A MULTISCALE HDG METHOD FOR SECOND ORDER ELLIPTIC EQUATIONS. PART I. POLYNOMIAL AND HOMOGENIZATION-BASED MULTISCALE SPACES

YALCHIN EFENDIEV *, RAYTCHO LAZAROV †, AND KE SHI ‡

Abstract. We introduce a finite element method for numerical upscaling of second order elliptic equations with highly heterogeneous coefficients. The method is based on a mixed formulation of the problem and the concepts of the domain decomposition and the hybrid discontinuous Galerkin methods. The method utilizes three different scales: (1) the scale of the partition of the domain of the problem, (2) the sale of partition of the boundaries of the subdomains (related to the corresponding space of Lagrange multipliers), and (3) the fine grid scale that is assumed to resolve the scale of the heterogeneous variation of the coefficients. Our proposed method gives a flexible framework that (1) couples independently generated multiscale basis functions in each coarse patch (2) provides a stable global coupling independent of local discretization, physical scales and contrast (3) allows avoiding any constraints (c.f., [8]) on coarse spaces. In this paper, we develop and study a multiscale HDG method that uses polynomial and homogenization-based multiscale spaces. These coarse spaces are designed for problems with scale separation. In our consequent paper, we plan to extend our flexible HDG framework to more challenging multiscale problems with non-separable scales and high contrast and consider enriched coarse spaces that use appropriate local spectral problems.

1. Introduction. In this paper we consider the following second order elliptic differential equation for the unknown function $u(x)$

$$-\nabla\cdot(\kappa(x)\nabla u) = f(x), \quad x\in\Omega \tag{1.1}$$

with homogeneous Dirichlet boundary conditions. Here $\kappa(x)\geq\kappa_0>0$ and $c(x)\geq 0$ are highly heterogeneous coefficients and Ω is a bounded polyhedral domain in $\mathbb{R}^n$, $n=2,3$. The presented in this paper methods are targeting applications of equation (1.1) to flows in porous media. Other possible applications are diffusion and transport of passive chemicals or heat transfer in heterogeneous media.

Flows in porous media appear in many industrial, scientific, engineering, and environmental applications. One common characteristic of these diverse areas is that porous media are intrinsically multiscale and typically display heterogeneities over a wide range of length-scales. Depending on the goals, solving the governing equations of flows in porous media might be sought at: (a) A coarse scale (e.g., if only the global pressure drop for a given flow rate is needed, and no other fine scale details of the solution are important), (b) A coarse scale enriched with some desirable fine scale details, and (c) The fine scale (if computationally affordable and practically desirable).

In naturally occurring materials, e.g. soil or rock, the permeability is small in granite formations (say 10^{-15} cm^2), medium in oil reservoirs, (say 10^{-7} cm^2 to 10^{-9} cm^2), and large in highly fractured or in vuggy media (say 10^{-3} cm^2). Numerical solution of such problems is a challenging task that has attracted a substantial attention in the scientific and engineering community.

In the last decade a number of numerical upscaling schemes that fall into the class of model reduction methods have been developed and used in various applications in geophysics and engineering related to problems in highly heterogeneous media. These include Galerkin multiscale finite element (e.g., [3, 13, 20, 18, 19]), mixed multiscale finite element (e.g., [1, 2, 5, 4]), the multiscale finite volume (see, e.g., [12, 23] mortar multiscale (see e.g., [8, 9], and variational multiscale (see e.g., [22]) methods. In the paper we present a general framework to design of numerical upscaling method based based on subgrid approximation using the hybrid discontinuous Galerkin finite element method (HDG) for second order elliptic equations. One of the earliest subgrid (variational multiscale) methods for Darcy's problem in a mixed form have been developed by Arbogast in [3], see also [25].

In order to fix the main ideas and to derive the numerical upscaling method we shall consider model

*Department of Mathematics, Texas A&M University, College Station, TX 77843, email: efendiev@math.tamu.edu
†Department of Mathematics, Texas A&M University, College Station, TX 77843, email: lazarov@math.tamu.edu
‡Department of Mathematics, Texas A&M University, College Station, TX 77843, email: shike@math.tamu.edu.

equation (1.1) with homogeneous Dirichlet boundary condition in a mixed form:

$$\alpha \boldsymbol{q} + \nabla u = 0 \qquad \text{in } \Omega, \tag{1.2a}$$

$$\nabla \cdot \boldsymbol{q} = f \qquad \text{in } \Omega \tag{1.2b}$$

$$u = 0 \qquad \text{on } \partial\Omega. \tag{1.2c}$$

Here $\alpha(x) = \kappa(x)^{-1}$ $\Omega \subset \mathbb{R}^n$ $(n = 2, 3)$ is a bounded polyhedral domain, $f \in L^2(\Omega)$.

In the paper we present a multiscale finite element approximation of the mixed system (1.2a) – (1.2c) based on the hybridized discontinuous Galerkin method. Multiscale methods have gained substantial popularity in the last decade. We can consider them as a procedure of numerical upscaling that extends the capabilities of the mathematical theory of homogenization to more general cases including materials with non-periodic properties, non-separable scales, and/or random coefficients.

The first efficient mixed multiscale finite element methods were devised by Arbogast in [7] as multi-block grid approximations using the framework of mortaring technique. Mortaring techniques (e.g. see the pioneering work [10]) were introduced to accommodate methods that can be defined in separate sub-domains that could have been independently meshed. This technique introduces an auxiliary space for a Lagrange multiplier associated with a continuity constraint on the approximate solution. The classical mortaring, devised for the needs of domain domain decomposition methods, has been adapted recently as multiscale finite element approximations, e.g. [4, 8, 9]. In a two-scale (two-grid, fine and coarse) method the aim is to resolve the local heterogeneities on the fine grid introduced on each coarse block and then "glue" these approximations together via mortar spaces, that play the role of Lagrange multipliers, defined of the boundaries of the coarse partition. In order to design a stable method the mortar spaces have to satisfy proper inf-sup condition. This approach shown to be well suited for problems with heterogeneous media and a number of efficient methods and implementations have been proposed, studied, and used for solving a variety of applied problems, see, e.g. [3, 4, 6, 8].

The multiscale finite element method in this paper is based on discretization of the domain Ω by using three different scales. First the domain is split into a number of non-overlapping subdomains with characteristic size L. This partition, denoted by $\mathcal{P}$, represents a coarse scale at which the global features of the solution are captured, but the local features are not resolved. Each subdomain is partitioned into finite elements with size h. This partition, denoted by $\mathcal{T}_h$, represents the scale at which the heterogeneities of the media are well represented and the local features of the solution can be resolved. Finally, on each interface of two adjacent subdomains from the partition $\mathcal{P}$ we introduce an additional partition $\mathcal{E}_H$ with characteristic size H. This partition is used to introduce the space of the Lagrange multipliers, which will provide the means of gluing together of the fine-grid approximations that are introduced on each subdomain using the fine-grid partition $\mathcal{T}_h$. Three scales partition of the domains have been used in the mortar multiscale finite element methods by Arbogast and Xiao in [9], where the scale L represents the size a cell at which a homogenized solution exits.

The hybridization of the finite element methods as outlined in [16] provides ample possibilities for "gluing" together various finite element approximations. The mechanism of this "glue" is based on the notion of *numerical trace* and *numerical flux*. Numerical trace is single valued function on the finite element interfaces and belongs to certain Lagrange multiplier space. This is also the space at which the global problem is formulated. The stability is ensured by a proper choice of the *numerical flux*, that involves a parameter τ, which proves stabilization of the scheme and some other desired properties (e.g. superconvergence) (for details, see, e.g. [15, 16, 17]). For multiscale methods, local basis functions are constructed independently in each coarse region. For this reason, approaches are needed that can flexibly couple these local multiscale local solutions without any constraints. In previous works, mortar multiscale methods are proposed to couple local basis functions; however, they require additional constraints on mortar spaces. The proposed approaches can avoid any constraints on coarse spaces and provide a flexible "gluing" procedure for coupling multiscale basis functions. In this paper, we focus on polynomial and homogenization based multiscale spaces and study their stability and convergence properties.

The paper is organized as follows. In Sect. 2 we introduce the necessary notations and describe the multiscale FEM based on the framework of hybridizable discontinuous Galerkin method. In Subsection

2.4 we recast the two scale method into a hybridized form which essentially reduces to a symmetric and positive definite system for the Lagrange multipliers associated with the trace of the solution of the interfaces. In Section 2.5 we show that under reasonable assumptions on the finite dimensional spaces and the mesh, the two-scale method has unique solution.

In Section 3 we present the error analysis for the multiscale method. In Subsection 3.1 we introduce a number of projection operators related to the finite dimensional spaces (that are used later) and also a special projection operator related to the hybridizable discontinuous Galerkin FEM. Further, we state the approximation properties of these projections in terms of the scales of the various partitions of the domain. In Subsections 4.2 and 4.3 we derive error estimates for the flux $\boldsymbol{q}$ and the pressure u. Finally in Section 5, we study a new class of non-polynomial space for the numerical trace for special case of heterogeneous media with periodic arrangement of the coefficients. This space was proposed by Arbogast and Xiao [9] as a space for the mortar method for constructing multiscale finite element approximations and uses information of the local solution on the periodic cell. We show that the proposed multi-scale method is well posed and has proper approximation properties.

2. Multiscale Finite Element Method. Now we present the multiscale finite element approximation of the system (1.2a) – (1.2c). For this we shall need partition of the domain into finite elements, the corresponding finite elements spaces, and some notation from Sobolev spaces.

2.1. Sobolev spaces and their norms. Throughout the paper we shall use the standard notations for Sobolev spaces and their norm on the domain Ω, subdomains $D \subset \Omega$ or their boundaries. For example, $\|v\|_{s,D}$, $|v|_{s,D}$, $\|v\|_{s,\partial D}$, $|v|_{s,\partial D}$, $s > 0$, denote the Sobolev norms and semi-norms on D and its boundary ∂D. For s integer the Sobolev spaces are Hilbert spaces and the norms are defined by the L^2-norms of their weak derivatives up to order s. For s non integer the spaces are defined by interpolation [21]. For $s = 0$ instead of $\|v\|_{0,D}$ we shall use $\|v\|_D$.

Further, we shall use various inequalities between norms and semi-norms related to embedding of Sobolev spaces. If $D \subset \Omega$ and $diam(D) = d$ then we have the following inequalities:

$$\|v\|_{\partial D}^2 \leq C \left(d \|\nabla v\|_D^2 + d^{-1} \|v\|_D^2 \right). \tag{2.1}$$

We remark that since in this paper we are using three different scales (L, H, h) of partition of the domain we shall use these inequalities for domains of sizes L, H or h.

2.2. Partition of the domain. The finite element spaces that are used in the proposed method are defined below. They involve three different meshes. Let $\mathcal{P}$ be a disjoint polygonal partition of the domain Ω which allows nonconforming decomposition, see e.g. Figure 1, and let the maximal diameter of all $T \in \mathcal{P}$ be L. Let $\mathcal{T}_h(T)$ be quasi-uniform conforming triangulations of T with maximum element diameter h_T, denote by $\mathcal{T}_h = \cup_{T\in\mathcal{P}} \mathcal{T}_h(T)$, and let $h = \max_{T\in\mathcal{P}} h_T$. Let $\mathcal{E}_T$ denote the set of all edges/faces of the triangulation $\mathcal{T}_h(T)$ and $\mathcal{E}_h := \cup_{T\in\mathcal{P}} \mathcal{E}_T$. We also set $\partial\mathcal{T}_h = \cup_{K\in\mathcal{T}_h} \partial K$. Consistently in this paper we shall denote by T the subdomains of the partition $\mathcal{P}$ while by K we shall denoted the finite elements of the fine partition $\mathcal{T}_h$.

We call F a interface of the partition $\mathcal{P}$ if F is either shared by two neighboring subdomains, $F = \bar{T}_1 \cap \bar{T}_2$ or $F = \bar{T} \cap \partial\Omega$. For each interface F, let $\mathcal{T}_F$ be a quasi-uniform partition of F with maximum element diameter H. Set $\mathcal{E}_H = \cup_{F\in\mathcal{E}} \mathcal{T}_F$ and $\mathcal{E}_h^0 := \{F \in \mathcal{E}_h : \quad F \cap \partial T = \emptyset \text{ for any } T \in \mathcal{P}\}$.

Thus, we have three scales: (1) L – the maximum size of the of the subdomains $T \in \mathcal{P}$, (2) H – the size of the partition of the boundaries of $T \in \mathcal{P}$, and finally (3) the scale of the fine-grid mesh – the maximum diameter h of the finite elements introduced in each subdomain $T \in \mathcal{P}$. In this paper we shall assume that the $diam(\Omega) = 1$ and $0 < h << H \leq L \leq 1$.

Below is a summary of the above notation by grouping them into categories according to the scale

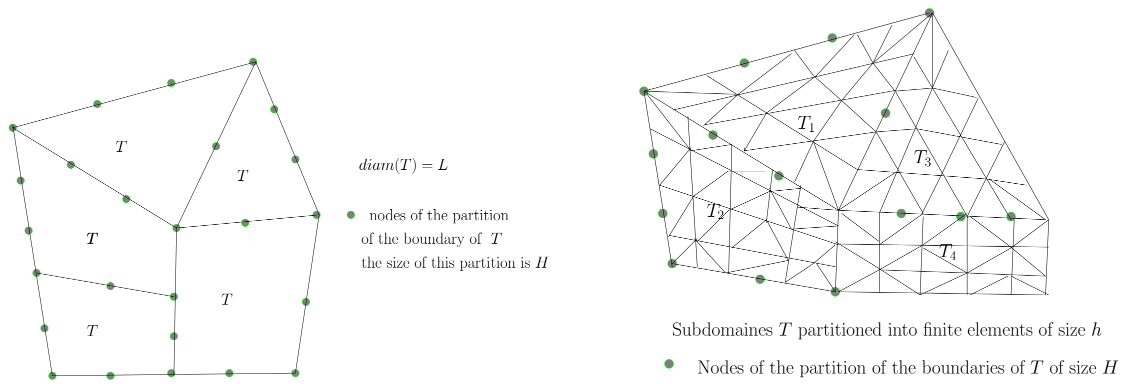

FIG. 1. *Partition of Ω: (left) on the boundaries of the subdomains T of size L an additional mesh of size H is shown; (right) on each subdomain T a fine mesh is introduced*

they represent:

(a) partition of the domain Ω into subdomains T (scale L):

$$\mathcal{P} := \text{the set of all subdomains } T$$
$$\partial\mathcal{P} := \cup_{T\in\mathcal{P}}\partial T$$

(b) partition of the boundaries of subdomains T (scale H):

$$\mathcal{E}_H(T) := \text{ the set of all coarse edges/faces of a subdomain } T\in\mathcal{P}$$
$$\mathcal{E}_H := \text{ the partition all edges/faces of the boundaries } \partial T,\, T\in\mathcal{P}$$

(c) partition of each subdomain $T\in\mathcal{P}$ into finite elements (scale h):

$$\mathcal{T}_h(T) := \text{fine grid triangulations of a subdomain } \; T\in\mathcal{P}$$
$$\mathcal{E}_h(T) := \text{ the set of all edges/faces of the triangulation } \mathcal{T}_h(T)$$
$$\mathcal{E}_h^0(T) := \text{ the set of all interior edges/faces of the triangulation } \mathcal{T}_h(T)\ (\equiv \mathcal{E}_h(T)\cap T)$$
$$\partial\mathcal{T}_T := \cup_{K\in\mathcal{T}_T}\partial K$$

(d) globally defined meshes on Ω:

$$\mathcal{T}_h := \cup_{T\in\mathcal{P}}\mathcal{T}_h(T)$$
$$\partial\mathcal{T}_h := \cup_{K\in\mathcal{T}_h}\partial K$$
$$\mathcal{E}_h := \cup_{T\in\mathcal{P}}\mathcal{E}_h(T)$$
$$\mathcal{E}_h^0 := \text{the set of all } F\in\mathcal{E}_h:\ F\in\mathcal{E}_h,\ F \text{ does not intersect } \partial T \text{ for any } T\in\mathcal{T}_h(T)$$
$$\mathcal{E}_{h,H} := \mathcal{E}_h^0\cup\mathcal{E}_H.$$

Note that the scale H is associated only with the partition of the boundaries of the subdomains T of the partition $\mathcal{P}$.

2.3. Multiscale FEM. The methods we are interested in seek an approximation to $(u, \boldsymbol{q}, u|_{\mathcal{E}_h})$ by the hybridized discontinuous Galerkin finite element method. For this purpose we need finite element spaces for these quantities consisting of piece-wise polynomial functions. Namely, we introduce

$$\begin{aligned}
W_h &:= \{w\in L^2(\mathcal{T}_h):\ w|_K\in W(K),\ K\in\mathcal{T}_h\},\\
\boldsymbol{V}_h &:= \{\boldsymbol{v}\in \boldsymbol{L}^2(\mathcal{T}_h):\ \boldsymbol{v}|_K\in \boldsymbol{V}(K),\ K\in\mathcal{T}_h\},\\
M_{h,H} &:= M_h^0\oplus M_H,
\end{aligned}$$

where the spaces M_h^0, M_H are defined as

$$\begin{aligned}
M_h^0 &:= \{\mu\in L^2(\mathcal{E}_{h,H}):\ \text{ for } F\in\mathcal{E}_h^0\ \ \mu|_F\in M_h(F),\ \text{ and }\ \mu|_{\mathcal{E}_H}=0\},\\
M_H &:= \{\mu\in L^2(\mathcal{E}_{h,H}):\ \text{ for } F\in\mathcal{E}_H\ \ \mu|_F\in M_H(F),\ \ \text{ and }\ \mu|_{\mathcal{E}_h^0\cup\partial\Omega}=0\}.
\end{aligned}$$

Now the hybridizable multicsale DG FEM reads as follows: find $(u_h, \boldsymbol{q}_h, \widehat{u}_{h,H})$ in the space $W_h \times \boldsymbol{V}_h \times M_{h,H}$ that satisfies the following weak problem

$$
\begin{aligned}
&(\alpha \boldsymbol{q}_h\,,\, \boldsymbol{v})_{\mathcal{T}_h} - (u_h\,,\, \nabla\cdot \boldsymbol{v})_{\mathcal{T}_h} + \langle \widehat{u}_{h,H}\,,\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &&= 0 \quad \forall \boldsymbol{v}\in \boldsymbol{V}_h, &&\text{(2.2a)}\\
&-(\boldsymbol{q}_h\,,\, \nabla w)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n}\,,\, w\rangle_{\partial\mathcal{T}_h} &&= (f\,,\, w)_{\mathcal{T}_h} \quad \forall w\in W_h, &&\text{(2.2b)}\\
&\langle \widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n}, \mu\rangle_{\partial\mathcal{T}_h} &&= 0 \quad \forall \mu\in M_{h,H}, &&\text{(2.2c)}\\
&\widehat{u}_{h,H} &&= 0 \quad \text{on } \partial\Omega. &&\text{(2.2d)}
\end{aligned}
$$

Since by the requirement $\widehat{u}_{h,H}\in M_{h,H}$ the last equation is trivially satisfied and therefore it is redundant. However, we prefer to have it written explicitly for later use in the error analysis.

For $\mathcal{T} = \mathcal{T}_h, \mathcal{T}_h(T)$, we write $(\eta\,,\,\zeta)_{\mathcal{T}} := \sum_{K\in\mathcal{T}}(\eta,\zeta)_K$, where $(\eta,\zeta)_D$ denotes the integral of $\eta\zeta$ over the domain $D\subset\mathbb{R}^n$. We also write $\langle\eta\,,\,\zeta\rangle_{\partial\mathcal{T}} := \sum_{K\in\mathcal{T}}\langle\eta\,,\,\zeta\rangle_{\partial K}$, where $\langle\eta\,,\,\zeta\rangle_{\partial D}$ denotes the integral of $\eta\zeta$ over the boundary of the domain $D\subset\mathbb{R}^{n-1}$. The definition of the method is completed with the definition of the normal component of the numerical trace:

$$
\widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n} = \boldsymbol{q}_h\cdot\boldsymbol{n} + \tau(u_h - \widehat{u}_{h,H}) \quad \text{on } \partial\mathcal{T}_h. \tag{2.3}
$$

On each $K\in\mathcal{T}_h$, the stabilization parameter τ is non-negative constant on each $F\in\partial K$ and we assume that $\tau>0$ on at least one face $F^*\in\partial K$. By taking particular choices of the local spaces $\boldsymbol{V}(K)$, $W(K)$ and $M_h(F), M_H(F)$, and the *linear local stabilization* operator τ, various mixed ($\tau=0$) and HDG ($\tau\neq 0$) methods are obtained. For a number of such choices we refer to [16, 17]. We note that on each fine element $K\in\mathcal{T}_h$ the local spaces $W(K)\times\boldsymbol{V}(K)\times M_h(F)$ can be any set of the spaces presented in [17, Tables 1 – 9]. It could be any classical mixed elements or the HDG elements defined on different triangulations. In Table 1 we give examples of local spaces for the classical mixed element and HDG element defined on a simplex.

TABLE 1
Possible choices for the finite element spaces for K a simplex.

method	$\boldsymbol{V}(K)$	$W(K)$	$M_h(F),\ F\in\partial K$	$M_H(F),\ F\in\mathcal{E}_H$
$\mathbf{BDFM}_{k+1}$	$\{\boldsymbol{q}\in\boldsymbol{P}^{k+1}(K):$ $\boldsymbol{q}\cdot\boldsymbol{n}\vert_{\partial K}\in P^k(F),\ \forall F\in\partial K\}$	$P^k(K)$	$P^k(F)$	$P^l(F)$
$\mathbf{RT}_k$	$\boldsymbol{P}^k(K)\oplus\boldsymbol{x}\widetilde{P}^k(K)$	$P^k(K)$	$P^k(F)$	$P^l(F)$
$\mathbf{HDG}_k$	$\boldsymbol{P}^k(K)$	$P^k(K)$	$P^k(F)$	$P^l(F)$

One feature of our formulation is that the choice of the space $M_H(F)$ is totally free. In this paper, we will consider two different choices. The first choice is the space of piece-wise polynomials defined in (3.1), while the second is the space uses multiscale functions defined in (5.4). In general it can consist of any function spaces.

2.4. The upscaled structure of the method. The main feature of this method is that it could be implemented in such a way that we need to solve certain global system on the coarse mesh $\mathcal{T}_H$ only. To show this possibility, we split (2.2c) into two equations by testing separately with $\mu\in M_h^0$ and $\mu\in M_H$ so that

$$
\langle\widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n},\mu\rangle_{\partial\mathcal{T}_h} = 0 \quad \forall\mu\in M_h^0 \quad\text{and}\quad \langle\widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n},\mu\rangle_{\partial\mathcal{T}_h} = \langle\widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n},\mu\rangle_{\partial\mathcal{T}_H} = 0 \quad \forall\mu\in M_H. \tag{2.4}
$$

Here $\langle\eta\,,\,\zeta\rangle_{\partial\mathcal{T}_H} := \sum_{T\in\mathcal{T}_h(T)}^n \int_{\partial T}\eta\zeta ds$. On any subdomain T, given the boundary data of $\widehat{u}_{h,H} = \xi_H$ for $\xi_H\in M_H(F), F\in\mathcal{E}_H(T)$, we can solve for $(\boldsymbol{q}_h, u_h, \widehat{u}_{h,H})|_T$ by restricting the equations (2.2a)–(2.2c) on this particular T:

$$\begin{aligned} (\alpha \boldsymbol{q}_h \,,\, \boldsymbol{v})_{\mathcal{T}_h(T)} - (u_h \,,\, \nabla\cdot \boldsymbol{v})_{\mathcal{T}_h(T)} + \langle \widehat{u}_{h,H} \,,\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h(T)} &= 0,\\ -(\boldsymbol{q}_h \,,\, \nabla w)_{\mathcal{T}_h(T)} \qquad\qquad + \langle \widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n} \,,\, w\rangle_{\partial\mathcal{T}_h(T)} &= (f \,,\, w)_{\mathcal{T}_h(T)},\\ \langle \widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n}, \mu\rangle_{\partial\mathcal{T}_h(T)} &= 0,\\ \widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n} &= \boldsymbol{q}_h\cdot\boldsymbol{n} + \tau(u_h - \widehat{u}_{h,H}) \quad \text{on} \quad \partial\mathcal{T}_h(T)\\ \widehat{u}_{h,H} &= \xi_H \quad \text{on } \partial T, \end{aligned}$$

for all $(w, \boldsymbol{v}, \mu) \in W_h|_T \times \boldsymbol{V}_h|_T \times M_h^0|_{\mathcal{E}_h^o(T)}$. In fact the above local system is the regular HDG methods defined on T. From [17] we already know that this system is stable. Hence, this HDG solver defines a global *affine* mapping from M_H to $W_h \times \boldsymbol{V}_h \times M_h^0$. The solution can be further split into two parts, namely,

$$(\boldsymbol{q}_h, u_h, \widehat{u}_{h,H}) = (\boldsymbol{q}_h(f), u_h(f), \widehat{u}_{h,H}(f)) + (\boldsymbol{q}_h(\xi_H), u_h(\xi_H), \widehat{u}_{h,H}(\xi_H))$$

where $(\boldsymbol{q}_h(f), u_h(f), \widehat{u}_{h,H}(f))$ satisfies

$$\begin{aligned} (\alpha \boldsymbol{q}_h(f) \,,\, \boldsymbol{v})_{\mathcal{T}_h(T)} - (u_h(f) \,,\, \nabla\cdot \boldsymbol{v})_{\mathcal{T}_h(T)} + \langle \widehat{u}_{h,H}(f) \,,\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h(T)} &= 0,\\ -(\boldsymbol{q}_h(f) \,,\, \nabla w)_{\mathcal{T}_h(T)} \qquad\qquad + \langle \widehat{\boldsymbol{q}}_{h,H}(f)\cdot\boldsymbol{n} \,,\, w\rangle_{\partial\mathcal{T}_h(T)} &= (f \,,\, w)_{\mathcal{T}_h(T)},\\ \langle \widehat{\boldsymbol{q}}_{h,H}(f)\cdot\boldsymbol{n}, \mu\rangle_{\partial\mathcal{T}_h(T)} &= 0,\\ \widehat{u}_{h,H} &= 0 \quad \text{on } \partial T, \end{aligned}$$

for all $(w, \boldsymbol{v}, \mu) \in W_h|_T \times \boldsymbol{V}_h|_T \times M_h^0|_{\mathcal{E}_h^o(T)}$ and $(\boldsymbol{q}_h(\xi_H), u_h(\xi_H), \widehat{u}_{h,H}(\xi_H))$ satisfies

$$\begin{aligned} (\alpha \boldsymbol{q}_h(\xi_H) \,,\, \boldsymbol{v})_{\mathcal{T}_h(T)} - (u_h(\xi_H) \,,\, \nabla\cdot \boldsymbol{v})_{\mathcal{T}_h(T)} + \langle \widehat{u}_{h,H}(\xi_H) \,,\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h(T)} &= 0,\\ -(\boldsymbol{q}_h(\xi_H) \,,\, \nabla w)_{\mathcal{T}_h(T)} \qquad\qquad + \langle \widehat{\boldsymbol{q}}_{h,H}(\xi_H)\cdot\boldsymbol{n} \,,\, w\rangle_{\partial\mathcal{T}_h(T)} &= 0,\\ \langle \widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n}, \mu\rangle_{\partial\mathcal{T}_h(T)} &= 0,\\ \widehat{u}_{h,H}(\xi_H) &= \xi_H \quad \text{on } \partial T, \end{aligned}$$

for all $(w, \boldsymbol{v}, \mu) \in W_h|_T \times \boldsymbol{V}_h|_T \times M_h^0|_{\mathcal{E}_h^o(T)}$.

Then the second equation (2.4) reduces to

$$a(\xi_H, \mu) = l(\mu) \qquad \text{for all } \mu \in M_H, \tag{2.5}$$

where the bilinear form $a(\xi_H, \mu) : M_H \times M_H \to R$ and the linear form $l(\mu) : M_H \to R$ are defined as

$$a(\xi_H, \mu) := \langle \widehat{\boldsymbol{q}}_{h,H}(\xi_H)\cdot\boldsymbol{n} \,,\, \mu\rangle_{\partial\mathcal{T}_H} \quad \text{and} \quad l(\mu) := a(f, \mu) = \langle \widehat{\boldsymbol{q}}_{h,H}(f)\cdot\boldsymbol{n} \,,\, \mu\rangle_{\partial\mathcal{T}_H}. \tag{2.6}$$

Remark 2.1. *The same procedure can be applied also for the case of non-homogeneous data $u = g$ on $\partial\Omega$. However, the presentation of this case is much more cumbersome. In order to simplify the notations and to highlight the main features of this method we have assumed homogeneous Dirichlet boundary data.*

2.5. Existence of the solution of the FEM. The framework is general in terms of flexibility in the choice of the local spaces. However, in order to ensure the solvability of the system, we need some assumptions.

Assumption 2.2. *For any $K \in \mathcal{T}_h$, F^* an arbitrary face of K, and $\mu \in M_h(F), F \in \partial K$, there exists a element $\boldsymbol{Z} \in \boldsymbol{V}(K)$ such that*

$$\begin{aligned} (\boldsymbol{Z}, \nabla w) &= 0, \quad \textit{for all } w \in W(K),\\ \boldsymbol{Z}\cdot\boldsymbol{n}|_F &= \mu, \quad \textit{for all } F \in \partial K\backslash F^*. \end{aligned}$$

This assumption is trivially satisfied by all classical mixed finite elements, e.g. **RT, BDM, BDDF**, etc. For these elements one can simply define $\boldsymbol{Z} = \boldsymbol{\Pi}_h \boldsymbol{Q}$, where $\boldsymbol{Q}$ is any solution of the problem:

$$\nabla \cdot \boldsymbol{Q} = 0 \quad \text{in } K \quad \text{and} \quad \boldsymbol{Q} \cdot \boldsymbol{n} = \mu \quad \text{on } \partial K,$$

where $\boldsymbol{\Pi}_h$ is the Fortin projection to the mixed elements (see, e.g. [11]). For the case of simplex triangulations and HDG elements, we refer the reader to [15, Lemma 3.2]. The proof for other HDG elements are very similar to the case of simplicial elements considered in [15].

Further, we need an assumption on the stabilization parameter τ:

Assumption 2.3. *On each $F_H \in \mathcal{T}_E$, for any T adjacent to F_H, i.e. $\bar{T} \cap F_H \neq \emptyset$, there exists at least one element $K \in \mathcal{T}_T$ adjacent to F_H, such that the stabilization operator $\tau > 0$ on $F^* = F_H \cap \partial K$.*

We are now ready to show the solvability of the method.

Theorem 2.4. *Let Assumptions 2.2 and 2.3 be satisfied. Then for any f, the system* (2.2) *has a unique solution.*

Proof. Notice that the system (2.2) is a square system. It suffices to show that the homogeneous system has only the trivial solution. From (2.2d) we see that $\widehat{u}_{h,H} = 0$ on $\partial\Omega$. Now assume that $(u_h, \boldsymbol{q}_h, \widehat{u}_{h,H})$ is any solution of (2.2). Setting $(w, \boldsymbol{v}, \mu) = (u_h, \boldsymbol{q}_h, \widehat{u}_{h,H})$ in (2.2a)-(2.2c) and adding all equations, we get after some algebraic manipulation,

$$(\alpha \boldsymbol{q}_h \,,\, \boldsymbol{q}_h)_{\mathcal{T}_h} - \langle \boldsymbol{q}_h \cdot \boldsymbol{n} - \widehat{\boldsymbol{q}}_{h,H} \cdot \boldsymbol{n} \,,\, u_h - \widehat{u}_{h,H} \rangle_{\partial \mathcal{T}_h} = 0.$$

By the definition of the numerical traces (2.3), we have

$$(\alpha \boldsymbol{q}_h \,,\, \boldsymbol{q}_h)_{\mathcal{T}_h} + \langle \tau (u_h - \widehat{u}_{h,H}) \,,\, u_h - \widehat{u}_{h,H} \rangle_{\partial \mathcal{T}_h} = 0$$

and since $\tau \geq 0$ we get

$$\boldsymbol{q}_h = 0, \qquad \tau(u_h - \widehat{u}_{h,H}) = 0 \tag{2.7}$$

and (2.2a) becomes

$$-(u_h \,,\, \nabla \cdot \boldsymbol{v})_{\mathcal{T}_h} + \langle \widehat{u}_{h,H} \,,\, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0, \quad \text{for all } \boldsymbol{v} \in \boldsymbol{V}_h.$$

Now we take this over an element K and after integration by parts we get

$$(\nabla u_h \,,\, \boldsymbol{v})_K + \langle \widehat{u}_{h,H} - u_h \,,\, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial K} = 0, \quad \text{for all } \boldsymbol{v} \in \boldsymbol{V}(K). \tag{2.8}$$

Since $\tau > 0$ on $F^* \in \partial K$ than the second equality (2.7) implies that $u_h - \widehat{u}_{h,H} = 0 \quad \text{on} \quad F^*$. Next, by Assumption 2.2, there is $\boldsymbol{v} \in \boldsymbol{V}(K)$ such that

$$\begin{aligned} (\boldsymbol{v}, \nabla w) &= 0, && \text{for all } w \in W(K), \\ \boldsymbol{v} \cdot \boldsymbol{n}|_F &= P_\partial^h \widehat{u}_{h,H} - u_h, && \text{for all } F \in \partial K \backslash F^*. \end{aligned}$$

where for $K \in \mathcal{T}_h$, $P_\partial^h : L^2(F) \to M_h(F)$ is the local L^2-orthogonal projection onto $M_h(F)$, for all $F \in \partial K$. Inserting such $\boldsymbol{v}$ in (2.8), we get

$$0 = (\nabla u_h \,,\, \boldsymbol{v})_K + \langle \widehat{u}_{h,H} - u_h \,,\, \boldsymbol{v} \cdot \boldsymbol{n} \rangle_{\partial K} = \langle P_\partial^h \widehat{u}_{h,H} - u_h \,,\, P_\partial^h \widehat{u}_{h,H} - u_h \rangle_{\partial K \backslash F^*}.$$

This implies that $P_\partial^h \widehat{u}_{h,H} - u_h = 0$ on $\partial K \backslash F^*$. Since on F^*, $P_\partial^h \widehat{u}_{h,H} - u_h = P_\partial^h (\widehat{u}_{h,H} - u_h) = 0$ we get

$$P_\partial^h \widehat{u}_{h,H} - u_h = 0 \quad \text{on } \partial K, \quad \text{for all } K \in \mathcal{T}_h. \tag{2.9}$$

Moreover, this means that $(\nabla u_h \,,\, \boldsymbol{v})_K = 0$ for all $\boldsymbol{v} \in \boldsymbol{V}(K)$. Taking $\boldsymbol{v} = \nabla u_h$, we have u_h is piecewise constant on each $K \in \mathcal{T}_h$. The above equation shows that $u_h = P_\partial^h \widehat{u}_{h,H}$ on ∂K. On each T, $\mathcal{T}_T$

is a conforming triangulation, so this implies that for any interior face $F \in \mathcal{E}_h^0$ shared by two neighboring elements K^+, K^-, the local spaces satisfy $M_h(F^+) = M_h(F^-)$ and hence $P_\partial^h \widehat{u}_{h,H}$ coincides from both sides. This implies that in fact $u_h = C_T$ in each subdomain T and $\widehat{u}_{h,H}|_{\mathcal{E}_h^0 \cap T} = C_T$.

Next, on each $F_H \in \mathcal{E}_H$, we assume $F_H \subset \bar{T}_1 \cap \bar{T}_2$, if $F_H \subset \partial\Omega$ then $F_H \subset \partial T_1$. By Assumption 2.3 there exists $K_1 \in \mathcal{T}_{T_1}, K_2 \in \mathcal{T}_{T_2}$ adjacent to F_H such that $\tau > 0$ on $F_i = \partial K_i \cap F_H$, $i = 1, 2$. By (2.7), we have

$$\widehat{u}_{h,H} - u_h = 0 \quad \text{on } F_i, \quad i = 1, 2.$$

This implies that $\widehat{u}_{h,H}|_{F_H} = C_{T_1} = C_{T_2}$. Hence we have $C_T = C$ for all T, which means that $u_h = C$ over the domain Ω and $\widehat{u}_{h,H}|_{\mathcal{E}_h} = C$. Finally, by the fact that $\widehat{u}_{h,H} = 0$ on $\partial\Omega$, we must have $u_h = \widehat{u}_{h,H} = C = 0$ and this completes the proof. ☐

In [8], in order to ensure the solvability of the mortar methods, the key assumption (roughly speaking) is that on $\mathcal{E}$ the fine scale space M_h should be rich enough comparing with the coarse scale space M_H. In this paper, since the stabilization is achieved by the parameter τ we prove stability under the assumption that on each $F_H \in \mathcal{E}_H$ the parameter τ is strictly positive on some portion of F_H. We do not need any conditions between the local spaces $M_h(F_h)$ and $M_H(F_H)$.

3. Error Analysis. In this section we derive error estimates for the proposed above method. We would like stress on two important points of this method. First, in the most general case we have three different scales in our partitioning. The error estimates should reflect this generality of the setting. Second, upon different choices of the spaces and the stabilization strategy (i.e. the choice of the parameter τ) we can get different convergence rates. For example, to obtain error estimates of optimal order we have to make some additional assumptions. All these are discussed in this section. For the sake of simplicity, we assume that the nonzero stabilization parameter τ is constant on all element $K \in \mathcal{T}_h$. In this section and the one follows, we only consider the method with the coarse space defined by polynomials, that is:

$$M_H(F) = P^l(F), \quad \text{for all} \quad F \in \mathcal{E}_H. \tag{3.1}$$

3.1. Preliminary Results. We present the main results in this section. In order to carry out a priori error estimates, we need some additional assumptions on the scheme. The first assumption is identical to *Assumption A* in [17], in order to be self-consistent, we still present it here:

ASSUMPTION 3.1. *The local spaces satisfy the following inclusion property:*

$$W(K)|_F \subset M_h(F) \quad \textit{for all } F \in \partial K, \tag{3.2a}$$

$$\boldsymbol{V}(K) \cdot \boldsymbol{n}|_F \subset M_h(F) \quad \textit{for all } F \in \partial K. \tag{3.2b}$$

On each element $K \in \mathcal{E}_h$, there exist local projection operators

$$\Pi_W : \ H^1(K) \to W(K) \quad \textit{and} \quad \boldsymbol{\Pi}_V : \ \boldsymbol{H}_{div}(K) \to \boldsymbol{V}(K)$$

associated with the spaces $W(K)$, $\boldsymbol{V}(K)$, $M_h(F)$ defined by:

$$(u, w)_K = (\Pi_W u, w)_K \qquad \textit{for all } w \in \nabla \cdot \boldsymbol{V}(K), \tag{3.3a}$$

$$(\boldsymbol{q}, \boldsymbol{v})_K = (\boldsymbol{\Pi}_V \boldsymbol{q}, \boldsymbol{v})_K \qquad \textit{for all } \boldsymbol{v} \in \nabla W(K), \tag{3.3b}$$

$$\langle \boldsymbol{q} \cdot \boldsymbol{n} + \tau u \, , \, \mu \rangle_F = \langle \boldsymbol{\Pi}_V \boldsymbol{q} \cdot \boldsymbol{n} + \tau \Pi_W u \, , \, \mu \rangle_F \quad \textit{for all } \mu \in M_h(F), F \in \partial K. \tag{3.3c}$$

ASSUMPTION 3.2. *On each fine element K, the stabilization operator τ is strictly positive on only one face $F \in \partial K$.*

ASSUMPTION 3.3. *For any element K adjacent to the skeleton $\partial\mathcal{P}$ on a face F, shared by K and $\partial\mathcal{P}$, τ is strictly positive, i.e. $\tau|_F > 0$.*

The above suggested local spaces $W(K) \times \boldsymbol{V}(K) \times M_h(F)$ or any set of local spaces presented in [17] satisfy Assumption 3.1. Moreover, assumptions 3.2 and 3.3 are the key to obtain optimal approximation results. In fact, without these two assumptions, we can still get some error estimates. However, the result will have a term with negative power of h which is not desirable since h is the finest scale. We will discuss this issue at the end of Section 4.2.

As a consequence of Assumption 3.2 and 3.3, the triangulation of each subdomain has to satisfy the requirement that *each fine scale finite element $K \in \mathcal{T}_h$ can share at most one face with the coarse skeleton $\mathcal{E}_H$.* This requirement implies that we need to put at least two fine elements to fill a corner of any subdomain. This suggests that we should use triangular (2D) or tetrahedral (3D) elements. In what follows, we restrict the choice of local spaces to be in Table 1. Notice that here we exclude the famous $\mathbf{BDM}_k$ space from the table. Roughly speaking, the reason is that in the case of $\mathbf{BDM}_k$ element, the local space $W(K) = P^{k-1}(K)$ is too small to provide a key property for the optimality of the error bound, see Lemma 4.2.

In [17] it has been shown that for any $(u, \boldsymbol{q}) \in H^1(K) \times \boldsymbol{H}_{div}(K)$, the projection $(\Pi_W u, \boldsymbol{\Pi}_V \boldsymbol{q}) \in W(K) \times \boldsymbol{V}(K)$ exists and is unique. Moreover, for all elements listed in Table 1, the projection has the following approximation property:

Lemma 3.4. *If the local spaces $\boldsymbol{V}(K), W(K)$ are mixed element spaces $\mathbf{RT}_k$ or $\mathbf{BDFM}_{k+1}$, then*

$$\|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_K \le Ch^s(\|\boldsymbol{q}\|_{s,K} + \tau\|u\|_{s,K}) \quad \textit{and} \quad \|u - \Pi_W u\|_K \le Ch^s\|u\|_{s,K}$$

and if the local spaces $\boldsymbol{V}(K), W(K)$ are $\mathbf{HDG}_k$ spaces, then

$$\|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_K \le Ch^s(\|\boldsymbol{q}\|_{s,K}) \quad \textit{and} \quad \|u - \Pi_W u\|_K \le Ch^s(\|u\|_{s,K} + \tau^{-1}\|\boldsymbol{q}\|_{s,K})$$

for all $1 \le s \le k+1$.

Further in our analysis we shall need some auxiliary projections and their properties:

$$\begin{array}{llcll} P_\partial^H : & L^2(F) & \to & M_H(F), & \langle P_\partial^H u, \mu\rangle_F = \langle u, \mu\rangle_F \quad \forall F \in \mathcal{E}_H, \\ P_\partial^h : & L^2(F) & \to & M_h(F), & \langle P_\partial^h u, \mu\rangle_F = \langle u, \mu\rangle_F \quad \forall F \in \mathcal{E}_h^0, \\ P_M : & L^2(\mathcal{E}_{h,H}) & \to & M_{h,H}, & \text{with} \quad P_M = \left\{ \begin{array}{ll} P_\partial^H & \text{on } \mathcal{E}_H, \\ P_\partial^h & \text{on } \mathcal{E}_h^0, \end{array}\right. \\ \mathcal{I}_H^0 : & C(\Omega) & \to & M_H^c & \text{with} \quad M_H^c \subset M_H, \end{array} \tag{3.4}$$

where M_H^c is the subset of M_H of continuous functions and $\mathcal{I}_H^0$ the Lagrange (nodal) interpolation operator.

From the last equation (3.3c) and the definitions (3.4) of the projection operators, we have

$$P_h^\partial(\boldsymbol{q} \cdot \boldsymbol{n}) + \tau P_h^\partial u = \boldsymbol{\Pi}_V \boldsymbol{q} \cdot \boldsymbol{n} + \tau \Pi_W u, \quad \text{for all } F \in \partial\mathcal{T}_h. \tag{3.5}$$

In the analysis, we will need the following useful approximation properties of the projections P_∂^h, P_∂^h and the interpolation operator $\mathcal{I}_H^0$:

Lemma 3.5. *For any $T \in \mathcal{T}$ and any smooth enough function u we have*

$$\|u - P_\partial^h u\|_{\partial T} \le CL^{-\frac{1}{2}} h^s \|u\|_{s+1,T}, \qquad 0 \le s \le k+1, \tag{3.6}$$

$$\|u - P_\partial^H u\|_{\partial T} \le CL^{-\frac{1}{2}} H^t \|u\|_{t+1,T}, \qquad 0 \le t \le l+1, \tag{3.7}$$

$$\|u - \mathcal{I}_H^0 u\|_{\frac{1}{2},\partial T} \le CL^{-\frac{1}{2}} H^{t-\frac{1}{2}} \|u\|_{t+1,T} \qquad 0 \le t \le l+1. \tag{3.8}$$

Here the constant C solely depends on the shape of the domain T but not if its size.

Remark 3.6. *The regularity assumptions $H^{s+1}(H^{t+1})$ can be weakened to $H^{s+\frac{1}{2}+\epsilon}(H^{t+\frac{1}{2}+\epsilon})$ for any $\epsilon > 0$ without reducing the approximation order. However, the above estimates make the presentation more transparent and shorter.*

Proof. (of Lemma 3.5) First we note the following standard estimates for the error on any edge/face $F \subset \partial T$, see [14]:

$$\|u - P_\partial^h u\|_F \leq Ch^s |u|_{s,F}, \qquad s \text{ integer}, \quad 0 \leq s \leq k+1, \tag{3.9a}$$
$$\|u - P_\partial^H u\|_F \leq CH^t |u|_{t,F}, \qquad t \text{ integer}, \quad 0 \leq t \leq l+1, \tag{3.9b}$$
$$\|(I - P_\partial^h)(\boldsymbol{q}\cdot\boldsymbol{n})\|_F \leq Ch^s |\boldsymbol{q}\cdot\boldsymbol{n}|_{s,F}, \qquad s \text{ integer}, \quad 0 \leq s \leq k+1, \tag{3.9c}$$
$$\|u - \mathcal{I}_H^0 u\|_{t,\partial T} \leq CH^{s-t} |u|_{s,\partial T},\ s,t \text{ integer},\ 1 < s \leq l+1,\ 0 \leq t \leq 1. \tag{3.9d}$$

All three inequalities can be obtained by a similar scaling argument. Here we only present the proof of the first one of them. Assume F is one of the faces of the element $T \in \mathcal{T}$. By (3.9a), we have

$$\begin{aligned}
\|u - P_\partial^h u\|_{\partial T} &\leq Ch^s |u|_{s,\partial T} \\
&\leq Ch^s (L^{-\frac{1}{2}} |u|_{s,T} + L^{\frac{1}{2}} |u|_{s+1,T}) \quad \text{by the trace inequality (2.1)}, \\
&\leq CL^{-\frac{1}{2}} h^s \|u\|_{s+1,T},
\end{aligned}$$

for all integer $0 \leq s \leq k+1$. The case of s non integer follows by interpolation and the other two are proven in a similar way. We note that the factor $L^{-\frac{1}{2}}$ related to the scale of the subdomains T. If the size of T is $O(1)$ then these estimates are well known. ∎

REMARK 3.7. *Note that this projections Π_W and $\boldsymbol{\Pi}_V$ are connected through the boundary equation* (3.3c). *Of course, for $\boldsymbol{H}_{div}$-conforming finite element spaces, we can take $\tau = 0$ and these two projections coincide with those of the mixed FEM. In particular, $\boldsymbol{\Pi}_V$ is well defined, see, e.g. [11, Section III.3.3].*

3.2. Main Result. We are now ready to state two main results for the methods, which proofs will postponed until Section 4. First we present the estimate for the vector variable $\boldsymbol{q}$ the the weighted norm

$$\|w\|_{\alpha,\Omega}^2 = \int_\Omega \alpha |w|^2 dx.$$

THEOREM 3.8. *Let the local spaces $W(K) \times \boldsymbol{V}(K) \times M_h(F)$ are any from Table 1 and let Assumption 3.2, 3.3 be satisfied. Then we have*

$$\begin{aligned}
\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\alpha,\Omega} &\leq C\|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\alpha,\Omega} + CH^{t-\frac{1}{2}} L^{-\frac{1}{2}} \|u\|_{t+1} \\
&\quad + C\tau H^t L^{-\frac{1}{2}} \|u\|_{t+1} + C\tau h^s L^{-\frac{1}{2}} \|u\|_{s+1} + C\tau^{-\frac{1}{2}} h^s L^{-\frac{1}{2}} \|\boldsymbol{q}\|_{s+1},
\end{aligned}$$

for all $0 \leq s \leq k+1$, $0 \leq t \leq l+1$ with constants C independent of $u, \boldsymbol{q}, h, H$, and L.

Next we state the result regarding the error $u - u_h$. It is valid under a typical elliptic regularity property we state next. Let $(\boldsymbol{\theta}, \phi)$ is the solution of the *dual* problem:

$$\alpha\boldsymbol{\theta} + \nabla\phi = 0 \quad \text{in } \Omega, \tag{3.10a}$$
$$\nabla\cdot\boldsymbol{\theta} = e_u \quad \text{in } \Omega, \tag{3.10b}$$
$$\phi = 0 \quad \text{on } \partial\Omega. \tag{3.10c}$$

We assume that we have full H^2-regularity,

$$\|\phi\|_{2,\Omega} + \|\boldsymbol{\theta}\|_{1,\Omega} \leq C\|e_u\|_\Omega, \tag{3.11}$$

where C only depends on the domain Ω.

THEOREM 3.9. *Let the conditions of Theorem 3.8 be satisfied. In addition, assume full elliptic regularity,* (3.11), *and the local space $W(K)$ contains piecewise linear functions for each $K \in \mathcal{T}_h$. Then* *for all* $1 \le s \le k+1,\ 1 \le t \le l+1$ *we have*

$$
\begin{aligned}
\|u-u_h\|_\Omega \le & \|u-\Pi_W u\|_\Omega \\
& + C\mathcal{C}\left(\|\boldsymbol{q}-\boldsymbol{\Pi}_V\boldsymbol{q}\|_{\alpha,\Omega} + (1+\tau H^{\frac12})H^{t-\frac12}L^{-\frac12}\|u\|_{t+1} + \tau h^s L^{-\frac12}\|u\|_{s+1} + \tau^{-\frac12}h^s L^{-\frac12}\|\boldsymbol{q}\|_{s+1}\right) \\
& + CH^{\frac32}L^{-\frac12}h^s(\|\boldsymbol{q}\|_{s+1} + \tau\|u\|_{s+1}) \\
& + CH^t L^{-\frac12}\left(H^{\frac32}\|\boldsymbol{q}\|_{t+1} + (h^{\frac12}+\tau H^{\frac32})\|u\|_{t+1}\right),
\end{aligned}
$$

where $\mathcal{C} := C_\alpha h + H + h^{\frac12}\tau^{-\frac12} + \tau^{\frac12}H^{\frac32}$, *and the constants* C *are independent of* $u, \boldsymbol{q}, h, H, L$.

The above two results are based on a general framework which utilizes three different scales L, H, h and a stabilization parameter τ. The richness of the proposed setup gives a flexibility that allows us to modify the method to fit different scenarios. On the other hand, it is hard to see the convergent rates of the methods based on this general setup. Now we discuss the results in details under some practical conditions. Here will simply assume the the coefficient α is uniformly bounded.

• **Case 1:** $L = \mathcal{O}(1)$. Basically, this means that the subdomains $T \in \mathcal{P}$ have the same scale as the original domain Ω. In this case, if we take $\tau = 1$, by the above two theorems and Lemma 3.4, we may summarize the order of convergence as follows:

$$\|\boldsymbol{q}-\boldsymbol{q}_h\|_\Omega = \mathcal{O}(H^{l+\frac12} + h^{k+1}) \quad \text{and} \quad \|u-u_h\|_\Omega = \mathcal{O}(H^{l+\frac32}\max\{1, h^{\frac12}H^{-1}\} + h^{k+1}).$$

In this case, our method is very close to the mortar methods introduced in [8]. Indeed, the mortar methods have the following convergence rate:

$$\|\boldsymbol{q}-\boldsymbol{q}_h\|_\Omega = \mathcal{O}(H^{l+\frac12} + h^{k+1}) \quad \text{and} \quad \|u-u_h\|_\Omega = \mathcal{O}(H^{l+\frac32} + h^{k+1}).$$

We can see that both methods have exactly the same order of convergence for $\boldsymbol{q}$. For the unknown u, the HDG methods have an extra term $\max\{1, h^{\frac12}H^{-1}\}$. This suggests that HDG method little weaker approximation property if $h > H^2$. This is due to the stabilization operator in the formulation. However, the advantage of the stabilization is that we don't need any assumption between the spaces M_h and M_H.

If we choose $\tau = H^{-1}$, then the constant $\mathcal{C} = \mathcal{O}(H)$, combining Lemma 3.4, Theorem 3.8, 3.9, we obtain the following convergence rate:

$$\|\boldsymbol{q}-\boldsymbol{q}_h\|_\Omega = \mathcal{O}(H^{l} + h^{k+1}H^{-1}) \quad \text{and} \quad \|u-u_h\|_\Omega = \mathcal{O}(H^{l+1} + h^{k+1}).$$

We can see that in this situation, the convergence rates for both $\boldsymbol{q}$ and u are slightly degenerated.

• **Case 2:** $H = L$. From the practical point of view, this assumption suggests that we don't further divide the edges of the subdomains $T \in \mathcal{P}$. In this case, we also present the convergence rates by taking $\tau = 1, \tau = H^{-1}$, respectively.

For $\tau = 1$, the order of convergence is:

$$\|\boldsymbol{q}-\boldsymbol{q}_h\|_\Omega = \mathcal{O}(H^{l} + h^{k+1}H^{-\frac12}) \quad \text{and} \quad \|u-u_h\|_\Omega = \mathcal{O}(H^{l+1}\max\{1, h^{\frac12}H^{-1}\} + h^{k+1}).$$

For $\tau = H^{-1}$, the order of convergence is:

$$\|\boldsymbol{q}-\boldsymbol{q}_h\|_\Omega = \mathcal{O}(H^{l-\frac12} + h^{k+1}H^{-\frac32}) \quad \text{and} \quad \|u-u_h\| = \mathcal{O}(H^{l+\frac12} + h^{k+1}H^{-\frac12}).$$

Similar as in **Case 1**, the convergence rates for both unknowns are worse if we choose $\tau = H^{-1}$.

We can see that if we choose the stabilization parameter τ inappropriately, the numerical solution does not even converge. On the other hand, if all other parameters are pre-assigned, we can follow a simple calculation to determine the optimal value of τ for the methods. We will illustrate this strategy with following setting: we assume that the polynomial degrees k, l are given, $L = H$, $h = H^\alpha$ ($\alpha > 1$), the local spaces are HDG spaces. Then the order of convergence for $\boldsymbol{q}$ solely depends on τ. Namely, it can be written as $\|\boldsymbol{q}-\boldsymbol{q}_h\| = \mathcal{O}(h^{k+1}+H^l+\tau H^{l+\frac{1}{2}}+\tau h^{k+1}H^{-\frac{1}{2}}+\tau^{-\frac{1}{2}}h^{k+1}H^{-\frac{1}{2}})$. Applying the relation $h=H^\alpha$ and setting $\tau=H^\gamma$, we obtain: $\|\boldsymbol{q}-\boldsymbol{q}_h\| = \mathcal{O}(H^{f(\gamma)})$, where $f(\gamma)$ will be the minimum of the $\alpha(k+1), l, \gamma+k+\frac{1}{2}, \gamma+\alpha(k+1)-\frac{1}{2}$, and $-\frac{\gamma}{2}+\alpha(k+1)-\frac{1}{2}\}$. The above function is continuous with respect to γ. It is obvious that $f(\gamma)<0$ if $|\gamma|>2\alpha(k+1)$. Therefore the absolute maximum of $f(\gamma)$ appears in the interval $(-2\alpha(k+1), 2\alpha(k+1))$. Assume that $f(\gamma)$ achieves its maximum at $\gamma=\gamma^*$, we can take $\tau=H^{\gamma^*}$ to get optimal convergence rate for $\boldsymbol{q}$. This strategy can be applied to u as well.

4. Proof of the main results.

Now we prove the main results of the paper stated in Theorems 3.8 and 3.9. The proofs follow the technique developed in [17] for the hybridizable discontinuous Galerkin method and is done in several steps, by establishing first an estimate for the vector variable $\boldsymbol{q}$ and then for scalar variable u.

4.1. Error equations.

We begin by obtaining the error equations we shall use in the analysis. The main idea is to work with the following projection errors:

$$
\begin{aligned}
\boldsymbol{e}_q &:= \boldsymbol{\Pi}_V\boldsymbol{q}-\boldsymbol{q}_h,\\
e_u &:= \Pi_W u - u_h,\\
\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} &:= P_M(\boldsymbol{q}\cdot\boldsymbol{n}) - \widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n},\\
e_{\widehat{u}} &:= P_M u - \widehat{u}_{h,H}.
\end{aligned}
$$

Further, we define

$$
\begin{aligned}
\delta_u &:= u - \Pi_W u,\\
\boldsymbol{\delta}_q &:= \boldsymbol{q}-\boldsymbol{\Pi}_V\boldsymbol{q}.
\end{aligned}
$$

Lemma 4.1. *Under the Assumption 3.1, we have*

$$
\begin{aligned}
&\text{(4.1a)} \quad (\alpha\boldsymbol{e}_q\,,\,\boldsymbol{v})_{\mathcal{T}_h} - (e_u\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + \langle e_{\widehat{u}}\,,\,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = -(\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{v})_{\mathcal{T}_h} - \langle (I-P_M)u\,,\,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h},\\
&\text{(4.1b)} \quad -(\boldsymbol{e}_q\,,\,\nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,w\rangle_{\partial\mathcal{T}_h} = -\langle (I-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,w\rangle_{\partial\mathcal{T}_h},\\
&\text{(4.1c)} \quad \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}, \mu\rangle_{\partial\mathcal{T}_h} = 0,\\
&\text{(4.1d)} \quad e_{\widehat{u}}|_{\partial\Omega} = 0,
\end{aligned}
$$

for all $(w,\boldsymbol{v},\mu)\in W_h\times\boldsymbol{V}_h\times M_{h,H}$. Here I is the identity operator. Moreover,

$$
\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} = \boldsymbol{e}_q\cdot\boldsymbol{n} + \tau(e_u - e_{\widehat{u}}) - (P_h^\partial - P_M)(\boldsymbol{q}\cdot\boldsymbol{n}+\tau u) \quad \text{on} \quad \partial\mathcal{T}_h. \tag{4.2}
$$

Proof. Let us begin by noting that the exact solution $(u,\boldsymbol{q})$ obviously satisfies

$$
\begin{aligned}
(\alpha\boldsymbol{q}\,,\,\boldsymbol{v})_{\mathcal{T}_h} - (u\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + \langle u\,,\,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
-(\boldsymbol{q}\,,\,\nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{q}\cdot\boldsymbol{n}\,,\,w\rangle_{\partial\mathcal{T}_h} &= (f\,,\,w)_{\mathcal{T}_h},\\
\langle \boldsymbol{q}\cdot\boldsymbol{n}, \mu\rangle_{\partial\mathcal{T}_h} &= 0,
\end{aligned}
$$

for all $(w,\boldsymbol{v},\mu)\in W_h\times\boldsymbol{V}_h\times M_{h,H}$. By the orthogonality properties (3.3a) and (3.3b) of the projection $\Pi=(\boldsymbol{\Pi}_V,\Pi_W)$, we obtain that

$$
\begin{aligned}
(\alpha\boldsymbol{q}\,,\,\boldsymbol{v})_{\mathcal{T}_h} - (\Pi_W u\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + \langle u\,,\,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= 0,\\
-(\boldsymbol{\Pi}_V\boldsymbol{q}\,,\,\nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{q}\cdot\boldsymbol{n}\,,\,w\rangle_{\partial\mathcal{T}_h} &= (f\,,\,w)_{\mathcal{T}_h},\\
\langle \boldsymbol{q}\cdot\boldsymbol{n}, \mu\rangle_{\partial\mathcal{T}_h} &= 0,
\end{aligned}
$$

for all $(w, \boldsymbol{v}, \mu) \in W_h \times \boldsymbol{V}_h \times M_{h,H}$. Moreover, since P_M is the L^2-projection into $M_{h,H}$, we get,

$$
\begin{aligned}
(\alpha \boldsymbol{q}\,,\, \boldsymbol{v})_{\mathcal{T}_h} - (\Pi_W u\,,\, \nabla\cdot \boldsymbol{v})_{\mathcal{T}_h} + \langle P_M u\,,\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= -\langle (I-P_M)u\,,\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h},\\
-(\boldsymbol{\Pi}_V \boldsymbol{q}\,,\, \nabla w)_{\mathcal{T}_h} + \langle P_M(\boldsymbol{q}\cdot\boldsymbol{n})\,,\, w\rangle_{\partial\mathcal{T}_h} &= (f\,,\, w)_{\mathcal{T}_h} - \langle (I-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\, w\rangle_{\partial\mathcal{T}_h},\\
\langle P_M(\boldsymbol{q}\cdot\boldsymbol{n}), \mu\rangle_{\partial\mathcal{T}_h} &= 0,
\end{aligned}
$$

for all $(w, \boldsymbol{v}, \mu) \in W_h \times \boldsymbol{V}_h \times M_{h,H}$. Subtracting the four equations defining the weak formulation of the HDG method (2.2) from the above equations, respectively, we obtain the equations for the projection of the errors. The last error equation (4.1d) is due to the definition of $\widehat{u}_{h,H}$ on $\partial\Omega$.

It remains to prove the identity (4.2) for $\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}$. On each $F \in \partial K$, $K \in \mathcal{T}_h$ after using the the definition of numerical traces (2.3) we get

$$
\begin{aligned}
\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} - \boldsymbol{e}_q\cdot\boldsymbol{n} &= P_M(\boldsymbol{q}\cdot\boldsymbol{n}) - \widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n} - (\boldsymbol{\Pi}_V\boldsymbol{q}\cdot\boldsymbol{n} - \boldsymbol{q}_h\cdot\boldsymbol{n})\\
&= P_M(\boldsymbol{q}\cdot\boldsymbol{n}) - \boldsymbol{\Pi}_V\boldsymbol{q}\cdot\boldsymbol{n} - (\widehat{\boldsymbol{q}}_{h,H}\cdot\boldsymbol{n} - \boldsymbol{q}_h\cdot\boldsymbol{n})\\
&= P_h^{\partial}(\boldsymbol{q}\cdot\boldsymbol{n}) - \boldsymbol{\Pi}_V\boldsymbol{q}\cdot\boldsymbol{n} + \tau(u_h - \widehat{u}_{h,H}) + (P_M - P_{\partial}^h)(\boldsymbol{q}\cdot\boldsymbol{n}).
\end{aligned}
$$

Then using the property of the projection Π_W defined in (3.5) the equality reduces to

$$
\begin{aligned}
\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} - \boldsymbol{e}_q\cdot\boldsymbol{n} &= \tau(-P_h^{\partial}u + \Pi_W u) + \tau(u_h - \widehat{u}_{h,H}) + (P_M - P_{\partial}^h)(\boldsymbol{q}\cdot\boldsymbol{n})\\
&= \tau(-P_M u + \Pi_W u) + \tau(u_h - \widehat{u}_{h,H}) + (P_M - P_{\partial}^h)(\boldsymbol{q}\cdot\boldsymbol{n} + \tau u)\\
&= \tau(e_u - e_{\widehat{u}}) + (P_M - P_{\partial}^h)(\boldsymbol{q}\cdot\boldsymbol{n} + \tau u)
\end{aligned}
$$

and this completes the proof. ∎

4.2. Estimate for $\boldsymbol{q}-\boldsymbol{q_h}$. For the error estimate of $\boldsymbol{e}_q$ we need to following lemma:

LEMMA 4.2. *Let the Assumptions 3.2, 3.3 hold. Then*
(a) on each subdomain $T \in \mathcal{P}$, $\boldsymbol{e}_q \in \boldsymbol{H}(div, T)$;
(b) $\|\nabla\cdot\boldsymbol{e}_q\|_T = 0$, for all $T \in \mathcal{P}$;
(c) $\boldsymbol{e}_q\cdot\boldsymbol{n}|_F = \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}|_F$, for all $F \in \mathcal{E}_h^0(T)$.

Proof. Now take any $T \in \mathcal{P}$. To prove that $\boldsymbol{e}_q$ is $\boldsymbol{H}_{div}$-conforming in T, we need to show that $\boldsymbol{e}_q\cdot\boldsymbol{n}$ is continuous across all interior interfaces $F \in \mathcal{E}_h^0(T)$. By the error equation (4.1c), we know that $\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}$ is single valued on all interior interfaces due to the fact that $\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}$ and the test function μ are in the same space $M_h(F)$. Hence, it is suffices to show that

$$
\boldsymbol{e}_q\cdot\boldsymbol{n}|_F = \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}|_F, \quad \forall\, F \in \mathcal{E}_h^0(T).
$$

First of all, on each interior face $P_{\partial}^h = P_M$, together with (4.2), we have

$$
\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} = \boldsymbol{e}_q\cdot\boldsymbol{n} + \tau(e_u - e_{\widehat{u}}), \quad \forall\, F \in \mathcal{E}_h^0(T). \tag{4.3}
$$

From here we can see that $\boldsymbol{e}_q\cdot\boldsymbol{n}|_F = \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}|_F$ if $\tau|_F = 0$. We only need to show that

$$
\tau(e_u - e_{\widehat{u}})|_{F^*} = 0, \quad \forall\, F^* \in \partial K, \quad F^* \cap \mathcal{E}_H = \emptyset. \tag{4.4}
$$

On any K adjacent with $\mathcal{E}_H$, by our assumptions, $\tau > 0$ on F^* where F^* is on the boundary of T. So on the other faces $\tau = 0$ and hence $\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} = \boldsymbol{e}_q\cdot\boldsymbol{n}$.

Let us consider an arbitrary interior element K with $\tau > 0$ on F^*. We restrict the error equation (4.1b) on K, integrating by parts, we have

$$
(\nabla\cdot\boldsymbol{e}_q\,,\, w)_K + \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} - \boldsymbol{e}_q\cdot\boldsymbol{n}\,,\, w\rangle_{\partial K} = -\langle (I-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\, w\rangle_{\partial K}.
$$

By (4.3) and the fact that $P_M = P_{\partial}^h$ on ∂K, we have

$$
(\nabla\cdot\boldsymbol{e}_q\,,\, w)_K + \langle \tau(e_u - e_{\widehat{u}})\,,\, w\rangle_{\partial K} = 0.
$$

Since $\tau > 0$ only on F^*, we have

$$(\nabla \cdot \boldsymbol{e}_q \,,\, w)_K + \langle\, \tau(e_u - e_{\widehat{u}}) \,,\, w \,\rangle_{F^*} = 0.$$

Now let $w \in P^k(K)$ be such that

$$(w \,,\, r)_K = (\nabla \cdot \boldsymbol{e}_q \,,\, r)_K, \qquad \forall\, r \in P^{k-1}(K), \tag{4.5a}$$

$$\langle w \,,\, \mu\rangle_{F^*} = \langle e_u - e_{\widehat{u}} \,,\, \mu\rangle_{F^*} \qquad \forall\, \mu \in P^k(F^*). \tag{4.5b}$$

One can easily see that such $w \in P^k(K)$ exists and is unique. Indeed, this is a square system for the coefficients of the polynomial w and it is sufficient to show that the homogeneous system has only a trivial solution. On F^* the equation $\langle w, \mu\rangle_{F^*} = 0$ represents a square homogeneous system for the trace $w|_{F^*} \in P^k(F^*)$. This ensures that the trace is identically zero on F^*. Without loss of generality we can assume that F^* is in the hyperplane $x_1 = 0$. Then obviously $w = x_1\tilde{w}$ with $\tilde{w} \in P^{k-1}(K)$ and now $(x_1\tilde{w},\ r)_K = 0$ for all $r \in P^{k-1}(K)$ implies $\tilde{w} = 0$. Then we plug w into the above error equation and notice that $\nabla \cdot \boldsymbol{e}_q \in P^{k-1}(K),\ e_u - e_{\widehat{u}} \in P^k(F^*)$ to get

$$(\nabla \cdot \boldsymbol{e}_q \,,\, \nabla \cdot \boldsymbol{e}_q)_K + \langle\, \tau(e_u - e_{\widehat{u}}) \,,\, e_u - e_{\widehat{u}} \,\rangle_{F^*} = 0.$$

This implies

$$\nabla \cdot \boldsymbol{e}_q|_K = 0, \quad e_u - e_{\widehat{u}}|_{F^*} = 0$$

and hence, $\boldsymbol{e}_q \cdot \boldsymbol{n}|_F = \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}|_F$ for all $F \in \mathcal{E}_h^0(K)$. Consequently, $\boldsymbol{e}_q \in \boldsymbol{H}(div, T)$ for all $T \in \mathcal{P}$.

To finish, we still need to show that $\nabla \cdot \boldsymbol{e}_q|_K = 0$ when K is adjacent with the boundary of T. Similarly as interior element K, error equation (4.1b) gives

$$(\nabla \cdot \boldsymbol{e}_q \,,\, w)_K + \langle\, \tau(e_u - e_{\widehat{u}}) \,,\, w \,\rangle_{F^*} = -\langle\, (I - P_M)(\boldsymbol{q} \cdot \boldsymbol{n}) \,,\, w \,\rangle_{F^*}.$$

Take w to be again the unique element in $P^k(K)$ such that

$$(w \,,\, r)_K = (\nabla \cdot \boldsymbol{e}_q \,,\, r)_K \quad \forall\, r \in P^{k-1}(K) \quad \text{and} \quad \langle w \,,\, \mu\rangle_{F^*} = 0 \quad \forall\, \mu \in P^k(F^*).$$

The second equation implies that $w = 0$ on F^*, so we have

$$(\nabla \cdot \boldsymbol{e}_q \,,\, w)_K + \langle \tau(e_u - e_{\widehat{u}}) \,,\, w\rangle_{F^*} = (\nabla \cdot \boldsymbol{e}_q \,,\, \nabla \cdot \boldsymbol{e}_q)_K = 0 \quad \to \quad \nabla \cdot \boldsymbol{e}_q = 0.$$

This completes the proof. ∎

REMARK 4.3. *The above proof cannot be applied for* $\mathbf{BDM}_k$*. Namely, a key step is the special choice of* w *which satisfies (4.5). In the case of* $\mathbf{BDM}_k$*,* w *is in a smaller space* $P^{k-1}(K)$*, hence the existence of* w *is no longer valid.*

We are now ready to obtain an upper bound of the L^2-norm of $\boldsymbol{e}_q$. We first prove the following Lemma.

LEMMA 4.4. *Under Assumption 3.1, we have*

$$\begin{aligned}\|\boldsymbol{e}_q\|^2_{\alpha,\Omega} + \|e_u - e_{\widehat{u}}\|^2_{\tau,\partial\mathcal{T}_h} = &-(\alpha\boldsymbol{\delta}_q \,,\, \boldsymbol{e}_q)_{\mathcal{T}_h} - \langle (I - P_\partial^H)u \,,\, \boldsymbol{e}_q \cdot \boldsymbol{n}\rangle_{\partial\mathcal{T}_H} \\ &+ \langle P_\partial^h u - P_\partial^H u \,,\, \tau(e_u - e_{\widehat{u}})\rangle_{\partial\mathcal{T}_H} - \langle (I - P_\partial^h)(\boldsymbol{q} \cdot \boldsymbol{n}) \,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_H},\end{aligned}$$

where

$$\|\boldsymbol{e}_q\|^2_{\alpha,\Omega} := (\alpha\boldsymbol{e}_q \,,\, \boldsymbol{e}_q)_{\mathcal{T}_h}, \quad \text{and} \quad \|e_u - e_{\widehat{u}}\|^2_{\tau,\mathcal{E}_h} = \langle \tau(e_u - e_{\widehat{u}}) \,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h}.$$

Proof. By the error equation (4.1d) we know that $e_{\widehat{u}} \in M^0_{h,H}$. Taking $(\boldsymbol{v}, w, \mu) = (\boldsymbol{e}_q, e_u, e_{\widehat{u}})$ in the error equations (4.1a)-(4.1c) respectively and adding, we get, after some algebraic manipulation,

$$\begin{aligned}\|\boldsymbol{e}_q\|^2_{\alpha,\Omega} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h} =& - (\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{e}_q)_{\mathcal{T}_h} - \langle (I-P_M)u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &- \langle (I-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\, e_u\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Inserting the identity (4.2) in the above equation, we get

$$\begin{aligned}\|\boldsymbol{e}_q\|^2_{\alpha,\Omega} + \|e_u - e_{\widehat{u}}\|^2_{\tau,\partial\mathcal{T}_h} =& - (\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{e}_q)_{\mathcal{T}_h} - \langle (I-P_M)u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &- \langle (I-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\, e_u\rangle_{\partial\mathcal{T}_h}\\ &+ \langle (P^h_\partial - P_M)(\boldsymbol{q}\cdot\boldsymbol{n} + \tau u)\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\\ =& - (\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{e}_q)_{\mathcal{T}_h} - \langle (I-P_M)u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &- \langle (I-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h} + \langle (P^h_\partial - P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\\ &+ \langle (P^h_\partial - P_M)(\tau u)\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\end{aligned}$$

Now using the fact that $e_{\widehat{u}}$ is single valued on $\mathcal{E}_h$ and $e_{\widehat{u}} = 0$ on $\partial\Omega$ we get

$$\begin{aligned}\|\boldsymbol{e}_q\|^2_{\alpha,\Omega} + \|e_u - e_{\widehat{u}}\|^2_{\tau,\partial\mathcal{T}_h} =& - (\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{e}_q)_{\mathcal{T}_h} - \langle (I-P_M)u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &+ \langle P^h_\partial u - P_M u\,,\,\tau(e_u - e_{\widehat{u}})\rangle_{\partial\mathcal{T}_h} - \langle (I-P^h_\partial)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Finally noticing that on each $F \in \partial\mathcal{T}_h, F\cap\mathcal{E}_H = \emptyset$

$$P_M = P^h_\partial, \quad e_u|_F, e_{\widehat{u}}|_F, \boldsymbol{e}_q\cdot\boldsymbol{n}|_F \in M_h(F),$$

we get the identity

$$\begin{aligned}&-\langle (I-P_M)u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle P^h_\partial u - P_M u\,,\,\tau(e_u - e_{\widehat{u}})\rangle_{\partial\mathcal{T}_h} - \langle (I-P^h_\partial)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\\ &\quad = - \langle (I-P^H_\partial)u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} + \langle P^h_\partial u - P^H_\partial u\,,\,\tau(e_u - e_{\widehat{u}})\rangle_{\partial\mathcal{T}_H} - \langle (I-P^h_\partial)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\, e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_H},\end{aligned}$$

which completes the proof. ∎

Now we are ready to present our first estimate for $\boldsymbol{e}_q$:

THEOREM 4.5. *If Assumption 3.1 -3.3 hold, then we have*

$$\begin{aligned}\|\boldsymbol{e}_q\|_{\alpha,\Omega} + \|e_u - e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h} \le& C_\alpha\|\boldsymbol{\delta}_q\|_{\alpha,\Omega} + C(H^{t-\frac{1}{2}}L^{-\frac{1}{2}} + \tau H^t L^{-\frac{1}{2}})\|u\|_{t+1,\Omega}\\ &+ C\tau h^s L^{-\frac{1}{2}}\|u\|_{s+1,\Omega} + C\tau^{-\frac{1}{2}} h^s L^{-\frac{1}{2}}\|\boldsymbol{q}\|_{s+1,\Omega},\end{aligned}$$

for all $0 \le s \le k+1,\ 0 \le t \le l+1$*. The constants* C *are independent of the mesh size* h, H*.* C_α *solely depends on* α*.*

Proof. We recall that by definition $\|\mu\|^2_{t,\partial\mathcal{T}_H} := \sum_{T\in\mathcal{P}}\|\mu\|^2_{t,\partial T}$. We begin by giving an alternative expression for $\langle u - P^H_\partial u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H}$. Using that $\mathcal{I}^0_H u - P^H_\partial u \in M_H$ and the equation (4.1c) we get

$$\begin{aligned}\langle u - P^H_\partial u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} &= \langle u - \mathcal{I}^0_H u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} + \langle \mathcal{I}^0_H u - P^H_\partial u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H}\\ &= \langle u - \mathcal{I}^0_H u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} + \langle \mathcal{I}^0_H u - P^H_\partial u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H}.\end{aligned}$$

Then further using (4.2) we get

$$\begin{aligned}\langle u - P^H_\partial u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} =& \langle u - \mathcal{I}^0_H u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} - \langle \mathcal{I}^0_H u - P^H_\partial u\,,\,\tau(e_u - e_{\widehat{u}})\rangle_{\partial\mathcal{T}_H}\\ &+ \langle \mathcal{I}^0_H u - P^H_\partial u\,,\,(P^h_\partial - P_M)(\boldsymbol{q}\cdot\boldsymbol{n} + \tau u)\rangle_{\partial\mathcal{T}_H}.\end{aligned}$$

Then using Lemma 4.2 we get the estimate:

$$\begin{aligned}\langle u-\mathcal{I}_H^0 u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} &= \sum_{T\in\mathcal{P}}\|u-\mathcal{I}_H^0 u\|_{\frac12,\partial T}\|\boldsymbol{e}_q\|_{H(div,T)}\\ &= \sum_{T\in\mathcal{P}}\|u-\mathcal{I}_H^0 u\|_{\frac12,\partial T}\|\boldsymbol{e}_q\|_{T)}\\ &\le CH^{t-\frac12}L^{-\frac12}\|u\|_{t+1,\Omega}\|\boldsymbol{e}_q\|_{\Omega},\end{aligned}$$

for all $0\le t\le l+1$, where in the last step we used Lemma 3.5.

By the previous identity and Lemma 4.4, we have

$$\begin{aligned}\|\boldsymbol{e}_q\|_{\alpha,\Omega}^2+\|e_u-e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h}^2 &= -(\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{e}_q)_{\mathcal{T}_h}-\langle u-\mathcal{I}_H^0 u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H}\\ &+\langle \mathcal{I}_H^0 u-P_\partial^H u\,,\,\tau(e_u-e_{\widehat{u}})\rangle_{\partial\mathcal{T}_H}-\langle \mathcal{I}_H^0 u-P_\partial^H u\,,\,(P_\partial^h-P_\partial^H)(\boldsymbol{q}\cdot\boldsymbol{n}+\tau u)\rangle_{\partial\mathcal{T}_H}\\ &+\langle P_\partial^h u-P_\partial^H u\,,\,\tau(e_u-e_{\widehat{u}})\rangle_{\partial\mathcal{T}_H}-\langle (I-P_\partial^h)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,e_u-e_{\widehat{u}}\rangle_{\partial\mathcal{T}_H}\\ &\le \|\boldsymbol{e}_q\|_{\alpha,\Omega}\|\boldsymbol{\delta}_q\|_{\alpha,\Omega}+CH^{t-\frac12}L^{-\frac12}\|u\|_{t+1,\Omega}\|\boldsymbol{e}_q\|_{\Omega}+\tau^{\frac12}\|\mathcal{I}_H^0 u-P_\partial^H u\|_{\partial\mathcal{T}_H}\|e_u-e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_H}\\ &+\|\mathcal{I}_H^0 u-P_\partial^H u\|_{\partial\mathcal{T}_H}\left(\|(I-P_\partial^h)(\boldsymbol{q}\cdot\boldsymbol{n})\|_{\partial\mathcal{T}_H}+\tau\|u-P_\partial^h u\|_{\partial\mathcal{T}_H}\right)\\ &+\tau^{\frac12}\|P_\partial^h u-P_\partial^H u\|_{\partial\mathcal{T}_H}\|e_u-e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_H}+\tau^{-\frac12}\|(I-P_\partial^h)(\boldsymbol{q}\cdot\boldsymbol{n})\|_{\partial\mathcal{T}_H}\|e_u-e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_H}.\end{aligned}$$

By using Young's inequality and Lemma 3.5, after some algebraic manipulations, we obtain

$$\begin{aligned}\|\boldsymbol{e}_q\|_{\alpha,\Omega}+\|e_u-e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h} &\le C_\alpha\|\boldsymbol{\delta}_q\|+C(H^{t-\frac12}L^{-\frac12}+\tau H^t L^{-\frac12})\|u\|_{t+1,\Omega}\\ &+C\tau h^s L^{-\frac12}\|u\|_{s+1,\Omega}+C\tau^{-\frac12}h^s L^{-\frac12}\|\boldsymbol{q}\|_{s+1,\Omega},\end{aligned}$$

for all $0\le t\le l+1,\ 0\le s\le k+1$. This completes the proof. ∎

As a consequence, by triangle inequality, we immediately have the estimate for $\boldsymbol{q}-\boldsymbol{q}_h$:

$$\begin{aligned}\|\boldsymbol{q}-\boldsymbol{q}_h\|_{\alpha,\Omega} &\le (C_\alpha+1)\|\boldsymbol{q}-\boldsymbol{\Pi}_V\boldsymbol{q}\|+CH^{t-\frac12}L^{-\frac12}\|u\|_{t+1,\Omega}\\ &+C\tau H^t L^{-\frac12}\|u\|_{t+1,\Omega}+C\tau h^s L^{-\frac12}\|u\|_{s+1,\Omega}+C\tau^{-\frac12}h^s L^{-\frac12}\|\boldsymbol{q}\|_{s+1,\Omega},\end{aligned}$$

for $1\le s\le k+1, 1\le t\le l+1$.

From this estimate we see that we may have various scenarios in choosing the scales L and H and the stabilization parameter τ. Some of these were discussed in Subsection 3.2 For example, we take $\tau=\mathcal{O}(1)$ and assume $L=\mathcal{O}(1)$, then $\|\boldsymbol{e}_q\|_0, \|e_u-e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h}$ has the order as $\mathcal{O}(h^{k+1}+H^{l+\frac12})$, which is the same as the result in [8].

REMARK 4.6. *It important to note that the fact that $\mathbf{e}_q\in H_{div}$ is essential in obtaining an optimal order of convergence. If $\mathbf{e}_q$ were not H_{div}-conforming, then we will have convergence rate $\mathcal{O}(h^{-\frac12}H^{l+1})$. In the proof, H_{div}-conformity of the vector field $\mathbf{e}_q$ depends essentially on the fact that τ is single faced. It will be interesting to see what kind of numerical result we have if this assumption is failed.*

4.3. Estimate for $u-u_h$. Using a standard elliptic duality argument, we have the following result:

LEMMA 4.7. *We have*

$$\|e_u\|_{\Omega}^2=\mathbb{S}_1+\mathbb{S}_2+\mathbb{S}_3+\mathbb{S}_4, \tag{4.6}$$

where

$$\begin{aligned}\mathbb{S}_1&=-(\alpha\boldsymbol{e}_q\,,\,\boldsymbol{\theta}-\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h}+(\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h},\\ \mathbb{S}_2&=\langle e_u-e_{\widehat{u}}\,,\,\boldsymbol{\theta}\cdot\boldsymbol{n}-P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h}-\langle \boldsymbol{e}_q\cdot\boldsymbol{n}-\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi-P_h^\partial\phi\rangle_{\partial\mathcal{T}_h},\\ \mathbb{S}_3&=-\langle (P_h^\partial-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,P_h^\partial\phi\rangle_{\partial\mathcal{T}_h}-\langle (P_h^\partial-P_M)u\,,\,P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h},\\ \mathbb{S}_4&=-\langle \boldsymbol{e}_q\cdot\boldsymbol{n}\,,\,\phi-\mathcal{I}_H^0\phi\rangle_{\partial\mathcal{T}_H}+\langle \boldsymbol{e}_q\cdot\boldsymbol{n}-\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi-\mathcal{I}_H^0\phi\rangle_{\partial\mathcal{T}_H}.\end{aligned}$$

Proof. We begin by using the second equation (3.10b) of the dual problem to write that

$$\begin{aligned}(e_u\,,\,e_u)_{\mathcal{T}_h} &= (e_u\,,\,\nabla\cdot\boldsymbol{\theta})_{\mathcal{T}_h}\\ &= (e_u\,,\,\nabla\cdot\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\alpha\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\nabla\phi)_{\mathcal{T}_h},\end{aligned}$$

by the first equation (3.10a) of the dual problem. This implies that

$$\begin{aligned}(e_u\,,\,e_u)_{\mathcal{T}_h} = {}& (e_u\,,\,\nabla\cdot\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} - (\alpha\boldsymbol{e}_q\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\nabla\Pi_W\phi)_{\mathcal{T}_h}\\ &+ (e_u\,,\,\nabla\cdot(\boldsymbol{\theta}-\boldsymbol{\Pi}_V\boldsymbol{\theta}))_{\mathcal{T}_h} - (\alpha\boldsymbol{e}_q\,,\,\boldsymbol{\theta}-\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\nabla(\phi-\Pi_W\phi))_{\mathcal{T}_h}.\end{aligned}$$

Taking $\boldsymbol{v} := \boldsymbol{\Pi}_V\boldsymbol{\theta}$ in the first error equation, (4.1a), and $w := \Pi_W\phi$ in the second, (4.1b), we obtain

$$\begin{aligned}(e_u\,,\,e_u)_{\mathcal{T}_h} = {}& (\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} + \langle e_{\widehat{u}}\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\Pi_W\phi\rangle_{\partial\mathcal{T}_h}\\ &+ (e_u\,,\,\nabla\cdot(\boldsymbol{\theta}-\boldsymbol{\Pi}_V\boldsymbol{\theta}))_{\mathcal{T}_h} - (\alpha\boldsymbol{e}_q\,,\,\boldsymbol{\theta}-\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\nabla(\phi-\Pi_W\phi))_{\mathcal{T}_h}\\ &+ \langle (I-P_M)u\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle (I-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,\Pi_W\phi\rangle_{\partial\mathcal{T}_h}\end{aligned}$$

and, after simple algebraic manipulations we get

$$(e_u\,,\,e_u)_{\mathcal{T}_h} = -(\alpha\boldsymbol{e}_q\,,\,\boldsymbol{\theta}-\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} + (\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} + \mathbb{T}, \tag{4.7}$$

where

$$\begin{aligned}\mathbb{T} :={}& \langle e_{\widehat{u}}\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\Pi_W\phi\rangle_{\partial\mathcal{T}_h}\\ &+ \langle (I-P_M)u\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle (I-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,\Pi_W\phi\rangle_{\partial\mathcal{T}_h}\\ &+ (e_u\,,\,\nabla\cdot(\boldsymbol{\theta}-\boldsymbol{\Pi}_V\boldsymbol{\theta}))_{\mathcal{T}_h} - (\boldsymbol{e}_q\,,\,\nabla(\phi-\Pi_W\phi))_{\mathcal{T}_h}.\end{aligned}$$

Integrating by parts for the last two terms and applying the projection properties (3.3a), (3.3b) we have,

$$\begin{aligned}\mathbb{T} ={}& \langle e_{\widehat{u}}\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\Pi_W\phi\rangle_{\partial\mathcal{T}_h}\\ &+ \langle (I-P_M)u\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle (I-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,\Pi_W\phi\rangle_{\partial\mathcal{T}_h}\\ &+ \langle e_u\,,\,(\boldsymbol{\theta}-\boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n}\,,\,\phi-\Pi_W\phi\rangle_{\partial\mathcal{T}_h} := \mathbb{T}_1 + \mathbb{T}_2,\end{aligned}$$

where

$$\begin{aligned}\mathbb{T}_1 :={}& \langle e_u - e_{\widehat{u}}\,,\,(\boldsymbol{\theta}-\boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi-\Pi_W\phi\rangle_{\partial\mathcal{T}_h}\\ \mathbb{T}_2 :={}& \langle (Id-P_M)u\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle (Id-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,\Pi_W\phi\rangle_{\partial\mathcal{T}_h}\\ &+ \langle e_{\widehat{u}}\,,\,\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

We will estimate $\mathbb{T}_1, \mathbb{T}_2$ separately. First we transform $\mathbb{T}_1$ by adding and subtracting the terms $P_h^{\partial}\boldsymbol{\theta}\cdot\boldsymbol{n}$ and $P_h^{\partial}\phi$ to get

$$\begin{aligned}\mathbb{T}_1 ={}& \langle e_u - e_{\widehat{u}}\,,\,P_h^{\partial}(\boldsymbol{\theta}\cdot\boldsymbol{n}) - \boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,P_h^{\partial}\phi - \Pi_W\phi\rangle_{\partial\mathcal{T}_h}\\ &+ \langle e_u - e_{\widehat{u}}\,,\,\boldsymbol{\theta}\cdot\boldsymbol{n} - P_h^{\partial}(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi - P_h^{\partial}\phi\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Then using the identity $\boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n} = \tau(\boldsymbol{\Pi}_V u - P_h^{\partial}u)$, a simple consequence of the projection property (3.5), and error equation (4.2) we get

$$\begin{aligned}\mathbb{T}_1 = {}& \langle e_u - e_{\widehat{u}}\,,\,\boldsymbol{\theta}\cdot\boldsymbol{n} - P_h^{\partial}(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi - P_h^{\partial}\phi\rangle_{\partial\mathcal{T}_h}\\ &- \langle (P_h^{\partial} - P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,P_h^{\partial}\phi - \Pi_W\phi\rangle_{\partial\mathcal{T}_h} - \langle (P_h^{\partial} - P_M)u\,,\,\tau(P_h^{\partial}\phi - \Pi_W\phi)\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Next, we transform the expression $\mathbb{T}_2$ by taking into account that $\langle e_{\widehat{u}}\,,\,\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}=0$ and using the fact that $\Pi_W u|_F, \boldsymbol{\Pi}_V\boldsymbol{q}\cdot\boldsymbol{n}|_F\in M(F)$ for any $F\in\partial\mathcal{T}_h$:

$$\begin{aligned}\mathbb{T}_2 &= \langle (I-P_M)u\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle (I-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,\Pi_W\phi\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi\rangle_{\partial\mathcal{T}_h}\\ &= \langle (P_h^\partial-P_M)u\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} - \langle (P_M^\partial-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,\Pi_W\phi\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Now combining $\mathbb{T}_1,\mathbb{T}_2$ and using the property (3.5) of the projections $\boldsymbol{\Pi}_V$ and Π_W, we get

$$\begin{aligned}\mathbb{T} =&\langle e_u-e_{\widehat{u}}\,,\,\boldsymbol{\theta}\cdot\boldsymbol{n}-P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n}-\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi-P_h^\partial\phi\rangle_{\partial\mathcal{T}_h}\\ &- \langle (P_h^\partial-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,P_h^\partial\phi\rangle_{\partial\mathcal{T}_h} - \langle (P_h^\partial-P_M)u\,,\,P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

To obtain the final estimate, on each interior face $F\in\mathcal{E}_h^0$, $M_{h,H}|_F=M_h(F)$ and $\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\in M_h(F)$, by (4.1c), we note that $\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}$ is single valued on $F\in\mathcal{E}_h^0$. Then we rewrite the last term as follows:

$$\begin{aligned}\langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi\rangle_{\partial\mathcal{T}_h} &= \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi\rangle_{\partial\mathcal{T}_H} = \langle \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi-\mathcal{I}_H^0\phi\rangle_{\partial\mathcal{T}_H}\\ &= \langle \boldsymbol{e}_q\cdot\boldsymbol{n}\,,\,\phi-\mathcal{I}_H^0\phi\rangle_{\partial\mathcal{T}_H} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n}-\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi-\mathcal{I}_H^0\phi\rangle_{\partial\mathcal{T}_H},\end{aligned}$$

where $\mathcal{I}_H^0$ is a C^0 Lagrange interpolant defined in (3.4). At the final step we have used the error equation (4.1c) and the fact that $\phi|_{\partial\Omega}=\mathcal{I}_H^0\phi|_{\partial\phi}=0$

Inserting the above expression into $\mathbb{T}_1+\mathbb{T}_2$, we finally obtain:

$$\begin{aligned}\|e_u\|_\Omega^2 =&-(\alpha\boldsymbol{e}_q\,,\,\boldsymbol{\theta}-\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} + (\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h}\\ &+\langle e_u-e_{\widehat{u}}\,,\,\boldsymbol{\theta}\cdot\boldsymbol{n}-P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n}-\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi-P_h^\partial\phi\rangle_{\partial\mathcal{T}_h}\\ &- \langle (P_h^\partial-P_M)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,P_h^\partial\phi\rangle_{\partial\mathcal{T}_h} - \langle (P_h^\partial-P_M)u\,,\,P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h}\\ &- \langle \boldsymbol{e}_q\cdot\boldsymbol{n}\,,\,\phi-\mathcal{I}_H^0\phi\rangle_{\partial\mathcal{T}_H} + \langle \boldsymbol{e}_q\cdot\boldsymbol{n}-\boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\,,\,\phi-\mathcal{I}_H^0\phi\rangle_{\partial\mathcal{T}_H},\end{aligned}$$

which completes the proof. ∎

Notice that in Lemma 4.7 we established bounds for the projection errors: $\boldsymbol{\theta}\cdot\boldsymbol{n}-P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})$, $\phi-P_h^\partial\phi$, and $\phi-\mathcal{I}_H^0\phi$. However, here we cannot apply the trace estimates of Lemma 3.5 for these terms since the solution of the dual problem $(\phi,\boldsymbol{\theta})$ is only in $H^2(\Omega)\times\boldsymbol{H}^1(\Omega)$. Alternatively, we will bound these terms based on the following result:

Lemma 4.8. *If the function $(\phi,\boldsymbol{\theta})\in H^2(T)\times\boldsymbol{H}^1(T)$, then we have*

$$\|\phi-P_h^\partial\phi\|_{\partial T}\le Ch^{\frac{3}{2}}\|\phi\|_{2,T}, \tag{4.8a}$$

$$\|\boldsymbol{\theta}\cdot\boldsymbol{n}-P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\|_{\partial T}\le Ch^{\frac{1}{2}}\|\boldsymbol{\theta}\|_{1,T}, \tag{4.8b}$$

$$\|\phi-\mathcal{I}_H^0\phi\|_{\partial T}\le CH^{\frac{3}{2}}\|\phi\|_{2,T}, \tag{4.8c}$$

$$\|\phi-\mathcal{I}_H^0\phi\|_{\frac{1}{2},\partial T}\le CH\|\phi\|_{2,T}. \tag{4.8d}$$

Proof. For (4.8a), on each $K\in\mathcal{T}_h$, let $\Pi_h\phi$ denotes the L^2-projection of ϕ onto the local space $W(K)$. We have

$$\begin{aligned}\|\phi-P_h^\partial\phi\|_{\partial T}\le&\sum_{K\in\mathcal{T}_h(T)}\|\phi-P_h^\partial\phi\|_{\partial K}\le\sum_{K\in\mathcal{T}_h(T)}\|\phi-\Pi_h\phi\|_{\partial K}\\ &\le C\sum_{K\in\mathcal{T}_h(T)}(h^{-\frac{1}{2}}\|\phi-\Pi_h\phi\|_K+h^{\frac{1}{2}}\|\nabla(\phi-\Pi_h\phi)\|_K)\quad\text{by the trace inequality (2.1)},\\ &\le C\sum_{K\in\mathcal{T}_h(T)}h^{\frac{3}{2}}\|\phi\|_{2,K}\le Ch^{\frac{3}{2}}\|\phi\|_{2,T}.\end{aligned}$$

The trace inequality (4.8b) can be proven in a similar way.

To prove (4.8c), we proceed as follows: on each subdomain T, based on the partition $\mathcal{E}_H(T)$, we can generate a conforming shape-regular triangulation $\mathcal{T}_H(T)$. Therefore, we can also extend the boundary interpolation $\mathcal{I}_H^0\phi$ to the whole domain T, denoted by $\widetilde{\mathcal{I}_H^0}\phi$. We have

$$\|\phi - \mathcal{I}_H^0\phi\|_{\partial T} \le \sum_{K_H \in \mathcal{T}_H(T)} \|\phi - \widetilde{\mathcal{I}_H^0}\phi\|_{\partial T}\|_{\partial K_H} \le CH^{\frac{3}{2}}\|\phi\|_{2,T}.$$

In the last step we've applied the same trace inequality (2.1) and the standard interpolation approximation property.

For the last inequality, by the definition of $\|\cdot\|_{\frac{1}{2},\partial T}$, we have:

$$\|\phi - \mathcal{I}_H^0\phi\|_{\frac{1}{2},\partial T} = \min_{\tilde{w}\in H^1(T), \tilde{w}|_{\partial T} = \phi - \mathcal{I}_H^c\phi} \|w\|_{1,T} \le \|\phi - \widetilde{\mathcal{I}_H^c}\phi\|_{1,T} \le CH\|\phi\|_{2,T}.$$

This completes the proof. ∎

Now we are ready to establish the final estimate for e_u.

Theorem 4.9. *Let the assumptions of Theorem 4.5 hold and, in addition, let the local space $W(K)$ contain piecewise linear functions for any $K \in \mathcal{T}_h$, i.e. $k \ge 1$. Then*

$$\begin{aligned}\|e_u\|_\Omega \le\; & (C_\alpha h + CH)\|\boldsymbol{e}_q\|_{\alpha,\Omega} + C_\alpha h\|\boldsymbol{\delta}_q\|_{\alpha,\Omega} + C(h^{\frac{1}{2}}\tau^{-\frac{1}{2}} + \tau^{\frac{1}{2}}H^{\frac{3}{2}})\|e_u - e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h} \\ & + CH^{\frac{3}{2}}L^{-\frac{1}{2}}h^s(\|\boldsymbol{q}\|_{s+1} + \tau\|u\|_{s+1}) + CH^tL^{-\frac{1}{2}}(H^{\frac{3}{2}}\|\boldsymbol{q}\|_{t+1} + (h^{\frac{1}{2}} + \tau H^{\frac{3}{2}})\|u\|_{t+1}),\end{aligned}$$

for all $1 < s \le k+1, 0 \le t \le l+1$.

Proof. We will estimate the terms $\mathbb{S}_j$, $j = 1, \cdots, 4$ in the error-norm representation (4.6) separately. By taking $\bar{\boldsymbol{\theta}}$ as the average of $\boldsymbol{\theta}$ for each $K \in \mathcal{T}_h$ we get

$$\begin{aligned}\mathbb{S}_1 &= -(\alpha\boldsymbol{e}_q, \boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} + (\alpha\boldsymbol{\delta}_q, \boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} \\ &= -(\alpha\boldsymbol{e}_q, \boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} - (\alpha\boldsymbol{\delta}_q, \boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} + (\alpha\boldsymbol{\delta}_q, \boldsymbol{\theta} - \bar{\boldsymbol{\theta}})_{\mathcal{T}_h} \\ &\le C_\alpha\|\boldsymbol{e}_q\|_{\alpha,\Omega}\|\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta}\|_0 + C_\alpha\|\boldsymbol{\delta}_q\|_{\alpha,\Omega}\|\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta}\|_0 + C_\alpha\|\boldsymbol{\delta}_q\|_{\alpha,\Omega}\|\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}\|_0 \\ &\le C_\alpha h(\|\boldsymbol{e}_q\|_{\alpha,\Omega} + \|\boldsymbol{\delta}_q\|_{\alpha,\Omega}\|\boldsymbol{\theta}\|_1 \le C_\alpha h(\|\boldsymbol{e}_q\|_{\alpha,\Omega} + \|\boldsymbol{\delta}_q\|_{\alpha,\Omega})\|e_u\|_0. \qquad \text{by (3.11)}\end{aligned}$$

Next we consider the term $\mathbb{S}_2$. First using the fact that $(e_u - e_{\widehat{u}})|_F, \boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}|_F \in M_h(F)$, for all $F \in \partial\mathcal{T}_h, F \cap \mathcal{E}_H = \emptyset$ we transform the integrals over the boundaries of the elements of the fine mesh into integrals on the boundaries on the coarse mesh only:

$$\begin{aligned}\mathbb{S}_2 &= \langle e_u - e_{\widehat{u}}, \boldsymbol{\theta}\cdot\boldsymbol{n} - P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}, \phi - P_h^\partial\phi\rangle_{\partial\mathcal{T}_h} \\ &= \langle e_u - e_{\widehat{u}}, \boldsymbol{\theta}\cdot\boldsymbol{n} - P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_H} - \langle \boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}, \phi - P_h^\partial\phi\rangle_{\partial\mathcal{T}_H}.\end{aligned}$$

Further, using the approximation property (4.8a)-(4.8b) and the regularity assumption (3.11) we get

$$\begin{aligned}\mathbb{S}_2 &\le \|e_u - e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h}\tau^{-\frac{1}{2}}\|\boldsymbol{\theta}\cdot\boldsymbol{n} - P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\|_{\partial\mathcal{T}_H} + \|\boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_H}\|\phi - P_h^\partial\phi\|_{\partial\mathcal{T}_H} \\ &\le Ch^{\frac{1}{2}}\tau^{-\frac{1}{2}}\|e_u - e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h}\|\boldsymbol{\theta}\|_1 + Ch^{\frac{3}{2}}\|\boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_H}\|\phi\|_2 \\ &\le Ch^{\frac{1}{2}}\tau^{-\frac{1}{2}}\|e_u - e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h}\|e_u\|_0 + Ch^{\frac{3}{2}}\|\boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_H}\|e_u\|_0.\end{aligned}$$

Now we consider $\mathbb{S}_3$. Using $\langle(I - P_H^\partial)(\boldsymbol{q}\cdot\boldsymbol{n}), \phi\rangle_{\partial\mathcal{T}_H} = \langle(I - P_H^\partial)u, \boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} = 0$, the approximation properties (3.7), (4.8a), (4.8b), and regularity assumption (3.11) we get

$$\begin{aligned}\mathbb{S}_3 &= -\langle(I - P_M)(\boldsymbol{q}\cdot\boldsymbol{n}), P_h^\partial\phi\rangle_{\partial\mathcal{T}_h} - \langle(I - P_M)u, P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_h} \\ &= \langle(I - P_H^\partial)(\boldsymbol{q}\cdot\boldsymbol{n}), \phi - P_h^\partial\phi\rangle_{\partial\mathcal{T}_H} - \langle(I - P_H^\partial)u, \boldsymbol{\theta}\cdot\boldsymbol{n} - P_h^\partial(\boldsymbol{\theta}\cdot\boldsymbol{n})\rangle_{\partial\mathcal{T}_H} \\ &\le CH^th^{\frac{3}{2}}L^{-\frac{1}{2}}\|\boldsymbol{q}\|_{t+1}\|\phi\|_2 + CH^th^{\frac{1}{2}}L^{-\frac{1}{2}}\|u\|_{t+1}\|\boldsymbol{\theta}\|_1 \\ &\le CH^th^{\frac{3}{2}}L^{-\frac{1}{2}}\|\boldsymbol{q}\|_{t+1}\|e_u\|_0 + CH^th^{\frac{1}{2}}L^{-\frac{1}{2}}\|u\|_{t+1}\|e_u\|_0,\end{aligned}$$

for any $0 \le t \le l+1$.

Next, we estimate the last term. Using (4.8c), (4.8d) and the regularity assumption (3.11), we get

$$\begin{aligned}
\mathbb{S}_4 &= -\langle \boldsymbol{e}_q \cdot \boldsymbol{n} \,,\, \phi - \mathcal{I}_H^0 \phi \rangle_{\partial \mathcal{T}_H} + \langle \boldsymbol{e}_q \cdot \boldsymbol{n} - \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n} \,,\, \phi - \mathcal{I}_H^0 \phi \rangle_{\partial \mathcal{T}_H} \\
&\le \sum_{T \in \mathcal{P}} \|\phi - \mathcal{I}_H^0 \phi\|_{\frac{1}{2}, \partial T} (\|\boldsymbol{e}_q\|_{0,T} + \|\nabla \cdot \boldsymbol{e}_q\|_{0,T}) + \|\boldsymbol{e}_q \cdot \boldsymbol{n} - \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\|_{\partial \mathcal{T}_H} \|\phi - \mathcal{I}_H^0 \phi\|_{\partial \mathcal{T}_H} \\
&\le \sum_{T \in \mathcal{P}} \|\phi - \mathcal{I}_H^0 \phi\|_{\frac{1}{2}, \partial T} \|\boldsymbol{e}_q\|_{0,T} + \|\boldsymbol{e}_q \cdot \boldsymbol{n} - \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\|_{\partial \mathcal{T}_H} \|\phi - \mathcal{I}_H^0 \phi\|_{\partial \mathcal{T}_H} \quad \text{by Lemma 4.2,} \\
&\le CH \left(\|\boldsymbol{e}_q\|_0 + H^{\frac{1}{2}} \|\boldsymbol{e}_q \cdot \boldsymbol{n} - \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\|_{\partial \mathcal{T}_h} \right) \|\phi\|_2 \\
&\le CH \left(\|\boldsymbol{e}_q\|_0 + H^{\frac{1}{2}} \|\boldsymbol{e}_q \cdot \boldsymbol{n} - \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\|_{\partial \mathcal{T}_h} \right) \|e_u\|_0.
\end{aligned}$$

Finally, we show how to bound $\|\boldsymbol{e}_q \cdot \boldsymbol{n} - \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\|_{\partial \mathcal{T}_h}$. By Lemma 4.2, 3.5 and (4.2), we have

$$\begin{aligned}
\|\boldsymbol{e}_q \cdot \boldsymbol{n} - \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\|_{\partial \mathcal{T}_h} &= \|\boldsymbol{e}_q \cdot \boldsymbol{n} - \boldsymbol{e}_{\widehat{q}} \cdot \boldsymbol{n}\|_{\partial \mathcal{T}_H} = \|\tau (e_u - e_{\widehat{u}}) - (P_h^\partial - P_M)(\boldsymbol{q} \cdot \boldsymbol{n} + \tau u)\|_{\partial \mathcal{T}_H} \\
&\le \tau^{\frac{1}{2}} \|e_u - e_{\widehat{u}}\|_{\tau, \partial \mathcal{T}_H} + \|(P_h^\partial - P_H^\partial)(\boldsymbol{q} \cdot \boldsymbol{n} + \tau u)\|_{\partial \mathcal{T}_H} \\
&\le \tau^{\frac{1}{2}} \|e_u - e_{\widehat{u}}\|_{\tau, \partial \mathcal{T}_H} + C h^s L^{-\frac{1}{2}} (\|\boldsymbol{q}\|_{s+1} + \tau \|u\|_{s+1}) + C H^t L^{-\frac{1}{2}} (\|\boldsymbol{q}\|_{t+1} + \tau \|u\|_{t+1}),
\end{aligned}$$

for all $1 \le s \le k+1; 0 \le t \le l+1$.

Combining the estimates for $\mathbb{S}_1 - \mathbb{S}_4$ and grouping the similar terms we get

$$\begin{aligned}
\|e_u\|_0 \le{}& (C_\alpha h + CH) \|\boldsymbol{e}_q\|_{\alpha, \Omega} + C_\alpha h \|\boldsymbol{\delta}_q\|_{\alpha, \Omega} + C (h^{\frac{1}{2}} \tau^{-\frac{1}{2}} + \tau^{\frac{1}{2}} H^{\frac{3}{2}}) \|e_u - e_{\widehat{u}}\|_{\tau, \partial \mathcal{T}_h} \\
&+ C H^{\frac{3}{2}} L^{-\frac{1}{2}} h^s (\|\boldsymbol{q}\|_{s+1} + \tau \|u\|_{s+1}) + C H^t L^{-\frac{1}{2}} (H^{\frac{3}{2}} \|\boldsymbol{q}\|_{t+1} + (h^{\frac{1}{2}} + \tau H^{\frac{3}{2}}) \|u\|_{t+1}),
\end{aligned}$$

for all $1 < s \le k+1, 0 \le t \le l+1$. ☐

5. Multiscale HDG methods. In this section, we will consider the problem involving multiscale features. Namely, let us assume that the *permeability* coefficient α has two separated scales,

$$\alpha(x) = \alpha(x, x/\epsilon), \tag{5.1}$$

where x is called the slowly varying variable and x/ϵ is called the fast varying variable. Under this assumption, the exact solution $u, \boldsymbol{q}$ also has two scales. Therefore, the derivatives of $u, \boldsymbol{q}$ also depends on the small scale ϵ. In fact, the exact solution $(u, \boldsymbol{q})$ asymptotically behaves ([24]) as

$$\|D^k u\|_\Omega = \mathcal{O}(\epsilon^{-(k-1)}), \quad \|D^k \boldsymbol{q}\|_\Omega = \mathcal{O}(\epsilon^{-k}),$$

for all $k \ge 1$. Here $D^k u$ and $D^k \boldsymbol{q}$ denote any k-th partial derivative of u and $\boldsymbol{q}$, respectively. Then, if we set the nonzero τ to be constant 1, by Theorem 3.8, the velocity error becomes:

$$\|\boldsymbol{q} - \boldsymbol{q}_h\|_\Omega \le C \left[\left(\frac{h}{\epsilon} \right)^s + \left(\frac{H}{\epsilon} \right)^{t - \frac{1}{2}} L^{-\frac{1}{2}} \epsilon^{-\frac{1}{2}} + \left(\frac{h}{\epsilon} \right)^s L^{-\frac{1}{2}} \epsilon^{-1} \right].$$

For a multiscale finite element method, the relation between all scales should be $h < \epsilon < H < L$. The above estimate is no longer valuable since $\frac{H}{\epsilon} > 1$. The error for u also has similar problem. In fact, this is a typical drawback for methods using polynomials for both fine and coarse scales. If we look at the estimate carefully, we can see the trouble appears on the term $\frac{H}{\epsilon}$ only. The scale H is solely associated with the coarse space M_H. This suggests that we should define the space M_H in a more appropriate way so that its approximation property is independent of the scale ϵ. This reasoning has been used by Arbogast and Xiao [9] to design a mortar multiscale finite element method that overcomes this deficiency of the standard muliscale method. Their construction is based on the idea of involving the three scales we have used in our considerations. However, instead of using mortar spaces to glue the approximations on the coarse grid, here we use the same mechanism that is provided by the hybridization of the discontinuous Galerkin method.

5.1. Homogenization results. In a very special case of periodic arrangement of the heterogeneous coefficient we propose to use non-polynomial spaces for the Lagrange multipliers that are based on the concept of existence of smooth solution of a homogenized problem and using the first order correction from the homogenization theory.

We first review some classical homogenization results. For more details, we refer readers to [24, 18]. We assume that $\alpha(x, y)$ is periodic in y with the unite cell $\boldsymbol{Y} = [0, 1]^d$ as its period. The homogenized problem is defined as

$$\alpha_0 \nabla u_0 + \boldsymbol{q}_0 = 0 \quad \text{in } \Omega, \tag{5.2a}$$

$$\nabla \cdot \boldsymbol{q}_0 = f \quad \text{in } \Omega, \tag{5.2b}$$

$$u_0 = 0 \quad \text{on } \partial\Omega. \tag{5.2c}$$

Here the homogenized tensor α_0 is defined as

$$\alpha_0^{ij} = \int_{\boldsymbol{Y}} \alpha(x, y)(\delta_{ij} + \frac{\partial \chi_j}{\partial y_i})dy, \quad i, j = 1, 2, \ldots, d.$$

Here $\chi_j, j = 1, 2, \ldots, d$ are the periodic solutions of the following cell problems:

$$\nabla_y \cdot [\alpha(x, y)(\nabla_y \chi_k(x, y) + \boldsymbol{e}_k)] = 0, \quad \text{in } \Omega \times \boldsymbol{Y}, \;\; k = 1, 2, \ldots, d,$$

where $\boldsymbol{e}_k$ is the standard unit vector in $\mathbb{R}^d$. Then for the first order corrector

$$u_\epsilon := u_0 + \epsilon \boldsymbol{\chi} \cdot \nabla u_0$$

with $\boldsymbol{\chi} = (\chi_1, \ldots, \chi_d)$ we have the following result:

LEMMA 5.1. *If $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$, then there is some constant C independent of ϵ, such that*

$$\|u - u_\epsilon\|_0 \leq C\epsilon \|u_0\|_2.$$

Moreover, if $u_0 \in H^2(\Omega) \cap H_0^1(\Omega) \cap W^{1,\infty}(\Omega)$, then we have (e.g., [24])

$$\|\nabla(u - u_\epsilon)\|_0 \leq C(\epsilon \|u_0\|_2 + \sqrt{\epsilon} \|\nabla u_0\|_\infty).$$

5.2. A multiscale coarse space M_H. Using the above basic results form homogenization of heterogeneous differential operators we shall design our multiscale method. As before, we introduce finite element partitioning of the domain. In this setting we assume that the partitions are such that

$$h < \epsilon << H < L \leq 1. \tag{5.3}$$

In this section we shall use the same polynomial spaces $\mathbf{V}(K)$ and $W(K)$ as before. The difference will be in the choice of the coarse space M_H. Here we shall follow the work of Arbogast and Xiao [9], where this construction was used for the mortar finite element method.

For each $F \in \mathcal{E}_H$, let $\bar{F}$ denotes a rectangular neighborhood of F, we define the local space as

$$M_H(F) := \{\mu \in L^2(F) \; : \; \mu = (1 + \epsilon \boldsymbol{\chi} \cdot \nabla)p|_F, \text{for } p \in P^l(\bar{F})\}. \tag{5.4}$$

Notice that, the local space $M_H(F)$ involves both, local cell solutions $\boldsymbol{\chi}$ and polynomial space $P(\bar{F})$. Therefore, its dimension is larger than $P^l(F)$. Simple considerations show that its dimension will depend on the structure of $\boldsymbol{\chi}$ and will between $2l$ and $3l$.

The coarse space M_H is then defined as:

$$M_H := \{\mu \in L^2(\mathcal{E}_{h,H}) \; : \; \text{for } F \in \mathcal{E}_H, \; \mu|_F \in M_H(F), \quad \text{and } \mu|_{\mathcal{E}_h^0 \cup \partial\Omega} = 0\}$$

On each $F \in \mathcal{E}_H$, we define the following projection of u_ϵ on M_H:

$$x \in F \in \mathcal{E}_H: \qquad \mathcal{I}u_\epsilon(x) = (1 + \epsilon\boldsymbol{\chi} \cdot \nabla)\mathcal{I}_H u_0(x).$$

Here $\mathcal{I}_H u_0$ is defined on $\bar{F}$ as the orthogonal L^2-projection of u_0 into $P^l(\bar{F})$. It has the following standard approximation property for $1 \le t \le l+1$ and $F \in \mathcal{E}_H$:

$$\|u_0 - \mathcal{I}_H u_0\|_{r,\bar{F}} \le CH^{t-r}\|u_0\|_{t,\bar{F}},$$

for all $0 \le r \le t \le l+1$. Also, for any function $\xi \in H^1(\bar{F})$, we have the following two trace inequalities:

$$\|\phi\|_{0,F} \le H^{-\frac{1}{2}}\|\xi\|_{0,\bar{F}} + H^{\frac{1}{2}}\|\nabla\xi\|_{0,\bar{F}}, \quad \|\phi\|_{\frac{1}{2},F} \le H^{-1}\|\xi\|_{0,\bar{F}} + \|\nabla\xi\|_{0,\bar{F}}.$$

The above two inequalities can be obtained by a simple scaling argument. Combining the trace inequalities and the approximation property of the interpolation, we get the following estimates:

$$\|u_0 - \mathcal{I}_H u_0\|_{r,F} \le CH^{t-\frac{1}{2}-r}\|u_0\|_{t,\bar{F}}, \tag{5.5a}$$

$$\|\nabla(u_0 - \mathcal{I}_H u_0)\|_{r,F} \le CH^{t-\frac{3}{2}-r}\|u_0\|_{t,\bar{F}}, \tag{5.5b}$$

After summation over the faces $F \in \partial\mathcal{T}_H$ these estimates produce the following bounds:

$$\|u_0 - \mathcal{I}_H u_0\|_{r,\partial\mathcal{T}_H} \le CH^{t-\frac{1}{2}-r}\|u_0\|_{t}, \tag{5.6a}$$

$$\|\nabla(u_0 - \mathcal{I}_H u_0)\|_{r,\partial\mathcal{T}_H} \le CH^{t-\frac{3}{2}-r}\|u_0\|_{t}, \tag{5.6b}$$

for all $1 \le t \le l+1, r = 0, \frac{1}{2}$.

We have the following approximation result:

Lemma 5.2. *Let* (5.1) *hold and let the homogenized solution be sufficiently smooth, i.e.* $u_0 \in H^t(\Omega)$. *Then for* $1 \le t \le l+1$

$$\|u - P_\partial^H u\|_{\partial\mathcal{T}_H} \le \|u - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H} \le C(\epsilon L^{-\frac{1}{2}}\|u_0\|_2 + \sqrt{\epsilon}L^{-\frac{1}{2}}\|\nabla u_0\|_\infty + H^{t-\frac{1}{2}}\|u_0\|_t).$$

Proof. The bound $\|u - P_\partial^H u\|_{\partial\mathcal{T}_H} \le \|u - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H}$ is obvious, since P_∂^H is an orthogonal projection on M_H. Then by the triangle inequality, we have

$$\|u - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H} \le \|u - u_\epsilon\|_{\partial\mathcal{T}_H} + \|u_\epsilon - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H}.$$

Now we estimate the two terms on the right hand side separately. By the trace inequality (2.1) and the approximation property of homogenized solution established in Lemma 5.1, we have

$$\|u - u_\epsilon\|_{\partial\mathcal{T}_H} \le CL^{-\frac{1}{2}}\|u - u_\epsilon\|_{1,\Omega} \le CL^{-\frac{1}{2}}(\epsilon\|u_0\|_2 + \sqrt{\epsilon}\|\nabla u_0\|_\infty).$$

The second term is bounded by using the approximation property (5.6) in the following manner:

$$\begin{aligned}\|u_\epsilon - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H} &= \|(1 + \epsilon\boldsymbol{\chi} \cdot \nabla)(u_0 - \mathcal{I}_H u_0)\|_{\partial\mathcal{T}_H} \\ &\le \|u_0 - \mathcal{I}_H u_0\|_{\partial\mathcal{T}_H} + \epsilon\|\boldsymbol{\chi}\|_\infty\|\nabla(u_0 - \mathcal{I}_H u_0)\|_{\partial\mathcal{T}_H} \\ &\le CH^{t-\frac{1}{2}}\|u_0\|_t + C\epsilon H^{t-\frac{3}{2}}\|u_0\|_t \le CH^{t-\frac{1}{2}}\|u_0\|_t,\end{aligned}$$

for $1 \le t \le l+1$. This completes the proof. ∎

We can prove a better estimate for smoother u_0. To get such a result we need an additional assumption on the space M_H:

Assumption 5.3. *The space* $P^l(\bar{\mathcal{E}}_H) = \cup_{F\in\mathcal{E}_H} P^l(\bar{F})$ *has a subspace which provides an approximation of order* $\mathcal{O}(H^{l+1})$ *for smooth* u_0 *and the restriction on* $\mathcal{E}_H$: $P^l(\bar{\mathcal{E}}_H)|_{\mathcal{E}_H} \subset M_H$ *is* C^1-*conforming over the coarse skeleton* $\mathcal{E}_H$.

If this assumption holds, we can define the interpolation of u_ϵ to be:

$$\mathcal{I}^1 u_\epsilon := (1 + \epsilon \boldsymbol{\chi} \cdot \nabla) \mathcal{I}^1_H u_0,$$

where $\mathcal{I}^1_H u_0$ is the C^1-interpolation of u_0 onto $P^l(\bar{\mathcal{E}}_H)$. Under this assumption, the interpolation $\mathcal{I}^1_H u_0$ has the same approximation property (5.6) as $\mathcal{I}_H u_0$.

In order to improve our estimates, we will need the following approximation result,

LEMMA 5.4. *In addition to Assumption 5.3, assume the homogenized solution of* (5.2) u_0 *belongs to* $H^{l+1}(\Omega) \cap H^2(\Omega) \cap W^{1,\infty}(\Omega)$. *Then on each subdomain* $T \subset \Omega$ *and* $1 \le t \le l+1$ *we have*

$$\langle u - \mathcal{I}^1 u_\epsilon \,,\, \boldsymbol{e}_q \cdot \boldsymbol{n} \rangle_{\partial T} \le C(H^{t-2} \|u_0\|_{t,T} + \epsilon \|u_0\|_{2,T} + \sqrt{\epsilon} \|\nabla u_0\|_{\infty,T}) \|\boldsymbol{e}_q\|_T.$$

Proof. Similar as in the proof of Lemma 5.2, we begin by splitting the term as:

$$\langle u - \mathcal{I}^1 u_\epsilon \,,\, \boldsymbol{e}_q \cdot \boldsymbol{n} \rangle_{\partial T} = \langle u - u_\epsilon \,,\, \boldsymbol{e}_q \cdot \boldsymbol{n} \rangle_{\partial T} + \langle u_\epsilon - \mathcal{I}^1 u_\epsilon \,,\, \boldsymbol{e}_q \cdot \boldsymbol{n} \rangle_{\partial T}.$$

For the first term, we apply the Stokes' Theorem:

$$\begin{aligned} \langle u - u_\epsilon \,,\, \boldsymbol{e}_q \cdot \boldsymbol{n} \rangle_{\partial T} &= \|u - u_\epsilon\|_{1,T} \|\boldsymbol{e}_q\|_{H(div,T)} = \|u - u_\epsilon\|_{1,T} \|\boldsymbol{e}_q\|_T && \text{by Lemma 4.2,} \\ &\le C(\epsilon \|u_0\|_{2,T} + \sqrt{\epsilon} \|\nabla u_0\|_{\infty,T}) \|\boldsymbol{e}_q\|_T && \text{by Lemma 5.1.} \end{aligned}$$

For the second term, by the definition of the interpolation $\mathcal{I}^1 u_\epsilon$, we have:

$$|\langle u_\epsilon - \mathcal{I}^1 u_\epsilon \,,\, \boldsymbol{e}_q \cdot \boldsymbol{n} \rangle_{\partial T}| \le |\langle u_0 - \mathcal{I}^1_H u_0 \,,\, \boldsymbol{e}_q \cdot \boldsymbol{n} \rangle_{\partial T}| + \epsilon \, |\langle \nabla (u_0 - \mathcal{I}^1_H u_0) \,,\, \boldsymbol{\chi} \otimes \boldsymbol{e}_q \cdot \boldsymbol{n} \rangle_{\partial T}|$$

due to Assumption 5.3, we have $u_0 - \mathcal{I}^c_H u_0 \in H^{\frac{2}{3}}(\partial T)$, so we can apply Stokes' Theorem on both terms and obtain

$$\begin{aligned} |\langle u_\epsilon - \mathcal{I}^1 u_\epsilon \,,\, \boldsymbol{e}_q \cdot \boldsymbol{n} \rangle_{\partial T}| &\le \|u_0 - \mathcal{I}^1_H u_0\|_{\frac{1}{2},\partial T} \|\boldsymbol{e}_q\|_{H(div),T} + \epsilon \, \|\nabla (u_0 - \mathcal{I}^1_H u_0)\|_{\frac{1}{2},\partial T} \|\boldsymbol{\chi} \otimes \boldsymbol{e}_q\|_{H(div),T} \\ &\le \|u_0 - \mathcal{I}^1_H u_0\|_{\frac{1}{2},\partial T} \|\boldsymbol{e}_q\|_T \\ &\quad + \epsilon \, \|\nabla (u_0 - \mathcal{I}^1_H u_0)\|_{\frac{1}{2},\partial T} (\|\boldsymbol{\chi}\|_{\infty,T} \|\boldsymbol{e}_q\|_T + \|\nabla \boldsymbol{\chi}\|_{\infty,T} \|\boldsymbol{e}_q\|_T), \end{aligned}$$

here we used the fact that $\|\nabla \cdot \boldsymbol{e}_q\|_T = 0$. Finally, under the assumption (5.1), we have the classic result: $\|\boldsymbol{\chi}\|_{\infty,T} = \mathcal{O}(1)$, $\|\nabla \boldsymbol{\chi}\|_{\infty,T} = \mathcal{O}(\epsilon^{-1})$, see [24]. Applying these results and the approximation result (5.6), we have

$$\begin{aligned} |\langle u_\epsilon - \mathcal{I}^1 u_\epsilon \,,\, \boldsymbol{e}_q \cdot \boldsymbol{n} \rangle_{\partial T}| &\le C H^{t-1} \|u_0\|_{t,T} \|\boldsymbol{e}_q\|_{0,T} + C \epsilon H^{t-2} \|u_0\|_{t,T} (1 + \epsilon^{-1}) \|\boldsymbol{e}_q\|_T \\ &\le C H^{t-2} \|u_0\|_{t,T} \|\boldsymbol{e}_q\|_T, \end{aligned}$$

which completes the proof. ∎

5.3. Estimate for $\boldsymbol{q} - \boldsymbol{q}_h$. We are now ready to establish the following result.

THEOREM 5.5. *Let the coefficient* α *satisfy* (5.1). *Then there exists a constant* C*, independent of* h, H, L *and* ϵ*, such that*

$$\begin{aligned} \|\boldsymbol{e}_q\|_{\alpha,\Omega} + \|e_u - e_{\widehat{u}}\|_{\tau,\mathcal{T}_h} &\le C \|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\alpha,\Omega} + C h^s L^{-\frac{1}{2}} (\tau^{-\frac{1}{2}} \|\boldsymbol{q}\|_{s+1} + \tau^{\frac{1}{2}} \|u\|_{s+1}) \\ &\quad + C(1 + \tau^{\frac{1}{2}} L^{-\frac{1}{2}}) (\epsilon \|u_0\|_2 + \sqrt{\epsilon} \|\nabla u_0\|_\infty) \\ &\quad + C(\tau^{\frac{1}{2}} + h^{-\frac{1}{2}}) H^{t-\frac{1}{2}} \|u_0\|_t, \end{aligned} \tag{5.7}$$

for all $1 \le s \le k+1$, $1 \le t \le l+1$.

Moreover, if Assumption 5.3 holds, then the following improved estimate holds:

$$
\begin{aligned}
\|\boldsymbol{e}_q\|_{\alpha,\Omega} + \|e_u - e_{\widehat{u}}\|_{\tau,\mathcal{T}_h} \le{}& C\|\boldsymbol{q} - \boldsymbol{\Pi}_V \boldsymbol{q}\|_{\alpha,\Omega} + Ch^s L^{-\frac{1}{2}}(\tau^{-\frac{1}{2}}\|\boldsymbol{q}\|_{s+1} + \tau^{\frac{1}{2}}\|u\|_{s+1}) \\
&+ C(1+\tau^{\frac{1}{2}}L^{-\frac{1}{2}})(\epsilon\|u_0\|_2 + \sqrt{\epsilon}\|\nabla u_0\|_\infty) \\
&+ C(\tau^{\frac{1}{2}} + H^{-\frac{3}{2}})H^{t-\frac{1}{2}}\|u_0\|_t,
\end{aligned}
\tag{5.8}
$$

for all $1 \le s \le k+1$, $1 \le t \le l+1$.

Proof. By Lemma 4.4 we have

$$
\begin{aligned}
\|\boldsymbol{e}_q\|^2_{\alpha,\Omega} + \|e_u - e_{\widehat{u}}\|^2_{\tau,\partial\mathcal{T}_h} ={}& -(\alpha\boldsymbol{\delta}_q\,,\,\boldsymbol{e}_q)_{\mathcal{T}_h} - \langle (I-P^H_\partial)u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} \\
&+ \langle P^h_\partial u - P^H_\partial u\,,\,\tau(e_u - e_{\widehat{u}})\rangle_{\partial\mathcal{T}_H} - \langle (I-P^h_\partial)(\boldsymbol{q}\cdot\boldsymbol{n})\,,\,e_u - e_{\widehat{u}}\rangle_{\partial\mathcal{T}_H} \\
={}& T_1 + T_2 + T_3 + T_4.
\end{aligned}
$$

We will estimate these four terms separately. For the first term we easily get

$$
|T_1| \le \|\alpha\|_\infty \|\boldsymbol{\delta}_q\|_{\alpha,\Omega}\|\boldsymbol{e}_q\|_{\alpha,\Omega} \quad \text{for all} \quad 1 \le s \le k+1.
$$

Further, using the approximation properties of P^h_∂ and P^H_∂ established in Lemma 3.5 we get

$$
\begin{aligned}
|T_3| &= |\langle P^h_\partial u - P^H_\partial u\,,\,\tau(e_u - e_{\widehat{u}})\rangle_{\partial\mathcal{T}_H}| \\
&= |\langle u - P^h_\partial u\,,\,\tau(e_u - e_{\widehat{u}})\rangle_{\partial\mathcal{T}_H} - \langle u - P^H_\partial u\,,\,\tau(e_u - e_{\widehat{u}})\rangle_{\partial\mathcal{T}_H}| \\
&\le \tau^{\frac{1}{2}}(\|u - P^h_\partial u\|_{\partial\mathcal{T}_H} + \|u - P^H_\partial u\|_{\partial\mathcal{T}_H})\|e_u - e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h} \\
&\le C\tau^{\frac{1}{2}}(L^{-\frac{1}{2}}h^s\|u\|_{s+1} + \|u - P^H_\partial u\|_{\partial\mathcal{T}_H})\|e_u - e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h} \qquad \text{by (3.6)}
\end{aligned}
$$

and bound the term $\|u - P^H_\partial u\|_{\partial\mathcal{T}_H}$ using Lemma 5.2.

In a similar way we estimate T_4:

$$
|T_4| \le \tau^{-\frac{1}{2}}\|(I-P^h_\partial)(\boldsymbol{q}\cdot\boldsymbol{n})\|_{\partial\mathcal{T}_H}\|e_u - e_{\widehat{u}}\|_{\tau,\mathcal{T}_H} \le C\tau^{-\frac{1}{2}}L^{-\frac{1}{2}}h^s\|\boldsymbol{q}\|_{s+1}\|e_u - e_{\widehat{u}}\|_{\tau,\mathcal{T}_h}.
$$

Finally, for T_2, if we simply apply Cauchy-Schwarz inequality and trace inequality, we will have a term of order $\sqrt{\frac{\epsilon}{h}}$. This estimate is not desirable since $h < \epsilon$. To bound T_2 in a better way, we first rewrite it using the error equation (4.1c):

$$
\begin{aligned}
T_2 &= -\langle u - \mathcal{I}u_\epsilon\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} - \langle \mathcal{I}u_\epsilon - P^H_\partial u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} \\
&= -\langle u - \mathcal{I}u_\epsilon\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} - \langle \mathcal{I}u_\epsilon - P^H_\partial u\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n} - \boldsymbol{e}_{\widehat{q}}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H}.
\end{aligned}
$$

Then using the equation (4.2) we get

$$
\begin{aligned}
T_2 = {}& -\langle u - \mathcal{I}u_\epsilon\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} - \langle \mathcal{I}u_\epsilon - P^H_\partial u\,,\,\tau(e_u - e_{\widehat{u}})\rangle_{\partial\mathcal{T}_H} \\
&+ \langle \mathcal{I}u_\epsilon - P^H_\partial u\,,\,(P^h_\partial - P^H_\partial)(\boldsymbol{q}\cdot\boldsymbol{n} + \tau u)\rangle_{\partial\mathcal{T}_H} := T_{21} + T_{22} + T_{23}.
\end{aligned}
$$

Now we estimate these three terms separately. Now recall that $\|\nabla\cdot\boldsymbol{e}_q\|_T = 0$ for $T \in \mathcal{P}$ by Lemma 4.2. Thus using divergence theorem we get

$$
\begin{aligned}
T_{21} &= \langle u - u_\epsilon\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} + \langle u_\epsilon - \mathcal{I}u_\epsilon\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} \\
&= \sum_{T\in\mathcal{T}}(\nabla(u - u_\epsilon), \boldsymbol{e}_q)_T + \langle u_\epsilon - \mathcal{I}u_\epsilon\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} \\
&\le \sum_{T\in\mathcal{T}}\|u - u_\epsilon\|_{1,T}\|\boldsymbol{e}_q\|_T + \langle (1 + \epsilon\boldsymbol{\chi}\cdot\nabla)(u_0 - \mathcal{I}_H u_0)\,,\,\boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} \\
&\le \|u - u_\epsilon\|_1\|\boldsymbol{e}_q\| + (\|u_0 - \mathcal{I}_H u_0\|_{\partial\mathcal{T}_H} + \epsilon\|\boldsymbol{\chi}\|_\infty\|\nabla(u_0 - \mathcal{I}_H u_0)\|_{\partial\mathcal{T}_H})\|\boldsymbol{e}_q\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_H}.
\end{aligned}
$$

Now using the trace inequality (2.1) (for $D = T$) and (5.6) we get

$$
\begin{aligned}
|T_{21}| &\le C\|\boldsymbol{e}_q\|_0 \left\{ \|u-u_\epsilon\|_1 + h^{-\frac{1}{2}}\left(\|u_0 - \mathcal{I}_H u_0\|_{\partial\mathcal{T}_H} + \epsilon\|\nabla(u_0 - \mathcal{I}_H u_0)\|_{\partial\mathcal{T}_H}\right)\right\} \\
&\le C\|\boldsymbol{e}_q\|_0 \left\{ \epsilon\|u_0\|_2 + \sqrt{\epsilon}\|\nabla u_0\|_\infty + h^{-\frac{1}{2}} H^{t-\frac{1}{2}}\|u_0\|_t \right\}.
\end{aligned}
$$

For estimating T_{22} we apply Cauchy-Schwarz and triangle inequalities

$$
|T_{22}| \le \tau^{\frac{1}{2}}\|\mathcal{I}u_\epsilon - P_\partial^H u\|_{\partial\mathcal{T}_H}\|e_u - e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_H} \le 2\tau^{\frac{1}{2}}\|u - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H}\|e_u - e_{\widehat{u}}\|_{\tau,\partial\mathcal{T}_h}
$$

and then bound $\|u - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H}$ using Lemma 5.2. For T_{23}, we have

$$
\begin{aligned}
|T_{23}| &= |\langle \mathcal{I}u_\epsilon - P_\partial^H u\,,\, (P_\partial^h - I)(\boldsymbol{q}\cdot\boldsymbol{n} + \tau u)\rangle_{\partial\mathcal{T}_H}| \\
&\le 2\|u - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H}(\|\boldsymbol{q}\cdot\boldsymbol{n} - P_\partial^h(\boldsymbol{q}\cdot\boldsymbol{n})\|_{\partial\mathcal{T}_H} + \tau\|u - P_\partial^h u\|_{\partial\mathcal{T}_H}) \\
&\le CL^{-\frac{1}{2}}h^s(\|\boldsymbol{q}\|_{s+1} + \tau\|u\|_{s+1})\|u - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H} \\
&\le C\left\{ (h^s L^{-\frac{1}{2}}(\tau^{-\frac{1}{2}}\|\boldsymbol{q}\|_{s+1} + \tau^{\frac{1}{2}}\|u\|_{s+1}))^2 + (\tau^{\frac{1}{2}}\|u - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H})^2 \right\}.
\end{aligned}
$$

Then we estimate the term $\|u - \mathcal{I}u_\epsilon\|_{\partial\mathcal{T}_H}$ by Lemma 5.2 and get (5.7) by combining the estimates for T_1, T_2, T_3, T_4.

The estimate (5.8) we shall prove under the Assumption 5.3, namely, a C^1-conforming interpolate $\mathcal{I}_H^1 u_\epsilon$ exists and it converges to u_ϵ with order $\mathcal{O}(H^{l+1})$. Then we can replace the L^2 interpolate $\mathcal{I}u_\epsilon$ by $\mathcal{I}_H^1 u_\epsilon$ and then apply Lemma 5.4 to T_{21} to get

$$
T_{21} = \langle u - \mathcal{I}_H^1 u_\epsilon\,,\, \boldsymbol{e}_q\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H} \le C\left(H^{t-2}\|u_0\|_t + \epsilon\|u_0\|_2 + \sqrt{\epsilon}\|\nabla u_0\|_\infty\right)\|\boldsymbol{e}_q\|_0,
$$

which completes the proof. ∎

As a consequence of Theorem 5.5, we immediately obtain an L^2 estimate for $\boldsymbol{q} - \boldsymbol{q}_h$:

COROLLARY 5.6. *Suppose we have the same assumption as Theorem 5.5.Then we have*

$$
\begin{aligned}
\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\alpha,\Omega} &\le C\|\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q}\|_{\alpha,\Omega} + Ch^s L^{-\frac{1}{2}}(\tau^{-\frac{1}{2}}\|\boldsymbol{q}\|_{s+1} + \tau^{\frac{1}{2}}\|u\|_{s+1}) \\
&\quad + C(1 + \tau^{\frac{1}{2}}L^{-\frac{1}{2}})(\epsilon\|u_0\|_2 + \sqrt{\epsilon}\|\nabla u_0\|_\infty) + C(\tau^{\frac{1}{2}} + h^{-\frac{1}{2}})H^{t-\frac{1}{2}}\|u_0\|_t,
\end{aligned}
$$

for all $1 \le s \le k+1$, $1 \le t \le l+1$.

Moreover, if Assumption 5.3 holds, then we have the following improved estimate

$$
\begin{aligned}
\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\alpha,\Omega} &\le C\|\boldsymbol{q} - \boldsymbol{\Pi}_V\boldsymbol{q}\|_{\alpha,\Omega} + Ch^s L^{-\frac{1}{2}}(\tau^{-\frac{1}{2}}\|\boldsymbol{q}\|_{s+1} + \tau^{\frac{1}{2}}\|u\|_{s+1}) \\
&\quad + C(1 + \tau^{\frac{1}{2}}L^{-\frac{1}{2}})(\epsilon\|u_0\|_2 + \sqrt{\epsilon}\|\nabla u_0\|_\infty) + C(\tau^{\frac{1}{2}} + H^{-\frac{3}{2}})H^{t-\frac{1}{2}}\|u_0\|_t,
\end{aligned}
$$

for all $1 \le s \le k+1$, $1 \le t \le l+1$.

5.4. Estimate for $u - u_h$. In Section 3, we use a standard duality argument to get an a priori estimate for $u - u_h$. It is based on the full H^2 regularity assumption (3.11) of the adjoint equation (3.10). When the permeability coefficient α has two separated scales, the regularity assumption is no longer valid. Instead, we consider the following adjoint problem:

(5.9a) $$\boldsymbol{\theta} + \nabla\phi = 0 \quad \text{in } \Omega,$$

(5.9b) $$\nabla\cdot\boldsymbol{\theta} = e_u \quad \text{in } \Omega,$$

(5.9c) $$\phi = 0 \quad \text{on } \partial\Omega.$$

We assume the above problem has full H^2 regularity:

(5.10) $$\|\phi\|_2 + \|\boldsymbol{\theta}\|_1 \le C\|e_u\|_0,$$

where C only depends on the domain Ω.

We are ready to state the estimate for $u - u_h$:

THEOREM 5.7. *Suppose the same assumptions as Theorem 5.5. In addition, we also assume the H^2 regularity* (5.10) *holds. Then we have*

$$\|e_u\|_0 \le C\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\alpha,\Omega} + C\Big(\frac{h}{\tau L}\Big)^{\frac{1}{2}}\|e_u - e_{\widehat{u}}\|_{\tau,\mathcal{T}_h} + C\Big(\frac{h}{L}\Big)^{\frac{1}{2}}\|u - P_\partial^H u\|_{\partial\mathcal{T}_H}.$$

Proof. We begin by the fact that

$$\|e_u\|_0^2 = (e_u\,,\,\nabla\cdot\boldsymbol{\theta})_{\mathcal{T}_h} = (e_u\,,\,\nabla\cdot\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} + (e_u\,,\,\nabla\cdot(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h}$$

Then integrating by parts on the second term and using the property (3.3b) we get

$$\|e_u\|_0^2 = (e_u\,,\,\nabla\cdot\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} + \langle e_u\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.$$

Taking $\boldsymbol{v} = \boldsymbol{\Pi}_V\boldsymbol{\theta}$ in the error equation (4.1a), we have

$$(e_u\,,\,\nabla\cdot\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} = (\alpha(\boldsymbol{q} - \boldsymbol{q}_h)\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} + \langle e_{\widehat{u}}\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle (I - P_M)u\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.$$

Now using the fact that $e_{\widehat{u}}, (I - P_M)u$ are single valued on $\partial\mathcal{T}_h$, so $\langle e_{\widehat{u}}\,,\,\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \langle (I - P_M)u\,,\,\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0$ after some algebraic manipulation, we obtain:

$$\begin{aligned}\|e_u\|_0^2 &= (\alpha(\boldsymbol{q} - \boldsymbol{q}_h)\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} + \langle (I - P_M)u\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &\quad + \langle e_{\widehat{u}}\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle e_u\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}\\ &= (\alpha(\boldsymbol{q} - \boldsymbol{q}_h)\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} - \langle (I - P_M)u\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} + \langle e_u - e_{\widehat{u}}\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

We now estimate the above three terms separately.

$$\begin{aligned}(\alpha(\boldsymbol{q} - \boldsymbol{q}_h)\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} &= (\alpha(\boldsymbol{q} - \boldsymbol{q}_h)\,,\,\boldsymbol{\theta})_{\mathcal{T}_h} - (\alpha(\boldsymbol{q} - \boldsymbol{q}_h)\,,\,\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h}\\ &\le C(\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\alpha,\Omega}\|\boldsymbol{\theta}\|_0 + \|\boldsymbol{q} - \boldsymbol{q}_h\|_{\alpha,\Omega}\|\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta}\|_0)\\ &\le C(\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\alpha,\Omega}\|e_u\|_0 + h\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\alpha,\Omega}\|\phi\|_1)\end{aligned}$$

so that after using the full regularity assumption (5.10) and Lemma 3.4, we get

$$(\alpha(\boldsymbol{q} - \boldsymbol{q}_h)\,,\,\boldsymbol{\Pi}_V\boldsymbol{\theta})_{\mathcal{T}_h} \le C\|\boldsymbol{q} - \boldsymbol{q}_h\|_{\alpha,\Omega}\|e_u\|_0.$$

In order to bound the other two terms, we need the following bound:

$$\|(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_H} \le Ch^{\frac{1}{2}}L^{-\frac{1}{2}}\|e_u\|_0. \tag{5.11}$$

On each $T \in \mathcal{P}$, we use $\mathcal{I}\boldsymbol{\theta}$ to denote the Clément interpolation from $H^1(T)$ into $\boldsymbol{V}(T)$. Then we have

$$\begin{aligned}\|(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_H} &\le \|\mathcal{I}\boldsymbol{\theta}\cdot\boldsymbol{n} - \boldsymbol{\Pi}_V\boldsymbol{\theta}\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_H} + \|(\boldsymbol{\theta} - \mathcal{I}\boldsymbol{\theta})\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_H}\\ &\le Ch^{-\frac{1}{2}}\|\mathcal{I}\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta}\|_0 + \|\boldsymbol{\theta} - \mathcal{I}\boldsymbol{\theta}\|_{\partial\mathcal{T}_H}\\ &\le Ch^{-\frac{1}{2}}(\|\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta}\|_0 + \|\boldsymbol{\theta} - \mathcal{I}\boldsymbol{\theta}\|_0) + \sum_{T\in\mathcal{P}} Ch^{\frac{1}{2}}\|\boldsymbol{\theta}\|_{1,T}\\ &\le Ch^{\frac{1}{2}}(\|\phi\|_1 + \|\boldsymbol{\theta}\|_1) + CL^{-\frac{1}{2}}h^{\frac{1}{2}}\|\boldsymbol{\theta}\|_1 \le Ch^{\frac{1}{2}}L^{-\frac{1}{2}}\|e_u\|_0\end{aligned}$$

the last step is due to the regularity condition (5.10). We now estimate the other two terms. First, we have the equality

$$\langle (I - P_M)u\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = \langle u - P_\partial^H u\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H},$$

due to the fact that $P_M = P_\partial^h$ on $F \in \mathcal{E}_h^0$ and $u - P_M u$ is single valued on $\partial \mathcal{T}_h$. We apply Cauchy-Schwarz inequality to obtain:

$$\begin{aligned}\langle (I - P_M)u\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &\le \|u - P_\partial^H\|_{\partial\mathcal{T}_H}\|(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_H}\\ &\le Ch^{\frac12}L^{-\frac12}\|u - P_\partial^H u\|_{\partial\mathcal{T}_H}\|e_u\|_0.\end{aligned}$$

Finally, consider any interior edge $F \in \mathcal{E}_h^0 \cap \partial\mathcal{T}_h$. If $\tau > 0$ on F, by the identity (4.4), we have

$$\langle e_u - e_{\widehat{u}}\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_F = 0.$$

On the other hand, if $\tau = 0$ on F, by the definition of the projection (3.3c), we still have the above identity. This applies

$$\begin{aligned}\langle e_u - e_{\widehat{u}}\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &= \langle e_u - e_{\widehat{u}}\,,\,(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_H}\\ &\le \tau^{-\frac12}\|e_u - e_{\widehat{u}}\|_{\tau,\mathcal{T}_h}\|(\boldsymbol{\theta} - \boldsymbol{\Pi}_V\boldsymbol{\theta})\cdot\boldsymbol{n}\|_{\partial\mathcal{T}_H}\\ &\le C\tau^{-\frac12}h^{\frac12}L^{-\frac12}\|e_u - e_{\widehat{u}}\|_{\tau,\mathcal{T}_h}\|e_u\|_0,\end{aligned}$$

by the estimate (5.11). The proof is complete by combining the above three estimates. ☐

As a consequence of Theorem 5.7, Theorem 5.5 and Lemma 5.2, we immediately have the following estimate for $u - u_h$:

COROLLARY 5.8. *Let the assumptions of Theorem 5.7 are fulfilled. Then*

$$\begin{aligned}\|u - u_h\|_0 \le Ch^s(\|u\|_s + \tau^{-1}\|\boldsymbol{q}\|_s) &+ C\Big(\frac{h}{L}\Big)^{\frac12}\Big(\epsilon\|u_0\|_2 + \sqrt{\epsilon}\|\nabla u_0\|_\infty + H^{t-1\frac12}\|u_0\|_t\Big)\\ &+ C\Big(1 + \Big(\frac{h}{\tau L}\Big)^{\frac12}\Big)(\|\boldsymbol{e}_q\|_0 + \|e_u - e_{\widehat{u}}\|_{\tau,\mathcal{T}_h}),\end{aligned}$$

for all $1 \le s \le k+1, 1 \le t \le l+1$.

6. Conclusions. In this paper, we introduce a hybrid discontinuous Galerkin method for solving multiscale elliptic equations. This is a first paper in a serious of two papers. In the present paper, we consider polynomial and homogenization-based coarse-grid spaces and lay a foundation of hybrid discontinuous Galerkin methods for solving multiscale flow equations. Our method gives a framework that (1) couples independently generated multiscale basis functions in each coarse patch (2) provides a stable global coupling independent of local discretization, physical scales and contrast (3) allows avoiding any constraints on coarse spaces. Though the coarse spaces in the paper are designed for problems with scale separation, the above properties of our framework are important for extending the method to more challenging multiscale problems with non-separable scales and high contrast. This is a subject of the subsequent paper.

7. Acknowledgements. Y. Efendiev's work is partially supported by the DOE and NSF (DMS 0934837 and DMS 0811180). R. Lazarov's research was supported in parts by NSF (DMS-1016525).

This publication is based in part on work supported by Award No. KUS-C1-016-04, made by King Abdullah University of Science and Technology (KAUST).

REFERENCES

[1] J.E. Aarnes. On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation. *SIAM J. Multiscale Modeling and Simulation*, 2:421–439, 2004.
[2] J.E. Aarnes and Y. Efendiev. Mixed multiscale finite element for stochastic porous media flows. *SIAM J. Sci. Comput.*, 30 (5):2319–2339, 2008.
[3] T. Arbogast. Analysis of a two-scale, locally conservative subgrid upscaling for elliptic problems. *SIAM J. Numer. Anal.*, 42(2):576–598 (electronic), 2004.

[4] T. Arbogast. Homogenization-based mixed multiscale finite elements for problems with anisotropy. *Multiscale Model. Simul.*, 9(2):624–653, 2011.

[5] T. Arbogast and K.J. Boyd. Subgrid upscaling and mixed multiscale finite elements. *SIAM J. Numer. Anal.*, 44(3):1150–1171 (electronic), 2006.

[6] T. Arbogast and D.S. Brunson. A computational method for approximating a Darcy-Stokes system governing a vuggy porous medium. *Comput. Geosci.*, 11(3):207–218, 2007.

[7] T. Arbogast, L.C. Cowsar, M.F. Wheeler, and I. Yotov. Mixed finite element methods on nonmatching multiblock grids. *SIAM J. Numer. Anal.*, 37(4):1295–1315, 2000.

[8] T. Arbogast, G. Pencheva, M.F. Wheeler, and I. Yotov. A multiscale mortar mixed finite element method. *Multiscale Model. Simul.*, 6(1):319–346, 2007.

[9] T. Arbogast and H. Xiao. A miltiscale mortar mixed space based homogenization for heterogeneous elliptic problems. *SIAM J. Numerical Analysis*, 51 (1)(1):377–399, 2013.

[10] C. Bernardi, Y. Maday, and A.T. Patera. Domain decomposition by the mortar element method. In HansG. Kaper, Marc Garbey, and GailW. Pieper, editors, *Asymptotic and Numerical Methods for Partial Differential Equations with Critical Parameters*, volume 384 of *NATO ASI Series*, pages 269–286. Springer Netherlands, 1993.

[11] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.

[12] Y. Chen and L.J. Durlofsky. Adaptive local-global upscaling for general flow scenarios in heterogeneous formations. *Transp. Porous Media*, 62(2):157–185, 2006.

[13] C.-C. Chu, I. G. Graham, and T.-Y. Hou. A new multiscale finite element method for high-contrast elliptic interface problems. *Math. Comp.*, 79(272):1915–1955, 2010.

[14] P. G. Ciarlet. *The finite element method for elliptic problems.* Elsevier North-Holland, New York, 1978.

[15] B. Cockburn, B. Dong, and J. Guzmán. A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems. *Math. Comp.*, 77(264):1887–1916, 2008.

[16] 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. Analysis*, 47(2):1319–1365, 2009.

[17] B. Cockburn, W. Qiu, and K. Shi. Conditions for superconvergence of HDG methods for second-order elliptic problems. *Math. Comp.*, 81(279):1327–1353, 2012.

[18] Y. Efendiev and T.-Y. Hou. *Multiscale Finite element Method*. Springer-Verlag, Berlin, 2012.

[19] Y.R. Efendiev, T.Y. Hou, and X.-H. Wu. Convergence of a nonconforming multiscale finite element method. *Siam J. Numer. Anal.*, 37(3):888–910, 2000.

[20] B. Engquist and Weinan E. The heterogeneous multi-scale methods. *Commun. Math. Sci.*, 1:87–132, 2003.

[21] P. Grisvard. *Elliptic problems in nonsmooth domains*. Pitman, Boston, MA, 1985.

[22] T. Hughes, G. Feijoo, L. Mazzei, and J. Quincy. The variational multiscale method - a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Engrg.*, 166:3–24, 1998.

[23] P. Jenny, S.H. Lee, and H. Tchelepi. Multi-scale finite volume method for elliptic problems in subsurface flow simulation. *J. Comput. Phys.*, 187:47–67, 2003.

[24] V.V. Jikov, S.M. Kozlov, and O.A. Oleinik. *Homogenization of Differential Operators and Integral Functionals.* Springer, 1st edition, 1994.

[25] J. Nolen, G. Papanicolaou, and O. Pironneau. A framework for adaptive multiscale methods for elliptic problems. *Multiscale Model. Simul.*, 7(1):171–196, 2008.