MATHEMATICS OF COMPUTATION
Volume 00, Number 0, Pages 000–000
S 0025-5718(XX)0000-0

hp-DGFEM FOR SECOND-ORDER MIXED ELLIPTIC PROBLEMS IN POLYHEDRA

DOMINIK SCHÖTZAU, CHRISTOPH SCHWAB, AND THOMAS P. WIHLER

Abstract. We prove exponential rates of convergence of hp-version discontinuous Galerkin (dG) interior penalty finite element methods for second-order elliptic problems with mixed Dirichlet-Neumann boundary conditions in axiparallel polyhedra. The dG discretizations are based on axiparallel, σ-geometric anisotropic meshes of mapped hexahedra and anisotropic polynomial degree distributions of μ-bounded variation. We consider piecewise analytic solutions which belong to a larger analytic class than those for the pure Dirichlet problem considered in [11, 12]. For such solutions, we establish the exponential convergence of a nonconforming dG interpolant given by local L^2-projections on elements away from corners and edges, and by suitable local low-order quasi-interpolants on elements at corners and edges. Due to the appearance of non-homogeneous, weighted norms in the analytic regularity class, new arguments are introduced to bound the dG consistency errors in elements abutting on Neumann edges. The non-homogeneous norms also entail some crucial modifications of the stability and quasi-optimality proofs, as well as of the analysis for the anisotropic interpolation operators. The exponential convergence bounds for the dG interpolant constructed in this paper generalize the results of [11, 12] for the pure Dirichlet case.

1. Introduction

Consider an open, bounded and axiparallel polyhedron $\Omega \subset \mathbb{R}^3$ with Lipschitz boundary $\Gamma = \partial\Omega$ that consists of a finite union of plane faces $\Gamma_{\boldsymbol{\iota}}$ indexed by $\boldsymbol{\iota} \in \mathcal{J}$. The faces $\Gamma_{\boldsymbol{\iota}}$ are assumed to be bounded, plane polygons whose sides form the (open) edges of Ω. The set $\{\Gamma_{\boldsymbol{\iota}}\}_{\boldsymbol{\iota}\in\mathcal{J}}$ is partitioned into a subset of Dirichlet faces $\{\Gamma_{\boldsymbol{\iota}}\}_{\boldsymbol{\iota}\in\mathcal{J}_D}$ and a subset of Neumann faces $\{\Gamma_{\boldsymbol{\iota}}\}_{\boldsymbol{\iota}\in\mathcal{J}_N}$, with corresponding (disjoint) index sets $\mathcal{J}_D$ and $\mathcal{J}_N$, respectively (i.e., $\mathcal{J} = \mathcal{J}_D \dot\cup \mathcal{J}_N$). Then we consider the diffusion equation

$$-\Delta u = f \qquad \text{in } \Omega, \tag{1.1}$$

$$\gamma_0(u) = 0 \qquad \text{on } \Gamma_{\boldsymbol{\iota}} \subset \partial\Omega, \quad \boldsymbol{\iota} \in \mathcal{J}_D, \tag{1.2}$$

$$\gamma_1(u) = 0 \qquad \text{on } \Gamma_{\boldsymbol{\iota}} \subset \partial\Omega, \quad \boldsymbol{\iota} \in \mathcal{J}_N, \tag{1.3}$$

where the operators γ_0 and γ_1 denote the trace and (co)normal derivative operators, respectively. With the Sobolev space $H^1_D(\Omega) := \{v \in H^1(\Omega) : v|_{\Gamma_{\boldsymbol{\iota}}} = 0,\ \boldsymbol{\iota} \in \mathcal{J}_D\}$

2010 *Mathematics Subject Classification.* 65N30.

Key words and phrases. hp-dGFEM, second-order elliptic problems in 3D polyhedra, mixed Dirichlet-Neumann boundary conditions, exponential convergence.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), the European Research Council AdG grant STAHDPDE 247277, and the Swiss National Science Foundation (SNF).

and the continuous bilinear form $a(u,v) := \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}\boldsymbol{x}$, the weak formulation of problem (1.1)–(1.3) is to find $u \in H^1_D(\Omega)$ such that

$$a(u,v) = \int_\Omega f v \, \mathrm{d}\boldsymbol{x} \qquad \forall v \in H^1_D(\Omega) \ . \tag{1.4}$$

For every $f \in H^1_D(\Omega)^\star$, the dual space of $H^1_D(\Omega)$, problem (1.4) admits a weak solution $u \in H^1_D(\Omega)$. The solution is unique if $\mathcal{J}_D \neq \emptyset$, and unique up to constants if $\mathcal{J}_D = \emptyset$ (in which case we also require the compatibility condition $\int_\Omega f \, \mathrm{d}\boldsymbol{x} = 0$).

This paper is a continuation of our work [11, 12] on *hp*-version discontinuous Galerkin (dG) finite element methods (FEM) for second-order elliptic boundary-value problems in polyhedral domains $\Omega \subset \mathbb{R}^3$. In [11], we showed the well-posedness, stability and consistency of *hp*-version interior penalty (IP) discontinuous Galerkin discretizations of (1.1) for the pure Dirichlet case, that is, for the case where $\mathcal{J} = \mathcal{J}_D$, $\mathcal{J}_N = \emptyset$, and the homogeneous essential boundary conditions (1.2) are posed on all of $\partial\Omega$. For axiparallel configurations, we then used these results in [12] to prove exponential rates of convergence in the number of degrees of freedom, for *hp*-dG discretizations on appropriate combinations of σ-geometric anisotropic meshes and $\mathfrak{s}$-linearly increasing anisotropic elemental polynomial degrees; see also [15] for related work on linear elasticity.

In this work, we consider and analyze *hp*-dG methods for the case $\mathcal{J}_N \neq \emptyset$. Although the *hp*-error analysis will be along the lines of [11, 12], there are significant differences. As shown in [3], the solutions of mixed Dirichlet-Neumann or pure Neumann problems for second-order, elliptic boundary-value problems in polyhedral domains with piecewise analytic data belong to countably normed Sobolev spaces $N^m_{\boldsymbol{\beta}}(\Omega)$ with *non-homogeneous weights*. In the case of homogeneous Dirichlet conditions (i.e., when $\mathcal{J}_N = \emptyset$), these spaces coincide with the (smaller) spaces $M^m_{\boldsymbol{\beta}}(\Omega)$ for which we proved exponential convergence in [12]. When $\mathcal{J}_N \neq \emptyset$, however, we have the *strict inclusion* $N^m_{\boldsymbol{\beta}}(\Omega) \supsetneq M^m_{\boldsymbol{\beta}}(\Omega)$, due to the different structure of the weights near Neumann edges (where two Neumann faces $\Gamma_{\boldsymbol{\iota}}$, $\boldsymbol{\iota} \in \mathcal{J}_N$, intersect). Compared to [12], this entails new technical difficulties, and requires *essential* modifications in the stability and consistency analyses, and in the choice of the anisotropic *hp*-interpolation operators.

We show that, for solutions to problem (1.1)–(1.3) belonging to the countably normed Sobolev spaces $N^m_{\boldsymbol{\beta}}(\Omega)$, the *hp*-dG approximations are well-defined and satisfy the Galerkin orthogonality property. Hence, the dG energy error can be bounded by suitable consistency terms involving a discontinuous elemental polynomial interpolation operator. The main result of this paper is the construction and analysis of a *non-conforming* dG *hp*-interpolant given by local L^2-projections on elements away from corners and edges, and by local low-order quasi-interpolants on elements at corners and edges, which allows us to bound the consistency terms at exponential rates of convergence. That is, we prove that *hp*-dGFEM achieves *exponential convergence* with respect to the dG energy error, i.e., asymptotic convergence rate bounds of the form $C\exp(-b\sqrt[5]{N})$, where N is the number of degrees of freedom, and where $b, C > 0$ are independent of N. An extensive numerical study of various aspects of these theoretical results will be presented in a forthcoming paper.

We point out that, although we use ideas and notation from [11, 12], the proof of exponential convergence in the present paper is self-contained, and that the results are in several respects stronger than the analysis in [12]: exponential convergence is shown for larger classes of solutions, and for a non-conforming dG interpolant which requires much less smoothness of the solutions than that in [12] (merely L^2-regularity for the L^2-projections, and $W^{1,1}$-regularity for the quasi-interpolants), thereby generalizing the analysis in [12] in the pure Dirichlet case, as well as providing an alternative proof for it. The main reason for using L^2-projections is that they allow us to separately analyze the errors in edge-perpendicular and edge-parallel directions, which is crucial in the appearance of Neumann boundary conditions. However, this is purchased at the expense of additional powers of the maximal polynomial degree (as compared to [12]) appearing in the consistency error bounds; these are subsequently absorbed into the exponentially small terms.

The outline of the article is as follows: In Section 2, we recapitulate analytic regularity results for solutions to (1.1)–(1.3) from [3] (which extend the pioneering work [2] in two dimensions to the three-dimensional case). In Section 3, we define *hp*-dG finite element spaces on *σ-geometric meshes* of hexahedral elements with possibly anisotropic and *$\mathfrak{s}$-linearly increasing polynomial degree distributions*. In Section 4, we focus on the dG discretizations and discuss their consistency and stability. In Section 5, we introduce the non-conforming dG interpolant which will be used in our analysis. In Section 6, we present dG-norm error estimates for this interpolant, and state our exponential convergence result (Theorem 6.2). Section 7 is devoted to the proof of this result.

The notation employed throughout this paper is consistent with [11, 12]. In particular, we shall frequently use the function

$$\Psi_{q,r} = \frac{\Gamma(q+1-r)}{\Gamma(q+1+r)}, \qquad 0 \le r \le q,\ q,r \in \mathbb{N}, \tag{1.5}$$

where Γ is the Gamma function satisfying $\Gamma(m+1) = m!$, for any $m \in \mathbb{N}$. Moreover, we shall use the notations "$\lesssim$" or "$\simeq$" to mean an inequality or an equivalence containing generic positive multiplicative constants which are independent of the discretization and regularity parameters, as well as of the geometric refinement level, but which may depend on the geometric refinement ratio σ and on the slope parameter $\mathfrak{s}$.

2. Regularity

In this section, we specify the regularity for solutions of (1.1)–(1.3). We follow [3], based on the notations already introduced in [11, 12].

2.1. **Subdomains and weights.** We denote by $\mathcal{C}$ the set of corners $\boldsymbol{c}$, and by $\mathcal{E}$ the set of (open) edges $\boldsymbol{e}$ of Ω. The singular set of Ω is then given by

$$\mathcal{S} := \left(\bigcup_{\boldsymbol{c}\in\mathcal{C}} \boldsymbol{c}\right) \cup \left(\bigcup_{\boldsymbol{e}\in\mathcal{E}} \boldsymbol{e}\right) \subset \Gamma. \tag{2.1}$$

For $\boldsymbol{c} \in \mathcal{C}$, $\boldsymbol{e} \in \mathcal{E}$, and $\boldsymbol{x} \in \Omega$, we define the following distance functions:

$$r_{\boldsymbol{c}}(\boldsymbol{x}) = |\boldsymbol{x} - \boldsymbol{c}|, \qquad r_{\boldsymbol{e}}(\boldsymbol{x}) = \inf_{\boldsymbol{y}\in\boldsymbol{e}} |\boldsymbol{x} - \boldsymbol{y}|, \qquad \rho_{\boldsymbol{ce}}(\boldsymbol{x}) = r_{\boldsymbol{e}}(\boldsymbol{x})/r_{\boldsymbol{c}}(\boldsymbol{x}). \tag{2.2}$$

As in [11, Section 2.1], the vertices of Ω are assumed to be separated. For each corner $\boldsymbol{c} \in \mathcal{C}$, we denote by $\mathcal{E}_{\boldsymbol{c}} := \{\, \boldsymbol{e} \in \mathcal{E} : \boldsymbol{c} \cap \overline{\boldsymbol{e}} \neq \emptyset \,\}$ the set of all edges of Ω which meet at $\boldsymbol{c}$. Similarly, for any $\boldsymbol{e} \in \mathcal{E}$, the set of corners of $\boldsymbol{e}$ is given by

$\mathcal{C}_{\boldsymbol{e}} := \{\, \boldsymbol{c} \in \mathcal{C} \,:\, \boldsymbol{c} \cap \overline{\boldsymbol{e}} \neq \emptyset \,\}$. Then, for $\varepsilon > 0$, $\boldsymbol{c} \in \mathcal{C}$, $\boldsymbol{e} \in \mathcal{E}$ respectively $\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}}$, we define the neighborhoods

$$
\begin{aligned}
\omega_{\boldsymbol{c}} &= \{\, \boldsymbol{x} \in \Omega \,:\, r_{\boldsymbol{c}}(\boldsymbol{x}) < \varepsilon \ \wedge\ \rho_{\boldsymbol{ce}}(\boldsymbol{x}) > \varepsilon \quad \forall \boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}} \,\}, \\
\omega_{\boldsymbol{e}} &= \{\, \boldsymbol{x} \in \Omega \,:\, r_{\boldsymbol{e}}(\boldsymbol{x}) < \varepsilon \ \wedge\ r_{\boldsymbol{c}}(\boldsymbol{x}) > \varepsilon \quad \forall \boldsymbol{c} \in \mathcal{C}_{\boldsymbol{e}} \,\}, \\
\omega_{\boldsymbol{ce}} &= \{\, \boldsymbol{x} \in \Omega \,:\, r_{\boldsymbol{c}}(\boldsymbol{x}) < \varepsilon \ \wedge\ \rho_{\boldsymbol{ce}}(\boldsymbol{x}) < \varepsilon \,\}.
\end{aligned} \tag{2.3}
$$

By choosing $\varepsilon > 0$ sufficiently small as in [11], the domain Ω can be partitioned into four *disjoint* subdomains, $\overline{\Omega} = \overline{\Omega_{\mathcal{C}} \,\dot\cup\, \Omega_{\mathcal{E}} \,\dot\cup\, \Omega_{\mathcal{C}\mathcal{E}} \,\dot\cup\, \Omega_0}$, referred to as *corner*, *edge*, *corner-edge* and *interior* neighborhoods of Ω, respectively, where

$$
\Omega_{\mathcal{C}} = \bigcup_{\boldsymbol{c} \in \mathcal{C}} \omega_{\boldsymbol{c}}, \qquad \Omega_{\mathcal{E}} = \bigcup_{\boldsymbol{e} \in \mathcal{E}} \omega_{\boldsymbol{e}}, \qquad \Omega_{\mathcal{C}\mathcal{E}} = \bigcup_{\boldsymbol{c} \in \mathcal{C}} \bigcup_{\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}}} \omega_{\boldsymbol{ce}}, \tag{2.4}
$$

and $\Omega_0 := \Omega \setminus \overline{\Omega_{\mathcal{C}} \cup \Omega_{\mathcal{E}} \cup \Omega_{\mathcal{C}\mathcal{E}}}$.

It will be useful to tag Dirichlet corners, as well as to distinguish Dirichlet and Neumann edges. To that end, we introduce the sets:

$$
\begin{aligned}
\mathcal{C}_D &:= \left\{\, \boldsymbol{c} \in \mathcal{C} \,:\, \exists \boldsymbol{\iota} \in \mathcal{J}_D \text{ with } \boldsymbol{c} \cap \overline{\Gamma}_{\boldsymbol{\iota}} \neq \emptyset \,\right\}, \\
\mathcal{E}_D &:= \left\{\, \boldsymbol{e} \in \mathcal{E} \,:\, \exists \boldsymbol{\iota} \in \mathcal{J}_D \text{ with } \boldsymbol{e} \cap \overline{\Gamma}_{\boldsymbol{\iota}} \neq \emptyset \,\right\},
\end{aligned} \tag{2.5}
$$

and set $\mathcal{E}_N := \mathcal{E} \setminus \mathcal{E}_D$. Corners in $\mathcal{C}_D$ and edges in $\mathcal{E}_D$ abut at at least one Dirichlet face $\Gamma_{\boldsymbol{\iota}}$ for $\boldsymbol{\iota} \in \mathcal{J}_D$. Note that we possibly have $\mathcal{E}_N = \emptyset$. Hence, the edge neighborhood $\Omega_{\mathcal{E}}$ in (2.4) can be further partitioned into:

$$
\Omega_{\mathcal{E}} = \Omega_{\mathcal{E}_D} \,\dot\cup\, \Omega_{\mathcal{E}_N}, \qquad \text{with} \quad \Omega_{\mathcal{E}_D} = \bigcup_{\boldsymbol{e} \in \mathcal{E}_D} \omega_{\boldsymbol{e}}, \quad \Omega_{\mathcal{E}_N} = \bigcup_{\boldsymbol{e} \in \mathcal{E}_N} \omega_{\boldsymbol{e}}. \tag{2.6}
$$

2.2. **Weighted Sobolev spaces.** To each $\boldsymbol{c} \in \mathcal{C}$ and $\boldsymbol{e} \in \mathcal{E}$ we associate a corner and an edge exponent $\beta_{\boldsymbol{c}}, \beta_{\boldsymbol{e}} \in \mathbb{R}$, respectively. We collect these quantities in the weight exponent vector $\boldsymbol{\beta} = \{\beta_{\boldsymbol{c}} \,:\, \boldsymbol{c} \in \mathcal{C}\} \cup \{\beta_{\boldsymbol{e}} \,:\, \boldsymbol{e} \in \mathcal{E}\} \in \mathbb{R}^{|\mathcal{C}|+|\mathcal{E}|}$. Inequalities of the form $\boldsymbol{\beta} < 1$ and expressions like $\boldsymbol{\beta} \pm s$, where $s \in \mathbb{R}$, are to be understood componentwise. We shall often use the notation

$$
b_{\boldsymbol{c}} := -1 - \beta_{\boldsymbol{c}}, \quad \boldsymbol{c} \in \mathcal{C}\,, \qquad b_{\boldsymbol{e}} := -1 - \beta_{\boldsymbol{e}}, \quad \boldsymbol{e} \in \mathcal{E}\,. \tag{2.7}
$$

To review the analytic regularity results of [3] for solutions to (1.1)–(1.3), we choose local coordinate systems in $\omega_{\boldsymbol{e}}$ and $\omega_{\boldsymbol{ce}}$, for $\boldsymbol{e} \in \mathcal{E}$ respectively $\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}}$, such that the edge $\boldsymbol{e}$ corresponds to the direction $(0,0,1)$. Then, we indicate quantities transversal to $\boldsymbol{e}$ by $(\cdot)^{\perp}$, and quantities parallel to $\boldsymbol{e}$ by $(\cdot)^{\|}$. In particular, if $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{N}_0^3$ is a multi-index of order $|\boldsymbol{\alpha}| = \alpha_1 + \alpha_2 + \alpha_3$, then we write $\boldsymbol{\alpha} = (\boldsymbol{\alpha}^{\perp}, \alpha^{\|})$ with $\boldsymbol{\alpha}^{\perp} = (\alpha_1, \alpha_2)$ and $\alpha^{\|} = \alpha_3$, and denote the partial derivative operator $\mathsf{D}^{\boldsymbol{\alpha}}$ by $\mathsf{D}^{\boldsymbol{\alpha}} = \mathsf{D}_{\perp}^{\boldsymbol{\alpha}^{\perp}} \mathsf{D}_{\|}^{\alpha^{\|}}$, where $\mathsf{D}_{\perp}^{\boldsymbol{\alpha}^{\perp}}$ and $\mathsf{D}_{\|}^{\alpha^{\|}}$ signify the derivative operators in the perpendicular and parallel directions, respectively. We also denote by $\mathsf{D}_{\perp}$ and $\mathsf{D}_{\perp}^2$ the gradient and the Hessian operator in edge-perpendicular direction, respectively, and set $\mathsf{D}_{\|} = \mathsf{D}_{\|}^1$.

The solution u of (1.1)–(1.3) belongs to a scale $N_{\boldsymbol{\beta}}^m(\Omega)$ of countably normed spaces which are, in the case $\mathcal{J}_N \neq \emptyset$ under consideration here, strictly larger than the scale $M_{\boldsymbol{\beta}}^m(\Omega)$ of spaces considered in [12] for the pure Dirichlet case (i.e., for $\mathcal{J} = \mathcal{J}_D$). For $k \geq 0$, we define the semi-norm

$$
\begin{aligned}
|u|^2_{N^k_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)} := \sum_{|\boldsymbol{\alpha}|=k} \Big\{ &\|\mathsf{D}^{\boldsymbol{\alpha}} u\|^2_{L^2(\Omega_0)} \\
&+ \sum_{\boldsymbol{c}\in\mathcal{C}_D} \left\| r_{\boldsymbol{c}}^{\beta_{\boldsymbol{c}}+|\boldsymbol{\alpha}|} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\omega_{\boldsymbol{c}})} + \sum_{\boldsymbol{c}\in\mathcal{C}\setminus\mathcal{C}_D} \left\| r_{\boldsymbol{c}}^{\max\{\beta_{\boldsymbol{c}}+|\boldsymbol{\alpha}|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\omega_{\boldsymbol{c}})} \\
&+ \sum_{\boldsymbol{e}\in\mathcal{E}_D} \left\| r_{\boldsymbol{e}}^{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\omega_{\boldsymbol{e}})} + \sum_{\boldsymbol{e}\in\mathcal{E}_N} \left\| r_{\boldsymbol{e}}^{\max\{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\omega_{\boldsymbol{e}})} \\
&+ \sum_{\boldsymbol{c}\in\mathcal{C}_D} \sum_{\boldsymbol{e}\in\mathcal{E}_{\boldsymbol{c}}\cap\mathcal{E}_D} \left\| r_{\boldsymbol{c}}^{\beta_{\boldsymbol{c}}+|\boldsymbol{\alpha}|} \rho_{\boldsymbol{ce}}^{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\omega_{\boldsymbol{ce}})} \\
&+ \sum_{\boldsymbol{c}\in\mathcal{C}_D} \sum_{\boldsymbol{e}\in\mathcal{E}_{\boldsymbol{c}}\cap\mathcal{E}_N} \left\| r_{\boldsymbol{c}}^{\beta_{\boldsymbol{c}}+|\boldsymbol{\alpha}|} \rho_{\boldsymbol{ce}}^{\max\{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\omega_{\boldsymbol{ce}})} \\
&+ \sum_{\boldsymbol{c}\in\mathcal{C}\setminus C_D} \sum_{\boldsymbol{e}\in\mathcal{E}_{\boldsymbol{c}}\cap\mathcal{E}_D} \left\| r_{\boldsymbol{c}}^{\max\{\beta_{\boldsymbol{c}}+|\boldsymbol{\alpha}|,0\}} \rho_{\boldsymbol{ce}}^{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\omega_{\boldsymbol{ce}})} \\
&+ \sum_{\boldsymbol{c}\in\mathcal{C}\setminus\mathcal{C}_D} \sum_{\boldsymbol{e}\in\mathcal{E}_{\boldsymbol{c}}\cap\mathcal{E}_N} \left\| r_{\boldsymbol{c}}^{\max\{\beta_{\boldsymbol{c}}+|\boldsymbol{\alpha}|,0\}} \rho_{\boldsymbol{ce}}^{\max\{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\omega_{\boldsymbol{ce}})} \Big\}.
\end{aligned} \tag{2.8}
$$

For $m > k_{\boldsymbol{\beta}}$, with

$$k_{\boldsymbol{\beta}} := -\min\{\min_{\boldsymbol{c}\in\mathcal{C}} \beta_{\boldsymbol{c}}, \min_{\boldsymbol{e}\in\mathcal{E}} \beta_{\boldsymbol{e}}\}, \tag{2.9}$$

we write $N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)$ for the space of functions u such that $\|u\|_{N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)} < \infty$, with the norm $\|u\|^2_{N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)} := \sum_{k=0}^m |u|^2_{N^k_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)}$. For subdomains $K \subseteq \Omega$ we shall denote by $|\cdot|_{N^k_{\boldsymbol{\beta}}(K;\mathcal{C}_D,\mathcal{E}_D)}$ the semi-norm (2.8) with all domains of integration replaced by their intersections with $K \subseteq \Omega$, and likewise we shall use the norm $\|\cdot\|_{N^m_{\boldsymbol{\beta}}(K;\mathcal{C}_D,\mathcal{E}_D)}$.

The spaces $N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)$ are monotonic with respect to the sets $\mathcal{C}_D$, $\mathcal{E}_D$: for $\emptyset \subseteq \mathcal{C}_D \subseteq \mathcal{C}$ and $\emptyset \subseteq \mathcal{E}_D \subseteq \mathcal{E}$, we have

$$M^m_{\boldsymbol{\beta}}(\Omega) := N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C},\mathcal{E}) \subseteq N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D) \subseteq N^m_{\boldsymbol{\beta}}(\Omega;\emptyset,\emptyset) =: N^m_{\boldsymbol{\beta}}(\Omega), \tag{2.10}$$

where $M^m_{\boldsymbol{\beta}}(\Omega)$ is the weighted Sobolev space obtained as the closure of $C_0^\infty(\Omega)$ with respect to the norm $\|\cdot\|_{M^m_{\boldsymbol{\beta}}(\Omega)} = \|\cdot\|_{N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C},\mathcal{E})}$.

2.3. **Regularity of weak solutions.** We adopt the following classes of analytic functions from [3].

Definition 2.1. For subdomains $K \subseteq \Omega$ and subsets $\emptyset \subseteq \mathcal{C}' \subseteq \mathcal{C}$, $\emptyset \subseteq \mathcal{E}' \subseteq \mathcal{E}$, the space $B_{\boldsymbol{\beta}}(K;\mathcal{C}',\mathcal{E}')$ consists of all functions u such that $u \in N^m_{\boldsymbol{\beta}}(K;\mathcal{C}',\mathcal{E}')$ for $m > k_{\boldsymbol{\beta}}$, with $k_{\boldsymbol{\beta}}$ as in (2.9), and such that there exists a constant $C_u > 0$ with the property that $|u|_{N^k_{\boldsymbol{\beta}}(K;\mathcal{C}',\mathcal{E}')} \leq C_u^{k+1} k!$ for all $k > k_{\boldsymbol{\beta}}$.

Remark 2.2. The analytic class $B_{\boldsymbol{\beta}}(\Omega) = B_{\boldsymbol{\beta}}(\Omega;\emptyset,\emptyset)$ is closely related to the countably normed spaces $B^\ell_\beta(\Omega)$ introduced by Babuška and Guo in [2, 7, 8]: if the edge and corner exponents $\beta_{ij} \in (0,1)$ and $\beta_m \in (0, 1/2)$ introduced in [2, 7, 8] satisfy $\beta_{ij} = \beta_{\boldsymbol{e}} + \ell$ and $\beta_m = \beta_{\boldsymbol{c}} + \ell$ for every $\boldsymbol{c} \in \mathcal{C}$ and $\boldsymbol{e} \in \mathcal{E}$, then $B^\ell_\beta(\Omega) = B_{\boldsymbol{\beta}}(\Omega)$. By (2.10), we also have $A_{\boldsymbol{\beta}}(\Omega) = B_{\boldsymbol{\beta}}(\Omega;\mathcal{C},\mathcal{E})$, where $A_{\boldsymbol{\beta}}(\Omega)$ is the analytic class considered in [12] for the pure Dirichlet problem; see also [3].

We have the following regularity result (see [3, Theorem 7.3]).

Proposition 2.3. *There are bounds* $b_{\mathcal{E}}, b_{\mathcal{C}} > 0$ *(depending on* Ω *and on the space* $H^1_D(\Omega)$*) such that, for* $\boldsymbol{b}$ *satisfying*

$$0 < b_{\boldsymbol{c}} < b_{\mathcal{C}}, \quad 0 < b_{\boldsymbol{e}} < b_{\mathcal{E}}, \qquad \boldsymbol{c} \in \mathcal{C}, \ \boldsymbol{e} \in \mathcal{E}, \tag{2.11}$$

any weak solution $u \in H^1_D(\Omega)$ *defined* (1.4) *of problem* (1.1)–(1.3) *satisfies:*

$$f \in B_{1-\boldsymbol{b}}(\Omega; \mathcal{C}, \mathcal{E}_D) \implies u \in B_{-1-\boldsymbol{b}}(\Omega; \mathcal{C}, \mathcal{E}_D)\ . \tag{2.12}$$

Remark 2.4. In the analytic regularity (2.12) all the corner weights have the same structure as in the pure Dirichlet case, even if $\boldsymbol{c} \in \mathcal{C} \setminus \mathcal{C}_D$ is a "Neumann corner" (where only Neumann faces meet in $\mathcal{E}_N$). This also means that in the analytic class $B_{-1-\boldsymbol{b}}(\Omega; \mathcal{C}, \mathcal{E}_D)$ only six out of the nine terms in the weighted semi-norms in (2.8) suffice to characterize the regularity of u (since $\mathcal{C} \setminus \mathcal{C}_D = \emptyset$ in $B_{-1-\boldsymbol{b}}(\Omega; \mathcal{C}, \mathcal{E}_D)$). Corner weights do *not* imply homogeneous Dirichlet boundary conditions since by Hardy's inequality $\left\{u \in H^1(\Omega) : \ r_{\boldsymbol{c}}^{-1} u \in L^2(\Omega) \ \forall \boldsymbol{c} \in \mathcal{C}\right\} = H^1(\Omega)$ for bounded Lipschitz domains Ω in $\mathbb{R}^3$. For edges $\boldsymbol{e} \in \mathcal{E}$, the two cases $\boldsymbol{e} \in \mathcal{E}_D$ and $\boldsymbol{e} \notin \mathcal{E}_D$ must be distinguished. An inspection of the terms in the semi-norm (2.8) reveals that the assumptions in Remark 2.5 force the solution to zero weakly *at Dirichlet edges* $\boldsymbol{e} \in \mathcal{E}_D$. On the other hand, the structure of the weights $r_{\boldsymbol{e}}^{\max\{\beta_{\boldsymbol{e}} + |\boldsymbol{\alpha}^\perp|, 0\}}$ associated with *Neumann* edges $\boldsymbol{e} \in \mathcal{E}_N$ in the fifth and seventh terms in (2.8) allows for nonzero traces of $u \in H^1_D(\Omega)$ at such edges. Indeed, by taking $|\boldsymbol{\alpha}^\perp| = 0$ and recalling that $\beta_{\boldsymbol{e}} < -1$ by Remark 2.5, we see that $r_{\boldsymbol{e}}^{\max\{\beta_{\boldsymbol{e}} + |\boldsymbol{\alpha}^\perp|, 0\}} \equiv 1$, and thus, no restriction on $u \in H^1_D(\Omega)$ is imposed by this weight function along the associated edge $\boldsymbol{e}$.

Remark 2.5. In the following and without loss of generality, we may and will assume that in (2.11) there holds $0 < b_{\boldsymbol{c}}, \ b_{\boldsymbol{e}} < 1$ for $\boldsymbol{c} \in \mathcal{C}, \ \boldsymbol{e} \in \mathcal{E}$ (i.e., $b_{\mathcal{C}} = b_{\mathcal{E}} = 1$). Then, we have $\beta_{\boldsymbol{c}}, \beta_{\boldsymbol{e}} \in (-2, -1)$ in (2.7). Consequently, $\kappa_\beta \in (1, 2)$ in (2.9), and the regularity property in Definition 2.1 holds for $k \geq 2$. Moreover, for $|\boldsymbol{\alpha}^\perp| \geq 2$, we have $\max\{-1 - b_{\boldsymbol{e}} + |\boldsymbol{\alpha}^\perp|, 0\} = -1 - b_{\boldsymbol{e}} + |\boldsymbol{\alpha}^\perp|$. In addition, we shall assume that, for any polyhedron Ω and right-hand side f in the classes considered here, there exists *some* $\theta \in (0, 1)$ such that the weak solution $u \in H^1_D(\Omega)$ belongs to $H^{1+\theta}(\Omega)$. For example, by [6, Theorems 2.2 and 2.4], this global regularity property is satisfied for $f \in L^2(\Omega)$ under certain restrictions on the angles between Dirichlet and Neumann faces. In [9, Theorems 4.3.2, 8.3.9 and 8.3.10], this regularity is verified for the pure Dirichlet and pure Neumann problem of (1.1). In [9, Theorem 8.1.7], a global regularity result in $N^2_{\boldsymbol{\beta}}$-spaces is shown for the mixed boundary conditions (1.2), (1.3), with bounds $b_{\mathcal{E}}, b_{\mathcal{C}} > 0$ in (2.11) characterized in terms of spectra of certain operator pencils for the Beltrami operator on spherical triangles, also for second order elliptic systems such as the Lamé-system. For the Dirichlet problem of that system, exponential convergence of a *hp*-dG discretization was shown in [15].

3. Discontinuous finite element spaces

In this section, we review the construction of *hp*-version dG spaces from [11, 12] in the axiparallel setting. The spaces are based on σ-geometric anistropic meshes and $\mathfrak{s}$-linearly increasing anisotropic polynomial degree distributions.

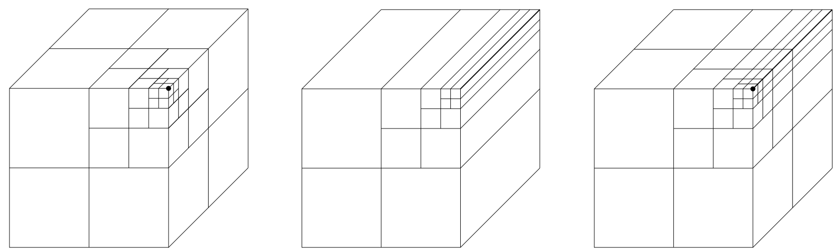

FIGURE 1. Examples of three basic geometric mesh subdivisions in the reference patch $\widetilde{Q}$ with subdivision ratio $\sigma = 1/2$: isotropic refinement towards the corner $\boldsymbol{c}$ (left), anisotropic refinement towards the edge $\boldsymbol{e}$ (center), and anisotropic refinement towards the edge-corner pair $\boldsymbol{ce}$ (right). The corner $\boldsymbol{c}$ and the edge $\boldsymbol{e}$ are shown in boldface.

3.1. **Geometric meshes and polynomial degree distributions.** To construct geometric meshes, we start from a coarse regular *quasiuniform* partition $\mathcal{M}^0 = \{Q_j\}_{j=1}^J$ of Ω into J convex axiparallel hexahedra, which we also call *patches*. Throughout, we shall assume that the initial mesh $\mathcal{M}^0$ is sufficiently fine so that an element $K \in \mathcal{M}^0$ has non-trivial intersection with at most one corner $\boldsymbol{c} \in \mathcal{C}$, and either none, one or several edges $\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}}$ meeting in $\boldsymbol{c}$. We assume further that the partition $\mathcal{M}^0$ is geometrically exact and conforming with the partition of $\partial\Omega$ into Dirichlet and Neumann faces. Each axiparallel element $Q_j = G_j(\widetilde{Q}) \in \mathcal{M}^0$ is the image under an affine mapping G_j of the *reference patch* $\widetilde{Q} = (-1, 1)^3$, given as the composition of isotropic dilations and translations. As in [11, 12], with each patch $Q_j \in \mathcal{M}^0$, we associate one of four types of *geometric reference patch meshes* on $\widetilde{Q}$, as constructed in [11, Section 3.3] in terms of four different *hp*-extensions (Ex1)–(Ex4). More specifically, whenever Q_j abuts at the singular set $\mathcal{S}$, we assign to Q_j one of the geometrically refined reference mesh patches shown in Figure 1. Here, we also allow for simultaneous refinement towards several edges in the corner-edge case shown in Figure 1 (right). The geometric refinements on the reference patches are characterized by (i) a fixed parameter $\sigma \in (0, 1)$ defining the *subdivision ratio* of the geometric refinements and (ii) the index $\ell \in \mathbb{N}$ defining the *number of refinements*. Interior patches Q_j, which have empty intersection with $\mathcal{S}$, are left unrefined, i.e., $Q_j = G_j(\widetilde{Q})$. If we denote by $\widetilde{\mathcal{M}}_j = \{\widetilde{K}\}$ the axiparallel reference mesh on $\widetilde{Q}$ associated with Q_j, then the corresponding partition $\mathcal{M}_j$ on patch Q_j will be given by $\mathcal{M}_j := \{ K : K = G_j(\widetilde{K}), \widetilde{K} \in \widetilde{\mathcal{M}}_j\}$.

For fixed parameters $\sigma \in (0, 1)$ and $\ell \in \mathbb{N}$, a geometric mesh $\mathcal{M} = \mathcal{M}_\sigma^{(\ell)}$ in Ω is now given by the disjoint union $\mathcal{M} := \bigcup_{j=1}^J \mathcal{M}_j$. Here, it is important to note that the geometric refinements $\mathcal{M}_j$ in the patches Q_j have to be suitably selected and oriented in order to achieve a proper geometric refinement towards corners and edges of Ω. Each axiparallel element $K \in \mathcal{M}$ in a geometric mesh $\mathcal{M}$ is the image of the reference cube $\widehat{K}$ under an element mapping $K = \Phi_K(\widehat{K})$, where Φ_K is the composition of the corresponding patch map G_j with an anisotropic dilation-translation. We collect all mappings Φ_K in the *mapping vector* $\boldsymbol{\Phi}(\mathcal{M}) := \{ \Phi_K : K \in \mathcal{M} \}$.

Following [11, Section 3], we may partition a geometric mesh $\mathcal{M}_\sigma^{(\ell)}$ into interior elements $\mathfrak{O}_\sigma^\ell$ away from $\mathcal{S}$ and into the terminal layer elements $\mathfrak{T}_\sigma^\ell$ at $\mathcal{S}$. That is,

$$\mathcal{M}_\sigma^{(\ell)} := \mathfrak{O}_\sigma^\ell \,\dot\cup\, \mathfrak{T}_\sigma^\ell, \tag{3.1}$$

with $\mathfrak{O}_\sigma^\ell := \{\, K \in \mathcal{M}_\sigma^{(\ell)} \,:\, \overline{K} \cap \mathcal{S} = \emptyset \,\}$ and $\mathfrak{T}_\sigma^\ell := \{\, K \in \mathcal{M}_\sigma^{(\ell)} \,:\, \overline{K} \cap \mathcal{S} \neq \emptyset \,\}$. We further partition the terminal layer $\mathfrak{T}_\sigma^\ell$ into $\mathfrak{T}_\sigma^\ell := \mathfrak{T}_\mathcal{C}^\ell \,\dot\cup\, \mathfrak{T}_\mathcal{E}^\ell$, where

$$\mathfrak{T}_\mathcal{C}^\ell := \bigcup_{\boldsymbol{c}\in\mathcal{C}} \mathfrak{T}_{\boldsymbol{c}}^\ell, \quad \mathfrak{T}_{\boldsymbol{c}}^\ell := \{\, K \in \mathfrak{T}_\sigma^\ell \,:\, \overline{K} \cap \boldsymbol{c} \neq \emptyset \,\}, \tag{3.2}$$

$$\mathfrak{T}_\mathcal{E}^\ell := \bigcup_{\boldsymbol{e}\in\mathcal{E}} \mathfrak{T}_{\boldsymbol{e}}^\ell, \quad \mathfrak{T}_{\boldsymbol{e}}^\ell := \{\, K \in \mathfrak{T}_\sigma^\ell \setminus \mathfrak{T}_\mathcal{C}^\ell \,:\, (\overline{K} \cap \boldsymbol{e})^\circ \text{ is an entire edge of } K \,\}. \tag{3.3}$$

For $\mathcal{M}^0$ sufficiently fine, we may assume that $\mathfrak{T}_{\boldsymbol{c}}^\ell$ consists of at most a finite number of terminal layer elements $K \in \mathfrak{T}_\sigma^\ell$.

With each element K of a geometric mesh $\mathcal{M}_\sigma^{(\ell)}$, we associate a polynomial degree vector $\boldsymbol{p}_K = (p_{K,1}, p_{K,2}, p_{K,3}) \in \mathbb{N}_0^3$. Its components correspond to the coordinate directions in $\widehat{K} = \Phi_K^{-1}(K)$. The polynomial degree is called *isotropic* if $p_{K,1} = p_{K,2} = p_{K,3} = p_K$. We combine the elemental polynomial degrees $\boldsymbol{p}_K$ into the *polynomial degree vector* $\boldsymbol{p}(\mathcal{M}) := \{\, \boldsymbol{p}_K \,:\, K \in \mathcal{M} \,\}$, and define $\boldsymbol{p}_{\max} := \max_{K\in\mathcal{M}} |\boldsymbol{p}_K|$, with $|\boldsymbol{p}_K| := \max_{i=1}^3 p_{K,i}$. We remark that, in addition to the mesh refinements, the extensions (Ex1)–(Ex4) introduced in [11] also provide appropriate polynomial degree distributions that increase $\mathfrak{s}$-linearly away from the singular set $\mathcal{S}$ for a *slope parameter* $\mathfrak{s} > 0$.

For an axiparallel element $K \in \mathcal{M}_\sigma^{(\ell)}$, we set $h_K := \operatorname{diam}(K)$, and denote by $h_K^\perp$ and $h_K^\parallel$ the elemental diameters of K transversal respectively parallel to the singular edge $\boldsymbol{e} \in \mathcal{E}$ nearest to K; cp. [11]. As shown in [12, Propositions 3.2 and 3.4], these quantities are closely related to the *local* distances:

$$d_K^{\boldsymbol{c}} := \operatorname{dist}(K, \boldsymbol{c}) = \inf_{\boldsymbol{x}\in K} r_{\boldsymbol{c}}(\boldsymbol{x}), \qquad d_K^{\boldsymbol{e}} := \operatorname{dist}(K, \boldsymbol{e}) = \inf_{\boldsymbol{x}\in K} r_{\boldsymbol{e}}(\boldsymbol{x}). \tag{3.4}$$

Consequently, we may write $K \in \mathcal{M}_\sigma^{(\ell)}$ in the product form

$$K := K^\perp \times K^\parallel\,, \tag{3.5}$$

where $K^\perp$ is an axiparallel and shape-regular rectangle with $\operatorname{diam}(K^\perp) \simeq h_K^\perp$ in edge-perpendicular direction, and $K^\parallel$ is an interval of length $h_K^\parallel$ in edge-parallel direction. In fact, in our analysis we may assume without loss of generality that $K = (0, h_K^\perp)^2 \times (0, h_K^\parallel)$; cp. [12, Section 5.1.4]. Analogously, we then choose $p_{K,1} = p_{K,2} =: p_K^\perp$, $p_{K,3} =: p_K^\parallel$, and write $\boldsymbol{p}_K = (p_K^\perp, p_K^\parallel)$.

For a fixed subdivision ratio $\sigma \in (0,1)$, we call the sequence $\mathfrak{M}_\sigma = \{\mathcal{M}_\sigma^{(\ell)}\}_{\ell\geq 1}$ of geometric meshes a *σ-geometric mesh family*; see [11, Definition 3.4]. As before, we shall refer to the index ℓ as *refinement level*. Geometric mesh families satisfy a bounded variation property with respect to the local mesh sizes; cp. [11, Section 3.3.3]. To review it, let $\mathfrak{M}_\sigma$ be a σ-geometric mesh family. For any $\mathcal{M} \in \mathfrak{M}_\sigma$, we define the set of all interior faces in $\mathcal{M}$ by $\mathcal{F}_I(\mathcal{M}) := \{\, f = (\partial K^\flat \cap \partial K^\sharp)^\circ \neq \emptyset \,:\, K^\flat, K^\sharp \in \mathcal{M} \,\}$. Similarly, the sets of Dirichlet and Neumann boundary faces are denoted by $\mathcal{F}_D(\mathcal{M})$ and $\mathcal{F}_N(\mathcal{M})$, respectively. We shall always assume that boundary faces belong to exactly one boundary plane $\Gamma_{\boldsymbol{\iota}}$ for $\boldsymbol{\iota} \in \mathcal{J}$. In addition, let $\mathcal{F}(\mathcal{M}) = \mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M}) \cup \mathcal{F}_N(\mathcal{M})$ denote the set of all (smallest) faces of $\mathcal{M}$. When clear from the context, we omit the dependence on $\mathcal{M}$, and simply write $\mathcal{F}_I$,

$\mathcal{F}_D$, $\mathcal{F}_N$, and $\mathcal{F}$, respectively. Furthermore, for an element $K \in \mathcal{M}$, we denote the set of its faces by $\mathcal{F}_K = \{ f \in \mathcal{F}(\mathcal{M}) : f \subset \partial K \}$. For $K \in \mathcal{M}$ and $f \in \mathcal{F}_K$, we denote by $h_{K,f}^{\perp}$ the height of K over the face f, i.e., the diameter of element K in the direction transversal to f. Then there is a constant $\mu \in (0,1)$ (only depending on σ, $\mathcal{M}^0$) such that

$$\mu \leq h_{K^\sharp,f}^{\perp} / h_{K^\flat,f}^{\perp} \leq \mu^{-1}, \qquad \forall \mathcal{M} \in \mathfrak{M}_\sigma,\ \forall f \in \mathcal{F}_I(\mathcal{M}). \tag{3.6}$$

3.2. **Finite element spaces.** Let $\mathfrak{M}_\sigma = \{\mathcal{M}_\sigma^{(\ell)}\}_{\ell \geq 1}$ be a σ-geometric mesh family in Ω. For a geometric mesh $\mathcal{M} = \mathcal{M}_\sigma^{(\ell)}$ in this family, let $\boldsymbol{\Phi}(\mathcal{M})$ and $\boldsymbol{p}(\mathcal{M})$ be the associated element mapping and elemental polynomial degree vectors, as introduced above. We then define the generic discontinuous hp-version finite element space by

$$V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p}) = \left\{ v \in L^2(\Omega) : v|_K \in \mathbb{Q}_{\boldsymbol{p}_K}(K),\ K \in \mathcal{M} \right\}. \tag{3.7}$$

Here, the local approximation spaces are defined as follows. First, on the reference element $\widehat{K}$ and for a degree vector $\boldsymbol{p} = (p_1, p_2, p_3)$, the tensor-product polynomial space $\mathbb{Q}_{\boldsymbol{p}}(\widehat{K})$ is given by $\mathbb{Q}_{\boldsymbol{p}}(\widehat{K}) = \mathbb{P}_{p_1}(\widehat{I}) \otimes \mathbb{P}_{p_2}(\widehat{I}) \otimes \mathbb{P}_{p_3}(\widehat{I})$, with $\mathbb{P}_p(\widehat{I})$ denoting the space of all polynomials of degree at most $p \geq 0$ on the reference interval $\widehat{I} = (-1,1)$. Second, on a generic element $K \in \mathcal{M}$ and with the element mapping $\Phi_K : \widehat{K} \to K$, we set $\mathbb{Q}_{\boldsymbol{p}}(K) := \{ v \in L^2(K) : v|_K \circ \Phi_K \in \mathbb{Q}_{\boldsymbol{p}}(\widehat{K}) \}$.

We now introduce two families of hp-finite element spaces for the discontinuous Galerkin methods; both yield exponentially convergent approximations and are based on a σ-geometric mesh family $\mathfrak{M}_\sigma = \{\mathcal{M}_\sigma^{(\ell)}\}_{\ell \geq 1}$. The *first family of hp-dG subspaces* is defined by

$$V_\sigma^\ell := V(\mathcal{M}_\sigma^{(\ell)}, \boldsymbol{\Phi}(\mathcal{M}_\sigma^{(\ell)}), \boldsymbol{p}_1(\mathcal{M}_\sigma^{(\ell)})), \qquad \ell \geq 1, \tag{3.8}$$

where the elemental polynomial degree vectors p_K in $\boldsymbol{p}_1(\mathcal{M}_\sigma^{(\ell)})$ are isotropic and uniform, given on each element $K \in M_\sigma^{(\ell)}$ as $p_K = \max\{3, \ell\}$. The *second family of hp-dG subspaces* is chosen as

$$V_{\sigma,\mathfrak{s}}^\ell := V(\mathcal{M}_\sigma^{(\ell)}, \boldsymbol{\Phi}(\mathcal{M}_\sigma^{(\ell)}), \boldsymbol{p}_2(\mathcal{M}_\sigma^{(\ell)})), \qquad \ell \geq 1, \tag{3.9}$$

for an increment parameter $\mathfrak{s} > 0$. Here the polynomial degree vectors $\boldsymbol{p}_2(\mathcal{M}_\sigma^{(\ell)})$ are linearly increasing with slope $\mathfrak{s}$ away from $\mathcal{S}$, i.e., specifically, the polynomial degrees $\boldsymbol{p}_K^{\perp}$ and $p_K^{\parallel}$ within each element $K \in \mathcal{M}_\sigma^{(\ell)}$ increase linearly with the number of mesh layers between that element and the closest edge $\boldsymbol{e} \in \mathcal{E}$ respectively the closed corner $\boldsymbol{c} \in \mathcal{C}$ of Ω, with the factor of proportionality being the slope parameter $\mathfrak{s} > 0$; see [11, Section 3] for more details. In the pure Neumann case ($\mathcal{J}_D = \emptyset$) we consider the factor spaces $\widetilde{V}_\sigma^\ell := V_\sigma^\ell / \mathbb{R}$ and $\widetilde{V}_{\sigma,\mathfrak{s}}^\ell := V_{\sigma,\mathfrak{s}}^\ell / \mathbb{R}$, respectively.

4. Discontinuous Galerkin discretization

In this section we present the hp-dG discretizations of (1.1)–(1.3) for which we shall prove exponential convergence. In addition, we shall adapt the stability and approximation results from [11, Section 4] to mixed boundary conditions. Throughout, $\mathcal{M} \in \mathfrak{M}_\sigma$ denotes a generic σ-geometric mesh.

4.1. **Trace operators and trace discretization parameters.** We shall first recall the jump and average operators over faces; cp. [11, 12]. For this purpose, consider an interior face $f \in \mathcal{F}_I(\mathcal{M})$ shared by two elements $K^\sharp, K^\flat \in \mathcal{M}$. Furthermore, let v respectively $\boldsymbol{w}$ be a scalar respectively vector-valued function that is sufficiently smooth inside the elements $K^\sharp, K^\flat$. Then we define the following jumps and averages of v and $\boldsymbol{w}$ along f:

$$\llbracket v \rrbracket = v|_{K^\sharp}\boldsymbol{n}_{K^\sharp} + v|_{K^\flat}\boldsymbol{n}_{K^\flat} \qquad\qquad \langle\!\langle v \rangle\!\rangle = {}^1\!/\!{}_2\left(v|_{K^\sharp} + v|_{K^\flat}\right)$$
$$\llbracket \boldsymbol{w} \rrbracket = \boldsymbol{w}|_{K^\sharp}\cdot\boldsymbol{n}_{K^\sharp} + \boldsymbol{w}|_{K^\flat}\cdot\boldsymbol{n}_{K^\flat} \qquad\qquad \langle\!\langle \boldsymbol{w} \rangle\!\rangle = {}^1\!/\!{}_2\left(\boldsymbol{w}|_{K^\sharp} + \boldsymbol{w}|_{K^\flat}\right).$$

Here, for an element $K \in \mathcal{M}$, we denote by $\boldsymbol{n}_K$ the outward unit normal vector on ∂K. For a Dirichlet boundary face $f \in \mathcal{F}_D(\mathcal{M})$ belonging to $K \in \mathcal{M}$, we let $\llbracket v \rrbracket = v|_K\boldsymbol{n}_\Omega$, $\llbracket \boldsymbol{w} \rrbracket = \boldsymbol{w}|_K \cdot \boldsymbol{n}_\Omega$, and $\langle\!\langle v \rangle\!\rangle = v|_K$, $\langle\!\langle \boldsymbol{w} \rangle\!\rangle = \boldsymbol{w}|_K$, where $\boldsymbol{n}_\Omega$ is the outward unit normal vector on $\partial\Omega$.

In analogy to the definition of $h^\perp_{K,f}$ in Section 3.1, we denote by $p^\perp_{K,f}$ the polynomial degree of $\boldsymbol{p}_K$ transversal to an elemental face $f \in \mathcal{F}_K$, $K \in \mathcal{M}$, defined as the corresponding component of $\Phi_K^{-1}(K)$. With this definitions, we introduce the trace discretization parameters $\mathtt{h}, \mathtt{p} \in L^\infty(\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M}))$ by setting $\mathtt{h}_f := \mathtt{h}|_f := \min\left\{h^\perp_{K^\sharp,f}, h^\perp_{K^\flat,f}\right\}$, and $\mathtt{p}_f := \mathtt{p}|_f := \max\left\{p^\perp_{K^\sharp,f}, p^\perp_{K^\flat,f}\right\}$, for any interior face $f \in \mathcal{F}_I(\mathcal{M})$ shared by $\partial K^\sharp$ and $\partial K^\flat$. For a Dirichlet boundary face $f \in \mathcal{F}_D(\mathcal{M})$ shared by ∂K and $\Gamma_{\boldsymbol{\iota}}$, $\boldsymbol{\iota} \in \mathcal{J}_D$, we set accordingly $\mathtt{h}_f := \mathtt{h}|_f = h^\perp_{K,f}$, $\mathtt{p}_f := \mathtt{p}|_f = p^\perp_{K,f}$.

4.2. **Interior penalty dGFEM.** The problem (1.1)–(1.3) will be discretized using an interior penalty (IP) discontinuous Galerkin finite element method. For an *hp*-dG finite element space $V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p})$ and a parameter $\theta \in \mathbb{R}$, we define the *hp*-discontinuous Galerkin approximation u_{DG} by

$$u_{\mathrm{DG}} \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p}): \qquad a_{\mathrm{DG}}(u_{\mathrm{DG}}, v) = \int_\Omega f v \,\mathrm{d}\boldsymbol{x} \qquad \forall\, v \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p}), \tag{4.1}$$

where the bilinear form $a_{\mathrm{DG}}(v, w)$ is given by

$$\begin{aligned} a_{\mathrm{DG}}(v,w) := &\int_\Omega \nabla_h v \cdot \nabla_h w \,\mathrm{d}\boldsymbol{x} - \int_{\mathcal{F}_I\cup\mathcal{F}_D} \langle\!\langle \nabla_h v \rangle\!\rangle \cdot \llbracket w \rrbracket \,\mathrm{d}s \\ &+ \theta \int_{\mathcal{F}_I\cup\mathcal{F}_D} \langle\!\langle \nabla_h w \rangle\!\rangle \cdot \llbracket v \rrbracket \,\mathrm{d}s + \gamma \int_{\mathcal{F}_I\cup\mathcal{F}_D} \mathtt{j}\,\llbracket v \rrbracket \cdot \llbracket w \rrbracket \,\mathrm{d}s. \end{aligned}$$

Here, ∇_h is the elementwise gradient operator, and $\gamma > 0$ is a stabilization parameter that will be chosen sufficiently large. Furthermore, $\mathtt{j}$ is defined as

$$\mathtt{j}\,|_f = \mathtt{p}_f^2 \mathtt{h}_f^{-1}, \qquad f \in \mathcal{F}_I \cup \mathcal{F}_D. \tag{4.2}$$

Finally, the parameter θ allows us to describe a whole range of interior penalty methods: for $\theta = -1$ we obtain the standard symmetric interior penalty (SIP) method while for $\theta = 1$ the non-symmetric (NIP) version is obtained; cp. [1] and the references therein.

To address the well-posedness of the *hp*-dGFEM, we use the standard dG norm:

$$\|\!|v|\!\|^2_{\mathrm{DG}} := \int_\Omega |\nabla_h v|^2 \,\mathrm{d}\boldsymbol{x} + \gamma \int_{\mathcal{F}_I\cup\mathcal{F}_D} \mathtt{j}\,|\llbracket v \rrbracket|^2 \,\mathrm{d}s, \tag{4.3}$$

for any $v \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p}) + H^1(\Omega)$. In the pure Neumann case ($\mathcal{F}_D = \emptyset$), $\|\!|\cdot|\!\|_{\mathrm{DG}}$ is a norm on the subspace $(V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p}) + H^1(\Omega))/\mathbb{R}$.

4.3. **Galerkin orthogonality and stability.** In order to show the well-posedness of the dG formulation (4.1), we first establish the Galerkin orthogonality of the dG discretization (4.1).

Proposition 4.1. *Suppose that the solution u to problem* (1.1)–(1.3) *belongs to the weighted space $N^2_{-1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D)$, where $\boldsymbol{b}$ is a weight vector satisfying* (2.11)*. Then, the dG approximation $u_{\mathrm{DG}} \in V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$ in* (4.1) *satisfies $a_{\mathrm{DG}}(u-u_{\mathrm{DG}},v)=0$ for any $v \in V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$.*

Proof. The proof is similar to the one of [11, Theorem 4.9], and follows from the fact that the solution u satisfies $a_{\mathrm{DG}}(u,v)=\int_\Omega fv\,\mathrm{d}\boldsymbol{x}$, for any $v \in V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$. To prove this identity, we first note that, for any $u \in N^2_{-1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D)$ and $v \in V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$, there holds the Green's formula

$$-\int_K v\Delta u\,\mathrm{d}\boldsymbol{x} = \int_K \nabla u\cdot\nabla_h v\,\mathrm{d}\boldsymbol{x} - \int_{\partial K}(\nabla u\cdot\boldsymbol{n}_K)v\,\mathrm{d}s, \qquad \forall\, K\in\mathcal{M}, \tag{4.4}$$

where in the case $\partial K\cap\partial\Omega\neq\emptyset$, the boundary term has to be understood as a pairing in $L^1(\partial K)\times L^\infty(\partial K)$. The formula (4.4) is proved along the lines of [11, Lemma 4.8] with the aid of the trace inequality in [11, Lemma 4.2] (with $t=1$). Employing (4.4), the term $\int_\Omega \nabla u\cdot\nabla_h v\,\mathrm{d}\boldsymbol{x}$ can be integrated by parts on each element, thereby revealing that $-\int_\Omega v\Delta u\,\mathrm{d}\boldsymbol{x} = \int_\Omega fv\,\mathrm{d}\boldsymbol{x}$. Here, the remaining boundary and inter-element flux terms vanish since $[\![u]\!]|_f=0$ along all $f\in\mathcal{F}_D\cup\mathcal{F}_I$, and that $[\![\nabla u]\!]|_f=0$ on all interior faces $f\in\mathcal{F}_I$. The proof of the latter identity is similar to the proof of [11, Lemma 4.7]. □

Moreover, the following proposition results from minor modifications of the proofs of the corresponding stability results presented in [11, Theorem 4.4] for the pure Dirichlet case,

Proposition 4.2. *For any degree vector $\boldsymbol{p}(\mathcal{M})$, the bilinear form a_{DG} is continuous and coercive on $V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$: there exist constants $0<C_2\le C_1<\infty$ independent of the refinement level ℓ, the local mesh sizes and the local polynomial degree vectors such that $|a_{\mathrm{DG}}(v,w)|\le C_1|\!|\!|v|\!|\!|_{\mathrm{DG}}|\!|\!|w|\!|\!|_{\mathrm{DG}}$ for all $v,w\in V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$, and such that, for $\gamma>0$ sufficiently large independent of the refinement level ℓ, the local mesh sizes and the local polynomial degree vectors, we have $a_{\mathrm{DG}}(v,v)\ge C_2|\!|\!|v|\!|\!|^2_{\mathrm{DG}}$ for all $v\in V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$. In particular, there exists a unique solution u_{DG} of* (4.1) *(unique up to constants in the pure Neumann case).*

5. Non-conforming approximation

In this section, we specify the dG interpolant, upon which our error analysis will be based, and discuss its properties (Section 5.4). To that end, we first prove auxiliary results for elemental L^2-projections (Sections 5.1 and 5.2), as well as for a low-order quasi-interpolant (Section 5.3). Finally, we show an anisotropic jump estimate for our dG interpolant (Section 5.5), which will be essential to control the non-homogeneous weights in (2.8) near Neumann edges.

5.1. **L^2-projections.** We denote by $\widehat{\pi}_p$ the L^2-projection onto $\mathbb{P}_p(\widehat{I})$ on the reference interval $\widehat{I}=(-1,1)$.

Lemma 5.1. *Let $p\ge 0$ and $u\in H^j(\widehat{I})$ for $j\in\mathbb{N}_0$. Then we have the bound*

$$\|(\widehat{\pi}_p u)^{(j)}\|_{L^2(\widehat{I})} \le C\max\{1,p\}^{2j}\|u^{(j)}\|_{L^2(\widehat{I})}\ , \tag{5.1}$$

where $C > 0$ is a constant depending only on j.

Proof. The L^2-stability of $\widehat{\pi}_p$ on $\widehat{I}$, that is the case $j = 0$, is clear and the inequality holds with constant $C = 1$. Next, consider the case $j \geq 1$. For $0 \leq p < j$, we have $(\widehat{\pi}_p u)^{(j)} \equiv 0$ and (5.1) is satisfied. Then, for $p \geq j$, we have that $(\widehat{\pi}_p u)^{(j)} \in \mathbb{P}_{p-j}(\widehat{I})$, and with the L^2-projection $\widehat{\pi}_{j-1} u \in \mathbb{P}_{j-1}(\widehat{I})$ there holds

$$\|(\widehat{\pi}_p u)^{(j)}\|_{L^2(\widehat{I})} = \|(\widehat{\pi}_p u - \widehat{\pi}_{j-1} u)^{(j)}\|_{L^2(\widehat{I})} = \|(\widehat{\pi}_p (u - \widehat{\pi}_{j-1} u))^{(j)}\|_{L^2(\widehat{I})} .$$

Hence, applying the inverse inequality from [13, Theorem 3.91] and the L^2-stability of $\widehat{\pi}_p$ yield

$$\|(\widehat{\pi}_p (u - \widehat{\pi}_{j-1} u))^{(j)}\|_{L^2(\widehat{I})} \leq C_{\mathrm{inv},j} p^{2j} \|u - \widehat{\pi}_{j-1}(u)\|_{L^2(\widehat{I})}.$$

Combining this estimate with a Poincaré-type inequality in $H^j(\widehat{I})/\mathbb{P}_{j-1}(\widehat{I})$ gives

$$\|(\widehat{\pi}_p u)^{(j)}\|_{L^2(\widehat{I})} \leq C_{\mathrm{inv},j} p^{2j} C_{\mathrm{Poinc},j} \|u^{(j)}\|_{L^2(\widehat{I})},$$

which is the desired estimate. $\square$

We now conclude the following approximation result for the L^2-projector $\widehat{\pi}_p$, with bounds which are explicit in the polynomial degree p and the regularity order s.

Lemma 5.2. *For any $3 \leq s \leq p$ and $u \in H^{s+1}(\widehat{I})$, we have*

$$\|u - \widehat{\pi}_p u\|^2_{H^2(\widehat{I})} \lesssim p^8 \Psi_{p-1,s-1} \|u^{(s+1)}\|^2_{L^2(\widehat{I})} , \tag{5.2}$$

with $\Psi_{p-1,s-1}$ defined in (1.5).

Proof. From [4, Section 8], it follows that for every $p \geq 3$ there exists a projector $\widehat{\pi}_{p,2} : H^2(\widehat{I}) \to \mathbb{P}_p(\widehat{I})$ that satisfies $(\widehat{\pi}_{p,2} u)^{(2)} = \widehat{\pi}_{p-2} u^{(2)}$ and $(\widehat{\pi}_{p,2})^{(j)} u(\pm 1) = u^{(j)}(\pm 1)$ for $j = 0, 1$. The projector $\widehat{\pi}_{p,2}$ is stable in $H^2(\widehat{I})$. Moreover, for any $3 \leq s \leq p$ and $u \in H^{s+1}(\widehat{I})$, there holds the approximation bound

$$\|u - \widehat{\pi}_{p,2} u\|^2_{H^2(\widehat{I})} \lesssim \Psi_{p-1,s-1} \|u^{(s+1)}\|^2_{L^2(\widehat{I})} . \tag{5.3}$$

By the triangle inequality, the fact that $\widehat{\pi}_p$ reproduces polynomials, and by the stability estimate (5.1), we see that

$$\|u - \widehat{\pi}_p u\|_{H^2(\widehat{I})} \leq \|u - \widehat{\pi}_{p,2} u\|_{H^2(\widehat{I})} + \|\widehat{\pi}_p (u - \widehat{\pi}_{p,2} u)\|_{H^2(\widehat{I})} \lesssim p^4 \|u - \widehat{\pi}_{p,2} u\|_{H^2(\widehat{I})}. \tag{5.4}$$

Referring to (5.3) yields the assertion for any $u \in H^{s+1}(\widehat{I})$. $\square$

Let now $\widehat{K} = (-1,1)^3$ be the reference element. In analogy to (3.5), we write $\widehat{K} = \widehat{K}^\perp \times \widehat{K}^\parallel$, with $\widehat{K}^\perp = (-1,1)^2$ and $\widehat{K}^\parallel = (-1,1)$. For a polynomial degree vector $\boldsymbol{p} = (p^\perp, p^\parallel)$ and $\widehat{v} : \widehat{K} \to \mathbb{R}$, the L^2-projection $\widehat{\Pi}_{\boldsymbol{p}}$ of $\widehat{v}$ into $\mathbb{Q}_{\boldsymbol{p}}(\widehat{K}) = \mathbb{Q}_{p^\perp}(\widehat{K}^\perp) \otimes \mathbb{Q}_{p^\parallel}(\widehat{K}^\parallel)$ is given by:

$$\widehat{\Pi}_{\boldsymbol{p}} \widehat{v} := \left(\widehat{\pi}^{(1)}_{p^\perp} \otimes \widehat{\pi}^{(2)}_{p^\perp} \otimes \widehat{\pi}^{(3)}_{p^\parallel}\right) \widehat{v} = \left(\widehat{\Pi}^\perp_{p^\perp} \otimes \widehat{\Pi}^\parallel_{p^\parallel}\right) \widehat{v}, \tag{5.5}$$

where the one-dimensional L^2-projections act in directions $\widehat{x}_1$, $\widehat{x}_2$, and $\widehat{x}_3$, and where we use the short-hand notation $\widehat{\Pi}^\perp_{p^\perp}$ and $\widehat{\Pi}^\perp_{p^\parallel}$ to denote the L^2-projections on $\widehat{K}$ in perpendicular and parallel direction, respectively. Moreover, in this setting we also introduce the tensor-product Sobolev space

$$H^2_{\mathrm{mix}}(\widehat{K}) := H^2(\widehat{K}^\perp) \otimes H^2(\widehat{K}^\parallel) = H^2(\widehat{I}) \otimes H^2(\widehat{I}) \otimes H^2(\widehat{I}), \tag{5.6}$$

endowed with the standard tensor-product norm $\|\cdot\|_{H^2_{\mathrm{mix}}(\widehat{K})}$.

Next, we provide approximation results of the L^2-projection (5.5) for a possibly anisotropic axiparallel hexahedron, *separately in edge-perpendicular and edge-parallel* direction. To state them, consider the element $K = (0, h^\perp)^2 \times (0, h^\|)$ with element mapping $\Phi_K : \widehat{K} \to K$ and a polynomial degree vector $\boldsymbol{p} = (p^\perp, p^\|)$. Consider the function $v : K \to \mathbb{R}$, and let $\widehat{v} := v \circ \Phi_K$.

Proposition 5.3. *In the above setting, let $\widehat{\eta}^\perp = \widehat{v} - \widehat{\Pi}^\perp_{p^\perp} \widehat{v}$ and $\widehat{\eta}^\| = \widehat{v} - \widehat{\Pi}^\|_{p^\|} \widehat{v}$. For the approximation error $\widehat{\eta}^\perp$ (on $\widehat{K}$) in edge-perpendicular direction there holds*

$$\|\widehat{\eta}^\perp\|^2_{H^2_{\mathrm{mix}}(\widehat{K})} \lesssim (p^\perp)^{16}\, E^\perp_{p^\perp, s^\perp}(K; v),$$

for any $3 \le s^\perp \le p^\perp$, with

$$E^\perp_{p^\perp,s^\perp}(K;v) := \Psi_{p^\perp-1,s^\perp-1} \sum_{\substack{s^\perp+1 \le |\boldsymbol{\alpha}^\perp| \le s^\perp+3 \\ 0 \le \alpha^\| \le 2}} (h^\perp)^{2|\boldsymbol{\alpha}^\perp|-2} (h^\|)^{2\alpha^\|-1} \|\mathsf{D}^{\boldsymbol{\alpha}^\perp}_\perp \mathsf{D}^{\alpha^\|}_\| v\|^2_{L^2(K)}.$$

For the approximation error $\eta^\|$ (on $\widehat{K}$) in edge-parallel direction there holds

$$\|\widehat{\mathsf{D}}^{\boldsymbol{\alpha}^\perp}_\perp \widehat{\mathsf{D}}^{\alpha^\|}_\| \widehat{\eta}^\|\|^2_{L^2(\widehat{K})} \lesssim (p^\|)^8\, \Psi_{p^\|-1,s^\|-1} (h^\perp)^{2|\boldsymbol{\alpha}^\perp|-2} (h^\|)^{2s^\|+1} \|\mathsf{D}^{\boldsymbol{\alpha}^\perp}_\perp \mathsf{D}^{s^\|+1}_\| v\|^2_{L^2(K)},$$

for any $|\boldsymbol{\alpha}^\perp| \ge 0$, $0 \le \alpha^\| \le 2$, and $3 \le s^\| \le p^\|$.

Proof. Estimate for $\widehat{\eta}^\perp$: From (5.5), we have

$$\widehat{\eta}^\perp = \widehat{v} - \widehat{\pi}^{(1)}_{p^\perp} \otimes \widehat{\pi}^{(2)}_{p^\perp} \widehat{v} = (\widehat{v} - \widehat{\pi}^{(1)}_{p^\perp} \widehat{v}) + \widehat{\pi}^{(1)}_{p^\perp} \big(\widehat{v} - \widehat{\pi}^{(2)}_{p^\perp} \widehat{v}\big)\ .$$

Hence, by the triangle inequality and the stability properties in (5.1), we find that

$$\|\widehat{\eta}^\perp\|^2_{H^2_{\mathrm{mix}}(\widehat{K})} \lesssim (p^\perp)^8 \sum_{i=1}^{2} \|\widehat{v} - \widehat{\pi}^{(i)}_{p^\perp} \widehat{v}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})}\ .$$

The one-dimensional approximation properties in Lemma 5.2 now imply that

$$\begin{aligned}\|\widehat{\eta}^\perp\|^2_{H^2_{\mathrm{mix}}(\widehat{K})} \lesssim (p^\perp)^{16} \Psi_{p^\perp-1,s^\perp-1} \Big(&\sum_{0 \le \alpha_2^\perp, \alpha^\| \le 2} \|\widehat{\mathsf{D}}^{(s^\perp+1, \alpha_2^\perp, \alpha^\|)} \widehat{v}\|^2_{L^2(\widehat{K})} \\ &+ \sum_{0 \le \alpha_1^\perp, \alpha^\| \le 2} \|\widehat{\mathsf{D}}^{(\alpha_1^\perp, s^\perp+1, \alpha^\|)} \widehat{v}\|^2_{L^2(\widehat{K})} \Big)\ .\end{aligned}$$

This bound and a scaling argument as in [12, Section 5.1.4] yield the desired bound for $\widehat{\eta}^\perp$.

Estimate for $\widehat{\eta}^\|$: The bound for $\widehat{\eta}^\|$ is a direct consequence of the one-dimensional result in Lemma 5.2 (applied in edge-parallel direction), again combined with a scaling argument as in [12, Section 5.1.4]. □

5.2. **One-dimensional geometric meshes.** In this section, we provide auxiliary exponential convergence results for elementwise L^2-projections on one-dimensional geometric meshes. To that end, on the domain $\omega = (0, 1)$, we consider a sequence $\{\mathcal{T}^{(\ell)}_\sigma\}_{\ell \ge 1}$ of geometric meshes $\mathcal{T}^{(\ell)}_\sigma = \{I_j\}_{j=1}^{\ell+1}$ with $\ell + 1$ elements which are geometrically graded towards the origin with grading factor $0 < \sigma < 1$. The elements are given by $I_1 = (0, \sigma^\ell)$ and $I_j = (\sigma^{\ell+2-j}, \sigma^{\ell+1-j})$ for $2 \le j \le \ell + 1$. The size of element I_j is given by

$$h_j := \sigma^{\ell+1-j}(1-\sigma), \qquad 2 \le j \le \ell+1, \tag{5.7}$$

which implies that there is a constant κ solely depending on σ such that

$$\kappa^{-1} h_j \le |x| \le \kappa h_j, \qquad x \in I_j,\ 2 \le j \le \ell+1\ . \tag{5.8}$$

For a slope parameter $\mathfrak{s} > 0$, we define on $\mathcal{T}_\sigma^{(\ell)}$ a $\mathfrak{s}$-linear polynomial degree vector $\boldsymbol{p}$ of length $\ell+1$ given by $\boldsymbol{p} = (p_1, ..., p_{\ell+1})$, with $p_j = \max\{3, \lceil \mathfrak{s} j \rceil\}$, $j = 1, 2, ..., \ell+1$, and set $|\boldsymbol{p}| = \max_{j=1}^{\ell+1} p_j$. We then consider the one-dimensional hp-version discontinuous finite element space

$$S^{\boldsymbol{p}}(\omega; \mathcal{T}_\sigma^{(\ell)}) = \left\{ u \in L^2(\omega) \,:\, u|_{I_j} \in \mathbb{P}_{p_j}(I_j),\ j = 1, 2, ..., \ell+1 \right\}. \tag{5.9}$$

Then, we denote by $\pi_{\boldsymbol{p}}$ the L^2-projection onto the space $S^{\boldsymbol{p}}(\omega; \mathcal{T}_\sigma^{(\ell)})$, defined on each element I_j as the (scaled) L^2-projection π_{p_j}; cp. Section 5.1. For a sufficiently smooth function $u : \omega \to \mathbb{R}$, we define the approximation error by $\eta := u - \pi_{\boldsymbol{p}} u$, and introduce the elemental error quantity:

$$T_j[\eta] := h_j^{-2} \|\eta\|^2_{L^2(I_j)} + \|\eta'\|^2_{L^2(I_j)} + h_j^2 \|\eta''\|^2_{L^2(I_j)}. \tag{5.10}$$

Proposition 5.4. *For a weight exponent $\beta > 0$, let $u : \omega \to \mathbb{R}$ be such that*

$$\| |x|^{-1-\beta+s} u^{(s)} \|_{L^2(\omega)} \le C_u^{s+1} \Gamma(s+1) \qquad \forall s \ge 0. \tag{5.11}$$

Then for ℓ sufficiently large, we have $\sum_{j=2}^{\ell+1} T_j[\eta] \le C \exp(-2b\ell)$, with constants $b, C > 0$ independent of ℓ.

Proof. Fix an element $I_j \in \mathcal{T}_\sigma^{(\ell)}$ for $2 \le j \le \ell+1$. A straightforward scaling argument yields $T_j[\eta] \simeq \left({}^{h_j}\!/_2\right)^{-1} \|\widehat{\eta}\|^2_{H^2(\widehat{I})}$, where as usual we denote by $\widehat{\eta}$ the pullback of $\eta|_{I_j}$ to the reference interval $\widehat{I} = (-1, 1)$. Therefore the approximation bound (5.2) implies that

$$T_j[\eta] \lesssim |\boldsymbol{p}|^8 \left({}^{h_j}\!/_2\right)^{-1} \Psi_{p_j-1, s_j-1} \|\widehat{u}^{(s_j+1)}\|^2_{L^2(\widehat{I})},$$

for any $3 \le s_j \le p_j$. Scaling the right-hand side above back to element I_j results in

$$T_j[\eta] \lesssim |\boldsymbol{p}|^8 \left({}^{h_j}\!/_2\right)^{2s_j} \Psi_{p_j-1, s_j-1} \|u^{(s_j+1)}\|^2_{L^2(I_j)}. \tag{5.12}$$

Moreover, by the equivalence (5.8),

$$\|u^{(s_j+1)}\|^2_{L^2(I_j)} \simeq h_j^{2+2\beta-2(s_j+1)} \| |x|^{-1-\beta+(s_j+1)} u^{(s_j+1)} \|^2_{L^2(I_j)}. \tag{5.13}$$

By combining (5.12), (5.13) with (5.11), we find that

$$\begin{aligned} T_j[\eta] &\lesssim |\boldsymbol{p}|^8 h_j^{2\beta} 2^{-2s_j} \Psi_{p_j-1, s_j-1} \| |x|^{-1-\beta+(s_j+1)} u^{(s_j+1)} \|^2_{L^2(I_j)} \\ &\lesssim |\boldsymbol{p}|^8 h_j^{2\beta} \left({}^{C_u}\!/_2\right)^{2s_j} \Psi_{p_j-1, s_j-1} \Gamma(s_j+2)^2\ , \end{aligned} \tag{5.14}$$

for any integer index $3 \le s_j \le p_j$. An interpolation argument as in [12, Lemma 5.8] shows that the bound (5.14) holds for any real $s_j \in [3, p_j]$.

Next, we sum the bound (5.14) over all layers $2 \le j \le \ell+1$. In view of (5.7), we obtain

$$\sum_{j=2}^{\ell+1} T_j[\eta] \lesssim |\boldsymbol{p}|^8 \left(\sum_{j=2}^{\ell+1} \sigma^{2(\ell+1-j)\beta} \min_{s_j \in [3, p_j]} \left[C^{2s_j} \Psi_{p_j-1, s_j-1} \Gamma(s_j+2)^2 \right] \right).$$

In [12, Lemma 5.12], it has been shown that terms of the form as in the bracket on the right-hand side above can be bounded by $C \exp(-2b(\ell+1))$. By possibly increasing the constant $C > 0$ and by reducing the value of b, the algebraic factor $|\boldsymbol{p}|^8$ can be absorbed into the exponential convergence bound. $\square$

Similarly, we obtain the following result.

Proposition 5.5. *For a weight exponent $\beta > 0$, let $u : \omega \to \mathbb{R}$ be such that*

$$\| |x|^{-\beta+s} u^{(s)} \|_{L^2(\omega)} \le C_u^{s+2} \Gamma(s+2) \qquad \forall s \ge 0. \tag{5.15}$$

Then for ℓ sufficiently large, we have $\sum_{j=2}^{\ell+1} \|\eta\|^2_{L^2(I_j)} \le C \exp(-2b\ell)$, with constants $b, C > 0$ independent of ℓ.

Proof. Fix an element $I_j \in \mathcal{T}_\sigma^{(\ell)}$ for $2 \le j \le \ell+1$. Scaling gives $\|\eta\|^2_{L^2(I_j)} = {}^{h_j}/2 \|\widehat{\eta}\|^2_{L^2(\widehat{I})}$. Then, the approximation bound (5.2), a scaling argument, the equivalence (5.8), and the regularity assumption (5.15) yield, for $3 \le s_j \le p_j$,

$$\begin{aligned}
\|\eta\|^2_{L^2(I_j)} &\lesssim |\boldsymbol{p}|^8 \, ({}^{h_j}/2) \, \Psi_{p_j-1,s_j-1} \|\widehat{u}^{(s_j+1)}\|^2_{L^2(\widehat{I})} \\
&\lesssim |\boldsymbol{p}|^8 \Psi_{p_j-1,s_j-1} \, ({}^{h_j}/2)^{2s_j+2} \, \|u^{(s_j+1)}\|^2_{L^2(I_j)} \\
&\lesssim |\boldsymbol{p}|^8 \Psi_{p_j-1,s_j-1} \, ({}^{h_j}/2)^{2s_j+2} \, h_j^{2\beta-2s_j-2} \| |x|^{-\beta+s_j+1} u^{(s_j+1)} \|^2_{L^2(I_j)} \\
&\lesssim |\boldsymbol{p}|^8 \Psi_{p_j-1,s_j-1} \, ({}^{C_u}/2)^{2s_j} \, h_j^{2\beta} \Gamma(s_j+3)^2 \, .
\end{aligned}$$

From here, the desired estimate follows as in the proof of Proposition 5.4. $\square$

5.3. **A low-order $\mathbb{P}_1$-approximation operator.** We further require the following low-order quasi-interpolation operator considered in [5]. Let $\mathfrak{K} \subset \mathbb{R}^d$ be a bounded, convex polygonal ($d = 2$) or convex polyhedral ($d = 3$) domain which is shape-regular, with diameter $h_{\mathfrak{K}}$, and whose barycenter is given by $\boldsymbol{x}_{\mathfrak{K}} := \frac{1}{|\mathfrak{K}|} \int_{\mathfrak{K}} \boldsymbol{x} \, \mathrm{d}\boldsymbol{x} \in \mathfrak{K}$, where $|\mathfrak{K}|$ denotes the volume of $\mathfrak{K}$. Then, by definition of $\boldsymbol{x}_{\mathfrak{K}}$,

$$\int_{\mathfrak{K}} (\boldsymbol{x} - \boldsymbol{x}_{\mathfrak{K}}) \, \mathrm{d}\boldsymbol{x} = \boldsymbol{0} \, . \tag{5.16}$$

Define the quasi-interpolation operator $\mathcal{I}_1 : W^{1,1}(\mathfrak{K}) \to \mathbb{P}_1(\mathfrak{K})$ by

$$\mathcal{I}_1 v := \Pi_0 v + (\boldsymbol{x} - \boldsymbol{x}_{\mathfrak{K}}) \cdot \boldsymbol{\Pi}_0(\nabla v), \tag{5.17}$$

where $\mathbb{P}_1(\mathfrak{K})$ denotes the polynomials of total degree at most 1 on $\mathfrak{K}$, and where Π_0 and $\boldsymbol{\Pi}_0$ denote element averages, i.e., the L^2- projections onto $\mathbb{P}_0(\mathfrak{K})$ and $\mathbb{P}_0(\mathfrak{K})^d$, $d = 2, 3$, respectively.

Lemma 5.6. *For the quasi-interpolation operator $\mathcal{I}_1$ defined in (5.17), there holds:*

(1) $\nabla(\mathcal{I}_1 v) \equiv \boldsymbol{\Pi}_0(\nabla v)$ *on $\mathfrak{K}$ for all $v \in W^{1,1}(\mathfrak{K})$.*
(2) $\int_{\mathfrak{K}} (v - \mathcal{I}_1 v) \, \mathrm{d}\boldsymbol{x} = 0$ *and* $\int_{\mathfrak{K}} \nabla(v - \mathcal{I}_1 v) \, \mathrm{d}\boldsymbol{x} = \boldsymbol{0}$ *for all $v \in W^{1,1}(\mathfrak{K})$.*
(3) $\mathcal{I}_1$ *reproduces polynomials in $\mathbb{P}_1(\mathfrak{K})$.*
(4) *For $1 \le q \le \infty$, the quasi-interpolant $\mathcal{I}_1$ is $W^{1,q}(\mathfrak{K})$-stable:*

$$\forall v \in W^{1,q}(\mathfrak{K}) : \quad \|\nabla(\mathcal{I}_1 v)\|_{L^q(\mathfrak{K})} \le \|\nabla v\|_{L^q(\mathfrak{K})} \, .$$

(5) *For $v \in H^1(\mathfrak{K})$, there holds*

$$\|v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}} \|\nabla v\|_{L^2(\mathfrak{K})}, \qquad \|v - \mathcal{I}_1 v\|_{L^2(\partial\mathfrak{K})} \lesssim h_{\mathfrak{K}}^{1/2} \|\nabla v\|_{L^2(\mathfrak{K})} \, .$$

(6) *For $v \in H^2(\mathfrak{K})$, there holds*

$$\|v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} + h_{\mathfrak{K}} \|\nabla(v - \mathcal{I}_1 v)\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}}^2 |v|_{H^2(\mathfrak{K})} \, .$$

(7) *Let $d = 2$, $\boldsymbol{c}$ a corner of $\mathfrak{K}$, and $r(\boldsymbol{x}) = |\boldsymbol{x} - \boldsymbol{c}|$. If $v \in H^1(\mathfrak{K})$ and $\sum_{|\boldsymbol{\alpha}|=2} \|r^\beta \mathsf{D}^{\boldsymbol{\alpha}} v\|_{L^2(\mathfrak{K})} < \infty$ for a weight exponent $0 < \beta < 1$, then there holds*

$$\|v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} + h_{\mathfrak{K}} \|\nabla(v - \mathcal{I}_1 v)\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}}^{2-\beta} \sum_{|\boldsymbol{\alpha}|=2} \|r^\beta \mathsf{D}^{\boldsymbol{\alpha}} v\|_{L^2(\mathfrak{K})} .$$

Proof. We prove this lemma item per item.

Item (1): The first item follows immediately from the definition of $\mathcal{I}_1$ in (5.17).

Item (2): Note that, by definition and item (1),

$$v - \mathcal{I}_1 v = (v - \Pi_0 v) - (\boldsymbol{x} - \boldsymbol{x}_{\mathfrak{K}}) \cdot \boldsymbol{\Pi}_0(\nabla v), \qquad \nabla(v - \mathcal{I}_1 v) = \nabla v - \boldsymbol{\Pi}_0(\nabla v).$$

Integrating these identities over $\mathfrak{K}$, the desired properties follow from (5.16) and from the fact that $\int_{\mathfrak{K}} (v - \Pi_0 v)\,\mathrm{d}\boldsymbol{x} = 0$ and $\int_{\mathfrak{K}} (\nabla v - \boldsymbol{\Pi}_0(\nabla v)\,\mathrm{d}\boldsymbol{x} = 0$.

Item (3): For $v \in \mathbb{P}_1(K)$, we see that, with item (1), $\nabla(\mathcal{I}_1 v) = \boldsymbol{\Pi}_0(\nabla v) = \nabla v$. Hence, $\mathcal{I}_1 v = v + c$, for a constant c. With item (2), we find that $c = 0$.

Item (4): For $1 \leq q < \infty$, the $W^{1,q}(\mathfrak{K})$-stability property results by noticing that $\boldsymbol{\Pi}_0(\nabla v)$ is constant, and from Hölder's inequality:

$$\begin{aligned}\|\nabla(\mathcal{I}_1 v)\|_{L^q(\mathfrak{K})} &= \|\boldsymbol{\Pi}_0(\nabla v)\|_{L^q(\mathfrak{K})} = |\mathfrak{K}|^{1/q} \left| |\mathfrak{K}|^{-1} \int_{\mathfrak{K}} \nabla v\,\mathrm{d}\boldsymbol{x} \right| \\ &\leq |\mathfrak{K}|^{1/q-1} \|\nabla v\|_{L^q(\mathfrak{K})} \|1\|_{L^{q/(q-1)}(\mathfrak{K})} \leq \|\nabla v\|_{L^q(\mathfrak{K})} .\end{aligned}$$

For $q = \infty$ the proof is similar.

Item (5): To prove the $L^2(\mathfrak{K})$-bound, we use (5.17) and the stability in item (4):

$$\begin{aligned}\|v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} &\leq \|v - \Pi_0 v\|_{L^2(\mathfrak{K})} + \|\Pi_0 v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} \\ &= \|v - \Pi_0 v\|_{L^2(\mathfrak{K})} + \|(\boldsymbol{x} - \boldsymbol{x}_{\mathfrak{K}}) \cdot \boldsymbol{\Pi}_0(\nabla v)\|_{L^2(\mathfrak{K})} \\ &\lesssim \|v - \Pi_0 v\|_{L^2(\mathfrak{K})} + h_{\mathfrak{K}} \|\nabla v\|_{L^2(\mathfrak{K})} .\end{aligned}$$

From the Poincaré inequality on $H^1(\mathfrak{K})/\mathbb{R}$, we have $\|v - \Pi_0 v\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}} \|\nabla v\|_{L^2(\mathfrak{K})}$, and thus, the $L^2(\mathfrak{K})$-bound follows.

To prove the $L^2(\partial\mathfrak{K})$-bound, we invoke the trace inequality from [11, Lemma 4.2] (with $t = 2$) for the *isotropic* element $\mathfrak{K}$:

$$\|v - \mathcal{I}_1 v\|_{L^2(\partial\mathfrak{K})} \lesssim h_{\mathfrak{K}}^{-1/2} \|v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} + h_{\mathfrak{K}}^{1/2} \|\nabla(v - \mathcal{I}_1 v)\|_{L^2(\mathfrak{K})} .$$

For the first term, we employ the previous $L^2(\mathfrak{K})$-bound. For the second term, we employ the triangle inequality and the stability bound in item (4). We readily arrive at $\|v - \mathcal{I}_1 v\|_{L^2(\partial\mathfrak{K})} \lesssim h_{\mathfrak{K}}^{1/2} \|\nabla v\|_{L^2(\mathfrak{K})}$.

Item (6): By items (2), (3), we can employ the Poincaré inequality twice, together with scaling, to obtain

$$\|v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} + h_{\mathfrak{K}} \|\nabla(v - \mathcal{I}_1 v)\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}}^2 |v - \mathcal{I}_1 v|_{H^2(\mathfrak{K})} = h_{\mathfrak{K}}^2 |v|_{H^2(\mathfrak{K})}.$$

Item (7): Similarly to before, we find that $\|v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}} \|\nabla(v - \mathcal{I}_1 v)\|_{L^2(\mathfrak{K})}$. To further bound this term we apply item (1) with the Poincaré inequalities of [10, Proposition 27] or [14, Corollary A.2.11] to find that

$$\|\nabla(v - \mathcal{I}_1 v)\|_{L^2(\mathfrak{K})} = \|\nabla v - \boldsymbol{\Pi}_0(\nabla v)\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}}^{1-\beta} \sum_{|\boldsymbol{\alpha}|=2} \left\| r^\beta \mathsf{D}^{|\boldsymbol{\alpha}|} v \right\|_{L^2(\mathfrak{K})} .$$

This completes the proof. □

5.4. **A non-conforming dG interpolant.** We now specify a dG interpolant Π as follows. Let $v \in H^1(\Omega)$, and let $V(\mathcal{M}_\sigma^{(\ell)}, \boldsymbol{\Phi}(\mathcal{M}_\sigma^{(\ell)}), \boldsymbol{p}(\mathcal{M}_\sigma^{(\ell)}))$ be an hp-dG space based on a geometric mesh $\mathcal{M}_\sigma^{(\ell)}$. For an element $K = K^\perp \times K^\parallel \in \mathcal{M}_\sigma^{(\ell)}$ with polynomial degree vector $\boldsymbol{p}_K = (p_K^\perp, p_K^\parallel)$, we choose Πv elementwise and with respect to the partition $\mathcal{M}_\sigma^{(\ell)} = \mathfrak{O}_\sigma^\ell \,\dot\cup\, \mathfrak{T}_\mathcal{C}^\ell \,\dot\cup\, \mathfrak{T}_\mathcal{E}^\ell$ (introduced in Section 3.1) as:

$$(\Pi v)_K = \Pi_K v|_K := \begin{cases} \Pi_{\boldsymbol{p}_K}(v|_K) = \Pi^\perp_{p_K^\perp} \otimes \Pi^\parallel_{p_K^\parallel}(v|_K) & \text{if } K \in \mathfrak{O}_\sigma^\ell, \\ \mathcal{I}_1(v|_K) & \text{if } K \in \mathfrak{T}_\mathcal{C}^\ell, \\ \mathcal{I}_1^\perp \otimes \Pi_{p_K^\parallel}(v|_K) & \text{if } K \in \mathfrak{T}_\mathcal{E}^\ell. \end{cases} \tag{5.18}$$

The operator $\Pi_{\boldsymbol{p}_K}$ is the (scaled) L^2-projection onto $\mathbb{Q}_{\boldsymbol{p}_K}(K)$ given by

$$\Pi_{\boldsymbol{p}_K}(v|_K) := \left(\widehat{\Pi}_{\boldsymbol{p}_K}(v \circ \Phi_K)\right) \circ \Phi_K^{-1}, \tag{5.19}$$

with $\widehat{\Pi}_{\boldsymbol{p}_K}$ the reference projection in (5.5) and Φ_K the element mapping. As on the reference element, it tensorizes into projections $\Pi^\perp_{p_K^\perp}$ and $\Pi^\parallel_{p_K^\parallel}$ in edge-perpendicular and edge-parallel direction, respectively. The operator $\mathcal{I}_1$ is the three-dimensional quasi-interpolant (5.17) for the isotropic corner elements, whereas $\mathcal{I}_1^\perp$ is the two-dimensional $\mathbb{P}_1$-interpolant (5.17) applied in edge-perpendicular direction.

It is evident that, on $K \in \mathfrak{O}_\sigma^\ell \,\dot\cup\, \mathfrak{T}_\mathcal{E}^\ell$, the interpolant Π in (5.18) has tensor-product structure. For simplicity, we shall then write $\Pi = \Pi^\perp \otimes \Pi^\parallel$ or $\Pi_K = \Pi_K^\perp \otimes \Pi_K^\parallel$ (to indicate the dependence on element K).

Lemma 5.7. *On elements $K \in \mathfrak{O}_\sigma^\ell \,\dot\cup\, \mathfrak{T}_\mathcal{E}^\ell$ with $K = K^\perp \times K^\parallel$, the tensor-product interpolant $\Pi_K = \Pi_K^\perp \otimes \Pi_K^\parallel$ introduced in* (5.18) *satisfies:*

(1) *The operator $\Pi_K^\parallel$ is the L^2-projection in edge-parallel direction into polynomials in $\mathbb{P}_{p_K^\parallel}(K^\parallel)$, and $\Pi_K^\perp$ is an approximation operator from $H^1(K^\perp)$ into $\mathbb{Q}_{p_K^\perp}(K^\perp)$ for $K \in \mathfrak{O}_\sigma^\ell$, respectively into $\mathbb{P}_1(K^\perp)$ for $K \in \mathfrak{T}_\mathcal{E}^\ell$.*
(2) *The operator $\Pi_K^\perp$ reproduces polynomials in $\mathbb{Q}_{p_K^\perp}(K^\perp)$ for $K \in \mathfrak{O}_\sigma^\ell$, respectively in $\mathbb{P}_1(K^\perp)$ for $K \in \mathfrak{T}_\mathcal{E}^\ell$.*
(3) *The operator $\Pi_K^\perp$ satisfies the approximation property:*
$$\|v - \Pi_K^\perp v\|^2_{L^2(\partial K^\perp)} \lesssim h_K^\perp \|\mathsf{D}_\perp v\|^2_{L^2(K^\perp)}, \qquad v \in H^1(K^\perp),$$

Proof. The first two properties follow by construction and Lemma 5.6, item (3). The trace approximation bound in item (3) is a standard result for the two-dimensional L^2-projection $\Pi_K^\perp = \Pi^\perp_{p_K^\perp}$ in (5.18). For $\Pi_K^\perp = \mathcal{I}_1^\perp$ in (5.18) this follows from Lemma 5.6, item (5). □

5.5. **An anisotropic jump estimate.** The following bound is crucial for controlling the consistency errors in anisotropic elements near Neumann edges.

Proposition 5.8. *Consider an interior face $f = (\partial K_1 \cap \partial K_2)^\circ$, which is parallel to the nearest edge $\boldsymbol{e} \in \mathcal{E}$ and shared by two axiparallel elements $K_1 = K_1^\perp \times K^\parallel$ and $K_2 = K_2^\perp \times K^\parallel$. Here, $K_1^\perp$ and $K_2^\perp$ are two shape-regular and possibly non-matching rectangles in edge-perpendicular direction, and $K^\parallel$ is a one-dimensional interval in edge-parallel direction. We assume that the bounded variation property* (3.6) *holds over the face f. For elemental polynomial degree vectors given by $\boldsymbol{p}_{K_i} = (p_{K_i}^\perp, p^\parallel)$,* *let* $\Pi = \Pi^\perp \otimes \Pi^\parallel$ *be a tensor-product dG interpolation operator as in* (5.18) *satisfying properties (1)–(3) in Lemma 5.7 over* $\{K_1, K_2\}$*. For* $v \in H^1((\overline{K}_1 \cup \overline{K}_2)^\circ)$*, let* $\eta = v - \Pi v$*,* $\eta^\perp = v - \Pi^\perp v$ *and* $\eta^\parallel = v - \Pi^\parallel v$*. Then there holds*

$$\mathtt{h}_f^{-1} \| [\![\eta]\!] \|^2_{L^2(f)} \lesssim \| \mathsf{D}_\perp \eta^\perp \|^2_{L^2(K_1)} + \| \mathsf{D}_\perp \eta^\perp \|^2_{L^2(K_2)}. \tag{5.20}$$

Proof. Since $\Pi^\perp$ reproduces polynomials in perpendicular direction, we see that

$$\eta^\perp - \Pi^\perp \eta^\perp = (v - \Pi^\perp v) - \Pi^\perp (v - \Pi^\perp v) = v - \Pi^\perp v = \eta^\perp, \tag{5.21}$$

on $\{K_1, K_2\}$. Noting that $[\![\eta]\!] = [\![\Pi v]\!]$ and $\Pi^\parallel_{K_1} v|_{K_1} = \Pi^\parallel_{K_2} v|_{K_2}$ on f and with (5.21), we obtain

$$\begin{aligned}
\| [\![\eta]\!] \|^2_{L^2(f)} &= \int_f \left(\Pi_{K_1} v|_{K_1} - \Pi_{K_2} v|_{K_2} \right)^2 \,\mathrm{d}s \\
&= \int_f \left((\Pi^\perp_{K_1} \otimes \Pi^\parallel_{K_1} v|_{K_1} - \Pi^\parallel_{K_1} v|_{K_1}) - (\Pi^\perp_{K_2} \otimes \Pi^\parallel_{K_2} v|_{K_2} - \Pi^\parallel_{K_2} v|_{K_2}) \right)^2 \,\mathrm{d}s \\
&\lesssim \int_f \left(\Pi^\parallel_{K_1} \eta^\perp|_{K_1} \right)^2 \,\mathrm{d}s + \int_f \left(\Pi^\parallel_{K_2} \eta^\perp|_{K_2} \right)^2 \,\mathrm{d}s \\
&\lesssim \int_f \left(\Pi^\parallel_{K_1} (\eta^\perp - \Pi^\perp_{K_1} \eta^\perp)|_{K_1} \right)^2 \,\mathrm{d}s + \int_f \left(\Pi^\parallel_{K_2} (\eta^\perp - \Pi^\perp_{K_2} \eta^\perp)|_{K_2} \right)^2 \,\mathrm{d}s.
\end{aligned}$$

By the definition of $\mathtt{h}_f$ and the bounded variation property (3.6), we remark that $\mathtt{h}_f \simeq h^\perp_{K_1,f} \simeq h^\perp_{K_2,f} \simeq h^\perp_{K_1} \simeq h^\perp_{K_2}$. Hence, the trace estimate in Lemma 5.7, item (3), in edge-perpendicular direction and the stability of the L^2-projection $\Pi^\parallel$ in edge-parallel direction (cp. Lemma 5.7, item (1)) readily yield

$$\| [\![\eta]\!] \|^2_{L^2(f)} \lesssim \mathtt{h}_f \left(\| \mathsf{D}_\perp \eta^\perp \|^2_{L^2(K_1)} + \| \mathsf{D}_\perp \eta^\perp \|^2_{L^2(K_2)} \right),$$

which completes the proof. □

6. Error analysis and exponential convergence

In this section, we first derive error estimates for the specific dG interpolant Π defined in (5.18). We than state our main exponential convergence bound.

6.1. Splitting of errors and consistency terms.

Let u be the solution of (1.1)–(1.3), and Π the dG interpolant defined in (5.18) on a geometric mesh $\mathcal{M} = \mathcal{M}^{(\ell)}_\sigma$. In the sequel, we shall denote by η the approximation error

$$\eta := u - \Pi u, \qquad \text{for } K \in \mathcal{M}. \tag{6.1}$$

We will separately consider the errors in edge-perpendicular and edge-parallel directions. Recall that $\Pi = \Pi^\perp \otimes \Pi^\parallel$ on $K \in \mathfrak{O}^\ell_\sigma \,\dot\cup\, \mathfrak{T}^\ell_{\mathcal{E}}$; cp. Lemma 5.7. We set

$$\eta^\perp := u - \Pi^\perp u, \qquad \eta^\parallel := u - \Pi^\parallel u, \qquad \text{for } K \in \mathfrak{O}^\ell_\sigma \,\dot\cup\, \mathfrak{T}^\ell_{\mathcal{E}}. \tag{6.2}$$

For $K \in \mathfrak{O}^\ell_\sigma \,\dot\cup\, \mathfrak{T}^\ell_{\mathcal{E}}$, we write $\eta = (u - \Pi^\parallel u) + \Pi^\parallel (u - \Pi^\perp u) = \eta^\parallel + \Pi^\parallel \eta^\perp$, with $\Pi^\parallel$ an L^2-projection; cp. Lemma 5.7. Hence, the stability result (5.1) yields

$$\| \mathsf{D}^{\boldsymbol{\alpha}^\perp}_\perp \mathsf{D}^{\alpha^\parallel}_\parallel \eta \|^2_{L^2(K)} \lesssim (p^\parallel_K)^{4\alpha^\parallel} \left(\| \mathsf{D}^{\boldsymbol{\alpha}^\perp}_\perp \mathsf{D}^{\alpha^\parallel}_\parallel \eta^\parallel \|^2_{L^2(K)} + \| \mathsf{D}^{\boldsymbol{\alpha}^\perp}_\perp \mathsf{D}^{\alpha^\parallel}_\parallel \eta^\perp \|^2_{L^2(K)} \right), \tag{6.3}$$

for any $K \in \mathfrak{O}^\ell_\sigma \,\dot\cup\, \mathfrak{T}^\ell_{\mathcal{E}}$, $\boldsymbol{\alpha}^\perp \in \mathbb{N}^2_0$ and $0 \le \alpha^\parallel \le 2$.

Next, we introduce various consistency terms, in acccordance with the partition of $\mathcal{M}_\sigma^{(\ell)} = \mathfrak{O}_\sigma^\ell \,\dot\cup\, \mathfrak{T}_\mathcal{C}^\ell \,\dot\cup\, \mathfrak{T}_\mathcal{E}^\ell$, with $\mathfrak{T}_\mathcal{C}^\ell = \cup_{\boldsymbol{c}\in\mathcal{C}}\mathfrak{T}_{\boldsymbol{c}}^\ell$ and $\mathfrak{T}_\mathcal{C}^\ell = \cup_{\boldsymbol{e}\in\mathcal{E}}\mathfrak{T}_{\boldsymbol{e}}^\ell$; cp. Section 3.1. We define

$$\Upsilon_{\mathfrak{O}_\sigma^\ell}[\eta] := \sum_{K\in\mathfrak{O}_\sigma^\ell} T_\mathfrak{O}^K[\eta], \quad \Upsilon_{\mathfrak{T}_{\boldsymbol{c}}^\ell}[\eta] := \sum_{K\in\mathfrak{T}_{\boldsymbol{c}}^\ell} T_{\boldsymbol{c}}^K[\eta], \quad \Upsilon_{\mathfrak{T}_{\boldsymbol{e},i}^\ell}[\eta] := \sum_{K\in\mathfrak{T}_{\boldsymbol{e}}^\ell} T_{\boldsymbol{e},i}^K[\eta], \tag{6.4}$$

for $i = 1, 2$, with

$$\begin{aligned}
T_\mathfrak{O}^K[\eta] &:= (h_K^\parallel)^{-2}\|\eta\|_{L^2(K)}^2 + \|\nabla\eta\|_{L^2(K)}^2 + (h_K^\perp)^2\|\mathsf{D}_\perp^2\eta\|_{L^2(K)}^2 + (h_K^\parallel)^2\|\mathsf{D}_\parallel^2\eta\|_{L^2(K)}^2,\\
T_{\boldsymbol{c}}^K[\eta] &:= h_K^{-2}\|\eta\|_{L^2(K)}^2 + \|\nabla\eta\|_{L^2(K)}^2 + h_K^{-1}|\eta|_{W^{2,1}(K)}^2,\\
T_{\boldsymbol{e},1}^K[\eta] &:= (h_K^\parallel)^{-2}\|\eta\|_{L^2(K)}^2 + \|\nabla\eta\|_{L^2(K)}^2 + (h_K^\parallel)^2\|\mathsf{D}_\parallel^2\eta\|_{L^2(K)}^2,\\
T_{\boldsymbol{e},2}^K[\eta] &:= |K|^{-1}(h_K^\perp)^2\|\mathsf{D}_\perp^2\eta\|_{L^1(K)}^2.
\end{aligned}$$

In addition, for a Dirichlet boundary edge $\boldsymbol{e}\in\mathcal{E}_D$, we set

$$\Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^\ell}[\eta] := \sum_{K\in\mathfrak{T}_{\boldsymbol{e}}^\ell} T_{\boldsymbol{e},D}^K[\eta], \qquad T_{\boldsymbol{e},D}^K[\eta] = (h_K^\perp)^{-2}\|\eta\|_{L^2(K)}^2. \tag{6.5}$$

An analogous term does not arise for Neumann boundary edges $\boldsymbol{e}\in\mathcal{E}_N$ (which are not present in the dG bilinear form $a_{\mathrm{DG}}(v,w)$).

6.2. **Error estimates.** We now establish the following error bound for the dG-energy norm error.

Theorem 6.1. *Let $u \in B_{-1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D)$ be the solution of (1.1)–(1.3), and let u_{DG} be the DG approximation obtained from (4.1) with a sufficiently large penalty parameter $\gamma > 0$ in the dG space V_σ^ℓ in (3.8), respectively in $V_{\sigma,\mathfrak{s}}^\ell$ in (3.9). Let Πu be the dG interpolant selected in (5.18). Then for the approximation errors in (6.1), (6.2) there holds the bound*

$$\begin{aligned}
\|u-u_{\mathrm{DG}}\|_{\mathrm{DG}}^2 \le C\boldsymbol{p}_{\max}^{12}\Bigg(&\Upsilon_{\mathfrak{O}_\sigma^\ell}[\eta^\perp] + \Upsilon_{\mathfrak{O}_\sigma^\ell}[\eta^\parallel] + \sum_{\boldsymbol{c}\in\mathcal{C}}\Upsilon_{\mathfrak{T}_{\boldsymbol{c}}^\ell}[\eta]\\
&+\sum_{\boldsymbol{e}\in\mathcal{E}}\Big(\Upsilon_{\mathfrak{T}_{\boldsymbol{e},1}^\ell}[\eta^\perp] + \Upsilon_{\mathfrak{T}_{\boldsymbol{e},1}^\ell}[\eta^\parallel] + \Upsilon_{\mathfrak{T}_{\boldsymbol{e},2}^\ell}[\eta]\Big) + \sum_{\boldsymbol{e}\in\mathcal{E}_D}\Big(\Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^\ell}[\eta^\perp] + \Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^\ell}[\eta^\parallel]\Big)\Bigg).
\end{aligned}$$

The constant $C > 0$ is independent of the refinement level ℓ, the local mesh sizes and the local polynomial degree vectors.

Proof. We write $u - u_{\mathrm{DG}} = \eta + \xi$, with $\eta = u - \Pi u$ as in (6.1), and $\xi := \Pi u - u_{\mathrm{DG}}$. Then,

$$\begin{aligned}
\|u-u_{\mathrm{DG}}\|_{\mathrm{DG}}^2 &\le 2\left(\|\xi\|_{\mathrm{DG}}^2 + \|\eta\|_{\mathrm{DG}}^2\right)\\
&\lesssim \|\xi\|_{\mathrm{DG}}^2 + \boldsymbol{p}_{\max}^2\Big(\|\nabla_h\eta\|_{L^2(\Omega)}^2 + \sum_{f\in\mathcal{F}_D\cup\mathcal{F}_I}\mathtt{h}_f^{-1}\|[\![\eta]\!]\|_{L^2(f)}^2\Big).
\end{aligned} \tag{6.6}$$

To bound $\|\xi\|_{\mathrm{DG}}^2$ in (6.6), we employ the coercivity in Proposition 4.2 and the Galerkin orthogonality in Proposition 4.1. We find that

$$\|\xi\|_{\mathrm{DG}}^2 \lesssim -a_{\mathrm{DG}}(\eta,\xi) =: T_1 + T_2, \tag{6.7}$$

where

$$T_1 = \int_\Omega \nabla_h \eta \cdot \nabla_h \xi \,\mathrm{d}\boldsymbol{x} - \theta \int_{\mathcal{F}_I \cup \mathcal{F}_D} \langle\!\langle \nabla_h \xi \rangle\!\rangle \cdot [\![\eta]\!] \,\mathrm{d}s - \gamma \int_{\mathcal{F}_I \cup \mathcal{F}_D} \mathtt{j}\, [\![\eta]\!] \cdot [\![\xi]\!] \,\mathrm{d}s,$$

$$T_2 - \int_{\mathcal{F}_I \cup \mathcal{F}_D} \langle\!\langle \nabla_h \eta \rangle\!\rangle \cdot [\![\xi]\!] \,\mathrm{d}s \; .$$

The term T_1 is bounded using the Cauchy-Schwarz inequality:

$$\begin{aligned} |T_1| &\lesssim \boldsymbol{p}_{\max} \Big(\|\nabla_h \eta\|^2_{L^2(\Omega)} + \Big\| \mathtt{h}^{-1/2} [\![\eta]\!] \Big\|^2_{L^2(\mathcal{F}_I \cup \mathcal{F}_D)} \Big)^{1/2} \\ &\quad \times \Big(\|\nabla_h \xi\|^2_{L^2(\Omega)} + \Big\| \mathtt{j}^{\,-1/2} \langle\!\langle \nabla_h \xi \rangle\!\rangle \Big\|^2_{L^2(\mathcal{F}_I \cup \mathcal{F}_D)} + \Big\| \mathtt{j}^{\,1/2} [\![\xi]\!] \Big\|^2_{L^2(\mathcal{F}_I \cup \mathcal{F}_D} \Big)^{1/2}. \end{aligned}$$

Estimating the term involving $\langle\!\langle \nabla_h \xi \rangle\!\rangle$ as in the proof of [11, Theorem 4.10], with the aid of [11, Lemma 4.3a)], we obtain

$$|T_1| \lesssim \boldsymbol{p}_{\max} |\!|\!| \xi |\!|\!|_{\mathrm{DG}} \Big(\|\nabla_h \eta\|^2_{L^2(\Omega)} + \Big\| \mathtt{h}^{-1/2} [\![\eta]\!] \Big\|^2_{L^2(\mathcal{F}_I \cup \mathcal{F}_D)} \Big)^{1/2} . \tag{6.8}$$

Next, we bound T_2. There holds

$$\begin{aligned} |T_2| &\leq \sum_{f \in \mathcal{F}_I \cup \mathcal{F}_D} \int_f |\langle\!\langle \nabla_h \eta \rangle\!\rangle \cdot \boldsymbol{n}_f| |[\![\xi]\!]| \,\mathrm{d}s \\ &\lesssim \sum_{f \in \mathcal{F}_I \cup \mathcal{F}_D} \|\mathtt{j}^{\,-1/2} \langle\!\langle \nabla_h \eta \rangle\!\rangle \cdot \boldsymbol{n}_f\|_{L^1(f)} \|\mathtt{j}^{\,1/2} [\![\xi]\!]\|_{L^\infty(f)}, \end{aligned}$$

where $\boldsymbol{n}_f$ is an orthonormal vector on f pointing in a preset direction. Therefore, using [11, Lemma 4.3b)] and the bounded variation property (3.6), it follows that

$$\begin{aligned} |T_2| &\lesssim \boldsymbol{p}^2_{\max} \sum_{f \in \mathcal{F}_I \cup \mathcal{F}_D} |f|^{-1/2} \|\mathtt{j}^{\,-1/2} \langle\!\langle \nabla_h \eta \rangle\!\rangle \cdot \boldsymbol{n}_f\|_{L^1(f)} \|\mathtt{j}^{\,1/2} [\![\xi]\!]\|_{L^2(f)} \\ &\lesssim \boldsymbol{p}^2_{\max} |\!|\!| \xi |\!|\!|_{\mathrm{DG}} \Big(\sum_{f \in \mathcal{F}_I \cup \mathcal{F}_D} |f|^{-1} \|\mathtt{j}^{\,-1/2} \langle\!\langle \nabla_h \eta \rangle\!\rangle \cdot \boldsymbol{n}_f\|^2_{L^1(f)} \Big)^{1/2} \\ &\lesssim \boldsymbol{p}^2_{\max} |\!|\!| \xi |\!|\!|_{\mathrm{DG}} \Big(\sum_{K \in \mathcal{M}} \sum_{f \in (\mathcal{F}_I \cup \mathcal{F}_D) \cap \mathcal{F}_K} |f|^{-1} h^\perp_{K,f} \|\nabla_h \eta \cdot \boldsymbol{n}_K\|^2_{L^1(f)} \Big)^{1/2} . \end{aligned}$$

Since $|\nabla \eta \cdot \boldsymbol{n}_K| = |\partial_{K,f,\perp} \eta|$ on $f \in \mathcal{F}_K$, with $\partial_{K,f,\perp}$ denoting the partial derivative in direction transversal to f, and $|K| \simeq |f| h^\perp_{K,f}$, applying the anisotropic trace inequality [11, Lemma 4.2] (with $t = 1$) yields

$$\begin{aligned} |T_2| \lesssim \boldsymbol{p}^2_{\max} |\!|\!| \xi |\!|\!|_{\mathrm{DG}} \Big(&\sum_{K \in \mathcal{M}} |K|^{-1} \|\nabla \eta\|^2_{L^1(K)} \\ &+ \sum_{K \in \mathcal{M}} \sum_{f \in (\mathcal{F}_I \cup \mathcal{F}_D) \cap \mathcal{F}_K} |K|^{-1} (h^\perp_{K,f})^2 \|\partial^2_{K,f,\perp} \eta\|^2_{L^1(K)} \Big)^{1/2}. \end{aligned}$$

By Hölder's inequality, we conclude that $|K|^{-1} \|\nabla \eta\|^2_{L^1(K)} \leq \|\nabla \eta\|^2_{L^2(K)}$. Since all elements K are axiparallel hexahedra, there are only two cases, $f \,\|\, \boldsymbol{e}$ and $f \perp \boldsymbol{e}$, where $\boldsymbol{e}$ is the edge nearest to $f \in \mathcal{F}_K$. In the former case, there holds $(h^\perp_{K,f})^2 \|\partial^2_{K,f,\perp} \eta\|^2_{L^1(K)} \leq (h^\perp_K)^2 \|\mathsf{D}^2_\perp \eta\|^2_{L^1(K)}$, in the latter $(h^\perp_{K,f})^2 \|\partial^2_{K,f,\perp} \eta\|^2_{L^1(K)} =$

$(h_K^{\|})^2\|\mathsf{D}_{\|}^2\eta\|_{L^1(K)}^2$. Therefore,

$$|T_2| \lesssim \boldsymbol{p}_{\max}^2 |\!|\!|\xi|\!|\!|_{\mathrm{DG}} \Big(\|\nabla_h \eta\|_{L^2(\Omega)}^2 + \sum_{K\in\mathcal{M}} \big(|K|^{-1}(h_K^{\perp})^2\|\mathsf{D}_{\perp}^2\eta\|_{L^1(K)}^2 + |K|^{-1}(h_K^{\|})^2\|\mathsf{D}_{\|}^2\eta\|_{L^1(K)}^2 \big) \Big)^{1/2}.$$

Combining this estimate with (6.7), (6.8), and dividing the resulting inequality by $|\!|\!|\xi|\!|\!|_{\mathrm{DG}}$ gives a bound for $|\!|\!|\xi|\!|\!|_{\mathrm{DG}}$. Squaring it and taking into account (6.6) give

$$|\!|\!|u-u_{\mathrm{DG}}|\!|\!|_{\mathrm{DG}}^2 \lesssim \boldsymbol{p}_{\max}^4 \left(\|\nabla_h\eta\|_{L^2(\Omega)}^2 + \sum_{f\in\mathcal{F}_I\cup\mathcal{F}_D} \mathtt{h}_f^{-1} \|[\![\eta]\!]\|_{L^2(f)}^2 + \sum_{K\in\mathcal{M}} |K|^{-1} \left((h_K^{\perp})^2\|\mathsf{D}_{\perp}^2\eta\|_{L^1(K)}^2 + (h_K^{\|})^2\|\mathsf{D}_{\|}^2\eta\|_{L^1(K)}^2 \right) \right)^{1/2}. \tag{6.9}$$

It remains to bound the jumps of η over $f\in\mathcal{F}_I\cup\mathcal{F}_D$. To this end, we distinguish three cases:

Case 1: If $f\perp \boldsymbol{e}$, $f\in\mathcal{F}_I$, is an *interior face transversal to the closest edge* $\boldsymbol{e}\in\mathcal{E}$, shared by two elements K_1 and K_2, with $\mathtt{h}_f\simeq h_{K_1,f}^{\perp}\simeq h_{K_2,f}^{\perp}\simeq h_{K_1}^{\|}\simeq h_{K_2}^{\|}$, cp. property (3.6), we use the trace estimate [11, Lemma 4.2] (with $t=2$) to obtain

$$\mathtt{h}_f^{-1}\|[\![\eta]\!]\|_{L^2(f)}^2 \lesssim \sum_{i=1}^{2} \left(h_{K_i}^{\|})^{-2}\|\eta\|_{L^2(K_i)}^2 + \|\nabla\eta\|_{L^2(K_i)}^2 \right).$$

The same bound is applied over interior faces shared by shape-regular elements K_1 and K_2, where $h_{K_1}^{\|}\simeq h_{K_1}\simeq h_{K_2}\simeq h_{K_2}^{\|}$.

Case 2: If $f\,\|\,\boldsymbol{e}$, $f\in\mathcal{F}_I$, is an *interior face parallel to the closest edge* $\boldsymbol{e}\in\mathcal{E}$, shared by two *anisotropic* elements K_1 and K_2, with $\mathtt{h}_f\simeq h_{K_1,f}^{\perp}\simeq h_{K_2,f}^{\perp}\simeq h_{K_1}^{\perp}\simeq h_{K_2}^{\perp}$, cp. (3.6), and with the same edge-parallel polynomial degree $p^{\|}$ as in Proposition 5.8 (see also [11]), then we apply the anisotropic jump estimate in (5.20) to find that

$$\mathtt{h}_f^{-1}\|[\![\eta]\!]\|_{L^2(f)}^2 \lesssim \|\mathsf{D}_{\perp}\eta^{\perp}\|_{L^2(K_1)}^2 + \|\mathsf{D}_{\perp}\eta^{\perp}\|_{L^2(K_2)}^2 \lesssim \|\nabla\eta^{\perp}\|_{L^2(K_1)}^2 + \|\nabla\eta^{\perp}\|_{L^2(K_2)}^2 .$$

Case 3: If $f\in\mathcal{F}_D$ is a *Dirichlet boundary face*, we again invoke the trace estimate [11, Lemma 4.2] (with $t=2$) to obtain, for $f\in\mathcal{F}_K$,

$$\mathtt{h}_f^{-1}\|[\![\eta]\!]\|_{L^2(f)}^2 \simeq (h_K^{\perp})^{-1}\|\eta\|_{L^2(f)}^2 \lesssim (h_K^{\perp})^{-2}\|\eta\|_{L^2(K)}^2 + \|\nabla\eta\|_{L^2(K)}^2 .$$

Inserting these jump bounds into estimate (6.9) results in

$$\begin{aligned} |\!|\!|u-u_{\mathrm{DG}}|\!|\!|_{\mathrm{DG}}^2 \lesssim{}& \boldsymbol{p}_{\max}^4 \sum_{K\in\mathcal{M}} \big((h_K^{\|})^{-2}\|\eta\|_{L^2(K)}^2 + \|\nabla\eta\|_{L^2(K)}^2 \big) \\ &+ \boldsymbol{p}_{\max}^4 \sum_{K\in\mathcal{M}} \big(|K|^{-1}(h_K^{\perp})^2\|\mathsf{D}_{\perp}^2\eta\|_{L^1(K)}^2 + |K|^{-1}(h_K^{\|})^2\|\mathsf{D}_{\|}^2\eta\|_{L^1(K)}^2 \big) \\ &+ \boldsymbol{p}_{\max}^4 \sum_{K\in\mathcal{M}\setminus\mathfrak{T}_{\mathcal{C}}^{\ell}} \|\nabla\eta^{\perp}\|_{L^2(K)}^2 + \boldsymbol{p}_{\max}^4 \sum_{\boldsymbol{e}\in\mathcal{E}_D} \Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^{\ell}}[\eta] . \end{aligned}$$

Recalling the partition $\mathcal{M}_{\sigma}^{(\ell)} = \mathfrak{O}_{\sigma}^{\ell} \,\dot\cup\, \mathfrak{T}_{\mathcal{E}}^{\ell} \,\dot\cup\, \mathfrak{T}_{\mathcal{C}}^{\ell}$, we estimate the $L^1(K)$-norms of $\mathsf{D}_{\perp}^2\eta$ (for $K\in\mathfrak{O}_{\sigma}^{\ell}\cup\mathfrak{T}_{\mathcal{C}}^{\ell}$), and $\mathsf{D}_{\|}^2\eta$ (for $K\in\mathfrak{O}_{\sigma}^{\ell}$) by their $L^2(K)$-norms using Hölder's inequality. Moreover, noting that elements in $\mathfrak{T}^\ell_{\mathcal{C}}$ are isotropic with $h_K \simeq h^\perp_K \simeq h^\parallel_K$ and $|K| \simeq h^3_K$, yields

$$\|u - u_{\mathrm{DG}}\|^2_{\mathrm{DG}} \lesssim \boldsymbol{p}^4_{\max}\Big(\Upsilon_{\mathfrak{O}^\ell_\sigma}[\eta] + \sum_{c\in\mathcal{C}} \Upsilon_{\mathfrak{T}^\ell_{\boldsymbol{c}}}[\eta] + \sum_{K\in\mathcal{M}\setminus\mathfrak{T}^\ell_{\mathcal{C}}} \|\nabla\eta^\perp\|^2_{L^2(K)}\Big)$$
$$+ \boldsymbol{p}^4_{\max} \sum_{\boldsymbol{e}\in\mathcal{E}} \Big(\Upsilon_{\mathfrak{T}^\ell_{\boldsymbol{e},1}}[\eta] + \Upsilon_{\mathfrak{T}^\ell_{\boldsymbol{e},2}}[\eta]\Big) + \boldsymbol{p}^4_{\max} \sum_{\boldsymbol{e}\in\mathcal{E}_D} \Upsilon_{\mathfrak{T}^\ell_{\boldsymbol{e},D}}[\eta]\ .$$

By property (6.3), we have

$$\Upsilon_{\mathfrak{O}^\ell_\sigma}[\eta] \lesssim \boldsymbol{p}^8_{\max}\big(\Upsilon_{\mathfrak{O}^\ell_\sigma}[\eta^\perp] + \Upsilon_{\mathfrak{O}^\ell_\sigma}[\eta^\parallel]\big), \quad \Upsilon_{\mathfrak{T}^\ell_{\boldsymbol{e},1}}[\eta] \lesssim \boldsymbol{p}^8_{\max}\big(\Upsilon_{\mathfrak{T}^\ell_{\boldsymbol{e},1}}[\eta^\perp] + \Upsilon_{\mathfrak{T}^\ell_{\boldsymbol{e},1}}[\eta^\parallel]\big),$$

as well as $\Upsilon_{\mathfrak{T}^\ell_{\boldsymbol{e},D}}[\eta] \lesssim \Upsilon_{\mathfrak{T}^\ell_{\boldsymbol{e},D}}[\eta^\perp] + \Upsilon_{\mathfrak{T}^\ell_{\boldsymbol{e},D}}[\eta^\parallel]$. This implies the assertion. ☐

6.3. **Exponential convergence.** We are now ready to state the main result of this paper.

Theorem 6.2. *Let the solution u of the boundary-value problem* (1.1)–(1.3) *in the axiparallel polyhedron $\Omega \subset \mathbb{R}^3$ belong to the analytic space $B_{-1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D)$, as in Proposition 2.3 and with a weight exponent vector $\boldsymbol{b}$ satisfying* (2.11). *Let the assumptions in Remark 2.5 be satisfied.*

Furthermore, let $\mathfrak{M}_\sigma = \{\mathcal{M}^{(\ell)}_\sigma\}_{\ell\geq 1}$ be a family of axiparallel σ-geometric meshes as introduced in Section 3.1, and consider the hp-dG discretizations in (4.1) *based on the sequences of approximating subspaces V^ℓ_σ and $V^\ell_{\sigma,\mathfrak{s}}$ defined in* (3.8) *respectively* (3.9)*, with the associated polynomial degree distributions $\boldsymbol{p}_1(\mathcal{M}^{(\ell)}_\sigma)$ (constant and uniform) respectively $\boldsymbol{p}_2(\mathcal{M}^{(\ell)}_\sigma)$ ($\mathfrak{s}$-linear and anisotropic). All polynomial degrees are assumed greater than or equal to* 3 *in interior elements $K \in \mathfrak{O}^\ell_\sigma$.*

Then for $\ell \geq 1$, the hp-dG approximation u_{DG} is well-defined, and as $\ell \to \infty$, the approximate solutions u_{DG} satisfy the error estimate

$$\|u - u_{DG}\|_{\mathrm{DG}} \leq C \exp\left(-b\sqrt[5]{N}\right), \tag{6.10}$$

where $N = \dim(V(\mathcal{M}^{(\ell)}_\sigma, \boldsymbol{\Phi}(\mathcal{M}^{(\ell)}_\sigma), \boldsymbol{p}(\mathcal{M}^{(\ell)}_\sigma)))$ denotes the number of degrees of freedom of the discretization for any of the two spaces V^ℓ_σ or $V^\ell_{\sigma,\mathfrak{s}}$.

The constants $b > 0$ and $C > 0$ are independent of N, but depend on σ, $\mathcal{M}^0$, θ, γ, $\min \boldsymbol{b} > 0$, and on which of the polynomial degree vectors $\boldsymbol{p}_1(\mathcal{M}^{(\ell)}_\sigma)$ or $\boldsymbol{p}_2(\mathcal{M}^{(\ell)}_\sigma)$ are used.

Remark 6.3. The assumption that polynomial degrees are greater than or equal to 3 in interior elements is purely technical; cp. Lemma 5.2. As for the pure Dirichlet case, we do not expect this assumption to be relevant in practice. This is corroborated by preliminary numerical tests which will be presented in a forthcoming computational study.

Remark 6.4. In particular, the *hp*-dG interpolant constructed to prove Theorem 6.2 yields an exponential approximation bound of the discretization error in the dG norm as in (6.10) for any $u \in B_{-1-\boldsymbol{b}}(\Omega;\emptyset,\emptyset)$.

Remark 6.5. We note that Theorem 6.2 remains true in the pure Neumann case. Indeed, the *hp*-approximation analysis on geometric meshes presented in this work as applied to the *hp*-dGFEM (4.1) with $\mathcal{F}_D(\mathcal{M}) = \emptyset$ and based on the *hp*-space $V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})/\mathbb{R}$ leads to the bound (6.10) as well. This simply follows from the fact that all the interpolants in our error analysis reproduce constant functions.

The proof of Theorem 6.2 will be detailed in Section 7, by proving that all the consistency terms in Theorem 6.1 are exponentially small for the hp-dG interpolant Π in (5.18).

7. Proof of Theorem 6.2

A geometric edge mesh $\mathcal{M}_\sigma^{(\ell)}$ consists of a finite number of patches $\{\mathcal{M}_j\}_{j=1}^J$. This makes it possible to bound the error terms in Theorem 6.1 separately on each patch $\mathcal{M}_j$. Moreover, due to the simple structure of the patch mappings, the weighted Sobolev space $N_{\boldsymbol{\beta}}^k(\mathcal{M}_j;\mathcal{C},\mathcal{E}_D)$, as restricted to a physical patch $\mathcal{M}_j$, can be identified with an equivalent space, which features the same regularity and is equipped with equivalent norms, on the associated reference patch $\widetilde{\mathcal{M}}_j$. Hence, it is sufficient to limit the proof of the exponential convergence bounds to geometrically refined reference patches as shown in Figure 1 (unrefined patches can be treated similarly to [12, Section 5.2.1]). Furthermore, by superposition arguments as in [12], it is enough to show exponential convergence bounds for *reference corner, edge, and corner-edge meshes* $\widehat{\mathcal{M}}_{\boldsymbol{c}}^\ell$, $\widehat{\mathcal{M}}_{\boldsymbol{e}}^\ell$ *and* $\widehat{\mathcal{M}}_{\boldsymbol{ce}}^\ell$, respectively, in the context of a single corner $\boldsymbol{c}$ and/or a single edge $\boldsymbol{e}$. Finally, since the meshes $\widehat{\mathcal{M}}_{\boldsymbol{c}}^\ell$ and $\widehat{\mathcal{M}}_{\boldsymbol{e}}^\ell$ can be viewed as collections of certain elements of $\widehat{\mathcal{M}}_{\boldsymbol{ce}}^\ell$, it is sufficient to consider *a single* reference corner-edge mesh $\widehat{\mathcal{M}}_{\boldsymbol{ce}}^\ell$, where $\boldsymbol{e}$ is either a Neumann or a Dirichlet edge.

7.1. **Reference corner-edge mesh.** We consider the reference patch $(0,1)^3$ with corner $\boldsymbol{c} = (\boldsymbol{0},0)$ and the single edge $\boldsymbol{e} = \{\boldsymbol{0}\}\times\omega_{\boldsymbol{c}}^{\|} \in \mathcal{E}_{\boldsymbol{c}}$, $\omega_{\boldsymbol{c}}^{\|} = (0,1)$, originating from it; cp. Figure 1 (right). We introduce the reference geometric corner-edge mesh $\widehat{\mathcal{M}}_{\boldsymbol{ce}}^\ell$ in $(0,1)^3$ by

$$\widehat{\mathcal{M}}_{\boldsymbol{ce}}^\ell = \bigcup_{j=1}^{\ell+1}\bigcup_{i=1}^{j}\widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{ij}, \tag{7.1}$$

where the sets $\widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{ij}$ stand for layers of elements with identical scaling properties; cp. [12, Section 5.2.4]. The decomposition in (7.1) is not a partition, in general: elements may be contained in several layers (but whose number is uniformly bounded with respect to ℓ). The index j indicates the number of the geometric mesh layers in edge-parallel direction along the edge $\omega_{\boldsymbol{c}}^{\|}$, whereas the index i indicates the number of mesh layers in direction perpendicular to $\omega_{\boldsymbol{c}}^{\|}$.

In agreement to Section 3.1, we split $\widehat{\mathcal{M}}_{\boldsymbol{ce}}^\ell$ into interior elements away from $\boldsymbol{c}$ and $\boldsymbol{e}$, boundary layer elements along $\boldsymbol{e}$ (but away from $\boldsymbol{c}$), and a corner element, $\widehat{\mathcal{M}}_{\boldsymbol{ce}}^\ell = \widehat{\mathfrak{O}}_{\boldsymbol{ce}}^\ell \,\dot\cup\, \widehat{\mathfrak{T}}_{\boldsymbol{e}}^\ell \,\dot\cup\, \widehat{\mathfrak{T}}_{\boldsymbol{c}}^\ell$, with

$$\widehat{\mathfrak{O}}_{\boldsymbol{ce}}^\ell := \bigcup_{j=2}^{\ell+1}\bigcup_{i=2}^{j}\widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{ij}, \qquad \widehat{\mathfrak{T}}_{\boldsymbol{e}}^\ell := \bigcup_{j=2}^{\ell+1}\widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{1j}, \qquad \widehat{\mathfrak{T}}_{\boldsymbol{c}}^\ell := \widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{11}. \tag{7.2}$$

In particular, an interior element $K \in \widehat{\mathfrak{O}}_{\boldsymbol{ce}}^\ell$ belongs to $\widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{ij}$ if it satisfies

$$r_{\boldsymbol{e}}|_K \simeq d_K^{\boldsymbol{e}} \simeq h_K^\perp \simeq \sigma^{\ell+1-i}, \qquad r_{\boldsymbol{c}}|_K \simeq d_K^{\boldsymbol{c}} \simeq h_K^{\|} \simeq \sigma^{\ell+1-j}, \tag{7.3}$$

for $2 \le i \le j \le \ell+1$. The terminal layers $\widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{1j}$ consist of elements $K \in \widehat{\mathfrak{T}}_{\boldsymbol{e}}^\ell$ with

$$r_{\boldsymbol{e}}|_K \simeq d_K^{\boldsymbol{e}} \lesssim h_K^\perp \simeq \sigma^{\ell}, \qquad r_{\boldsymbol{c}}|_K \simeq d_K^{\boldsymbol{c}} \simeq h_K^{\|} \simeq \sigma^{\ell+1-j}, \tag{7.4}$$

for $2 \le j \le \ell+1$. Finally, an element in the layer $\widehat{\mathfrak{T}}^\ell_{\boldsymbol{c}} = \widehat{\mathfrak{L}}^{11}_{\boldsymbol{ce}}$ is isotropic with $r_{\boldsymbol{e}}|_K \simeq d^{\boldsymbol{e}}_K \lesssim h_K \simeq \sigma^\ell$, and $r_{\boldsymbol{c}}|_K \simeq d^{\boldsymbol{c}}_K \lesssim h_K \simeq \sigma^\ell$. The sets $\widehat{\mathfrak{L}}^{1j}_{\boldsymbol{ce}}$ and $\widehat{\mathfrak{L}}^{11}_{\boldsymbol{ce}}$ are in fact singletons, and $K \in \widehat{\mathfrak{L}}^{1j}_{\boldsymbol{ce}}$ can be written as

$$K_j = K^\perp \times K^\parallel_j, \qquad 2 \le j \le \ell+1, \tag{7.5}$$

where $K^\perp = (0,\sigma^\ell)^2$, and the sequence $\{K^\parallel_j\}_{j=2}^{\ell+1}$ forms a one-dimensional geometric mesh $\mathcal{T}^{(\ell)}_\sigma$ along the edge $\omega^\parallel_{\boldsymbol{c}} = (0,1)$ as in Section 5.2; moreover, the corner element $K \in \widehat{\mathfrak{T}}^\ell_{\boldsymbol{c}}$ is given by $K = (0,\sigma^\ell)^3$. In agreement with Section 3.1, we consider $\mathfrak{s}$-linearly increasing polynomial degree distributions on $\widehat{\mathcal{M}}^\ell_{\boldsymbol{ce}}$ that satisfy, for $1 \le i \le j \le \ell+1$,

$$\forall\, K \in \widehat{\mathfrak{L}}^{ij}_{\boldsymbol{ce}}: \qquad \boldsymbol{p}_K = (p^\perp_i, p^\parallel_j) \simeq (\max\{\lceil \mathfrak{s} i\rceil, 3\}, \max\{\lceil \mathfrak{s} j\rceil, 3\})\,. \tag{7.6}$$

Analogous to the definition of the reference corner-edge mesh $\widehat{\mathcal{M}}^\ell_{\boldsymbol{ce}}$, we introduce the *reference corner mesh* $\widehat{\mathcal{M}}^\ell_{\boldsymbol{c}}$ and the *reference edge mesh* $\widehat{\mathcal{M}}^\ell_{\boldsymbol{e}}$, cp. Figure 1. For the purpose of deriving the ensuing exponential convergence estimates it is important that, without loss of generality, *the geometric meshes* $\widehat{\mathcal{M}}^\ell_{\boldsymbol{c}}$ *and* $\widehat{\mathcal{M}}^\ell_{\boldsymbol{e}}$ *can be characterized as collections of certain elements* $K \in \widehat{\mathcal{M}}^\ell_{\boldsymbol{ce}}$. More precisely, for $\ell \ge 2$, and with $\widehat{\mathfrak{L}}^{ij}_{\boldsymbol{ce}}$ as in (7.1), we define

$$\widehat{\mathcal{M}}^\ell_{\boldsymbol{c}} := \widehat{\mathfrak{O}}^\ell_{\boldsymbol{c}} \cup \widehat{\mathfrak{T}}^\ell_{\boldsymbol{c}}\,, \qquad \widehat{\mathfrak{O}}^\ell_{\boldsymbol{c}} := \bigcup_{j=2}^{\ell+1} \widehat{\mathfrak{L}}^{jj}_{\boldsymbol{ce}}\,, \qquad \widehat{\mathfrak{T}}^\ell_{\boldsymbol{c}} := \widehat{\mathfrak{L}}^{11}_{\boldsymbol{ce}}\,, \tag{7.7}$$

$$\widehat{\mathcal{M}}^\ell_{\boldsymbol{e}} := \widehat{\mathfrak{O}}^\ell_{\boldsymbol{e}} \cup \widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}\,, \qquad \widehat{\mathfrak{O}}^\ell_{\boldsymbol{e}} := \bigcup_{i=2}^{\ell+1} \widehat{\mathfrak{L}}^{i,\ell+1}_{\boldsymbol{ce}}\,, \qquad \widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}} := \widehat{\mathfrak{L}}^{1,\ell+1}_{\boldsymbol{ce}}\,. \tag{7.8}$$

Here, we remark that we abuse notation slightly in that the definition of $\widehat{\mathfrak{O}}^\ell_{\boldsymbol{c}}$ and $\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}$ in (7.7) and (7.8) differs from (7.2).

In the sequel, we denote the domains formed by all elements in $\widehat{\mathcal{M}}^\ell_{\boldsymbol{ce}}$, $\widehat{\mathcal{M}}^\ell_{\boldsymbol{c}}$ and $\widehat{\mathcal{M}}^\ell_{\boldsymbol{e}}$ by $\widehat{\Omega}^\ell_{\boldsymbol{ce}}$, $\widehat{\Omega}^\ell_{\boldsymbol{c}}$ and $\widehat{\Omega}^\ell_{\boldsymbol{e}}$, respectively. Let now $\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}} \cap \mathcal{E}_N$ be a Neumann edge. By the regularity property (2.12), the definition of the weighted semi-norm (2.8), and for exponents $b_{\boldsymbol{c}}, b_{\boldsymbol{e}}$, we introduce the *corner-edge semi-norm on* $\widehat{\Omega}^\ell_{\boldsymbol{ce}}$ as

$$|u|^2_{\widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} := \sum_{|\boldsymbol{\alpha}|=k} \left\| r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+|\boldsymbol{\alpha}|} \rho_{\boldsymbol{ce}}^{\max\{-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\widehat{\Omega}^\ell_{\boldsymbol{ce}})}, \quad k \ge 0. \tag{7.9}$$

Under the assumption $b_{\boldsymbol{c}}, b_{\boldsymbol{e}} \in (0,1)$ as in Remark 2.5 and for $\alpha^\parallel \ge 0$, the norms on the right-hand side of (7.9) take the form:

$$\begin{cases} \|r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+\alpha^\parallel} \mathsf{D}^{\alpha^\parallel}_\parallel u\|^2_{L^2(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} & |\boldsymbol{\alpha}^\perp| = 0, \\ \|r_{\boldsymbol{c}}^{-b_{\boldsymbol{c}}+\alpha^\parallel} \mathsf{D}^{\boldsymbol{\alpha}^\perp}_\perp \mathsf{D}^{\alpha^\parallel}_\parallel u\|^2_{L^2(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} & |\boldsymbol{\alpha}^\perp| = 1, \\ \|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}+\alpha^\parallel} r_{\boldsymbol{e}}^{-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|} \mathsf{D}^{\boldsymbol{\alpha}^\perp}_\perp \mathsf{D}^{\alpha^\parallel}_\parallel u\|^2_{L^2(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} & |\boldsymbol{\alpha}^\perp| \ge 2. \end{cases} \tag{7.10}$$

For $m > k_{\boldsymbol{\beta}}$ as in (2.9), the corresponding weighted spaces $\widehat{N}^m_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$ are defined as in Section 2.2 with respect to the norm $\|\cdot\|^2_{\widehat{N}^m_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} = \sum_{k=0}^m |\cdot|^2_{\widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})}$.

Under the analytic regularity property in Proposition 2.3, the solution u to problem (1.1)–(1.3), localized and scaled to $\widehat{\Omega}^\ell_{\boldsymbol{ce}}$, belongs to $B_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$, that is, we have $u \in \widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$ for $k > k_{\boldsymbol{\beta}}$ and there is a constant $d_u > 0$ such that

$$|u|_{\widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} \leq d_u^{k+1} k! \,, \qquad k > k_{\boldsymbol{\beta}} \,. \tag{7.11}$$

In the reference corner domain $\widehat{\Omega}^\ell_{\boldsymbol{c}}$ and the reference edge mesh $\widehat{\Omega}^\ell_{\boldsymbol{e}}$ defined in (7.7) and (7.8), respectively, expressions analogous (but simpler) to (7.9) result: since $\rho_{\boldsymbol{ce}}|_{\widehat{\Omega}^\ell_{\boldsymbol{c}}} = \mathcal{O}(1)$, we introduce the *corner semi-norm on* $\widehat{\Omega}^\ell_{\boldsymbol{c}}$ by:

$$|u|^2_{\widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{c}})} := \sum_{|\boldsymbol{\alpha}|=k} \left\| r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+|\boldsymbol{\alpha}|} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\widehat{\Omega}^\ell_{\boldsymbol{c}})} , \qquad k \geq 0. \tag{7.12}$$

In the reference Neumann edge mesh $\widehat{\Omega}^\ell_{\boldsymbol{e}}$, since $r_{\boldsymbol{c}}|_{\widehat{\Omega}^\ell_{\boldsymbol{e}}} = \mathcal{O}(1)$, we define the *edge semi-norm on* $\widehat{\Omega}^\ell_{\boldsymbol{e}}$ as:

$$|u|^2_{\widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{e}})} = \sum_{|\boldsymbol{\alpha}|=k} \left\| r_{\boldsymbol{e}}^{\max\{-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\widehat{\Omega}^\ell_{\boldsymbol{e}})} , \qquad k \geq 0. \tag{7.13}$$

The weighted spaces $\widehat{N}^m_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{c}})$ and $\widehat{N}^m_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{e}})$ are defined as before, for $m > k_\beta$.

7.2. **Exponential convergence at Neumann edges.** In this section, we establish exponential convergence estimates for the consistency bounds in Theorem 6.1 over the *reference corner-edge submesh* $\widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}} \,\dot\cup\, \widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}$ and for a Neumann edge $\boldsymbol{e}$. Due to (7.7) and (7.8), *the required exponential convergence bounds for the basic geometric (sub)meshes* $\widehat{\mathfrak{O}}^\ell_{\boldsymbol{c}}$ *and* $\widehat{\mathcal{M}}^\ell_{\boldsymbol{e}}$ *will follow as a special case.*

For a function $u \in B_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$, cp. (7.11), we define the elemental dG approximation operator Π on $\widehat{\mathcal{M}}^\ell_{\boldsymbol{ce}}$ as in (5.18) and in accordance with the partition in (7.1). As in (6.1), we then write $\eta = u - \Pi u$ for the dG approximation error $\eta = u - \Pi u$ on $\widehat{\mathcal{M}}^\ell_{\boldsymbol{ce}}$. On the submesh $\widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}} \,\dot\cup\, \widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}$, we set $\eta^\perp = u - \Pi^\perp u$, $\eta^\| = u - \Pi^\| u$, analogous to (6.2). In view of the error estimates in Theorem 6.1, we will now bound the contributions $\Upsilon_{\widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}}}$, $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},1}}$, $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},2}}$, and $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{c}}}$, where these terms are defined exactly as in (6.4) (but over the reference mesh $\widehat{\mathcal{M}}^\ell_{\boldsymbol{ce}}$).

Theorem 7.1. *Let* $\boldsymbol{e} \in \mathcal{E}_N$ *be a Neumann edge,* $u \in B_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$*, cp.* (7.11)*, with weight exponents* $b_{\boldsymbol{c}}, b_{\boldsymbol{e}}$ *as in Remark 2.5. Then in the setting of Section 7.1 and for* ℓ *sufficiently large, there exist constants* $b, C > 0$ *such that*

$$\Upsilon_{\widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}}}[\eta^\perp] + \Upsilon_{\widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}}}[\eta^\|] + \Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},1}}[\eta^\perp] + \Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},1}}[\eta^\|] + \Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},2}}[\eta] \leq C \exp(-2b\ell) \,. \tag{7.14}$$

Analogous exponential bounds hold for the consistency terms $\Upsilon_{\widehat{\mathfrak{O}}^\ell_{\boldsymbol{c}}}$ *and* $\Upsilon_{\widehat{\mathfrak{O}}^\ell_{\boldsymbol{e}}}$.

The remainder of this subsection is devoted to the proof of Theorem 7.1. We proceed in several steps. Note that the proofs for the terms $\Upsilon_{\widehat{\mathfrak{O}}^\ell_{\boldsymbol{c}}}$, $\Upsilon_{\widehat{\mathfrak{O}}^\ell_{\boldsymbol{e}}}$ are analogous and will not be detailed; see (7.7) and (7.8).

7.2.1. *Convergence of* $\Upsilon_{\widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}}}[\eta^\perp]$. The following auxiliary result holds.

Lemma 7.2. *Let* $K \in \widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}}$*,* $v : K \to \mathbb{R}$*, and* $\widehat{v} = v \circ \Phi_K^{-1} \in H^2_{\mathrm{mix}}(\widehat{K})$*; cp.* (5.6)*. Then there holds* $T^K_{\mathfrak{O}}[v] \lesssim h^\|_K \|\widehat{v}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})}$.

Proof. The scaling properties in [12, Section 5.1.4] imply

$$(h_K^{\|})^{-2}\|v\|^2_{L^2(K)} + \|\mathsf{D}_{\|}v\|^2_{L^2(K)} + (h_K^{\|})^2\|\mathsf{D}_{\|}^2 v\|^2_{L^2(K)} \lesssim (h_K^{\perp})^2(h_K^{\|})^{-1}\|\widehat{v}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})},$$

as well as $\|\mathsf{D}_{\perp}v\|^2_{L^2(K)} + (h_K^{\perp})^2\|\mathsf{D}_{\perp}^2 v\|^2_{L^2(K)} \lesssim h_K^{\|}\|\widehat{v}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})}$. These scalings and the fact that $h_K^{\perp} \lesssim h_K^{\|}$ for $K \in \widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}$ yield the result. $\square$

Proposition 7.3. *For ℓ sufficiently large, there are constants $b, C > 0$ such that $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\perp}] \leq C\exp(-2b\ell)$.*

Proof. We consider an element $K \in \widehat{\mathfrak{L}}^{ij}_{\boldsymbol{ce}}$ with $2 \leq j \leq \ell+1$ and $2 \leq i \leq j$, according to (7.2). With Proposition 5.3, Lemma 7.2 and (7.6), we obtain

$$T^K_{\mathfrak{O}}[\eta^{\perp}] \lesssim |\boldsymbol{p}_K|^{16}\, h_K^{\|}\,\Psi_{p_i^{\perp}-1,s_i^{\perp}-1} \times \sum_{\substack{s_i^{\perp}+1\leq|\boldsymbol{\alpha}^{\perp}|\leq s_i^{\perp}+3\\ 0\leq\alpha^{\|}\leq 2}} (h_K^{\perp})^{2|\boldsymbol{\alpha}^{\perp}|-2}(h_K^{\|})^{2\alpha^{\|}-1}\|\mathsf{D}_{\perp}^{\boldsymbol{\alpha}^{\perp}}\mathsf{D}_{\|}^{\alpha^{\|}}u\|^2_{L^2(K)},$$

for any $3 \leq s_i^{\perp} \leq p_i^{\perp}$. Thanks to the equivalences (7.4) on K, we may insert the appropriate weights as in (7.10) to obtain

$$\|\mathsf{D}_{\perp}^{\boldsymbol{\alpha}^{\perp}}\mathsf{D}_{\|}^{\alpha^{\|}}u\|^2_{L^2(K)} \simeq (d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}-2\alpha^{\|}}(d_K^{\boldsymbol{e}})^{2+2b_{\boldsymbol{e}}-2|\boldsymbol{\alpha}^{\perp}|} \times \|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}+\alpha^{\|}} r_{\boldsymbol{e}}^{-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^{\perp}|}\mathsf{D}_{\perp}^{\boldsymbol{\alpha}^{\perp}}\mathsf{D}_{\|}^{\alpha^{\|}}u\|^2_{L^2(K)}\,.$$

From the analytic regularity (7.11) there exists a constant $C > 0$ such that

$$T^K_{\mathfrak{O}}[\eta^{\perp}] \lesssim \boldsymbol{p}_{\max}^{16}\,\Psi_{p_i^{\perp}-1,s_i^{\perp}-1}(d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}}(d_K^{\boldsymbol{e}})^{2b_{\boldsymbol{e}}}C^{2s_i^{\perp}}\Gamma(s_i^{\perp}+6)^2, \tag{7.15}$$

for all $3 \leq s_i^{\perp} \leq p_i^{\perp}$. Summing (7.15) over all layers in $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}$ in (7.2) with the use of (7.3) results in

$$\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\perp}] \lesssim \boldsymbol{p}_{\max}^{16}\sum_{j=2}^{\ell+1}\sigma^{2(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})(\ell+1-j)}\sum_{i=2}^{j}\sigma^{2b_{\boldsymbol{e}}(\ell+1-i)}\Psi_{p_i^{\perp}-1,s_i^{\perp}-1}C^{2s_i^{\perp}}\Gamma(s_i^{\perp}+6)^2\,.$$

By interpolating to real parameters $s_i^{\perp} \in [3, p_i^{\perp}]$ as in [12, Lemma 5.8], this sum is of exactly the same form as $S^{\perp}$ in the proof of [12, Proposition 5.17], and the assertion now follows from the arguments there and after adjusting the constants to absorb the algebraic loss in $\boldsymbol{p}_{\max}$. $\square$

7.2.2. *Convergence of* $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\|}]$. To establish the analog of Proposition 7.3 in edge-parallel direction, we make use of the following auxiliary estimate.

Lemma 7.4. *Let $K \in \widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}$ and $3 \leq s_K^{\|} \leq p_K^{\|}$. Then there holds $T^K_{\mathfrak{O}}[\eta^{\|}] \lesssim (p_K^{\|})^8\,(S_{1,K}[u] + S_{2,K}[u])$, with*

$$S_{1,K}[u] \lesssim \Psi_{p_K^{\|}-1,s_K^{\|}-1}(d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}}\|u\|^2_{\widehat{N}^{s_K^{\|}+2}_{-1-\boldsymbol{b}}(K)},$$

$$S_{2,K}[u] \lesssim \Psi_{p_K^{\|}-1,s_K^{\|}-1}(d_K^{\boldsymbol{e}})^{2b_{\boldsymbol{e}}}(d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}}\|u\|^2_{\widehat{N}^{s_K^{\|}+3}_{-1-\boldsymbol{b}}(K)}.$$

Proof. From scaling arguments as in [12, Section 5.1.4] and the approximation property in Proposition 5.3 for $\eta^{\parallel}$ (with $|\boldsymbol{\alpha}^{\perp}|=0$), we obtain

$$\begin{aligned}&(h_K^{\parallel})^{-2}\|\eta^{\parallel}\|^2_{L^2(K)}+\|\mathsf{D}_{\parallel}\eta^{\parallel}\|^2_{L^2(K)}+(h_K^{\parallel})^2\|\mathsf{D}^2_{\parallel}\eta^{\parallel}\|^2_{L^2(K)}\\&\lesssim (h_K^{\perp})^2(h_K^{\parallel})^{-1}\sum_{0\le\alpha^{\parallel}\le 2}\|\widehat{\mathsf{D}}_{\parallel}^{\alpha^{\parallel}}\widehat{\eta}\|^2_{L^2(\widehat{K})}\lesssim (p_K^{\parallel})^8\Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1}(h_K^{\parallel})^{2s_K^{\parallel}}\|\mathsf{D}_{\parallel}^{s_K^{\parallel}+1}u\|^2_{L^2(K)}\,.\end{aligned}$$

Then, we insert the corner weight with the aid of (7.10), (7.3) to find that

$$\begin{aligned}\|\mathsf{D}_{\parallel}^{s_K^{\parallel}+1}u\|^2_{L^2(K)}&\simeq (d_K^{\boldsymbol{c}})^{2+2b_{\boldsymbol{c}}-2s_K^{\parallel}-2}\|r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+s_K^{\parallel}+1}\mathsf{D}_{\parallel}^{s_K^{\parallel}+1}u\|^2_{L^2(K)}\\&\lesssim (d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}-2s_K^{\parallel}}|u|^2_{\widehat{N}^{s_K^{\parallel}+1}_{-1-\boldsymbol{b}}(K)}.\end{aligned}$$

Combining the two estimates above shows that

$$(h_K^{\parallel})^{-2}\|\eta^{\parallel}\|^2_{L^2(K)}+\|\mathsf{D}_{\parallel}\eta^{\parallel}\|^2_{L^2(K)}+(h_K^{\parallel})^2\|\mathsf{D}^2_{\parallel}\eta^{\parallel}\|^2_{L^2(K)}\lesssim (p_K^{\parallel})^8S_{1,K}[u].$$

With $|\boldsymbol{\alpha}^{\perp}|=1$ and (7.10), we conclude analogously that

$$\begin{aligned}&\|\mathsf{D}_{\perp}\eta^{\parallel}\|^2_{L^2(K)}\lesssim h_K^{\parallel}\|\widehat{\mathsf{D}}_{\perp}\widehat{\eta}^{\parallel}\|^2_{L^2(\widehat{K})}\\&\lesssim (p_K^{\parallel})^8\Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1}(h_K^{\parallel})^{2s_K^{\parallel}+2}\|\mathsf{D}_{\perp}\mathsf{D}_{\parallel}^{s_K^{\parallel}+1}u\|^2_{L^2(K)}\\&\lesssim (p_K^{\parallel})^8\Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1}(d_K^{\boldsymbol{c}})^{2s_K^{\parallel}+2}(d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}-2s_K^{\parallel}-2}\|r_{\boldsymbol{c}}^{-b_{\boldsymbol{c}}+s_K^{\parallel}+1}\mathsf{D}_{\perp}\mathsf{D}_{\parallel}^{s_K^{\parallel}+1}u\|^2_{L^2(K)}\\&\lesssim (p_K^{\parallel})^8\Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1}(d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}}|u|^2_{\widehat{N}^{s_K^{\parallel}+2}_{-1-\boldsymbol{b}}(K)}\lesssim (p_K^{\parallel})^8S_{1,K}[u].\end{aligned}$$

It remains to bound the term $(h_K^{\perp})^2\|\mathsf{D}^2_{\perp}\eta^{\parallel}\|^2_{L^2(K)}$. To do so, we proceed along the same lines, for $|\boldsymbol{\alpha}^{\perp}|=2$. With (7.10), we obtain

$$\begin{aligned}&(h_K^{\perp})^2\|\mathsf{D}^2_{\perp}\eta^{\parallel}\|^2_{L^2(K)}\lesssim h_K^{\parallel}\|\widehat{\mathsf{D}}^2_{\perp}\widehat{\eta}^{\parallel}\|^2_{L^2(\widehat{K})}\\&\lesssim (p_K^{\parallel})^8\Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1}(h_K^{\perp})^2(h_K^{\parallel})^{2s_K^{\parallel}+2}\|\mathsf{D}^2_{\perp}\mathsf{D}_{\parallel}^{s_K^{\parallel}+1}u\|^2_{L^2(K)}\\&\lesssim (p_K^{\parallel})^8\Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1}(d_K^{\boldsymbol{e}})^{2b_{\boldsymbol{e}}}(d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}}\|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}+s_K^{\parallel}+1}r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}^2_{\perp}\mathsf{D}_{\parallel}^{s_K^{\parallel}+1}u\|^2_{L^2(K)}\\&\lesssim (p_K^{\parallel})^8S_{2,K}[u],\end{aligned}$$

which finishes the proof. □

Proposition 7.5. *For ℓ sufficiently large, there are constants $b,C>0$ such that $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\parallel}]\le C\exp(-2b\ell)$.*

Proof. By summing the result of Lemma 7.4 over all layers of $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}$ and noticing (7.3), (7.6), as well as the analytic regularity (7.11), we conclude that $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\parallel}]\lesssim \boldsymbol{p}^8_{\max}\big(S_1+S_2\big)$, where the sums S_1 and S_2 are given by

$$\begin{aligned}S_1&=\sum_{j=2}^{\ell+1}\sum_{i=2}^{j}\Psi_{p_j^{\parallel}-1,s_j^{\parallel}-1}\sigma^{2(\ell+1-j)b_{\boldsymbol{c}}}C^{2s_j^{\parallel}}\Gamma(s_j^{\parallel}+3)^2,\\S_2&=\sum_{j=2}^{j}\sum_{i=2}^{j}\Psi_{p_j^{\parallel}-1,s_j^{\parallel}-1}\sigma^{2(\ell+1-i)b_{\boldsymbol{e}}}\sigma^{2(\ell+1-j)(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})}C^{2s_j^{\parallel}}\Gamma(s_j^{\parallel}+4)^2.\end{aligned}$$

The terms in the first sum S_1 are independent of the inner index i. Hence, by interpolation to real parameters $s_j^{\|} \in [3, p_j^{\|}]$ as in [12, Lemma 5.8], by applying [12, Lemma 5.12], and after possibly adjusting constants, we conclude $S_1 \lesssim \ell \exp(-2b_1(\ell+1)) \lesssim \exp(-2b_2\ell)$. The second sum S_2 can be estimated in exactly the same manner as the sum $S^{\|}$ in the proof of [12, Proposition 5.17], and we obtain $S_2 \lesssim \exp(-2b_3\ell)$. Adjusting the constants to absorb the algebraic factor $\boldsymbol{p}_{\max}^8$ yields the assertion. □

7.2.3. *Convergence of* $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},1}}[\eta^\perp]$. We show the following auxiliary bounds for $\eta^\perp$, by using properties of the quasi-interpolation operator $\mathcal{I}_1^\perp$ in edge-perpendicular direction on $\mathfrak{K} = K^\perp$.

Lemma 7.6. *Let $K = K^\perp \times K_j^{\|}$, $j \geq 2$, be an element in the terminal layer $\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}$ of the form* (7.5). *Then there holds*

$$(h_K^{\|})^{-2}\|\eta^\perp\|^2_{L^2(K)} + \|\mathsf{D}_\perp\eta^\perp\|^2_{L^2(K} \lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}\|u\|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}\,, \tag{7.16}$$

and

$$(h_K^{\|})^{2}\|\mathsf{D}^2_{\|}\eta^\perp\|^2_{L^2(K)} + \|\mathsf{D}_{\|}\eta^\perp\|^2_{L^2(K)} \lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}\|u\|^2_{\widehat{N}^4_{-1-\boldsymbol{b}}(K)}\,. \tag{7.17}$$

Proof. To show (7.16), let $s = 0, 1$ and $|\boldsymbol{\alpha}^\perp| = s$. By Lemma 5.6, item (7) (with $\beta = 1 - b_{\boldsymbol{e}}$), we get

$$(h_K^{\|})^{2(s-1)}\|\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}\eta^\perp\|^2_{L^2(K)} \lesssim (h_K^{\|})^{2s-2}(h_K^\perp)^{4-2s-2(1-b_{\boldsymbol{e}})}\|r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}^2_\perp u\|^2_{L^2(K)}\,.$$

From the equivalences in (7.4), we further obtain

$$\begin{aligned}\|r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}^2_\perp u\|^2_{L^2(K)} &\lesssim (h_K^{\|})^{-2(b_{\boldsymbol{e}}-b_{\boldsymbol{c}})}\|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}}r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}^2_\perp u\|^2_{L^2(K)}\\ &\lesssim (h_K^{\|})^{-2b_{\boldsymbol{e}}+2b_{\boldsymbol{c}}}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}\,.\end{aligned}$$

Thus, combining these estimates and expressing the mesh sizes in terms of σ, cp. (7.4), we see that

$$\begin{aligned}(h_K^{\|})^{2(s-1)}\|\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}\eta^\perp\|^2_{L^2(K)} &\lesssim (h_K^{\|})^{2s-2-2b_{\boldsymbol{e}}+2b_{\boldsymbol{c}}}(h_K^\perp)^{2-2s+2b_{\boldsymbol{e}}}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}\\ &\simeq \sigma^{2b_{\boldsymbol{c}}(\ell+1-j)+2b_{\boldsymbol{e}}(j-1)}\sigma^{2j(1-s)+2(s-1)}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}\\ &\lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)},\end{aligned}$$

which yields (7.16).

To prove (7.17), we proceed similarly and obtain, for $s = 1, 2$,

$$\begin{aligned}(h_K^{\|})^{2(s-1)}\|\mathsf{D}^s_{\|}\eta^\perp\|^2_{L^2(K)} &\lesssim (h_K^{\|})^{2s-2}(h_K^\perp)^{4-2(1-b_{\boldsymbol{e}})}\|r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}^2_\perp\mathsf{D}^s_{\|}u\|^2_{L^2(K)}\\ &\lesssim (h_K^{\|})^{-2-2(b_{\boldsymbol{e}}-b_{\boldsymbol{c}})}(h_K^\perp)^{2+2b_{\boldsymbol{e}}}\|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}+s}r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}^2_\perp\mathsf{D}^s_{\|}u\|^2_{L^2(K)}\\ &\lesssim \sigma^{2b_{\boldsymbol{c}}(\ell+1-j)+2b_{\boldsymbol{e}}(j-1)}\sigma^{2(j-1)}|u|^2_{\widehat{N}^{s+2}_{-1-\boldsymbol{b}}(K)}\\ &\lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}|u|^2_{\widehat{N}^{s+2}_{-1-\boldsymbol{b}}(K)}\,.\end{aligned}$$

This completes the proof. □

As a consequence of the preceding lemma, we have the following approximation bound in perpendicular direction.

Proposition 7.7. *For ℓ sufficiently large, there are constants $b, C > 0$ such that $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},1}}[\eta^\perp] \le C\exp(-2b\ell)$.*

Proof. From Lemma 7.6 we find that $T^K_{\boldsymbol{e},1}[\eta^\perp] \lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}\|u\|^2_{\widehat{N}^4_{-1-\boldsymbol{b}}(K)}$, for any $K \in \widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}$. The assertion now follows by summing this estimate over all elements $K \in \widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}$ (i.e., over $2 \le j \le \ell+1$) and by suitably adjusting constants. $\square$

7.2.4. *Convergence of* $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},1}}[\eta^\parallel]$. A similar estimate holds for the approximation error $\eta^\parallel$ in direction parallel to $\boldsymbol{e}$.

Proposition 7.8. *For ℓ sufficiently large, there are constants $b, C > 0$ such that $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},1}}[\eta^\parallel] \le C\exp(-2b\ell)$.*

Proof. We note that, by (7.10), the functions u and $\mathsf{D}_\perp u$ satisfy, respectively,

$$\|r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+\alpha^\parallel}\mathsf{D}_\parallel^{\alpha^\parallel}u\|_{L^2(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} \le C^{\alpha^\parallel+1}\Gamma(\alpha^\parallel+1), \qquad \alpha^\parallel \ge 2,$$
$$\|r_{\boldsymbol{c}}^{-b_{\boldsymbol{c}}+\alpha^\parallel}\mathsf{D}_\parallel^{\alpha^\parallel}\mathsf{D}_\perp u\|_{L^2(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} \le C^{\alpha^\parallel+2}\Gamma(\alpha^\parallel+2), \qquad \alpha^\parallel \ge 1\ .$$

In view of (7.4), (7.5), these properties correspond to the one-dimensional analytic regularity assumptions considered in (5.11) and (5.15), respectively. Moreover, due to (7.6), the polynomial degrees $p^\parallel_K$ are $\mathfrak{s}$-linearly increasing away from the corner $\boldsymbol{c}$. Hence, Proposition 5.4 respectively Proposition 5.5, and the tensor-product structure of the elements yield

$$\sum_{K\in\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}}\Big((h^\parallel_K)^{-2}\|\eta^\parallel\|^2_{L^2(K)} + \|\mathsf{D}_\parallel\eta^\parallel\|^2_{L^2(K)} + (h^\parallel_K)^2\|\mathsf{D}^2_\parallel\eta^\parallel\|^2_{L^2(K)}\Big) \lesssim \exp(-2b\ell),$$

respectively, $\sum_{K\in\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}}\|\mathsf{D}_\perp\eta^\parallel\|^2_{L^2(K)} \lesssim \exp(-2b\ell)$. This completes the proof. $\square$

7.2.5. *Convergence of* $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},2}}[\eta]$. Finally, we bound the term in $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},2}}[\eta]$.

Proposition 7.9. *Let u be in $\widehat{N}^2_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$.*

1. *For $K \in \widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}$, there holds $T^K_{\boldsymbol{e},2}[\eta] \lesssim (h^\perp_K)^{2b_{\boldsymbol{e}}}(h^\parallel_K)^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}$.*
2. *For ℓ sufficiently large, there are constants $b, C > 0$ such that $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},2}}[\eta] \le C\exp(-2b\ell)$.*

Proof. To show item (1), we note that, by Hölder's inequality and due to the fact that $b_{\boldsymbol{c}}, b_{\boldsymbol{e}} \in (0,1)$, cp. Remark 2.5, and $\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}\eta = \mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}u$ for $|\boldsymbol{\alpha}^\perp| = 2$ (since $\mathcal{I}^\perp_1 u \in \mathbb{P}_1(K^\perp)$), there holds

$$\sum_{|\boldsymbol{\alpha}^\perp|=2}\|\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}\eta\|^2_{L^1(K)} \lesssim \|r_{\boldsymbol{c}}^{-1+b_{\boldsymbol{c}}}\rho_{\boldsymbol{ce}}^{-1+b_{\boldsymbol{e}}}\|^2_{L^2(K)}\sum_{|\boldsymbol{\alpha}^\perp|=2}\|r_{\boldsymbol{c}}^{1-b_{\boldsymbol{c}}}\rho_{\boldsymbol{ce}}^{1-b_{\boldsymbol{e}}}\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}\eta\|^2_{L^2(K)}$$
$$\le \|r_{\boldsymbol{c}}^{b_{\boldsymbol{c}}-b_{\boldsymbol{e}}}r_{\boldsymbol{e}}^{-1+b_{\boldsymbol{e}}}\|^2_{L^2(K)}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})}.$$

Then, employing (7.4) in direction parallel to $\boldsymbol{e}$ yields $\|r_{\boldsymbol{c}}^{b_{\boldsymbol{c}}-b_{\boldsymbol{e}}}r_{\boldsymbol{e}}^{-1+b_{\boldsymbol{e}}}\|^2_{L^2(K)} \simeq (h^\parallel_K)^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}}\|r_{\boldsymbol{e}}^{-1+b_{\boldsymbol{e}}}\|^2_{L^2(K)}$. Since $|K| \simeq h^\parallel_K(h^\perp_K)^2$, we also have $\|r_{\boldsymbol{e}}^{-1+b_{\boldsymbol{e}}}\|^2_{L^2(K)} \lesssim |K|(h^\perp_K)^{2b_{\boldsymbol{e}}-2}$. Therefore, referring to (7.4) yields

$$T^K_{\boldsymbol{e},2}[\eta] \lesssim \sigma^{2b_{\boldsymbol{c}}(\ell+1-j)+2b_{\boldsymbol{e}}(j-1)}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)} \lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)},$$

which yields item (1). Summing this last bound over all elements $K \in \widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}$ yields item (2). $\square$

7.2.6. *Conclusion.* The proof of the exponential convergence bound (7.14) on the *hp*-dG interpolation error η in Theorem 7.1 follows now straightforwardly by using the above results.

7.3. **Exponential convergence at Dirichlet edges.** For the reference corner-edge patch $\widehat{\mathcal{M}}^{\ell}_{\boldsymbol{ce}}$ in the setting of Section 7.1, we then consider the case where $\boldsymbol{e} \in \mathcal{E}_D$ is a Dirichlet edge, and establish the analog of Theorem 7.1 for the submesh $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}} \dot\cup \widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}$. Note that, if $\boldsymbol{e}$ is a Dirichlet edge, we also need to estimate $\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},D}}$ given as in (6.5). According to (2.8) and [3], the solution regularity in the Dirichlet case is characterized by the *homogeneous corner-edge semi-norms*

$$|u|^2_{\widehat{M}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})} = \sum_{|\boldsymbol{\alpha}|=k} \left\| r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+|\boldsymbol{\alpha}|} \rho_{\boldsymbol{ce}}^{-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^{\perp}|} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})}, \qquad k \geq 0. \tag{7.18}$$

For $m > k_{\boldsymbol{\beta}}$, the weighted spaces $\widehat{M}^m_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ are defined accordingly. We say a function $u \in H^1(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ belongs to $A_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ if $u \in \widehat{M}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$, for $k > k_{\boldsymbol{\beta}}$, and there is a constant $d_u > 0$ such that

$$|u|_{\widehat{M}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})} \leq d_u^{k+1} k!\,, \qquad \forall k > k_{\boldsymbol{\beta}}\,. \tag{7.19}$$

While exponential convergence for solutions with regularity in this family of spaces was already shown in [12], we present an alternative argument, based on the preceding analysis of the Neumann case.

Corollary 7.10. *Let $\boldsymbol{e} \in \mathcal{E}_D$ be a Dirichlet edge, $u \in A_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$, cp.* (7.19)*, with weight exponents $b_{\boldsymbol{c}}, b_{\boldsymbol{e}}$ as in Remark 2.5. Then in the setting of Section 7.1 and for ℓ sufficiently large, there exist constants $b, C > 0$ such that*

$$\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\perp}] + \Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\|}] + \Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},1}}[\eta^{\perp}] + \Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},1}}[\eta^{\|}] + \Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},2}}[\eta] \leq C \exp(-2b\ell). \tag{7.20}$$

Analogous exponential bounds holds for the consistency term $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{e}}}$. In addition, there holds

$$\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},D}}[\eta^{\perp}] + \Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},D}}[\eta^{\|}] \leq C \exp(-2b\ell). \tag{7.21}$$

Proof. There holds $|u|_{\widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})} \leq |u|_{\widehat{M}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})}$ for $k \geq 0$, and $u \in A_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ implies $u \in B_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$. Hence, the bound (7.20) follows from Theorem 7.1.

To bound (7.21), let K be in $\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}$. Then, by Lemma 5.6, item (7), the definition of the corner-edge semi-norm (7.18), and the equivalences (7.4), we find that

$$\begin{aligned}
(h_K^{\perp})^{-2} \|\eta^{\perp}\|^2_{L^2(K)} &\lesssim (h_K^{\perp})^{2b_{\boldsymbol{e}}} (h_K^{\|})^{2(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})} \|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}} r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}} \mathsf{D}^2_{\perp} u\|^2_{L^2(K)} \\
&\lesssim (h_K^{\perp})^{2b_{\boldsymbol{e}}} (h_K^{\|})^{2(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})} |u|^2_{\widehat{M}^2_{-1-\boldsymbol{b}}(K)}
\end{aligned}$$

In direction parallel to $\boldsymbol{e}$, we proceed similarly: The stability of the L^2-projection and inserting appropriate weights in accordance with the definition of the homogeneous semi-norm (7.18) and by employing property (7.4) readily yield

$$(h_K^{\perp})^{-2} \|\eta^{\|}\|^2_{L^2(K)} \lesssim (h_K^{\perp})^{-2} \|u\|^2_{L^2(K)} \lesssim (h_K^{\perp})^{2b_{\boldsymbol{e}}} (h_K^{\|})^{2(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})} |u|^2_{\widehat{M}^0_{-1-\boldsymbol{b}}(K)}.$$

Therefore, expressing the mesh sizes in terms of σ, cp. (7.4), implies

$$(h_K^{\perp})^{-2}\big(\|\eta^{\perp}\|^2_{L^2(K)} + \|\eta^{\|}\|^2_{L^2(K)}\big) \lesssim \sigma^{2(\ell+1-j)b_{\boldsymbol{c}}+2(j-1)b_{\boldsymbol{e}}}\|u\|^2_{\widehat{M}^2_{-1-\boldsymbol{b}}(K)} \lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}\|u\|^2_{\widehat{M}^2_{-1-\boldsymbol{b}}(K)} .$$

Summing the above bound over all elements in $\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}$ implies the asserted bound. $\square$

7.4. **Exponential convergence for corner elements.** It remains to show exponential convergence for the consistency terms $T^{K_{\boldsymbol{c}}}_{\boldsymbol{c}}[\eta]$ in (6.4) associated with the single corner elements $K_{\boldsymbol{c}} \in \widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{c}}$ of the reference corner-edge mesh $\widehat{\mathcal{M}}^{\ell}_{\boldsymbol{ce}}$, cp. (7.2), and the reference corner mesh $\widehat{\mathcal{M}}^{\ell}_{\boldsymbol{c}}$ in (7.7). We detail this for $K_{\boldsymbol{c}} \in \widehat{\mathcal{M}}^{\ell}_{\boldsymbol{ce}}$, the estimates in $K_{\boldsymbol{c}} \in \widehat{\mathcal{M}}^{\ell}_{\boldsymbol{c}}$ being analogous. The element $K_{\boldsymbol{c}}$ is shape-regular and of diameter $h_{\boldsymbol{c}} = \mathcal{O}(\sigma^{\ell})$. The quasi-interpolant $\mathcal{I}_1$ defined in Lemma 5.6 is used for $\mathfrak{K} = K_{\boldsymbol{c}}$, and $\eta|_{K_{\boldsymbol{c}}} = u|_{K_{\boldsymbol{c}}} - \mathcal{I}_1(u|_{K_{\boldsymbol{c}}})$. By Lemma 5.6, there holds $\|\eta\|_{L^2(K_{\boldsymbol{c}})} \lesssim h_{\boldsymbol{c}}\|\nabla\eta\|_{L^2(K_{\boldsymbol{c}})} = h_{\boldsymbol{c}}\|\nabla u - \mathbf{\Pi}_0(\nabla u)\|_{L^2(K_{\boldsymbol{c}})}$. We conclude that

$$T^{K_{\boldsymbol{c}}}_{\boldsymbol{c}}[\eta] \lesssim \|\nabla u - \mathbf{\Pi}_0(\nabla u)\|^2_{L^2(K_{\boldsymbol{c}})} + h_{\boldsymbol{c}}^{-1}|u|^2_{W^{2,1}(K_{\boldsymbol{c}})} . \tag{7.22}$$

The following bound is a standard h-version approximation result.

Lemma 7.11. *Assume that $u \in H^{1+\theta}(K_{\boldsymbol{c}})$ for some $\theta \in (0,1)$; cp. Remark 2.5. Then we have $\|\nabla(u-\mathcal{I}_1 u)\|_{L^2(K_{\boldsymbol{c}})} = \|\nabla u - \mathbf{\Pi}_0(\nabla u)\|_{L^2(K_{\boldsymbol{c}})} \lesssim h^{\theta}_{\boldsymbol{c}}\|u\|_{H^{1+\theta}(K_{\boldsymbol{c}})}$.*

It remains to bound the term $h_{\boldsymbol{c}}^{-1}|u|^2_{W^{2,1}(K_{\boldsymbol{c}})}$ in (7.22).

Lemma 7.12. *Let $u \in \widehat{N}^2_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ with weight exponents $b_{\boldsymbol{c}}, b_{\boldsymbol{e}} \in (0,1)$ as in Remark 2.5. Then we have $|u|_{W^{2,1}(K_{\boldsymbol{c}})} \lesssim h_{\boldsymbol{c}}^{1/2+b_{\boldsymbol{c}}}|u|_{\widehat{N}^2_{-1-\boldsymbol{b}}(K_{\boldsymbol{c}})}$.*

Proof. Let $|\boldsymbol{\alpha}| = 2$. Then there holds

$$\|\mathsf{D}^{\boldsymbol{\alpha}}u\|_{L^1(K_{\boldsymbol{c}})} \leq \|r_{\boldsymbol{c}}^{1+b_{\boldsymbol{c}}-|\boldsymbol{\alpha}|}\rho_{\boldsymbol{ce}}^{-\max(-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^{\perp}|,0)}\|_{L^2(K_{\boldsymbol{c}})} \times \|r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+|\boldsymbol{\alpha}|}\rho_{\boldsymbol{ce}}^{\max(-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^{\perp}|,0)}\mathsf{D}^{\boldsymbol{\alpha}}u\|_{L^2(K_{\boldsymbol{c}})} .$$

Introducing spherical polar coordinates on $K_{\boldsymbol{c}}$, we bound the last term by

$$\|r_{\boldsymbol{c}}^{1+b_{\boldsymbol{c}}-|\boldsymbol{\alpha}|}\rho_{\boldsymbol{ce}}^{-\max(-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^{\perp}|,0)}\|_{L^2(K_{\boldsymbol{c}})} \lesssim h_{\boldsymbol{c}}^{5/2+b_{\boldsymbol{c}}-|\boldsymbol{\alpha}|} \lesssim h_{\boldsymbol{c}}^{1/2+b_{\boldsymbol{c}}}, \qquad |\boldsymbol{\alpha}^{\perp}| = 0,1,$$

as well as

$$\|r_{\boldsymbol{c}}^{-1+b_{\boldsymbol{c}}}\rho_{\boldsymbol{ce}}^{-1+b_{\boldsymbol{e}}}\|_{L^2(K_{\boldsymbol{c}})} = \|r_{\boldsymbol{c}}^{b_{\boldsymbol{c}}-b_{\boldsymbol{e}}}r_{\boldsymbol{e}}^{-1+b_{\boldsymbol{e}}}\|_{L^2(K_{\boldsymbol{c}})} \lesssim h_{\boldsymbol{c}}^{1/2+b_{\boldsymbol{c}}}, \qquad |\boldsymbol{\alpha}^{\perp}| = 2.$$

Hence, we arrive at $|u|_{W^{2,1}(K_{\boldsymbol{c}})} \lesssim h_{\boldsymbol{c}}^{1/2+b_{\boldsymbol{c}}}|u|_{\widehat{N}^2_{-1-\boldsymbol{b}}(K_{\boldsymbol{c}})}$. $\square$

By inserting the estimates in the previous two lemmas into (7.22) we obtain the following error bound in corner elements.

Proposition 7.13. *Let $\boldsymbol{e}$ be a Neumann edge, $u \in \widehat{N}^2_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ and let the assumptions of Remark 2.5 be satisfied. For ℓ sufficiently large, there exist constants $b, C > 0$ such that $\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{c}}}[\eta] \leq C\exp(-2b\ell)$. The same result holds for a Dirichlet edge $\boldsymbol{e}$ and $u \in \widehat{M}^2_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$.*

7.5. **Conclusion of Theorem 6.2.** The exponential convergence of hp-dGFEM, Theorem 6.2, follows now immediately from the error bounds in Theorem 6.1, and from the fact that, by our analysis in this section, all terms on the right-hand side of the estimate in Theorem 6.1 converge exponentially in the number of mesh layers ℓ over the reference meshes $\widehat{\mathcal{M}}^{\ell}_{\boldsymbol{c}}$, $\widehat{\mathcal{M}}^{\ell}_{\boldsymbol{e}}$, and $\widehat{\mathcal{M}}^{\ell}_{\boldsymbol{ce}}$. The general result will then follow upon noting that a geometric mesh $\mathcal{M}^{(\ell)}_{\sigma}$ is obtained by a finite superposition of (scaled and translated versions of) these reference meshes. Furthermore, for the number of degrees of freedom in either of the hp-dG spaces in (3.8) and (3.9) there holds $N \simeq \ell^5 + \mathcal{O}(\ell^4)$, which yields the desired estimate (6.10).

References

1. D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal. **39** (2001), 1749–1779.
2. I. Babuška and B. Q. Guo, *Regularity of the solution of elliptic problems with piecewise analytic data. I. Boundary value problems for linear elliptic equation of second order*, SIAM J. Math. Anal. **19** (1988), no. 1, 172–203. MR MR924554 (89a:35039)
3. M. Costabel, M. Dauge, and S. Nicaise, *Analytic regularity for linear elliptic systems in polygons and polyhedra*, Math. Models Methods Appl. Sci. **22** (2012), no. 8, 1250015.
4. M. Costabel, M. Dauge, and Ch. Schwab, *Exponential convergence of hp-FEM for Maxwell's equations with weighted regularization in polygonal domains*, Math. Models Methods Appl. Sci. **15** (2005), no. 4, 575–622.
5. V. Dolejší, M. Feistauer, and Ch. Schwab, *A finite volume discontinuous Galerkin scheme for nonlinear convection-diffusion problems*, Calcolo **39** (2002), no. 1, 1–40. MR 1901200 (2003d:65091)
6. C. Ebmeyer, *Mixed boundary value problems for nonlinear elliptic systems with p–structure in polyhedral domains*, Math. Nachr. **236** (2002), 91–108.
7. B. Q. Guo, *The h-p version of the finite element method for solving boundary value problems in polyhedral domains*, Boundary Value Problems and Integral Equations in Nonsmooth Domains, Lecture Notes in Pure and Appl. Math., vol. 167, Dekker, New York, 1995, pp. 101–120.
8. B. Q. Guo and I. Babuška, *Regularity of the solutions for elliptic problems on nonsmooth domains in $\mathbb{R}^3$. I. Countably normed spaces on polyhedral domains*, Proc. Roy. Soc. Edinburgh Sect. A **127** (1997), no. 1, 77–126.
9. V. Maz'ya and J. Rossmann, *Elliptic equations in polyhedral domains*, Mathematical Surveys and Monographs, vol. 162, American Mathematical Society, Providence, RI, 2010. MR 2641539 (2011h:35002)
10. D. Schötzau and Ch. Schwab, *Exponential convergence in a Galerkin least squares hp-FEM for Stokes flow*, IMA J. Numer. Anal. **21** (2001), 53–80.
11. D. Schötzau, Ch. Schwab, and T. P. Wihler, *hp-dGFEM for second-order elliptic problems in polyhedra I: Stability on geometric meshes*, SIAM J. Numer. Anal. **51** (2013), no. 3, 1610–1633. MR 3061472
12. ———, *hp-dGFEM for second-order elliptic problems in polyhedra II: Exponential convergence*, SIAM J. Numer. Anal. **51** (2013), no. 4, 2005–2035. MR 3073653
13. Ch. Schwab, *p- and hp-finite element methods*, Numerical Mathematics and Scientific Computation, Oxford University Press, 1998. MR 1695813 (2000d:65003)
14. T. P. Wihler, *Discontinuous galerkin fem for elliptic problems in polygonal domains*, Ph.D. thesis, Swiss Federal Institute of Technology Zurich, 2002, Diss. ETH No. 14973.
15. T. P. Wihler and M. Wirz, *Mixed hp-discontinuous Galerkin FEM for linear elasticity and Stokes flow in three dimensions*, Math. Models Methods Appl. Sci. **22** (2012), no. 8, 1250016.

Mathematics Department, University of British Columbia, Vancouver, BC V6T 1Z2, Canada
E-mail address: `schoetzau@math.ubc.ca`

Seminar of Applied Mathematics, ETH Zürich, 8092 Zürich, Switzerland
E-mail address: `schwab@math.ethz.ch`

Mathematisches Institut, Universität Bern, 3012 Bern, Switzerland
E-mail address: `wihler@math.unibe.ch`