SIAM J. NUMER. ANAL.
Vol. 57, No. 3, pp. 1058–1081

ANISOTROPIC ERROR ESTIMATES OF THE LINEAR NONCONFORMING VIRTUAL ELEMENT METHODS*

SHUHAO CAO† AND LONG CHEN†

Abstract. A refined a priori error analysis of the lowest-order (linear) nonconforming virtual element method (VEM) for approximating a model Poisson problem is developed in both 2D and 3D. A set of new geometric assumptions is proposed on the shape regularity of polytopal meshes. A new error equation for the lowest-order (linear) nonconforming VEM is derived for any choice of stabilization, and a new stabilization using a projection on an extended element patch is introduced for the error analysis on anisotropic elements.

Key words. virtual element methods, polytopal finite elements, anisotropic error analysis, nonconforming method

AMS subject classifications. 65N12, 65N15, 65N30, 46E35

DOI. 10.1137/18M1196455

1. Introduction.

In this paper, we develop a modified nonconforming virtual element method (VEM), together with a new way to perform the a priori error analysis for a model Poisson equation. The new analysis incorporates several new geometry assumptions on polytopal partitions in both 2D and 3D.

To approximate multiphysics problems involving complex geometrical features using the finite element method (FEM) in 2D and 3D, how to encode this geometric information into the discretization is a challenge. In a specific problem's interest, common practices include either to generate a body/interface-fitted mesh by cutting a shape-regular background mesh or to build cut-aware approximation spaces/variational forms (stencils) on the unfitted background mesh. Some notable methods utilizing the latter idea include eXtended FEM (see, e.g., [29, 39]), fictitious domain FEM [31], cut FEM [18], and immersed FEM [33].

One resolution combining the advantages of both approaches in 3D was proposed in [22] by using polyhedral meshes rather than the tetrahedral ones. It avoids manually tweaking problematic tetrahedra like slivers with four vertices nearly coplanar, which is usually an unavoidable problem in generating a body-fitted mesh from a background mesh, especially when the mesh is fine.

Since arbitrary-shaped polygons or polyhedra are now introduced into the partition, it requires that the underlying FEMs can handle these kinds of general meshes. There are several classes of modifications of classical numerical methods to work on the polytopal meshes, including mimetic finite difference (MFD) methods [15, 9], generalized barycentric coordinates [30], compatible discrete operator schemes [13], composite/agglomerated discontinuous Galerkin finite element methods (DGFEMs) [2], hybridizable discontinuous Galerkin (HDG) methods [21, 24], hybrid high-order (HHO) methods [28, 27], weak Galerkin (WG) methods [40, 36], discontinuous Petrov–Galerkin (PolyDPG) methods [3], etc. Among them, the VEM introduced in [5] proposed a universal framework for constructing approximation spaces and proving

*Received by the editors June 25, 2018; accepted for publication (in revised form) February 25, 2019; published electronically May 9, 2019.
http://www.siam.org/journals/sinum/57-3/M119645.html
Funding: This work was supported by National Science Foundation grant DMS-1418934.
†Department of Mathematics, University of California Irvine, Irvine, CA 92697 (scao@math.uci.edu, chenlong@math.uci.edu).

the optimal-order convergence on polytopal meshes. Until now, VEMs for elliptic problems have been developed with elaborated details (see, e.g., [1, 7, 16, 4, 19, 6]).

The nonconforming FEM for elliptic problems, better known as the Crouzeix–Raviart element, was introduced in [25]. It is nonconforming in the sense that the approximation polynomial space is not a subspace of the underlying Sobolev space corresponding to the continuous weak formulation. Its VEM counterpart was constructed in [4]; see also [34]. The degrees of freedom (DoFs) of a nonconforming VEM function on an element K are the natural dual to this function's values according to a Neumann boundary value problem on K, which are induced by the integration by parts. When a locally constructed stabilization term satisfies the patch test, the convergence in a broken H^1-seminorm is obtained through a systematized approach by showing the norm equivalence for the VEM functions between the broken Sobolev norm and the norm induced by the bilinear form [4].

Establishing the norm equivalence above requires geometric constraints on the shape regularity of the mesh. Almost all VEM error analyses to date are performed on star-shaped elements, and the mostly used assumptions are (1) every element K and every face $F \subset \partial K$ are star-shaped with the chunkiness parameter uniformly bounded above; and (2) no short edge/no small face, i.e., $h_F \eqsim h_K$ for every face $F \subset \partial K$. In the former condition, the so-called chunkiness parameter of a star-shaped domain E is the ratio of the diameter of E over the radius of the largest inscribed ball with respect to which E is star-shaped, which may become unbounded for anisotropic elements or anisotropic faces in 3D star-shaped elements.

Recently, some refined VEM error analyses (see [10, 14]) have removed the "no short edge" assumption in the 2D conforming VEM by introducing a new tangential derivative-type stabilization first proposed in [41]. In the 3D case [14], the removal of the "no small face" assumption comes at a price, in that the convergence constant depends on the log of the ratio of the longest edge and the shortest edge on a face of a polyhedral element, which also appears in the 2D analysis using the traditional DoF-type stabilization. This factor seems nonremovable due to the norm equivalence being used in these approaches, and it excludes anisotropic elements and/or isotropic elements with anisotropic faces with high aspect ratios in 3D (see, e.g., Figure 4). However, in a variety of numerical tests, some of which even use the traditional stabilization that is suboptimal in theory, VEM performs robustly regardless of these seemingly artificial geometric constraints in situations like random-control-points Voronoi meshes, irregular concave meshes, a polygon degenerating to a line, and interface/crack-fitted meshes (see [6, 8, 11, 12, 22, 26, 35]). Especially, anisotropic elements and/or elements with anisotropic faces pose no bottlenecks to the convergence of VEM numerically.

In an effort to partially explain the robustness of VEM regarding the shape regularity of the mesh, in [20], an a priori error analysis for the lowest-order conforming VEM is conducted based on a mesh dependent norm $|||\cdot|||$ induced by the bilinear form, which is weaker than an H^1-seminorm. The main instrument is an error equation similar to the ones used in the error analysis in discontinuous Galerkin (DG)-type methods, thus bypassing the norm equivalence. In this way, fewer geometric constraints are required than in the error analysis using the norm equivalence. However, results in [20] are restricted to 2D, and the anisotropic error analysis is restricted to a special class of elements cut from a shape-regular mesh. In particular, long edges in an anisotropic element are required to be paired in order to control the interpolation error in different directions. A precise quantitative characterization of such anisotropic meshes, on which the analysis can be applied, is not explicitly given in [20].

In this paper, we follow this approach and derive an error equation for the lowest-order nonconforming VEM. Thanks to the natural definition of DoFs, the nonconforming interpolation defined using DoFs brings no error into the error estimate in the sense that $|||u - u_I||| = 0$, compared with the error estimates of the conforming interpolant being proved using an intricate edge-pairing technique in [20]. As a result, under geometric conditions introduced in [20], the anisotropic error analysis can be extended to the lowest-order nonconforming VEM in both 2D and 3D.

The findings in this paper strengthen our opinion: one of the reasons why VEM is immune to badly shaped elements is that the approximation to the gradient of an H^1-function is handled by the projection of the gradient of a VEM function, not the exact gradient of it. On the other hand, the flexibility of the VEM framework allows us to modify the stabilization from the one used in [4], tailored for the anisotropic elements, in two ways: (1) the weight is changed from the size of each face, respectively, to the diameter of an extended element patch; (2) the stabilization stencil enlarges to this extended element patch, and its form remains the same with the original DoF-type integral, in which the penalization now computes the difference of the VEM functions and their projections onto this extended element patch, not the underlying anisotropic element. In this way, the anisotropic elements can be integrated into the analysis naturally using the tools improved from the results in [40, 32], and an optimal-order convergence can be proved in this mesh dependent norm $|||\cdot|||$. Our stabilization has a spirit similar to the so-called ghost penalty method introduced in [17] for fictitious domain methods.

When extending the geometric conditions in [20] from 2D to 3D in section 3, some commonly used tools in finite element analysis, including various trace inequalities and Poincaré inequalities, for simplices are revisited for polyhedron elements. The conditions that these inequalities hold serve as a motivation to propose a set of constraints as minimal as possible on the shapes of elements. In this regard, Assumptions **B**–**C** are proposed with more local geometric conditions than the star-shaped condition, which in our opinion is a more "global"-oriented condition for a certain element. Moreover, the hourglass condition in Assumption **C** allows the approximation on "nice" hourglass-shaped elements, which further relaxes a constraint in the conforming case in [20], in which vertices have to be artificially added to make hourglass-shaped elements isotropic.

As mentioned earlier, the way to deal with an anisotropic element is to assume one can embed this element into an isotropic extended element patch in Assumption **D**. However, the current analysis forbids the existence of a cube/square being cut into thin slabs, in which the number of cuts $\to \infty$ when $h \to 0$. From the standpoint of the implementation, the total number of anisotropic elements cannot make up a significant portion of all elements in practice, as the enlarged stencil for the modified stabilization makes the stiffness matrix denser.

This paper is organized as follows: In section 2, the linear nonconforming VEM together with our modification is introduced. Section 3 discusses the aforementioned set of new geometric assumptions in 2D and 3D. In section 4, we derive a new error equation and an a priori error bound for the linear nonconforming VEM. Last, in section 5, we study how to alter the assembling procedure in the implementation.

For convenience, $x \lesssim y$ and $z \gtrsim w$ are used to represent $x \leq c_1 y$ and $z \geq c_2 w$, respectively, and $a \eqsim b$ means $a \lesssim b$ and $a \gtrsim b$. The constants involved are independent of the mesh size h. When there exists a certain dependence of these relations on certain geometric properties, such a dependence shall be stated explicitly.

2. Nonconforming VEMs. In this section, we shall introduce the linear nonconforming virtual element space and the corresponding discretization of a model Poisson equation. In order to deal with anisotropic elements, we shall propose a new stabilization term.

Let Ω be a bounded polytopal domain in $\mathbb{R}^d$ $(d = 2, 3)$, and consider the model Poisson equation in the weak form with data $f \in L^2(\Omega)$: find $u \in H_0^1(\Omega)$ such that

$$a(u, v) := (\nabla u, \nabla v) = (f, v) \quad \forall v \in H_0^1(\Omega). \tag{2.1}$$

Provided with the mesh satisfying the assumptions to be discussed in section 3, the goal of this subsection is to build the following discretization using a bilinear form $a_h(\cdot, \cdot)$ in a VEM approximation space V_h on a given mesh $\mathcal{T}_h$, which approximates the original bilinear form $a(\cdot, \cdot)$:

$$\text{find } u_h \in V_h \text{ such that } a_h(u_h, v_h) = \langle f, v_h \rangle \quad \forall v_h \in V_h, \tag{2.2}$$

where $\langle f, v_h \rangle \approx (f, v_h)$ can be thought of as a numerical quadrature.

2.1. Notation. Throughout the paper, the standard notation $(\cdot, \cdot)_D$ is used to denote the L^2-inner product on a domain/hyperplane D, and the subscript is omitted when $D = \Omega$. For every geometrical object D and for every integer $k \geq 0$, $\mathbb{P}_k(D)$ denotes the set of polynomials of degree $\leq k$ on D. The average of an L^1-integrable function or vector field v over D, endowed with the usual Lebesgue measure, is denoted by $\overline{v}^D = |D|^{-1} \int_D v$, where $|D| = \text{meas}(D)$.

To approximate problem (2.1), first Ω is partitioned into a polytopal mesh $\mathcal{T}_h$, each polytopal element being either a simple polygon $(d = 2)$ or a simple polyhedron $(d = 3)$. The set of elements contained in a subset $D \subset \Omega$ is denoted by $\mathcal{T}_h(D) := \{K \in \mathcal{T}_h : K \subset \bar{D}\}$. $h := \max_{K \in \mathcal{T}_h} h_K$ stands for the mesh size, with $h_D := \operatorname{diam} D$ for any bounded geometric object D. Denote by $\text{conv}(D)$ the convex hull of D. The term "face" F is usually used to refer to the $(d-1)$-flat face of a d-dimensional polytope in this partition $(d = 2, 3)$. For the $d = 2$ case, a face refers to an edge unless being specifically stated otherwise. The set of all the faces in $\mathcal{T}_h$ is denoted by $\mathcal{F}_h$. The set of faces on the boundary of an element K is denoted by $\mathcal{F}_h(K)$, and $n_K := |\mathcal{F}_h(K)|$ is the number of faces on the boundary of K. More generally, $\mathcal{F}_h(D) := \{F \in \mathcal{F}_h : F \subset \bar{D}\}$ denotes faces restricted to a bounded domain D. With help from the context, $\boldsymbol{n}_F$ denotes the outward unit normal vector of face F with respect to the element K. On an interior face $F \in \mathcal{F}_h$ that is shared by two elements $K^{\pm}$, define the jump of any function v as $[\![v]\!]_F = v^- - v^+$ on F, where $v^{\pm} = \lim_{\epsilon \to 0} v(\boldsymbol{x} - \epsilon \boldsymbol{n}_F^{\pm})$, and $\boldsymbol{n}_F^{\pm}$ represents the outward unit normal vector with respect to $K^{\pm}$. On a boundary face $F \subset \partial\Omega$, $[\![v]\!]_F := v|_F$.

For a bounded Lipschitz domain D, $\|\cdot\|_{0,D}$ denotes the L^2-norm, and $|\cdot|_{s,D}$ is the $H^s(D)$-seminorm. When $D = \Omega$ is the whole domain, the subscript Ω will be omitted.

2.2. Nonconforming VEM spaces. The lowest-order, i.e., the linear nonconforming, VEM [4] is the main focus of this article. The linear nonconforming VEM has rich enough content to demonstrate anisotropic meshes' local impact on the a priori error analysis, and yet it is elegantly simple enough to be understood without many technicalities. Our main goal is to develop the tools for the linear nonconforming VEM to improve the anisotropic error analysis for the VEM.

The lowest-order nonconforming virtual element space V_h, restricted on an element K, can be defined as follows [4]:

$$V_h(K) := \left\{ v \in H^1(K) : \Delta v = 0 \text{ in } K, \nabla v \cdot \boldsymbol{n}\big|_F \in \mathbb{P}_0(F) \ \forall F \in \mathcal{F}_h(K) \right\}. \tag{2.3}$$

The set of DoFs for the local space $V_h(K)$ contains the average of $v_h \in V_h(K)$ on every face $F \in \mathcal{F}_h(K)$:

$$\chi_F(v_h) = \frac{1}{|F|}\int_F v_h \, \mathrm{d}S. \tag{2.4}$$

Denote this set of DoFs by $\mathcal{N}(K) = \{\chi_F, F \in \mathcal{F}_h(K)\}$ with cardinality $|\mathcal{N}(K)| = n_K$; then one can easily verify that $(K, V_h(K), \mathcal{N}(K))$ forms a finite element triple in the sense of Chapter 2.3 in [23] (see [4]).

The global nonconforming VEM space V_h can then be defined as

$$V_h = \left\{ v \in L^2(\Omega) : \ v\big|_K \in V_h(K) \ \forall K \in \mathcal{T}_h, \ \int_F [\![v]\!]_F \, \mathrm{d}S = 0 \ \forall F \in \mathcal{F}_h \right\}. \tag{2.5}$$

The canonical interpolation $v_I\big|_K \in V_h(K)$ in the nonconforming VEM local space of $v \in H^1(K)$ is defined using the DoFs:

$$\chi_F(v) = \chi_F(v_I) \quad \forall F \in \mathcal{F}_h(K). \tag{2.6}$$

The canonical interpolation $v_I \in V_h$ is then defined using the global DoFs:

$$\chi_F(v) = \chi_F(v_I) \quad \forall F \in \mathcal{F}_h. \tag{2.7}$$

2.3. Local projections. The shape functions in $V_h(K)$ do not have to be formed explicitly in assembling the stiffness matrix. Based on the construction in (2.3), locally on an element K, a certain shape function is the solution to a Neumann boundary value problem, the exact pointwise value of which is unknown. Instead, for $u_h, v_h \in V_h(K)$, some computable quantities based on the DoFs of u_h and v_h are used to compute $a_h(u_h, v_h)$, which approximates the original continuous bilinear form $a(u_h, v_h)$. We now explore what quantities can be computed explicitly using DoFs.

First, the L^2-projection $Q_F : v \mapsto Q_F v \in \mathbb{P}_0(F)$ for any $v \in L^1(F)$ to piecewise constant space on a face F is defined as

$$\big(v - Q_F v, q\big)_F = 0 \quad \forall q \in \mathbb{P}_0(F). \tag{2.8}$$

For a VEM function $v_h \in V_h(K)$, this projection can be directly derived from the DoFs (2.4) since $Q_F(v_h) = \chi_F(v_h)$ by definition. In contrast, the L^2-projection $Q_K : L^1(K) \to \mathbb{P}_0(K)$, defined as

$$\big(v - Q_K v, q\big)_K = 0 \quad \forall q \in \mathbb{P}_0(K), \tag{2.9}$$

is not computable for $v_h \in V_h(K)$ by using only the DoFs of v_h.

On an element K, we can also compute an elliptic projection to the linear polynomial space: for any $v \in H^1(K)$, $\Pi_K v \in \mathbb{P}_1(K)$ satisfies

$$(\nabla \Pi_K v, \nabla q)_K = (\nabla v, \nabla q)_K \quad \forall q \in \mathbb{P}_1(K). \tag{2.10}$$

By choosing $q = x_i$, $i = 1, \ldots, d$, one can easily verify that $\nabla \Pi_K v = Q_K(\nabla u)$. Namely, $\nabla \Pi_K v$ is the best constant approximation of ∇u in K.

As an H^1-semi-inner product is used in (2.10), $\Pi_K v$ is unique up to a constant. The constant kernel will be eliminated by the following constraint:

$$\int_{\partial K} \Pi_K v \, \mathrm{d}S = \int_{\partial K} v \, \mathrm{d}S = \sum_{F \in \mathcal{F}_h(K)} \chi_F(v)|F|. \tag{2.11}$$

Using integration by parts and the fact that $\Delta q = 0$, ∇q being constant for $q \in \mathbb{P}_1(K)$, the right-hand side of (2.10) can be written as

$$(\nabla v, \nabla q)_K = (v, \nabla q \cdot \boldsymbol{n})_{\partial K} = \sum_{F \in \mathcal{F}_h(K)} \nabla q \cdot \boldsymbol{n}_F \, \chi_F(v) |F|. \tag{2.12}$$

Thus for a VEM function $v_h \in V_h(K)$, $\Pi_K v_h$ can be computed by the DoFs of v_h.

The following lemma shows that Π_K mapping depends only on DoFs. In this regard, the elliptic projection Π_K works in a more natural way for nonconforming VEM local space, thanks to the choice of DoFs being the natural dual from the integration by parts.

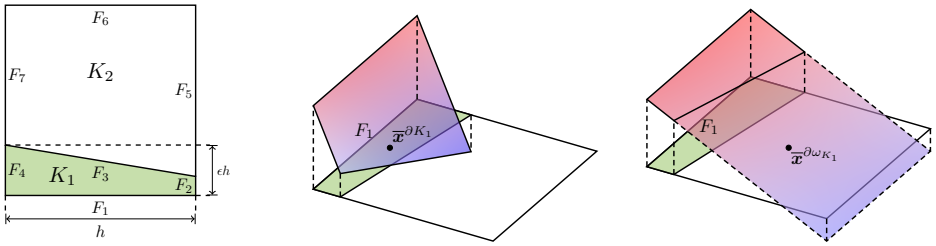
LEMMA 2.1. *For $v, w \in H^1(K)$, where $K \in \mathcal{T}_h$, if for all $F \in \mathcal{F}_h(K)$, $\chi_F(v) = \chi_F(w)$, then $\Pi_K v = \Pi_K w$.*

Proof. This is a direct consequence of the definition of Π_K in view of (2.11)–(2.12). ∎

To incorporate the possibility of the anisotropic analysis, we shall define an extended element patch containing K:

$$\omega_K := \bigcup\nolimits_{\alpha \in A} K_\alpha,$$

where $A = A(K)$ is an index set related to K such that $K \subseteq \omega_K$, $K_\alpha \in \mathcal{T}_h$ for all $\alpha \in A$, and ω_K is isotropic in the sense of Assumptions **A**–**C**, the existence of which shall be elaborated upon in section 3; for example, see Figure 1a. When K itself is isotropic, $\omega_K = K$.



(a) (b) (c)

FIG. 1. *An illustration of the extended element patch and the elliptic projections on it. Let $\epsilon \to 0$ as $h \to 0$.* (a) *K_1 is anisotropic, and $\omega_{K_1} = K_1 \cup K_2$ is isotropic.* (b) *$\Pi_{K_1} \phi_{F_1}$ in* (2.10) *has a sharp gradient.* (c) *$\Pi_{\omega_{K_1}} \phi_{F_1}$ in* (2.13) *has a smoother gradient over ω_{K_1} and is used only in the stabilization term on ∂K_1, not on ∂K_2.*

We define a discrete H^1-type projection on ω_K as follows: given a $v_h \in V_h$,

$$(\nabla \Pi_{\omega_K} v_h, \nabla q)_{\omega_K} = \sum_{K \in \mathcal{T}_h(\omega_K)} (\nabla v_h, \nabla q)_K \quad \forall q \in \mathbb{P}_1(\omega_K). \tag{2.13}$$

Notice here by the continuity condition in (2.5) that it is straightforward to verify that using integration by parts, for $q \in \mathbb{P}_1(\omega_K)$, on any $F \in \mathcal{F}_h(\omega_K)$, $\nabla q \cdot \boldsymbol{n}\big|_F \in \mathbb{P}_0(F)$,

and $\Delta q = 0$, we have

$$\sum_{K\in\mathcal{T}_h(\omega_K)} (\nabla v_h, \nabla q)_K = \sum_{K\in\mathcal{T}_h(\omega_K)} (v_h, \nabla q\cdot \boldsymbol{n})_{\partial K}$$
$$= \sum_{F\in\mathcal{F}_h(\partial\omega_K)} (v_h, \nabla q\cdot\boldsymbol{n})_F + \sum_{\substack{F\in\mathcal{F}_h(\omega_K),\\ F\not\subset\partial\omega_K}} \big(\llbracket v_h\rrbracket_F, \nabla q\cdot\boldsymbol{n}\big)_F = \sum_{F\in\mathcal{F}_h(\partial\omega_K)} (v_h, \nabla q\cdot\boldsymbol{n})_F,$$

whose right-hand side can be evaluated using the DoFs of v_h similar to (2.12).

When $K \neq \omega_K$, the constraint for Π_{ω_K}, as well as for Π_K (cf. (2.11)), is chosen as the average on the boundary of ω_K: for $v_h \in V_h$, and $\Pi_K v_h$ extended to be defined everywhere as the same polynomial in the closure of ω_K,

$$\int_{\partial\omega_K} \Pi_K v_h \,\mathrm{d}S = \int_{\partial\omega_K} \Pi_{\omega_K} v_h \,\mathrm{d}S = \int_{\partial\omega_K} v_h \,\mathrm{d}S, \tag{2.14}$$

which are both computable using DoFs of v_h.

In summary, although we do not have access to the pointwise value of $v_h \in V_h(K)$, we can find its average on each face and a linear polynomial $\Pi_K v_h$ inside K, whose gradient is the best piecewise constant approximation of the elementwise gradient of v_h. When needed, we can compute another linear polynomial $\Pi_{\omega_K} v_h$ on an extended patch ω_K (see, e.g., Figure 1c), the implementation details of which we refer the reader to section 5.

2.4. Discretization. As the H^1-projection, $\big(\nabla\Pi_K u_h, \nabla\Pi_K v_h\big)_K$ is a good approximation of $(\nabla u_h, \nabla v_h)_K$. However, $\big(\nabla\Pi_K u_h, \nabla\Pi_K v_h\big)_K$ alone will not lead to a stable method, as $|\ker(\Pi_K)| = \dim(V_h(K)) - \dim \mathbb{P}_1(K) \geq 0$ and the equality holds only if K is a simplex. The so-called stabilization term is needed to have a well-posed discretization. The principle of designing a stabilization is two-fold [5]:

1. *Consistency.* $S_K(u, v)$ should vanish when either u or v is in $\mathbb{P}_1(K)$. This can be ensured to use the slice operator $(\mathrm{I} - \Pi_K)$ in the inputs of $S_K(\cdot,\cdot)$ beforehand.
2. *Stability and continuity.* $S_K(\cdot,\cdot)$ is chosen so that the following norm equivalence holds:

$$a(v, v) \lesssim a_h(v, v) \lesssim a(v, v) \quad \forall v \in V_h. \tag{2.15}$$

The original bilinear form used in [4] for problem (2.2) is as follows: for $u_h, v_h \in V_h$,

$$a_h^{\mathrm{orig}}(u_h, v_h) := \sum_{K\in\mathcal{T}_h} \big(\nabla\Pi_K u_h, \nabla\Pi_K v_h\big)_K + \sum_{K\in\mathcal{T}_h} S_K^{\mathrm{orig}}\big((\mathrm{I} - \Pi_K)u_h, (\mathrm{I} - \Pi_K)v_h\big),$$

where the stabilization term $S_K^{\mathrm{orig}}(\cdot,\cdot)$ penalizes the difference between the VEM space and the polynomial projection using DoFs (2.4), while gluing the local spaces together using a weak continuity condition in (2.5): for $u_h, v_h \in V_h$,

$$S_K^{\mathrm{orig}}(u_h, v_h) := \sum_{F\in\mathcal{F}_h(K)} h_F^{d-2} \chi_F(u_h)\chi_F(v_h). \tag{2.16}$$

The dependence of constants in the norm equivalence (2.15) on the geometry of the element K is, however, not carefully studied in the literature. Especially on anisotropic

elements, constants hidden in (2.15) could be very large. In the 2D and 3D cases when every face $F \in \mathcal{F}_h(K)$ is shape-regular, we have the following relation:

$$S_K^{\text{orig}}(u_h, v_h) \eqsim \sum_{F\in\mathcal{F}_h(K)} h_F^{-1}\big(Q_F u_h, Q_F v_h\big)_F. \tag{2.17}$$

Inspired by this equivalence, we shall use a modified bilinear form: for $u_h, v_h \in V_h$,

$$a_h(u_h, v_h) := \sum_{K\in\mathcal{T}_h} \Big\{ \big(\nabla\Pi_K u_h, \nabla\Pi_K v_h\big)_K + \underbrace{S_K\big(u_h - \Pi_{\omega_K} u_h, v_h - \Pi_{\omega_K} v_h\big)}_{(\mathfrak{s})} \Big\}. \tag{2.18}$$

In (2.18), the stabilization on element K is

$$(\mathfrak{s}) := \sum_{F\in\mathcal{F}_h(K)} h_{\omega_K}^{-1}\big(Q_F(u_h - \Pi_{\omega_K} u_h), Q_F(v_h - \Pi_{\omega_K} v_h)\big)_F, \tag{2.19}$$

which penalizes the difference between a VEM function with its projection Π_{ω_K} on the boundary of K. To allow faces with small h_F, the weight is changed to $h_{\omega_K}^{-1}$ as well.

Now a nonconforming VEM discretization of (2.1) is as follows: for the bilinear form (2.18), find $u_h \in V_h$ such that

$$a_h(u_h, v_h) = \sum_{K\in\mathcal{T}_h} \big(f, \Pi_K v_h\big)_K \quad \forall v_h \in V_h. \tag{2.20}$$

In section 4, we shall derive a general error equation for the difference of the VEM approximation u_h to the interpolation u_I under the bilinear form induced norm and present an a priori error bound.

3. Geometric assumptions and inequalities. In this section, we explore some constraints to put on the meshes $\mathcal{T}_h$ in order that problem (2.20) yields a sensible a priori error estimate.

An element $K \in \mathcal{T}_h$ shall be categorized into either "isotropic" or "anisotropic" using some of the following assumptions on the geometry of the mesh. In the following assumptions, the uniformity of the constants is with respect to the mesh size $h \to 0$ in a family of meshes $\{\mathcal{T}_h\}$.

3.1. Isotropic elements. First, recall that n_K represents the number of faces as well as the number of DoFs in the element K. For both isotropic or anisotropic elements, the following assumption shall be fulfilled.

A. For $K \in \mathcal{T}_h$, the number of faces n_K is uniformly bounded.

Second, for a simple polygon/polyhedron that is not self-intersecting, a height l_F, measuring how far from F one can advance to the interior of K in its inward normal direction, determines to what degree of smoothness a function defined on F can be extended into the interior of K.

Without loss of generality, the presentation is based on the dimension $d = 3$ here, after which the case $d = 2$ follows naturally. For a given flat face $F \in \mathcal{F}_h(K)$, we choose a local Cartesian coordinate (ξ, η, τ) such that the face F is on the $\tau = 0$ plane. For any $x_F \in F$, $\boldsymbol{x}_F = \xi \boldsymbol{t}_{F,1} + \eta \boldsymbol{t}_{F,2}$, where $\boldsymbol{t}_{F,1}$ and $\boldsymbol{t}_{F,2}$ are two orthogonal unit vectors that span the hyperplane on which the face F lies.

The positive τ-direction is chosen such that it is the inward normal of F. Now define

$$\delta_F := \inf\Big\{\tau \in \mathbb{R}^+ :\ K \cap \big(F \times (\tau, +\infty)\big) = \varnothing\Big\}. \tag{3.1}$$

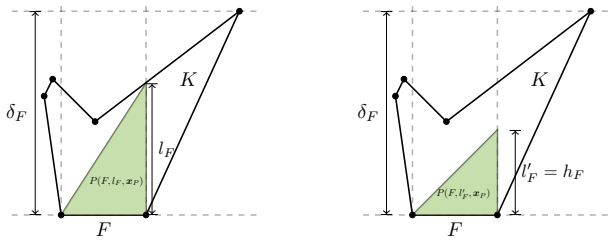


FIG. 2. (a) $l_F \geq \gamma_1 h_F$ *with* $\gamma_1 > 1$. (b) *A rescaled* $P(F, l'_F, \boldsymbol{x}_P)$ *with* $l'_F = h_F$.

As K is a simply polyhedral, $\delta_F > 0$, although it can be very small.

A pyramid with base F, apex $\boldsymbol{x}_P$, and height $l = \operatorname{dist}(\boldsymbol{x}_P, F)$ is defined as follows:

$$P(F, l, \boldsymbol{x}_P) := \{\boldsymbol{x} : \ \boldsymbol{x} = (1-t)\boldsymbol{x}_F + t\boldsymbol{x}_P, t \in (0,1), \boldsymbol{x}_F \in F\}. \tag{3.2}$$

Then an inward height l_F associated with face F can be defined as follows:

$$l_F := \sup\left\{l \in \mathbb{R}^+ : \exists\, P(F, l, \boldsymbol{x}_P) \subset K \cap \big(F \times (0, \delta_F]\big)\right\}. \tag{3.3}$$

Here the prism $F \times (0, \delta_F]$ is used to ensure the dihedral angles are bounded by $\pi/2$ between F and the side faces of the pyramid $P(F, l_F, \boldsymbol{x}_P)$.

When $d = 2$, as K is nondegenerate (there are no self-intersecting edges) and bounded, $0 < \delta_F < +\infty$ and $0 < l_F \leq \delta_F$ (see Figure 2a, for example). When $d = 3$, the existence of such a pyramid $P(F, l_F, \boldsymbol{x}_P)$ is unclear since F itself can be nonconvex. To be able to deal with such a case, we impose the following assumption.

B. (Height condition) There exists a constant $\gamma_1 > 0$, such that for all $F \in \mathcal{F}_h(K)$, it has a partition $F = \bigcup_{\beta \in B_1} F_\beta$ with $|B_1|$ uniformly bounded, such that each F_β satisfies the height condition $l_{F_\beta} \geq \gamma_1 h_F$ and consequently $l_F := \min_{\beta \in B_1} l_{F_\beta} \geq \gamma_1 h_F$.

In Figure 3a, the bottom edge satisfies the height condition **B** only when the decomposition argument is added in the assumption. In Figure 3c, for the whole front face F without decomposition, no such pyramid in (3.2) exists to yield a sensible l_F in (3.3) since there exist points outside K in the line connecting the apex of the pyramid with a point on F.

Without loss of generality, one can assume that the constant in Assumption **B** satisfies $0 < \gamma_1 \leq 1$ when Assumption **B** is used as a premise of a proposition in later sections. The reason is that, when **B** holds, one can always rescale the height l_F to $l'_F = \gamma'_1 h_F$, for any $0 < \gamma'_1 \leq \gamma_1$, while the new pyramid $P(F, l'_F, \boldsymbol{x}_P)$ is still in K. When $\gamma_1 > 1$, we can simply set $\gamma'_1 = 1$ to be the new γ_1. See the illustration in Figure 2b for an example in 2D.

Furthermore, in Figure 3a, it shows a "good" hourglass-shaped element. To avoid small hourglass-type bumps from an element (see, e.g., Figure 3b), the following assumption is imposed:

C. (Hourglass condition) For all $F \in \mathcal{F}_h(K)$, it has a partition $F = \bigcup_{\beta \in B_2} F_\beta$ with $|B_2|$ uniformly bounded, such that each F_β satisfies the hourglass condition. For all $\beta \in B_2$, there exists a convex subset $K_\beta \subseteq K$ with $h_{K_\beta} \eqsim h_K$, such that $P(F_\beta, l_{F_\beta}, \boldsymbol{x}_P) \subset K_\beta$.

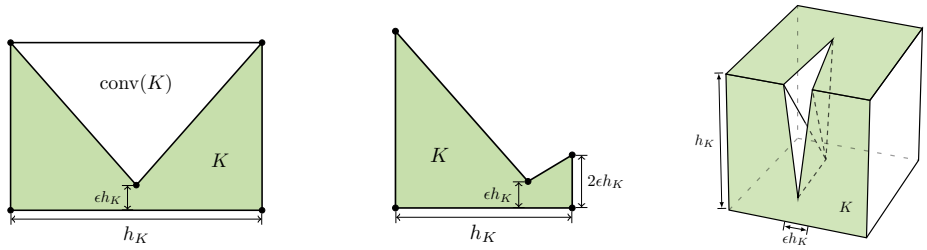


FIG. 3. $\epsilon \to 0$ *as* $h \to 0$. (a) K *has the hourglass shape and is not an isotropic element in the sense of the geometry assumptions in* [20]. *Yet this* K *is isotropic under Assumptions* **A**–**C**. (b) K *has a small hourglass-type bump which is ruled out by Assumption* **C**. (c) K *with a crack is isotropic, and it has two faces satisfying Assumptions* **A**–**C** *in the sense of decompositions.*

Now we say an element K is *isotropic* if Assumptions **A**–**C** hold for K, with the partitions $\{F_\beta\}_{\beta\in B_1} = \{F_\beta\}_{\beta\in B_2}$ for the same face F in Assumptions **B**–**C**. Otherwise, it is called *anisotropic*. As we mentioned earlier, isotropy and anisotropy can be formulated for 2D polygons using the height condition and hourglass condition for edges. A complication in 3D meshes is that for an isotropic polyhedron, we may have an anisotropic face or a tiny face. In both cases, $|F| \ll h_K^2$ (see Figure 4 for examples of polyhedral elements). Henceforth, when Assumptions **B** and/or **C** are met, we denote $P_F := P(F, l_F, \boldsymbol{x}_P)$, and whether the decomposition is used or not should be clear from the context.

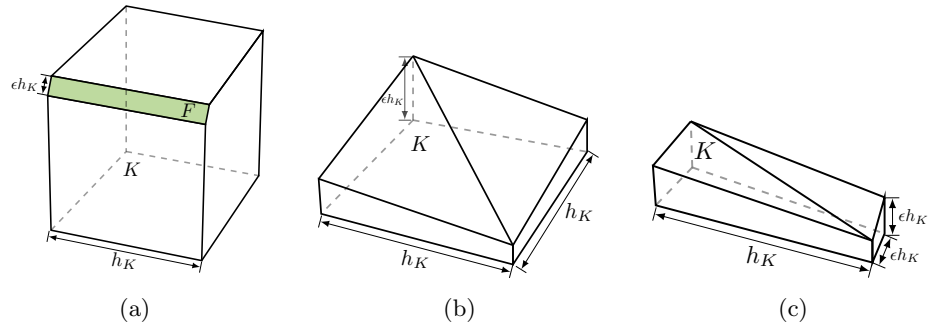


FIG. 4. $\epsilon \to 0$ *as* $h \to 0$. (a) K *is a cube without a prismatic slit; the marked face is anisotropic, yet the element is isotropic.* (b) K *itself is anisotropic; all four side faces are anisotropic.* (c) K *is anisotropic.*

LEMMA 3.1 (scale of the volume/area for isotropic elements). *If* K *is isotropic in the sense of Assumptions* **A**–**C**, *then* $|K| \eqsim h_K^d$.

Proof. Obviously, $|K| \lesssim h_K^d$ by the definition of diameter. It suffices to bound the volume $|K|$ below by h_K^d. We choose the face F with the largest area on ∂K. By Assumption **A**, $|F| \gtrsim |\partial K|$. With slight abuse of the order of the presentation, by the trace inequality with $v = 1$ in Lemma 3.3, we have $|F| = \|v\|_{0,F}^2 \lesssim h_K^{-1} \|v\|_{0,K}^2 = h_K^{-1}|K|$. Hence $|K| \gtrsim h_K|F| \gtrsim h_K|\partial K|$, and the lemma follows from the isoperimetric inequality $|\partial K| \gtrsim |K|^{(d-1)/d}$. ∎

3.2. Anisotropic elements. For anisotropic elements, by definition, there exist faces such that the height condition and/or hourglass condition are violated. The case $l_F \ll h_F$ can be caused by either the nonconvexity of F and K or the chunkiness parameter of K being large.

To be able to use the trace inequalities on a face in an anisotropic element, the following condition on this element K is proposed:

D. There exists an isotropic extended element patch ω_K consisting of elements in $\mathcal{T}_h$ such that the following hold:
1. $K \subseteq \omega_K$;
2. $h_{\omega_K} \leq \gamma_2 h$ with a uniform constant $\gamma_2 > 0$;
3. for all $F \in \mathcal{F}_h(K)$, F satisfies Assumptions **B** and **C** toward ω_K;
4. $n_{\omega_K} := \left|\{K' \in \mathcal{T}_h(\omega_K)\}\right|$ is uniformly bounded above.

By the construction of ω_K and the definition of an isotropic element, the height condition in Assumption **B** and the hourglass condition in Assumption **C** are met for every face $F \in \mathcal{F}_h(\partial\omega_K)$. With Assumption **D**, one can then lift a function defined on a boundary face $F \in \mathcal{F}_h(K)$ to the isotropic element ω_K.

3.3. Finite overlapping of convex hulls. For polytopal meshes, we impose the following conditions on the convex hull of K for isotropic elements or ω_K for anisotropic elements:

E. There exists a uniform constant $\gamma_3 > 0$ such that for each $K \in \mathcal{T}_h$,

$$|\{K' \in \mathcal{T}_h : \operatorname{conv}(\omega_{K'}) \cap \operatorname{conv}(\omega_K) \neq \varnothing\}| \leq \gamma_3.$$

It can be verified that Assumption **E** is ensured if for any vertex there are a uniformly bounded number of polytopal elements surrounding this vertex.

3.4. Trace inequalities. When using a trace inequality, one should be extremely careful, as the constant depends on the shape of the domain. In this subsection, we shall re-examine several trace inequalities with more explicit analyses on the geometric conditions.

LEMMA 3.2 (a trace inequality on a face in a polytopal element). *Suppose for the face F there exists a triangle/pyramid $P_F := P(F, l_F, \boldsymbol{x}_P) \subset K \cap \big(F \times (0, \delta_F]\big)$ with height l_F; then the following trace inequality holds for any $v \in H^1(K)$:*

$$\|v\|_{0,F} \lesssim l_F^{-1/2} \|v\|_{0,P_F} + \left(h_F l_F^{-1/2} + l_F^{1/2}\right) \|\nabla v\|_{0,P_F}. \tag{3.4}$$

Consequently, if, furthermore, the height condition **B** *is satisfied, it holds that for $P_F := \bigcup_{\beta \in B_1} P_{F_\beta}$,*

$$\|v\|_{0,F} \lesssim h_F^{-1/2} \|v\|_{0,P_F} + h_F^{1/2} \|\nabla v\|_{0,P_F}. \tag{3.5}$$

Proof. We first consider the case $l_F \eqsim h_F$ when $d = 3$. By Lemma A.3 in [40],

$$\|v\|_{0,F}^2 \lesssim h_F^{-1} \|v\|^2_{0,P_{\frac{1}{2}}(F,l_F,\boldsymbol{x}_P)} + h_F \|\nabla v\|^2_{0,P_{\frac{1}{2}}(F,l_F,\boldsymbol{x}_P)}, \tag{3.6}$$

where $P_{\frac{1}{2}}(F, l_F, \boldsymbol{x}_P) := \{\boldsymbol{x} : \ \boldsymbol{x} = (1-t)\boldsymbol{x}_F + t\boldsymbol{x}_P,\, t \in (0, 1/2),\, \boldsymbol{x}_F \in F\}$. The motivation to truncate the pyramid P_F to a prismatoid is that the Jacobian of the mapping from the prismatoid to the prism is bounded.

For a general case, without loss of generality, we assume $l_F \leq h_F$ since otherwise one can set $l_F = h_F$ first and (3.4) still holds by (3.6): we consider the mapping

$(\boldsymbol{x}_F, \tau) \mapsto (\boldsymbol{x}_F, \tau l_F/h_F)$, denote $P := P_{\frac{1}{2}}(F, l_F, \boldsymbol{x}_P)$, and let $\nabla_{\boldsymbol{x}_F}$ be the gradient taken with respect to (ξ, η) in F's local coordinate system. Then a straightforward change of variable computation yields

$$\|v\|_{0,F}^2 \lesssim l_F^{-1} \|v\|_{0,P}^2 + h_F^2 l_F^{-1} \|\nabla_{\boldsymbol{x}_F} v\|_{0,P}^2 + l_F \|\partial_\tau v\|_{0,P}^2 . \tag{3.7}$$

When $d = 2$, a similar scaling argument can be found in [20, Lemma 6.3] and the estimate there should be corrected to $\|v\|_{0,e}^2 \lesssim l_e^{-1} \|v\|_{0,P}^2 + h_e^2 l_e^{-1} \|\partial_x v\|_{0,P}^2 + l_e \|\partial_y v\|_{0,P}^2$ for an edge e. As a result, (3.4) holds.

When F satisfies Assumption **B**, $F = \bigcup_{\beta \in B_1} F_\beta$, each of F_β satisfies the height condition with disjoint pyramids $P(F_\beta, l_{F_\beta}, \boldsymbol{x}_\beta)$. One can rescale all l_{F_β} to be $l_F := \min_{\beta \in B_1} l_{F_\beta}$, and $P(F_\beta, l_F) \subset P(F_\beta, l_{F_\beta}) \subset K \cap \big(F_\beta \times (0, \delta_{F_\beta}]\big)$. Thus under Assumption **B**, $\|v\|_{0,F}^2 = \sum_{\beta \in B_1} \|v\|_{0,F_\beta}^2$ can be estimated by a simple summation of (3.7). ∎

As we mentioned before, even for an isotropic element, it may contain a face F with $h_F \ll h_K$, and thus the factor $h_F^{-1/2}$ in the trace inequality (3.5) may be uncontrollable. Next, we shall use the hourglass condition **C** to replace $h_F^{-1/2}$ by a smaller factor $h_K^{-1/2}$.

LEMMA 3.3 (a trace inequality on a face satisfying the height condition and the hourglass condition). *If a face $F \in \mathcal{F}_h(K)$ satisfies the height condition* **B** *and the hourglass condition* **C**, *then for any $v \in H^1(K)$ it holds that*

$$\|v\|_{0,F} \lesssim h_K^{-1/2} \|v\|_{0,K} + h_K^{1/2} \|\nabla v\|_{0,K} . \tag{3.8}$$

Proof. As the final inequality (3.8) can be trivially generalized from each F_β to $F = \bigcup_\beta F_\beta$ using the same argument as the one in Lemma 3.2, we consider only one face F_β in the decomposition, which shall be denoted by F subsequently in the proof. First, Assumption **B** implies the validity of the trace inequality (3.5). When Assumption **C** is met, let K_F be the convex subset of K containing P_F. Since (3.8) holds trivially if $h_F \eqsim h_K$, it suffices to consider the case when $h_F \ll h_K$. Due to the convexity of K_F, $\mathrm{conv}(P_F) \subset K_F$, without loss of generality we can assume that P_F is convex. Moreover, we recall that the rescaling argument facilitated by Assumption **B** allows us to set $h_{P_F} \eqsim h_F$. By $K_F \subseteq K$, it suffices to show that

$$h_F^{-1} \|v\|_{0,P_F}^2 \lesssim h_K^{-1} \|v\|_{0,K_F}^2 + h_K \|\nabla v\|_{0,K_F}^2 . \tag{3.9}$$

The proof of the inequality above shall be completed via a density argument for $v \in C^1(K_F)$. By Assumption **C**, there exists a point $\boldsymbol{a} \in K_F$, such that $\mathrm{dist}(\boldsymbol{a}, P_F) \eqsim h_{K_F} \eqsim h_K$ and $\mathrm{conv}(\boldsymbol{a}, P_F) \subset K_F$ (see, e.g., Figure 5). Now, thanks to the convexity of P_F, it has the following local polar coordinate representation using $\boldsymbol{a}$ as the origin:

$$P_F = \{\boldsymbol{x} = \boldsymbol{x}(r, \boldsymbol{\omega}) = r\boldsymbol{\omega} : r_1(\boldsymbol{\omega}) \le r \le r_2(\boldsymbol{\omega}),\ \boldsymbol{\omega} \in A_F \subset \mathbb{S}^{d-1},\ d = 2, 3\}, \tag{3.10}$$

and $\mathrm{conv}(\boldsymbol{a}, P_F) = \{r\boldsymbol{\omega} : 0 \le r \le r_2(\boldsymbol{\omega}),\ \boldsymbol{\omega} \in A_F\}$.

For $\boldsymbol{x}(r, \boldsymbol{\omega}) \in P_F$, we denote $v(r, \boldsymbol{\omega}) := v\big(\boldsymbol{x}(r, \boldsymbol{\omega})\big)$. Now for a fixed surface variable $\boldsymbol{\omega}$, $|r_1(\boldsymbol{\omega}) - r_2(\boldsymbol{\omega})| \lesssim h_F$, and $r_i(\boldsymbol{\omega}) \eqsim h_K$. Moreover, we can choose ρ and constant c bounded away from 0 independent of h_K such that $ch_K = \rho < r_1(\boldsymbol{\omega})$, and thus $|r_2(\boldsymbol{\omega}) - \rho| \gtrsim h_K$. The mean value theorem implies that there exists a $\xi \in (\rho, r_2(\boldsymbol{\omega}))$ such that

$$v^2(\xi, \boldsymbol{\omega}) = \frac{1}{|r_2(\boldsymbol{\omega}) - \rho|} \int_\rho^{r_2(\boldsymbol{\omega})} v^2(t, \boldsymbol{\omega})\, \mathrm{d}t \lesssim \frac{1}{h_K} \int_\rho^{r_2(\boldsymbol{\omega})} v^2(t, \boldsymbol{\omega})\, \mathrm{d}t. \tag{3.11}$$

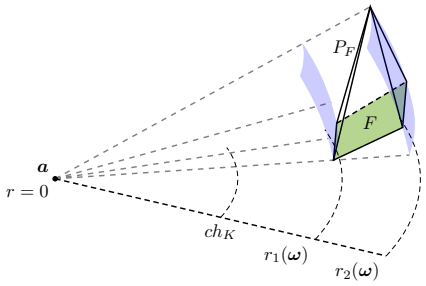


FIG. 5. *An illustration of* $\operatorname{conv}(\boldsymbol{a}, P_F) \subset K_F$.

By the fundamental theorem of calculus, Young's inequality, and the inequality above, we have

$$\begin{aligned} v^2(r,\boldsymbol{\omega}) &= v^2(\xi,\boldsymbol{\omega}) + \int_\xi^r \partial_t\big(v^2(t,\boldsymbol{\omega})\big)\,\mathrm{d}t \le v^2(\xi,\boldsymbol{\omega}) + h_K^{-1}\int_\xi^r v^2\,\mathrm{d}t + h_K\int_\xi^r |\partial_t v|^2\,\mathrm{d}t \\ &\lesssim h_K^{-1}\int_\rho^{r_2(\boldsymbol{\omega})} v^2\,\mathrm{d}t + h_K\int_\rho^{r_2(\boldsymbol{\omega})} |\partial_t v|^2\,\mathrm{d}t, \end{aligned}$$

where we note that ξ is not present in the final inequality above, and thus this inequality holds for any $\boldsymbol{\omega} \in A_F$. Integrating the inequality above with respect to $r^{d-1}\,\mathrm{d}r$ and using the fact that $|r_1(\boldsymbol{\omega}) - r_2(\boldsymbol{\omega})| \le h_{P_F} \eqsim h_F$, we have

$$\begin{aligned} \int_{r_1(\boldsymbol{\omega})}^{r_2(\boldsymbol{\omega})} v^2(r,\boldsymbol{\omega}) r^{d-1}\,\mathrm{d}r &\lesssim \int_{r_1(\boldsymbol{\omega})}^{r_2(\boldsymbol{\omega})} \left(h_K^{-1}\int_\rho^{r_2(\boldsymbol{\omega})} v^2\,\mathrm{d}t + h_K\int_\rho^{r_2(\boldsymbol{\omega})} |\partial_t v|^2\,\mathrm{d}t \right) r^{d-1} dr \\ &\lesssim h_F h_K^{d-1}\left(h_K^{-1}\int_\rho^{r_2(\boldsymbol{\omega})} v^2\,\mathrm{d}t + h_K\int_\rho^{r_2(\boldsymbol{\omega})} |\partial_t v|^2\,\mathrm{d}t \right) := (\dagger). \end{aligned}$$

As $t > \rho = ch_K$ in the integrals and c is bounded away from 0, the factor h_K^{d-1} can be moved into the integrals above, and thus (†) can be bounded by

$$(\dagger) \lesssim h_F\left(h_K^{-1}\int_\rho^{r_2(\boldsymbol{\omega})} v^2(t,\boldsymbol{\omega}) t^{d-1}\,\mathrm{d}t + h_K\int_\rho^{r_2(\boldsymbol{\omega})} |\partial_t v(t,\boldsymbol{\omega})|^2 t^{d-1}\,\mathrm{d}t \right).$$

Last, $\boldsymbol{x} = r\boldsymbol{\omega}$ together with $|\boldsymbol{\omega}| = 1$ implies $|\partial_t v(t,\boldsymbol{\omega})| \le |\nabla v|$, and integrating both sides of the integral above with respect to surface measure $\mathrm{d}\boldsymbol{\omega}$ on A_F and rearranging the factors yield

$$h_F^{-1}\|v\|_{0,P_F}^2 \lesssim h_K\|v\|_{0,\operatorname{conv}(\boldsymbol{a},P_F)}^2 + h_K\|\nabla v\|_{0,\operatorname{conv}(\boldsymbol{a},P_F)}^2.$$

Consequently, (3.9) is valid since $\operatorname{conv}(\boldsymbol{a}, P_F) \subset K_F$, and the lemma follows. ∎

Remark 3.4. When K is uniformly star-shaped, we can choose the vertex $\boldsymbol{a}$ as the center of the largest inscribed ball for all faces F. With Assumptions **B** and **C**, the vertex $\boldsymbol{a}$ could vary for different faces, and thus a more flexible geometry is allowed. See Figures 3a and 3c for examples satisfying Assumptions **A**–**C** but that are not uniformly star-shaped.

3.5. Poincaré inequalities. In this subsection, we review Poincaré–Friedrichs inequalities with a constant depending only on the diameter of the domain but not on the shape.

LEMMA 3.5 (Poincaré inequality of a linear polynomial on a face). *On any face $F \in \mathcal{F}_h(K)$, for a linear polynomial $q \in \mathbb{P}_1(F)$, the estimate*

$$\left\|q - \overline{q}^F\right\|_{0,F} \leq h_F \left\|\nabla_F q\right\|_{0,F},$$

where ∇_F denotes the surface gradient on F.

Proof. Here we use the local Cartesian coordinate $\xi \boldsymbol{t}_{F,1} + \eta \boldsymbol{t}_{F,2} =: \boldsymbol{x} \in F$ in defining (3.1); then for $q \in \mathbb{P}_1(F)$, $q = \boldsymbol{x} \cdot \nabla_F q + c$, where $\nabla_F q$ is a constant vector. The lemma then follows from the a direct calculation: $q - \overline{q}^F = (\boldsymbol{x} - \overline{\boldsymbol{x}}^F) \cdot \nabla_F q$, and

$$\left\|q - \overline{q}^F\right\|_{0,F}^2 \leq \int_F |\boldsymbol{x} - \overline{\boldsymbol{x}}^F|^2 \, |\nabla_F q|^2 \, \mathrm{d}S \leq h_F^2 \left\|\nabla_F q\right\|_{0,F}^2. \tag{3.12}$$

∎

LEMMA 3.6 (Poincaré inequality of a linear polynomial on the patch). *For a linear polynomial $q \in \mathbb{P}_1(\omega_K)$ such that $\int_{\partial\omega_K} q \, \mathrm{d}S = 0$, where ω_K satisfies Assumption* **D**, *the following estimate holds with a constant independent of the geometries of K or ω_K:*

$$\|q\|_{0,K} \leq h_{\omega_K} \|\nabla q\|_{0,K}.$$

Proof. For $q \in \mathbb{P}_1(\omega_K)$, $q = \boldsymbol{x} \cdot \nabla q + c$ with a constant c. By the fact that the constraint is imposed on the boundary integral on $\partial\omega_K$, similar to the previous lemma, it can be verified that $q = (\boldsymbol{x} - \overline{\boldsymbol{x}}^{\partial\omega_K}) \cdot \nabla q$, where

$$\overline{\boldsymbol{x}}^{\partial\omega_K} = \frac{1}{|\partial\omega_K|} \int_{\partial\omega_K} \boldsymbol{x} \, \mathrm{d}S = \sum_{F \in \mathcal{F}_h(\partial\omega_K)} \frac{|F|}{|\partial\omega_K|} \overline{\boldsymbol{x}}^F,$$

where $\overline{\boldsymbol{x}}^F \in \mathrm{conv}(F)$, and hence $\overline{\boldsymbol{x}}^{\partial\omega_K} \in \mathrm{conv}(\omega_K)$. As a result, $|\boldsymbol{x} - \overline{\boldsymbol{x}}^{\partial\omega_K}| \leq h_{\omega_K}$, and we have

$$\|q\|_{0,K}^2 \leq \int_K |\boldsymbol{x} - \overline{\boldsymbol{x}}^{\partial\omega_K}|^2 \, |\nabla q|^2 \, \mathrm{d}\boldsymbol{x} \leq h_{\omega_K}^2 \|\nabla q\|_{0,K}^2. \tag{3.13}$$

∎

Notice that the constraint is imposed on the boundary of a bigger patch ω_K but the inequality holds on a smaller region K. When using this inequality in the a priori error estimate in order to get the optimal rate of convergence, the constant will be dependent only on γ_2 introduced in Assumption **D**.

For the approximation property of the polynomial projection, we opt to use the Poincaré inequality on a convex domain, thus to utilize the convex hull of a possible nonconvex element.

LEMMA 3.7 (Poincaré inequality on the convex hull). *Let ω be a bounded simple polygon/polyhedron; then the following Poincaré inequality holds for any $v \in H^1\big(\mathrm{conv}(\omega)\big)$:*

$$\|v - \overline{v}^\omega\|_{0,\omega} \leq \frac{h_\omega}{\pi} \|\nabla v\|_{0,\mathrm{conv}(\omega)}. \tag{3.14}$$

Proof. As $\overline{v}^\omega$ is the best constant approximation in $L^2(\omega)$-norm,

$$\|v - \overline{v}^\omega\|_{0,\omega} \leq \left\|v - \overline{v}^{\mathrm{conv}(\omega)}\right\|_{0,\omega} \leq \left\|v - \overline{v}^{\mathrm{conv}(\omega)}\right\|_{0,\mathrm{conv}(\omega)} \leq \frac{h_\omega}{\pi} \|\nabla v\|_{0,\mathrm{conv}(\omega)}.$$

In the last step, the Poincaré inequality on a convex set [37] is used. ∎

We then establish a similar result when the constraint is posed on the boundary integral.

LEMMA 3.8 (Poincaré inequality with zero boundary average on an isotropic polyhedron). *Let K be a polypotal element satisfying Assumptions* **A**–**B***; then, for any $v \in H^1(\mathrm{conv}(K))$, the following Poincaré inequality holds:*

$$\left\|v - \overline{v}^{\partial K}\right\|_{0,K} \lesssim h_K \left\|\nabla v\right\|_{0,\mathrm{conv}(K)}. \tag{3.15}$$

Proof. First, the triangle inequality implies that

$$\left\|v - \overline{v}^{\partial K}\right\|_{0,K} \le \left\|v - \overline{v}^{K}\right\|_{0,K} + \left\|\overline{v}^{K} - \overline{v}^{\partial K}\right\|_{0,K}, \tag{3.16}$$

where the first term can be estimated by Lemma 3.7. Rewriting the second term above and using the Cauchy–Schwarz inequality yield

$$\begin{aligned}\left\|\overline{v}^{K} - \overline{v}^{\partial K}\right\|_{0,K} &= |K|^{1/2}\left|\frac{1}{|\partial K|}\int_{\partial K}(\overline{v}^{K} - v)\,\mathrm{d}S\right| \\ &\le \frac{|K|^{1/2}}{|\partial K|}\sum_{F\in\mathcal{F}_h(K)}|F|^{1/2}\left\|v - \overline{v}^{K}\right\|_{0,F}.\end{aligned} \tag{3.17}$$

By the trace inequality in Lemma 3.2,

$$\left\|v - \overline{v}^{K}\right\|_{0,F} \le h_F^{-1/2}\left\|v - \overline{v}^{K}\right\|_{0,K} + h_F^{1/2}\left\|\nabla v\right\|_{0,K}. \tag{3.18}$$

Applying the Poincaré inequality in Lemma 3.7 on $\left\|v - \overline{v}^{K}\right\|_{0,K}$ and the fact that $h_F \le h_K$ yield

$$\left\|\overline{v}^{K} - \overline{v}^{\partial K}\right\|_{0,K} \le \frac{h_K|K|^{1/2}}{|\partial K|}\sum_{F\in\mathcal{F}_h(K)}\left(\frac{|F|}{h_F}\right)^{1/2}\left\|\nabla v\right\|_{0,\mathrm{conv}(K)}. \tag{3.19}$$

As $|F| \lesssim h_F^{d-1}$ and $|\partial K| = \sum_{F\in\mathcal{F}_h(K)}|F|$,

$$\sum_{F\in\mathcal{F}_h(K)}\left(\frac{|F|}{h_F}\right)^{1/2} \lesssim \sum_{F\in\mathcal{F}_h(K)}|F|^{\frac{d-2}{2(d-1)}} \le n_K^{\frac{d}{2(d-1)}}|\partial K|^{\frac{d-2}{2(d-1)}}.$$

Then

$$\frac{|K|^{1/2}}{|\partial K|}\sum_{F\in\mathcal{F}_h(K)}\left(\frac{|F|}{h_F}\right)^{1/2} \lesssim C(n_K)\frac{|K|^{1/2}}{|\partial K|^{d/2(d-1)}} \lesssim C(n_K),$$

where in the last step we used the isoperimetric inequality $|K| \le C_d|\partial K|^{d/(d-1)}$. ∎

4. A priori error analysis. The error analysis will be performed under a mesh dependent norm $|||\cdot|||$ induced by $a_h(\cdot,\cdot)$, i.e., for $v \in H_0^1(\Omega) + V_h$,

$$|||v|||^2 := a_h(v,v) = \sum_{K\in\mathcal{T}_h}\left(\left\|\nabla\Pi_K v\right\|_K^2 + h_{\omega_K}^{-1}\sum_{F\in\mathcal{F}_h(K)}\left\|Q_F(v - \Pi_{\omega_K}v)\right\|_F^2\right), \tag{4.1}$$

which is weaker than the H^1-seminorm $|\cdot|_1$ upon which the conventional VEM analysis is built. We denote the local norm on K as $|||\cdot|||_K$. As all projections Π_K, Π_{ω_K}, and Q_F can be computed using only the DoFs (see (2.12) and (2.13)), it is straightforward to verify that $|||v - v_I||| = 0$ for the interpolant v_I defined using the DoFs in (2.7).

4.1. A mesh dependent norm. First, the following lemma is needed for proving that $|||\cdot|||$ is a norm on V_h which bounds the projection Π_{ω_K} in an extended element measured in the H^1-seminorm.

LEMMA 4.1 (bound of the projection Π_{ω_K} on the patch). *Let $\overline{\omega} = \overline{\cup_{\alpha\in A} K_\alpha}$ for $v \in H^1(\omega)$; then the following estimate holds:*

$$\|\nabla \Pi_\omega v\|_{0,\omega}^2 \le \sum_{\alpha\in A} \|\nabla \Pi_{K_\alpha} v\|_{0,K_\alpha}^2 . \tag{4.2}$$

Proof. By definition (2.10), for any $q \in \mathbb{P}_1(\omega)$, $q\big|_{K_\alpha} \in \mathbb{P}_1(K_\alpha)$, and thus by the Cauchy–Schwarz inequality and an ℓ^2–ℓ^2 Hölder inequality, we have

$$\begin{aligned} (\nabla \Pi_\omega v, \nabla q)_\omega &= (\nabla v, \nabla q)_\omega = \sum_{\alpha\in A} (\nabla v, \nabla q)_{K_\alpha} = \sum_{\alpha\in A} (\nabla \Pi_{K_\alpha} v, \nabla q)_{K_\alpha} \\ &\le \sum_{\alpha\in A} \|\nabla \Pi_{K_\alpha} v\|_{0,K_\alpha} \|\nabla q\|_{0,K_\alpha} \le \left(\sum_{\alpha\in A} \|\nabla \Pi_{K_\alpha} v\|_{0,K_\alpha}^2 \right)^{1/2} \|\nabla q\|_{0,\omega} . \end{aligned} \tag{4.3}$$

The lemma then follows from letting $q = \Pi_\omega v$. ∎

LEMMA 4.2. *$|||\cdot|||$ defines a norm on the nonconforming VEM space V_h.*

Proof. Since each component of $|||\cdot|||$ supports the triangle inequality and is scalable, it suffices to verify that if $|||v_h||| = 0$ for $v_h \in V_h$, then $v_h \equiv 0$. By definition, $a_h(v_h, v_h) = |||v_h|||^2 = 0$ implies that

$$\nabla \Pi_K v_h = \mathbf{0} \quad \forall K \in \mathcal{T}_h; \quad Q_F(v_h - \Pi_{\omega_K} v_h) = 0 \text{ on } F \quad \forall F \in \mathcal{F}_h(K). \tag{4.4}$$

Without loss of generality, we assume that ω_K consists of K and K' sharing a face, which covers the case of $\omega_K = K$, while it can be generalized to the case where ω_K contains three or more elements.

First, by Lemma 4.1, $\nabla \Pi_{\omega_K} v_h = \mathbf{0}$ since $\nabla \Pi_K v_h = \nabla \Pi_{K'} v_h = \mathbf{0}$. Restricting ourselves on K, consider the following quantity:

$$\begin{aligned} \|\nabla v_h\|_{0,K}^2 &= \big(\nabla v_h, \nabla(v_h - \Pi_{\omega_K} v_h)\big)_K \\ &= -\big(\Delta v_h, v_h - \Pi_{\omega_K} v_h\big)_K + \langle \nabla v_h \cdot \boldsymbol{n}, v_h - \Pi_{\omega_K} v_h \rangle_{\partial K} \\ &= \sum_{F\in\mathcal{F}_h(K)} \big(\nabla v_h \cdot \boldsymbol{n}, Q_F(v_h - \Pi_{\omega_K} v_h)\big)_F = 0. \end{aligned} \tag{4.5}$$

In the last step, $\Delta v_h = 0$ in K is used. Since $\nabla v_h \cdot \boldsymbol{n} \in \mathbb{P}_0(F)$, the L^2-projection Q_F can be inserted into the pair.

As a result of (4.5), in every K, $\nabla v_h = \mathbf{0}$, and thus $v_h =$ constant. Finally, by the boundary condition and the continuity condition in (2.5), $v_h \equiv 0$. ∎

4.2. A priori error estimates on isotropic elements. Next, an error equation is developed for the lowest-order nonconforming VEM following [20, 36], and the a priori error analysis on isotropic elements is established.

LEMMA 4.3 (an error equation). *Let u_h and u_I be the solution to problem* (2.20) *and the canonical interpolation in* (2.7)*, respectively, and let u_π be any piecewise linear*

polynomial on $\mathcal{T}_h$; for any $v_h \in V_h$ and stabilization $S_K(\cdot,\cdot)$, it holds that

$$
\begin{aligned}
a_h(u_h - u_I, v_h) = &\sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_h(K)} \langle \nabla(u-u_\pi)\cdot \boldsymbol{n}, Q_F v_h - \Pi_K v_h\rangle_F \\
&- \sum_{K\in\mathcal{T}_h} S_K\big(u_I - \Pi_{\omega_K} u_I, v_h - \Pi_{\omega_K} v_h\big).
\end{aligned} \tag{4.6}
$$

Proof. Using the VEM discretization problem (2.20), the original PDE $-\Delta u = f$, the definition of the elliptic projection (2.10), and integration by parts, we have

$$
\begin{aligned}
&a_h(u_h - u_I, v_h) \\
&= \sum_{K\in\mathcal{T}_h} \big(f, \Pi_K v_h\big)_K - a_h(u_I, v_h) = \sum_{K\in\mathcal{T}_h} \big(-\Delta u, \Pi_K v_h\big)_K - a_h(u_I, v_h) \\
&= \sum_{K\in\mathcal{T}_h} \big(\nabla\Pi_K u, \nabla \Pi_K v_h\big)_K - \sum_{K\in\mathcal{T}_h} \langle \nabla u\cdot\boldsymbol{n}, \Pi_K v_h\rangle_{\partial K} - a_h(u_I, v_h) \\
&= \sum_{K\in\mathcal{T}_h} \big(\nabla\Pi_K (u-u_I), \nabla \Pi_K v_h\big)_K + \sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_h(K)} \langle \nabla u\cdot\boldsymbol{n}, Q_F v_h - \Pi_K v_h\rangle_F \\
&\quad - \sum_{K\in\mathcal{T}_h} S_K\big(u_I - \Pi_{\omega_K} u_I, v_h - \Pi_{\omega_K} v_h\big).
\end{aligned} \tag{4.7}
$$

We note that in the derivation above, on each face F, $Q_F v_h$, which is single-valued on F, can be freely inserted into boundary integrals since the interelement jump of $\nabla u\cdot\boldsymbol{n}$ on F vanishes by the assumption that $f \in L^2(\Omega)$.

Moreover, since $\chi_F(u-u_I) = 0$ for all faces F by (2.6), by Lemma 2.1 we have $\Pi_K(u-u_I) = 0$, and the first term in (4.7) vanishes. Last, using the fact that in the lowest-order case, since $u_\pi \in \mathbb{P}_1(K)$, $\Delta u_\pi = 0$, the following zero term can be inserted into the boundary integral in (4.7) to get (4.6):

$$
\begin{aligned}
&\sum_{F\in\partial K} \langle \nabla u_\pi\cdot\boldsymbol{n}, Q_F v_h - \Pi_K v_h\rangle_F = \langle \nabla u_\pi\cdot\boldsymbol{n}, v_h - \Pi_K v_h\rangle_{\partial K} \\
&= \big(\Delta u_\pi, v_h - \Pi_K v_h\big)_K + \big(\nabla u_\pi, \nabla(v_h - \Pi_K v_h)\big)_K = 0.
\end{aligned}
$$

∎

Lemma 4.4 (an a priori error estimate on isotropic meshes). *Under the same setting with Lemma 4.3, when the mesh $\mathcal{T}_h$ satisfies Assumptions* **A**–**B**, *it holds that*

$$
\begin{aligned}
|\!|\!| u_h - u_I |\!|\!|^2 \lesssim &\sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_h(K)} h_K \left\|\nabla\big(u - \Pi_K u\big)\cdot\boldsymbol{n}\right\|_{0,F}^2 \\
&+ \sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_h(K)} h_K^{-1} \left\|u - \Pi_K u\right\|_{0,F}^2.
\end{aligned} \tag{4.8}
$$

Proof. As $\mathcal{T}_h$ contains only isotropic elements, $\omega_K = K$ for all $K \in \mathcal{T}_h$. Let $u_\pi = \Pi_K u$ and $v_h = u_h - u_I$ in (4.6). Similarly to the proof of Lemma 4.3, definition (2.4) of DoFs with (2.6) implies that $Q_F u = Q_F u_I$, and hence $\Pi_K u_I = \Pi_K u$ by

Lemma 2.1. As a result, the stabilization term in (4.6) can be estimated as follows:

(4.9)
$$\begin{aligned} &S_K\big(u_I - \Pi_K u_I, v_h - \Pi_K v_h\big) \\ &\le \sum_{F\in\mathcal{F}_h(K)} h_K^{-1} \left\|Q_F(u_I - \Pi_K u_I)\right\|_{0,F} \left\|Q_F(v_h - \Pi_K v_h)\right\|_{0,F} \\ &\le \left(\sum_{F\in\mathcal{F}_h(K)} h_K^{-1} \left\|Q_F(u - \Pi_K u)\right\|_{0,F}^2\right)^{1/2} \left(\sum_{F\in\mathcal{F}_h(K)} h_K^{-1} \left\|Q_F(v_h - \Pi_K v_h)\right\|_{0,F}\right)^{1/2}, \end{aligned}$$

in which the first term can be estimated by $\|Q_F(u - \Pi_K u)\|_{0,F} \le \|u - \Pi_K u\|_{0,F}$, and the second term is a part of $|\!|\!|v_h|\!|\!|$. For the boundary integral term in (4.6), after using the Cauchy–Schwarz inequality on each face F,

(4.10)
$$\langle \nabla(u - \Pi_K u)\cdot \boldsymbol{n}, Q_F v_h - \Pi_K v_h\rangle_F \le \left\|\nabla(u - \Pi_K u)\cdot \boldsymbol{n}\right\|_{0,F} \left\|Q_F v_h - \Pi_K v_h\right\|_{0,F},$$

we assign $h_K^{1/2}$ to the first term and $h_K^{-1/2}$ to the second term in (4.10) and apply the triangle inequality as follows:

$$\left\|Q_F v_h - \Pi_K v_h\right\|_{0,F} \le \left\|Q_F(v_h - \Pi_K v_h)\right\|_{0,F} + \left\|Q_F\Pi_K v_h - \Pi_K v_h\right\|_{0,F}. \tag{4.11}$$

Consequently, the first term above, together with the weight $h_K^{-1/2}$, is now a part of $|\!|\!|v_h|\!|\!|$. Applying the Poincaré inequality for the linear polynomial $\Pi_K v_h$ on face F in Lemma 3.5 on the second term above, together with $|F|h_F \lesssim |F|l_F \le |K|$ implied by the height condition **B**, leads to

$$h_K^{-1/2} \left\|Q_F\Pi_K v_h - \Pi_K v_h\right\|_{0,F} \le h_F^{1/2} \left\|\nabla_F \Pi_K v_h\right\|_{0,F} \lesssim \left\|\nabla \Pi_K v_h\right\|_{0,K}, \tag{4.12}$$

which is a part of $|\!|\!|v_h|\!|\!|$. Last, summing up (4.10) in the ℓ^2-sense yields the lemma. ∎

With the a priori error estimate in Lemma 4.4, it suffices to estimate the two terms from estimate (4.8). First, we estimate $(u - \Pi_K u)$ in the following lemma.

Lemma 4.5 (error estimate of Π_K on an isotropic element). *When K satisfies Assumptions* **A**–**C**, *for $u \in H^2\big(\operatorname{conv}(K)\big)$ it holds that*

$$h_K^{-1} \left\|u - \Pi_K u\right\|_{0,K} + \left\|\nabla(u - \Pi_K u)\right\|_{0,K} \lesssim h_K |u|_{2,\operatorname{conv}(K)}. \tag{4.13}$$

Proof. Since $\nabla\Pi_K u = \overline{\nabla u}^K$, the estimate in the second term follows from the Poincaré inequality in Lemma 3.7. For the first term, by constraint (2.11), applying the Poincaré inequality in Lemma 3.8 and the triangle inequality lead to

$$\begin{aligned} &h_K^{-1} \left\|u - \Pi_K u\right\|_{0,K} \lesssim \left\|\nabla(u - \Pi_K u)\right\|_{0,\operatorname{conv}(K)} \\ &\le \left\|\nabla u - \overline{\nabla u}^{\operatorname{conv}(K)}\right\|_{0,\operatorname{conv}(K)} + \left\|\overline{\nabla u}^{\operatorname{conv}(K)} - \overline{\nabla u}^K\right\|_{0,\operatorname{conv}(K)}. \end{aligned} \tag{4.14}$$

For the second term above, the Cauchy–Schwarz inequality, $|\operatorname{conv}(K)| \lesssim h_K^d$, and $|K| \eqsim h_K^d$ in Lemma 3.1 imply that

$$\begin{aligned} &\left\|\overline{\nabla u}^{\operatorname{conv}(K)} - \overline{\nabla u}^K\right\|_{0,\operatorname{conv}(K)} = |\operatorname{conv}(K)|^{1/2} \left|\frac{1}{|K|}\int_K \left(\nabla u - \overline{\nabla u}^{\operatorname{conv}(K)}\right)\right| \\ &\le \frac{|\operatorname{conv}(K)|^{1/2}}{|K|^{1/2}} \left\|\nabla u - \overline{\nabla u}^{\operatorname{conv}(K)}\right\|_{0,K} \lesssim \left\|\nabla u - \overline{\nabla u}^{\operatorname{conv}(K)}\right\|_{0,\operatorname{conv}(K)}. \end{aligned} \tag{4.15}$$

Consequently, the desired estimate follows from applying Lemma 3.7 on $\operatorname{conv}(K)$ and the fact that the diameter of $\operatorname{conv}(K)$ is h_K. ∎

LEMMA 4.6 (error estimate of the normal derivative of Π_K). *For $K \in \mathcal{T}_h$, provided that every F satisfies Assumptions* **B**–**C**, *the following error estimate holds on a face $F \in \mathcal{F}_h(K)$ for $u \in H^2\big(\operatorname{conv}(K)\big)$:*

$$h_K^{1/2} \|\nabla(u - \Pi_K u) \cdot \boldsymbol{n}\|_{0,F} \lesssim h_K |u|_{2,\operatorname{conv}(K)}. \tag{4.16}$$

Proof. By Assumptions **B**–**C**, we apply trace inequality (3.8) toward K:

$$h_K^{1/2} \|\nabla(u - \Pi_K u) \cdot \boldsymbol{n}\|_{0,F} \lesssim \|\nabla(u - \Pi_K u)\|_{0,K} + h_K |u|_{2,K}. \tag{4.17}$$

The lemma then follows from Lemma 4.5. ∎

LEMMA 4.7 (error estimate of Π_K on a face). *For $K \in \mathcal{T}_h$, provided that K satisfies Assumptions* **A**–**C**, *the following error estimate holds on a face $F \in \mathcal{F}_h(K)$ for $u \in H^2\big(\operatorname{conv}(K)\big)$:*

$$h_K^{-1/2} \|u - \Pi_K u\|_{0,F} \lesssim h_K |u|_{2,\operatorname{conv}(K)}. \tag{4.18}$$

Proof. Since for every $F \in \mathcal{F}_h(K)$, F satisfies Assumptions **B**–**C** with respect to K, by the trace inequality in Lemma 3.3, we have

$$h_K^{-1/2} \|u - \Pi_K u\|_{0,F} \lesssim h_K^{-1} \|u - \Pi_K u\|_{0,K} + \|\nabla(u - \Pi_K u)\|_{0,K}, \tag{4.19}$$

which yields the desired estimate by Lemma 4.5. ∎

Now the a priori convergence result for the lowest-order nonconforming VEM on an isotropic mesh can be summarized as follows.

THEOREM 4.8 (convergence on isotropic meshes). *Assume that the mesh $\mathcal{T}_h$ is isotropic in the sense of Assumptions* **A**–**E**. *When the solution u to* (2.1) *satisfies $u \in H^2(\Omega)$, the following error estimate holds for the solution u_h to* (2.20)*:*

$$|\!|\!| u - u_h |\!|\!| \lesssim h \|u\|_{2,\Omega}. \tag{4.20}$$

Proof. First, we apply Stein's extension theorem [38, Theorem 6.5] to $u \in H^2(\Omega)$ to get a function $u_E \in H^2(\mathbb{R}^d)$, $u_E|_\Omega = u|_\Omega$, and $\|u_E\|_{2,\mathbb{R}^d} \le C(\Omega)\|u\|_{2,\Omega}$. With this extension, $u_E \in H^2\big(\operatorname{conv}(K)\big)$ for any $K \in \mathcal{T}_h$.

Second, the estimates from Lemmas 4.6 and 4.7 are plugged into Lemma 4.4, and Assumption **A** ensures that these estimates are summed up bounded times on a fixed element. Meanwhile, Assumption **E** implies that the integral on the overlap $\operatorname{conv}(K) \cap \operatorname{conv}(K')$ is repeated bounded times for neighboring $K, K' \in \mathcal{T}_h$. Therefore,

$$|\!|\!| u_I - u_h |\!|\!|^2 \lesssim \sum_{K \in \mathcal{T}_h} h_K^2 |u_E|^2_{2,\operatorname{conv}(K)} \lesssim h^2 |u_E|^2_{2,\operatorname{conv}(\Omega)} \lesssim h^2 \|u\|^2_{2,\Omega}. \tag{4.21}$$

As $|\!|\!| u - u_I |\!|\!| = 0$ by the construction of u_I and (4.1), the theorem follows. ∎

4.3. A priori error estimates on anisotropic elements. In the vanilla error equation (4.6), a boundary term that involves $\Pi_K v_h$ is present. For an anisotropic element K, key estimates including (4.12), (4.17), and (4.19) will become problematic where Assumptions **B**–**C** are violated. Instead, the boundary term will be lifted to its isotropic extended element patch ω_K, and thus in the next lemma we aim to replace $\Pi_K v$ by $\Pi_{\omega_K} v$ in the error equation (4.6) taking the anisotropic elements into account.

LEMMA 4.9 (expanded error equation). *Under the same setting with Lemma* 4.3*, it holds that*

$$
\begin{aligned}
a_h(u_h - u_I, v_h) = &\sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_h(K)} \langle \nabla(u-u_\pi)\cdot\boldsymbol{n}, Q_F v_h - \Pi_{\omega_K} v_h\rangle_F \\
&- \sum_{K\in\mathcal{T}_h} S_K\big(u_I - \Pi_{\omega_K} u_I, v_h - \Pi_{\omega_K} v_h\big) \\
&+ \sum_{K\in\mathcal{T}_h} \big(\nabla(u-u_\pi), \nabla\big(\Pi_{\omega_K} v_h - \Pi_K v_h\big)\big)_K \\
&- \sum_{K\in\mathcal{T}_h} (f, \Pi_{\omega_K} v_h - \Pi_K v_h)_K.
\end{aligned} \tag{4.22}
$$

Proof. Starting with (4.6), we only need to expand the difference term as follows:

$$
\begin{aligned}
&\sum_{F\in\mathcal{F}_h(K)} \langle \nabla(u-u_\pi)\cdot\boldsymbol{n}, \Pi_{\omega_K} v_h - \Pi_K v_h\rangle_F \\
&= \big(\Delta(u-u_\pi), \Pi_{\omega_K} v_h - \Pi_K v_h\big)_K + \big(\nabla(u-u_\pi), \nabla(\Pi_{\omega_K} v_h - \Pi_K v_h)\big)_K \\
&= -(f, \Pi_{\omega_K} v_h - \Pi_K v_h)_K + \big(\nabla(u-u_\pi), \nabla(\Pi_{\omega_K} v_h - \Pi_K v_h)\big)_K.
\end{aligned}
$$

∎

For the last term in (4.22) involving difference in an L^2-inner product, Poincaré inequalities with appropriate constraints can be applied to change it to the energy norm.

LEMMA 4.10 (difference between projections). *If Assumption* **D** *is met for* K*, denote* $\|\!|v_h|\!\|_{\omega_K}^2 := \sum_{K\in\mathcal{T}_h(\omega_K)} \|\!|v_h|\!\|_K^2$*; then*

$$
\|\Pi_K v_h - \Pi_{\omega_K} v_h\|_{0,K} \lesssim h_{\omega_K} \|\!|v_h|\!\|_{\omega_K}. \tag{4.23}
$$

Proof. As we choose the constraint $\int_{\partial\omega_K} \Pi_K v_h = \int_{\partial\omega_K} \Pi_{\omega_K} v_h = \int_{\partial\omega_K} v_h$, and $\Pi_K v_h - \Pi_{\omega_K} v_h$ is a linear polynomial on K, the estimate is a direct consequence of Lemmas 3.6 and 4.1. ∎

LEMMA 4.11 (a priori error estimate using the expanded error equation). *Under the same setting with Lemma* 4.3*, when* $\mathcal{T}_h$ *satisfies Assumptions* **A**–**D***, it holds that*

$$
\begin{aligned}
\|\!|u_h - u_I|\!\|^2 \lesssim &\sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_h(K)} h_{\omega_K} \big\|\nabla\big(u - \Pi_{\omega_K} u\big)\cdot\boldsymbol{n}\big\|_{0,F}^2 \\
&+ \sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_h(K)} h_{\omega_K}^{-1} \|u - \Pi_{\omega_K} u\|_{0,F}^2 \\
&+ \sum_{K\in\mathcal{T}_h} \|\nabla(u - \Pi_{\omega_K} u)\|_{0,K}^2 + \sum_{K\in\mathcal{T}_h} h_{\omega_K}^2 \|f\|_{0,K}^2.
\end{aligned} \tag{4.24}
$$

Proof. We proceed similarly with the proof of Lemma 4.4 by choosing $v_h = u_h - u_I$ yet letting $u_\pi = \Pi_{\omega_K} u$ instead in (4.22). The four terms in (4.22) shall be estimated in a backward order. For the fourth term, by the Cauchy–Schwarz inequality and applying Lemma 4.10, we have

$$
\big(f, (\Pi_K - \Pi_{\omega_K}) v_h\big)_K \le \|f\|_{0,K} \|(\Pi_K - \Pi_{\omega_K}) v_h\|_K \lesssim h_{\omega_K} \|f\|_{0,K} \|\!|v_h|\!\|_{\omega_K}. \tag{4.25}
$$

The third term can be estimated in a similar fashion by applying the Cauchy–Schwarz inequality and applying Lemma 4.1 to get

$$
\|\nabla(\Pi_{\omega_K} v_h - \Pi_K v_h)\|_{0,K} \le \|\nabla\Pi_{\omega_K} v_h\|_{0,\omega_K} + \|\nabla\Pi_K v_h\|_{0,K} \lesssim \|\!|v_h|\!\|_{\omega_K}. \tag{4.26}
$$

For the second term, which is the stabilization, an argument similar to that in (4.9) in the proof of Lemma 4.4 can be used. By $\mathcal{F}_h(\omega_K) \subset \mathcal{F}_h$, Lemma 2.1 implies that $\Pi_{\omega_K} u_I = \Pi_{\omega_K} u$, which leads to an estimate similar to the second term in (4.8), and the difference is that Π_K and h_K are replaced in (4.9) by Π_{ω_K} and h_{ω_K}, respectively.

The first term of (4.22) is treated similarly with (4.10) and (4.11); then, since ω_K is isotropic, the rest of the proof, in which Π_K and h_K are replaced by Π_{ω_K} and h_{ω_K}, proceeds exactly the same as with (4.12):

$$h_{\omega_K}^{-1/2} \|Q_F \Pi_{\omega_K} v_h - \Pi_{\omega_K} v_h\|_{0,F} \lesssim h_F^{1/2} \|\nabla_F \Pi_{\omega_K} v_h\|_{0,F} \lesssim \|\nabla \Pi_{\omega_K} v_h\|_{0,\omega_K}, \tag{4.27}$$

and finally the lemma follows from Lemma 4.1. ∎

With the a priori error estimate in Lemma 4.11, it suffices to estimate term by term in (4.24). Since now it involves only the error of the projection on the isotropic extended patch ω_K, the estimates in Lemmas 4.5, 4.6, and 4.7 can be reused by replacing the K with ω_K, both of which are isotropic.

The next theorem summarizes an a priori convergence result that incorporates possible anisotropic elements (cf. Theorem 4.8), and we remark that Assumption **D** includes the scenarios when Assumptions **B**–**C** are met as $\omega_K = K$.

THEOREM 4.12 (convergence on possible anisotropic meshes). *Assume that the mesh $\mathcal{T}_h$ satisfies Assumptions* **A**–**E**. *When the solution u to problem* (2.1) *satisfies $u \in H^2(\Omega)$, the following error estimate holds for the solution u_h to problem* (2.20)*:*

$$|||u - u_h||| \lesssim h \|u\|_{2,\Omega}. \tag{4.28}$$

Proof. We proceed exactly like Theorem 4.8 by extending u to $H^2(\mathbb{R}^d)$ first. The estimate in Lemma 4.5 can be changed straightforwardly on ω_K:

$$\|\nabla(u - \Pi_{\omega_K} u)\|_{0,\omega_K} \lesssim h_{\omega_K} |u|_{2,\mathrm{conv}(\omega_K)}. \tag{4.29}$$

Since ω_K satisfies Assumptions **B**–**C** by Assumption **D**, the estimates in Lemmas 4.7 and 4.6 are changed accordingly on ω_K as well:

$$h_{\omega_K}^{-1/2} \|u - \Pi_{\omega_K} u\|_{0,F} \lesssim h_{\omega_K} |u|_{2,\mathrm{conv}(\omega_K)}, \tag{4.30}$$

$$h_{\omega_K}^{1/2} \big\|\nabla\big(u - \Pi_{\omega_K} u\big) \cdot \boldsymbol{n}\big\|_{0,F} \lesssim h_{\omega_K} |u|_{2,\mathrm{conv}(\omega_K)}. \tag{4.31}$$

After these estimates are plugged into Lemma 4.11, Assumptions **A**–**E** are applied in the same way with Theorem 4.8, except now we consider the integral overlap on patches $\mathrm{conv}(\omega_K) \cap \mathrm{conv}(\omega_{K'})$ for neighboring elements. Upon using the fact that $\|f\|_{0,K} = \|\Delta u\|_{0,K} \le |u|_{2,K}$, we obtain

$$|||u_I - u_h|||^2 \lesssim \sum_{K \in \mathcal{T}_h} h_{\omega_K}^2 |u_E|_{2,\mathrm{conv}(\omega_K)}^2 \lesssim h^2 |u_E|_{2,\mathrm{conv}(\Omega)}^2 \lesssim h^2 \|u\|_{2,\Omega}^2, \tag{4.32}$$

and the rest of the proof is the same as the one in Theorem 4.8. ∎

5. Concluding remarks and future study. The error analysis in this paper further relaxes and extends to 3D the geometry constraints for the linear VEM's conforming counterpart in [20], and Assumptions **B**–**C** can be generalized for arbitrary dimension. Since the stabilization is of a weighted L^2 type, unlike the analysis in [20] bridging the stabilization with a discrete $H^{1/2}$-norm on the boundary, the versatility

of the VEM framework allows the stabilization in the nonconforming VEM to be more flexible and localizable.

As a result, even for the isotropic case in 3D, the current analysis allows a tiny face and anisotropic face, provided that the element is isotropic in the sense of Assumptions **B**–**C**, in addition to two alternative shape-regularity conditions in Assumptions **A**–**E**. In our view, being "isotropic" for an element is a localized property near a face, in that the tangential direction and the normal direction of this face are comparable by the height condition **B**. Furthermore, the hourglass condition **C** can be viewed as a localized star-shaped condition for face F, where K_F can be different for different F. The convexity of K_F allows any line connecting a point in K_F to a face of P_F to be entirely in K_F (cf. Remark 3.4), which makes K_F's role similar to the inscribed ball to which K is uniformly star-shaped in the traditional VEM analysis. Meanwhile, the existence of concave faces is allowed in the decomposition sense.

One of the major factors facilitating the new analysis is the introduction of the stabilization on an extended element patch in the discretization (2.18). We state some of the concerns regarding the implementation using a 2D example in the following subsections.

5.1. Implementation remarks on the extended patch. When $\mathcal{T}_h$ is a body-fitted mesh generated by cutting a shape-regular background grid, ω_K, which is only needed for certain anisotropic cut elements $\{K\}$, can be naturally chosen as the patch joining K with one of K's nearest neighbors in the background mesh. Here we shall illustrate using the elements $K_1, K_2 \in \mathcal{T}_h$ in Figure 1a, which are cut from a Cartesian mesh in 2D.

When $\mathcal{T}_h$ is not generated from cutting a background shape-regular mesh, the situation is much more complicated, as the search for a possible extended element patch may produce more overhead. To pin down ω_K for an anisotropic K, one possible procedure is to estimate the chunkiness parameter of K first: computing $|K|$ and diameter h_F of an edge or a face $F \subset \partial K$, if the ratio $h_F/|K|^{1/d}$ is bigger than a threshold, then K shall be treated as anisotropic. Starting from an anisotropic element, we can join its immediate neighbor sharing an edge or a face with K to form ω_K, such that ω_K has the minimum chunkiness parameter among all possible unions with neighbors. Last, this test is repeated when necessary until ω_K passes the test.

In the implementation, using the data structure for polyhedral elements [22], one can use an array to store all faces and another to store the indices of the polyhedra to which the every face belongs. In this regard, the elements are represented by these two arrays, and the merging of neighboring elements is very efficient; we refer the reader to [22, section 3.2] for technical details.

5.2. Implementation of the new stabilization. In the elementwise assembling of the matrix corresponding to the bilinear form (2.18), we shall separate the terms of the projected gradient part $\left(\nabla\Pi_K u_h, \nabla\Pi_K v_h\right)_K$ and the stabilization part $S_K\big(u_h - \Pi_{\omega_K} u_h, v_h - \Pi_{\omega_K} v_h\big)$. The former remains unchanged from the unmodified formulation. We focus on the implementation of the stabilization term.

For each anisotropic element K, assume that we have found an extended patch ω_K which itself is also represented as a polytopal element. Then Π_{ω_K} can be realized by a matrix $\mathbf{\Pi}_{\omega_K}$ of size $(d+1)\times n_{\omega_K}$. The L^2-projection Q_F on $F \in \mathcal{F}_h(\partial K)$ applied to a linear polynomial is realized by the DoF matrix $\mathbf{D}$ of size $n_K \times (d+1)$. See [7] for detailed formulations of matrices $\mathbf{\Pi}_{\omega_K}$ and $\mathbf{D}$. Denote by $\bar{\mathbf{I}} = (\mathbf{I} \quad \mathbf{0})_{n_K \times n_{\omega_K}}$ the extended identity matrix. The stabilization on K can be realized by an $n_{\omega_K} \times n_{\omega_K}$

local matrix

$$h_{\omega_K}^{-1}(\bar{\mathbf{I}} - \mathbf{D}\boldsymbol{\Pi}_{\omega_K})^\intercal \mathrm{diag}(|F_1|, \ldots, |F_{n_K}|)(\bar{\mathbf{I}} - \mathbf{D}\boldsymbol{\Pi}_{\omega_K}). \tag{5.1}$$

Thus the standard assembling procedure looping over all elements can be applied to assemble a global one.

As a comparison, the original stabilization using Π_K is a matrix of size $n_K \times n_K$ and in the form $h_K^{-1}(\mathbf{I} - \mathbf{D}\boldsymbol{\Pi}_K)^\intercal \mathrm{diag}(|F_1|, \ldots, |F_{n_K}|)(\mathbf{I} - \mathbf{D}\boldsymbol{\Pi}_K)$. We note that one effect of enlarging the element is that the stabilization matrix (5.1) is denser or, equivalently, the stencil is larger.

Acknowledgments. We appreciate an anonymous reviewer for bringing up several insightful questions which improved an early version of the paper, and Dr. Xuehai Huang for careful proofreading.

REFERENCES

[1] B. Ahmad, A. Alsaedi, F. Brezzi, L. D. Marini, and A. Russo, *Equivalent projectors for virtual element methods*, Comput. Math. Appl., 66 (2013), pp. 376–391.

[2] P. F. Antonietti, A. Cangiani, J. Collis, Z. Dong, E. H. Georgoulis, S. Giani, and P. Houston, *Review of discontinuous Galerkin finite element methods for partial differential equations on complicated domains*, in Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations, Springer, Cham, 2016, pp. 281–310.

[3] A. V. Astaneh, F. Fuentes, J. Mora, and L. Demkowicz, *High-order polygonal discontinuous Petrov–Galerkin (POLYDPG) methods using ultraweak formulations*, Comput. Methods Appl. Mech. Engrg., 332 (2018), pp. 686–711.

[4] B. Ayuso de Dios, K. Lipnikov, and G. Manzini, *The nonconforming virtual element method*, ESAIM Math. Model. Numer. Anal., 50 (2016), pp. 879–904.

[5] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo, *Basic principles of virtual element methods*, Math. Models Methods Appl. Sci., 23 (2012), pp. 1–16.

[6] L. Beirão da Veiga, F. Brezzi, L. Marini, and A. Russo, *Virtual element method for general second-order elliptic problems on polygonal meshes*, Math. Models Methods Appl. Sci., 26 (2016), pp. 729–750.

[7] L. Beirão da Veiga, F. Brezzi, L. D. Marini, and A. Russo, *The hitchhiker's guide to the virtual element method*, Math. Models Methods Appl. Sci., 24 (2014), pp. 1541–1573.

[8] L. Beirão da Veiga, F. Dassi, and A. Russo, *High-order virtual element method on polyhedral meshes*, Comput. Math. Appl., 74 (2017), pp. 1110–1122.

[9] L. Beirao da Veiga, K. Lipnikov, and G. Manzini, *The Mimetic Finite Difference Method for Elliptic Problems*, MS&A. Model. Simul. Appl. 11, Springer, Cham, 2014.

[10] L. Beirão da Veiga, C. Lovadina, and A. Russo, *Stability analysis for the virtual element method*, Math. Models Methods Appl. Sci., 27 (2017), pp. 2557–2594.

[11] M. F. Benedetto, S. Berrone, S. Pieraccini, and S. Scialò, *The virtual element method for discrete fracture network simulations*, Comput. Methods Appl. Mech. Engrg., 280 (2014), pp. 135–156.

[12] S. Berrone and A. Borio, *Orthogonal polynomials in badly shaped polygonal elements for the virtual element method*, Finite Elem. Anal. Des., 129 (2017), pp. 14–31.

[13] J. Bonelle and A. Ern, *Analysis of compatible discrete operator schemes for elliptic problems on polyhedral meshes*, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 553–581.

[14] 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.

[15] F. Brezzi, A. Buffa, and K. Lipnikov, *Mimetic finite differences for elliptic problems*, ESAIM Math. Model. Numer. Anal., 43 (2009), pp. 277–295.

[16] F. Brezzi, R. S. Falk, and L. D. Marini, *Basic principles of mixed virtual element methods*, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 1227–1240.

[17] E. Burman, *Ghost penalty*, C. R. Math. Acad. Sci. Paris, 348 (2010), pp. 1217–1220.

[18] 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.

[19] A. Cangiani, G. Manzini, and O. J. Sutton, *Conforming and nonconforming virtual element methods for elliptic problems*, IMA J. Numer. Anal., 37 (2017), pp. 1317–1354.
[20] S. Cao and L. Chen, *Anisotropic error estimates of the linear virtual element method on polygonal meshes*, SIAM J. Numer. Anal., 56 (2018), pp. 2913–2939, https://doi.org/10.1137/17M1154369.
[21] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau, *An a priori error analysis of the local discontinuous Galerkin method for elliptic problems*, SIAM J. Numer. Anal., 38 (2000), pp. 1676–1706, https://doi.org/10.1137/S0036142900371003.
[22] 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.
[23] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Stud. Math. Appl. 4, North–Holland, Amsterdam, New York, Oxford, 1978.
[24] B. Cockburn, D. A. Di Pietro, and A. Ern, *Bridging the hybrid high-order and hybridizable discontinuous Galerkin methods*, ESAIM Math. Model. Numer. Anal., 50 (2016), pp. 635–650.
[25] M. Crouzeix and P.-A. Raviart, *Conforming and nonconforming finite element methods for solving the stationary Stokes equations* I, Rev. Française Automat. Informat. Recherche Opérationnelle. Sér. Rouge, 7 (1973), pp. 33–75.
[26] F. Dassi and L. Mascotto, *Exploring high-order three dimensional virtual elements: Bases and stabilizations*, Comput. Math. Appl., 75 (2018), pp. 3379–3401.
[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 S. Lemaire, *An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators*, Comput. Methods Appl. Math., 14 (2014), pp. 461–472.
[29] J. Dolbow and T. Belytschko, *A finite element method for crack growth without remeshing*, Internat. J. Numer. Methods Engrg., 46 (1999), pp. 131–150.
[30] A. Gillette, A. Rand, and C. Bajaj, *Error estimates for generalized barycentric interpolation*, Adv. Comput. Math., 37 (2012), pp. 417–439.
[31] J. Haslinger and Y. Renard, *A new fictitious domain approach inspired by the extended finite element method*, SIAM J. Numer. Anal., 47 (2009), pp. 1474–1499, https://doi.org/10.1137/070704435.
[32] J. Li, J. M. Melenk, B. Wohlmuth, and J. Zou, *Optimal a priori estimates for higher order finite elements for elliptic interface problems*, Appl. Numer. Math., 60 (2010), pp. 19–37.
[33] Z. Li, T. Lin, and X. Wu, *New Cartesian grid methods for interface problems using the finite element formulation*, Numer. Math., 96 (2003), pp. 61–98.
[34] X. Liu, J. Li, and Z. Chen, *A nonconforming virtual element method for the Stokes problem on general meshes*, Comput. Methods Appl. Mech. Engrg., 320 (2017), pp. 694–711.
[35] L. Mascotto, *Ill-conditioning in the virtual element method: Stabilizations and bases*, Numer. Methods Partial Differential Equations, 34 (2018), pp. 1258–1281.
[36] L. Mu, J. Wang, and X. Ye, *A weak Galerkin finite element method with polynomial reduction*, J. Comput. Appl. Math., 285 (2015), pp. 45–58.
[37] L. E. Payne and H. F. Weinberger, *An optimal Poincaré inequality for convex domains*, Arch. Rational Mech. Anal., 5 (1960), pp. 286–292.
[38] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Vol. 2, Princeton University Press, Princeton, NJ, 1970.
[39] N. Sukumar, N. Moës, B. Moran, and T. Belytschko, *Extended finite element method for three-dimensional crack modelling*, Internat. J. Numer. Methods Engrg., 48 (2000), pp. 1549–1570.
[40] J. Wang and X. Ye, *A weak Galerkin mixed finite element method for second order elliptic problems*, Math. Comp., 83 (2014), pp. 2101–2126.
[41] P. Wriggers, W. Rust, and B. Reddy, *A virtual element method for contact*, Comput. Mech., 58 (2016), pp. 1039–1050.