mod. Meth. Appl. Sci. 17 (01) (2007) 33–62. doi:10.1142/S0218202507001826.

[8] B. Cockburn, C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, SIAM J. Numer. Anal. 35 (6) (1998) 2440–2463 (electronic). doi:10.1137/S0036142997316712.

[9] D. A. Di Pietro, A. Ern, Mathematical aspects of discontinuous Galerkin methods, Vol. 69 of Mathématiques & Applications (Berlin) [Mathematics & Applications], Springer, Heidelberg, 2012. doi:10.1007/978-3-642-22980-0.

[10] J. Hesthaven, T. Warburton, Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications, 1st Edition, Vol. 54 of Texts in Applied Mathematics, Springer, New York, 2007.

[11] B. Rivière, Discontinuous Galerkin methods for solving elliptic and parabolic equations, Vol. 35 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008, theory and implementation.

[12] M. Ainsworth, R. Rankin, Fully computable bounds for the error in nonconforming finite element approximations of arbitrary order on triangular elements, SIAM J. Numer. Anal. 46 (6) (2008) 3207–3232. doi:10.1137/07070838X.

[13] D. Sármány, F. Izsák, J. J. W. van der Vegt, Optimal penalty parameters for symmetric discontinuous Galerkin discretisations of the time-harmonic Maxwell equations, J. Sci. Comput. 44 (3) (2010) 219–254. doi:10.1007/s10915-010-9366-1.

[14] S. C. Brenner, L. Owens, L.-Y. Sung, A weakly over-penalized symmetric interior penalty method., Electron. Trans. Numer. Anal. 30 (2008) 107–127.

[15] S. C. Brenner, L. Owens, L.-Y. Sung, Higher order weakly over-penalized symmetric interior penalty methods, J. Comput. Appl. Math. 236 (11) (2012) 2883–2894. doi:10.1016/j.cam.2012.01.025.

[16] S. C. Brenner, L. Owens, A W-cycle algorithm for a weakly over-penalized interior penalty method, Comput. Methods Appl. Mech. Engrg. 196 (37-40) (2007) 3823–3832. doi:10.1016/j.cma.2007.02.011.

[17] T. Gudi, A new error analysis for discontinuous finite element methods for linear elliptic problems, Math. Comp. 79 (272) (2010) 2169–2189. doi:10.1090/S0025-5718-10-02360-4.

[18] A. Ern, J.-L. Guermond, Theory and Practice of Finite Elements, Vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004.

[19] A. Buffa, C. Ortner, Compact embeddings of broken Sobolev spaces and applications, IMA J. Numer. Anal. 29 (4) (2009) 827–855. doi:10.1093/imanum/drn038.

[20] D. A. Di Pietro, A. Ern, Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier-Stokes equations, Math. Comp. 79 (271) (2010) 1303–1330. doi:10.1090/S0025-5718-10-02333-1.

[21] G. Csörgő, F. Izsák, Energy norm error estimates for averaged discontinuous Galerkin methods in 1 dimension, Int. J. Numer. Anal. Model. 11 (3) (2014) 567–586.

[22] B. Cockburn, M. Luskin, C.-W. Shu, E. Süli, Enhanced accuracy by post-processing for finite element methods for hyperbolic equations, Math. Comp. 72 (242) (2003) 577–606 (electronic). doi:10.1090/S0025-5718-02-01464-3.

[23] J. King, H. Mirzaee, J. K. Ryan, R. M. Kirby, Smoothness-Increasing Accuracy-Conserving (SIAC) Filtering for Discontinuous Galerkin Solutions: Improved Errors Versus Higher-Order Accuracy, J. Sci. Comput. 53 (2012) 129–149. doi:10.4208/jcm.2009.27.5.013.

[24] L. Ji, Y. Xu, J. K. Ryan, Accuracy-enhancement of discontinuous Galerkin solutions for convection-diffusion equations in multiple-dimensions, Math. Comp. 81 (280) (2012) 1929–1950 (electronic).

[25] H. Mirzaee, J. King, J. Ryan, R. Kirby, Smoothness-increasing accuracy-conserving filters for discontinuous Galerkin solutions over unstructured triangular meshes, SIAM J. Sci. Comput. 35 (1) (2013) A212–A230. doi:10.1137/120874059.

[26] H. Mirzaee, J. K. Ryan, R. M. Kirby, Smoothness-increasing accuracy-conserving (SIAC) filters for discontinuous Galerkin J. Sci. Comput. 58 (3) (2014) 690–704. doi:10.1007/s10915-013-9748-2.
URL http://dx.doi.org/10.1007/s10915-013-9748-2

[27] M. A. Peletier, R. Planqué, M. Röger, Sobolev regularity via the convergence rate of convolutions and Jensen's inequality, Ann. Sc. Norm. Super. Pisa Cl. Sci. 6 (4) (2007) 499–510.

[28] P. R. Halmos, Measure Theory, D. Van Nostrand Company, Inc., New York, N. Y., 1950.

[29] T. Titkos, A simple proof of the Lebesgue decomposition theorem, Amer. Math. MonthlyTo appear.

[30] H. Attouch, G. Buttazzo, G. Michaille, Variational analysis in Sobolev and BV spaces, Vol. 6 of MPS/SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006, applications to PDEs and optimization.

[31] F. Hirsch, G. Lacombe, Elements of Functional Analysis, Vol. 192 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1999, translated from the 1997 French original by Silvio Levy.

[32] L. C. Evans, Partial differential equations, 2nd Edition, Vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2010.