An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators

Daniele A. Di Pietro*[1], Alexandre Ern†[2], and Simon Lemaire‡[3]

[1] University of Montpellier 2, I3M, 34057 Montpellier CEDEX 5, France
[2]University of Paris-Est, CERMICS (ENPC), 6–8 avenue Blaise Pascal, 77455 Marne-la-Vallée CEDEX 2, France
[3]ENPC, Laboratoire Navier, 6–8 avenue Blaise Pascal, 77455 Marne-la-Vallée CEDEX 2, France

April 13, 2014

Abstract

We develop an arbitrary-order primal method for diffusion problems on general polyhedral meshes. The degrees of freedom are scalar-valued polynomials of the same order at mesh elements and faces. The cornerstone of the method is a local (element-wise) discrete gradient reconstruction operator. The design of the method additionally hinges on a least-squares penalty term on faces weakly enforcing the matching between local element- and face-based degrees of freedom. The scheme is proved to optimally converge in the energy norm and in the L^2-norm of the potential for smooth solutions. In the lowest-order case, equivalence with the Hybrid Finite Volume method is shown. The theoretical results are confirmed by numerical experiments up to order 4 on several polygonal meshes.

Keywords Diffusion, general meshes, arbitrary-order, gradient reconstruction

1 Introduction

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, denote an open bounded connected polygonal/polyhedral domain. In this work we propose an arbitrary-order primal hybrid method for the model diffusion problem

$$\begin{aligned} -\triangle u &= f \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega, \end{aligned} \tag{1}$$

*daniele.di-pietro@univ-montp2.fr
†ern@cermics.enpc.fr
‡Corresponding author; simon.lemaire87@gmail.com

where u denotes the potential and f a forcing term. More general boundary conditions and diffusion tensors could be considered, but we stick to the simpler case (1) for ease of presentation. For $X \subset \overline{\Omega}$, we respectively denote by $(\cdot,\cdot)_X$ and $\|\cdot\|_X$ the standard inner product and norm in $L^2(X)$, with the convention that the subscript is omitted whenever $X = \Omega$ and that the same notation is used in $L^2(X)^d$. Classically, the weak formulation of (1) consists, for $f \in L^2(\Omega)$, in seeking $u \in U_0 := H_0^1(\Omega)$ such that

$$(\nabla u, \nabla v) = (f, v) \qquad \forall v \in U_0. \tag{2}$$

Approximation methods on general polyhedral meshes have received an increasing attention over the last few years, motivated by applications (e.g., in geosciences) where the mesh cannot be easily adapted to the needs of the numerical scheme. To handle general discretizations, a wide range of new numerical methods has been developed. We can cite, e.g., the Mimetic Finite Difference (MFD) [5], the Hybrid Finite Volume (HFV) [14], and the Mixed Finite Volume (MFV) [11] methods, that have been proved to be closely related in [12]. Another example is the recent framework of Compatible Discrete Operator (CDO) schemes [4], for which a correspondence is established with nodal MFD discretizations for vertex-based CDO schemes and with MFV for cell-based CDO schemes. All these methods share the particularity of being lowest-order, which may be sufficient for most practical cases. However, the emphasis has been recently set on the design of higher-order discretizations capable of handling comparably general meshes. Results in this direction include the polygonal and extended Finite Element (FE) methods [16, 17] where nonpolynomial shape functions are considered. Furthermore, high-order MFD schemes have been recently analyzed in [3]. Even more recently, the Virtual Element Method (VEM) was introduced in [2], broadening the ideas underpinning the MFD approach and, at the same time, allowing one to design arbitrary-order conforming finite element methods on polyhedral meshes, without the need to specify the additional nonpolynomial shape functions. The VEM also allows for higher-order continuity conditions between neighboring elements.

In the present work, we show how the schemes based on local reconstruction operators originally developed in the context of linear elasticity in [9] apply to the design of an arbitrary-order primal (as opposed to the mixed case considered in [8]) method for the model diffusion problem (1). The resulting scheme can be viewed as a high-order extension of the HFV method, or of the generalized Crouzeix–Raviart method introduced in [10] in the context of linear elasticity. For a given polynomial degree $k \geqslant 0$, we select as degrees of freedom (DOFs) scalar-valued polynomials at mesh elements and faces up to degree k. The associated interpolation operator maps potentials in $H^1(\Omega)$ to their moments up to degree k at elements and faces. Then, the method is defined in two steps: (i) we devise a local discrete gradient reconstruction operator of order k in terms of the local DOFs by solving an inexpensive problem inside each element; (ii) we design a least-squares penalty term that is local to one element (i.e., it does not affect the stencil of the method), that weakly enforces the matching between local element- and face-based DOFs, and that preserves the order of the gradient reconstruction. A key ingredient in the design of the penalty term is a potential reconstruction of order $(k+1)$ obtained by correcting element DOFs by a higher-order term inferred from the discrete gradient reconstruction. A major difference with respect to VEM is that nodal unknowns are not present, which results in a more compact stencil (especially in three space dimensions). Additionally, as usual with the present choice of DOFs, static condensation allows one to solve a global system in terms of face unknowns only. The Hybrid Discontinuous Galerkin (HDG) method [6], which is devised on different ideas, also hinges on element- and face-based DOFs. The main difference is that the present choice of DOFs only involves scalar-valued polynomials attached to elements, which yields significant computational savings in the static condensation, especially in three space dimensions.

The paper is organized as follows. In Section 2 we introduce the method: we recall the notion of admissible mesh sequence, introduce the spaces of DOFs, derive the gradient reconstruction, present the discrete problem, study its well-posedness, and link our scheme to the HFV method in the lowest-order case. In Section 3 we perform the error analysis. We prove a convergence rate for smooth solutions of order $(k+1)$ in the energy norm and of order $(k+2)$ in the L^2-norm of the error, respectively. Incidentally, this latter result provides an optimal L^2-norm potential error estimate for HFV schemes with arbitrary penalty parameter (as opposed to [10], where a specific choice yielding continuity of the potential reconstruction at face barycenters is considered). Finally, in Section 4, we present numerical examples up to order 4 on various polygonal meshes confirming the theoretical predictions.

2 Description of the method

2.1 Admissible mesh sequences

We recall the notion of admissible mesh sequence of [7, Chapter 1]. Let $\mathcal{H} \subset \mathbb{R}^+_*$ denote a countable set of meshsizes having 0 as its unique accumulation point. We consider h-refined mesh sequences $(\mathcal{T}_h)_{h\in\mathcal{H}}$ where, for all $h \in \mathcal{H}$, $\mathcal{T}_h$ is a finite collection of nonempty disjoint open polygons/polyhedra (the elements) $\mathcal{T}_h = \{T\}$ such that $\overline{\Omega} = \bigcup_{T\in\mathcal{T}_h} \overline{T}$ and $h = \max_{T\in\mathcal{T}_h} h_T$ (h_T stands for the diameter of the element T). We call a face any hyperplanar closed connected subset F of $\overline{\Omega}$ with positive $(d-1)$-dimensional measure and such that (i) either there exist $T_1, T_2 \in \mathcal{T}_h$ such that $F \subset \partial T_1 \cap \partial T_2$ (and F is an interface) or (ii) there exists $T \in \mathcal{T}_h$ such that $F \subset \partial T \cap \partial\Omega$ (and F is a boundary face). Interfaces are collected in the set $\mathcal{F}_h^{\rm i}$, boundary faces in $\mathcal{F}_h^{\rm b}$, and we let $\mathcal{F}_h := \mathcal{F}_h^{\rm i} \cup \mathcal{F}_h^{\rm b}$. The diameter of a face $F \in \mathcal{F}_h$ is denoted by h_F. For all $T \in \mathcal{T}_h$, $\mathcal{F}_T := \{F \in \mathcal{F}_h \mid F \subset \partial T\}$ denotes the set of faces lying on the boundary of T and, for all $F \in \mathcal{F}_T$, $\boldsymbol{n}_{TF}$ is the unit normal to F pointing out of T. Finally, the l-dimensional Lebesgue measure, $0 \leqslant l \leqslant d$, is denoted by $|\cdot|_l$.

Definition 1 (Admissible mesh sequence)**.** *The mesh sequence $(\mathcal{T}_h)_{h\in\mathcal{H}}$ is admissible if, for all $h \in \mathcal{H}$, $\mathcal{T}_h$ admits a matching simplicial submesh $\mathfrak{T}_h$ and there exists a real number $\varrho > 0$ independent of h such that, for all $h \in \mathcal{H}$, (i) for all simplex $S \in \mathfrak{T}_h$ of diameter h_S and inradius r_S, $\varrho h_S \leqslant r_S$ holds and (ii) for all $T \in \mathcal{T}_h$, and all $S \in \mathfrak{T}_h$ such that $S \subset T$, $\varrho h_T \leqslant h_S$ holds.*

In what follows, we often abbreviate as $a \lesssim b$ the inequality $a \leqslant Cb$ with $C > 0$ independent of h but possibly depending on the mesh regularity parameter ϱ.

We next recall some basic results valid for admissible mesh sequences. First, according to [7, Lemma 1.42], for all $h \in \mathcal{H}$, all $T \in \mathcal{T}_h$, and all $F \in \mathcal{F}_T$, h_F is comparable to h_T in the sense that

$$\varrho^2 h_T \leqslant h_F \leqslant h_T. \tag{3}$$

Moreover, [7, Lemma 1.41] shows that there exists an integer N_∂ depending on ϱ and d such that

$$\forall h \in \mathcal{H}, \qquad \max_{T\in\mathcal{T}_h} \mathrm{card}(\mathcal{F}_T) \leqslant N_\partial. \tag{4}$$

There also exist real numbers C_{tr} and $C_{\mathrm{tr,c}}$ depending on ϱ but independent of h such that the following discrete and continuous trace inequalities hold for all $T \in \mathcal{T}_h$ and $F \in \mathcal{F}_T$, cf. [7, Lemmata 1.46 and 1.49]:

$$\|v\|_F \leqslant C_{\mathrm{tr}} h_F^{-1/2} \|v\|_T \qquad \forall v \in \mathbb{P}_d^l(T), \tag{5}$$

$$\|v\|_{\partial T} \leqslant C_{\mathrm{tr,c}} \left(h_T^{-1}\|v\|_T^2 + h_T \|\boldsymbol{\nabla} v\|_T^2\right)^{1/2} \qquad \forall v \in H^1(T), \tag{6}$$

where, for X being an n-dimensional subset of $\overline{\Omega}$ ($n \leqslant d$), $\mathbb{P}_n^l(X)$ is spanned by the restrictions to X of n-variate polynomials of total degree $\leqslant l$. Using [7, Lemma 1.40] together with the results of [13], one can prove that there exists a real number C_{app} depending on ϱ and l but independent of h such that, for all $T \in \mathcal{T}_h$, denoting by π_T^l the L^2-orthogonal projector on $\mathbb{P}_d^l(T)$, the following holds: For all $s \in \{1, \dots, l+1\}$, and all $v \in H^s(T)$,

$$|v - \pi_T^l v|_{H^m(T)} + h_T^{1/2} |v - \pi_T^l v|_{H^m(\partial T)} \leqslant C_{\mathrm{app}} h_T^{s-m} |v|_{H^s(T)} \qquad \forall m \in \{0, \dots, (s-1)\}. \tag{7}$$

Finally, the following Poincaré inequality is valid for all $T \in \mathcal{T}_h$ and all $v \in H^1(T)$ such that $\int_T v = 0$:

$$\|v\|_T \leqslant C_{\mathrm{P}} h_T \|\boldsymbol{\nabla} v\|_T, \tag{8}$$

where $C_{\mathrm{P}} = \pi^{-1}$ for convex elements (cf. [1]). For more general element shapes, C_{P} can be estimated in terms of ϱ.

2.2 Degrees of freedom

Let a polynomial degree $k \geqslant 0$ be fixed. For all $T \in \mathcal{T}_h$, we define the local space of DOFs as follows:

$$\mathsf{U}_T^k := \mathbb{P}_d^k(T) \times \left\{ \bigtimes_{F\in\mathcal{F}_T} \mathbb{P}_{d-1}^k(F) \right\}. \tag{9}$$

The global space of DOFs is obtained by patching interface values in (9):

$$\mathsf{U}_h^k := \left\{ \bigtimes_{T\in\mathcal{T}_h} \mathbb{P}_d^k(T) \right\} \times \left\{ \bigtimes_{F\in\mathcal{F}_h} \mathbb{P}_{d-1}^k(F) \right\}. \tag{10}$$

Boundary conditions can be embedded in the discrete space (10) by letting

$$\mathsf{U}_{h,0}^k := \left\{\mathsf{v}_h = ((\mathsf{v}_T)_{T\in\mathcal{T}_h}, (\mathsf{v}_F)_{F\in\mathcal{F}_h}) \in \mathsf{U}_h^k \mid \mathsf{v}_F \equiv 0\ \forall F \in \mathcal{F}_h^{\mathrm{b}}\right\}. \tag{11}$$

For all $T \in \mathcal{T}_h$, we denote by $\mathsf{L}_T : \mathsf{U}_h^k \to \mathsf{U}_T^k$ the restriction operator that maps the global DOFs in U_h^k to the corresponding local DOFs in U_T^k. The local interpolation operator $\mathsf{I}_T^k : H^1(T) \to \mathsf{U}_T^k$ is such that, for all $v \in H^1(T)$,

$$\mathsf{I}_T^k v := (\pi_T^k v, (\pi_F^k v)_{F\in\mathcal{F}_T}), \tag{12}$$

where π_F^k is the L^2-orthogonal projector on $\mathbb{P}_{d-1}^k(F)$. The corresponding global interpolation operator $\mathsf{I}_h^k : H^1(\Omega) \to \mathsf{U}_h^k$ is such that, for all $v \in H^1(\Omega)$,

$$\mathsf{I}_h^k v := ((\pi_T^k v)_{T\in\mathcal{T}_h}, (\pi_F^k v)_{F\in\mathcal{F}_h}). \tag{13}$$

When applied to functions in U_0, I_h^k maps onto $\mathsf{U}_{h,0}^k$.

2.3 Local gradient reconstruction

For all $T \in \mathcal{T}_h$, we define the local gradient reconstruction operator $\boldsymbol{G}_T^k : \mathsf{U}_T^k \to \boldsymbol{\nabla}\mathbb{P}_d^{k+1,0}(T)$ (where, for $l \geqslant 1$, $\mathbb{P}_d^{l,0}(T)$ stands for the space of d-variate polynomial functions of total degree $\leqslant l$ that have zero average on T) such that, for all $\mathsf{v} := (\mathsf{v}_T, (\mathsf{v}_F)_{F\in\mathcal{F}_T}) \in \mathsf{U}_T^k$ and all $w \in \mathbb{P}_d^{k+1,0}(T)$,

$$(\boldsymbol{G}_T^k\mathsf{v}, \boldsymbol{\nabla} w)_T = (\boldsymbol{\nabla}\mathsf{v}_T, \boldsymbol{\nabla} w)_T + \sum_{F\in\mathcal{F}_T} (\mathsf{v}_F - \mathsf{v}_T, \boldsymbol{\nabla} w\cdot\boldsymbol{n}_{TF})_F. \tag{14}$$

Since $\boldsymbol{G}_T^k\mathsf{v} \in \boldsymbol{\nabla}\mathbb{P}_d^{k+1,0}(T)$ means that there is $v \in \mathbb{P}_d^{k+1,0}(T)$ such that $\boldsymbol{G}_T^k\mathsf{v} = \boldsymbol{\nabla} v$, (14) corresponds to the local (well-posed) Neumann problem

$$(\boldsymbol{\nabla} v, \boldsymbol{\nabla} w)_T = (\boldsymbol{\nabla}\mathsf{v}_T, \boldsymbol{\nabla} w)_T + \sum_{F\in\mathcal{F}_T} (\mathsf{v}_F - \mathsf{v}_T, \boldsymbol{\nabla} w\cdot\boldsymbol{n}_{TF})_F. \tag{15}$$

Solving (15) requires to invert the local stiffness matrix inside each element, which can be performed effectively via a Cholesky factorization.

Remark 2 (Compatibility condition)**.** *Observing that the right-hand side of the (discrete) Neumann problem* (15) *satisfies the usual compatibility condition for test functions in* $\mathbb{P}_d^0(T)$*, we infer that* (14) *and* (15) *hold in fact for all* $w \in \mathbb{P}_d^{k+1}(T)$*.*

We next introduce the potential reconstruction operator $p_T^k : \mathsf{U}_T^k \to \mathbb{P}_d^{k+1}(T)$ such that, for all $\mathsf{v} \in \mathsf{U}_T^k$,

$$\boldsymbol{\nabla} p_T^k\mathsf{v} := \boldsymbol{G}_T^k\mathsf{v}, \qquad \int_T p_T^k\mathsf{v} := \int_T \mathsf{v}_T. \tag{16}$$

Lemma 3 (Approximation properties for $p_T^k \mathsf{I}_T^k$)**.** *There exists a real number $C > 0$, depending on ϱ but independent of h_T such that, for all $v \in H^{k+2}(T)$,*

$$\|v - p_T^k \mathsf{I}_T^k v\|_T + h_T^{1/2} \|v - p_T^k \mathsf{I}_T^k v\|_{\partial T} \\ + h_T \|\boldsymbol{\nabla}(v - p_T^k \mathsf{I}_T^k v)\|_T + h_T^{3/2} \|\boldsymbol{\nabla}(v - p_T^k \mathsf{I}_T^k v)\|_{\partial T} \leqslant C h_T^{k+2} \|v\|_{H^{k+2}(T)}. \quad (17)$$

Proof. Let $v \in H^{k+2}(T)$. Integrating by parts the right-hand side of (14) and using the definition (16) of the operator p_T^k together with the definition (12) of the interpolation operator I_T^k yields, for all $w \in \mathbb{P}_d^{k+1}(T)$ (cf. Remark 2),

$$(\boldsymbol{\nabla} p_T^k \mathsf{I}_T^k v, \boldsymbol{\nabla} w)_T = (\boldsymbol{G}_T^k \mathsf{I}_T^k v, \boldsymbol{\nabla} w)_T = -(\pi_T^k v, \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla} w))_T + \sum_{F \in \mathcal{F}_T} (\pi_F^k v, \boldsymbol{\nabla} w \cdot \boldsymbol{n}_{TF})_F \\ = -(v, \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla} w))_T + \sum_{F \in \mathcal{F}_T} (v, \boldsymbol{\nabla} w \cdot \boldsymbol{n}_{TF})_F,$$

since $\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla} w) \in \mathbb{P}_d^{k-1}(T) \subset \mathbb{P}_d^k(T)$ and $\boldsymbol{\nabla} w_{|F} \cdot \boldsymbol{n}_{TF} \in \mathbb{P}_{d-1}^k(F)$. Performing a second integration by parts leads to

$$(\boldsymbol{\nabla} v - \boldsymbol{\nabla} p_T^k \mathsf{I}_T^k v, \boldsymbol{\nabla} w)_T = 0 \qquad \forall w \in \mathbb{P}_d^{k+1}(T). \quad (18)$$

The orthogonality condition (18) implies that

$$\|\boldsymbol{\nabla}(v - p_T^k \mathsf{I}_T^k v)\|_T = \inf_{z \in \mathbb{P}_d^{k+1}(T)} \|\boldsymbol{\nabla} v - \boldsymbol{\nabla} z\|_T \lesssim h_T^{k+1} \|v\|_{H^{k+2}(T)}, \quad (19)$$

where we have used the approximation property (7) of π_T^{k+1} (with $s = k+2$ and $m = 1$). Then, from (19), the fact that $\int_T p_T^k \mathsf{I}_T^k v = \int_T \pi_T^k v = \int_T v$ owing to the second relation in (16) and the definition (12) of the local interpolation operator, and the Poincaré inequality (8), we infer that

$$\|v - p_T^k \mathsf{I}_T^k v\|_T \lesssim h_T^{k+2} \|v\|_{H^{k+2}(T)}. \quad (20)$$

The consecutive use of the continuous trace inequality (6) and (19)-(20) yields

$$h_T \|v - p_T^k \mathsf{I}_T^k v\|_{\partial T}^2 \lesssim \|v - p_T^k \mathsf{I}_T^k v\|_T^2 + h_T^2 \|\boldsymbol{\nabla}(v - p_T^k \mathsf{I}_T^k v)\|_T^2 \lesssim h_T^{2(k+2)} \|v\|_{H^{k+2}(T)}^2.$$

Finally, the bound on $h_T^{3/2} \|\boldsymbol{\nabla}(v - p_T^k \mathsf{I}_T^k v)\|_{\partial T}$ is obtained by introducing $\pm \pi_T^k \boldsymbol{\nabla} v$ inside the norm, using the triangle inequality, and concluding with the approximation property (7) of π_T^k (applied componentwise to $\boldsymbol{\nabla} v$ with $s = k+1$ and $m = 0$), the discrete trace inequality (5), the bound (4) on card($\mathcal{F}_T$), the mesh regularity property (3), the fact that $\boldsymbol{\nabla} p_T^k \mathsf{I}_T^k v \in [\mathbb{P}_d^k(T)]^d$ so that $\|\pi_T^k \boldsymbol{\nabla} v - \boldsymbol{\nabla} p_T^k \mathsf{I}_T^k v\|_T \leqslant \|\boldsymbol{\nabla}(v - p_T^k \mathsf{I}_T^k v)\|_T$, and (19). □

2.4 Discrete problem and well-posedness

To discretize the left-hand side of (2), we introduce the following bilinear forms on $\mathsf{U}_h^k \times \mathsf{U}_h^k$:

$$a_h(\mathsf{u}_h, \mathsf{v}_h) := \sum_{T \in \mathcal{T}_h} a_T(\mathsf{L}_T \mathsf{u}_h, \mathsf{L}_T \mathsf{v}_h), \qquad s_h(\mathsf{u}_h, \mathsf{v}_h) := \sum_{T \in \mathcal{T}_h} s_T(\mathsf{L}_T \mathsf{u}_h, \mathsf{L}_T \mathsf{v}_h), \quad (21)$$

where, for all $T \in \mathcal{T}_h$, the local bilinear forms a_T and s_T on $\mathsf{U}_T^k \times \mathsf{U}_T^k$ are such that

$$a_T(\mathsf{u},\mathsf{v}) := (\boldsymbol{G}_T^k \mathsf{u}, \boldsymbol{G}_T^k \mathsf{v})_T + s_T(\mathsf{u},\mathsf{v}), \quad s_T(\mathsf{u},\mathsf{v}) := \sum_{F\in\mathcal{F}_T} \frac{1}{h_F} (\pi_F^k(\mathsf{u}_F - P_T^k\mathsf{u}), \pi_F^k(\mathsf{v}_F - P_T^k\mathsf{v}))_F, \tag{22}$$

where the local potential reconstruction $P_T^k : \mathsf{U}_T^k \to \mathbb{P}_d^{k+1}(T)$ is defined such that, for all $\mathsf{v} \in \mathsf{U}_T^k$,

$$P_T^k\mathsf{v} := \mathsf{v}_T + (p_T^k\mathsf{v} - \pi_T^k p_T^k \mathsf{v}). \tag{23}$$

The term in parentheses can be interpreted as a higher-order correction of the element unknown v_T derived from the discrete gradient reconstruction operator (14) (note that this correction is independent of the second relation in (16)). The stabilization bilinear form s_T defined by (22) introduces a least-squares penalty of the L^2-orthogonal projection on $\mathbb{P}_{d-1}^k(F)$ of the difference between v_F and $(P_T^k\mathsf{v})_{|F}$, cf. Remark 6 below. Introducing the global discrete gradient operator $\boldsymbol{G}_h^k : \mathsf{U}_h^k \to \times_{T\in\mathcal{T}_h} \boldsymbol{\nabla}\mathbb{P}_d^{k+1,0}(T)$ such that, for all $\mathsf{v}_h \in \mathsf{U}_h^k$,

$$(\boldsymbol{G}_h^k \mathsf{v}_h)_{|T} := \boldsymbol{G}_T^k \mathsf{L}_T \mathsf{v}_h \qquad \forall T \in \mathcal{T}_h, \tag{24}$$

we can reformulate the bilinear form a_h defined in (21) as

$$a_h(\mathsf{u}_h, \mathsf{v}_h) = (\boldsymbol{G}_h^k \mathsf{u}_h, \boldsymbol{G}_h^k \mathsf{v}_h) + s_h(\mathsf{u}_h, \mathsf{v}_h).$$

We define the local and global energy semi-norms as follows:

$$\|\mathsf{v}\|_{a,T}^2 := a_T(\mathsf{v},\mathsf{v}) \quad \forall \mathsf{v} \in \mathsf{U}_T^k, \qquad \|\mathsf{v}_h\|_{a,h}^2 := \sum_{T\in\mathcal{T}_h} \|\mathsf{L}_T\mathsf{v}_h\|_{a,T}^2 \quad \forall \mathsf{v}_h \in \mathsf{U}_h^k, \tag{25}$$

and observe that, owing to (21),

$$a_h(\mathsf{v}_h, \mathsf{v}_h) = \|\mathsf{v}_h\|_{a,h}^2. \tag{26}$$

The forcing term in (2) is discretized by means of the linear form on U_h^k such that

$$l_h(\mathsf{v}_h) := \sum_{T\in\mathcal{T}_h} (f, \mathsf{v}_T)_T. \tag{27}$$

The discrete problem reads: Find $\mathsf{u}_h \in \mathsf{U}_{h,0}^k$ such that, for all $\mathsf{v}_h \in \mathsf{U}_{h,0}^k$,

$$a_h(\mathsf{u}_h, \mathsf{v}_h) = l_h(\mathsf{v}_h). \tag{28}$$

The stability of the method is expressed in terms of the following $H_0^1(\Omega)$-like discrete norm:

$$\|\mathsf{v}_h\|_{1,h}^2 := \sum_{T\in\mathcal{T}_h} \|\mathsf{L}_T\mathsf{v}_h\|_{1,T}^2 \quad \forall \mathsf{v}_h \in \mathsf{U}_{h,0}^k, \qquad \|\mathsf{v}\|_{1,T}^2 := \|\boldsymbol{\nabla}\mathsf{v}_T\|_T^2 + |\mathsf{v}|_{1,\partial T}^2 \quad \forall \mathsf{v} \in \mathsf{U}_T^k, \tag{29}$$

where $|\mathsf{v}|_{1,\partial T}^2 := \sum_{F\in\mathcal{F}_T} h_F^{-1} \|\mathsf{v}_F - \mathsf{v}_T\|_F^2$. Since the homogeneous Dirichlet boundary condition is embedded in the discrete space $\mathsf{U}_{h,0}^k$ defined in (11), the map $\|\cdot\|_{1,h}$ defines a norm on $\mathsf{U}_{h,0}^k$. Using a discrete Poincaré inequality in broken polynomial spaces, see [7, Corollary 5.4] and references therein, it is possible to show that for all $\mathsf{v}_h \in \mathsf{U}_{h,0}^k$, $\|v_h\| \lesssim \|\mathsf{v}_h\|_{1,h}$, where v_h is the piecewise polynomial function such that $v_{h|T} := \mathsf{v}_T$, for all $T \in \mathcal{T}_h$.

Lemma 4 (Norm equivalence). *There exists $\eta > 0$ such that, for all $T \in \mathcal{T}_h$ and all $\mathsf{v} \in \mathsf{U}_T^k$,*

$$\eta^{-1}\|\mathsf{v}\|_{1,T}^2 \leqslant \|\mathsf{v}\|_{a,T}^2 \leqslant \eta\|\mathsf{v}\|_{1,T}^2. \tag{30}$$

Consequently, for all $\mathsf{v}_h \in \mathsf{U}_h^k$,

$$\eta^{-1}\|\mathsf{v}_h\|_{1,h}^2 \leqslant \|\mathsf{v}_h\|_{a,h}^2 \leqslant \eta\|\mathsf{v}_h\|_{1,h}^2. \tag{31}$$

Proof. Let $T \in \mathcal{T}_h$ and $\mathsf{v} \in \mathsf{U}_T^k$. Taking $w = \mathsf{v}_T \in \mathbb{P}_d^k(T) \subset \mathbb{P}_d^{k+1}(T)$ in (14) (cf. Remark 2) yields

$$\|\nabla \mathsf{v}_T\|_T^2 = (\boldsymbol{G}_T^k\mathsf{v}, \nabla \mathsf{v}_T)_T - \sum_{F\in\mathcal{F}_T}(\mathsf{v}_F - \mathsf{v}_T, \nabla \mathsf{v}_T\cdot\boldsymbol{n}_{TF})_F \leqslant \|\boldsymbol{G}_T^k\mathsf{v}\|_T^2 + \frac{1}{2}\|\nabla \mathsf{v}_T\|_T^2 + N_\partial C_{\mathrm{tr}}^2|\mathsf{v}|_{1,\partial T}^2, \tag{32}$$

where we have used the Cauchy–Schwarz and Young's inequalities, and applied the discrete trace inequality (5) and the bound (4) on $\mathrm{card}(\mathcal{F}_T)$ for the face term. Owing to (32), we infer that

$$\|\nabla \mathsf{v}_T\|_T^2 \lesssim \|\boldsymbol{G}_T^k\mathsf{v}\|_T^2 + |\mathsf{v}|_{1,\partial T}^2. \tag{33}$$

Let now $F \in \mathcal{F}_T$. Adding and subtracting $\pi_F^k P_T^k\mathsf{v}$, and using the triangle inequality yields

$$h_F^{-1/2}\|\mathsf{v}_F - \mathsf{v}_T\|_F \leqslant h_F^{-1/2}\|\pi_F^k(\mathsf{v}_F - P_T^k\mathsf{v})\|_F + h_F^{-1/2}\|\pi_F^k(p_T^k\mathsf{v} - \pi_T^k p_T^k\mathsf{v})\|_F, \tag{34}$$

where we have used the fact that both v_F and $\mathsf{v}_{T|F}$ belong to $\mathbb{P}_{d-1}^k(F)$ together with the definition (23) of P_T^k. The second term on the right-hand side of (34) can be estimated as

$$h_F^{-1/2}\|\pi_F^k(p_T^k\mathsf{v} - \pi_T^k p_T^k\mathsf{v})\|_F \leqslant C_{\mathrm{tr}} h_F^{-1}\|p_T^k\mathsf{v} - \pi_T^k p_T^k\mathsf{v}\|_T \leqslant C_{\mathrm{tr}} C_{\mathrm{app}}\varrho^{-2}\|\boldsymbol{G}_T^k\mathsf{v}\|_T,$$

where we have used the discrete trace inequality (5), the approximation property (7) of π_T^k (with $s = 1$ and $m = 0$), the definition (16) of $\nabla p_T^k\mathsf{v}$, and the mesh regularity property (3). Hence,

$$h_F^{-1/2}\|\mathsf{v}_F - \mathsf{v}_T\|_F \lesssim h_F^{-1/2}\|\pi_F^k(\mathsf{v}_F - P_T^k\mathsf{v})\|_F + \|\boldsymbol{G}_T^k\mathsf{v}\|_T. \tag{35}$$

Squaring (35), summing over $F \in \mathcal{F}_T$, and using the bound (4) on $\mathrm{card}(\mathcal{F}_T)$ yields

$$|\mathsf{v}|_{1,\partial T}^2 \lesssim s_T(\mathsf{v}, \mathsf{v}) + \|\boldsymbol{G}_T^k\mathsf{v}\|_T^2. \tag{36}$$

The first inequality in (30) follows from (33) and (36). Turning to the second inequality, we deduce from (14), the Cauchy–Schwarz inequality, the discrete trace inequality (5), and the bound (4) on $\mathrm{card}(\mathcal{F}_T)$ that

$$\|\boldsymbol{G}_T^k\mathsf{v}\|_T = \sup_{w\in\mathbb{P}_d^{k+1,0}(T)} \frac{(\boldsymbol{G}_T^k\mathsf{v}, \nabla w)_T}{\|\nabla w\|_T} \leqslant \|\nabla \mathsf{v}_T\|_T + \sum_{F\in\mathcal{F}_T} C_{\mathrm{tr}} h_F^{-1/2}\|\mathsf{v}_F - \mathsf{v}_T\|_F \lesssim \|\mathsf{v}\|_{1,T}.$$

Moreover, for all $F \in \mathcal{F}_T$, the triangle inequality and the fact that π_F^k is a projector yield

$$h_F^{-1/2}\|\pi_F^k(\mathsf{v}_F - P_T^k\mathsf{v})\|_F \leqslant h_F^{-1/2}\|\mathsf{v}_F - \mathsf{v}_T\|_F + h_F^{-1/2}\|\mathsf{v}_T - P_T^k\mathsf{v}\|_F,$$

and, owing to the definition (23) of P_T^k, the second term on the right-hand side is equal to $h_F^{-1/2}\|p_T^k \mathsf{v} - \pi_T^k p_T^k \mathsf{v}\|_F$. Using the mesh regularity property (3), the discrete trace inequality (5), and the approximation property (7) of π_T^k (with $s = 1$ and $m = 0$), we infer that

$$h_F^{-1/2}\|p_T^k \mathsf{v} - \pi_T^k p_T^k \mathsf{v}\|_F \lesssim h_T^{-1}\|p_T^k \mathsf{v} - \pi_T^k p_T^k \mathsf{v}\|_T \lesssim \|\nabla p_T^k \mathsf{v}\|_T = \|\boldsymbol{G}_T^k \mathsf{v}\|_T.$$

Combining the above bounds yields the second inequality in (30). Finally, summing (30) over $T \in \mathcal{T}_h$ proves (31). □

Corollary 5 (Well-posedness)**.** *Problem* (28) *is well-posed.*

Proof. Combining (26) with Lemma 4 yields $a_h(\mathsf{v}_h, \mathsf{v}_h) = \|\mathsf{v}_h\|_{a,h}^2 \geqslant \eta^{-1}\|\mathsf{v}_h\|_{1,h}^2$. Since $\|\cdot\|_{1,h}$ is a norm on $\mathsf{U}_{h,0}^k$, the well-posedness results from the Lax–Milgram Lemma. □

Remark 6 (Stabilization)**.** *The design of the local stabilization bilinear form s_T is tailored to ensure control of the $\|\cdot\|_{1,T}$-norm as reflected by the first inequality in* (30)*, and, at the same time, yield the same convergence order as the gradient reconstruction, cf. the proof of Theorem 8 below, in particular* (46)*. This is the reason why s_T is not set to be, e.g., $s_T(\mathsf{u}, \mathsf{v}) := \sum_{F\in\mathcal{F}_T} h_F^{-1}(\mathsf{u}_F - \mathsf{u}_T, \mathsf{v}_F - \mathsf{v}_T)_F$ (this choice trivially ensures control of the $|\cdot|_{1,\partial T}$-seminorm), but the (projections of the) high-order potential reconstructions $\pi_F^k P_T^k \mathsf{u}$ and $\pi_F^k P_T^k \mathsf{v}$ are used in place of u_T and v_T, respectively.*

2.5 Link with the HFV method for $k = 0$

In the lowest-order case ($k = 0$), the proposed method shares strong links with the HFV method of [14]. We assume in this section that, for all $h \in \mathcal{H}$ and all $T \in \mathcal{T}_h$, T is star-shaped with respect to its barycenter $\boldsymbol{x}_T$ and, for all $F \in \mathcal{F}_T$, one has

$$\tilde{\varrho} h_T \leqslant d_{T,F} \leqslant h_T, \tag{37}$$

where $d_{T,F}$ denotes the orthogonal distance between $\boldsymbol{x}_T$ and F and $\tilde{\varrho} > 0$ is a mesh regularity parameter independent of h. The HFV discretization of problem (1) reads: Find $\mathsf{u}_h \in \mathsf{U}_{h,0}^0$ such that, for all $\mathsf{v}_h \in \mathsf{U}_{h,0}^0$,

$$\sum_{T\in\mathcal{T}_h} |T|_d(\mathcal{G}_T \mathsf{L}_T \mathsf{u}_h)\cdot(\mathcal{G}_T \mathsf{L}_T \mathsf{v}_h) + \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} |T_F|_d(\mathcal{R}_{T,F}\mathsf{L}_T \mathsf{u}_h)(\mathcal{R}_{T,F}\mathsf{L}_T \mathsf{v}_h) = \sum_{T\in\mathcal{T}_h} |T|_d f_T \mathsf{v}_T, \tag{38}$$

where T_F denotes the pyramid of base F and apex $\boldsymbol{x}_T$ such that $|T_F|_d = |F|_{d-1} d_{T,F}/d$, $f_T := |T|_d^{-1}\int_T f$ for all $T \in \mathcal{T}_h$ and, for all $\mathsf{v} \in \mathsf{U}_T^0$, denoting by $\boldsymbol{x}_F$ the barycenter of $F \in \mathcal{F}_T$,

$$\mathcal{G}_T \mathsf{v} := \frac{1}{|T|_d}\sum_{F\in\mathcal{F}_T} |F|_{d-1}(\mathsf{v}_F - \mathsf{v}_T)\boldsymbol{n}_{TF}, \qquad \mathcal{R}_{T,F}\mathsf{v} := \frac{d^{1/2}}{d_{T,F}}(\mathsf{v}_F - \mathsf{v}_T - \mathcal{G}_T \mathsf{v}\cdot(\boldsymbol{x}_F - \boldsymbol{x}_T)). \tag{39}$$

Proposition 7 (Equivalence with the HFV discretization for $k = 0$)**.** *The discrete problem* (38) *coincides with* (28) *for $k = 0$ and $\boldsymbol{x}_T$ the barycenter of T, up to the (uniformly comparable) change of scaling $d_{T,F} \leftarrow h_F$ in the penalty term $\mathcal{R}_{T,F}$ defined by* (39).

Proof. To derive an explicit expression for $\boldsymbol{G}_T^0\mathsf{v}$, we notice that

$$\boldsymbol{\nabla}\mathbb{P}_d^{1,0} = [\mathbb{P}_d^0]^d = \operatorname{span}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\},$$

where $(\boldsymbol{e}_i)_{i\in\{1,\ldots,d\}}$ denotes the canonical basis of $\mathbb{R}^d$. Then, for any $T \in \mathcal{T}_h$, testing (14) with $w_i \in \mathbb{P}_d^{1,0}(T)$ such that $\boldsymbol{\nabla} w_i = \boldsymbol{e}_i$ for all $i \in \{1, \ldots, d\}$, it is straightforward that, for any $\mathsf{v} \in \mathsf{U}_T^0$, $\boldsymbol{G}_T^0\mathsf{v} \equiv \mathcal{G}_T\mathsf{v}$. Additionally, for all $\boldsymbol{x} \in T$, $p_T^0\mathsf{v}(\boldsymbol{x}) \equiv \mathcal{G}_T\mathsf{v}\cdot(\boldsymbol{x} - \boldsymbol{x}_T) + |T|_d^{-1}\int_T \mathsf{v}_T \in \mathbb{P}_d^1(T)$, and hence $P_T^0\mathsf{v}(\boldsymbol{x}) \equiv \mathsf{v}_T + \mathcal{G}_T\mathsf{v}\cdot(\boldsymbol{x} - \boldsymbol{x}_T)$, whose restriction to F belongs to $\mathbb{P}_{d-1}^1(F)$. As a consequence, for all $F \in \mathcal{F}_T$, we infer that $\pi_F^0(\mathsf{v}_F - P_T^0\mathsf{v}) = \mathsf{v}_F - \mathsf{v}_T - \mathcal{G}_T\mathsf{v}\cdot(\boldsymbol{x}_F - \boldsymbol{x}_T) = d_{T,F}d^{-1/2}\mathcal{R}_{T,F}\mathsf{v}$. Plugging these expressions into (21), (22), and (27), and comparing (28) with (38), we observe that the only difference between the two discretizations lies in the scaling choice for the least-squares penalty term (the scaling is $d_{T,F}^{-1}$ in (38) and h_F^{-1} in (28)). The two choices are uniformly comparable owing to (37) and (3). □

3 Error analysis

3.1 Energy-norm error estimate

Theorem 8 (Discrete error estimate)**.** *Let $u \in U_0$ and $\mathsf{u}_h \in \mathsf{U}_{h,0}^k$ denote the unique solutions to* (2) *and* (28) *respectively, and assume the additional regularity $u \in H^{k+2}(\Omega)$. Then, letting $\widehat{\mathsf{u}}_h := \mathsf{I}_h^k u$, there exists a real number $C > 0$ depending on ϱ but independent of h such that*

$$\eta^{-1/2}\|\widehat{\mathsf{u}}_h - \mathsf{u}_h\|_{1,h} \leqslant \|\widehat{\mathsf{u}}_h - \mathsf{u}_h\|_{a,h} \leqslant Ch^{k+1}\|u\|_{H^{k+2}(\Omega)}. \tag{40}$$

Proof. The first inequality in (40) results from the first inequality in (31). Moreover, using (26), (31), and the fact that $\widehat{\mathsf{u}}_h - \mathsf{u}_h \in \mathsf{U}_{h,0}^k$ yields

$$\|\widehat{\mathsf{u}}_h - \mathsf{u}_h\|_{a,h} \leqslant \eta^{1/2}\frac{a_h(\widehat{\mathsf{u}}_h - \mathsf{u}_h, \widehat{\mathsf{u}}_h - \mathsf{u}_h)}{\|\widehat{\mathsf{u}}_h - \mathsf{u}_h\|_{1,h}} \leqslant \eta^{1/2} \sup_{\mathsf{v}_h\in\mathsf{U}_{h,0}^k,\,\|\mathsf{v}_h\|_{1,h}=1} a_h(\widehat{\mathsf{u}}_h - \mathsf{u}_h, \mathsf{v}_h).$$

Owing to (28), we infer that

$$\|\widehat{\mathsf{u}}_h - \mathsf{u}_h\|_{a,h} \leqslant \eta^{1/2} \sup_{\mathsf{v}_h\in\mathsf{U}_{h,0}^k,\,\|\mathsf{v}_h\|_{1,h}=1} \mathcal{E}_h(\mathsf{v}_h), \tag{41}$$

where $\mathcal{E}_h(\mathsf{v}_h) := a_h(\widehat{\mathsf{u}}_h, \mathsf{v}_h) - l_h(\mathsf{v}_h)$ is the consistency error. We derive a bound for this quantity for a generic $\mathsf{v}_h \in \mathsf{U}_{h,0}^k$. Recalling that $f = -\triangle u$ a.e. in Ω, an element-wise integration by parts in (27) yields

$$l_h(\mathsf{v}_h) = \sum_{T\in\mathcal{T}_h}(\boldsymbol{\nabla} u, \boldsymbol{\nabla}\mathsf{v}_T)_T + \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}(\mathsf{v}_F - \mathsf{v}_T, \boldsymbol{\nabla} u\cdot\boldsymbol{n}_{TF})_F, \tag{42}$$

where we have used the fact that the flux is continuous at interfaces and that the homogeneous Dirichlet boundary condition is embedded in $\mathsf{U}_{h,0}^k$ (cf. (11)) to introduce v_F in the second term on the right-hand side of (42). Choosing $w = \breve{u}_T := p_T^k \mathsf{L}_T \widehat{\mathsf{u}}_h = p_T^k \mathsf{I}_T^k(u_{|T})$ in the definition (14) of $\boldsymbol{G}_T^k \mathsf{L}_T \mathsf{v}_h$ for all $T \in \mathcal{T}_h$ (recall from Remark 2 that the zero-mean condition is not needed on w), and owing to (16) and (24), we infer that

$$(\boldsymbol{G}_h^k \widehat{\mathsf{u}}_h, \boldsymbol{G}_h^k \mathsf{v}_h) = \sum_{T \in \mathcal{T}_h} (\boldsymbol{\nabla} \breve{u}_T, \boldsymbol{\nabla} \mathsf{v}_T)_T + \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} (\mathsf{v}_F - \mathsf{v}_T, \boldsymbol{\nabla} \breve{u}_T \cdot \boldsymbol{n}_{TF})_F. \tag{43}$$

Combining (43) with (42), we arrive at

$$\begin{aligned} \mathcal{E}_h(\mathsf{v}_h) = \sum_{T \in \mathcal{T}_h} (\boldsymbol{\nabla}(\breve{u}_T - u), \boldsymbol{\nabla} \mathsf{v}_T)_T + \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} (\mathsf{v}_F - \mathsf{v}_T, (\boldsymbol{\nabla} \breve{u}_T - \boldsymbol{\nabla} u) \cdot \boldsymbol{n}_{TF})_F \\ + \sum_{T \in \mathcal{T}_h} s_T(\mathsf{L}_T \widehat{\mathsf{u}}_h, \mathsf{L}_T \mathsf{v}_h) := \mathfrak{T}_1 + \mathfrak{T}_2 + \mathfrak{T}_3. \end{aligned} \tag{44}$$

To estimate $\mathfrak{T}_1$ and $\mathfrak{T}_2$, we use the Cauchy–Schwarz inequality followed by the approximation property (17) of $p_T^k \mathsf{I}_T^k$ (and also the bound (4) on $\mathrm{card}(\mathcal{F}_T)$ for $\mathfrak{T}_2$). Recalling (29), we infer that

$$|\mathfrak{T}_1| + |\mathfrak{T}_2| \lesssim h^{k+1} \|u\|_{H^{k+2}(\Omega)} \|\mathsf{v}_h\|_{1,h}. \tag{45}$$

To estimate $\mathfrak{T}_3$, let $T \in \mathcal{T}_h$ and $F \in \mathcal{F}_T$. We observe that

$$\begin{aligned} h_F^{-1/2} \|\pi_F^k(\widehat{\mathsf{u}}_F - P_T^k \mathsf{L}_T \widehat{\mathsf{u}}_h)\|_F &= h_F^{-1/2} \|\pi_F^k(u - \breve{u}_T) - \pi_T^k(u - \breve{u}_T)\|_F \\ &\leqslant h_F^{-1/2} \|u - \breve{u}_T\|_F + h_F^{-1} C_{\mathrm{tr}} \|u - \breve{u}_T\|_T \lesssim h_T^{k+1} \|u\|_{H^{k+2}(T)}, \end{aligned} \tag{46}$$

where we have used the definitions (23) and (12) of P_T^k and I_T^k, respectively, the fact that $\pi_F^k \circ \pi_T^k = \pi_T^k$ on F, the discrete trace inequality (5), the approximation property (17) of $p_T^k \mathsf{I}_T^k$, and the mesh regularity property (3). Finally, using Cauchy–Schwarz inequality with (46), the bound (4) on $\mathrm{card}(\mathcal{F}_T)$, (25), and (31), we infer that

$$|\mathfrak{T}_3| \leqslant s_h(\widehat{\mathsf{u}}_h, \widehat{\mathsf{u}}_h)^{1/2} s_h(\mathsf{v}_h, \mathsf{v}_h)^{1/2} \lesssim h^{k+1} \|u\|_{H^{k+2}(\Omega)} \|\mathsf{v}_h\|_{a,h} \lesssim h^{k+1} \|u\|_{H^{k+2}(\Omega)} \|\mathsf{v}_h\|_{1,h}. \tag{47}$$

The conclusion of the proof then follows from (41), (44), (45), and (47). □

Corollary 9 (Error estimate on the exact gradient)**.** *Under the assumptions of Theorem 8, the following holds:*

$$\|\boldsymbol{\nabla} u - \boldsymbol{G}_h^k \mathsf{u}_h\| \leqslant C h^{k+1} \|u\|_{H^{k+2}(\Omega)}.$$

Proof. The triangle inequality and definition (25) yield

$$\|\boldsymbol{\nabla} u - \boldsymbol{G}_h^k \mathsf{u}_h\| \leqslant \|\boldsymbol{\nabla} u - \boldsymbol{G}_h^k \widehat{\mathsf{u}}_h\| + \|\boldsymbol{G}_h^k(\widehat{\mathsf{u}}_h - \mathsf{u}_h)\| \leqslant \|\boldsymbol{\nabla} u - \boldsymbol{G}_h^k \widehat{\mathsf{u}}_h\| + \|\widehat{\mathsf{u}}_h - \mathsf{u}_h\|_{a,h}.$$

Use (17) and (40) to estimate the terms on the right-hand side and conclude. □

3.2 L^2-norm error estimate

Adapting the techniques of [9, Section 4.2], we can also prove an optimal L^2-error estimate for the potential. To this end, we assume elliptic regularity in the following form: For all $g \in L^2(\Omega)$, the unique solution $z \in U_0$ to

$$(\nabla z, \nabla v) = (g, v) \qquad \forall v \in U_0, \tag{48}$$

satisfies the a priori estimate

$$\|z\|_{H^2(\Omega)} \leqslant C_{\text{ell}} \|g\|, \tag{49}$$

with a constant $C_{\text{ell}} > 0$ only depending on Ω.

Theorem 10 (L^2-error estimate for the potential)**.** *Under the assumptions of Theorem 8, assuming elliptic regularity* (49) *for problem* (1) *and that $f \in H^1(\Omega)$ for $k = 0$, there exists a real number $C > 0$ depending on the mesh regularity parameter ϱ but independent of h such that, for $k \geqslant 1$,*

$$\|\widehat{u}_h - u_h\| \leqslant C h^{k+2} \|u\|_{H^{k+2}(\Omega)}, \tag{50}$$

and for $k = 0$,

$$\|\widehat{u}_h - u_h\| \leqslant C h^2 \|f\|_{H^1(\Omega)}, \tag{51}$$

where $\widehat{u}_h, u_h$ are piecewise polynomial functions such that $\widehat{u}_{h|T} := \widehat{\mathsf{u}}_T = \pi_T^k u$ and $u_{h|T} := \mathsf{u}_T$ for all $T \in \mathcal{T}_h$.

Proof. We only sketch the proof, referring to [9, Section 4.2] for further insight. Let z solve (48) with $g := \widehat{u}_h - u_h$. Set $\widehat{\mathsf{z}}_h := \mathsf{I}_h^k z$ and $\mathsf{e}_h := \widehat{\mathsf{u}}_h - \mathsf{u}_h$. A straightforward computation shows that

$$\begin{aligned}
\|\widehat{u}_h - u_h\|^2 = &\sum_{T \in \mathcal{T}_h} \left\{ (\nabla \mathsf{e}_T, \nabla z)_T + \sum_{F \in \mathcal{F}_T} (\mathsf{e}_F - \mathsf{e}_T, \nabla z \cdot \boldsymbol{n}_{TF})_F - a_T(\mathsf{L}_T \mathsf{e}_h, \mathsf{L}_T \widehat{\mathsf{z}}_h) \right\} \\
&+ \sum_{T \in \mathcal{T}_h} \left\{ -(f, \pi_T^k z)_T + a_T(\mathsf{L}_T \widehat{\mathsf{u}}_h, \mathsf{L}_T \widehat{\mathsf{z}}_h) \right\} := \mathfrak{T}_1 + \mathfrak{T}_2.
\end{aligned}$$

To bound $\mathfrak{T}_1$, we observe that, with $\delta_T(z) := z_{|T} - p_T^k \mathsf{I}_T^k z_{|T}$,

$$\mathfrak{T}_1 = \sum_{T \in \mathcal{T}_h} \left\{ (\nabla \mathsf{e}_T, \nabla \delta_T(z))_T + \sum_{F \in \mathcal{F}_T} (\mathsf{e}_F - \mathsf{e}_T, \nabla \delta_T(z) \cdot \boldsymbol{n}_{TF})_F - s_T(\mathsf{L}_T \mathsf{e}_h, \mathsf{L}_T \widehat{\mathsf{z}}_h) \right\},$$

whence we infer that

$$\begin{aligned}
|\mathfrak{T}_1| &\leqslant \left\{ \|\mathsf{e}_h\|_{1,h}^2 + s_h(\mathsf{e}_h, \mathsf{e}_h) \right\}^{1/2} \left\{ \sum_{T \in \mathcal{T}_h} \left[\|\nabla \delta_T(z)\|_T^2 + h_T \|\nabla \delta_T(z)\|_{\partial T}^2 \right] + s_h(\widehat{\mathsf{z}}_h, \widehat{\mathsf{z}}_h) \right\}^{1/2} \\
&\lesssim h^{k+1} \|u\|_{H^{k+2}(\Omega)} h \|z\|_{H^2(\Omega)} \lesssim h^{k+2} \|u\|_{H^{k+2}(\Omega)} \|\widehat{u}_h - u_h\|,
\end{aligned}$$

owing to the energy-norm error estimate and elliptic regularity, while the bound on $\delta_T(z)$ and $s_h(\widehat{\mathsf{z}}_h, \widehat{\mathsf{z}}_h)$ is shown as in the proofs of Lemma 3 and Theorem 8, respectively. Turning to $\mathfrak{T}_2$, we observe that $(f, \pi_T^k z)_T = (\pi_T^k f, z)_T$ and since $(f, z) = (\nabla u, \nabla z)$, we infer that

$$\mathfrak{T}_2 = (f - \pi_h^k f, z) - \sum_{T \in \mathcal{T}_h} \left\{ (\nabla u, \nabla z)_T - (\nabla p_T^k \mathsf{I}_T^k u, \nabla p_T^k \mathsf{I}_T^k z)_T \right\} + s_h(\widehat{\mathsf{u}}_h, \widehat{\mathsf{z}}_h),$$

where π_h^k denotes the global version of the local L^2-projector π_T^k. Denote by $\mathfrak{T}_{2,1}, \mathfrak{T}_{2,2}, \mathfrak{T}_{2,3}$ the three terms on the right-hand side. If $k \geqslant 1$, we can write $(f - \pi_h^k f, z) = (f - \pi_h^k f, z - \pi_h^1 z)$ so that

$$|\mathfrak{T}_{2,1}| \lesssim h^k \|f\|_{H^k(\Omega)} h^2 \|z\|_{H^2(\Omega)} \lesssim h^{k+2} \|u\|_{H^{k+2}(\Omega)} \|\widehat{u}_h - u_h\|,$$

while for $k = 0$, we write $(f - \pi_h^0 f, z) = (f - \pi_h^0 f, z - \pi_h^0 z)$ so that

$$|\mathfrak{T}_{2,1}| \lesssim h \|f\|_{H^1(\Omega)} h \|z\|_{H^1(\Omega)} \lesssim h^2 \|f\|_{H^1(\Omega)} \|\widehat{u}_h - u_h\|.$$

Concerning $\mathfrak{T}_{2,2}$, we exploit the orthogonality property (18) to infer that

$$\mathfrak{T}_{2,2} = -\sum_{T \in \mathcal{T}_h} (\nabla u - \nabla p_T^k \mathsf{I}_T^k u, \nabla z - \nabla p_T^k \mathsf{I}_T^k z)_T,$$

whence $|\mathfrak{T}_{2,2}| \lesssim h^{k+2} \|u\|_{H^{k+2}(\Omega)} \|\widehat{u}_h - u_h\|$. Finally, proceeding as above, $\mathfrak{T}_{2,3}$ is bounded similarly, and this concludes the proof. □

4 Numerical tests

We solve the Dirichlet problem in the unit square with

$$u = \sin(\pi x_1) \sin(\pi x_2),$$

and corresponding right-hand side $f = 2\pi^2 \sin(\pi x_1) \sin(\pi x_2)$ on the four mesh families depicted in Figure 1. The triangular, Cartesian, and Kershaw mesh families correspond, respectively, to the mesh families 1, 2, and 4.1 of the FVCA5 benchmark [15], whereas the (predominantly) hexagonal mesh family was first introduced in [10]. The implementation framework corresponds to the one described in [9, Section 5], to which we refer for further details. Figure 2 displays convergence results for the various mesh families and polynomial degrees up to 4. The measure for the gradient error is $\|\widehat{\mathsf{u}}_h - \mathsf{u}_h\|_{a,h}$ (in case of a nontrivial diffusion coefficient or tensor, this quantity is actually to be interpreted as a flux error), whereas the potential error is estimated as $\|\widehat{u}_h - u_h\|$. In all cases, the numerical results show asymptotic convergence rates that match those predicted by the theory. The apparent super-convergence on the Kershaw mesh family (cf. Figures 2e–2f) is linked to the fact that the mesh quality improves when refining. Genuine super-convergence is on the other hand observed for the gradient on the Cartesian mesh family up to polynomial degree 2, cf. Figure 2c. For $k = 4$, round-off errors start to surface in the last refinement iteration for the Cartesian mesh family, as confirmed by a convergence rate slightly lower than expected, cf. Figure 2d.

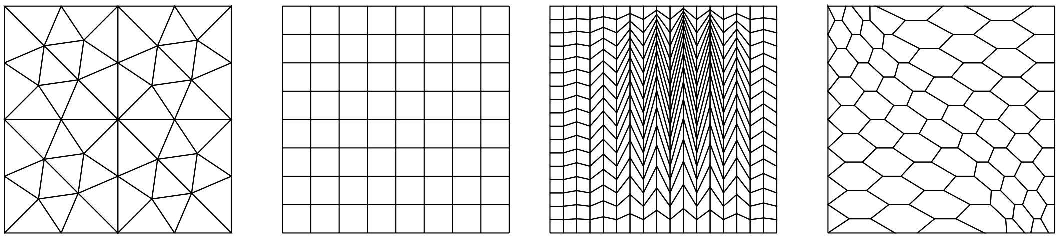


Figure 1: Triangular, Cartesian, Kershaw, and hexagonal-dominant meshes for the numerical tests of Section 4

References

[1] M. Bebendorf. A note on the Poincaré inequality for convex domains. *Z. Anal. Anwendungen*, 22(4):751–756, 2003.

[2] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. Basic principles of virtual element methods. *M3AS Math. Models Methods Appl. Sci.*, 199(23):199–214, 2013.

[3] L. Beirão da Veiga, K. Lipnikov, and G. Manzini. Arbitrary-order nodal mimetic discretizations of elliptic problems on general meshes. *SIAM J. Numer. Anal.*, 5(49):1737–1760, 2011.

[4] J. Bonelle and A. Ern. Analysis of compatible discrete operator schemes for elliptic problems on polyhedral meshes. *M2AN Math. Model. Numer. Anal.*, 2013. Published online. DOI: 10.1051/m2an/2013104.

[5] F. Brezzi, K. Lipnikov, and M. Shashkov. Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. *SIAM J. Numer. Anal.*, 43(5):1872–1896, 2005.

[6] 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(2):1319–1365, 2009.

[7] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications*. Springer-Verlag, Berlin, 2012.

[8] D. A. Di Pietro and A. Ern. A family of arbitrary order mixed methods for heterogeneous anisotropic diffusion on general meshes. Submitted. Preprint hal-00918482, 2013.

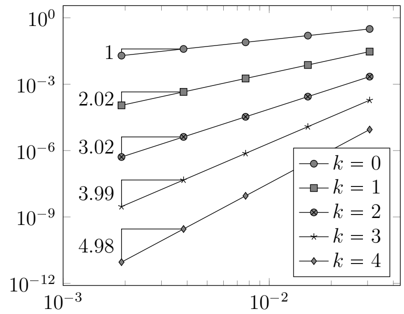
[9] D. A. Di Pietro and A. Ern. An arbitrary-order locking-free method for linear elasticity on general meshes. Submitted, 2014.

[10] D. A. Di Pietro and S. Lemaire. An extension of the Crouzeix–Raviart space to general meshes with application to quasi-incompressible linear elasticity and Stokes flow. *Math. Comp.*, 2014. Accepted for publication. Preprint hal-00753660.

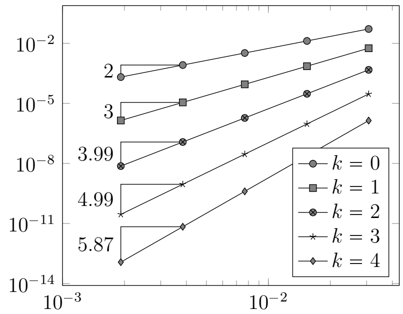
[11] J. Droniou and R. Eymard. A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numer. Math.*, 105:35–71, 2006.

[12] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *M3AS Mathematical Models and Methods in Applied Sciences*, 20(2):1–31, 2010.

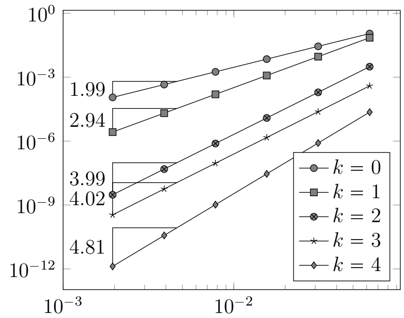
[13] T. Dupont and R. Scott. Polynomial approximation of functions in Sobolev spaces. *Math. Comp.*, 34(150):441–463, 1980.



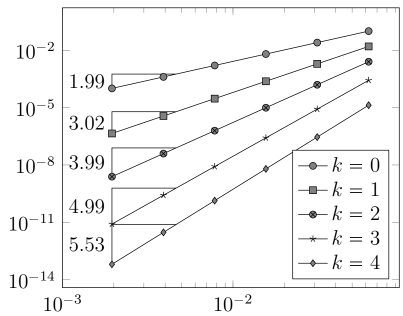
(a) Gradient error, triangular meshes



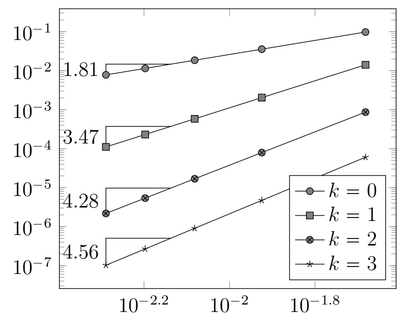
(b) Potential error, triangular meshes



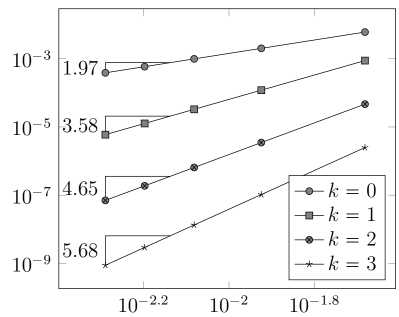
(c) Gradient error, Cartesian meshes



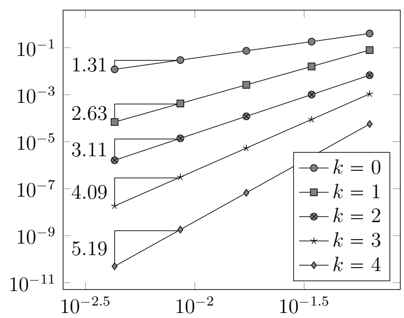
(d) Potential error, Cartesian meshes



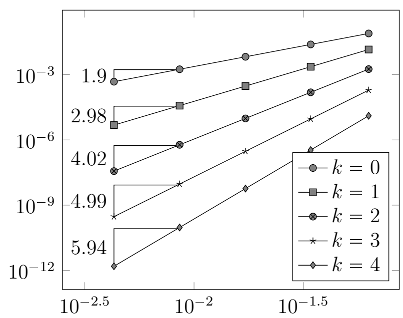
(e) Gradient error, Kershaw meshes



(f) Potential error, Kershaw meshes



(g) Gradient error, hexagonal meshes



(h) Potential error, hexagonal meshes

Figure 2: Errors vs. h

[14] R. Eymard, T. Gallouët, and R. Herbin. Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes. SUSHI: a scheme using stabilization and hybrid interfaces. *IMA J. Numer. Anal.*, 30(4):1009–1043, 2010.

[15] R. Herbin and F. Hubert. Benchmark on discretization schemes for anisotropic diffusion problems on general grids. In R. Eymard and J.-M. Hérard, editors, *Finite Volumes for Complex Applications V*, pages 659–692. John Wiley & Sons, 2008.

[16] A. Tabarraei and N. Sukumar. Conforming polygonal finite elements. *Int. J. Numer. Methods Engrg.*, 61(12):2045–2066, 2004.

[17] A. Tabarraei and N. Sukumar. Extended finite element method on polygonal and quadtree meshes. *Comput. Methods Appl. Mech. Engrg.*, 197(5):425–438, 2008.