hp -Version discontinuous Galerkin methods on essentially arbitrarily-shaped elements

Andrea Cangiani, Zhaonan Dong, Emmanuil H Georgoulis

▶ **To cite this version:**

Andrea Cangiani, Zhaonan Dong, Emmanuil H Georgoulis. hp -Version discontinuous Galerkin methods on essentially arbitrarily-shaped elements. 2021. hal-03109470v2

HAL Id: hal-03109470

https://inria.hal.science/hal-03109470v2

Preprint submitted on 19 Apr 2021 (v2), last revised 9 May 2021 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

hp-VERSION DISCONTINUOUS GALERKIN METHODS ON ESSENTIALLY ARBITRARILY-SHAPED ELEMENTS

ANDREA CANGIANI, ZHAONAN DONG, AND EMMANUIL H. GEORGOULIS

Abstract. We extend the applicability of the popular interior-penalty discontinuous Galerkin (dG) method discretizing advection-diffusion-reaction problems to meshes comprising extremely general, essentially arbitrarily-shaped element shapes. In particular, our analysis allows for *curved* element shapes, without the use of non-linear elemental maps. The feasibility of the method relies on the definition of a suitable choice of the discontinuity penalization, which turns out to be explicitly dependent on the particular element shape, but essentially independent on small shape variations. This is achieved upon proving extensions of classical trace and Markov-type inverse estimates to arbitrary element shapes. A further new $H^1 - L_2$-type inverse estimate on essentially arbitrary element shapes enables the proof of inf-sup stability of the method in a streamline-diffusion-like norm. These inverse estimates may be of independent interest. A priori error bounds for the resulting method are given under very mild structural assumptions restricting the magnitude of the local curvature of element boundaries. Numerical experiments are also presented, indicating the practicality of the proposed approach.

1. Introduction

Recent years have witnessed a coordinated effort to generalize mesh concepts in the context of Galerkin/finite element methods. A key argument has been that more general-shaped elements/cells can potentially lead to computational complexity reduction. This effort has given rise to a number of recent approaches: mimetic finite difference methods [11], virtual element methods [12, 13], various discontinuous Galerkin approaches such as interior penalty [20], hybridized DG [26] and the related hybrid high-order methods [27]. Earlier approaches involving non-polynomial approximation spaces, such as polygonal and other generalized finite element methods [58, 32], have also been developed and used by the engineering community. All the above numerical frameworks allow for polygonal/polyhedral element shapes (henceforth, collectively termed as *polytopic*) of varying levels of generality.

Simultaneously, various classes of fitted and unfitted grid methods for interface or transmission problems exploit generalized concepts of mesh elements in an effort to provide accurate representations of internal interfaces. Several unfitted finite element methods have been proposed in recent years: unfitted finite element methods [9], immersed finite element methods [37, 36], virtual element methods [25], unfitted penalty methods [10, 48, 43, 60, 23], see also [44] for unfitted discretization of the boundary, cutCell/cutFEM [16, 14], and unfitted hybrid high-order methods [15], to name just the few closer to the developments we shall be concerned with below. A central idea in the majority of these methods is the weak imposition of interface conditions in conjunction with some form of penalization, see, e.g., [38], an idea going back to [8]. These approaches are often combined with level set concepts [53] to describe the interfaces accurately. Nonetheless, in their practical implementation, the interface is typically represented via piecewise smooth polynomial approximations to the level sets.

The interior penalty discontinuous Galerkin (IP-dG) approach appears to allow for extreme generality with regard to element shapes/geometries. Indeed, in contrast to aforementioned families of general mesh methods, IP-dG can handle arbitrary number of faces per element with solid theoretical backing involving provable stability and convergence results; see [19, 20] for details. This property becomes extremely relevant upon realising that IP-dG (as well as other classical dG methods, such as LDG) associate local numerical degrees of freedom to the elements *only*, and not to other geometrical entities such as faces or vertices. As such, the nature and dimension of the local discretization space is independent of the number of vertices/faces. The latter observation implies also naturally a form of complexity reduction: classical total degree ('$P-$type') local polynomial spaces in physical coordinates are admissible on box-type or highly complex element shapes [21, 18, 20]. We refer to our monograph [20] for details on the admissible polygonal/polyhedral element shapes for which the IP-dG method is, provably, both stable and convergent. The mild element shape assumptions in [20] are such to ensure the validity of crucial generalizations of standard approximation results, such as inverse estimates, best approximation estimates, and extension theorems. Thus, the developments presented for IP-dG in [20] can be potentially ported also to other classical dG approaches within the unified framework of [4]; we also note the recent static condensation approach presented in [47] in this context.

The question, therefore, of further extending rigorously the applicability of *hp*-version IP-dG methods to meshes consisting of *curved* polygonal/polyhedral elements arises naturally. Indeed, such a development is expected to provide multifaceted advantages compared to current approaches, including, but not restricted to, the treatment of curved interfaces as done, e.g., in [48, 43, 60, 15, 23]. For instance, allowing for extremely general curved elements enables the *exact* representation of curved computational domains, e.g., arising directly from Computer Aided Design programs. Allowing also for arbitrary local polynomial degrees, provides the possibility of achieving required accuracy via local (polynomial) basis enrichment (p-version Galerkin approaches) without increased mesh-granularity. If, nonetheless, local mesh refinement is also required/desired, IP-dG methods can be immediately applied on refined curved elements without local re-parametrizations of the local Galerkin spaces. This is in contrast to the need to perform costly re-parametrisations upon mesh refinement in other approaches, e.g., Isogeometric Analysis [39] or, indeed, even to keep track of the domain-approximation variational crimes of standard finite element discretizations. Exact geometry representation can also be highly relevant in representing locally discontinuous/sharply changing PDE coefficients, e.g., in permeability pressure computations in porous media, coefficients defined via level-sets of smooth functions, or shape/topology optimization applications.

Furthermore, exact geometry representation is relevant in the p-version Galerkin context: to achieve spectral/exponential convergence for smooth PDE problems posed on general curved domains, we are required to use isoparametrically mapped elements. This is both cumbersome to implement and costly as the polynomial degree increases [50, 51]. A successful alternative to isoparametric maps is the use of non-linear maps on element patches [56, 49] to represent domain geometry. Nevertheless, if the elemental maps are not *a priori* provided, it is difficult to construct them in practice, especially in three dimensions.

Finally, curved element capabilities should ideally be developed *in conjunction* with the already developed highly general polytopic mesh IP-dG methods, allowing for instance elements with *arbitrary* number of faces. This is particularly pertinent in the contexts of adaptivity and multilevel solvers, which benefit from element

agglomeration [2, 3] to achieve coarser representations. With regard to adaptivity, mesh coarsening is essential in keeping the computation sizes at bay, at least in the case of evolution problems. The extreme coarsening capabilities via element agglomeration, therefore, have the potential in retaining structure, e.g., possible coefficient heterogeneities at the discrete level for instance.

It is, therefore, desirable to design and analyze IP-dG and related methods posed on meshes comprising of elements with *arbitrary* number of *curved* faces, under as mild geometric assumptions as possible. To address this central, in our view, question, this work aims at rigorously extending the applicability of IP-dG methods on meshes comprising of essentially arbitrarily *curve-shaped* polytopic elements with arbitrary number of faces per element; this includes, in particular, curved elements *not* exactly representable by (iso-)parametric polynomial element mappings.

The theoretical developments presented below regarding stability and a-priori error analysis of IP-dG methods hinge on new, to the best of our knowledge, extensions of known inverse and trace inequalities. More specifically, we extend the *hp*-version trace inverse estimate presented in [23], allowing for more general curved element shapes; see also [43] for an earlier, related result. Trace inverse estimates are crucial in the proof of stability of IP-dG methods and, simultaneously, determine the so-called discontinuity-penalization parameter for a given mesh. This is crucial on meshes of such generality: insufficient penalization results in loss of stability, while excessive penalization typically results in accuracy loss. Also, we prove new *hp*-version $L_\infty - L_2$ and $H^1 - L_2$ inverse inequalities on extremely general curved domains. Particular care has been given so that these new inverse estimates are 'shape-robust', in the sense that there is no hidden dependence of the element shape in the constants. We believe that these extensions of known inverse estimates to be of independent interest, due to their frequent use in the analysis of finite element methods.

The new inverse estimates are combined with ideas from the analysis of polytopic dG methods [20], resulting in significant generalization of the results presented therein. More specifically, by relaxing certain earlier coverability assumptions, (postulating the ability to cover tightly general-shaped elements by unions of simplices of similar size, cf. [20, Definition 10]) as well as by proving a new stability result for norms of polynomials under domain perturbations (Lemma 4.14), we prove stability and a new *hp*-version a priori error analysis for the IP-dG method on essentially arbitrary element shapes. The a priori error analysis follows closely the proof from [18]: upon establishing an inf-sup stability result of the method in a streamline-diffusion-like norm, standard Strang-type arguments with *hp*-best approximation results lead to an error bound. The inf-sup result justifies also the good stability properties of the method in convection-dominated problems. The theoretical tools presented may also be of interest in Nitsche-type formulations of unfitted grid interface methods. To emphasize the mesh-generality of the proposed approach, we shall refer to the framework presented below as *discontinuous Galerkin method on essentially arbitrarily-shaped elements (dG-EASE)*.

The remainder of this work is organised as follows. Upon describing the advection-diffusion-reaction model problem in Section 2, we introduce the *hp*-version interior penalty discontinuous Galerkin method in Section 3. We prove new inverse estimates in Section 4, along with the necessary *hp*-approximation results. In Section 5, we present the stability and a-priori error analysis. Finally, the performance of the dG methods is assessed in practice through a series of numerical experiments presented in Section 6.

2. Model problem

To highlight the versatility of dG-EASE, we consider the class of second–order partial differential equations with nonnegative characteristic form over an open bounded Lipschitz domain Ω in $\mathbb{R}^d$, $d \geq 1$, with boundary $\partial\Omega$. This class includes general advection-diffusion-reaction problems possibly of changing type, see, e.g., [20]. The model problem reads: find $u \in \mathcal{V}$ such that

$$-\nabla \cdot (a\nabla u) + \nabla \cdot (\mathbf{b}u) + cu = f \quad \text{in } \Omega, \tag{2.1}$$

for some suitable solution space $\mathcal{V}$, and $a = \{a_{ij}\}_{i,j=1}^d$, symmetric with $a_{ij} \in L_\infty(\Omega)$, so that at each $\mathbf{x}$ in $\bar{\Omega}$, we have

$$\sum_{i,j=1}^{d} a_{ij}(\mathbf{x})\xi_i\xi_j \geq 0, \qquad \text{for any} \quad \xi = (\xi_1, \ldots, \xi_d)^T \in \mathbb{R}^d; \tag{2.2}$$

also $\mathbf{b} = (b_1, \ldots, b_d)^T \in \left[W^1_\infty(\Omega)\right]^d$, $c \in L_\infty(\Omega)$ and $f \in L_2(\Omega)$.

To supplement (2.1) with suitable boundary conditions, following [52], we first subdivide the boundary $\partial\Omega$ into $\partial_0\Omega = \left\{\mathbf{x} \in \partial\Omega : \sum_{i,j=1}^d a_{ij}(\mathbf{x})n_i n_j > 0\right\}$, and $\partial\Omega\backslash\partial_0\Omega$ with $\mathbf{n} = (n_1, \ldots, n_d)^T$ denoting the unit outward normal vector to $\partial\Omega$. Loosely speaking, we may think of $\partial_0\Omega$ as being the 'elliptic' portion of the boundary $\partial\Omega$. We further split the 'hyperbolic' portion of the boundary $\partial\Omega\backslash\partial_0\Omega$, into inflow and outflow boundaries $\partial_-\Omega$ and $\partial_+\Omega$, respectively, by

$$\partial_-\Omega = \{\mathbf{x} \in \partial\Omega\backslash\partial_0\Omega : \mathbf{b}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) < 0\}, \ \partial_+\Omega = \{\mathbf{x} \in \partial\Omega\backslash\partial_0\Omega : \mathbf{b}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \geq 0\}.$$

If $\partial_0\Omega$ is nonempty, we shall further divide it into two disjoint subsets $\partial\Omega_{\mathrm{D}}$ and $\partial\Omega_{\mathrm{N}}$, with $\partial\Omega_{\mathrm{D}}$ nonempty and relatively open in $\partial\Omega$. It is evident from these definitions that $\partial\Omega = \partial\Omega_{\mathrm{D}} \cup \partial\Omega_{\mathrm{N}} \cup \partial_-\Omega \cup \partial_+\Omega$.

It is physically reasonable to assume that $\mathbf{b} \cdot \mathbf{n} \geq 0$ on $\partial\Omega_{\mathrm{N}}$, whenever $\partial\Omega_{\mathrm{N}}$ is nonempty; then, we impose the boundary conditions:

$$u = g_{\mathrm{D}} \ \text{ on } \partial\Omega_{\mathrm{D}} \cup \partial_-\Omega, \qquad \mathbf{n} \cdot (a\nabla u) = g_{\mathrm{N}} \ \text{ on } \partial\Omega_{\mathrm{N}}; \tag{2.3}$$

For an extension, allowing also for $\mathbf{b} \cdot \mathbf{n} < 0$ on $\partial\Omega_{\mathrm{N}}$, we refer to [22]. Additionally, we assume that there exists a positive constant γ_0 such that

$$c_0(\mathbf{x}) := \left(c(\mathbf{x}) + \frac{1}{2}\nabla \cdot \mathbf{b}(\mathbf{x})\right)^{1/2} \geq \gamma_0 \quad \text{a.e. } x \in \Omega. \tag{2.4}$$

For a proof of the well–posedness of (2.1), (2.3), subject to (2.4), we refer to [52, 42].

3. Discontinuous Galerkin method

We shall now define the interior penalty discontinuous Galerkin (dG) method posed on essentially arbitrarily-shaped elements. A key attribute of the method is the use of physical frame basis functions, i.e., the elemental bases consist of polynomials on the elements themselves, rather than mapped from a reference element. The implementation challenges arising from this non-standard choice with regard to construction of the resulting linear system will be discussed below.

3.1. **The mesh.** Let $\mathcal{T} = \{K\}$ be a subdivision of Ω into non-overlapping subsets (elements) $K \in \mathcal{T}$ with, possibly curved, Lipschitz boundaries and let $h_K := \mathrm{diam}(K)$. The mesh skeleton $\Gamma := \cup_{K\in\mathcal{T}}\partial K$ is subdivided into the internal part $\Gamma_{\mathrm{int}} := \Gamma\backslash\partial\Omega$ and boundary part $\partial\Omega$. We further explicitly assume that the $(d-1)$-dimensional Hausdorff measure of Γ is globally finite, thereby, not allowing for fractal-shaped elements.

We note immediately that we allow mesh elements $K \in \mathcal{T}$ which are essentially arbitrarily-shaped and with very general interfaces with neighbouring elements.

For instance, two elements may interface at a collection of $(d-1)$-dimensional (possibly curved) *faces*, as those shown in Figure 3.1. The precise assumptions on the admissible element shapes are given in Section 4 below.

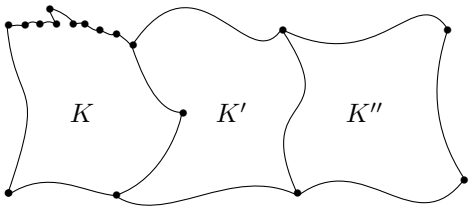

FIGURE 1. Curved elements K, K', K'' for $d = 2$ with possibly many curved faces; $\bullet$ denotes a vertex.

3.2. **Discontinuous Galerkin method.** We define the *hp*-version discontinuous finite element space $S^{\mathbf{p}}_{\mathcal{T}}$, subordinate to the mesh $\mathcal{T} = \{K\}$ and a polynomial degree vector $\mathbf{p} := \{p_K\}$, possibly different for each element K, by

$$S^{\mathbf{p}}_{\mathcal{T}} = \{v \in L^2(\Omega) : v|_K \in \mathcal{P}_{p_K}(K),\ K \in \mathcal{T}\}. \tag{3.1}$$

For any elemental face $F \subset \Gamma_{\text{int}}$, let K and K' be the two elements such that $F \subset \partial K \cap \partial K'$. The outward unit normal vectors on F of ∂K and $\partial K'$ are denoted by $\mathbf{n}_K$ and $\mathbf{n}_{K'}$, respectively. For a function $v : \Omega \to \mathbb{R}$ that may be discontinuous across Γ, we define the jump $[\![v]\!]$ and the average $\{v\}$ of v across F by

$$[\![v]\!] = v|_K \mathbf{n}_K + v|_{K'} \mathbf{n}_{K'}, \quad \{v\} = \frac{1}{2}\left(v|_K + v|_{K'}\right). \tag{3.2}$$

Similarly, for a vector valued function $\mathbf{w}$, piecewise smooth on $\mathcal{T}$, we define

$$[\![\mathbf{w}]\!] = \mathbf{w}|_K \cdot \mathbf{n}_K + \mathbf{w}|_{K'} \cdot \mathbf{n}_{K'}, \quad \{\mathbf{w}\} = \frac{1}{2}\left(\mathbf{w}|_K + \mathbf{w}|_{K'}\right).$$

When $F \subset \partial\Omega$, we set $\{v\} = v$, $[\![v]\!] = v\mathbf{n}$ and $[\![\mathbf{w}]\!] = \mathbf{w} \cdot \mathbf{n}$ with $\mathbf{n}$ denoting the outward unit normal to the boundary $\partial\Omega$.

For any element $K \in \mathcal{T}$, we define inflow and outflow parts of ∂K by

$$\partial_- K = \{\mathbf{x} \in \partial K :\ \mathbf{b}(\mathbf{x}) \cdot \mathbf{n}_K(\mathbf{x}) < 0\}, \quad \partial_+ K = \{\mathbf{x} \in \partial K : \mathbf{b}(\mathbf{x}) \cdot \mathbf{n}_K(\mathbf{x}) \geq 0\},$$

respectively, with $\mathbf{n}_K(\mathbf{x})$ denoting the unit outward normal vector to ∂K at $\mathbf{x} \in \partial K$. Further, we define the *upwind jump* of the (scalar-valued) function v across the inflow boundary $\partial_- K$ of $K \in \mathcal{T}$ by

$$\lfloor v \rfloor(\mathbf{x}) := \lim_{\epsilon \to 0^+} \Big(v(\mathbf{x} + \mathbf{b}(\mathbf{x})\epsilon) - v(\mathbf{x} - \mathbf{b}(\mathbf{x})\epsilon)\Big), \quad \text{when } \mathbf{x} \in \partial_- K \backslash \partial\Omega.$$

Finally, we define the broken gradient $\nabla_{\mathcal{T}} v$ of a function $v \in L_2(\Omega)$ with $v|_K \in H^1(K)$, for all $K \in \mathcal{T}$, element-wise by $\big(\nabla_{\mathcal{T}} v\big)|_K := \nabla(v|_K)$.

The *discontinuous Galerkin method on essentially arbitrarily-shaped elements* (dG-EASE for short) reads: find $u_h \in S^{\mathbf{p}}_{\mathcal{T}}$ such that

$$B(u_h, v_h) = \ell(v_h) \qquad \text{for all} \quad v_h \in S^{\mathbf{p}}_{\mathcal{T}}, \tag{3.3}$$

with the bilinear form $B(\cdot,\cdot) : S^{\mathbf{p}}_{\mathcal{T}} \times S^{\mathbf{p}}_{\mathcal{T}} \to \mathbb{R}$ defined as

$$B(u, v) := B_{\text{ar}}(u, v) + B_{\text{d}}(u, v),$$

where $B_{\text{ar}}(\cdot,\cdot)$ accounts for the advection and reaction terms:

$$\begin{aligned} B_{\text{ar}}(u, v) := &\int_\Omega \big(\nabla_{\mathcal{T}}(\mathbf{b} \cdot u) + cu\big)v \,\mathrm{d}\mathbf{x} - \sum_{K \in \mathcal{T}} \int_{\partial_- K \backslash \partial\Omega} (\mathbf{b} \cdot \mathbf{n}) \lfloor u \rfloor v \,\mathrm{d}s \\ &- \sum_{K \in \mathcal{T}} \int_{\partial_- K \cap (\partial\Omega_{\text{D}} \cup \partial_- \Omega)} (\mathbf{b} \cdot \mathbf{n}) uv \,\mathrm{d}s, \end{aligned} \tag{3.4}$$

and $B_{\mathrm{d}}(\cdot,\cdot)$ corresponds to the diffusion part:

$$
\begin{aligned}
B_{\mathrm{d}}(u,v) := &\int_\Omega a\nabla_{\mathcal{T}} u\cdot\nabla_{\mathcal{T}} v\,\mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{int}}\cup\partial\Omega_{\mathrm{D}}} \sigma[\![u]\!]\cdot[\![v]\!]\,\mathrm{d}s\\
&- \int_{\Gamma_{\mathrm{int}}\cup\partial\Omega_{\mathrm{D}}} \Big(\{\!\{a\nabla u\}\!\}\cdot[\![v]\!] + \{\!\{a\nabla v\}\!\}\cdot[\![u]\!]\Big)\,\mathrm{d}s,
\end{aligned}
\tag{3.5}
$$

while the linear functional $\ell : S^{\mathbf{p}}_{\mathcal{T}} \to \mathbb{R}$ is defined by

$$
\begin{aligned}
\ell(v) := &\int_\Omega fv\,\mathrm{d}\mathbf{x} - \sum_{K\in\mathcal{T}}\int_{\partial_- K\cap(\partial\Omega_{\mathrm{D}}\cup\partial_-\Omega)} (\mathbf{b}\cdot\mathbf{n})g_{\mathrm{D}} v\,\mathrm{d}s\\
&- \int_{\partial\Omega_{\mathrm{D}}} g_{\mathrm{D}}\big((a\nabla v)\cdot\mathbf{n} - \sigma v\big)\,\mathrm{d}s + \int_{\partial\Omega_{\mathrm{N}}} g_{\mathrm{N}} v\,\mathrm{d}s.
\end{aligned}
\tag{3.6}
$$

The nonnegative function $\sigma \in L_\infty(\Gamma_{\mathrm{int}}\cup\partial\Omega_{\mathrm{D}})$ appearing in (3.5) and (3.6) is the *discontinuity-penalization function*, whose precise definition, which depends on the diffusion tensor a and the discretization parameters, will be given below. We note that a 'good' choice of discontinuity penalization is instrumental for the stability of the method, while simultaneously not affecting the approximation properties in the general mesh setting considered herein.

For simplicity of presentation, we shall assume that the entries of the diffusion tensor a are element-wise constants on each element $K\in\mathcal{T}$, i.e.,

$$
a \in [S^{\mathbf{0}}_{\mathcal{T}}]^{d\times d}_{\mathrm{sym}},
\tag{3.7}
$$

Our results can be applied to the case of general $a\in[H^{1/2}(\Omega)]^{d\times d}_{\mathrm{sym}}$ by slightly modifying the bilinear form above as proposed originally in [35] and extended to polytopic meshes in [20]. In the following, $\sqrt{a}$ denotes the (positive semidefinite) square-root of the symmetric matrix a; further, $\bar{a}_K := |\sqrt{a}|^2_2|_K$, where $|\cdot|_2$ denotes the matrix-2–norm. Also, in the interest of accessibility, we shall not consider problems with high contrast diffusion tensors, with the usual weighted averaging modification of the method [17, 31, 28]; the extension to that setting is completely analogous to the analysis presented below.

Remark 3.1. The parameter σ is typically selected to be face-wise constant in the definition and implementation of IP-dG methods. To ensure that only physically correct penalization takes place, σ is chosen below to be proportional to the quantity $\mathbf{n}^T a\mathbf{n}$; see [35] for details. As such, σ will vary along a curved element face even for element-wise constant diffusion a, thereby justifying the terminology "penalization function" as opposed to the standard terminology "penalization function" from the literature. Further, the theory presented below can also be extended with minor modifications to curved faces F, such that $\mathbf{n}^T a\mathbf{n} > 0$ only on a strict subset of that face and $\mathbf{n}^T a\mathbf{n} = 0$ on the remaining part. That way one can reduce or even remove unphysical penalization on the hypersurfaces where the PDE may change type.

4. Inverse and approximation estimates

A key challenge in the error analysis presented below is the availability of inverse estimation and approximation results with uniform /explicit constants with respect to the shape of the elements in a given mesh.

A trace type inverse estimate for elements with one curved face has been recently proven in [23] under a shape-regularity assumption; see (4.1) below. Results in this direction have also appeared under various geometric assumptions in [60, 48, 15], among others. Here, we extend these results by proving trace-inverse estimates for elements that are locally star-shaped, Lipschitz domains (see Assumption 4.1 below). Moreover, given the importance of trace-inverse estimates for the stability

of interior penalty dG methods, the new estimate constant is expressed via explicit and practically verifiable, geometric quantities (Lemma 4.4 below).

In the same vein, we also extend the classical (Markov-type) $H^1 - L_2$ inverse estimate to elements with piecewise C^1, locally star-shaped boundaries (see Assumptions 4.1 and 4.3 below). The proof builds upon and extends on earlier ideas from [45] and [20]. Here, we are particularly concerned with explicit quantification of the respective constant for a given element geometry. We note that $H^1 - L_2$ inverse estimates are also relevant in the determination of penalty parameters in IP-dG methods for biharmonic operators [29].

Also, we revisit a key stability argument that enabled the use of 'degenerate' polytopic element shapes, i.e., ones containing very small/degenerating faces/edges compared to the element diameter, first proposed in [21]; see also [18, 20] for improvements. This result is crucial in offering a practical choice of the discontinuity-penalization parameter for general polytopic meshes. The stability argument was based on two ingredients: 1) control of integral norms of polynomials with respect to domain perturbations using [34, Lemma 3.7], and 2) an $L_\infty - L_2$ inverse estimate. To retain this capability in the current setting, we prove an extension of [34, Lemma 3.7] (see also [45, Lemma 6] for a related result) for generalized/curved prismatic elements; see Lemma 4.14 below. Moreover, we also prove an extension of the classical $L_\infty - L_2$ inverse estimate for generalized/curved prismatic elements. The latter two new estimates, in conjunction with a revised concept of *coverability* (compared to [21, 20]) are enough to provide extensions to previously known stability results for IP-dG within the present level of mesh generality.

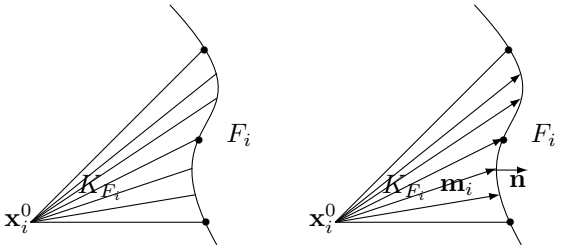

FIGURE 2. Elements $K \in \mathcal{T}^{tr}$ are assumed to satisfy Assumption 4.1 (a) (left) and (b) (right); $\bullet$ denotes a vertex.

Assumption 4.1. For each element $K \in \mathcal{T}$, we assume that K is a Lipschitz domain, and that we can subdivide ∂K into mutually exclusive subsets $\{F_i\}_{i=1}^{n_K}$ satisfying the following property: there exist respective sub-elements $K_{F_i} \equiv K_{F_i}(\mathbf{x}_i^0) \subset K$ with d planar faces meeting at one vertex $\mathbf{x}_i^0 \in K$, with $F_i \subset \partial K_{F_i}$, such that, for $i = 1, \ldots, n_K$,

(a) K_{F_i} is star-shaped with respect to $\mathbf{x}_i^0$. We refer to Figure 2(left) for an illustration for $d = 2$;

(b) $\mathbf{m}_i(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) > 0$ for $\mathbf{m}_i(\mathbf{x}) := \mathbf{x} - \mathbf{x}_i^0$, $\mathbf{x} \in K_{F_i}$, and $\mathbf{n}(\mathbf{x})$ the respective unit outward normal vector to F_i at $\mathbf{x} \in F_i$. (We refer to Figure 2(right) for an illustration for $d = 2$.)

Remark 4.2. Some remarks on the above (very mild) mesh assumption are in order:

(i) The sub-domains $\{F_i\}_{i=1}^{n_K}$ are *not* required to coincide with the faces of the element K: each F_i may be part of a face or may include one or more faces of K. Also, there is *no* requirement for $\{n_K\}_{K \in \mathcal{T}}$ to remain uniformly bounded across the mesh.

(ii) We can make Assumption 4.1(b) stronger by further postulating that: it is possible to fix the point $\mathbf{x}_i^0$ such that there exists a global constant $c_{sh} > 0$,

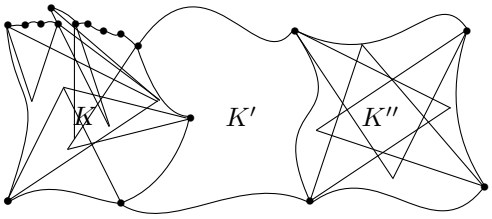

FIGURE 3. Curved elements K, K'' with, respectively, 8 and 4 sub-elements satisfying Assumption 4.1.

such that

$$\mathbf{m}_i(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \geq c_{sh} h_{K_{F_i}}; \tag{4.1}$$

this is the case, of course, for straight-faced polytopic elements, cf., [23, 60]. Note that (4.1) does *not* imply shape-regularity of the K_{F_i}'s; in particular K_{F_i}'s with 'small' F_i compared to the remaining (straight) faces of K_{F_i} are acceptable. Such anisotropic sub-elements K_{F_i}'s may be necessary to ensure that each K_{F_i} remains star-shaped when an element boundary's curvature is locally large; see, e.g., K_{F_i} in Figure 2 and a collection of both 'shape-regular' and 'anisotropic' K_{F_i}'s in Figure 3.

(iii) On certain geometrically extreme cases, satisfying Assumption 4.1 may require a small number of refinements of the elements $K \in \mathcal{T}$ of a given initial mesh.

(iv) F_i is not required to be connected. However, by splitting F_i to its connected subsets, re-indexing the F_i's to correspond to unique K_{F_i}, we can allocate one K_{F_i} to each F_i; we shall take the latter point of view in what follows to avoid further notational complexity. □

Assumption 4.3. We assume that the boundary ∂K of each element $K \in \mathcal{T}$ is the union of a finite (yet, arbitrarily large!) number of closed C^1 surfaces.

Assumption 4.1 is sufficient for the proof of the trace estimates presented below. Requiring both Assumptions 4.1 and 4.3 is sufficient for the validity of the $H^1 - L_2$ inverse estimate presented below.

4.1. **Basic trace estimates.** We now discuss the new trace-inverse estimate and a version of the standard Sobolev trace estimate for Lipschitz domains satisfying Assumption 4.1.

Lemma 4.4. *Let element $K \in \mathcal{T}$ be a Lipschitz domain satisfying Assumption 4.1. Then, for each $F_i \subset \partial K$, $i = 1, \ldots, n_K$, and for each $v \in \mathcal{P}_p(K)$, we have the inverse estimate:*

$$\|v\|^2_{F_i} \leq \frac{(p+1)(p+d)}{\min_{\mathbf{x} \in F_i}(\mathbf{m}_i \cdot \mathbf{n})} \|v\|^2_{K_{F_i}}. \tag{4.2}$$

Proof. We partition F_i into r $(d-1)$-dimensional curved simplices denoted by F_i^j, $j = 1, \ldots, r$, which are subordinate to the vertices possibly contained in F_i; r is large enough to accommodate this requirement. Further, we construct a partition of K_{F_i} into (curved) sub-elements K_i^j, by considering the simplices with one (curved) face F_i^j and the remaining vertex being $\mathbf{x}_i^0$; this is possible due to the star-shapedness of K_{F_i} with respect to $\mathbf{x}_i^0$ as per Assumption 4.1(a). We refer to Figure 4 for an illustration when $d = 2$. Notice that each F_i^j may include at most one constituent curved face of F_i, or part thereof.

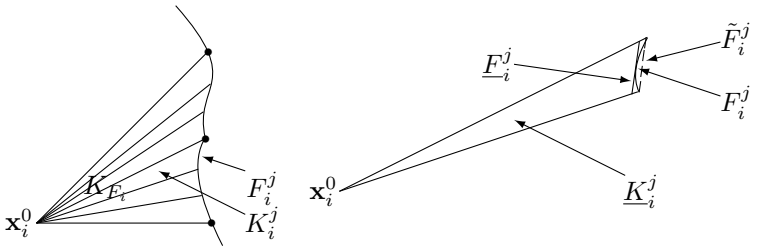

FIGURE 4. Partitioned curved sub-element $K_{F_i} \subset K \in \mathcal{T}$; $\bullet$ denotes a vertex of K (left); detail with $\underline{K}_i^j$ and related faces $\tilde{F}_i^j$ and $\underline{F}_i^j$ (right).

Let now $\tilde{F}_i^j$ denote the straight/planar related face defined by the $d-1$ vertices of F_i^j. Let also $\underline{K}_i^j$ be the largest straight-faced simplex contained in K_i^j with face $\underline{F}_i^j$ parallel to $\tilde{F}_i^j$ and the remaining faces being subsets of the straight faces of K_i^j. The Divergence Theorem implies

$$\begin{aligned}\int_{K_i^j\backslash\underline{K}_i^j} \nabla\cdot(v^2\mathbf{m}_i)\,\mathrm{d}\mathbf{x} &= \int_{\partial(K_i^j\backslash\underline{K}_i^j)} v^2\mathbf{m}_i\cdot\mathbf{n}_{\partial(K_i^j\backslash\underline{K}_i^j)}\,\mathrm{d}s\\ &= \int_{F_i^j} v^2\mathbf{m}_i\cdot\mathbf{n}\,\mathrm{d}s + \int_{\underline{F}_i^j} v^2\mathbf{m}_i\cdot\mathbf{n}_{\underline{F}_i^j}\,\mathrm{d}s,\end{aligned}$$

with $\mathbf{n}_\omega$ denoting the outward normal vector of a domain $\omega\subset\mathbb{R}^d$ and $\mathbf{m}_i$ as in Assumption 4.1(b), upon observing that $\mathbf{m}_i\cdot\mathbf{n}_{\partial(K_i^j\backslash\underline{K}_i^j)}=0$ on $\partial(K_i^j\backslash\underline{K}_i^j)\backslash(F_i^j\cup\underline{F}_i^j)$. Now, denoting by $|\cdot|_2$ the Euclidean distance in $\mathbb{R}^d$, the product rule and elementary estimates imply

$$\int_{K_i^j\backslash\underline{K}_i^j}\nabla\cdot(v^2\mathbf{m}_i)\,\mathrm{d}\mathbf{x} \le \Big(2\max_{K_i^j}|\mathbf{m}_i|_2\|v\nabla v\|_{L_\infty(K_i^j\backslash\underline{K}_i^j)} + d\|v\|^2_{L_\infty(K_i^j\backslash\underline{K}_i^j)}\Big)|K_i^j\backslash\underline{K}_i^j|,$$

noting that $\nabla\cdot\mathbf{m}_i = d$. The right-hand side of the above inequality converges to zero as $|K_i^j\backslash\underline{K}_i^j|\to 0$, which, in turn, is achieved as $r\to\infty$. Thus, Assumption 4.1(b) gives

$$\min_{\mathbf{x}\in F_i^j}(\mathbf{m}_i\cdot\mathbf{n})\|v\|^2_{F_i^j} \le \Big|\int_{F_i^j} v^2\mathbf{m}_i\cdot\mathbf{n}\,\mathrm{d}s\Big| \le \Big|\int_{\underline{F}_i^j} v^2\mathbf{m}_i\cdot\mathbf{n}_{\underline{F}_i^j}\,\mathrm{d}s\Big| + \epsilon,$$

for some $\epsilon=\mathcal{O}(|K_i^j\backslash\underline{K}_i^j|)$ as $r\to\infty$. Each of the finite F_i's is, in turn, image of a finite number of Lipschitz functions locally. Let L be the Lipschitz constant of a parametrisation of F_i^j with respect to $\underline{F}_i^j$, giving $|F_i^j|\le L|\underline{F}_i^j|$. At the same time, we have $|K_i^j\setminus\underline{K}_i^j|\le Lh_{\underline{F}_i^j}|\underline{F}_i^j|$, as the maximum Euclidean distance between $\underline{F}_i^j$ and $\tilde{F}_i^j$ is bounded from above by $Lh_{\underline{F}_i^j}$. Hence the area $|K_i^j\setminus\underline{K}_i^j|$ converges to zero faster than $|F_i^j|$ by an order of $h_{\underline{F}_i^j}$.

At the same time, a standard trace-inverse estimate on simplices, [59], yields

$$\Big|\int_{\underline{F}_i^j} v^2\mathbf{m}_i\cdot\mathbf{n}_{\underline{F}_i^j}\,\mathrm{d}s\Big| \le \max_{\mathbf{x}\in\underline{F}_i^j}(\mathbf{m}_i\cdot\mathbf{n}_{\underline{F}_i^j})\frac{(p+1)(p+d)|\underline{F}_i^j|}{d|\underline{K}_i^j|}\|v\|^2_{\underline{K}_i^j}.$$

Combining the above, we have that, for any $\delta>0$, there exists an r large enough such that

$$\begin{aligned}\|v\|^2_{F_i^j} &\le \frac{\max_{\mathbf{x}\in\underline{F}_i^j}(\mathbf{m}_i\cdot\mathbf{n}_{\underline{F}_i^j})}{\min_{\mathbf{x}\in F_i^j}(\mathbf{m}_i\cdot\mathbf{n})}\frac{(p+1)(p+d)|\underline{F}_i^j|}{d|\underline{K}_i^j|}\|v\|^2_{\underline{K}_i^j} + \frac{\epsilon}{\min_{\mathbf{x}\in F_i^j}(\mathbf{m}_i\cdot\mathbf{n})}\\ &\le (1+\delta)\frac{(p+1)(p+d)|\underline{F}_i^j|}{d|\underline{K}_i^j|}\|v\|^2_{\underline{K}_i^j} \le (1+\delta)\frac{(p+1)(p+d)|F_i^j|}{d|\underline{K}_i^j|}\|v\|^2_{K_i^j},\end{aligned}$$

as the first ratio on the first estimate tends to 1 as $r \to \infty$. In the last inequality we used the bound $|\underline{F}_i^j| \leq |F_i^j|$ and that $\underline{K}_i^j \subset K_i^j$. Another application of the Divergence Theorem and elementary calculations give

$$d|K_i^j| = \int_{K_i^j} \nabla \cdot \mathbf{m}_i \,\mathrm{d}\mathbf{x} = \int_{F_i^j} \mathbf{m}_i \cdot \mathbf{n} \,\mathrm{d}s \geq \min_{\mathbf{x}\in F_i^j} (\mathbf{m}_i \cdot \mathbf{n}) |F_i^j|,$$

or

$$d|\underline{K}_i^j| + d|K_i^j \backslash \underline{K}_i^j| \geq \min_{\mathbf{x}\in F_i^j} (\mathbf{m}_i \cdot \mathbf{n}) |F_i^j|,$$

or

$$\frac{|F_i^j|}{|\underline{K}_i^j|} \leq \frac{d}{\min_{\mathbf{x}\in F_i^j} (\mathbf{m}_i \cdot \mathbf{n})} \Big(1 + \frac{|K_i^j \backslash \underline{K}_i^j|}{|\underline{K}_i^j|}\Big) \leq \frac{(1+\delta)d}{\min_{\mathbf{x}\in F_i^j} (\mathbf{m}_i \cdot \mathbf{n})},$$

for any $\delta > 0$ when r is sufficiently large. Combining the above, we deduce

$$\|v\|_{F_i}^2 = \sum_{j=1}^{r} \|v\|_{F_i^j}^2 \leq \sum_{j=1}^{r} \frac{(1+\delta)^2(p+1)(p+d)}{\min_{\mathbf{x}\in F_i^j} (\mathbf{m}_i \cdot \mathbf{n})} \|v\|_{K_i^j}^2 \leq \frac{(1+\delta)^2(p+1)(p+d)}{\min_{\mathbf{x}\in F_{K_i}} (\mathbf{m}_i \cdot \mathbf{n})} \|v\|_{K_i}^2.$$

Taking, finally, $r \to \infty$, allows for $\delta \to 0$ and the result (4.2) follows. □

Remark 4.5. It is important to stress that the right-hand side of (4.2) is a function of $\mathbf{x}_i^0$ defining K_{F_i}. Since the closure of the original (curved) element K is compact in $\mathbb{R}^d$, it is possible to minimise the right-hand side of (4.2) by selecting an 'optimal' $\mathbf{x}_i^0$. Moreover, upon making the stronger assumption (4.1), we arrive at the familiar trace inverse estimate for star-shaped, shape-regular elements with piecewise smooth boundaries: $\|v\|_{\partial K}^2 \leq Cp^2 h_K^{-1} \|v\|_K^2$.

Example 4.6. Let $K = B(0,R) \subset \mathbb{R}^d$ be the ball of radius R centred at the origin. Then, selecting $F_1 = \partial K =: S(0,r)$, we have $\|v\|_{S(0,R)}^2 \leq (p+1)(p+d)R^{-1}\|v\|_{B(0,R)}^2$.

Within this geometric setting, we can specify the constants of the classical trace inequality for H^1-functions. The result below is a mild extension of [23, Lemma 4.1], (cf. also [60]) following closely the classical proof from [1].

Lemma 4.7. *Let $K \in \mathcal{T}$ be a Lipschitz domain satisfying Assumption 4.1. Then, for all $\zeta > 0$, we have the estimate*

$$\|v\|_{F_i}^2 \leq \frac{d+\zeta}{\min_{\mathbf{x}\in F_i} (\mathbf{m}_i \cdot \mathbf{n})} \|v\|_{K_{F_i}}^2 + \frac{\max_{\mathbf{x}\in F_i} |\mathbf{m}_i|_2^2}{\zeta \min_{\mathbf{x}\in F_i} (\mathbf{m}_i \cdot \mathbf{n})} \|\nabla v\|_{K_{F_i}}^2, \tag{4.3}$$

for all $v \in H^1(\Omega)$ and $i = 1, \ldots, n_K$.

Proof. The Divergence Theorem and the fact that $\mathbf{m}_i \cdot \mathbf{n} = 0$ on $\partial K_{F_i} \setminus F_i$ imply

$$\int_{F_i} v^2 \mathbf{m}_i \cdot \mathbf{n} \,\mathrm{d}s = \int_{K_{F_i}} \nabla \cdot (v^2 \mathbf{m}_i) \,\mathrm{d}\mathbf{x} \leq d\|v\|_{K_{F_i}}^2 + 2 \max_{\mathbf{x}\in F_i} |\mathbf{m}_i|_2 \|v\|_{K_{F_i}} \|\nabla v\|_{K_{F_i}},$$

from which the result already follows. □

Remark 4.8. Summing over $i = 1, \ldots, n_K$, assuming (4.1) and that $h_{K_{F_i}} \sim h_K$, (4.3) gives the classical trace estimate $\|v\|_{\partial K}^2 \leq C\big(h_K^{-1}\|v\|_K^2 + h_K \|\nabla v\|_K^2\big)$.

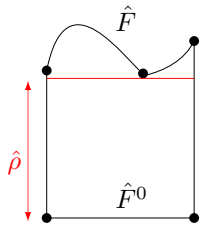

FIGURE 5. A reference generalized prism $\hat{K}$ for $d = 2$.

4.2. **Basic $H^1 - L_2$ inverse estimate.** $H^1 - L_2$ inverse estimates for polynomials on d-dimensional simplicial or box-like domains are proven via directional arguments, if explicit dependence on the polynomial degree is desired, see, e.g., [56]. Generalizations of these estimates on convex domains use an analogous method of proof [45]. Here, in the same spirit, we extend further the domain generality in $H^1 - L_2$ inverse estimates, by also employing directional arguments on curved prismatic subdomains; the general case then follows by covering general Lipschitz domains by these curved prisms.

Definition 4.9. Let $\hat{F}^0 := [0,1]^{d-1} \subset \mathbb{R}^d$ and $\phi : \hat{F}^0 \to \mathbb{R}$ a Lipschitz continuous scalar function. A *reference generalized prism* is a domain $\hat{K} \equiv \hat{K}_\phi \subset \mathbb{R}^d$ given by

$$\hat{K} \equiv \hat{K}_\phi := \{\mathbf{x} \in \mathbb{R}^d : 0 \le x_i \le 1, i = 1, \dots, d-1, 0 \le x_d \le \phi(x_1, \dots, x_{d-1})\},$$

with the properties: 1) $[0,1]^d \subset \hat{K}$, and 2) the flat base $\hat{F}^0$ is star-shaped with respect to the curved base

$$\hat{F} := \{\mathbf{x} \in \mathbb{R}^d : 0 \le x_i \le 1, i = 1, \dots, d-1, x_d = \phi(x_1, \dots, x_{d-1})\}.$$

Also, we set $\hat{\rho} := \sup\{\rho \ge 1 : \hat{F}^0 \times [0,\rho] \subset \hat{K}\}$ and $\hat{r} := \lfloor \max_{\mathbf{x} \in \hat{F}^0} \phi(\mathbf{x}) \rfloor + 1$. □

We refer to Figure 5 for an illustration.

Remark 4.10. A sufficient but, crucially, *not* necessary condition for $\hat{K} \equiv \hat{K}_\phi$ to be a reference generalized prism is that ϕ is a contraction. Since, however, $\hat{K}_\phi$ will be used in conjunction with affine maps below, it will become possible to consider ϕ with Lipschitz constants greater than one.

Remark 4.11. The 'height' $\hat{r}$ is a measure of anisotropy of the reference generalized prism. Note that we can take $\hat{\rho} = 1$ without essential loss of generality. Indeed, if $\hat{\rho} > 1$, the change of variables $x_d \to x_d/\hat{\rho}$ implies a modification of the Lipschitz function ϕ, reducing its Lipschitz constant. Star-shapedness with respect to $\hat{F}^0$ is also ensured (cf., Remark 4.10).

In light of the above remark, we consider the case $\hat{\rho} = 1$ only, in what follows.

Lemma 4.12. *Let $v \in \mathcal{P}_p(\hat{K})$, $p \in \mathbb{N}$, with $\hat{K} \subset \mathbb{R}^d$ a reference generalized prism. Then, we have the inverse estimate*

$$\|\nabla v\|^2_{\hat{K}} \le C^B_{\text{inv}} p^4 \|v\|^2_{\hat{K}}, \tag{4.4}$$

with $C^B_{\text{inv}} \equiv C^B_{\text{inv}}(d, \hat{r}) := 288(d-1)\hat{r}^2 + 12d$.

Proof. We begin by introducing some notation. Let F be a hyperplanar region in $\mathbb{R}^d$ and let vector $\mathbf{v} \in \mathbb{R}^d$. We define a *zone* $Z(F, \mathbf{v}) \subset \mathbb{R}^d$, to be the geometric locus given by $Z(F, \mathbf{v}) := \{\mathbf{z} + \beta\mathbf{v} : \mathbf{z} \in F, \beta \in \mathbb{R}\}$. (Thus, for instance, the domain $[0,1] \times \mathbb{R} = Z([0,1], (0,\alpha))$ for any $\alpha \in \mathbb{R}$.) Using this notation, we now construct $2(d-1)$ suitable subsets $\{\hat{K}^{x_j}_\pm\}_{j=1}^{d-1}$, so that the union of $\hat{K}_0 := [0,1]^d$ together with $\{\hat{K}^{x_j}_\pm\}_{j=1}^{d-1}$ cover $\hat{K}$.

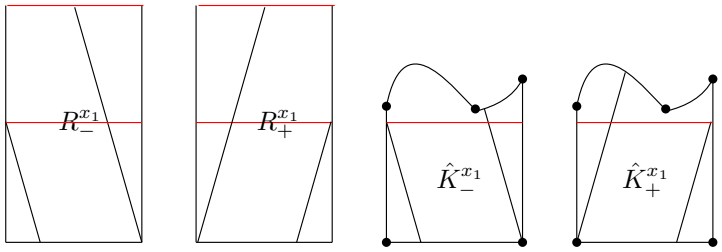

FIGURE 6. $R = [0,1] \times [0, \hat{r}]$ and $\hat{K}$, with $\hat{r} = 2$, and their respective truncated prisms $R_\pm^{x_1}$ and $\hat{K}_\pm^{x_1}$.

We first present the construction for $d = 2$ for accessibility. Set $\hat{F}_{0,-}^{x_1} := [(2\hat{r})^{-1}, 1]$ and $\hat{F}_{0,+}^{x_1} := [0, 1-(2\hat{r})^{-1}]$. Then, elementary geometric arguments reveal that the rectangle $R := [0,1] \times [0, \hat{r}]$ can be covered by the union of $R_0 := [0,1]^2$ and the 2 truncated prisms defined as:

$$R_\pm^{x_1} := R \cap Z\big(\hat{F}_{0,\pm}^{x_1}, \mathbf{v}_\pm^{x_1}\big), \text{ with } \mathbf{v}_\pm^{x_1} := \frac{(\pm 1/2, \hat{r})^T}{|(1/2, \hat{r})|},$$

with $|\cdot|$ denoting the standard Euclidean distance; we refer to Figure 6 for an illustration with $\hat{r} = 2$.

Correspondingly, for $d = 3$, $R := [0,1]^2 \times [0, \hat{r}]$ can be covered by $R_0 := [0,1]^3$ together with 2 'x_1-direction tilted', truncated prisms:

$$R_\pm^{x_1} := R \cap Z\big(\hat{F}_{0,\pm}^{x_1}, \mathbf{v}_\pm^{x_1}\big), \text{ with } \mathbf{v}_\pm^{x_1} := \frac{(\pm 1/2, 0, \hat{r})^T}{|(1/2, 0, \hat{r})|},$$

with $\hat{F}_{0,-}^{x_1} := [(2\hat{r})^{-1}, 1] \times [0,1]$ and $\hat{F}_{0,+}^{x_1} := [0, 1-(2\hat{r})^{-1}] \times [0,1]$ the respective prism bases. At the same time, R can be also covered by R_0 together with the 2 'x_2-direction tilted', truncated prisms:

$$R_\pm^{x_2} := R \cap Z\big(\hat{F}_{0,\pm}^{x_2}, \mathbf{v}_\pm^{x_2}\big), \text{ with } \mathbf{v}_\pm^{x_2} := \frac{(0, \pm 1/2, \hat{r})}{|(0, 1/2, \hat{r})^T|},$$

with $\hat{F}_{0,-}^{x_2} := [0,1] \times [(2\hat{r})^{-1}, 1]$ and $\hat{F}_{0,+}^{x_2} := [0,1] \times [0, 1-(2\hat{r})^{-1}]$ the respective prism *bases*. (We note the 'overloading' of notation with respect to dimension.) The construction for $d \geq 4$ follows in a completely analogous fashion by considering 2 'x_j-direction tilted', truncated prisms for each $j = 1, \ldots, d-1$.

Since $\hat{K} \subset [0,1]^{d-1} \times [0, \hat{r}]$, we consider the sets $\hat{K}_0 = R_0$ together with

$$\hat{K}_\pm^{x_j} := \hat{K} \cap R_\pm^{x_j},$$

for each fixed $j = 1, \ldots, d-1$; see Figure 6 for an illustration for $d = 2$ and $\hat{r} = 2$.

First, we observe the estimates

$$|v_{x_j}|^2 \leq 8\hat{r}^2 |v_{x_d}|^2 + 2(4\hat{r}^2 + 1)|\mathbf{v}_\pm^{x_j} \cdot \nabla v|^2, \tag{4.5}$$

for $j = 1, \ldots, d-1$. Using the latter, we have, respectively,

$$\begin{aligned} \|\nabla v\|_{\hat{K}}^2 - \|v_{x_d}\|_{\hat{K}}^2 - \sum_{j=1}^{d-1} \|v_{x_j}\|_{\hat{K}_0}^2 &\leq \sum_{j=1}^{d-1} \Big(\|v_{x_j}\|_{\hat{K}_+^{x_j}}^2 + \|v_{x_j}\|_{\hat{K}_-^{x_j}}^2 \Big) \\ &\leq \sum_{j=1}^{d-1} \Big(8\hat{r}^2 \|v_{x_d}\|_{\hat{K}}^2 + 2(4\hat{r}^2+1)\big(\|\mathbf{v}_+^{x_j} \cdot \nabla v\|_{\hat{K}_+^{x_j}}^2 + \|\mathbf{v}_-^{x_j} \cdot \nabla v\|_{\hat{K}_-^{x_j}}^2 \big) \Big). \end{aligned} \tag{4.6}$$

We now estimate each term on the right-hand side of (4.6). For $\mathrm{e}_d := (0, \ldots, 0, 1)^{\mathrm{T}}$, let $\ell_{\mathbf{x}} := \hat{K} \cap \{\mathbf{x} + \alpha \mathrm{e}_d : \alpha \in \mathbb{R}\}$, i.e., the vertical line contained in $\hat{K}$ and passing

through a point $\mathbf{x}$. Then, we have

$$\|v_{x_d}\|^2_{\hat{K}} = \int_{\hat{F}^0}\int_{\ell_{\mathbf{x}}} v^2_{x_d}\,\mathrm{d}x_d\,\mathrm{d}\mathbf{x} \leq 12p^4\|v\|^2_{\hat{K}}, \tag{4.7}$$

from Fubini's Theorem and an *one-dimensional* inverse estimate, see, e.g., [56, Theorem 3.91]. We set $\ell^{x_j}_{\mathbf{x},\pm} := \hat{K}\cap\{\mathbf{x}+\alpha\mathbf{v}^{x_j}_{\pm} : \alpha\in\mathbb{R}\}$. Then,

$$\|\mathbf{v}^{x_j}_{\pm}\cdot\nabla v\|^2_{\hat{K}^{x_j}_{\pm}} = \int_{\hat{F}^{x_j}_{0,\pm}}\int_{\ell^{x_j}_{\mathbf{x},\pm}} (\mathbf{v}^{x_j}_{\pm}\cdot\nabla v)^2\,\mathrm{d}\zeta\,\mathrm{d}\mathbf{x} \leq \frac{48\hat{r}^2p^4}{4\hat{r}^2+1}\|v\|^2_{\hat{K}},$$

for $j=1,\ldots,d$, upon noticing that the length of $\ell^{x_j}_{\mathbf{x},\pm}$ is bounded from below by $\sqrt{\hat{r}^2+1/4}/\hat{r}$ (the length of the portion of $\ell^{x_j}_{\mathbf{x},\pm}$ contained in R_0). Thus, (4.6) implies

$$\|\nabla v\|^2_{\hat{K}} - \|v_{x_d}\|^2_{\hat{K}} - \sum_{j=1}^{d-1}\|v_{x_j}\|^2_{\hat{K}_0} \leq 288p^4(d-1)\hat{r}^2\|v\|^2_{\hat{K}}.$$

The result already follows by combining the last estimate with (4.7) and the corresponding inverse estimates for $\|v_{x_j}\|^2_{[0,1]^d}$. □

Notice that (4.4) retrieves the known constant for $d=1$. If K is cuspoidal, (i.e., not Lipschitz) (4.4) does *not* hold in general; we refer to [46] for a counterexample.

Remark 4.13. Careful inspection of the above proof shows that, in fact, we have proven the sharper inverse estimate $\|\nabla v\|^2_{\hat{K}} \leq C^B_{\mathrm{inv}}p^4\hat{\rho}^{-2}\|v\|^2_{\hat{K}}$. In view of Remark 4.11, however, a linear scaling results into a modified $\hat{K}$ (and, possibly modified $\hat{r}$) for which (4.4) is sharp.

4.3. **Stability with respect to domain perturbation.** We now prove a stability result with respect to domain perturbation in the spirit of [34, Lemma 3.7] (see also [45, Lemma 6]).

Lemma 4.14. *Let $\hat{K}$ a reference generalized prism and consider its subset $\hat{K}_\epsilon := \hat{K}\cap(\hat{K}-\epsilon\mathrm{e}_d)$; here $A+z := \{x+z, x\in A\}$, for $A\subset\mathbb{R}^d$ and $z\in\mathbb{R}^d$. Then, for all $v\in\mathcal{P}_p(\hat{K})$, $p\in\mathbb{N}$, and for any $0<\epsilon\leq(8p)^{-2}$, we have the estimate*

$$\frac{1}{2}\|v\|^2_{\hat{K}} \leq \|v\|^2_{\hat{K}_\epsilon}. \tag{4.8}$$

Proof. Set $\ell_{\mathbf{x},\epsilon} := \hat{K}\cap(\ell_{\mathbf{x}}-\epsilon\mathrm{e}_d)$. Then, we have, respectively,

$$\begin{aligned}\|v\|^2_{\hat{K}\setminus\hat{K}_\epsilon} &= \int_{\hat{F}^0}\int_{\ell_{\mathbf{x}}\setminus\ell_{\mathbf{x},\epsilon}} v^2\,\mathrm{d}x_d\,\mathrm{d}\mathbf{x} \leq \epsilon\int_{\hat{F}^0}\|v\|^2_{L_\infty(\ell_{\mathbf{x}}\setminus\ell_{\mathbf{x},\epsilon})}\,\mathrm{d}\mathbf{x}\\ &\leq \epsilon\int_{\hat{F}^0}\|v\|^2_{L_\infty(\ell_{\mathbf{x}})}\,\mathrm{d}\mathbf{x} \leq 32\epsilon p^2\|v\|^2_{\hat{K}},\end{aligned} \tag{4.9}$$

by Markov's inequality (see, e.g., [56, Theorem 3.92]) since the length of $\ell_{\mathbf{x}}$ is bounded from below by one. Selecting now $0<\epsilon\leq(8p)^{-2}$, the result follows, by simply observing that $\|v\|^2_{\hat{K}} - \|v\|^2_{\hat{K}_\epsilon} = \|v\|^2_{\hat{K}\setminus\hat{K}_\epsilon}$. □

Example 4.15. Consider $\hat{K}$ with $\phi(\mathbf{x}) = 1+\mathbf{x}\cdot(\mathbf{1}-\mathbf{x})+(16p)^{-2}\sin(\alpha\pi\mathbf{1}\cdot\mathbf{x})$, for some $\alpha\geq 1$, with $\mathbf{1} := (1,1,\ldots,1)^T\in\mathbb{R}^{d-1}$. This is chosen so that $\phi(\mathbf{x})\in[1+\mathbf{x}\cdot(\mathbf{1}-\mathbf{x})-\epsilon, 1+\mathbf{x}\cdot(\mathbf{1}-\mathbf{x})+\epsilon]$, for ϵ within the range required for the statement of Lemma 4.14 to hold. For sufficiently large α, $\hat{K}$ is *not* star-shaped with respect to $\hat{F}^0=[0,1]^{d-1}$. Nevertheless, ϕ is sufficiently approximated by $\phi_2(\mathbf{x}) := 1+\mathbf{x}\cdot(\mathbf{1}-\mathbf{x})$, which is star-shaped with respect to $\hat{F}^0=[0,1]^{d-1}$. Thus, for $v\in\mathcal{P}_p(\hat{K})$, from Lemma 4.12, we have

$$\|\nabla v\|^2_{\hat{K}} \leq 2\|\nabla v\|^2_{\hat{K}_{\phi_2}} \leq 4C^B_{\mathrm{inv}}(d,2)p^4\|v\|^2_{\hat{K}}.$$

4.4. $L_\infty - L_2$ **inverse estimate.** We continue by proving an $L_\infty - L_2$ inverse estimate for reference generalized prisms.

Lemma 4.16. *Let $\hat{K} \equiv \hat{K}_\phi \subset \mathbb{R}^d$ a reference generalized prism. Then, the inverse estimate*

$$\|v\|^2_{L_\infty(\hat{K})} \leq (32p^2)^d \|v\|^2_{\hat{K}}, \tag{4.10}$$

holds for all $v \in \mathcal{P}_p(\hat{K})$, $p \in \mathbb{N}$.

Proof. Let $\mathbf{x}_{\max} \in \hat{K}$ such that $\|v\|_{L_\infty(\hat{K})} = |v(\mathbf{x}_{\max})|$. Then, either $\mathbf{x}_{\max} \in \hat{K}_0 = [0,1]^d \subset \hat{K}$ or $\mathbf{x}_{\max} \in \hat{K} \backslash \hat{K}_0$. Let now $\hat{K}(\mathbf{x}_{\max})$ be the pyramid with vertex $\mathbf{x}_{\max}$ and base $\hat{F}^0 = [0,1]^{d-1}$. If $\mathbf{x}_{\max} \in \hat{K}_0$, then, we have $\|v\|^2_{L_\infty(\hat{K})} \leq 32^d p^{2d} \|v\|^2_{\hat{K}}$ (see, e.g., [56, eq. (3.6.4)],) whereas if $\mathbf{x}_{\max} \in \hat{K} \backslash \hat{K}_0$, we have, respectively, $|v(\mathbf{x}_{\max})|^2 = \|v\|^2_{L_\infty(\hat{K})} = \|v\|^2_{L_\infty(\hat{K}(\mathbf{x}_{\max}))}$. Since $\mathbf{x}_{\max}$ is a vertex, we employ an one-dimensional trace inverse estimate [59] iteratively with respect to dimension, to deduce $|v(\mathbf{x}_{\max})| \leq \prod_{j=1}^d (p+j) \|v\|_{\hat{K}(\mathbf{x}_{\max})}$. Here we have used the fact that the dimensions of $\hat{K}(\mathbf{x}_{\max})$ are grater than 1. Combining the two cases and taking the maximum constant, the result already follows. □

Remark 4.17. An inspection of the proof of Lemma 4.16 show that the only geometric assumption needed is that the curved face $\hat{F}$ is given as the graph of a Lipschitz function ϕ. So Lemma 4.16 holds without assuming property 2) in Definition 4.9.

4.5. **Inverse estimates on general domains.** We now extend the above inverse estimates to general curved polytopic elements $K \in \mathcal{T}$. To that end, we shall relax the concept of p-coverability of polytopic elements introduced in [21], (see also [18, 20]) from simplicial coverings of general-shaped elements $K \in \mathcal{T}$, to coverings involving affinely mapped generalized prisms.

Definition 4.18. An element $K \in \mathcal{T}$ is said to be *p-coverable* with respect to $p \in \mathbb{N}$, if there exists a set of $m_K \in \mathbb{N}$ generalized prisms $\hat{K}_j$ and corresponding affine maps Φ_j, such that the mapped generalized prisms $\overline{K}_j := \Phi_j(\hat{K}_j)$, $j = 1, \ldots, m_K$, form a, possibly overlapping, covering of K with the additional properties

$$\operatorname{dist}(\partial \overline{K}_j, K) \leq \mathfrak{h}_{\overline{K}_j} (8p)^{-2} \tag{4.11}$$

and

$$|\overline{K}_j| \geq c_{as} |K|, \tag{4.12}$$

for all $j = 1, \ldots, m_K$, where $\mathfrak{h}_{\overline{K}_j} := \sup_{\mathbf{x} \in \hat{F}^0} |\Phi_j(\ell_{\mathbf{x},j})|$ and c_{as} is a positive constant, independent of K and of $\mathcal{T}$, with $\operatorname{dist}(\partial \overline{K}_j, K) := \sup_{\mathbf{x} \in \partial \overline{K}} \inf_{\mathbf{y} \in K} |\mathbf{x} - \mathbf{y}|$ the one-sided Hausdorff distance of $\partial \overline{K}_j$ from K, and $\ell_{\mathbf{x},j} := \hat{K}_j \cap \{\mathbf{x} + \alpha \mathrm{e}_d : \alpha \in \mathbb{R}\}$.

The motivation for the above definition is the stability result for polynomials with respect to domain perturbation given in Lemma 4.14 above. If K is p-coverable, (4.11) implies that there exists a covering of affinely mapped generalized prisms $\overline{K}_j$ and respective sub-prisms $\underline{K}_j := \overline{K}_{j,\epsilon}$, $0 < \epsilon \leq \mathfrak{h}_{\overline{K}_j} (8p)^{-2}$, $j = 1, \ldots, m_K$, such that $\underline{K}_j \subset K$. Then, we have, for any $v \in \mathcal{P}_p(K)$,

$$\frac{1}{2} \|v\|^2_{\overline{K}_j} \leq \|v\|^2_{\underline{K}_j} \leq \|v\|^2_K. \tag{4.13}$$

We now show that (4.11) is implied by Assumptions 4.1 and 4.3. Therefore, p-coverability for an element satisfying Assumptions 4.1 and 4.3 is ensured under the validity of (4.12) only.

Lemma 4.19. *Let $K \in \mathcal{T}$ satisfying Assumptions 4.1 and 4.3. Then, there exists a set of $m_K \in \mathbb{N}$ generalized prisms $\hat{K}_j$ and corresponding affine maps Φ_j, such that the mapped domains $\overline{K}_j := \Phi_j(\hat{K}_j)$, $j = 1, \ldots, m_K$, form a cover of K with the property*

$$\operatorname{dist}(\partial\overline{K}_j, K) \leq \mathfrak{h}_{\overline{K}_j}(8p)^{-2} \tag{4.14}$$

for all $j = 1, \ldots, m_K$, with the notation of Definition 4.18.

Proof. From Assumption 4.3, ∂K is comprised of a finite number of closed (co-dimension one) C^1 surfaces $(\partial K)_j$, $j = 1, \ldots, \tilde{z}_K$, for some $\tilde{z}_K \in \mathbb{N}$. By possibly further subdividing the $(\partial K)_j$'s into subsets, say, $(\partial K)_j$, $j = 1, \ldots, z_K$, Assumption 4.1, ensures that for each of $(\partial K)_j$ there exist a point $\mathbf{x}_j^0 \in K$ such that the curved simplex $K_{(\partial K)_j} \equiv K_{(\partial K)_j}(\mathbf{x}_j^0)$ is star-shaped with respect to $\mathbf{x}_j^0$ and that $g_j(\mathbf{x}) := (\mathbf{x} - \mathbf{x}_j^0) \cdot \mathbf{n}(\mathbf{x}) > 0$ for any $\mathbf{x} \in (\partial K)_j$. (More than one $(\partial K)_j$ are allowed to share the same $\mathbf{x}_j^0$.) Since $(\partial K)_j$ is C^1, g_j is continuous on $(\partial K)_j$ and, thus, there exists a positive number δ_j, such that $g_j(\mathbf{x}) \geq \delta_j$.

Now, for any $\tilde{\mathbf{x}}_j^0 \in \mathbb{R}^d$, with $|\mathbf{x}_j^0 - \tilde{\mathbf{x}}_j^0| < \delta_j$, we have

$$|g_j(\mathbf{x}) - (\mathbf{x} - \tilde{\mathbf{x}}_j^0) \cdot \mathbf{n}(\mathbf{x})| \leq |\mathbf{x}_j^0 - \tilde{\mathbf{x}}_j^0| < \delta_j,$$

and, therefore, $\tilde{g}_j(\mathbf{x}) := (\mathbf{x} - \tilde{\mathbf{x}}_j^0) \cdot \mathbf{n}(\mathbf{x}) > 0$ for any $\mathbf{x} \in (\partial K)_j$. Hence, $(\partial K)_j$ is star-shaped with respect to $B(\mathbf{x}_j^0, \delta_j)$ in K; that is any line connecting any point of $(\partial K)_j$ with a point of $B(\mathbf{x}_j^0, \delta_j)$ lies wholly in K. This implies that $K_{(\partial K)_j}$ is star-shaped in K with respect to any subset of $B(\mathbf{x}_j^0, \delta_j)$ and, in particular, with respect to any $(d-1)$-hypercube passing through $\mathbf{x}_j^0$ and contained in $B(\mathbf{x}_j^0, \delta_j)$. In general, however, $B(\mathbf{x}_j^0, \delta_j) \not\subset K$, but we have $\operatorname{dist}(\partial B(\mathbf{x}_j^0, \delta_j), K) \leq |\mathbf{x}_j^0 - \tilde{\mathbf{x}}_j^0| < \delta_j$. For the boundary pieces $(\partial K)_j$ with $B(\mathbf{x}_j^0, \delta_j) \subset K$, we fix δ_j to its largest possible value ensuring $B(\mathbf{x}_j^0, \delta_j) \subset K$. If, however, $B(\mathbf{x}_j^0, \delta_j) \not\subset K$, we select δ_j small enough, so that $\operatorname{dist}(\partial B(\mathbf{x}_j^0, \delta_j), K) \leq \mathfrak{h}_{\overline{K}_j}(8p)^{-2}$.

On the other hand, the C^1 smoothness of $(\partial K)_j$ ensures that there exists a finite tessellation comprising of diagonally scaled and rotated $(d-1)$-hypercubes approximating $(\partial K)_j$ to a desired accuracy, say $\delta_j/2$. Consider now the truncated prisms intersecting $(\partial K)_j$ defined uniquely by the 2^{d-1} vertices of each element of the tessellation and the 2^{d-1} vertices of a second $(d-1)$-hypercubical base contained in $B(\mathbf{x}_j^0, \delta_j)$ and passing through $\mathbf{x}_j^0$. The union of the latter generalized prisms covers $K_{(\partial K)_j}$ within a distance δ_j. Considering the corresponding construction for all j, we conclude the construction of a finite cover of K by affinely mapped generalized prisms such that (4.14) holds. □

Remark 4.20. The purpose of the construction in the proof of Lemma 4.19 is to assert the existence of at least *one* covering with the required properties, and not to construct the 'optimal' one.

Lemma 4.21. *Let $K \in \mathcal{T}$ Lipschitz satisfying Assumption 4.1. Then, for each $v \in \mathcal{P}_p(K)$, we have the inverse inequality*

$$\|v\|_{F_i}^2 \leq \mathcal{C}_{\mathrm{INV}}(p, K, F_i) \frac{(p+1)(p+d)|F_i|}{|K|} \|v\|_K^2, \tag{4.15}$$

with

$$\mathcal{C}_{\mathrm{INV}}(p, K, F_i) := \begin{cases} \min\left\{\mathcal{C}_{\mathrm{reg}}(K, F_i), 2c_{as}^{-1} 32^d p^{2(d-1)}\right\}, & K \text{ p-coverable,} \\ \mathcal{C}_{\mathrm{reg}}(K, F_i), & \text{otherwise,} \end{cases} \tag{4.16}$$

with $\mathcal{C}_{\mathrm{reg}}(K, F_i) := |K| / \left(|F_i| \sup_{\mathbf{x}_i^0 \in K} \min_{\mathbf{x} \in F_i} (\mathbf{m}_i \cdot \mathbf{n})\right)$, and $c_{as} > 0$ as in Def. 4.18.

Proof. If K is not p-coverable, using (4.2), we simply have

$$\|v\|_{F_i}^2 \leq \frac{(p+1)(p+d)}{\min_{\mathbf{x}\in F_i}(\mathbf{m}_i\cdot\mathbf{n})}\|v\|_{K_{F_i}}^2 \leq \frac{(p+1)(p+d)|F_i|}{|K|}\mathcal{C}_{\mathrm{reg}}(K,F_i)\|v\|_K^2. \tag{4.17}$$

If, on the other hand, K is p-coverable, then $K_{F_i}\subset K\subset \cup_{j=1,\ldots,m_K}\overline{K}_j$ and, thus,

$$\|v\|_{F_i}^2 \leq |F_i|\|v\|_{L_\infty(F_i)}^2 \leq |F_i|\|v\|_{L_\infty(K_{F_i})}^2 \leq |F_i| \max_{j=1,\ldots,m_K}\|v\|_{L_\infty(\overline{K}_j)}^2.$$

Now, Lemma 4.16 (together with an elementary scaling argument), along with (4.11) and (4.13), imply

$$\|v\|_{L_\infty(\overline{K}_j)}^2 \leq 32^d p^{2d}|\overline{K}_j|^{-1}\|v\|_{\overline{K}_j}^2 \leq 2(32p^d)^2 c_{as}^{-1}|K|^{-1}\|v\|_K^2.$$

Combining the last two estimates, taking the supremum over $\mathbf{x}_i^0\in K$, the inverse estimate constant is then given by the minimum of the two estimates. □

The above result generalizes both [20, Lemma 11] and [23, Lemma 4.9] in a number of ways. The coverings are now allowed to consist of curved domains. Also, elements with arbitrary number of (curved) faces are now admissible and an earlier hypothesis on uniform boundedness of m_K across the mesh has now been removed by a more careful analysis. Note that, when $K\in\mathcal{T}$ is a polytopic element with straight faces, Lemma 4.21 collapses to [20, Lemma 11] with improved constants.

Remark 4.22. The sub-division $\{F_i\}_{i=1}^{n_K}$ of the (curved) element boundary ∂K is typically *not* unique. We can seek to minimize the coefficient (4.16) by considering different candidates for $\{F_i\}_{i=1}^{n_K}$. However, such optimization would be practically beneficial only for rather "exotic" element shapes as, in most cases, we can simply resort to (4.1). Of course, extremely general curved "exotic" element shapes must be used only when deemed beneficial for the particular problem at hand. In such cases, a basic geometric study for improving the constant (4.16) (and, therefore, as we shall see below, the dG discontinuity-penalization function, cf. Remark 5.3 below) may be in order. In any case, Lemma 4.21 is sharp for each given subdivision $\{F_i\}_{i=1}^{n_K}$ and directly generalizes the inverse estimates in [20].

Next, we present an H^1-L_2-inverse inequality for polynomials on a general curved element $K\in\mathcal{T}$.

Lemma 4.23. *Let $K\in\mathcal{T}$ satisfy Assumptions 4.1 and 4.3. Then, for each $v\in\mathcal{P}_p(K)$, the inverse estimate*

$$\|\nabla v\|_K^2 \leq \mathcal{C}_{\mathrm{INV}}^B(p,K)\frac{p^4}{\rho_K^2}\|v\|_K^2, \tag{4.18}$$

holds, with ρ_ω denoting the radius of the largest inscribed circle of a domain $\omega\subset\mathbb{R}^d$, and

$$\mathcal{C}_{\mathrm{INV}}^B(p,K) := \begin{cases} 4\min\{\rho_{\mathrm{cov}}(K),\rho_{\mathrm{p-cov}}(p,K)\}, & \text{if } K \text{ p-coverable} \\ 4\rho_{\mathrm{cov}}(K), & \text{otherwise,}\end{cases} \tag{4.19}$$

with

$$\rho_{\mathrm{cov}}(K) := \sum_{j=1}^{m_K} C_{\mathrm{inv}}^B(d,\hat{r}_j)\Big(\frac{\hat{r}_j\rho_K}{\rho_{\overline{K}_j}}\Big)^2, \tag{4.20}$$

and

$$\rho_{\mathrm{p-cov}}(p,K) := c_{as}^{-1}(32)^d(p-1)^{2d}\max_{1\leq\ell\leq m_K} C_{\mathrm{inv}}^B(d,\hat{r}_\ell)\Big(\frac{\hat{r}_\ell\rho_K}{\rho_{\overline{K}_\ell}}\Big)^2, \tag{4.21}$$

for $\overline{K}_\ell$, $\ell=1,\ldots,m_K$, cover of K consisting of affinely mapped generalised prisms.

Proof. From Lemma 4.19, there exists a cover of K, consisting of affinely mapped generalized prisms $\overline{K}_j$, $j = 1, \dots, m_K$. Thus, for $v \in \mathcal{P}_p(K)$, Lemma 4.12, (with a standard affine scaling) and (4.13) imply:

$$\|\nabla v\|_K^2 \leq \sum_{j=1}^{m_K} \|\nabla v\|_{\overline{K}_j}^2 \leq \sum_{j=1}^{m_K} C_{\mathrm{inv}}^B(d, \hat{r}_j) \frac{2\hat{r}_j^2 p^4}{\rho_{\overline{K}_j}^2} \|v\|_{\overline{K}_j}^2 \leq 4 \sum_{j=1}^{m_K} C_{\mathrm{inv}}^B(d, \hat{r}_j) \frac{\hat{r}_j^2 p^4}{\rho_{\overline{K}_j}^2} \|v\|_K^2,$$

with $\hat{r}_j$ denoting the $\hat{r}$ of $\overline{K}_j$ as per Definition 4.9. Thus, we have

$$\|\nabla v\|_K^2 \leq \rho_{\mathrm{cov}}(K) \frac{p^4}{\rho_K^2} \|v\|_K^2. \tag{4.22}$$

Note that $\rho_{\mathrm{cov}}(K)$ grows with $\rho_K / \min_{j=1,\dots,m_K} \rho_{\overline{K}_j}$ growing.

On the other hand, if K is p-coverable, there exists $\ell \in \{1, \dots, m_K\}$ such that $\|\nabla v\|_K^2 \leq |K| \|\nabla v\|_{L_\infty(K)}^2 \leq |K| \|\nabla v\|_{L_\infty(\overline{K}_\ell)}^2$. Using now Lemmata 4.16 and 4.12 (with scaling), as well as (4.12), we deduce, respectively,

$$\begin{aligned} \|\nabla v\|_{L_\infty(\overline{K}_\ell)}^2 &\leq 32^d (p-1)^{2d} |\overline{K}_\ell|^{-1} \|\nabla v\|_{\overline{K}_\ell}^2 \\ &\leq 32^d (p-1)^{2d} p^4 C_{\mathrm{inv}}^B(d, \hat{r}_\ell) 2\hat{r}_\ell^2 \rho_{\overline{K}_\ell}^{-2} |\overline{K}_\ell|^{-1} \|v\|_{\overline{K}_\ell}^2. \end{aligned} \tag{4.23}$$

The result already follows by combining (4.22) and (4.23). □

In the special case of an element K being star-shaped with respect to a contained ball, we can have a more precise statement in terms of the constants involved.

Corollary 4.24. *Let $K \subset \mathbb{R}^d$ domain which is star-shaped with respect to a ball $B(\mathbf{x}, \rho_K) \subset K$, $\mathbf{x} \in K$. Then, for any $v \in \mathcal{P}_p(K)$, we have the inverse estimate*

$$\|\nabla v\|_K^2 \leq C(d) \Big(\frac{h_K}{\rho_K} \Big)^{d+5} \frac{p^4}{\rho_K^2} \|v\|_K^2,$$

for some universal constant $C(d) > 0$ that can be estimated explicitly. Thus, if additionally, K is shape-regular, i.e., $h_K \sim \rho_K$, we retrieve the classical inverse estimate $\|\nabla v\|_K^2 \leq C p^4 / h_K^2 \|v\|_K^2$, with C now also dependent on the shape-regularity constant.

Proof. We have $B(\mathbf{x}, \rho_K) \subset K \subset B(\mathbf{x}, h_K)$. A comparison of the area of the largest $(d-1)$-hypercube contained in $B(\mathbf{x}, \rho_K)$, given by $\rho_K^{d-1}/2^{(d-1)/2}$, with the surface of $B(\mathbf{x}, h_K)$, shows that we can cover K using $z := \lfloor 2^{(d+1)/2} \pi^{d/2} h_K^{d-1} / (\Gamma(d/2) \rho_K^{d-1}) \rfloor + 1$ mapped right generalized prisms K_j, $j = 1, \dots, z$, whose bases are given by rotations of the largest $(d-1)$-hypercube contained in $B(\mathbf{x}, \rho_K)$. So, we have

$$\|\nabla v\|_K^2 \leq \sum_{j=1}^{z} \|\nabla v\|_{K_j}^2 \leq \sum_{j=1}^{z} C_{\mathrm{inv}}^B(d, \hat{r}_j) \frac{2\hat{r}_j^2 p^4}{\rho_{K_j}^2} \|v\|_{K_j}^2;$$

here we have used scaling via $\Phi_j : \hat{K}_j \to K_j$ affine mapping $[0,1]^d$ to a rotation of the largest d-hypercube contained in $B(\mathbf{x}, \rho_K)$. Since each K_j is right, we have $\hat{r}_j \leq \lfloor h_K / \rho_{K_j} \rfloor + 1$, for all $j = 1, \dots, z$. Also, from the star-shapedness with respect to $B(\mathbf{x}, \rho_K)$, we have $\rho_{K_j} \geq \rho_K / 2$. Combining the above, we deduce

$$\|\nabla v\|_K^2 \leq 8 C_{\mathrm{inv}}^B(\lfloor h_K / \rho_{K_j} \rfloor + 1, d) (\lfloor h_K / \rho_{K_j} \rfloor + 1)^2 \frac{p^4}{\rho_K^2} \sum_{j=1}^{z} \|v\|_{K_j}^2.$$

Using the (pessimistic) bound $\sum_{j=1}^{z} \|v\|_{K_j}^2 \leq z \|v\|_K^2$, and combining the numerical constants, the result follows. □

The last result holds under weaker domain assumptions compared to [45, Theorem 1] and, in contrast to the main result in [46], it offers explicit dependence on the domain size in the case of piecewise C^1 domains. We also note [45, Theorem 3], which provides a similar bound for the special case of K being a $d-$ellipsoid, in conjunction with John's Ellipsoid Theorem. It is interesting to investigate the extension of the above inverse estimates with explicit constants to cuspoidal domains in the spirit of [46]; this will be considered elsewhere.

Example 4.25. We revisit Example 4.6 for $d = 2$, with $K = B(0, R)$ a circular element with radius R. Let $K_1 = B(0, R) \cap \{(x_1, x_2) : x_1 \in \mathbb{R},\ -\sqrt{2}R/2 \leq x_2 \leq \sqrt{2}R/2\}$ and $K_2 = B(0, R) \cap \{(x_1, x_2) : -\sqrt{2}R/2 \leq x_1 \leq \sqrt{2}R/2,\ x_2 \in \mathbb{R}\}$, so that $K = K_1 \cup K_2$. We further subdivide each K_i in half to form prisms with flat base; for each of these, we can select $\hat{r} = 2$. Thus, Lemma 4.23 implies

$$\|\nabla v\|_K^2 \leq \mathcal{C}_{\mathrm{INV}}^B \frac{p^4}{R^2} \|v\|_K^2,$$

with $\mathcal{C}_{\mathrm{INV}}^B \leq 2304$. The constant in this special case can be improved considerably upon taking advantage of the circle's symmetries.

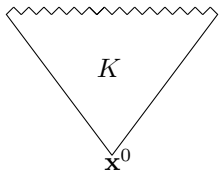

FIGURE 7. Example 4.26. $K \in \mathcal{T}$ with 'multiscale' boundary behaviour.

Example 4.26. Let $d = 2$, and consider the polygonal element $K \in \mathcal{T}$ with 'multiscale' boundary behaviour depicted in Figure 7. Denoting by r the length of each of the (equal length) n small faces and with h_K its diameter, we consider the case $r \ll h_K$. If $r < h_K/p^2$, we can cover K by one triangle , namely, the smallest simplex containing K. Then K is p-coverable and $\mathcal{C}_{\mathrm{INV}}^B(p, K)$ remains bounded, independently of n. Hence, when the two geometric scales h_K and r are significantly different, K is essentially a simplex in this context.

On the other hand, for p large enough and fixed r and n, we have $r > h_K/p^2$ and, hence, we cannot cover K as before. Instead, we consider a family of n non-overlapping simplices $\overline{K}_j \subset K$, each defined by one 'small' face of length r and the vertex $\mathbf{x}^0$. Then, we have $c_{as} = n^{-1}$ in Definition 4.18 and $\rho_{\overline{K}_j} \sim h_K/n$. Since also $\rho_K \sim h_K$ and $q_\ell = 1$, we compute $\mathcal{C}_{\mathrm{INV}}^B(p, K) \sim n^{-1}$. This is reasonable, as sufficiently high polynomial degree p basis functions can resolve the scale of the 'sawtooth' face ensemble.

4.6. **Best approximation estimates.** We now turn to *hp*-version polynomial approximation bounds over general domains. The setting here remains essentially unchanged compared to the case of just polytopic elements presented in [21, 20]. More specifically, under a mild set of covering assumptions and upon postulating the existence of so-called function space domain extension operators, we are able to apply *hp*-version best approximation results in various seminorms.

Definition 4.27. Given a mesh $\mathcal{T}$, we define a *covering* $\mathcal{T}^\sharp = \{\mathcal{K}\}$ of $\mathcal{T}$ to be a set of open shape-regular d–simplices $\mathcal{K}$, such that for each $K \in \mathcal{T}$, there exists a $\mathcal{K} \in \mathcal{T}^\sharp$ with $K \subset \mathcal{K}$. For a given $\mathcal{T}^\sharp$, we define the *covering domain* $\bar{\Omega}_\sharp := \cup_{\mathcal{K} \in \mathcal{T}^\sharp} \bar{\mathcal{K}}$.

For illustration, in Figure 8 we show a single two-dimensional curved element $K \in \mathcal{T}$, along with a covering simplex $\mathcal{K} \in \mathcal{T}^\sharp$ with $K \subset \mathcal{K}$.

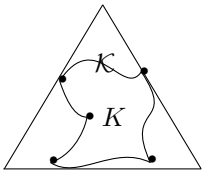

FIGURE 8. A simplex $\mathcal{K} \in \mathcal{T}^\sharp$ covering an element $K \in \mathcal{T}$.

Assumption 4.28. For a given mesh $\mathcal{T}$, we postulate the existence of a covering $\mathcal{T}^\sharp$, and of a (global) constant $\mathcal{O}_\Omega \in \mathbb{N}$, independent of the mesh parameters, such that

$$\max_{K\in\mathcal{T}} \mathrm{card}\Big\{K' \in \mathcal{T} : K' \cap \mathcal{K} \neq \emptyset,\ \mathcal{K} \in \mathcal{T}^\sharp \ \text{ such that } \ K \subset \mathcal{K}\Big\} \leq \mathcal{O}_\Omega.$$

For such $\mathcal{T}^\sharp$, we further assume that $h_\mathcal{K} := \mathrm{diam}(\mathcal{K}) \leq C_{\mathrm{diam}} h_K$, for all pairs $K \in \mathcal{T}$, $\mathcal{K} \in \mathcal{T}^\sharp$, with $K \subset \mathcal{K}$, for a (global) constant $C_{\mathrm{diam}} > 0$, uniformly with respect to the mesh size h_K.

Remark 4.29. Assumption 4.28 ensures the shape–regularity of the mesh covering $\mathcal{T}^\sharp$ only. Shape-regularity of the mesh $\mathcal{T}$ is *not* assumed. We refer to Figure 3.6 in [20] for an example on how these two concepts may differ considerably.

The validity of Assumption 4.28 allows for the use known hp–version approximation estimates on simplicial elements [6, 7, 56], on each $\mathcal{K}$ and, subsequently restrict the error over $K \subset \mathcal{K}$. However, it requires to extend the exact solution u into $\Omega_\sharp$ in a stable fashion. To that end, we shall use the following classical result.

Theorem 4.30 ([57])**.** *Let Ω be a domain with a Lipschitz boundary. Then there exists a linear extension operator $\mathfrak{E} : H^s(\Omega) \mapsto H^s(\mathbb{R}^d)$, $s \in \mathbb{N}_0$, such that $\mathfrak{E}v|_\Omega = v$ and $\|\mathfrak{E}v\|_{H^s(\mathbb{R}^d)} \leq C_\mathfrak{E}\|v\|_{H^s(\Omega)}$, where the constant $C_\mathfrak{E} > 0$ depends only on s, Ω.*

Subsequent refinements of the dependence of the constant $C_\mathfrak{E}$ on the domain shape in Theorem 4.30, have been presented for instance in [55, 24].

For the estimation of the best approximation error on the mesh skeleton $\Gamma_{\mathrm{int}} \cup \partial\Omega$, we require the trace estimate on curved domains from Lemma 4.7.

We now have all the ingredients to assert the validity of the following hp-approximation error bounds.

Lemma 4.31. *Let $K \in \mathcal{T}$ satisfy Assumptions 4.1 and 4.28, and let $\mathcal{K} \in \mathcal{T}^\sharp$ be the corresponding simplex with $K \subset \mathcal{K}$ as per Definition 4.27. Suppose that $v \in L_2(\Omega)$ is such that $\mathfrak{E}v|_\mathcal{K} \in H^{l_K}(\mathcal{K})$, for some $l_K \geq 0$, and that Assumption 4.28 is satisfied. Then, there exists an operator $\pi_p : H^{l_K}(\mathcal{K}) \to \mathcal{P}_p(\mathcal{K})$, such that*

$$\|v - \pi_p v\|_{H^q(K)} \leq C_1 \frac{h_K^{s_K - q}}{p^{l_K - q}} \|\mathfrak{E}v\|_{H^{l_K}(\mathcal{K})}, \quad l_K \geq 0, \tag{4.24}$$

for $0 \leq q \leq l_K$, and

$$\|v - \pi_p v\|_{F_i} \leq \mathcal{C}_{ap}^{1/2}(p, K, F_i)|F_i|^{1/2} \frac{h_K^{s_K - d/2}}{p^{l_K - 1/2}} \|\mathfrak{E}v\|_{H^{l_K}(\mathcal{K})},\ l_K > d/2, \tag{4.25}$$

with

$$\mathcal{C}_{ap}(p, K, F_i) := C_2 \min\Big\{h_K^d \big(|F_i| \sup_{\mathbf{x}_i^0 \in K} \min_{\mathbf{x}\in F_i} (\mathbf{m}_i \cdot \mathbf{n})\big)^{-1}, p^{d-1}\Big\},$$

$s_K = \min\{p + 1, l_K\}$, and $C_1, C_2 > 0$ constants depending only on the shape-regularity of $\mathcal{K}$, q, l_K, on C_{diam} (from Assumption 4.28) and on the domain Ω.

Proof. Let $\Pi_p : H^l(\mathcal{K}) \to \mathcal{P}_p(\mathcal{K})$ be a known optimal *hp*-version approximation operator on simplices, see, e.g., [6, 7, 56]. We define $\pi_p : H^l(\mathcal{K}) \to \mathcal{P}_p(\mathcal{K})$ by $\pi_p v := \Pi_p(\mathfrak{E}v)$. To prove (4.24), we begin by observing that

$$\|v - \pi_p v\|_{H^q(K)} = \|\mathfrak{E}v - \Pi_p(\mathfrak{E}v)\|_{H^q(K)} \leq \|\mathfrak{E}v - \Pi_p(\mathfrak{E}v)\|_{H^q(\mathcal{K})}.$$

Thus, Assumption 4.28 and standard *hp*-approximation estimates on simplices (e.g. [6, 7, 56] yield the desired bound; we refer to the proof of [20, Lemma 3.7] for a similar argument for polytopic elements.

To prove (4.25), we use the trace inequality (4.3) with $\zeta = p$ to deduce

$$\|v - \pi_p v\|_{F_i}^2 \leq C \frac{h_K}{\sup\limits_{\mathbf{x}_i^0 \in K} \min\limits_{\mathbf{x} \in F_i} (\mathbf{m}_i \cdot \mathbf{n})} \frac{h_K^{2s_K - 1}}{p^{2l_K - 1}} \|\mathfrak{E}v\|_{H^{l_K}(\mathcal{K})}^2, \tag{4.26}$$

noting that $\max_{\mathbf{x} \in F_i} |\mathbf{m}_i|_2^2 \leq h_K^2$. On the other hand, we observe that

$$\|v - \pi_p v\|_{F_i}^2 \leq |F_i| \|v - \pi_p v\|_{L_\infty(K_{F_i})}^2 \leq \|\mathfrak{E}v - \Pi_p(\mathfrak{E}v)\|_{L_\infty(\mathcal{K})}^2.$$

Hence, employing a classical *hp*-approximation estimate for the maximum norm error from [6, 7], (cf. also [20, Lemma 20] we arrive at

$$\|v - \pi_p v\|_{F_i}^2 \leq C |F_i| \frac{h_K^{2s_K - d}}{p^{2l_K - d}} \|\mathfrak{E}v\|_{H^{l_K}(\mathcal{K})}^2, \tag{4.27}$$

for $l_K > d/2$. The result follows by taking the minimum between the bound in (4.26) and the bound in (4.27). □

Remark 4.32. We note the correspondence between $\mathcal{C}_{\mathrm{INV}}(p, K, F_i)$ from Lemma 4.21 and $\mathcal{C}_{ap}(p, K, F_i)$, in the typical case $h_K^d \sim |K|$. The key attribute of both expressions is that they remain bounded for degenerating $|F_i|$, allowing for the estimates (4.15) and (4.25) to remain finite as $|F_i| \to 0$.

Remark 4.33. If the constant $\mathcal{C}_{ap}(p, K, F_i)$ in (4.25) is taken with the first term, then the approximation result (4.25) holds with $l_K > 1/2$.

5. A priori error analysis

We are now ready to briefly discuss a priori error bounds for sufficiently smooth exact solutions, thereby generalizing the respective results presented in [20] to the case of curved polytopic meshes. The line of argument is similar to the case of straight polytopic meshes presented in detail in [20].

A crucial ingredient of the analysis for the proof of stability of the dG-EASE method is the precise definition of the discontinuity-penalization function σ appearing in the method (3.3). It is important to define σ sufficiently large for stability, while at the same time not substantially larger than what is required, as it could potentially cause loss of accuracy and/or conditioning issues. Additionally, following [20], we provide a stronger inf-sup stability result with respect to a 'steamline-diffusion'-type augmented norm, when the wind $\mathbf{b}$ is non-zero. The size of the 'steamline-diffusion' coefficient depends crucially on Lemma 4.23, whose constant provides information on the stabilization capabilities of the method.

The dG norm for which we seek to prove a priori error bounds is given by $|||v||| := \big(|||v|||_{\mathrm{ar}}^2 + |||v|||_{\mathrm{d}}^2\big)^{1/2}$, where

$$|||v|||_{\mathrm{ar}}^2 = \|c_0 v\|^2 + \frac{1}{2} \sum_{K \in \mathcal{T}} \big(\|\sqrt{|\mathbf{b} \cdot \mathbf{n}|} \lfloor v \rfloor\|_{\partial_- K}^2 + \|\sqrt{|\mathbf{b} \cdot \mathbf{n}|} v\|_{\partial_+ K \cap \partial\Omega}^2 \big),$$

with c_0 given in (2.4), and $|||v|||_{\mathrm{d}}^2 = \|\sqrt{a} \nabla_{\mathcal{T}} v\|^2 + \|\sqrt{\sigma} [\![v]\!]\|_{\Gamma_{\mathrm{int}} \cup \partial\Omega_{\mathrm{D}}}^2$.

Definition 5.1. For a mesh $\mathcal{T}$, we define the set $\mathcal{F}_{\rm int}$ of *interfaces* $F \subset \Gamma_{\rm int}$ by

$$\mathcal{F}_{\rm int} := \{F \subset \Gamma_{\rm int} : \text{ there exist } K, K' \in \mathcal{T} \text{ with } F = \partial K \cap \partial K'\};$$

correspondingly, we set $\mathcal{F}_{\rm D} := \{F \subset \partial\Omega_{\rm D} : \text{ there exists } K \in \mathcal{T} \text{ with } F = \partial K \cap \partial\Omega_{\rm D}\}$. For notational compactness, we also define $\mathcal{F}_{\rm int,D} := \mathcal{F}_{\rm int} \cup \mathcal{F}_{\rm D}$. (Note that F may comprise of one or more faces of K, K'.) Moreover, each interface F may be contained in one or more F_i's of the elements K, K' as per Assumption 4.1. Thus, there exists a subset $\{F_i^K\}_{i\in I_F^K}$ with index set $I_F^K \subset \{1, \dots, n_K\}$, such that $F \subset \cup_{i\in I_K^F} F_i^K$ and, correspondingly, a set $I_F^{K'} \subset \{1, \dots, n_{K'}\}$ such that $F \subset \cup_{i\in I_{K'}^F} F_i^{K'}$.

For technical reasons (cf. [20] and the references therein), we shall make use of the following extensions $\tilde{B}_{\rm d} : (H^1(\Omega) + S_{\mathcal{T}}^{\mathbf{p}}) \times (H^1(\Omega) + S_{\mathcal{T}}^{\mathbf{p}}) \to \mathbb{R}$ and $\tilde{\ell} : (H^1(\Omega) + S_{\mathcal{T}}^{\mathbf{p}}) \to \mathbb{R}$ of the bilinear and linear forms (3.5) and (3.6), which are given replacing $\{\!\{a\nabla w\}\!\}$ and $\{\!\{a\nabla v\}\!\}$ with $\{\!\{a\Pi\nabla w\}\!\}$ and $\{\!\{a\Pi\nabla v\}\!\}$ in $B_{\rm d}$ and ℓ, respectively, where $\Pi : [L_2(\Omega)]^d \to [S_{\mathcal{T}}^{\mathbf{p}}]^d$ denotes the orthogonal L_2-projection operator onto the (vectorial) finite element space. Observe that $\tilde{B}_{\rm d}(w, v) = B_{\rm d}(w, v)$ and $\tilde{\ell}(v) = \ell(v)$ when $w, v \in S_{\mathcal{T}}^{\mathbf{p}}$. Similarly, we define $\tilde{B}(w, v) := \tilde{B}_{\rm d}(w, v) + B_{\rm ar}(w, v)$. Next, we discuss the coercivity and continuity of $\tilde{B}_{\rm d}$.

Lemma 5.2. *Let* (3.7) *hold and consider a mesh* $\mathcal{T}$ *satisfying Assumption 4.1. With the notation introduced in Definition 5.1, define the discontinuity-penalization function* $\sigma : \Gamma_{\rm int} \cup \partial\Omega_{\rm D} \to \mathbb{R}$ *for every interface* $F \in \mathcal{F}_{\rm int,D}$, $F = \partial K \cap \partial K'$, *by*

$$\sigma|_F := 2 \max_{\mathcal{K}\in\{K,K'\}} \Big\{ |I_F^{\mathcal{K}}| \max_{i\in I_F^{\mathcal{K}}} \big\{ \mathcal{C}_{\rm INV}(p_{\mathcal{K}}, \mathcal{K}, F_i^{\mathcal{K}}) |F_i^{\mathcal{K}}| \big\} \frac{\bar{a}_K (p_{\mathcal{K}} + 1)(p_{\mathcal{K}} + d)}{|\mathcal{K}|} \Big\}; \tag{5.1}$$

when $F \in \mathcal{F}_{\rm D}$ *we set* $K = K'$. *Then, for all* $w, v \in H^1(\Omega) + S_{\mathcal{T}}^{\mathbf{p}}$, *we have*

$$\tilde{B}_{\rm d}(w, w) \geq \frac{1}{2} |||w|||_{\rm d}^2 \quad \text{and} \quad \tilde{B}_{\rm d}(w, v) \leq 2 |||w|||_{\rm d} |||v|||_{\rm d}. \tag{5.2}$$

Proof. The idea of proof is standard and makes use of the trace inverse estimate developed above. The novel attribute here is the choice of σ which requires some care since the star-shapedness of each interface F may correspond to different boundary segments F_i in either side of the interface. To that end, for $w \in H^1(\Omega) + S_{\mathcal{T}}^{\mathbf{p}}$, we have $\tilde{B}_{\rm d}(w, w) \geq |||w|||_{\rm d}^2 - 2 \int_{\Gamma_{\rm int}\cup\partial\Omega_{\rm D}} \{\!\{a\Pi\nabla w\}\!\} \cdot [\![w]\!] \, {\rm d}s$.

Therefore, Lemma 4.21 and the stability of the orthogonal L_2-projection give

$$\|\Pi\sqrt{a}\nabla w\|_{F\cap\partial K}^2 \leq |I_F^K| \max_{i\in I_F^K} \big\{ \mathcal{C}_{\rm INV}(p_K, K, F_i^K) |F_i^K| \big\} \frac{(p_K + 1)(p_K + d)}{|K|} \|\sqrt{a}\nabla w\|_K^2,$$

Coercivity already follows by a Young's inequality. The proof of continuity is standard and, therefore, omitted for brevity. □

Remark 5.3. The stability of the dG-EASE method is guaranteed under extremely general mesh assumptions thanks to the judicious choice of the penalization function (5.1). As discussed also in Remark 4.22, the latter ultimately depends on the choice of subdivisions $\{F_i\}_{i=1}^{n_K}$ of ∂K appearing in Assumption 4.1. Of course, whenever possible, by simply following the recipe in Remark 4.2(ii), we can easily arrive at a practical value of the penalization function for general curved elements.

We shall additionally assume for simplicity of the presentation that

$$\mathbf{b} \cdot \nabla\xi \in S_{\mathcal{T}}^{\mathbf{p}}, \qquad \text{for all } \xi \in S_{\mathcal{T}}^{\mathbf{p}}, \tag{5.3}$$

as is a standard in this context, cf. [41] and also [20, Chapter 5]. Assumption (5.3) can be further relaxed at the expense of an additional mild suboptimality with respect to the polynomial degree $\mathbf{p}$; see [41, Remark 3.13] and Remark 5.8 below.

Theorem 5.4. *Let $\mathcal{T} = \{K\}$ a subdivision of $\Omega \subset \mathbb{R}^d$, consisting of, possibly curved, elements satisfying Assumptions 4.1, 4.3 and 4.28. Then, assuming that* (5.3) *holds and that the discontinuity-penalization function σ is given by* (5.1)*, there exists a constant $\Lambda_s > 0$, independent of h and of $\mathbf{p}$, such that:*

$$\inf_{w\in S^{\mathbf{p}}_{\mathcal{T}}\setminus\{0\}} \sup_{v\in S^{\mathbf{p}}_{\mathcal{T}}\setminus\{0\}} \frac{B(w,v)}{|||w|||_{\mathrm{s}}|||v|||_{\mathrm{s}}} \geq \Lambda_s, \tag{5.4}$$

with $|||v|||_{\mathrm{s}} := \Big(|||v|||^2 + \sum_{K\in\mathcal{T}} \lambda_K \|\mathbf{b}\cdot\nabla v\|^2_K\Big)^{1/2}$, *whereby*

$$\lambda_K := \frac{\min\left\{\dfrac{\rho_K}{\sqrt{\mathcal{C}^B_{\mathrm{INV}}(p_K,K)}}, \Big(\displaystyle\sum_{F\subset\partial K}\sum_{i\in I^K_F} \mathcal{C}_{\mathrm{INV}}(p_K,K,F^K_i)\frac{|F^K_i|}{|K|}\Big)^{-1}\right\}}{\max\{\|\mathbf{b}\|_{L_\infty(K)},\sigma_K\}(p_K+1)(p_K+d)},$$

for $K\in\mathcal{T}$, $p_K\geq 1$, $\sigma_K := \max\{\sigma^a_K,\sigma^b_K\}$, with $\sigma^a_K := \max_{F\subset\partial K}\sigma|_F$, and

$$\sigma^b_K := 2\max_{F\subset\partial K}\left\{\max_{\mathcal{K}\in\{K,K'\}}\left\{\sqrt{\mathcal{C}^B_{\mathrm{INV}}(p_{\mathcal{K}},\mathcal{K})}\frac{\bar{a}_{\mathcal{K}}(p_{\mathcal{K}}+1)(p_{\mathcal{K}}+d)}{\rho_{\mathcal{K}}}\right\}\right\}. \tag{5.5}$$

Proof. The proof follows in a completely analogous fashion to the proof of [20, Theorem 5.2] and is, therefore, largely omitted for brevity: the key idea is to set $v\equiv v(w) := w+\alpha w_s$, for $w\in S^{\mathbf{p}}_{\mathcal{T}}$ with $w_s|_K := \lambda_K\mathbf{b}\cdot\nabla w$, $K\in\mathcal{T}$, with $\mathbb{R}\ni\alpha>0$. Then, it is sufficient to prove that $|||v|||_{\mathrm{s}}\leq C^*|||w|||_{\mathrm{s}}$, and $B(w,v)\geq C_*|||w|||^2_{\mathrm{s}}$, and then to set $\Lambda_s = C_*/C^*$, for some $C_*, C^*>0$ constants independent of the discretization parameters. The last two conditions are proven by using the inverse estimates above along with a judicious use of α. □

Remark 5.5. Theorem 5.4 extends respective results on polytopic meshes from [18, 20], to general meshes consisting of polytopic and/or curved elements with arbitrary number of faces. Moreover, this choice removes a dependence of the inf-sup constant Λ_s on the inverse inequality constants C_{INV} and $\mathcal{C}^B_{\mathrm{INV}}$; cf. [18, 20].

Theorem 5.6. *Let $\mathcal{T} = \{K\}$ be a subdivision of $\Omega\subset\mathbb{R}^d$, consisting of general curved elements satisfying Assumptions 4.1, 4.3 and 4.28. Let also $\mathcal{T}^\sharp = \{\mathcal{K}\}$ an associated covering of $\mathcal{T}$ consisting of shape-regular simplices as per Definition 4.27. Assume that* (5.3) *holds. Assume that $u\in H^1(\Omega)$ the exact solution to* (2.1),(2.3)*, is such that $u|_K\in H^{l_K}(K)$, $l_K > 1+d/2$, for each $K\in\mathcal{T}$. Let $u_h\in S^{\mathbf{p}}_{\mathcal{T}}$, with $p_K\geq 1$, $K\in\mathcal{T}$, be the solution of* (3.3)*, with σ as in* (5.1)*. Then, we have*

$$|||u-u_h|||^2_{\mathrm{s}} \leq C\sum_{K\in\mathcal{T}}\frac{h_K^{2s_K}}{p_K^{2l_K}}\left(\mathcal{D}_K(F,\mathcal{C}_{ap},\lambda_K,p_K)+\mathcal{G}_K(F,\mathcal{C}_{\mathrm{INV}},\mathcal{C}_{ap},p_K)\right)\|\mathfrak{E}u\|^2_{H^{l_K}(\mathcal{K})},$$

with $s_K = \min\{p_K+1, l_K\}$,

$$\begin{aligned}\mathcal{D}_K(F,\mathcal{C}_{ap},\lambda_K,p_K) &= \|c_0\|^2_{L_\infty(K)} + \zeta_K^2 + \lambda_K^{-1} + \lambda_K\beta_K^2 p_K^2 h_K^{-2} + \bar{a}_K p_K^2 h_K^{-2}\\ &\quad + \beta_K h_K^{-d} p_K \sum_{F\subset\partial K}|I^K_F|\max_{i\in I^K_F}\{\mathcal{C}_{ap}(p_K,K,F^K_i)|F^K_i|\},\end{aligned}$$

and

$$\begin{aligned}\mathcal{G}_K(F,\mathcal{C}_{\mathrm{INV}},\mathcal{C}_{ap},p_K) &= \bar{a}_K^2 p_K^3 h_K^{-d-2}\sum_{F\subset\partial K\cap(\Gamma_{\mathrm{int}}\cup\partial\Omega_{\mathrm{D}})}\sigma^{-1}|I^K_F|\max_{i\in I^K_F}\{\mathcal{C}_{ap}(p_K,K,F^K_i)|F^K_i|\}\\ &\quad + \bar{a}_K^2 p_K^4 h_K^{-2}|K|^{-1}\sum_{F\subset\partial K\cap(\Gamma_{\mathrm{int}}\cup\partial\Omega_{\mathrm{D}})}\sigma^{-1}|I^K_F|\max_{i\in I^K_F}\{\mathcal{C}_{\mathrm{INV}}(p_K,K,F^K_i)|F^K_i|\}\\ &\quad + h_K^{-d}p_K\sum_{F\subset\partial K\cap(\Gamma_{\mathrm{int}}\cup\partial\Omega_{\mathrm{D}})}\sigma|I^K_F|\max_{i\in I^K_F}\{\mathcal{C}_{ap}(p_K,K,F^K_i)|F^K_i|\},\end{aligned}$$

$s_K = \min\{p_K + 1, l_K\}$, $\zeta_K := \|c/c_0\|_{L_\infty(K)}$, c_0 *is in* (2.4), $\beta_K := \|\mathbf{b}\|_{L_\infty(K)}$, *and* C *is a positive constant, which depends on the shape-regularity of* $\mathcal{T}^\sharp$, *but is independent of the discretization parameters.*

In the special case in which the coefficient a *is strictly positive definite a.e. in* Ω *while* $\mathbf{b} = \mathbf{0}$, *Assumption 4.3 can be removed from the hypotheses.*

Proof. The proof follows on very similar lines to the respective one for polytopic meshes and can be found in [20, Section 5.2]. □

The above *hp*–version *a priori* error bounds hold without any assumptions on the relative size of the faces F, $F \subset \partial K$, of a given curved element $K \in \mathcal{T}$. To aid the understanding of the rates of convergence resulting from the above results, we set $p_K = p \geq 1$, $h = \max_{K\in\mathcal{T}} h_K$, $s_K = s$, $s = \min\{p+1, l\}$, and $l > 1 + d/2$, and assume that $\mathrm{diam}(F) \sim h_K$, for all faces $F \subset \partial K$, $K \in \mathcal{T}$, so that $|F| \sim h_K^{(d-1)}$. Then, Theorem 5.6 reduces to

$$|||u - u_h|||_{\mathrm{d}} \leq C \frac{h^{s-1}}{p^{l-\frac{3}{2}}} \|u\|_{H^l(\Omega)},$$

i.e., it proves optimal convergence in h and suboptimal in p by $p^{1/2}$.

At the other end of the spectrum, consider the case of transport equation, i.e., when $a \equiv \mathbf{0}$. In this case, the dG norm $|||\cdot|||$ degenerates to $|||\cdot|||_{\mathrm{ar}}$; note that, then we have $\lambda_K = \mathcal{O}(h_K/p_K^2)$, and the *a priori* error bound in Theorem 5.6 reduces to

$$|||u - u_h|||_{\mathrm{ar}} \leq \frac{h^{s-\frac{1}{2}}}{p^{l-1}} \|u\|_{H^l(\Omega)}.$$

This bound is, again, optimal in h and suboptimal in p by $p^{1/2}$ and completely generalizes the error estimate derived in our previous work [18] to essentially arbitrarily-shaped meshes under the same assumption (5.3).

Remark 5.7. We remark on typical cases which result to simplified formulas for λ_K. Assuming $|K| \sim h_K^d$, $\rho_K \sim h_K$, and $|F| \sim h_K^{(d-1)}$ for an element $K \in \mathcal{T}$ and for its immediate neighbours, both constants $\mathcal{C}_{\mathrm{INV}}^B$ and $\mathcal{C}_{\mathrm{INV}}$ will be defined by the first term in the maxima in (4.16) and (4.19), respectively. Then, we deduce $\lambda_K \sim h_K/p_K^2$ for the important case of advection-dominated problems.

Remark 5.8. For general advection fields $\mathbf{b}$, the proof of the inf-sup condition needs to be modified by using a slightly different norm involving $\Pi(\mathbf{b}\cdot\nabla_{\mathcal{T}})$ instead of $(\mathbf{b}\cdot\nabla_{\mathcal{T}})$ in the s-norm, yielding an error bound which is optimal in h but suboptimal in p by $p^{3/2}$ for the purely hyperbolic problem. Of course, if we modify the method by including a streamline-diffusion stabilization term as done in [40], then an hp-optimal bound can be derived without enforcing (5.3).

6. Numerical examples

We test the dG-EASE approach through a series of numerical experiments using curved elements, ranging from basic domain approximation to highly complex element shapes arising from random element agglomeration of a fine background triangulation.

In the case of the agglomeration-constructed elements, the background (curved) triangulation is also used for the assembly step. In particular, the discontinuity-penalisation function $\sigma|_F$ is fixed following the recipe in (5.1) with the subdivisions $\{F_i\}_{i=1}^{n_K}$ of ∂K, $K \in \mathcal{T}$, appearing in Assumption 4.1, given by unions of faces of the background triangulation. Moreover, for simplicity the background triangulation is also used for integration, exploiting parallellization of the quadrature process [30], see also [20] for a more detailed discussion of implementation of such methods.

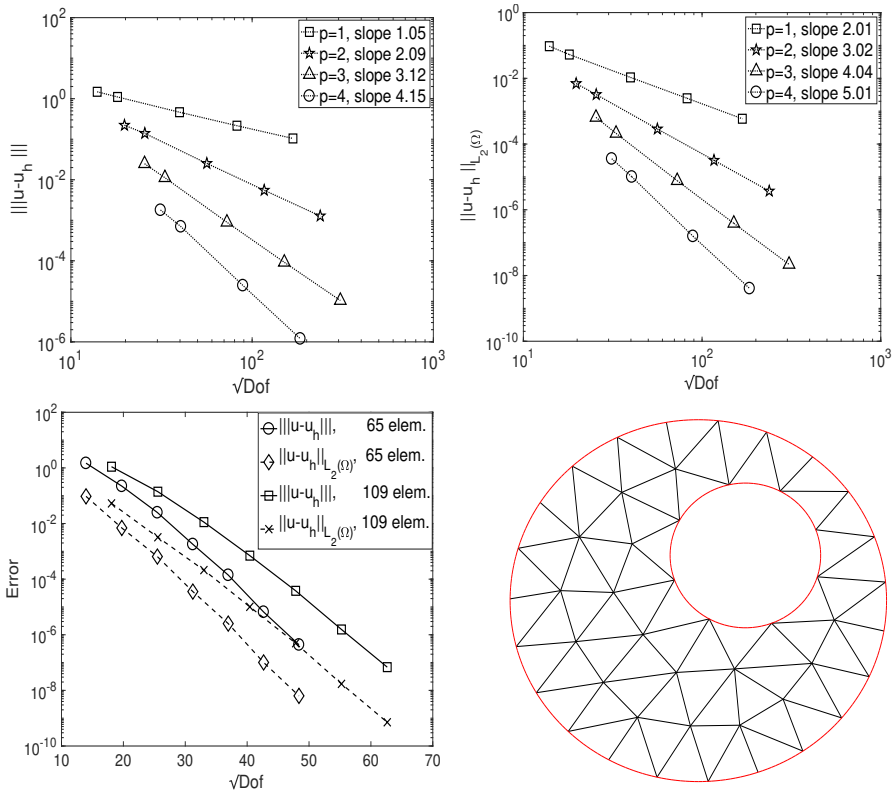

FIGURE 9. Example 1. Top: convergence history for $p = 1, 2, 3, 4$ in the $|||\cdot|||$ and $\|\cdot\|_\Omega$ norms against $\sqrt{\text{Dof}}$ for the curved triangular mesh exemplified in the bottom (right) plot with $65, 109, 527, 2266$, and 9411 elements, respectively. Bottom (left): Convergence history for $p = 1, 2, \ldots, 7$ in the $|||\cdot|||$ and $\|\cdot\|_\Omega$ against $\sqrt{\text{Dof}}$ for the meshes with 65 and 109 elements.

Nonetheless, very often it is possible to use substantially coarser subdivisions than the background triangulation the elements have been constructed from, e.g., a subdivision with one simplex per straight face.

For curved elements, the current implementation performs quadrature by a sufficiently fine sub-triangulation approximating the curved element, exactly as in the agglomerated-element case. We stress, however, that in this case the sub-triangulation is only used to generate the quadrature rules. These calculations are fully parallelizable : in [30] it is shown that quadrature cost becomes irrelevant if modern GPU architectures are used in the implementation of assembly. Of course, this is *not* the only possibility. For instance, domain-exact quadrature algorithms for many curved domains exist, see, e.g., [5] and the references therein for such algorithms.

6.1. **Example 1: curved elements.** We begin by testing the method on triangular elements with (non-parametric) curved faces. Specifically, we consider a two-dimensional diffusion problem with $a = I_{2\times 2}$, $I_{2\times 2}$ denoting the 2×2-identity matrix, $\mathbf{b} = (0, 0)^\top$, $c = 0$ and f so that $u(x_1, x_2) = \sin(\pi x_1)\sin(\pi x_2)$ in (2.1). We solve this problem on an irregular annular domain constructed as the unit disc centred at origin, with a circular hole centred at $(0.25, 0.25)$ and radius 0.4 removed; cf. Figure 9 for an illustration. We construct a sequence of domain-fitted curvilinear meshes as follows. First, using the mesh generator from [54], we construct a

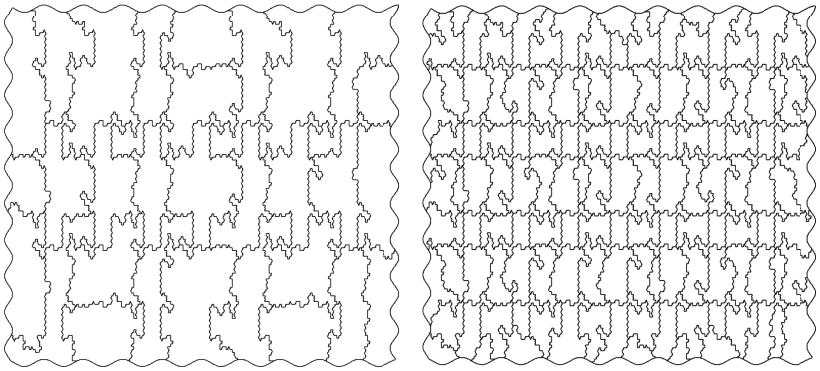

FIGURE 10. Example 2. The computational domain Ω and two meshes with 30 and 132 elements, respectively.

sequence of meshes approximating the domain Ω comprising of $65, 109, 527, 2266$, and 9411 quasi-uniform triangular elements, respectively. The 65-element mesh is shown in Figure 9. Then, exploiting the knowledge of the level-set function of $\partial\Omega$, elements containing straight faces approximating the curved boundary are marked. Finally, all marked elements are treated as curved triangular elements with two straight faces and one curved face described by the domain level-set function, thus *capturing the domain exactly.*

In Figure 9 (top row), we present the convergence history of $|||u - u_h|||$ and $\|u - u_h\|_\Omega$ against $\sqrt{\text{Dof}}$, with Dof the number of degrees of freedom on the aforementioned curvilinear meshes with $65, 109, 527, 2266$, and 9411 elements, for $p = 1, 2, 3, 4$, respectively. We clearly observe that, for each fixed p, all errors converge to zero at the optimal rates $\mathcal{O}(h^p)$ and $\mathcal{O}(h^{p+1})$, respectively, as the mesh size h tends to zero. In the two bottom plots in Figure 9, we also investigate the convergence history of the dG-EASE solution under p-refinement, using the two meshes with 65 and 109 curved elements, respectively, in linear-log scale. Here, we observe exponential convergence of all errors against $\sqrt{\text{Dof}}$.

6.2. **Example 2: convergence study.** We now investigate the convergence of dG-EASE on a highly complex mesh comprising of elements arising from agglomeration of a very fine background mesh, which also contains curved boundary-touching elements. Set $a = \epsilon I_{2\times 2}$ and $\epsilon = 0.01$, $\mathbf{b} = (1 - x_2, 1 - x_1)^\top$, $c = 2$ and f so that $u(x_1, x_2) = \sin(\pi x_1)\sin(\pi x_2)$ in (2.1) for $d = 2$, on a domain $\Omega \approx (0, 1)^2$ enclosed by a piecewise curved sinusoidal boundary; we refer to Figure 10 for an illustration. We impose non-homogeneous Dirichlet boundary conditions on $\partial\Omega$.

The mesh is constructed as follows. An initial curved mesh, fitted to the sinusoidal boundary via the level set approach described above, is subdivided into a very fine background subdivision consisting of approximately 500K sub-elements. The latter is, in turn, agglomerated into $30, 132, 555, 2151, 8337$ curved/polygonal elements using a standard mesh partitioning software. The parameters chosen in the partitioning software have been selected to yield a high-frequency 'sawtooth' vertical boundary for many of the agglomerated elements. We refer to Figure 10 for an illustration of the resulting meshes with 30 and 132 agglomerated elements.

In Figure 11, the convergence history for $p = 1, 2, 3, 4$ for the errors $|||u - u_h|||$ and $\|u - u_h\|_\Omega$ against $\sqrt{\text{Dof}}$ is presented for the aforementioned agglomerated meshes with $30, 132, 555, 2151, 8337$ elements. Here, we clearly observe that, for each fixed p, the dG- and $L_2(\Omega)$-norm errors converge to zero at the optimal rates $\mathcal{O}(h^p)$ and $\mathcal{O}(h^{p+1})$, respectively, as the mesh size h tends to zero. Further, we report also the

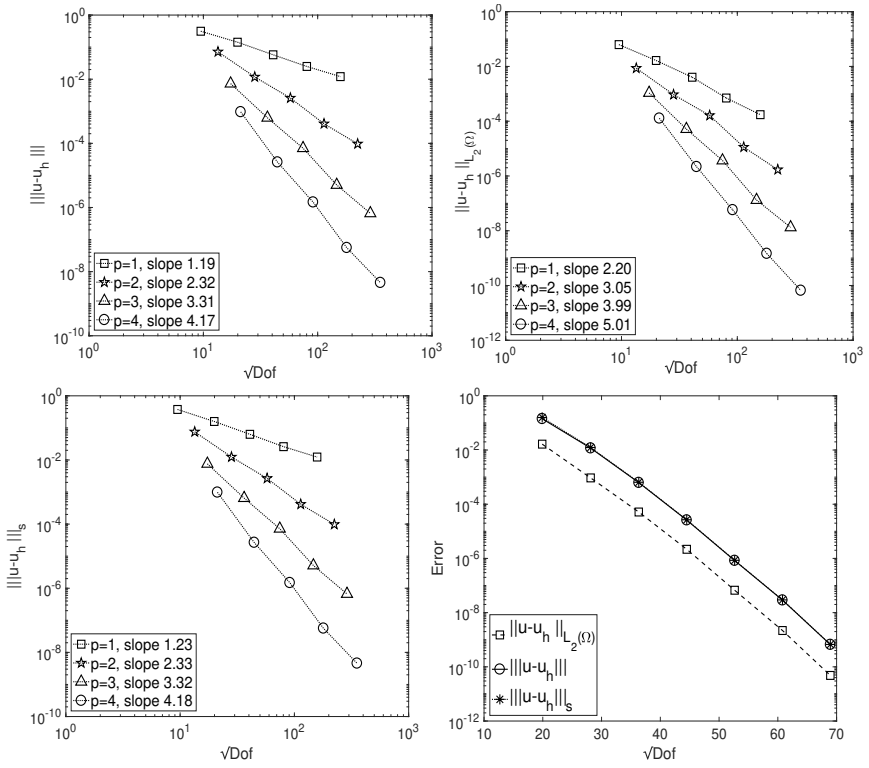

FIGURE 11. Example 2. Convergence history for $p = 1, 2, 3, 4$, in $|||\cdot|||$ (top-left), $\|\cdot\|_\Omega$(top-right) and $|||\cdot|||_{\mathrm{s}}$(bottom-left) against $\sqrt{\mathrm{Dof}}$ for the meshes exemplified in Figure 10 with $30, 132, 555, 2151, 8337$ elements, respectively. Bottom right: convergence history for $p = 1, 2, \ldots, 7$ in the three norms against $\sqrt{\mathrm{Dof}}$ for the mesh with 132 elements shown in Figure 10 (right).

error in the stronger 'streamline-diffusion' norm $|||u - u_h|||_{\mathrm{s}}$ in Figure 11; here we have chosen $\lambda_K = \mathcal{O}(\rho_K/p_K^2)$. For each fixed p, the errors converge to zero at the optimal rates $\mathcal{O}(h^p)$, as the mesh size h tends to zero.

Finally, in Figure 11 (bottom-right), we also investigate the convergence history of the dG-EASE solution under p-refinement, using the mesh with 132 elements shown in Figure 10(right). Here, we observe exponential convergence of the three norm errors against $\sqrt{\mathrm{Dof}}$. Interestingly, we observe that the difference between the errors $|||u - u_h|||$ and $|||u - u_h|||_{\mathrm{s}}$ is insignificant.

6.3. **Example 3: stability study.** We continue by assessing the stability of the dG-EASE method for convection-diffusion problems in the presence of unresolved lower-dimensional sharp solution layers. To this end, for $d = 2$, we set $a = \epsilon I_{2\times 2}$ and $\epsilon = 10^{-4}$, $\mathbf{b} = (1, 1)^\top$, $c = 0$ and $f = 1$ in (2.1). We solve this problem on a variant of the domain Ω from Example 2 above, in which circular internal pieces of the domain of various radii have been removed; we refer to Figure 12(left) for an illustration of the domain and sample mesh of essentially arbitrarily-shaped elements obtained using a completely analogous construction to that used in Example 2. We close the problem by prescribing homogeneous Dirichlet boundary conditions on $\partial\Omega$ (i.e., including the internal boundaries at the holes). We expect strong exponential boundary layers on the top and right portions of the curved boundary, as well as variable intensity layers at the outflow portions of the internal hole boundaries.

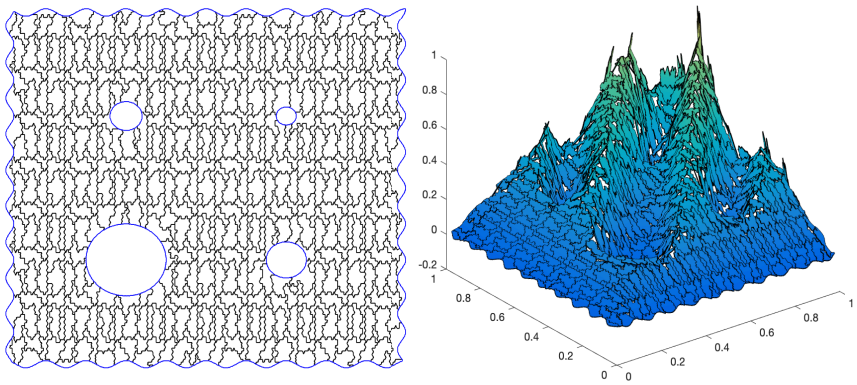

FIGURE 12. Example 3. Domain with holes and 531-element mesh (left). Corresponding solution obtained with $p = 1$ (right).

In Figure 12 (right), we provide the dG-EASE solution using $p = 1$ and the mesh of 531 elements shown on the left plot. This mesh is not fine enough to resolve the singularly perturbed behaviour in the vicinity of the outflow portions of the boundary. Nevertheless, the dG-EASE method provides a stable discretization with very localized, expected, oscillatory behaviour at the vicinity of the outflow boundary. The stable behaviour of dG-EASE with respect to the size of the Péclet number $Pe := \|\mathbf{b}\|/\epsilon$ is expected due to the upwind flux used in $B_{\text{ar}}(\cdot,\cdot)$; nonetheless, to the best of our knowledge, its performance in the context of elements with such geometrical shape generality has not been tested before in the literature. To highlight the behaviour of the method on different meshes, we report the dG-EASE solution, obtained with meshes composed of 129 and 2048 linear elements in Figure 13 (top). In both cases the mesh is not sufficiently fine to resolve the exponential boundary layer behaviour, while the finer mesh with 2048 linear elements sufficiently resolves the parabolic layers initiated at the holes.

Finally, we test the hyperbolic limit case by setting $\epsilon = 0$. The DG-EASE solution, shown in Figure 13 (bottom), remains stable and there is no oscillation around the outflow boundaries, as expected by a stabilised method.

6.4. **Example 4: changing type PDE across a curved interface.** To highlight a number of attractive features of the dG-EASE approach, we consider a coupled parabolic-hyperbolic partial differential equation, whose type changes across a sinusoidal interface Γ. Let $\Omega = \Omega_1 \cup \Omega_2 \cup \Gamma$ with

$$\begin{aligned}\Omega_1 &= \{(x_1,x_2) : -1 \le x_1 \le 1,\ A\sin(\omega\pi x_1) \le x_2 \le 1\},\\ \Omega_2 &= \{(x_1,x_2) : -1 \le x_1 \le 1,\ -1 \le x_2 \le A\sin(\omega\pi x_1)\},\end{aligned}$$

for $A, \omega > 0$ whose precise values will be given below; we refer to Figure 14 for an illustration. On this geometrical setting, we consider the problem:

$$\begin{cases} -x_1^2 u_{x_2x_2} + u_{x_1} + A\omega\pi\cos(\omega\pi x_1)u_{x_2} + u = 0, & \text{in } \Omega_1,\\ u_{x_1} + A\omega\pi\cos(\omega\pi x_1)u_{x_2} + u = 0, & \text{in } \Omega_2,\end{cases}$$

coupled with inflow and Dirichlet boundary conditions, so that the analytical solution is given by

$$u(x_1,x_2) = \begin{cases} \sin(\frac{\pi}{2}(1+x_2-A\sin(\omega\pi x_1)))\exp(-(x_1+\frac{\pi^2 x_1^3}{12})), & \text{in } \Omega_1,\\ \sin(\frac{\pi}{2}(1+x_2-A\sin(\omega\pi x_1)))\exp(-x_1). & \text{in } \Omega_2.\end{cases}$$

This problem is hyperbolic when $x_2 \le A\sin(\omega\pi x_1)$, $x_1 \in (-1,1)$, and parabolic otherwise. The normal flux of the exact solution is continuous across the interface

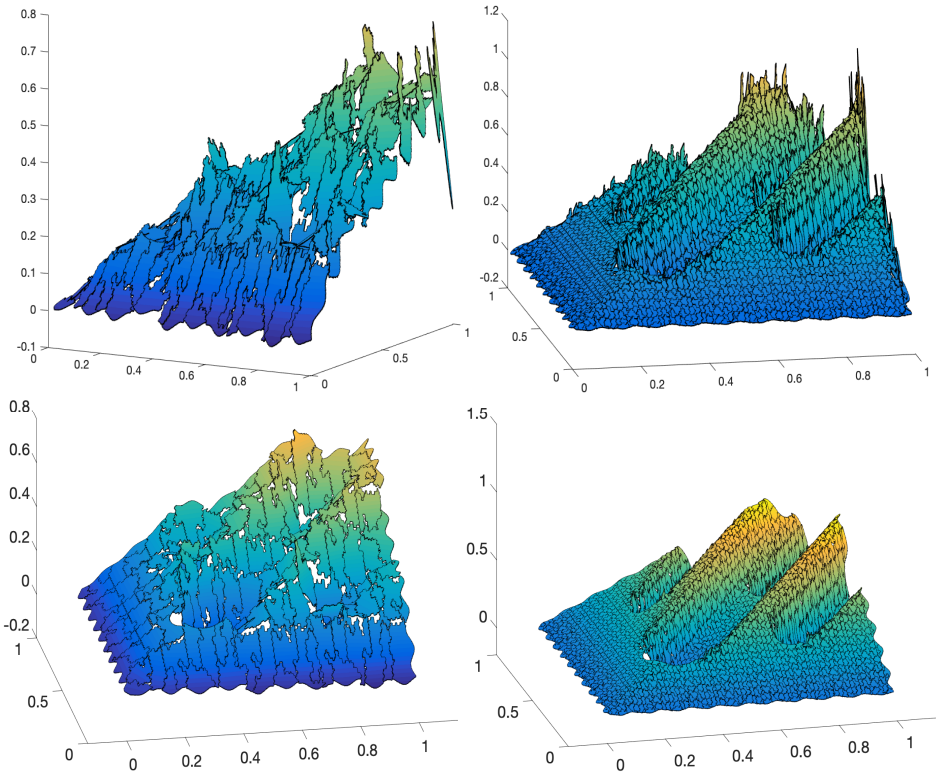

FIGURE 13. Example 3. Solutions computed for $\epsilon = 10^{-4}$ (top) and $\epsilon = 0$ (bottom) using linear elements. The meshes are composed of 129 (left) and 2048 (right) elements with compex shapes.

Γ with equation $x_2 = A\sin(\omega\pi x_1)$, while the solution itself has a discontinuity across the interface. This problem is a variant of an example from [33, 18]. As such, there is no discontinuity penalisation imposed at the interface Γ. Moreover, we point out that $\mathbf{b}\cdot\mathbf{n} = 0$ at the interface in this example.

Our aim is to highlight the performance of dG-EASE of arbitrary order, when the mesh is fitted with respect to a discontinuity of the exact solution. To that end, we focus on p-version convergence, using 64 rectangular elements with curved faces exactly fitting the interface; we refer to Figure 14 for an illustration with $A = 0.025$, and $\omega = 8$ and $\omega = 16$, respectively.

Interestingly, the mesh is *not* aligned with the inflow and outflow parts of the boundary $\partial\Omega$. This is due to the oscillating coefficient of the first order term. In Figure 14, the inflow parts of the boundary are marked in red; on these parts, inflow boundary conditions are imposed. Correspondingly, this pattern continues in the internal element faces in which the inflow parts of $\partial_- K$ are also not aligned with the faces. As such, face integral terms in the dG method may be computed only on parts of a face of a rectangular element. Nonetheless, the method is able to cope unaltered with this complication. The quadrature is implemented in the composite fashion described in Example 1 above.

We begin by setting $A = 0.025$ and $\omega = 8$. In Figure 15, we record the convergence history of the dG-EASE solution under p-refinement, using the mesh shown in Figure 14 (left) and $p = 1, \ldots, 13$. Although the elements are perfectly aligned with the interface Γ, the mesh is still coarse: each element includes roughly one full oscillation of the wind $\mathbf{b}$. Still we observe exponential convergence of $|||u - u_h|||$ and

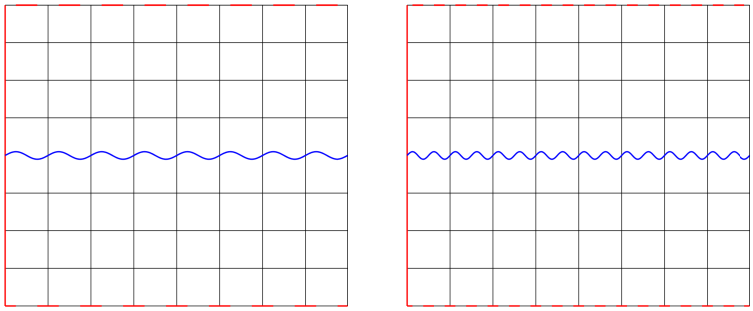

FIGURE 14. Example 4. The domain Ω with the 64 element mesh fitted to the interface (blue) for $A = 0.025$, and $\omega = 8$ (left) and $\omega = 16$ (right). The inflow parts of the boundary are drawn in red.

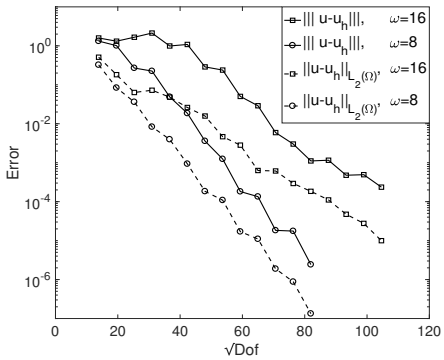

FIGURE 15. Example 4. Convergence history for $p = 1, 2, \ldots$ in the $|||\cdot|||$ and $\|\cdot\|_\Omega$ against $\sqrt{\text{Dof}}$ for the curved triangular mesh with 64 elements shown in Figure 14, for $A = 0.025$, $\omega = 8, 16$.

$\|u - u_h\|_\Omega$ errors against $\sqrt{\text{Dof}}$ under p-refinement. This result reinforces the claim that dG-EASE on perfectly aligned meshes with appropriate quadrature rules can lead to spectral accuracy. In contrast, if the mesh is not aligned exactly with the solution's discontinuity, the error is only expected to decay at an algebraic rate, according to standard best approximation results.

Next, we set $\omega = 16$ and we record the convergence history under p-refinement, for $p = 1, \ldots, 17$, for the fixed mesh from Figure 14(right). Here 64 elements constitute a very coarse mesh as, at the interface, there are now two full oscillations of the wind $\mathbf{b}$ per element. Again, we observe exponential convergence of $|||u - u_h|||$ and $\|u - u_h\|_\Omega$ against $\sqrt{\text{Dof}}$ under p-refinement, after an initial plateau for $p \leq 5$. This is expected as the dG-EASE approach is not designed as a multiscale framework. Nevertheless, for $p \geq 6$, exponential convergence is observed.

7. Acknowledgements

We are grateful to the anonymous referees and to the editor for their constructive comments which helped to improve this work substantially. AC gratefully acknowledges support from the MRC (MR/T017988/1), ZD from IACM-FORTH, Greece, and EHG from The Leverhulme Trust (RPG-2015- 306). This work was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant" (Proj. no. 3270).

References

[1] S. Agmon, *Lectures on elliptic boundary value problems*, Prepared for publication by B. Frank Jones, Jr. with the assistance of George W. Batten, Jr. Van Nostrand Mathematical Studies, No. 2, D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London, 1965.

[2] P. F. Antonietti, P. Houston, X. Hu, M. Sarti, and M. Verani, *Multigrid algorithms for hp-version interior penalty discontinuous Galerkin methods on polygonal and polyhedral meshes*, Calcolo, 54 (2017), pp. 1169–1198.

[3] P. F. Antonietti, P. Houston, G. Pennesi, and E. Süli, *An agglomeration-based massively parallel non-overlapping additive schwarz preconditioner for high-order discontinuous galerkin methods on polytopic grids*, Math. Comp., (2020).

[4] D. Arnold, F. Brezzi, B. Cockburn, and L. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2001), pp. 1749–1779.

[5] E. Artioli, A. Sommariva, and M. Vianello, *Algebraic cubature on polygonal elements with a circular edge*, Comput. Math. Appl., 79 (2020), pp. 2057–2066.

[6] I. Babuška and M. Suri, *The h-p version of the finite element method with quasi-uniform meshes*, RAIRO Modél. Math. Anal. Numér., 21 (1987), pp. 199–238.

[7] ———, *The optimal convergence rate of the p-version of the finite element method*, SIAM J. Numer. Anal., 24 (1987), pp. 750–776.

[8] I. Babuška, *The finite element method for elliptic equations with discontinuous coefficients*, Computing (Arch. Elektron. Rechnen), 5 (1970), pp. 207–213.

[9] J. W. Barrett and C. M. Elliott, *Fitted and unfitted finite-element methods for elliptic equations with smooth interfaces*, IMA J. Numer. Anal., 7 (1987), pp. 283–300.

[10] P. Bastian and C. Engwer, *An unfitted finite element method using discontinuous Galerkin*, Internat. J. Numer. Methods Engrg., 79 (2009), pp. 1557–1576.

[11] L. Beirão da Veiga, K. Lipnikov, and G. Manzini, *The mimetic finite difference method for elliptic problems*, vol. 11 of MS&A. Modeling, Simulation and Applications, Springer, Cham, 2014.

[12] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. Marini, and A. Russo, *Basic principles of virtual element methods*, Math. Models Methods Appl. Sci., 23 (2013), pp. 199–214.

[13] S. C. Brenner and L.-Y. Sung, *Virtual element methods on meshes with small edges or faces*, Math. Models Methods Appl. Sci., 28 (2018), pp. 1291–1336.

[14] E. Burman, S. Claus, P. Hansbo, M. G. Larson, and A. Massing, *CutFEM: discretizing geometry and partial differential equations*, Internat. J. Numer. Methods Engrg., 104 (2015), pp. 472–501.

[15] E. Burman and A. Ern, *An unfitted hybrid high-order method for elliptic interface problems*, SIAM J. Numer. Anal., 56 (2018), pp. 1525–1546.

[16] E. Burman and P. Hansbo, *Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method*, Appl. Numer. Math., 62 (2012), pp. 328–341.

[17] E. Burman and P. Zunino, *A domain decomposition method based on weighted interior penalties for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., 44 (2006), pp. 1612–1638.

[18] A. Cangiani, Z. Dong, E. Georgoulis, and P. Houston, *hp–Version discontinuous Galerkin methods for advection–diffusion–reaction problems on polytopic meshes*, ESAIM: M2AN, 50 (2016), pp. 699–725.

[19] A. Cangiani, Z. Dong, and E. H. Georgoulis, *hp-version space-time discontinuous Galerkin methods for parabolic problems on prismatic meshes*, SIAM J. Sci. Comput., 39 (2017), pp. A1251–A1279.

[20] A. Cangiani, Z. Dong, E. H. Georgoulis, and P. Houston, *hp-version discontinuous Galerkin methods on polygonal and polyhedral meshes*, SpringerBriefs in Mathematics, Springer, Cham, 2017.

[21] A. Cangiani, E. Georgoulis, and P. Houston, *hp–Version discontinuous Galerkin methods on polygonal and polyhedral meshes*, Math. Models Methods Appl. Sci., 24 (2014), pp. 2009–2041.

[22] A. Cangiani, E. Georgoulis, and M. Jensen, *Discontinuous Galerkin methods for mass transfer through semipermeable membranes*, SIAM J. Numer. Anal., 51 (2013), pp. 2911–2934.

[23] A. Cangiani, E. H. Georgoulis, and Y. A. Sabawi, *Adaptive discontinuous Galerkin methods for elliptic interface problems*, Math. Comp., 87 (2018), pp. 2675–2707.

[24] C. Carstensen and S. A. Sauter, *A posteriori error analysis for elliptic PDEs on domains with complicated structures*, Numer. Math., 96 (2004), pp. 691–721.

[25] L. Chen, H. Wei, and M. Wen, *An interface-fitted mesh generator and virtual element methods for elliptic interface problems*, J. Comput. Phys., 334 (2017), pp. 327–348.
[26] B. Cockburn, J. Gopalakrishnan, and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1319–1365.
[27] D. A. Di Pietro and A. Ern, *A hybrid high-order locking-free method for linear elasticity on general meshes*, Comput. Methods Appl. Mech. Engrg., 283 (2015), pp. 1–21.
[28] D. A. Di Pietro, A. Ern, and J.-L. Guermond, *Discontinuous Galerkin methods for anisotropic semidefinite diffusion with advection*, SIAM J. Numer. Anal., 46 (2008), pp. 805–831.
[29] Z. Dong, *Discontinuous galerkin methods for the biharmonic problem on polygonal and polyhedral meshes*, Int. J. Numer. Anal. Model., 16 (2019), pp. 825–846.
[30] Z. Dong, E. Georgoulis, and T. Kappas, *GPU-accelerated discontinuous Galerkin methods on polygonal and polyhedral meshes*, In preparation, (2020).
[31] A. Ern, A. F. Stephansen, and P. Zunino, *A discontinuous Galerkin method with weighted averages for advection-diffusion equations with locally small and anisotropic diffusivity*, IMA J. Numer. Anal., 29 (2009), pp. 235–256.
[32] T.-P. Fries and T. Belytschko, *The extended/generalized finite element method: an overview of the method and its applications*, Internat. J. Numer. Methods Engrg., 84 (2010), pp. 253–304.
[33] E. Georgoulis, *Discontinuous Galerkin methods on shape-regular and anisotropic meshes*, D.Phil. Thesis, University of Oxford, (2003).
[34] ———, *Inverse-type estimates on hp-finite element spaces and applications*, Math. Comp., 77 (2008), pp. 201–219.
[35] E. Georgoulis and A. Lasis, *A note on the design of hp-version interior penalty discontinuous Galerkin finite element methods for degenerate problems*, IMA J. Numer. Anal., 26 (2006), pp. 381–390.
[36] Y. Gong, B. Li, and Z. Li, *Immersed-interface finite-element methods for elliptic interface problems with nonhomogeneous jump conditions*, SIAM J. Numer. Anal., 46 (2007/08), pp. 472–495.
[37] R. Guo and T. Lin, *A group of immersed finite-element spaces for elliptic interface problems*, IMA J. Numer. Anal., 39 (2019), pp. 482–511.
[38] A. Hansbo and P. Hansbo, *An unfitted finite element method, based on Nitsche's method, for elliptic interface problems*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 5537–5552.
[39] P. Hennig, M. Kästner, P. Morgenstern, and D. Peterseim, *Adaptive mesh refinement strategies in isogeometric analysis—a computational comparison*, Comput. Methods Appl. Mech. Engrg., 316 (2017), pp. 424–448.
[40] P. Houston, C. Schwab, and E. Süli, *Stabilized hp-finite element methods for first-order hyperbolic problems*, SIAM J. Numer. Anal., 37 (2000), pp. 1618–1643.
[41] ———, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., 39 (2002), pp. 2133–2163.
[42] P. Houston and E. Süli, *Stabilised hp–finite element approximation of partial differential equations with nonnegative characteristic form*, Computing, 66 (2001), pp. 99–119.
[43] P. Huang, H. Wu, and Y. Xiao, *An unfitted interface penalty finite element method for elliptic interface problems*, Comput. Methods Appl. Mech. Engrg., 323 (2017), pp. 439–460.
[44] A. Johansson and M. G. Larson, *A high order discontinuous Galerkin Nitsche method for elliptic problems with fictitious boundary*, Numer. Math., 123 (2013), pp. 607–628.
[45] A. Kroó, *On Bernstein-Markov-type inequalities for multivariate polynomials in L_q-norm*, J. Approx. Theory, 159 (2009), pp. 85–96.
[46] ———, *Sharp L_p Markov type inequality for cuspidal domains in $\mathbb{R}^d$*, J. Approx. Theory, 250 (2020), pp. 105336, 6.
[47] A. Lozinski, *A primal discontinuous Galerkin method with static condensation on very general meshes*, Numer. Math., 143 (2019), pp. 583–604.
[48] R. Massjung, *An unfitted discontinuous Galerkin method applied to elliptic interface problems*, SIAM J. Numer. Anal., 50 (2012), pp. 3134–3162.
[49] J. M. Melenk, *hp-finite element methods for singular perturbations*, vol. 1796 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2002.
[50] V. Murti and S. Valliappan, *Numerical inverse isoparametric mapping in remeshing and nodal quantity contouring*, Comput. Struct., 22 (1986), pp. 1011–1021.
[51] V. Murti, Y. Wang, and S. Valliappan, *Numerical inverse isoparametric mapping in 3d fem*, Comput. Struct., 29 (1988), pp. 611–622.

[52] O. Oleinik and E. Radkevič, *Second Order Equations with Nonnegative Characteristic Form*, American Mathematical Society, 1973.
[53] S. Osher and R. Fedkiw, *Level set methods and dynamic implicit surfaces*, vol. 153 of Applied Mathematical Sciences, Springer-Verlag, New York, 2003.
[54] P. Persson and G. Strang, *A simple mesh generator in MATLAB*, SIAM Rev., 46 (2004), pp. 329–345.
[55] S. Sauter and R. Warnke, *Extension operators and approximation on domains containing small geometric details*, East-West J. Numer. Math., 7 (1999), pp. 61–77.
[56] C. Schwab, *p– and hp–Finite element methods: Theory and applications in solid and fluid mechanics*, Oxford University Press: Numerical mathematics and scientific computation, 1998.
[57] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton, University Press, Princeton, N.J., 1970.
[58] N. Sukumar and A. Tabarraei, *Conforming polygonal finite elements*, Internat. J. Numer. Methods Engrg., 61 (2004), pp. 2045–2066.
[59] T. Warburton and J. S. Hesthaven, *On the constants in hp-finite element trace inverse inequalities*, Comput. Methods Appl. Mech. Engrg., 192 (2003), pp. 2765–2773.
[60] H. Wu and Y. Xiao, *An unfitted hp-interface penalty finite element method for elliptic interface problems*, J. Comput. Math., 37 (2019), pp. 316–339.

SISSA, via Bonomea 265, 34136 Trieste, Italy
E-mail address: Andrea.Cangiani@sissa.it

1) Inria, 2 rue Simone Iff, 75589 Paris, France, and 2) CERMICS, Ecole des Ponts, 77455 Marne-la-Vallée 2, France
E-mail address: Zhaonan.Dong@inria.fr

1) School of Mathematics and Actuarial Science, University of Leicester, University Road, Leicester, LE1 7RH, United Kingdom, 2) Department of Mathematics, School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Zografou 15780, Greece, and 3) IACM-FORTH, Crete, Greece
E-mail address: Emmanuil.Georgoulis@le.ac.uk