Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.OA-2019-0190

Vol. **14**, No. 1, pp. 144-175
February 2021

# A New Ensemble HDG Method for Parameterized Convection Diffusion PDEs

Yong Yu[1], Gang Chen[1], Liangya Pi[2] and Yangwen Zhang[3,*]

[1] *School of Mathematics, Sichuan University, Chengdu, China*
[2] *Department of Mathematics and Statistical, Missouri University of Science and Technology, Rolla, MO, USA*
[3] *Department of Mathematics Science, University of Delaware, Newark, DE, USA*

Received 17 December 2019; Accepted (in revised version) 19 June 2020

**Abstract.** A new second order time stepping ensemble hybridizable discontinuous Galerkin method for parameterized convection diffusion PDEs with various initial and boundary conditions, body forces, and time depending coefficients is developed. For ensemble solutions in $L^\infty(0,T;L^2(\Omega))$, a superconvergent rate with respect to the freedom degree of the globally coupled unknowns for all the polynomials of degree $k \geq 0$ is established. The results of numerical experiments are consistent with the theoretical findings.

**AMS subject classifications**: 65M60

**Key words**: HDG method, ensemble methods, parameterized convection diffusion PDEs, numerical analysis.

## 1. Introduction

In this work, we propose a new second order time stepping ensemble hybridizable discontinuous Galerkin (HDG) method to efficiently simulate a group of parameterized convection diffusion equations on a Lipschitz polyhedral domain $\Omega \subset \mathbb{R}^d$ $(d \geq 2)$. For $j = 1, \ldots, J$, find $(\boldsymbol{q}_j, u_j)$ satisfying

$$\begin{aligned}
&c_j\boldsymbol{q}_j + \nabla u_j = 0 && \text{in } \Omega \times (0,T],\\
&\partial_t u_j + \nabla\cdot\boldsymbol{q}_j + \boldsymbol{\beta}_j\cdot\nabla u_j = f_j && \text{in } \Omega \times (0,T],\\
&u_j = g_j && \text{on } \partial\Omega \times (0,T],\\
&u_j(\cdot,0) = u_j^0 && \text{in } \Omega,
\end{aligned} \tag{1.1}$$

*Corresponding author. *Email addresses:* `yongyu@scu.edu.cn` (Y. Yu), `cglwdm@scu.edu.cn` (G. Chen), `lpp4f@mst.edu` (L. Pi), `ywzhangf@udel.edu` (Y. Zhang)

 

where

$$c_j := c_j(\boldsymbol{x}, t), \quad f_j := f_j(\boldsymbol{x}, t), \quad g_j := g_j(\boldsymbol{x}, t), \quad \boldsymbol{\beta}_j := \boldsymbol{\beta}_j(\boldsymbol{x}, t), \quad u_j^0 := u_j^0(\boldsymbol{x})$$

are given functions.

For many computational applications in real life, one needs to solve a group of PDEs with different input conditions, like the applications in petroleum engineering, which need to predict the transport properties of rock core-sample in centimeter scale. We need to capture the flow capacity of every single nanopore with different inputs, and the porous media of shale core-sample is composed of more than $10^6$ pores. However, to efficiently simulate a group of PDEs with different inputs is a great challenge.

A first order time stepping ensemble method was proposed by [16] to study a set of $J$ solutions of the Navier-Stokes equations with different initial conditions and forcing terms. The $J$ solutions are computed simultaneously by solving a linear system with one common coefficient matrix and multiple RHS vectors. This leads to a great computational efficiency in linear solvers when either the LU factorization (for small-scale systems) or a block iterative algorithm (for large-scale systems) is used. Later, a second order time stepping ensemble algorithm was designed in [14]. Recently, a new ensemble method was proposed to treat the PDEs which have different coefficients [11, 12]. The ensemble method has been applied to many different models [8–10, 15, 17, 18, 20]. It is worthwhile to mention that the previous works only obtained a *suboptimal* $L^\infty(0, T; L^2(\Omega))$ convergence rate for the ensemble solutions.

More recently, we proposed a first order time stepping ensemble hybridizable discontinuous Galerkin (HDG) method in [3] to study a group of convection diffusion PDEs with different initial conditions, boundary conditions, body forces and coefficients. We obtained an *optimal* $L^\infty(0, T; L^2(\Omega))$ convergence rate for the solutions on a simplex mesh, and we obtained a $L^2(0, T; L^2(\Omega))$ superconvergent rate if the polynomials of degree $k \geq 1$ and the coefficients of the PDEs are independent of time. This ensemble HDG method uses polynomials of degree $k$ for all variables, i.e., the flux variables $\boldsymbol{q}_j$ and the scalar variables $u_j$.

In this work, we devise a new second order time stepping ensemble HDG method for a group of convection diffusion PDEs. We use polynomials degree $k$ to approximate the fluxes and the numerical traces, and use polynomials degree $k + 1$ to approximate the scale variable. This method was proposed by [19] and later analyzed by [21] for a single steady elliptic PDEs, they obtained a superconvergent rate for the scalar variable for all $k \geq 0$. This HDG method has been extended to study the PDEs with a convection term by [23, 24].

In this paper, we first restore the superconvergence for $k = 0$ by modifying the stabilization function in [23]. Next, we show that the new ensemble HDG method can obtain a $L^\infty(0, T; L^2(\Omega))$ superconvergent rate for all $k \geq 0$ on a general polyhedron mesh and without assume the coefficients are independent of time. It is worth mentioning that this new ensemble HDG method keep the advantages of the ensemble methods, i.e., all realizations share one common coefficient matrix and multiple RHS vectors at each time step, which can be solved efficiently by some exist solvers as we mentioned previously.

The paper is organized as follows. We introduce the improved HDG formulation and the ensemble HDG method in Section 2. Next, we give some preliminary materials and prove the ensemble HDG method is conditionally stable in Section 3. Then we give a rigorous error analysis in Section 4. Finally, we provide some numerical experiments to confirm our theoretical result in Section 5.

## 2. The ensemble HDG formulation

The HDG methods were proposed by [6], which are based on a mixed formulation and introduce a numerical flux and a numerical trace to approximate the flux and the trace of the solution. The global system involves the numerical trace only since we can element-by-element eliminate the numerical flux and the solution. Therefore, the HDG methods have a significantly smaller number of globally coupled degrees of freedom comparing to DG methods. The HDG methods have been extended to many models [4,5,7,25,26]. We emphasize that the HDG method in this work is considered to be a superconvergent method. Specifically, if polynomials of degree $k \geq 0$ are used for the numerical traces (global system), then we can obtain $k+2$ order for the scalar variables [22–24]. Hence, from the viewpoint of globally coupled degrees of freedom, this method achieves superconvergence for the scalar variable.

To describe the ensemble HDG method, we introduce some notation. Let $\mathcal{T}_h$ be a collection of disjoint shape regular polyhedral $K$ that partition $\Omega$. Here by shape regular we refer to [2]. Let $\partial\mathcal{T}_h$ denote the set $\{\partial K : K \in \mathcal{T}_h\}$. For an element $K$ of the collection $\mathcal{T}_h$, let $e = \partial K \cap \partial\Omega$ denote the boundary face of $K$ if the $d-1$ Lebesgue measure of $e$ is non-zero. For two elements $K^+$ and $K^-$ of the collection $\mathcal{T}_h$, let $e = \partial K^+ \cap \partial K^-$ denote the interior face between $K^+$ and $K^-$ if the $d-1$ Lebesgue measure of $e$ is non-zero. Let $\mathcal{E}_h^o$ and $\mathcal{E}_h^\partial$ denote the sets of interior and boundary faces, respectively, and let $\mathcal{E}_h$ denote the union of $\mathcal{E}_h^o$ and $\mathcal{E}_h^\partial$. For each $K \in \mathcal{T}_h$, let $h_K$ denote the diameter of the smallest $d$-dimensional ball contain $K$, and $h = \max_{K\in\mathcal{T}_h} h_K$. We finally set

$$(w,v)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}(w,v)_K, \quad \langle\zeta,\rho\rangle_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}\langle\zeta,\rho\rangle_{\partial K},$$

where $(\cdot,\cdot)_K$ and $\langle\cdot,\cdot\rangle_{\partial K}$ denote the standard $L^2$ inner product.

For any integer $k \geq 0$, let $\mathcal{P}^k(K)$ denote the set of polynomials of degree at most $k$ on the element $K$. We recall the standard $L^2$ projection operators $\Pi_\ell : L^2(K) \to \mathcal{P}^\ell(K)$ and $P_M : L^2(e) \to \mathcal{P}^k(e)$ satisfying

$$(\Pi_\ell u, w)_K = (u,w)_K, \quad \forall w \in \mathcal{P}^\ell(K), \tag{2.1a}$$

$$\langle P_M u, \mu\rangle_e = \langle u,\mu\rangle_e, \quad \forall \mu \in \mathcal{P}^k(e). \tag{2.1b}$$

Moreover, the vector $L^2$ projection $\boldsymbol{\Pi}_\ell$ is defined similarly.

We consider the discontinuous finite element spaces:

$$
\begin{aligned}
\boldsymbol{V}_h &:= \left\{\boldsymbol{v} \in \left[L^2(\Omega)\right]^d : \boldsymbol{v}|_K \in \left[\mathcal{P}^k(K)\right]^d, \ \forall K \in \mathcal{T}_h\right\}, \\
W_h &:= \left\{w \in L^2(\Omega) : w|_K \in \mathcal{P}^{k+1}(K), \ \forall K \in \mathcal{T}_h\right\}, \\
M_h(g) &:= \left\{\mu \in L^2(\mathcal{E}_h) : \mu|_e \in \mathcal{P}^k(e), \ \forall e \in \mathcal{E}_h, \mu|_{\mathcal{E}_h^\partial} = P_M g\right\}.
\end{aligned}
$$

For $w_h \in W_h$ and $\boldsymbol{r}_h \in \boldsymbol{V}_h$, let $\nabla v_h$ and $\nabla \cdot \boldsymbol{r}_h$ denote the gradient of $w_h$ and the divergence of $\boldsymbol{r}_h$ applied piecewise on each element $K \in \mathcal{T}_h$.

## 2.1. The improved HDG method

Next, we consider the spatial semidiscretization for (1.1) by an improved HDG method. For all $j = 1, \ldots, J$, find $(\boldsymbol{q}_{jh}, u_{jh}, \widehat{u}_{jh}) \in \boldsymbol{V}_h \times W_h \times M_h(g_j)$ satisfying

$$
\begin{aligned}
&(c_j \boldsymbol{q}_{jh}, \boldsymbol{r}_j)_{\mathcal{T}_h} - (u_{jh}, \nabla \cdot \boldsymbol{r}_j)_{\mathcal{T}_h} + \langle \widehat{u}_{jh}, \boldsymbol{r}_j \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0, \\
&(\partial_t u_{jh}, w_j)_{\mathcal{T}_h} - (\boldsymbol{q}_{jh}, \nabla w_j)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{q}}_{jh} \cdot \boldsymbol{n}, w_j \rangle_{\partial \mathcal{T}_h} - (\boldsymbol{\beta}_j u_{jh}, \nabla w_j)_{\mathcal{T}_h} \\
&\qquad - ((\nabla \cdot \boldsymbol{\beta}_j) u_{jh}, w_j)_{\mathcal{T}_h} + \langle \boldsymbol{\beta}_j \cdot \boldsymbol{n} \widehat{u}_{jh}, w_j \rangle_{\partial \mathcal{T}_h} = (f_j, w_j)_{\mathcal{T}_h}, \\
&\langle \widehat{\boldsymbol{q}}_{jh} \cdot \boldsymbol{n}, \mu_j \rangle_{\partial \mathcal{T}_h} = 0
\end{aligned} \tag{2.2}
$$

for all $(\boldsymbol{r}_j, w_j, \mu_j) \in \boldsymbol{V}_h \times W_h \times M_h(0)$. The numerical traces on $\partial \mathcal{T}_h$ are defined by

$$
\widehat{\boldsymbol{q}}_{jh} \cdot \boldsymbol{n} = \boldsymbol{q}_{jh} \cdot \boldsymbol{n} + h_K^{-1}(P_M u_{jh} - \widehat{u}_{jh}). \tag{2.3}
$$

**Remark 2.1.** The stabilization functions in [23] are defined as following

$$
\widehat{\boldsymbol{q}}_{jh} \cdot \boldsymbol{n} = \boldsymbol{q}_{jh} \cdot \boldsymbol{n} + h_K^{-1}(P_M u_{jh} - \widehat{u}_{jh}) + \tau_j^C (u_{jh} - \widehat{u}_{jh}), \tag{2.4}
$$

where $\tau_j^C$ are positive stabilization functions defined on $\partial \mathcal{T}_h$. Comparing with our stabilization function (2.3), a upwind term in (2.4) was added to guarantee the well-posedness but destroy the superconvergence when $k = 0$ (see [23] for a single convection diffusion PDE and [13] for an optimal control problem).

## 2.2. The ensemble HDG formulation

It is easy to see that the system (2.2)-(2.3) has $J$ different coefficient matrices since $c_j^n$ and $\boldsymbol{\beta}_j^n$ are different for each $j$, the superscript $n$ denotes the function value at the time $t_n$. The main idea of the ensemble algorithms is change the variables $c_j^n$ and $\boldsymbol{\beta}_j^n$ into their ensemble means:

$$
\overline{c}^n = \frac{1}{J} \sum_{j=1}^{J} c_j^n, \quad \overline{\boldsymbol{\beta}}^n = \frac{1}{J} \sum_{j=1}^{J} \boldsymbol{\beta}_j^n. \tag{2.5}
$$

Next, we suppose the time domain is uniformly partition into $N$ steps and the time step is $\Delta t := \frac{T}{N}$. Let $t_n := n\Delta t$ for $n = 2, \ldots, N$, we define

$$\partial_t^+ w^n = \frac{1}{2\Delta t}\left(3w^n - 4w^{n-1} + w^{n-2}\right).$$

For all $j = 1, \ldots, J$ and $n = 2, \ldots, N$, our BDF-2 discretization plus second-order extrapolation on the deviation from the average state ensemble HDG method finds $(\boldsymbol{q}_j^n, u_j^n, \widehat{u}_j^n) \in \boldsymbol{V}_h \times W_h \times M_h(g_j)$ satisfying

$$\begin{aligned}&\left(\overline{c}^n \boldsymbol{q}_{jh}^n, \boldsymbol{r}_j\right)_{\mathcal{T}_h} - \left(u_{jh}^n, \nabla\cdot \boldsymbol{r}_j\right)_{\mathcal{T}_h} + \left\langle \widehat{u}_{jh}^n, \boldsymbol{r}_j\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h}\\ &=\Big(\left(\overline{c}^n - c_j^n\right)\left(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2}\right), \boldsymbol{r}_j\Big)_{\mathcal{T}_h}\end{aligned} \tag{2.6a}$$

for all $\boldsymbol{r}_j \in \boldsymbol{V}_h$, and

$$\begin{aligned}&\left(\partial_t^+ u_{jh}^n, w_j\right)_{\mathcal{T}_h} - \left(\boldsymbol{q}_{jh}^n, \nabla w_j\right)_{\mathcal{T}_h} + \left\langle \widehat{\boldsymbol{q}}_{jh}^n\cdot\boldsymbol{n}, w_j\right\rangle_{\partial\mathcal{T}_h}\\ &- \Big(\nabla\cdot\overline{\boldsymbol{\beta}}^n u_{jh}^n, w_j\Big)_{\mathcal{T}_h} - \Big(\overline{\boldsymbol{\beta}}^n u_{jh}^n, \nabla w_j\Big)_{\mathcal{T}_h} + \Big\langle (\overline{\boldsymbol{\beta}}^n\cdot\boldsymbol{n})\widehat{u}_{jh}^n, v_j\Big\rangle_{\partial\mathcal{T}_h}\\ =& \left(f_j^n, w_j\right)_{\mathcal{T}_h} - \Big(\left[\nabla\cdot\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\right]\left(2u_{jh}^{n-1} - u_{jh}^{n-2}\right), w_j\Big)_{\mathcal{T}_h}\\ &- \Big(\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2u_{jh}^{n-1} - u_{jh}^{n-2}\right), \nabla w_j\Big)_{\mathcal{T}_h}\\ &+ \Big\langle\left[\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\cdot\boldsymbol{n}\right]\left(2\widehat{u}_{jh}^{n-1} - \widehat{u}_{jh}^{n-2}\right), w_j\Big\rangle_{\partial\mathcal{T}_h}\end{aligned} \tag{2.6b}$$

for all $w_j \in W_h$, and

$$\left\langle \widehat{\boldsymbol{q}}_{jh}^n\cdot\boldsymbol{n}, \mu_j\right\rangle_{\partial\mathcal{T}_h} = 0 \tag{2.6c}$$

for all $\mu_j \in M_h(0)$, and the numerical fluxes are defined by

$$\widehat{\boldsymbol{q}}_{jh}^n\cdot\boldsymbol{n} = \boldsymbol{q}_{jh}^n\cdot\boldsymbol{n} + h_K^{-1}\left(P_M u_{jh}^n - \widehat{u}_{jh}^n\right). \tag{2.6d}$$

To start up the second order time stepping ensemble HDG system (2.6), besides the initial condition $(\boldsymbol{q}_{jh}^0, u_{jh}^0, \widehat{u}_{jh}^0)$, we need the information of $(\boldsymbol{q}_{jh}^1, u_{jh}^1, \widehat{u}_{jh}^1)$. We take the initial conditions $u_{jh}^0 = \Pi_{k+1}u_0$, $\boldsymbol{q}_{jh}^0 = -\frac{\nabla u_{jh}^0}{c_j^0}$. Since $u_{jh}^0$ is double-valued on $\mathcal{E}_h$, then the restriction of $u_{jh}^0$ on $\mathcal{E}_h$ is double valued. Therefore, we only take one as the initial condition for $\widehat{u}_{jh}^0$. Followed in [11], $(\boldsymbol{q}_{jh}^1, u_{jh}^1, \widehat{u}_{jh}^1)$ is computed by the following backward Euler ensemble HDG method

$$\left(\overline{c}^1 \boldsymbol{q}_{jh}^1, \boldsymbol{r}_j\right)_{\mathcal{T}_h} - \left(u_{jh}^1, \nabla\cdot\boldsymbol{r}_j\right)_{\mathcal{T}_h} + \left\langle \widehat{u}_{jh}^1, \boldsymbol{r}_j\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} = \Big(\left(\overline{c}^1 - c_j^1\right)\boldsymbol{q}_{jh}^0, \boldsymbol{r}_j\Big)_{\mathcal{T}_h}$$

for all $\boldsymbol{r}_j \in \boldsymbol{V}_h$, and

$$
\begin{aligned}
&\Big(\frac{1}{\Delta}(u_{jh}^1 - u_{jh}^0)t, w_j\Big)_{\mathcal{T}_h} - (\boldsymbol{q}_{jh}^1, \nabla w_j)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{q}}_{jh}^1 \cdot \boldsymbol{n}, w_j \rangle_{\partial\mathcal{T}_h} \\
&- \Big(\nabla \cdot \overline{\boldsymbol{\beta}}^1 u_{jh}^1, w_j\Big)_{\mathcal{T}_h} - \Big(\overline{\boldsymbol{\beta}}^1 u_{jh}^1, \nabla w_j\Big)_{\mathcal{T}_h} + \Big\langle (\overline{\boldsymbol{\beta}}^1 \cdot \boldsymbol{n}) \widehat{u}_{jh}^1, v_j \Big\rangle_{\partial\mathcal{T}_h} \\
= &(f_j^1, w_j)_{\mathcal{T}_h} - \Big([\nabla \cdot (\overline{\boldsymbol{\beta}}^1 - \boldsymbol{\beta}_j^1)] u_{jh}^0, w_j\Big)_{\mathcal{T}_h} - \Big((\overline{\boldsymbol{\beta}}^1 - \boldsymbol{\beta}_j^1) u_{jh}^0, \nabla w_j\Big)_{\mathcal{T}_h} \\
&+ \Big\langle [(\overline{\boldsymbol{\beta}}^1 - \boldsymbol{\beta}_j^1) \cdot \boldsymbol{n}] \widehat{u}_{jh}^0, w_j \Big\rangle_{\partial\mathcal{T}_h}
\end{aligned}
$$

for all $w_j \in W_h$.

The following equivalent system is derived to benefit the theoretical analysis.

**Lemma 2.1.** *System* (2.6a)-(2.6d) *is equivalent to the following system*

$$
\begin{aligned}
&(\overline{c}^n \boldsymbol{q}_{jh}^n, \boldsymbol{r}_j)_{\mathcal{T}_h} - (u_{jh}^n, \nabla \cdot \boldsymbol{r}_j)_{\mathcal{T}_h} + \langle \widehat{u}_{jh}^n, \boldsymbol{r}_j \cdot \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} \\
= &\Big((\overline{c}^n - c_j^n)\left(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2}\right), \boldsymbol{r}_j\Big)_{\mathcal{T}_h},
\end{aligned} \tag{2.7a}
$$

$$
\begin{aligned}
&(\partial_t^+ u_{jh}^n, w_j)_{\mathcal{T}_h} + (\nabla \cdot \boldsymbol{q}_{jh}^n, w_j)_{\mathcal{T}_h} - \langle \boldsymbol{q}_{jh}^n \cdot \boldsymbol{n}, \mu_j \rangle_{\partial\mathcal{T}_h} \\
&- \Big((\nabla \cdot \overline{\boldsymbol{\beta}}^n) u_{jh}^n, w_j\Big)_{\mathcal{T}_h} - \Big(\overline{\boldsymbol{\beta}}^n u_{jh}^n, \nabla w_j\Big)_{\mathcal{T}_h} + \Big\langle (\overline{\boldsymbol{\beta}}^n \cdot \boldsymbol{n}) \widehat{u}_{jh}^n, w_j \Big\rangle_{\partial\mathcal{T}_h} \\
&+ \Big\langle h_K^{-1}(P_M u_{jh}^n - \widehat{u}_{jh}^n), P_M w_j - \mu_j \Big\rangle_{\partial\mathcal{T}_h} \\
= &(f_j^n, w_j)_{\mathcal{T}_h} - \Big([\nabla \cdot (\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)]\left(2u_{jh}^{n-1} - u_{jh}^{n-2}\right), w_j\Big)_{\mathcal{T}_h} \\
&- \Big((\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)\left(2u_{jh}^{n-1} - u_{jh}^{n-2}\right), \nabla w_j\Big)_{\mathcal{T}_h} \\
&+ \Big\langle [(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n) \cdot \boldsymbol{n}]\left(2\widehat{u}_{jh}^{n-1} - \widehat{u}_{jh}^{n-2}\right), w_j \Big\rangle_{\partial\mathcal{T}_h}
\end{aligned} \tag{2.7b}
$$

*for all* $(\boldsymbol{r}_j, w_j, \mu_j) \in \boldsymbol{V}_h \times W_h \times M_h(0)$.

The proof of Lemma 2.1 is simply by substituting (2.6d) into (2.6a)-(2.6c), substracting (2.6c) from (2.6b) and using integration by parts.

## 3. Stability

Throughout the paper, we use the standard notation $W^{m,p}(D)$ for the Sobolev spaces on $D$ with norm $\|\cdot\|_{m,p,D}$ and seminorm $|\cdot|_{m,p,D}$. We use $H^m(D)$ instead of $W^{m,p}(D)$ when $p = 2$. We omit the index $p$ and $D$ in the corresponding norms and the seminorms when $p = 2$ or $D = \Omega$. Also, we omit the index $m$ when $m = 0$ in the corresponding norms. We denote by $C(0, T; W^{m,s}(\Omega))$ the Banach space of all continuous functions from $[0, T]$ into $W^{m,s}(\Omega)$. The definition of $L^p(0, T; W^{m,s}(\Omega))$ with $1 \leq p \leq \infty$ is similar.

To obtain the stability of (2.1) in this section, we assume $f_j \in C(0,T;L^2(\Omega))$, $g_j \in H^1(0,T;H^{1/2}(\partial\Omega))$, $u_j^0 \in L^2(\Omega)$ and the vector fields $\boldsymbol{\beta}_j \in C(0,T;[W^{1,\infty}(\Omega)]^d)$ and satisfying

$$\nabla \cdot \boldsymbol{\beta}_j \leq 0, \quad \boldsymbol{\beta}_j = \mathcal{O}(1). \tag{3.1}$$

These exists a positive constant $c_0$ such that the coefficients $c_j > c_0$, and $c_j \in C(0,T;L^\infty(\Omega))$, and the ensemble mean satisfy the following condition

$$\left|c_j^n - \overline{c}^n\right| < \frac{1}{3}\min\left\{\overline{c}^n, \overline{c}^{n-1}, \overline{c}^{n-2}\right\}, \quad n = 2,\ldots,N, \tag{3.2a}$$

$$\left|c_j^1 - \overline{c}^1\right| < \min\left\{\overline{c}^1, \overline{c}^0\right\}. \tag{3.2b}$$

The following error estimates for the $L^2$ projections are standard:

**Lemma 3.1.** *Suppose integers $k,\ell \geq 0$. There exists a constant $C$ independent of $K \in \mathcal{T}_h$ such that*

$$\|w - \Pi_\ell w\|_K \leq Ch^{\ell+1}|w|_{\ell+1,K}, \qquad \forall\, w \in H^{\ell+1}(K), \tag{3.3a}$$

$$\|w - P_M w\|_{\partial K} \leq Ch^{k+\frac{1}{2}}|w|_{k+1,K}, \quad \forall\, w \in H^{k+1}(K). \tag{3.3b}$$

We also use the following local inverse inequality:

$$\|w_h\|_{\partial K} \leq Ch_K^{-\frac{1}{2}}\|w_h\|_K, \quad \forall\, w_h \in W_h. \tag{3.4}$$

## 3.1. Preliminary material

Next, we give the following several lemmas, which will be frequently used in our analysis.

**Lemma 3.2.** *For any real numbers $a, b$ and $c$, we have*

$$\begin{aligned}&\frac{1}{2}(3a-4b+c)a\\ =&\frac{1}{4}\left[a^2+(2a-b)^2-b^2-(2b-c)^2\right]+\frac{1}{4}(a-2b+c)^2.\end{aligned}$$

**Lemma 3.3.** *For $\boldsymbol{\gamma} \in [W^{1,\infty}(\Omega)]^d$ and $w \in W_h$, we have*

$$(\boldsymbol{\gamma} w, \nabla w)_{\mathcal{T}_h} = \frac{1}{2}\langle \boldsymbol{\gamma}\cdot\boldsymbol{n} w, w\rangle_{\partial\mathcal{T}_h} - \frac{1}{2}(\nabla\cdot\boldsymbol{\gamma} w, w)_{\mathcal{T}_h}. \tag{3.5}$$

The proofs of Lemmas 3.2 and 3.3 are trivial and we omit them here.

**Lemma 3.4.** *Suppose the function $v := v(\boldsymbol{x},t)$ is smooth enough, then the following estimates hold true*

$$\left\|\partial_t^+ v^n\right\|_{\mathcal{T}_h}^2 \leq C\Delta t^{-1}\|\partial_t v\|^2_{[L^2(t_{n-2},t_n);L^2(\Omega)]}, \tag{3.6a}$$

$$\Delta t^4\left\|\partial_{tt}^+ v^n\right\|_{\mathcal{T}_h}^2 \leq C\Delta t^3\|\partial_{tt} v\|^2_{[L^2(t_{n-2},t_n);L^2(\Omega)]}, \tag{3.6b}$$

$$\left\|\partial_t v^n - \partial_t^+ v^n\right\|_{\mathcal{T}_h}^2 \leq C\Delta t^3\|\partial_{ttt} v\|^2_{[L^2(t_{n-2},t_n);L^2(\Omega)]}, \tag{3.6c}$$

*where*

$$\partial_{tt}^{+} v^n = \frac{1}{\Delta t^2}\left(v^n - 2v^{n-1} + v^{n-2}\right).$$

The proof of (3.6c) can be found in [11], the proofs of (3.6a)-(3.6b) are very similar to the proof of (3.6c) and hence we omit them.

The following lemma is very crucial for our analysis.

**Lemma 3.5.** *For* $\boldsymbol{\gamma} \in [W^{1,\infty}(\Omega)]^d$, $(w,\mu) \in W_h \times M_h(0)$, $\nabla \cdot \boldsymbol{\gamma} \le 0$ *and* $h$ *small enough, we have*

$$\begin{aligned}
&\left\|h_K^{-\frac{1}{2}}(P_M w - \mu)\right\|_{\partial\mathcal{T}_h}^2 - (\nabla\cdot\boldsymbol{\gamma} w, w)_{\mathcal{T}_h} - (\boldsymbol{\gamma} w, \nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{\gamma}\cdot\boldsymbol{n}\mu, w\rangle_{\partial\mathcal{T}_h}\\
\ge&\frac{1}{2}\left\|h_K^{-\frac{1}{2}}(P_M w - \mu)\right\|_{\partial\mathcal{T}_h}^2 - Ch\|\nabla w\|_{\mathcal{T}_h}^2. 
\end{aligned} \tag{3.7}$$

*Proof.* Using $\langle \boldsymbol{\gamma}\cdot\boldsymbol{n}\mu, \mu\rangle_{\partial\mathcal{T}_h} = 0$, $\nabla\cdot\boldsymbol{\gamma} \le 0$ and integration by parts, we have

$$\begin{aligned}
&-(\nabla\cdot\boldsymbol{\gamma} w, w)_{\mathcal{T}_h} - (\boldsymbol{\gamma} w, \nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{\gamma}\cdot\boldsymbol{n}\mu, w\rangle_{\partial\mathcal{T}_h}\\
=&-\frac{1}{2}\big\langle \boldsymbol{\gamma}\cdot\boldsymbol{n}(w-\mu), w-\mu\big\rangle_{\partial\mathcal{T}_h} - \frac{1}{2}(\nabla\cdot\boldsymbol{\gamma} w, w)_{\mathcal{T}_h} \qquad \text{by (3.5)}\\
=&-\frac{1}{2}\big\langle \boldsymbol{\gamma}\cdot\boldsymbol{n}(w-P_M w), w-P_M w\big\rangle_{\partial\mathcal{T}_h}\\
&-\big\langle \boldsymbol{\gamma}\cdot\boldsymbol{n}(P_M w-\mu), P_M w-\mu\big\rangle_{\partial\mathcal{T}_h}\\
&-\frac{1}{2}\big\langle \boldsymbol{\gamma}\cdot\boldsymbol{n}(P_M w-\mu), P_M w-\mu\big\rangle_{\partial\mathcal{T}_h} - \frac{1}{2}(\nabla\cdot\boldsymbol{\gamma} w, w)_{\mathcal{T}_h}\\
\ge&-C\left(h\|\nabla w\|_{\mathcal{T}_h}^2 + h^{\frac{1}{2}}\|\nabla w\|_{\mathcal{T}_h}\|P_M w-\mu\|_{\partial\mathcal{T}_h}\right)\\
&-\frac{1}{2}\big\langle \boldsymbol{\gamma}\cdot\boldsymbol{n}(P_M w-\mu), P_M w-\mu\big\rangle_{\partial\mathcal{T}_h} \qquad \text{by (3.4)}\\
\ge&-Ch\|\nabla w\|_{\mathcal{T}_h}^2 - \frac{1}{4}\left\|h_K^{-\frac{1}{2}}(P_M w-\mu)\right\|_{\partial\mathcal{T}_h}^2\\
&-\frac{1}{2}\big\langle \boldsymbol{\gamma}\cdot\boldsymbol{n}(P_M w-\mu), P_M w-\mu\big\rangle_{\partial\mathcal{T}_h}.
\end{aligned}$$

The mesh size $h$ small enough and $\boldsymbol{\gamma} \in [W^{1,\infty}(\Omega)]^d$ imply $\frac{1}{4}h_K^{-1} - \frac{1}{2}\boldsymbol{\gamma}\cdot\boldsymbol{n} \ge 0$, therefore,

$$\begin{aligned}
&\left\|h_K^{-\frac{1}{2}}(P_M w - \mu)\right\|_{\partial\mathcal{T}_h}^2 - (\nabla\cdot\boldsymbol{\gamma} w, w)_{\mathcal{T}_h} - (\boldsymbol{\gamma} w, \nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{\gamma}\cdot\boldsymbol{n}\mu, w\rangle_{\partial\mathcal{T}_h}\\
\ge&\frac{1}{2}\left\|h_K^{-\frac{1}{2}}(P_M w - \mu)\right\|_{\partial\mathcal{T}_h}^2 - Ch\|\nabla w\|_{\mathcal{T}_h}^2.
\end{aligned}$$

□

**Lemma 3.6.** *Let* $(\boldsymbol{q}_{jh}^n, u_{jh}^n, \widehat{u}_{jh}^n)$ *be the solution of* (2.7)*, then we have the following bound*

$$
\begin{aligned}
&\|\nabla u_{jh}^n\|_{\mathcal{T}_h} \\
\leq & C\left(\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\|_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\|_{\mathcal{T}_h} + \left\|h_K^{-\frac{1}{2}}(P_M u_{jh}^n - \widehat{u}_{jh}^n)\right\|_{\partial\mathcal{T}_h}\right).
\end{aligned}
$$

*Proof.* We take $\boldsymbol{r}_j = \nabla u_{jh}^n$ in Eq. (2.7a) and use integration by parts to get

$$
\begin{aligned}
\|\nabla u_{jh}^n\|_{\mathcal{T}_h}^2 =& -\left(\overline{c}^n\boldsymbol{q}_{jh}^n, \nabla u_{jh}^n\right)_{\mathcal{T}_h} + \left\langle u_{jh}^n - \widehat{u}_{jh}^n, \nabla u_{jh}^n\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} \\
&+ \left(\left(\overline{c}^n - c_j^n\right)\left(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2}\right), \nabla u_{jh}^n\right)_{\mathcal{T}_h} \\
=& -\left(\overline{c}^n\boldsymbol{q}_{jh}^n, \nabla u_{jh}^n\right)_{\mathcal{T}_h} + \left\langle P_M u_{jh}^n - \widehat{u}_{jh}^n, \nabla u_{jh}^n\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} \\
&+ \left(\left(\overline{c}^n - c_j^n\right)\left(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2}\right), \nabla u_{jh}^n\right)_{\mathcal{T}_h},
\end{aligned}
$$

then the desired result is followed by the Cauchy-Schwarz inequality and the local inverse inequality (3.4). ☐

**Lemma 3.7** (Discrete Poincaré-Friedrichs inequality)**.** *For all* $(w,\mu)\in W_h\times M_h(0)$*, we have*

$$
\|w\|_{\mathcal{T}_h} \leq C\|\nabla w\|_{\mathcal{T}_h} + C\left\|h_K^{-\frac{1}{2}}(w-\mu)\right\|_{\partial\mathcal{T}_h}.
$$

The proof of Lemma 3.7 is found in [2, Lemma 5].

**Lemma 3.8.** *For all* $\boldsymbol{\gamma}\in[W^{1,\infty}(\Omega)]^d$ *and* $(v,w,\widehat{v},\widehat{w})\in W_h\times W_h\times M_h(0)\times M_h(0)$*, we have*

$$
\begin{aligned}
&-(\nabla\cdot\boldsymbol{\gamma}w, v)_{\mathcal{T}_h} - (\boldsymbol{\gamma}w,\nabla v)_{\mathcal{T}_h} + \langle\boldsymbol{\gamma}\cdot\boldsymbol{n}\widehat{w}, v\rangle_{\partial\mathcal{T}_h} \\
\leq & C\left(\|w\|_{\mathcal{T}_h}^2 + \|v\|_{\mathcal{T}_h}^2 + \left\|h_K^{\frac{1}{2}}(P_M v - \widehat{v})\right\|_{\partial\mathcal{T}_h}^2\right) \\
&+ (\nabla\cdot[\boldsymbol{\Pi}_0\boldsymbol{\gamma}w], v)_{\mathcal{T}_h} - \langle\boldsymbol{\Pi}_0\boldsymbol{\gamma}\cdot\boldsymbol{n}w,\widehat{v}\rangle_{\partial\mathcal{T}_h} + \langle\boldsymbol{\gamma}\cdot\boldsymbol{n}(\widehat{w}-w), v-\widehat{v}\rangle_{\partial\mathcal{T}_h}. \qquad (3.8)
\end{aligned}
$$

*Proof.* We note that $\langle\boldsymbol{\gamma}\cdot\boldsymbol{n}\widehat{w},\widehat{v}\rangle_{\partial\mathcal{T}_h} = 0$, then

$$
\begin{aligned}
&-(\nabla\cdot\boldsymbol{\gamma}w, v)_{\mathcal{T}_h} - (\boldsymbol{\gamma}w,\nabla v)_{\mathcal{T}_h} + \langle\boldsymbol{\gamma}\cdot\boldsymbol{n}\widehat{w}, v\rangle_{\partial\mathcal{T}_h} \\
=&(\boldsymbol{\gamma}\cdot\nabla w, v)_{\mathcal{T}_h} + \langle\boldsymbol{\gamma}\cdot\boldsymbol{n}(\widehat{w}-w), v\rangle_{\partial\mathcal{T}_h} \qquad \text{by (3.5)} \\
=&(\boldsymbol{\gamma}\cdot\nabla w, v)_{\mathcal{T}_h} - \langle\boldsymbol{\gamma}\cdot\boldsymbol{n}w,\widehat{v}\rangle_{\partial\mathcal{T}_h} + \langle\boldsymbol{\gamma}\cdot\boldsymbol{n}(\widehat{w}-w), v-\widehat{v}\rangle_{\partial\mathcal{T}_h} \\
=&\left((\boldsymbol{\gamma}-\boldsymbol{\Pi}_0\boldsymbol{\gamma})\cdot\nabla w, v\right)_{\mathcal{T}_h} - \left\langle(\boldsymbol{\gamma}-\boldsymbol{\Pi}_0\boldsymbol{\gamma})\cdot\boldsymbol{n}w,\widehat{v}\right\rangle_{\partial\mathcal{T}_h} \\
&+ \left(\boldsymbol{\Pi}_0\boldsymbol{\gamma}\cdot\nabla w, v\right)_{\mathcal{T}_h} - \left\langle\boldsymbol{\Pi}_0\boldsymbol{\gamma}\cdot\boldsymbol{n}w,\widehat{v}\right\rangle_{\partial\mathcal{T}_h} + \left\langle\boldsymbol{\gamma}\cdot\boldsymbol{n}(\widehat{w}-w),(v-\widehat{v})\right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

We use integration by parts to get

$$\begin{aligned}
&-(\nabla\cdot\boldsymbol{\gamma} w, v)_{\mathcal{T}_h}-(\boldsymbol{\gamma} w,\nabla v)_{\mathcal{T}_h}+\langle\boldsymbol{\gamma}\cdot\boldsymbol{n}\widehat{w},v\rangle_{\partial\mathcal{T}_h}\\
=&-(\nabla\cdot(\boldsymbol{\gamma}-\boldsymbol{\Pi}_0\boldsymbol{\gamma})w,v)_{\mathcal{T}_h}-((\boldsymbol{\gamma}-\boldsymbol{\Pi}_0\boldsymbol{\gamma})\cdot\nabla v,w)_{\mathcal{T}_h}\\
&+\left\langle(\boldsymbol{\gamma}-\boldsymbol{\Pi}_0\boldsymbol{\gamma})\cdot\boldsymbol{n}w,v-\widehat{v}\right\rangle_{\partial\mathcal{T}_h}\\
&+\left(\boldsymbol{\Pi}_0\boldsymbol{\gamma}\cdot\nabla w,v\right)_{\mathcal{T}_h}-\left\langle(\boldsymbol{\Pi}_0\boldsymbol{\gamma}\cdot\boldsymbol{n}w,\widehat{v}\right\rangle_{\partial\mathcal{T}_h}\\
&+\left\langle\boldsymbol{\gamma}\cdot\boldsymbol{n}(\widehat{w}-w),(v-\widehat{v})\right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}$$

Since $\boldsymbol{\gamma}\in[W^{1,\infty}(\Omega)]^d$, then $\|\boldsymbol{\gamma}-\boldsymbol{\Pi}_0\boldsymbol{\gamma}\|_{0,\infty,K}\le Ch_K\|\boldsymbol{\gamma}\|_{1,\infty,K}$. Use the local inverse inequality (3.4) to get

$$\begin{aligned}
&-(\nabla\cdot\boldsymbol{\gamma} w,v)_{\mathcal{T}_h}-(\boldsymbol{\gamma} w,\nabla v)_{\mathcal{T}_h}+\langle\boldsymbol{\gamma}\cdot\boldsymbol{n}\widehat{w},v\rangle_{\partial\mathcal{T}_h}\\
\le& C\left(\|w\|_{\mathcal{T}_h}^2+\|v\|_{\mathcal{T}_h}^2+\left\|h_K^{\frac{1}{2}}(P_Mv-\widehat{v})\right\|_{\partial\mathcal{T}_h}^2\right)\\
&+\left(\nabla\cdot[\boldsymbol{\Pi}_0\boldsymbol{\gamma} w],v\right)_{\mathcal{T}_h}-\left\langle\boldsymbol{\Pi}_0\boldsymbol{\gamma}\cdot\boldsymbol{n}w,\widehat{v}\right\rangle_{\partial\mathcal{T}_h}+\left\langle\boldsymbol{\gamma}\cdot\boldsymbol{n}(\widehat{w}-w),v-\widehat{v}\right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}$$

This proves the desired result. □

## 3.2. Stability

Next, we prove the Ensemble HDG system (2.1) is conditionally stable. Unlike the previous works, we do not assume the Dirichlet boundary conditions are zeros. Hence, the proof here is more involved.

**Theorem 3.1.** *The ensemble HDG system* (2.1) *is conditionally stable, i.e. the condition stable under the assumption* (3.2). *In particular, for $j=1,\dots,J$, we have*

$$\begin{aligned}
&\max_{2\le n\le N}\|u_{jh}^n\|_{\mathcal{T}_h}^2+\Delta t\sum_{n=2}^N\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2\\
\le& C\Delta t\sum_{n=2}^N\left(\|f_j^n\|_{\mathcal{T}_h}^2+\|g_j^n\|_{\frac{1}{2},\partial\Omega}^2\right)\\
&+C\left(\|u_{jh}^0\|_{\mathcal{T}_h}^2+\|u_{jh}^1\|_{\mathcal{T}_h}^2+\Delta t\|\sqrt{\overline{c}^1}\boldsymbol{q}_{jh}^1\|_{\mathcal{T}_h}^2+\|\partial_t g_j\|_{L^2(0,T;H^{\frac{1}{2}}(\partial\Omega))}^2\right),
\end{aligned}$$

*and the constant $C$ depends on $\boldsymbol{\beta}_j$ and $c_j$.*

The proof of Theorem 3.1 follows by triangle inequality, the definition of $H^{\frac{1}{2}}$ norm and Lemma 3.10.

To deal with the inhomogeneous boundary condition in the stability analysis, we need some additional notation. Let $m_j\in H^1(0,T;H^1(\Omega))$ be an arbitrary function such that $m_j|_{\partial\Omega}=g_j$, and define

$$w_{jh}^n=u_{jh}^n-\Pi_{k+1}m_j^n,\quad \widehat{w}_{jh}^n=\widehat{u}_{jh}^n-P_Mm_j^n. \tag{3.9}$$

This implies $\widehat{w}_{jh}^n=0$ on $\mathcal{E}_h^\partial$. Now we give the estimate for $w_{jh}^n$.

**Lemma 3.9.** *Let* $(w_{jh}^n, \widehat{w}_{jh}^n)$ *be defined in* (3.9) *and* $(\boldsymbol{q}_{jh}^n, u_{jh}^n, \widehat{u}_{jh}^n)$ *be the solution of* (2.7)*, then we have the estimate*

$$\begin{aligned}\|\nabla w_{jh}^n\|_{\mathcal{T}_h} \leq & C\left(\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\|_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\|_{\mathcal{T}_h}\right) \\ & + C\Big\|h_K^{-\frac{1}{2}}(P_M w_{jh}^n - \widehat{w}_{jh}^n)\Big\|_{\partial\mathcal{T}_h} + C\|\nabla m_j^n\|_{\mathcal{T}_h}. \end{aligned} \tag{3.10}$$

*Proof.* By Lemma 3.6 and the triangle inequality, we get

$$\begin{aligned}\|\nabla w_{jh}^n\|_{\mathcal{T}_h} \leq & \|\nabla u_{jh}^n\|_{\mathcal{T}_h} + \|\nabla \Pi_{k+1} m_j^n\|_{\mathcal{T}_h} \\ \leq & C\bigg(\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\|_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\|_{\mathcal{T}_h} \\ & + \Big\|h_K^{-\frac{1}{2}}(P_M u_{jh}^n - \widehat{u}_{jh}^n)\Big\|_{\partial\mathcal{T}_h}\bigg) + C\|\nabla m_j^n\|_{\mathcal{T}_h} \\ \leq & C\left(\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\|_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\|_{\mathcal{T}_h}\right) \\ & + C\bigg(\Big\|h_K^{-\frac{1}{2}}(P_M w_{jh}^n - \widehat{w}_{jh}^n)\Big\|_{\partial\mathcal{T}_h} \\ & + \Big\|h_K^{-\frac{1}{2}}(P_M \Pi_{k+1} m_j^n - P_M m_j^n)\Big\|_{\partial\mathcal{T}_h}\bigg) + C\|\nabla m_j^n\|_{\mathcal{T}_h} \\ \leq & C\left(\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\|_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\|_{\mathcal{T}_h}\right) \\ & + C\Big\|h_K^{-\frac{1}{2}}(P_M w_{jh}^n - \widehat{w}_{jh}^n)\Big\|_{\partial\mathcal{T}_h} + C\|\nabla m_j^n\|_{\mathcal{T}_h}. \end{aligned}$$

□

**Lemma 3.10.** *Let* $(w_{jh}^n, \widehat{w}_{jh}^n)$ *be defined in* (3.9) *and* $(\boldsymbol{q}_{jh}^n, u_{jh}^n, \widehat{u}_{jh}^n)$ *be the solution of* (2.7)*, if the condition* (3.2) *holds, we have*

$$\begin{aligned}&\max_{2\leq n\leq N}\|w_{jh}^n\|_{\mathcal{T}_h}^2 + \Delta t\sum_{n=2}^N\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2 \\ \leq & C\Delta t\sum_{n=2}^N\left(\|f_j^n\|_{\mathcal{T}_h}^2 + \|\nabla m_j^n\|_{\mathcal{T}_h}^2\right) \\ & + C\left(\|w_{jh}^0\|_{\mathcal{T}_h}^2 + \|w_{jh}^1\|_{\mathcal{T}_h}^2 + \|\partial_t m_j\|_{L^2(0,T;L^2(\Omega))}^2 + \Delta t\|\sqrt{c^1}\boldsymbol{q}_{jh}^1\|_{\mathcal{T}_h}^2\right), \end{aligned}$$

*the constant* $C$ *in the above inequality depends on* $\boldsymbol{\beta}_j$ *and* $c_j$.

*Proof.* By the definitions of $w_{jh}^n$, $\widehat{w}_{jh}^n$ in (3.9), we can rewrite (2.7a) and (2.7b) as

$$\begin{aligned}&\left(\overline{c}^n\boldsymbol{q}_{jh}^n, \boldsymbol{r}_j\right)_{\mathcal{T}_h} - \left(w_{jh}^n, \nabla\cdot\boldsymbol{r}_j\right)_{\mathcal{T}_h} + \langle\widehat{w}_{jh}^n, \boldsymbol{r}_j\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ =&\Big(\left(\overline{c}^n - c_j^n\right)\left(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2}\right), \boldsymbol{r}_j\Big)_{\mathcal{T}_h} + \left(m_j^n, \nabla\cdot\boldsymbol{r}_j\right)_{\mathcal{T}_h} - \langle m_j^n, \boldsymbol{r}_j\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}, \end{aligned} \tag{3.11a}$$

$$
\begin{aligned}
&\left(\partial_t^+ w_{jh}^n, v_j\right)_{\mathcal{T}_h} + \left(\nabla\cdot \boldsymbol{q}_{jh}^n, v_j\right)_{\mathcal{T}_h} - \left\langle \boldsymbol{q}_{jh}^n\cdot \boldsymbol{n}, \widehat{v}_j\right\rangle_{\partial\mathcal{T}_h} - \left(\nabla\cdot\overline{\boldsymbol{\beta}}^n w_{jh}^n, v_j\right)_{\mathcal{T}_h}\\
&\quad - \left(\overline{\boldsymbol{\beta}}^n w_{jh}^n, \nabla v_j\right)_{\mathcal{T}_h} + \left\langle \overline{\boldsymbol{\beta}}^n\cdot\boldsymbol{n}, \widehat{w}_{jh}^n v_j\right\rangle_{\partial\mathcal{T}_h} + \left\langle h_K^{-\frac12}\left(P_M w_{jh}^n - \widehat{w}_{jh}^n\right), P_M v_j - \widehat{v}_j\right\rangle_{\partial\mathcal{T}_h}\\
&= \left(f_j^n, v_j\right)_{\mathcal{T}_h} - \left(\partial_t^+\left(\Pi_{k+1} w_j^n\right), v_j\right)_{\mathcal{T}_h} - \left(\left[\nabla\cdot\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\right]\left(2w_{jh}^{n-1} - w_{jh}^{n-2}\right), v_j\right)_{\mathcal{T}_h}\\
&\quad - \left(\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2w_{jh}^{n-1} - w_{jh}^{n-2}\right), \nabla v_j\right)_{\mathcal{T}_h} + \left\langle \left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\cdot\boldsymbol{n}, \left(2\widehat{w}_{jh}^{n-1} - \widehat{w}_{jh}^{n-2}\right)v_j\right\rangle_{\partial\mathcal{T}_h}\\
&\quad - \left(\left[\nabla\cdot\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\right]\left(2\Pi_{k+1} m_j^{n-1} - \Pi_{k+1} m_j^{n-2}\right), v_j\right)_{\mathcal{T}_h}\\
&\quad - \left(\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2\Pi_{k+1} m_j^{n-1} - \Pi_{k+1} m_j^{n-2}\right), \nabla v_j\right)_{\mathcal{T}_h}\\
&\quad + \left\langle \left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\cdot\boldsymbol{n}, \left(2P_M m_j^{n-1} - P_M m_j^{n-2}\right)v_j\right\rangle_{\partial\mathcal{T}_h}\\
&\quad + \left(\nabla\cdot\overline{\boldsymbol{\beta}}^n \Pi_{k+1} m_j^n, v_j\right)_{\mathcal{T}_h} + \left(\overline{\boldsymbol{\beta}}^n \Pi_{k+1} m_j^n, \nabla v_j\right)_{\mathcal{T}_h}\\
&\quad - \left\langle \overline{\boldsymbol{\beta}}^n\cdot\boldsymbol{n}, P_M m_j^n v_j\right\rangle_{\partial\mathcal{T}_h} - \left\langle h_K^{-1} P_M\left(\Pi_{k+1} m_j^n - m_j^n\right), P_M v_j - \widehat{v}_j\right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
\tag{3.11b}
$$

Now we take $(\boldsymbol{r}_j, v_j, \widehat{v}_j) = (\boldsymbol{q}_{jh}^n, w_{jh}^n, \widehat{w}_{jh}^n)$ in (3.11), add them together, use Lemma 3.2 and stability (3.7) with $(w,\mu,\boldsymbol{\gamma}) = (w_{jh}^n, \widehat{w}_{jh}^n, \overline{\boldsymbol{\beta}}^n)$ to get

$$
\begin{aligned}
&\frac{1}{4\Delta t}\left(\|w_{jh}^n\|_{\mathcal{T}_h}^2 + \|2w_{jh}^n - w_{jh}^{n-1}\|_{\mathcal{T}_h}^2 - \|w_{jh}^{n-1}\|_{\mathcal{T}_h}^2 - \|2w_{jh}^{n-1} - w_{jh}^{n-2}\|_{\mathcal{T}_h}^2\right)\\
&\quad + \frac{1}{4\Delta t}\|w_{jh}^n - 2w_{jh}^{n-1} + w_{jh}^{n-2}\|_{\mathcal{T}_h}^2 + \|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2 + \frac12\left\|h_K^{-\frac12}\left(P_M w_{jh}^n - \widehat{w}_{jh}^n\right)\right\|_{\partial\mathcal{T}_h}^2\\
&\le \left(\left(\overline{c}^n - c_j^n\right)\left(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2}\right), \boldsymbol{q}_{jh}^n\right)_{\mathcal{T}_h} + \left(m_j^n, \nabla\cdot\boldsymbol{q}_{jh}^n\right)_{\mathcal{T}_h} - \left\langle m_j^n, \boldsymbol{q}_{jh}^n\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h}\\
&\quad + \left(f_j^n, w_{jh}^n\right)_{\mathcal{T}_h} - \left(\partial_t^+\Pi_{k+1} m_j^n, w_{jh}^n\right)_{\mathcal{T}_h} + Ch\|\nabla w_{jh}^n\|_{\mathcal{T}_h}^2\\
&\quad - \left(\nabla\cdot\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2w_{jh}^{n-1} - w_{jh}^{n-2}\right), w_{jh}^n\right)_{\mathcal{T}_h}\\
&\quad - \left(\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2w_{jh}^{n-1} - w_{jh}^{n-2}\right), \nabla w_{jh}^n\right)_{\mathcal{T}_h}\\
&\quad + \left\langle \left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\cdot\boldsymbol{n}, \left(2\widehat{w}_{jh}^{n-1} - \widehat{w}_{jh}^{n-2}\right)w_{jh}^n\right\rangle_{\partial\mathcal{T}_h}\\
&\quad - \left(\nabla\cdot\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2\Pi_{k+1} m_j^{n-1} - \Pi_{k+1} m_j^{n-2}\right), w_{jh}^n\right)_{\mathcal{T}_h}\\
&\quad - \left(\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2\Pi_{k+1} m_j^{n-1} - \Pi_{k+1} m_j^{n-2}\right), \nabla w_{jh}^n\right)_{\mathcal{T}_h}\\
&\quad + \left\langle \left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\cdot\boldsymbol{n}, \left(2m_j^{n-1} - P_M m_j^{n-2}\right)w_{jh}^n\right\rangle_{\partial\mathcal{T}_h}\\
&\quad + \left(\nabla\cdot\overline{\boldsymbol{\beta}}^n \Pi_{k+1} m_j^n, w_{jh}^n\right)_{\mathcal{T}_h} + \left(\overline{\boldsymbol{\beta}}^n \Pi_{k+1} m_j^n, \nabla w_{jh}^n\right)_{\mathcal{T}_h}\\
&\quad - \left\langle \overline{\boldsymbol{\beta}}^n\cdot\boldsymbol{n}, P_M m_j^n w_{jh}^n\right\rangle_{\partial\mathcal{T}_h} - \left\langle h_K^{-\frac12} P_M\left(P_M m_j^n - m_j^n\right), P_M w_{jh}^n - \widehat{w}_{jh}^n\right\rangle_{\partial\mathcal{T}_h} =: \sum_{i=1}^{16} R_i.
\end{aligned}
$$

Next, we estimate $\{R_i\}_{i=1}^{16}$ term by term. By (3.2), there exists a constant $\kappa > 0$, such that

$$\left|\overline{c}^n - c_j^n\right| \leq \frac{\kappa}{3(\kappa+1)} \min\left\{\overline{c}^n, \overline{c}^{n-1}, \overline{c}^{n-2}\right\}. \tag{3.12}$$

Using the above condition (3.12) and the Young's inequality to have

$$\begin{aligned} R_1 &\leq \frac{\kappa}{3(\kappa+1)} \left(2\left\|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\right\|_{\mathcal{T}_h} + \left\|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\right\|_{\mathcal{T}_h}\right) \left\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\right\|_{\mathcal{T}_h} \\ &\leq \frac{\kappa}{3(\kappa+1)} \left(\frac{3}{2}\left\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\right\|_{\mathcal{T}_h}^2 + \left\|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\right\|_{\mathcal{T}_h}^2 + \frac{1}{2}\left\|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\right\|_{\mathcal{T}_h}^2\right). \end{aligned}$$

For the term $R_2 + R_3$, we use integration by parts to obtain

$$R_2 + R_3 = -\left(\nabla m_j^n, \boldsymbol{q}_{jh}^n\right) \leq \frac{1}{4(\kappa+1)}\left\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\right\|_{\mathcal{T}_h}^2 + C\left\|\nabla m_j^n\right\|_{\mathcal{T}_h}^2.$$

For the term $R_4$, we use the Cauchy-Schwarz inequality to get

$$R_4 \leq 2\left(\left\|w_{jh}^n\right\|_{\mathcal{T}_h}^2 + \left\|f_j^n\right\|_{\mathcal{T}_h}^2\right).$$

For the term $R_5$, by the definition of $\Pi_{k+1}$ in (2.1a) and we use the Cauchy-Schwarz inequality and the estimate (3.6a) to get

$$\begin{aligned} R_5 &= -\left(\partial_t^+ \Pi_{k+1} m_j^n, w_{jh}^n\right)_{\mathcal{T}_h} \\ &= -\left(\partial_t^+ m_j^n, w_{jh}^n\right)_{\mathcal{T}_h} \\ &\leq C\left(\left\|\partial_t^+ m_j^n\right\|_{\mathcal{T}_h}^2 + \left\|w_{jh}^n\right\|_{\mathcal{T}_h}^2\right) \\ &\leq C\Delta t^{-1}\left\|\partial_t m_j^n\right\|_{L^2(t_{n-2},t_n;L^2(\Omega))}^2 + C\left\|w_{jh}^n\right\|_{\mathcal{T}_h}^2. \end{aligned}$$

For the term $R_6$, by the estimate (3.10) and let $h$ sufficient small, one has

$$\begin{aligned} R_6 \leq & Ch\left(\left\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\right\|_{\mathcal{T}_h}^2 + \left\|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\right\|_{\mathcal{T}_h}^2 + \left\|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\right\|_{\mathcal{T}_h}^2\right) \\ & + Ch\left\|h_K^{-\frac{1}{2}}\left(P_M w_{jh}^n - \widehat{w}_{jh}^n\right)\right\|_{\partial\mathcal{T}_h}^2 + Ch\left\|\nabla m_j^n\right\|_{\mathcal{T}_h}^2 \\ \leq & \frac{1}{24(\kappa+1)}\left(\left\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\right\|_{\mathcal{T}_h}^2 + \left\|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\right\|_{\mathcal{T}_h}^2 + \left\|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\right\|_{\mathcal{T}_h}^2\right) \\ & + \frac{1}{16}\left\|h_K^{-\frac{1}{2}}\left(P_M w_{jh}^n - \widehat{w}_{jh}^n\right)\right\|_{\partial\mathcal{T}_h}^2 + C\left\|\nabla m_j^n\right\|_{\mathcal{T}_h}^2. \end{aligned}$$

For the term $R_7 + R_8 + R_9$, we let $(\boldsymbol{\gamma}, v, w, \widehat{v}, \widehat{w}) = (\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n, 2w_{jh}^{n-1} - w_{jh}^{n-2}, 2\widehat{w}_{jh}^{n-1} -$

$\widehat{w}_{jh}^{n-2}, w_{jh}^n, \widehat{w}_{jh}^n)$ in (3.8) to get

$$
\begin{aligned}
&R_7 + R_8 + R_9 \\
\le & C\left( \|w_{jh}^n\|_{\mathcal{T}_h}^2 + \|w_{jh}^{n-1}\|_{\mathcal{T}_h}^2 + \|w_{jh}^{n-2}\|_{\mathcal{T}_h}^2 + \left\| h_K^{\frac{1}{2}} (P_M w_{jh}^n - \widehat{w}_{jh}^n) \right\|_{\partial\mathcal{T}_h}^2 \right) \\
& + \left( \nabla \cdot \left[ \mathbf{\Pi}_0 (\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)(2w_{jh}^{n-1} - w_{jh}^{n-2}) \right], w_{jh}^n \right)_{\mathcal{T}_h} \\
& - \left\langle \left[ \mathbf{\Pi}_0^o (\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)(2w_{jh}^{n-1} - w_{jh}^{n-2}) \right] \cdot \boldsymbol{n}, \widehat{w}_{jh}^n \right\rangle_{\partial\mathcal{T}_h} \\
& + \left\langle (\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n) \cdot \boldsymbol{n} (2w_{jh}^{n-1} - w_{jh}^{n-2} - 2\widehat{w}_{jh}^{n-1} + \widehat{w}_{jh}^{n-2}), w_{jh}^n - \widehat{w}_{jh}^n \right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

Using (3.11a) with $\boldsymbol{r}_j = \mathbf{\Pi}_0(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)(2w_{jh}^{n-1} - w_{jh}^{n-2}) \in \boldsymbol{V}_h$, we get

$$
\begin{aligned}
&R_7 + R_8 + R_9 \\
\le & C\left( \|w_{jh}^n\|_{\mathcal{T}_h}^2 + \|w_{jh}^{n-1}\|_{\mathcal{T}_h}^2 + \|w_{jh}^{n-2}\|_{\mathcal{T}_h}^2 + \left\| h_K^{\frac{1}{2}} (P_M w_{jh}^n - \widehat{w}_{jh}^n) \right\|_{\partial\mathcal{T}_h}^2 \right) \\
& + \left( \overline{c}^n \boldsymbol{q}_{jh}^n, \mathbf{\Pi}_0 (\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)(2w_{jh}^{n-1} - w_{jh}^{n-2}) \right)_{\mathcal{T}_h} \\
& - \left( (\overline{c}^n - c_j^n)(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2}), \mathbf{\Pi}_0 (\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)(2w_{jh}^{n-1} - w_{jh}^{n-2}) \right)_{\mathcal{T}_h} \\
& + \left( \nabla m_j^n, \mathbf{\Pi}_0 (\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)(2w_{jh}^{n-1} - w_{jh}^{n-2}) \right)_{\mathcal{T}_h} \\
& + \left\langle (\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n) \cdot \boldsymbol{n} (2w_{jh}^{n-1} - w_{jh}^{n-2} - 2\widehat{w}_{jh}^{n-1} + \widehat{w}_{jh}^{n-2}), w_{jh}^n - \widehat{w}_{jh}^n \right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

For $h$ small enough, Cauchy-Schwarz inequality and Young's inequality give

$$
\begin{aligned}
&R_7 + R_8 + R_9 \\
\le & \frac{1}{24(\kappa+1)} \left( \|\sqrt{\overline{c}^n} \boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2 + \|\sqrt{\overline{c}^{n-1}} \boldsymbol{q}_{jh}^{n-1}\|_{\mathcal{T}_h}^2 + \|\sqrt{\overline{c}^{n-2}} \boldsymbol{q}_{jh}^{n-2}\|_{\mathcal{T}_h}^2 \right) \\
& + C\left( \|w_{jh}^n\|_{\mathcal{T}_h}^2 + \|w_{jh}^{n-1}\|_{\mathcal{T}_h}^2 + \|w_{jh}^{n-2}\|_{\mathcal{T}_h}^2 + \|\nabla m_j^n\|_{\mathcal{T}_h}^2 \right) \\
& + \frac{1}{16} \left\| h_K^{-\frac{1}{2}} (P_M w_{jh}^n - \widehat{w}_{jh}^n) \right\|_{\partial\mathcal{T}_h}^2.
\end{aligned}
$$

Using integration by parts and the estimate (3.10), we have

$$
\begin{aligned}
&R_{10} + R_{11} + R_{12} \\
\le & C\left( \|w_{jh}^n\|_{\mathcal{T}_h} + \|\nabla w_{jh}^n\|_{\mathcal{T}_h} \right) \left( \|\nabla m_j^{n-1}\|_{\mathcal{T}_h} + \|\nabla m_j^{n-2}\|_{\mathcal{T}_h} \right) \\
\le & \alpha \left( \|\nabla w_{jh}^n\|_{\mathcal{T}_h}^2 + \left\| h_K^{-\frac{1}{2}} (P_M w_{jh}^n - \widehat{w}_{jh}^n) \right\|_{\partial\mathcal{T}_h}^2 \right) \\
& + C_\alpha \left( \|\nabla m_j^{n-1}\|_{\mathcal{T}_h}^2 + \|\nabla m_j^{n-2}\|_{\mathcal{T}_h}^2 \right)
\end{aligned}
$$

$$
\begin{aligned}
\le & C\alpha\left(\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2+\|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\|_{\mathcal{T}_h}^2+\|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\|_{\mathcal{T}_h}^2\right)\\
&+C\alpha\left(\left\|h_K^{-\frac{1}{2}}\left(P_M w_{jh}^n-\widehat{w}_{jh}^n\right)\right\|_{\partial\mathcal{T}_h}^2+\|\nabla m_j^n\|_{\mathcal{T}_h}^2\right)\\
&+C_\alpha\left(\|\nabla m_j^{n-1}\|_{\mathcal{T}_h}^2+\|\nabla m_j^{n-2}\|_{\mathcal{T}_h}^2\right).
\end{aligned}
$$

Choosing $\alpha$ small enough, we get

$$
\begin{aligned}
& R_{10}+R_{11}+R_{12}\\
\le & \frac{1}{24(\kappa+1)}\left(\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2+\|\sqrt{\overline{c}^{n-1}}\boldsymbol{q}_{jh}^{n-1}\|_{\mathcal{T}_h}^2+\|\sqrt{\overline{c}^{n-2}}\boldsymbol{q}_{jh}^{n-2}\|_{\mathcal{T}_h}^2\right)\\
&+\frac{1}{16}\left\|h_K^{-\frac{1}{2}}\left(P_M w_{jh}^n-\widehat{w}_{jh}^n\right)\right\|_{\partial\mathcal{T}_h}^2\\
&+C\left(\|\nabla m_j^n\|_{\mathcal{T}_h}^2+\|\nabla m_j^{n-1}\|_{\mathcal{T}_h}^2+\|\nabla m_j^{n-2}\|_{\mathcal{T}_h}^2\right).
\end{aligned}
$$

We use integration by parts to get

$$
\begin{aligned}
& R_{13}+R_{14}+R_{15}\\
=&-\left(\overline{\beta}^n\cdot\nabla\Pi_{k+1}m_j^n, w_{jh}^n\right)_{\mathcal{T}_h}\le\|\overline{\beta}^n\|_{0,\infty}\|\nabla\Pi_{k+1}m_j^n\|_{\mathcal{T}_h}\|w_{jh}^n\|_{\mathcal{T}_h}\\
\le & C\|\nabla m_j^n\|_{\mathcal{T}_h}\|w_{jh}^n\|_{\mathcal{T}_h}\le C\left(\|\nabla m_j^n\|_{\mathcal{T}_h}^2+\|w_{jh}^n\|_{\mathcal{T}_h}^2\right),
\end{aligned}
$$

where $C$ depends on $\overline{\beta}^n$. We hide the dependence on $\boldsymbol{\beta}_j$ since we assume that $\boldsymbol{\beta}_j=\mathcal{O}(1)$ in (3.1). Therefore, by all the estimate above one gets

$$
\begin{aligned}
&\frac{1}{4\Delta t}\left(\|w_{jh}^n\|_{\mathcal{T}_h}^2+\|2w_{jh}^n-w_{jh}^{n-1}\|_{\mathcal{T}_h}^2-\|w_{jh}^{n-1}\|_{\mathcal{T}_h}^2-\|2w_{jh}^{n-1}-w_{jh}^{n-2}\|_{\mathcal{T}_h}^2\right)\\
&+\frac{1}{4\Delta t}\left(\|w_{jh}^n-2w_{jh}^{n-1}+w_{jh}^{n-2}\|_{\mathcal{T}_h}^2\right)+\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2+\frac{1}{2}\left\|h_K^{-\frac{1}{2}}\left(P_M w_{jh}^n-\widehat{w}_{jh}^n\right)\right\|_{\partial\mathcal{T}_h}^2\\
\le & C\left(\|w_{jh}^n\|_{\mathcal{T}_h}^2+\|w_{jh}^{n-1}\|_{\mathcal{T}_h}^2+\|w_{jh}^{n-2}\|_{\mathcal{T}_h}^2+\Delta t^{-1}\|\partial_t m_j^n\|_{L^2(t_{n-2},t_n;L^2(\Omega))}^2\right)\\
&+C\left(\|f_j^n\|_{\mathcal{T}_h}^2+\|\nabla m_j^n\|_{\mathcal{T}_h}^2+\|\nabla m_j^{n-1}\|_{\mathcal{T}_h}^2+\|\nabla m_j^{n-2}\|_{\mathcal{T}_h}^2\right)\\
&+\frac{2\kappa+1}{6(\kappa+1)}\left(\|\sqrt{c^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2+\|\sqrt{c^{n-1}}\boldsymbol{q}_{jh}^{n-1}\|_{\mathcal{T}_h}^2+\|\sqrt{c^{n-2}}\boldsymbol{q}_{jh}^{n-2}\|_{\mathcal{T}_h}^2\right)\\
&+\frac{1}{4}\left\|h_K^{-\frac{1}{2}}\left(P_M w_{jh}^n-\widehat{w}_{jh}^n\right)\right\|_{\partial\mathcal{T}_h}^2+\frac{1}{4(\kappa+1)}\|\sqrt{c^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2.
\end{aligned}
$$

We add last inequality from $n=2$ to $n=N$, rearrange it, and multiply $4\Delta t$ to get

$$
\begin{aligned}
&\max_{2\le n\le N}\|w_{jh}^n\|_{\mathcal{T}_h}^2+\Delta t\sum_{n=2}^N\|\sqrt{\overline{c}^n}\boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2\le C\Delta t\sum_{n=2}^N\left(\|f_j^n\|_{\mathcal{T}_h}^2+\|\nabla m_j^n\|_{\mathcal{T}_h}^2\right)\\
&\quad+C\left(\|w_{jh}^0\|_{\mathcal{T}_h}^2+\|w_{jh}^1\|_{\mathcal{T}_h}^2+\|\partial_t m_j\|_{L^2(0,T;L^2(\Omega))}^2+\Delta t\|\sqrt{\overline{c}^1}\boldsymbol{q}_{jh}^1\|_{\mathcal{T}_h}^2\right),
\end{aligned}
$$

then the result followed by Gronwall's inequality the $C$ in the above inequality depends on $\boldsymbol{\beta}_j$ and $c_j$. $\square$

## 4. Error analysis

The strategy of the error analysis for the Ensemble HDG method is based on [1]. Throughout, we assume the data, the solutions of (1.1) are smooth enough and the domain $\Omega$ is convex.

### 4.1. HDG elliptic projection

For any $t \in [0,T]$ and $j = 1,\dots,J$, let $(\overline{\boldsymbol{q}}_{jh}, \overline{u}_{jh}, \widehat{\overline{u}}_{jh}) \in \boldsymbol{V}_h \times W_h \times M_h(g_j)$ be the solution of the following steady state problem

$$\begin{aligned}
&\big(c_j\overline{\boldsymbol{q}}_{jh}, \boldsymbol{r}_j\big)_{\mathcal{T}_h} - \big(\overline{u}_{jh}, \nabla\cdot\boldsymbol{r}_j\big)_{\mathcal{T}_h} + \left\langle \widehat{\overline{u}}_{jh}, \boldsymbol{r}_j\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} = 0, && \text{(4.1a)}\\
&\big(\nabla\cdot\overline{\boldsymbol{q}}_{jh}, w_j\big)_{\mathcal{T}_h} - \langle\overline{\boldsymbol{q}}_{jh}\cdot\boldsymbol{n}, \mu_j\rangle_{\partial\mathcal{T}_h} - \big(\nabla\cdot\boldsymbol{\beta}_j\overline{u}_{jh}, w_j\big)_{\mathcal{T}_h}\\
&- \big(\boldsymbol{\beta}_j\overline{u}_{jh}, \nabla w_j\big)_{\mathcal{T}_h} + \left\langle \boldsymbol{\beta}_j\cdot\boldsymbol{n}, \widehat{\overline{u}}_{jh} w_j\right\rangle_{\partial\mathcal{T}_h}\\
&+ \left\langle h_K^{-1}\big(P_M\overline{u}_{jh} - \widehat{\overline{u}}_{jh}\big), P_M w_j - \mu_j\right\rangle_{\partial\mathcal{T}_h}\\
=&\big(f_j - \partial_t u_j, w_j\big)_{\mathcal{T}_h} && \text{(4.1b)}
\end{aligned}$$

for all $(\boldsymbol{r}_j, w_j, \mu_j) \in \boldsymbol{V}_h \times W_h \times M_h(0)$.

The proofs of the following estimations are presented in Appendix A.

**Theorem 4.1.** *For any $t \in [0,T]$ and $j = 1,\dots,J$, we have*

$$\begin{aligned}
&\big\|\boldsymbol{q}_j - \overline{\boldsymbol{q}}_{jh}\big\|_{\mathcal{T}_h} \le Ch^{k+1}, && \big\|\partial_t\boldsymbol{q}_j - \partial_t\overline{\boldsymbol{q}}_{jh}\big\|_{\mathcal{T}_h} \le Ch^{k+1},\\
&\big\|\partial_{tt}\boldsymbol{q}_j - \partial_{tt}\overline{\boldsymbol{q}}_{jh}\big\|_{\mathcal{T}_h} \le Ch^{k+1}, && \big\|u_j - \overline{u}_{jh}\big\|_{\mathcal{T}_h} \le Ch^{k+2},\\
&\big\|\partial_t u_j - \partial_t\overline{u}_{jh}\big\|_{\mathcal{T}_h} \le Ch^{k+2}, && \big\|\partial_{tt} u_j - \partial_{tt}\overline{u}_{jh}\big\|_{\mathcal{T}_h} \le Ch^{k+2},\\
&\Big\|h_K^{\frac{1}{2}}\big(\overline{u}_{jh} - \widehat{\overline{u}}_{jh}\big)\Big\|_{\partial\mathcal{T}_h} \le Ch^{k+1}, && \Big\|h_K^{\frac{1}{2}}\big(\partial_{tt}\overline{u}_{jh} - \partial_{tt}\widehat{\overline{u}}_{jh}\big)\Big\|_{\partial\mathcal{T}_h} \le Ch^{k+1}.
\end{aligned}$$

### 4.2. Main result

We can now state our main result for the ensemble HDG method.

**Theorem 4.2.** *If the condition* (3.2) *holds and the domain is convex, then we have the following error estimate*

$$\max_{1\le n\le N}\big\|u_j^n - u_{jh}^n\big\|_{\mathcal{T}_h} \le C\left(\Delta t^2 + h^{k+2}\right), \tag{4.2}$$

$$\sqrt{\Delta t \sum_{n=1}^{N} \left\|\sqrt{\overline{c}^n}(\boldsymbol{q}_j^n - \boldsymbol{q}_{jh}^n)\right\|_{\mathcal{T}_h}^2} \le C\left(\Delta t^2 + h^{k+1}\right). \tag{4.3}$$

**Remark 4.1.** To the best of our knowledge, all previous works only contain a *suboptimal* $L^\infty(0,T;L^2(\Omega))$ convergent rate for the ensemble solutions $u_j$. Only one other very recent work [3] contains an *optimal* $L^\infty(0,T;L^2(\Omega))$ convergent rate for the ensemble solutions $u_j$, and a $L^2(0,T;L^2(\Omega))$ superconvergent rate if the coefficients of the PDEs are independent of time and degree polynomial $k \ge 1$; our main result: Theorem 4.2 is the *first* time to obtain the $L^\infty(0,T;L^2(\Omega))$ supconvergent rate for the ensemble solutions $u_j$ for all $k \ge 0$ and without assume that the coefficients of the PDEs are independent of time. It is also the *first* time to obtain the superconvergent rate for a single convection diffusion PDE when $k = 0$.

### 4.3. Proof of Theorem 4.2

The proof of (4.2) with $n = 1$ is quite standard in backward Euler discretization, thus we omit it, and we prove (4.2) holds for all $n \ge 2$.

#### 4.3.1. The equations of the projection of the errors

**Lemma 4.1.** *For* $e_{jh}^{\boldsymbol{q}^n} = \boldsymbol{q}_{jh}^n - \overline{\boldsymbol{q}}_{jh}^n$, $e_{jh}^{u^n} = u_{jh}^n - \overline{u}_{jh}^n$, $e_{jh}^{\widehat{u}^n} = \widehat{u}_{jh}^n - \overline{\widehat{u}}_{jh}^n$, *for all* $j = 1, \dots, J$, *we have the following error equations*

$$\begin{aligned}
&\left(\overline{c}^n e_{jh}^{\boldsymbol{q}^n}, \boldsymbol{r}_j\right)_{\mathcal{T}_h} - \left(e_{jh}^{u^n}, \nabla\cdot\boldsymbol{r}_j\right)_{\mathcal{T}_h} + \left\langle e_{jh}^{\widehat{u}^n}, \boldsymbol{r}_j\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} \\
=&\Big(\left(\overline{c}^n - c_j^n\right)\left(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2} - \overline{\boldsymbol{q}}_{jh}^n\right), \boldsymbol{r}_j\Big)_{\mathcal{T}_h},
\end{aligned} \tag{4.4a}$$

$$\begin{aligned}
&\left(\partial_t^+ e_{jh}^{u^n}, w_j\right)_{\mathcal{T}_h} + \left(\nabla\cdot e_{jh}^{\boldsymbol{q}^n}, w_j\right)_{\mathcal{T}_h} - \left\langle e_{jh}^{\boldsymbol{q}^n}\cdot\boldsymbol{n}, \mu_j\right\rangle_{\partial\mathcal{T}_h} \\
&- \left(\nabla\cdot\overline{\boldsymbol{\beta}}^n e_{jh}^{u^n}, w_j\right)_{\mathcal{T}_h} - \left(\overline{\boldsymbol{\beta}}^n e_{jh}^{u^n}, \nabla w_j\right)_{\mathcal{T}_h} + \left\langle\overline{\boldsymbol{\beta}}^n\cdot\boldsymbol{n}, e_{jh}^{\widehat{u}^n} w_j\right\rangle_{\partial\mathcal{T}_h} \\
&+ \left\langle h_K^{-1}\left(P_M e_{jh}^{u^n} - e_{jh}^{\widehat{u}^n}\right), P_M w_j - \mu_j\right\rangle_{\partial\mathcal{T}_h} \\
=&-\Big(\nabla\cdot\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n\right), w_j\Big)_{\mathcal{T}_h} \\
&-\Big(\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n\right), \nabla w_j\Big)_{\mathcal{T}_h} \\
&+\Big\langle\left(\overline{\boldsymbol{\beta}}_j^n - \boldsymbol{\beta}_j^n\right)\cdot\boldsymbol{n}, \left(2\widehat{u}_{jh}^{n-1} - \widehat{u}_{jh}^{n-2} - \overline{\widehat{u}}_{jh}^n\right)w_j\Big\rangle_{\partial\mathcal{T}_h} + \left(\partial_t u_j^n - \partial_t^+\overline{u}_{jh}^n, w_j\right)_{\mathcal{T}_h}
\end{aligned} \tag{4.4b}$$

*for all* $(\boldsymbol{r}_j, w_j, \mu_j) \in \boldsymbol{V}_h \times W_h \times M_h(0)$ *and* $n = 1, \dots, N$.

The proof of Lemma 4.1 follows by subtracting Eq. (4.1) from Lemma 2.1.

### 4.3.2. Energy argument

We take $\boldsymbol{r}_j = \nabla e^{u^n}_{jh}$ in (4.4a) and use integration by parts to get the following lemma.

**Lemma 4.2.** *We have*

$$\begin{aligned}
&\|\nabla e^{u^n}_{jh}\|_{\mathcal{T}_h} + \left\|h_K^{-\frac{1}{2}}(e^{u^n}_{jh} - e^{\widehat{u}^n}_{jh})\right\|_{\partial\mathcal{T}_h} \\
\le & C\left(\|\sqrt{\overline{c}^n} e^{\boldsymbol{q}^n}_{jh}\|_{\mathcal{T}_h} + \left\|h_K^{-\frac{1}{2}}(P_M e^{u^n}_{jh} - e^{\widehat{u}^n}_{jh})\right\|_{\partial\mathcal{T}_h}\right) \\
&+ C\left\|(\overline{c}^n - c_j^n)(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2} - \overline{\boldsymbol{q}}_{jh}^n)\right\|_{\mathcal{T}_h}.
\end{aligned} \tag{4.5}$$

**Lemma 4.3.** *If the condition* (3.2) *holds and the domain is convex, then we have the following error estimate*

$$\max_{2\le n\le N} \|e^{u^n}_{jh}\|_{\mathcal{T}_h} + \sqrt{\Delta t \sum_{n=2}^N \|\sqrt{\overline{c}^n} e^{\boldsymbol{q}^n}_{jh}\|^2_{\mathcal{T}_h}} \le C\left(\Delta t^2 + h^{k+2}\right).$$

*Proof.* We take $(\boldsymbol{r}_j, w_j, \mu_j) = (e^{\boldsymbol{q}^n}_{jh}, e^{u^n}_{jh}, e^{\widehat{u}^n}_{jh})$ in (4.4), use the polarization identity (3.2), stability (3.7) with $(\boldsymbol{\gamma}, w, \mu) = (\boldsymbol{\beta}_j^n, e^{u^n}_{jh}, e^{\widehat{u}^n}_{jh})$, and add them together to get

$$\begin{aligned}
&\frac{1}{4\Delta t}\left(\|e^{u^n}_{jh}\|^2_{\mathcal{T}_h} + \|2e^{u^n}_{jh} - e^{u^{n-1}}_{jh}\|^2_{\mathcal{T}_h} - \|e^{u^{n-1}}_{jh}\|^2_{\mathcal{T}_h} - \|2e^{u^{n-1}}_{jh} - e^{u^{n-2}}_{jh}\|^2_{\mathcal{T}_h}\right) \\
&+ \frac{1}{4\Delta t}\|e^{u^n}_{jh} - 2e^{u^{n-1}}_{jh} + e^{u^{n-2}}_{jh}\|^2_{\mathcal{T}_h} + \|\sqrt{\overline{c}^n} e^{\boldsymbol{q}^n}_{jh}\|^2_{\mathcal{T}_h} + \frac{1}{2}\left\|h_K^{-\frac{1}{2}}(P_M e^{u^n}_{jh} - e^{\widehat{u}^n}_{jh})\right\|^2_{\partial\mathcal{T}_h} \\
\le & \left((\overline{c}^n - c_j^n)(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2} - \overline{\boldsymbol{q}}_{jh}^n), e^{\boldsymbol{q}^n}_{jh}\right) + \left(\partial_t u_j^n - \partial_t^+ \overline{u}_{jh}^n, e^{u^n}_{jh}\right)_{\mathcal{T}_h} \\
&- \left(\nabla\cdot(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)(2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n), e^{u^n}_{jh}\right)_{\mathcal{T}_h} \\
&- \left((\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)(2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n), \nabla e^{u^n}_{jh}\right)_{\mathcal{T}_h} \\
&+ \left\langle(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n)\cdot\boldsymbol{n}, (2\widehat{u}_{jh}^{n-1} - \widehat{u}_{jh}^{n-2} - \widehat{\overline{u}}_{jh}^n)e^{u^n}_{jh}\right\rangle_{\partial\mathcal{T}_h} + Ch\|\nabla\varepsilon^{u^n}_{jh}\|^2_{\mathcal{T}_h} \\
=: & \sum_{i=1}^6 R_i.
\end{aligned} \tag{4.6}$$

Next, we estimate $\{R_i\}_{i=1}^6$ term by term. For the first term $R_1$, since

$$2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2} - \overline{\boldsymbol{q}}_{jh}^n = 2e^{\boldsymbol{q}^{n-1}}_{jh} - e^{\boldsymbol{q}^{n-2}}_{jh} - \Delta t^2 \partial_{tt}^+ \overline{\boldsymbol{q}}_{jh}^n \tag{4.7}$$

we use condition (3.12) to get

$$\begin{aligned}
R_1 &= \left((\overline{c}^n - c_j^n)(2e^{\boldsymbol{q}^{n-1}}_{jh} - e^{\boldsymbol{q}^{n-2}}_{jh} - \Delta t^2 \partial_{tt}^+ \overline{\boldsymbol{q}}_{jh}^n), e^{\boldsymbol{q}^n}_{jh}\right) \\
&\le \frac{\kappa}{3(\kappa+1)}\left(\frac{3}{2}\|\sqrt{\overline{c}^n} e^{\boldsymbol{q}^n}_{jh}\|^2_{\mathcal{T}_h} + \|\sqrt{\overline{c}^{n-1}} e^{\boldsymbol{q}^{n-1}}_{jh}\|^2_{\mathcal{T}_h} + \frac{1}{2}\|\sqrt{\overline{c}^{n-1}} e^{\boldsymbol{q}^{n-2}}_{jh}\|^2_{\mathcal{T}_h}\right) \\
&\quad + \frac{1}{8(\kappa+1)}\|\sqrt{\overline{c}^n} e^{\boldsymbol{q}^n}_{jh}\|^2_{\mathcal{T}_h} + C\Delta t^4 \|\partial_{tt}^+ \overline{\boldsymbol{q}}_{jh}^n\|^2_{\mathcal{T}_h}.
\end{aligned}$$

For the term $R_2$, we have

$$
\begin{aligned}
R_2 &= \left(\partial_t^+\left(u_j^n - \overline{u}_{jh}^n\right) + \partial_t u_j^n - \partial_t^+ u_j^n, e_{jh}^{u^n}\right)_{\mathcal{T}_h} \\
&\le C\left(\left\|\partial_t^+(u_j^n - \overline{u}_{jh}^n)\right\|_{\mathcal{T}_h}^2 + \left\|\partial_t u_j^n - \partial_t^+ u_j^n\right\|_{\mathcal{T}_h}^2 + \left\|e_{jh}^{u^n}\right\|_{\mathcal{T}_h}^2\right).
\end{aligned}
$$

For the term $R_3 + R_4 + R_5$, Eq. (3.8) and (4.4a) give

$$
\begin{aligned}
&R_3 + R_4 + R_5 \\
\le& C\left(\left\|2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n\right\|_{\mathcal{T}_h}^2 + \left\|e_{jh}^{u^n}\right\|_{\mathcal{T}_h}^2 + \left\|h_K^{\frac{1}{2}}\left(P_M e_{jh}^{u^n} - e_{jh}^{\widehat{u}^n}\right)\right\|_{\partial\mathcal{T}_h}^2\right) \\
&+ \left(\nabla\cdot\left[\boldsymbol{\Pi}_0\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n\right)\right], e_{jh}^{u^n}\right)_{\mathcal{T}_h} \\
&- \left\langle \boldsymbol{\Pi}_0\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n\right)\cdot\boldsymbol{n}, e_{jh}^{\widehat{u}^n}\right\rangle_{\partial\mathcal{T}_h} \\
&+ \left\langle \left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\cdot\boldsymbol{n}\left(2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n - 2\widehat{u}_{jh}^{n-1} + \widehat{u}_{jh}^{n-2} + \widehat{\overline{u}}_{jh}^n\right), e_{jh}^{u^n} - e_{jh}^{\widehat{u}^n}\right\rangle_{\partial\mathcal{T}_h} \\
=& C\left(\left\|2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n\right\|_{\mathcal{T}_h}^2 + \left\|e_{jh}^{u^n}\right\|_{\mathcal{T}_h}^2 + \left\|h_K^{\frac{1}{2}}(P_M e_{jh}^{u^n} - e_{jh}^{\widehat{u}^n})\right\|_{\partial\mathcal{T}_h}^2\right) \\
&+ \left(\overline{c}^n e_{jh}^{\boldsymbol{q}^n}, \boldsymbol{\Pi}_0\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n\right)\right)_{\mathcal{T}_h} \\
&- \left(\left(\overline{c}^n - c_j^n\right)\left(2\boldsymbol{q}_{jh}^{n-1} - \boldsymbol{q}_{jh}^{n-2} - \overline{\boldsymbol{q}}_{jh}^n\right), \boldsymbol{\Pi}_0\left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\left(2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n\right)\right)_{\mathcal{T}_h} \\
&+ \left\langle \left(\overline{\boldsymbol{\beta}}^n - \boldsymbol{\beta}_j^n\right)\cdot\boldsymbol{n}\left(2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n - 2\widehat{u}_{jh}^{n-1} + \widehat{u}_{jh}^{n-2} + \widehat{\overline{u}}_{jh}^n\right), e_{jh}^{u^n} - e_{jh}^{\widehat{u}^n}\right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

Similar to (4.7), we have

$$
\begin{aligned}
&2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n = 2e_{jh}^{u^{n-1}} - e_{jh}^{u^{n-2}} - \Delta t^2\partial_{tt}^+\overline{u}_{jh}^n, \\
&2u_{jh}^{n-1} - u_{jh}^{n-2} - \overline{u}_{jh}^n - 2\widehat{u}_{jh}^{n-1} + \widehat{u}_{jh}^{n-2} + \widehat{\overline{u}}_{jh}^n \\
=& 2\left(e_{jh}^{u^{n-1}} - e_{jh}^{\widehat{u}^{n-1}}\right) - \left(e_{jh}^{u^{n-2}} - e_{jh}^{\widehat{u}^{n-2}}\right) - \Delta t^2\partial_{tt}^+\overline{u}_{jh}^n + \Delta t^2\partial_{tt}^+\widehat{\overline{u}}_{jh}^n.
\end{aligned}
$$

Therefore, when $h$ is small enough, we have

$$
\begin{aligned}
&R_3 + R_4 + R_5 \\
\le& \frac{1}{24(\kappa+1)}\left(\left\|\sqrt{\overline{c}^n} e_{jh}^{\boldsymbol{q}^n}\right\|_{\mathcal{T}_h}^2 + \left\|\sqrt{\overline{c}^{n-1}} e_{jh}^{\boldsymbol{q}^{n-1}}\right\|_{\mathcal{T}_h}^2 + \left\|\sqrt{\overline{c}^{n-2}} e_{jh}^{\boldsymbol{q}^{n-2}}\right\|_{\mathcal{T}_h}^2\right) \\
&+ \frac{1}{16}\left\|h_K^{-\frac{1}{2}}\left(P_M e_{jh}^{u^n} - e_{jh}^{\widehat{u}^n}\right)\right\|_{\mathcal{T}_h}^2 + C\left(\left\|e_{jh}^{u^n}\right\|_{\mathcal{T}_h}^2 + \left\|e_{jh}^{u^{n-1}}\right\|_{\mathcal{T}_h}^2 + \left\|e_{jh}^{u^{n-2}}\right\|_{\mathcal{T}_h}^2\right) \\
&+ \frac{1}{16}\left\|h_K^{-\frac{1}{2}}\left(P_M e_{jh}^{u^{n-1}} - e_{jh}^{\widehat{u}^{n-1}}\right)\right\|_{\mathcal{T}_h}^2 + \frac{1}{16}\left\|h_K^{-\frac{1}{2}}\left(P_M e_{jh}^{u^{n-2}} - e_{jh}^{\widehat{u}^{n-2}}\right)\right\|_{\mathcal{T}_h}^2 \\
&+ C\Delta t^4\left\|\partial_{tt}^+\overline{\boldsymbol{q}}_{jh}^n\right\|_{\mathcal{T}_h}^2 + C\Delta t^4\left\|\partial_{tt}^+\overline{u}_{jh}^n\right\|_{\mathcal{T}_h}^2 + C\Delta t^4\left\|h_K^{\frac{1}{2}}\partial_{tt}^+\left(\overline{u}_{jh}^n - \widehat{\overline{u}}_{jh}^n\right)\right\|_{\partial\mathcal{T}_h}^2.
\end{aligned}
$$

By the Cauchy-Schwarz inequality and $h$ small enough, by (4.5) and (4.7), we get

$$\begin{aligned} R_6 \leq & \frac{1}{24(\kappa+1)} \left( \|\sqrt{c^n} e_{jh}^{\boldsymbol{q}^n}\|_{\mathcal{T}_h}^2 + \|\sqrt{c^{n-1}} e_{jh}^{\boldsymbol{q}^{n-1}}\| + \|\sqrt{c^{n-1}} e_{jh}^{\boldsymbol{q}^{n-2}}\|_{\mathcal{T}_h}^2 \right) \\ & + C\Delta t^4 \|\partial_{tt}^+ \overline{\boldsymbol{q}}_{jh}^n\|_{\mathcal{T}_h}^2. \end{aligned}$$

We add (4.6) from $n=2$ to $n=N$ and use the above inequalities to get

$$\begin{aligned} & \max_{2\leq n\leq N} \|e_{jh}^{u^n}\|_{\mathcal{T}_h}^2 + \Delta t \sum_{n=2}^N \|\sqrt{c^n} e_{jh}^{\boldsymbol{q}^n}\|_{\mathcal{T}_h}^2 \\ \leq & C\Delta t^5 \sum_{n=2}^N \left( \|\partial_{tt}^+ \overline{u}_{jh}^n\|_{\mathcal{T}_h}^2 + \|\partial_{tt}^+ \overline{\boldsymbol{q}}_{jh}^n\|_{\mathcal{T}_h}^2 + \left\| h_K^{\frac{1}{2}} \partial_{tt}^+ (\overline{u}_{jh}^n - \widehat{\overline{u}}_{jh}^n) \right\|_{\partial\mathcal{T}_h}^2 \right), \\ & + C\Delta t \sum_{n=2}^N \left( \|\partial_t^+ (u_j^n - \overline{u}_{jh}^n)\|_{\mathcal{T}_h}^2 + \|\partial_t u_j^n - \partial_t^+ u_j^n\|_{\mathcal{T}_h}^2 \right) \\ & + C\Delta t \sum_{n=2}^N \|e_{jh}^{u^n}\|_{\mathcal{T}_h}^2 + C \left( \|e_{jh}^{u^0}\|_{\mathcal{T}_h}^2 + \|e_{jh}^{u^1}\|_{\mathcal{T}_h}^2 + \Delta t \|e_{jh}^{\boldsymbol{q}^1}\|_{\mathcal{T}_h}^2 \right). \end{aligned} \tag{4.8}$$

Next, we bound the terms on the right side of (4.8) by Lemma 3.4.

$$\begin{aligned} & \Delta t^5 \sum_{n=2}^N \|\partial_{tt}^+ \overline{u}_{jh}^n\|_{\mathcal{T}_h}^2 \leq C\Delta t^4 \|\partial_{tt} \overline{u}_{jh}\|_{L^2(0,T;L^2(\Omega))}^2, \\ & \Delta t^5 \sum_{n=2}^N \|\partial_{tt}^+ \overline{\boldsymbol{q}}_{jh}^n\|_{\mathcal{T}_h}^2 \leq C\Delta t^4 \|\partial_{tt} \overline{\boldsymbol{q}}_{jh}\|_{L^2(0,T;L^2(\Omega))}^2, \\ & \Delta t \sum_{n=2}^N \|\partial_t^+ (u_j^n - \overline{u}_{jh}^n)\|_{\mathcal{T}_h}^2 \leq C \|\partial_t (u_j - \overline{u}_{jh})\|_{L^2(0,T;L^2(\Omega))}^2, \\ & \Delta t \sum_{n=2}^N \|\partial_t u_j^n - \partial_t^+ u_j^n\|_{\mathcal{T}_h}^2 \leq C\Delta t^4 \|\partial_{ttt} u_j\|_{L^2(0,T;L^2(\Omega))}^2, \\ & \Delta t^5 \sum_{n=2}^N \left\| h_K^{\frac{1}{2}} \partial_{tt}^+ (\overline{u}_{jh}^n - \widehat{\overline{u}}_{jh}^n) \right\|_{\partial\mathcal{T}_h}^2 \leq C\Delta t^4 \left\| h_K^{\frac{1}{2}} \partial_{tt} (\overline{u}_{jh} - \widehat{\overline{u}}_{jh}) \right\|_{L^2(0,T;L^2(\partial\mathcal{T}_h))}^2. \end{aligned}$$

Gronwall's inequality, the estimates above, Theorem 4.1 applied to (4.8) and (4.2) give the desired result. □

As a consequence, a simple application of the triangle inequality for (4.3) and Theorem 4.1 give the proof of Theorem 4.2.

## 5. Numerical experiments

In this section, we present some numerical tests of the Ensemble HDG method for parameterized convection diffusion PDEs. A group of simulations are considered containing $J = 3$ members. Let $Eu_j$ be the error between the exact solution $u_j$ at the final time $T = 1$ and the Ensemble HDG solution $u_{jh}^N$, i.e., $Eu_j = \|u_j^N - u_{jh}^N\|_{\mathcal{T}_h}$. Let

$$E\boldsymbol{q}_j = \sqrt{\Delta t \sum_{n=1}^{N} \|\boldsymbol{q}_j^n - \boldsymbol{q}_{jh}^n\|_{\mathcal{T}_h}^2}.$$

We test the convergence rate of the Ensemble HDG method on a square domain $\Omega = [0,1] \times [0,1]$. In the first test, the data is chosen as

$$\begin{aligned} &c_1 = 1.1(1+t), && c_2 = 1.2(1+t), && c_3 = 1.3(1+t), \\ &\boldsymbol{\beta}_1 = [1,1], && \boldsymbol{\beta}_2 = [2,2], && \boldsymbol{\beta}_3 = [3,3], \\ &u_1 = e^{-t}\sin(x), && u_2 = \cos(t)\cos(x), && u_3 = e^{x-t}, \end{aligned}$$

and the initial conditions, boundary conditions, and source terms are chosen to match the exact solution of Eq. (1.1). It is easy to see that the coefficients $c_j$ satisfy the condition (3.2).

In order to confirm our theoretical results, we take $\Delta t = h$ when $k = 0$ and $\Delta t = h^{\frac{3}{2}}$ when $k = 1$. The approximation errors of the Ensemble HDG method are listed in Table 1 and the observed convergence rates match our theory.

## 6. Conclusion

In this work, we devised a new superconvergent Ensemble HDG method for parameterized convection diffusion PDEs. This new Ensemble HDG method shares one common coefficient matrix and multiple RHS vectors, which is more efficient than performing separate simulations. We obtained a $L^\infty(0,T;L^2(\Omega))$ superconvergent rate for the solutions for all polynomial degree $k \geq 0$. As far as we are aware, this is the first time in the literature, it is even the first time for a single convection diffusion PDE to obtain the superconvergence rate when $k = 0$.

## Acknowledgements

G. Chen was supported by National Natural Science Foundation of China (NSFC) (11801063), by China Postdoctoral Science Foundation(2018M633339, 2019T120808) and by the Fundamental Research Funds for the Central Universities (YJ202030). Y. Zhang was supported by US National Science Foundation (NSF) (DMS-1619904).

## Appendix A

In this section, we only give a proof of $\|\boldsymbol{q}_j - \overline{\boldsymbol{q}}_{jh}\|_{\mathcal{T}_h} \leq Ch^{k+1}$, $\|u_j - \overline{u}_{jh}\|_{\mathcal{T}_h} \leq Ch^{k+2}$ and $\|h_K^{\frac{1}{2}}(\overline{u}_{jh} - \widehat{\overline{u}}_{jh})\|_{\partial\mathcal{T}_h} \leq Ch^{k+1}$ since the rest are similar. To prove the rest,

Table 1: History of convergence.

| Degree | $\frac{h}{\sqrt{2}}$ | $E\boldsymbol{q}_1$ Error | Rate | $Eu_1$ Error | Rate |
|---|---|---|---|---|---|
| $k=0$ | $2^{-1}$ | 1.0360E+00 | | 4.1730E-01 | |
| | $2^{-2}$ | 5.3867E-01 | 0.94 | 6.0658E-02 | 2.78 |
| | $2^{-3}$ | 2.7518E-01 | 0.97 | 1.8030E-02 | 1.75 |
| | $2^{-4}$ | 1.3795E-01 | 1.00 | 4.8156E-03 | 1.90 |
| | $2^{-5}$ | 6.9056E-02 | 1.00 | 1.2012E-03 | 2.00 |
| $k=1$ | $2^{-1}$ | 3.5178E-01 | | 1.5269E-01 | |
| | $2^{-2}$ | 7.8269E-02 | 2.17 | 9.6593E-03 | 3.98 |
| | $2^{-3}$ | 1.9677E-02 | 1.99 | 1.2344E-03 | 2.97 |
| | $2^{-4}$ | 4.9408E-03 | 1.99 | 1.5697E-04 | 2.98 |
| | $2^{-5}$ | 1.2367E-03 | 2.00 | 1.9823E-05 | 2.99 |
| Degree | $\frac{h}{\sqrt{2}}$ | $E\boldsymbol{q}_2$ Error | Rate | $Eu_2$ Error | Rate |
| $k=0$ | $2^{-1}$ | 3.0237E-01 | | 1.8409E-01 | |
| | $2^{-2}$ | 1.7819E-01 | 0.76 | 4.3019E-02 | 2.10 |
| | $2^{-3}$ | 9.7785E-02 | 0.87 | 1.2796E-02 | 1.75 |
| | $2^{-4}$ | 5.1027E-02 | 0.94 | 3.5441E-03 | 1.85 |
| | $2^{-5}$ | 2.5807E-02 | 0.98 | 8.9715E-04 | 1.98 |
| $k=1$ | $2^{-1}$ | 1.2216E-01 | | 5.9224E-02 | |
| | $2^{-2}$ | 2.3969E-02 | 2.35 | 3.9697E-03 | 3.90 |
| | $2^{-3}$ | 5.4027E-03 | 2.15 | 3.8968E-04 | 3.35 |
| | $2^{-4}$ | 1.3536E-03 | 2.00 | 4.8216E-05 | 3.01 |
| | $2^{-5}$ | 3.3937E-04 | 2.00 | 6.0519E-06 | 2.99 |
| Degree | $\frac{h}{\sqrt{2}}$ | $E\boldsymbol{q}_3$ Error | Rate | $Eu_3$ Error | Rate |
| $k=0$ | $2^{-1}$ | 2.2660E-01 | | 8.5994E-02 | |
| | $2^{-2}$ | 1.2689E-01 | 0.84 | 2.4143E-02 | 1.83 |
| | $2^{-3}$ | 6.5402E-02 | 0.96 | 6.2378E-03 | 1.95 |
| | $2^{-4}$ | 3.2963E-02 | 0.99 | 1.5734E-03 | 1.99 |
| | $2^{-5}$ | 1.6515E-02 | 1.00 | 3.9432E-04 | 2.00 |
| $k=1$ | $2^{-1}$ | 6.5344E-02 | | 1.7573E-02 | |
| | $2^{-2}$ | 1.7278E-02 | 1.92 | 2.1733E-03 | 3.02 |
| | $2^{-3}$ | 4.3806E-03 | 1.98 | 2.6866E-04 | 3.02 |
| | $2^{-4}$ | 1.0990E-03 | 1.99 | 3.3473E-05 | 3.00 |
| | $2^{-5}$ | 2.7501E-04 | 2.00 | 4.1812E-06 | 3.00 |

we differentiate the error equations in Eq. (4.1) with respect to time $t$. It is worth mentioning that we do not need to assume that the coefficients are independent of time. However, we need to assume the coefficients are independent of time in the previous work [3].

To shorten lengthy equations, we define the following HDG operators $\mathscr{B}_j$ and $\mathscr{C}_j$:

$$
\begin{aligned}
&\mathscr{B}_j\big(\overline{\boldsymbol{q}}_{jh},\overline{u}_{jh},\widehat{\overline{u}}_{jh};\boldsymbol{r}_j,w_j,\mu_j\big)\\
=&\big(c_j\overline{\boldsymbol{q}}_{jh},\boldsymbol{r}_j\big)_{\mathcal{T}_h}-\big(\overline{u}_{jh},\nabla\cdot\boldsymbol{r}_j\big)_{\mathcal{T}_h}+\langle\widehat{\overline{u}}_{jh},\boldsymbol{r}_j\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}+\big(\nabla\cdot\overline{\boldsymbol{q}}_{jh},w_j\big)_{\mathcal{T}_h}
\end{aligned}
$$

$$
\begin{aligned}
&- \langle \overline{\boldsymbol{q}}_{jh} \cdot \boldsymbol{n}, \mu_j \rangle_{\partial \mathcal{T}_h} + \Big\langle h_K^{-1}\big(P_M \overline{u}_{jh} - \widehat{\overline{u}}_{jh}\big), P_M w_j - \mu_j \Big\rangle_{\partial \mathcal{T}_h} \\
&- \big(\boldsymbol{\beta}_j \overline{u}_{jh}, \nabla w_j\big)_{\mathcal{T}_h} - \big(\nabla \cdot \boldsymbol{\beta}_j \overline{u}_{jh}, w_j\big)_{\mathcal{T}_h} + \big\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n} \widehat{\overline{u}}_{jh}, w_j \big\rangle_{\partial \mathcal{T}_h}, \qquad \text{(A.1)}\\
&\mathscr{C}_j\big(\overline{\boldsymbol{q}}_{jh}, \overline{u}_{jh}, \widehat{\overline{u}}_{jh}; \boldsymbol{r}_j, w_j, \mu_j\big) \\
=&\big(c_j \overline{\boldsymbol{q}}_{jh}, \boldsymbol{r}_j\big)_{\mathcal{T}_h} - \big(\overline{u}_{jh}, \nabla \cdot \boldsymbol{r}_j\big)_{\mathcal{T}_h} + \big\langle \widehat{\overline{u}}_{jh}, \boldsymbol{r}_j \cdot \boldsymbol{n} \big\rangle_{\partial \mathcal{T}_h} + \big(\nabla \cdot \overline{\boldsymbol{q}}_{jh}, w_j\big)_{\mathcal{T}_h} \\
&- \langle \overline{\boldsymbol{q}}_{jh} \cdot \boldsymbol{n}, \mu_j \rangle_{\partial \mathcal{T}_h} + \Big\langle h_K^{-1}\big(P_M \overline{u}_{jh} - \widehat{\overline{u}}_{jh}\big), P_M w_j - \mu_j \Big\rangle_{\partial \mathcal{T}_h} \\
&+ \big(\boldsymbol{\beta}_j \overline{u}_{jh}, \nabla w_j\big)_{\mathcal{T}_h} - \big\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n} \widehat{\overline{u}}_{jh}, w_j \big\rangle_{\partial \mathcal{T}_h}.
\end{aligned}
$$

By the definition of (A.1), we can rewrite the HDG formulation of the system (4.1), as follows: find $(\overline{\boldsymbol{q}}_{jh}, \overline{u}_{jh}, \widehat{\overline{u}}_{jh}) \in \boldsymbol{V}_h \times W_h \times M_h(g_j)$ such that

$$
\mathscr{B}_j\big(\overline{\boldsymbol{q}}_{jh}, \overline{u}_{jh}, \widehat{\overline{u}}_{jh}; \boldsymbol{r}_j, w_j, \mu_j\big) = (f_j - \partial_t u_j, w_j)_{\mathcal{T}_h} \tag{A.2}
$$

for all $(\boldsymbol{r}_j, w_j, \mu_j) \in \boldsymbol{V}_h \times W_h \times M_h(0)$.

In the next lemmas, we present some basic properties of the operators $\mathscr{B}_j$ and $\mathscr{C}_j$.

**Lemma A.1.** *For any $(\overline{\boldsymbol{v}}_{jh}, \overline{w}_{jh}, \overline{\mu}_{jh}) \in \boldsymbol{V}_h \times W_h \times M_h(0)$, we have*

$$
\begin{aligned}
&\mathscr{B}_j\big(\overline{\boldsymbol{v}}_{jh}, \overline{w}_{jh}, \overline{\mu}_{jh}; \overline{\boldsymbol{v}}_{jh}, \overline{w}_{jh}, \overline{\mu}_{jh}\big) \\
=&\big(c_j \overline{\boldsymbol{v}}_{jh}, \overline{\boldsymbol{v}}_{jh}\big)_{\mathcal{T}_h} + \Big\langle h_K^{-1}\big(P_M \overline{w}_{jh} - \overline{\mu}_{jh}\big), P_M \overline{w}_{jh} - \overline{\mu}_{jh} \Big\rangle_{\partial \mathcal{T}_h} \\
&- \frac{1}{2}\Big\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n}\big(\overline{w}_{jh} - \overline{\mu}_{jh}\big), \overline{w}_{jh} - \overline{\mu}_{jh} \Big\rangle_{\partial \mathcal{T}_h} - \frac{1}{2}\big(\nabla \cdot \boldsymbol{\beta}_j \overline{w}_{jh}, \overline{w}_{jh}\big)_{\mathcal{T}_h}.
\end{aligned}
$$

**Lemma A.2.** *For any $(\overline{\boldsymbol{v}}_{jh}, \overline{w}_{jh}, \widehat{\overline{u}}_{jh}; \overline{\boldsymbol{p}}_{jh}, \overline{z}_{jh}, \widehat{\overline{z}}_{jh}) \in \boldsymbol{V}_h \times W_h \times M_h(0) \times \boldsymbol{V}_h \times W_h \times M_h(0)$, we have*

$$
\begin{aligned}
&\mathscr{B}_j\Big(\overline{\boldsymbol{v}}_{jh}, \overline{w}_{jh}, \widehat{\overline{u}}_{jh}; \overline{\boldsymbol{p}}_{jh}, -\overline{z}_{jh}, -\widehat{\overline{z}}_{jh}\Big) + \mathscr{C}_j\left(\overline{\boldsymbol{p}}_{jh}, \overline{z}_{jh}, \widehat{\overline{z}}_{jh}; -\overline{\boldsymbol{v}}_{jh}, \overline{w}_{jh}, \widehat{\overline{u}}_{jh}\right) \\
=&\Big\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n}\big(\overline{w}_{jh} - \widehat{\overline{w}}_{jh}\big), \overline{z}_{jh} - \widehat{\overline{z}}_{jh} \Big\rangle_{\partial \mathcal{T}_h}.
\end{aligned}
$$

*Proof.* By definition:

$$
\begin{aligned}
&\mathscr{B}_j\left(\overline{\boldsymbol{v}}_{jh}, \overline{w}_{jh}, \widehat{\overline{u}}_{jh}; \overline{\boldsymbol{p}}_{jh}, -\overline{z}_{jh}, -\widehat{\overline{z}}_{jh}\right) + \mathscr{C}_j\left(\overline{\boldsymbol{p}}_{jh}, \overline{z}_{jh}, \widehat{\overline{z}}_{jh}; -\overline{\boldsymbol{v}}_{jh}, \overline{w}_{jh}, \widehat{\overline{u}}_{jh}\right) \\
=&\big(c_j \overline{\boldsymbol{v}}_{jh}, \overline{\boldsymbol{p}}_{jh}\big)_{\mathcal{T}_h} - \big(\overline{w}_{jh}, \nabla \cdot \overline{\boldsymbol{p}}_{jh}\big)_{\mathcal{T}_h} + \big\langle \widehat{\overline{w}}_{jh}, \overline{\boldsymbol{p}}_{jh} \cdot \boldsymbol{n} \big\rangle_{\partial \mathcal{T}_h} - \big(\nabla \cdot \overline{\boldsymbol{v}}_{jh}, \overline{z}_{jh}\big)_{\mathcal{T}_h} \\
&- \Big\langle h_K^{-1}(P_M \overline{w}_{jh} - \widehat{\overline{w}}_{jh}), P_M \overline{z}_{jh} - \widehat{\overline{z}}_{jh} \Big\rangle_{\partial \mathcal{T}_h} + \big\langle \overline{\boldsymbol{v}}_{jh} \cdot \boldsymbol{n}, \widehat{\overline{z}}_{jh} \big\rangle_{\partial \mathcal{T}_h} \\
&+ \big(\boldsymbol{\beta}_j \overline{w}_{jh}, \nabla \overline{z}_{jh}\big)_{\mathcal{T}_h} + \big(\nabla \cdot \boldsymbol{\beta}_j \overline{w}_{jh}, \overline{z}_{jh}\big)_{\mathcal{T}_h} - \big\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n} \widehat{\overline{w}}_{jh}, \overline{z}_{jh} \big\rangle_{\partial \mathcal{T}_h} \\
&- \big(c_j \boldsymbol{p}_h, \overline{\boldsymbol{v}}_{jh}\big)_{\mathcal{T}_h} + \big(\overline{z}_{jh}, \nabla \cdot \overline{\boldsymbol{v}}_{jh}\big)_{\mathcal{T}_h} - \big\langle \widehat{\overline{z}}_{jh}, \overline{\boldsymbol{v}}_{jh} \cdot \boldsymbol{n} \big\rangle_{\partial \mathcal{T}_h} + \big(\nabla \cdot \overline{\boldsymbol{p}}_{jh}, \overline{w}_{jh}\big)_{\mathcal{T}_h} \\
&+ \Big\langle h_K^{-1}\big(P_M \overline{z}_{jh} - \widehat{\overline{z}}_{jh}\big), P_M \overline{w}_{jh} - \widehat{\overline{w}}_{jh} \Big\rangle_{\partial \mathcal{T}_h} + \big\langle \overline{\boldsymbol{p}}_{jh} \cdot \boldsymbol{n}, \widehat{\overline{w}}_{jh} \big\rangle_{\partial \mathcal{T}_h}
\end{aligned}
$$

$$
\begin{aligned}
&+ \left(\boldsymbol{\beta}_j \overline{z}_{jh}, \nabla \overline{w}_{jh}\right)_{\mathcal{T}_h} - \left\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n} \widehat{\overline{z}}_{jh}, \overline{w}_{jh}\right\rangle_{\partial\mathcal{T}_h} \\
=&\left(\boldsymbol{\beta}_j \overline{w}_{jh}, \nabla \overline{z}_{jh}\right)_{\mathcal{T}_h} + \left(\nabla\cdot\boldsymbol{\beta}_j \overline{w}_{jh}, \overline{z}_{jh}\right)_{\mathcal{T}_h} - \left\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n} \widehat{\overline{w}}_{jh}, \overline{z}_{jh}\right\rangle_{\partial\mathcal{T}_h} \\
&+ \left(\boldsymbol{\beta}_j \overline{z}_{jh}, \nabla \overline{w}_{jh}\right)_{\mathcal{T}_h} - \left\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n} \widehat{\overline{z}}_{jh}, \overline{w}_{jh}\right\rangle_{\partial\mathcal{T}_h} \\
=&\left\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n}\left(\overline{w}_{jh} - \widehat{\overline{w}}_{jh}\right), \overline{z}_{jh} - \widehat{\overline{z}}_{jh}\right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

□

## A.1 Proof of main result

### A.1.1 Step 1: Error equation

**Lemma A.3.** *For* $\varepsilon^{\boldsymbol{q}}_{jh} = \boldsymbol{\Pi}_k \boldsymbol{q}_j - \overline{\boldsymbol{q}}_{jh}$, $\varepsilon^{u}_{jh} = \Pi_{k+1} u_j - \overline{u}_{jh}$ *and* $\varepsilon^{\widehat{u}}_{jh} = P_M u_j - \widehat{\overline{u}}_{jh}$, *we have*

$$
\begin{aligned}
&\mathscr{B}_j\left(\varepsilon^{\boldsymbol{q}}_{jh}, \varepsilon^{u}_{jh}, \varepsilon^{\widehat{u}}_{jh}; \boldsymbol{r}_j, w_j, \mu_j\right) \\
=&\left\langle \left(\boldsymbol{\Pi}_k \boldsymbol{q}_j - \boldsymbol{q}_j\right)\cdot\boldsymbol{n}, w_j - \mu_j\right\rangle_{\partial\mathcal{T}_h} + \left\langle h_K^{-1}\left(\Pi_{k+1}u_j - u_j\right), P_M w_j - \mu_j\right\rangle_{\partial\mathcal{T}_h} \\
&- \left(\boldsymbol{\beta}\left(\Pi_{k+1}u_j - u_j\right), \nabla w_j\right)_{\mathcal{T}_h} - \left(\nabla\cdot\boldsymbol{\beta}\left(\Pi_{k+1}u_j - u_j\right), w_j\right)_{\mathcal{T}_h} \\
&+ \left\langle \boldsymbol{\beta}\cdot\boldsymbol{n}\left(P_M u_j - u_j\right), w_j - \mu_j\right\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
\tag{A.3}
$$

*Proof.* By the definition of operator $\mathscr{B}_j$ in (A.1), we have

$$
\begin{aligned}
&\mathscr{B}_j\left(\boldsymbol{\Pi}_k \boldsymbol{q}_j, \Pi_{k+1}u_j, P_M u_j, \boldsymbol{r}_j, w_j, \mu_j\right) \\
=&\left(c_j \boldsymbol{\Pi}_k \boldsymbol{q}_j, \boldsymbol{r}_j\right)_{\mathcal{T}_h} - \left(\Pi_{k+1}u_j, \nabla\cdot\boldsymbol{r}_j\right)_{\mathcal{T}_h} + \left\langle P_M u_j, \boldsymbol{r}_j\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} \\
&+ \left(\nabla\cdot\left(\boldsymbol{\Pi}_k \boldsymbol{q}_j\right), w_j\right)_{\mathcal{T}_h} + \left\langle h_K^{-1}\left(\Pi_{k+1}u_j - u_j\right), P_M w_j - \mu_j\right\rangle_{\partial\mathcal{T}_h} \\
&- \left\langle \boldsymbol{\Pi}_k \boldsymbol{q}_j\cdot\boldsymbol{n}, \mu_j\right\rangle_{\partial\mathcal{T}_h} - \left(\boldsymbol{\beta}_j \Pi_{k+1}u_j, \nabla w_j\right)_{\mathcal{T}_h} \\
&- \left(\nabla\cdot\boldsymbol{\beta}_j \Pi_{k+1}u_j, w_j\right)_{\mathcal{T}_h} + \left\langle \boldsymbol{\beta}_j\cdot\boldsymbol{n} P_M u_j, w_j\right\rangle_{\partial\mathcal{T}_h} \\
=&\left(c_j\left(\boldsymbol{\Pi}_k \boldsymbol{q}_j - \boldsymbol{q}_j\right), \boldsymbol{r}_j\right)_{\mathcal{T}_h} + \left(c_j \boldsymbol{q}_j, \boldsymbol{r}_j\right)_{\mathcal{T}_h} - \left(u_j, \nabla\cdot\boldsymbol{r}_j\right)_{\mathcal{T}_h} \\
&+ \left\langle u_j, \boldsymbol{r}_j\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} + \left\langle \boldsymbol{\Pi}_k \boldsymbol{q}_j\cdot\boldsymbol{n}, w_j\right\rangle_{\partial\mathcal{T}_h} - \left(\boldsymbol{\Pi}_k \boldsymbol{q}_j, \nabla w_j\right)_{\mathcal{T}_h} \\
&+ \left\langle h_K^{-1}\left(\Pi_{k+1}u_j - u_j\right), P_M w_j - \mu_j\right\rangle_{\partial\mathcal{T}_h} - \left\langle \boldsymbol{\Pi}_k \boldsymbol{q}_j\cdot\boldsymbol{n}, \mu_j\right\rangle_{\partial\mathcal{T}_h} \\
&- \left(\boldsymbol{\beta}_j\left(\Pi_{k+1}u_j - u_j\right), \nabla w_j\right)_{\mathcal{T}_h} + \left(\boldsymbol{\beta}_j \nabla u_j, w_j\right)_{\mathcal{T}_h} \\
&- \left(\nabla\cdot\boldsymbol{\beta}_j\left(\Pi_{k+1}u_j - u_j\right), w_j\right)_{\mathcal{T}_h} + \left\langle \boldsymbol{\beta}\cdot\boldsymbol{n}\left(P_M u_j - u_j\right), w_j - \mu_j\right\rangle_{\partial\mathcal{T}_h} \\
=&\left(c_j\left(\boldsymbol{\Pi}_k \boldsymbol{q}_j - \boldsymbol{q}_j\right), \boldsymbol{r}\right)_{\mathcal{T}_h} + \left(\boldsymbol{q}_j, \boldsymbol{r}_j\right)_{\mathcal{T}_h} - \left(u_j, \nabla\cdot\boldsymbol{r}_j\right)_{\mathcal{T}_h} + \left\langle u_j, \boldsymbol{r}_j\cdot\boldsymbol{n}\right\rangle_{\partial\mathcal{T}_h} \\
&+ \left\langle \boldsymbol{q}_j\cdot\boldsymbol{n}, w_j\right\rangle_{\partial\mathcal{T}_h} + \left\langle \left(\boldsymbol{\Pi}_k \boldsymbol{q}_j - \boldsymbol{q}_j\right)\cdot\boldsymbol{n}, w_j\right\rangle_{\partial\mathcal{T}_h} - \left(\boldsymbol{q}_j, \nabla w_j\right)_{\mathcal{T}_h}
\end{aligned}
$$

$$
\begin{aligned}
&+\Big\langle h_K^{-1}\big(\Pi_{k+1}u_j-u_j\big),P_Mw_j-\mu_j\Big\rangle_{\partial\mathcal{T}_h}-\Big\langle\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big)\cdot\boldsymbol{n},\mu_j\Big\rangle_{\partial\mathcal{T}_h}\\
&-\Big(\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j-u_j\big),\nabla w_j\Big)_{\mathcal{T}_h}+\big(\boldsymbol{\beta}_j\nabla u_j,w_j\big)_{\mathcal{T}_h}\\
&-\Big(\nabla\cdot\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j-u_j\big),w_j\Big)_{\mathcal{T}_h}+\Big\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(P_Mu_j-u_j\big),w_j-\mu_j\Big\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

Note that the exact state $u_j$ and exact flux $\boldsymbol{q}_j$ satisfy

$$
\begin{aligned}
&\big(c_j\boldsymbol{q}_j,\boldsymbol{r}_j\big)_{\mathcal{T}_h}-\big(u_j,\nabla\cdot\boldsymbol{r}_j\big)_{\mathcal{T}_h}+\langle u_j,\boldsymbol{r}_j\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h}=0,\\
-&\big(\boldsymbol{q}_j,\nabla w_j\big)_{\mathcal{T}_h}+\langle\boldsymbol{q}_j\cdot\boldsymbol{n},w_j\rangle_{\partial\mathcal{T}_h}+\big(\boldsymbol{\beta}_j\nabla u_j,w_j\big)_{\mathcal{T}_h}=\big(f_j-\partial_tu_j,w_j\big)_{\mathcal{T}_h},\\
&\langle\boldsymbol{q}_j\cdot\boldsymbol{n},\mu_j\rangle_{\partial\mathcal{T}_h}=0
\end{aligned}
$$

for all $(\boldsymbol{r}_j,w_j,\mu_j)\in\boldsymbol{V}_h\times W_h\times M_h(0)$. Then we have

$$
\begin{aligned}
&\mathscr{B}_j\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j,\Pi_{k+1}u_j,P_Mu_j,\boldsymbol{r}_j,w_j,\mu_j\big)\\
=&\Big(c_j\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big),\boldsymbol{r}_j\Big)_{\mathcal{T}_h}+\Big\langle\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big)\cdot\boldsymbol{n},w_j-\mu_j\Big\rangle_{\partial\mathcal{T}_h}\\
&+\big(f_j-\partial_tu_j,w_j\big)_{\mathcal{T}_h}+\Big\langle h_K^{-1}\big(\Pi_{k+1}u_j-u_j\big),P_Mw_j-\mu_j\Big\rangle_{\partial\mathcal{T}_h}\\
&-\Big(\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j-u_j\big),\nabla w_j\Big)_{\mathcal{T}_h}-\Big(\nabla\cdot\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j-u_j\big),w_j\Big)_{\mathcal{T}_h}\\
&+\Big\langle\boldsymbol{\beta}\cdot\boldsymbol{n}\big(P_Mu_j-u_j\big),w_j-\mu_j\Big\rangle_{\partial\mathcal{T}_h},
\end{aligned}
\tag{A.4}
$$

subtract (A.2) from (A.4), we have

$$
\begin{aligned}
&\mathscr{B}_j\left(\varepsilon_{jh}^{\boldsymbol{q}},\varepsilon_{jh}^{u},\varepsilon_{jh}^{\widehat{u}};\boldsymbol{r}_j,w_j,\mu_j\right)\\
=&\Big(c_j\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big),\boldsymbol{r}\Big)_{\mathcal{T}_h}+\Big\langle\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big)\cdot\boldsymbol{n},w_j-\mu_j\Big\rangle_{\partial\mathcal{T}_h}\\
&+\Big\langle h_K^{-1}\big(\Pi_{k+1}u_j-u_j\big),P_Mw_j-\mu_j\Big\rangle_{\partial\mathcal{T}_h}-\Big(\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j-u_j\big),\nabla w_j\Big)_{\mathcal{T}_h}\\
&-\Big(\nabla\cdot\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j-u_j\big),w_j\Big)_{\mathcal{T}_h}+\Big\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(P_Mu_j-u_j\big),w_j-\mu_j\Big\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

□

### A.1.2 Step 2: Estimate for $\varepsilon_h^q$

The proof of the following lemma is similar to a result established in [23] and hence is omitted.

**Lemma A.4.** *For all $j=1,\ldots,J$ and $(\varepsilon_{jh}^u,\varepsilon_{jh}^{\widehat{u}})\in W_h\times M_h(0)$, we have*

$$
\begin{aligned}
&\big\|\nabla\varepsilon_{jh}^u\big\|_{\mathcal{T}_h}+\Big\|h_K^{-\frac12}\big(\varepsilon_{jh}^u-\varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}\\
\le&C\big\|\varepsilon_{jh}^{\boldsymbol{q}}\big\|_{\mathcal{T}_h}+C\Big\|h_K^{-\frac12}\big(P_M\varepsilon_{jh}^u-\varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}+C\big\|\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big\|_{\mathcal{T}_h}.
\end{aligned}
$$

The next lemma is based on energy arguments.

**Lemma A.5.** *For $h$ small enough, we have*

$$
\begin{aligned}
&\|\varepsilon_{jh}^{\boldsymbol q}\|_{\mathcal T_h}^2 + \left\|h_K^{-\frac12}(P_M\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u})\right\|_{\partial\mathcal T_h}^2 \\
\le & C\|\boldsymbol\Pi_k\boldsymbol q_j - \boldsymbol q_j\|_{\mathcal T_h}^2 + C\left\|h_K^{\frac12}(\boldsymbol\Pi_k\boldsymbol q_j - \boldsymbol q_j)\right\|_{\partial\mathcal T_h}^2 + C\left\|h_K^{-\frac12}(\Pi_{k+1}u_j - u_j)\right\|_{\partial\mathcal T_h}^2 \\
&+ C\left\|h_K^{\frac12}(P_M u_j - u_j)\right\|_{\partial\mathcal T_h}^2 + C\|\Pi_{k+1}u_j - u_j\|_{\mathcal T_h}^2.
\end{aligned}
$$

*Proof.* First, the basic property of $\mathscr B_j$ in Lemma A.1 and use $\nabla\cdot\boldsymbol\beta_j \le 0$ to get

$$
\begin{aligned}
&\mathscr B_j\left(\varepsilon_{jh}^{\boldsymbol q}, \varepsilon_{jh}^u, \varepsilon_{jh}^{\widehat u}; \varepsilon_{jh}^{\boldsymbol q}, \varepsilon_{jh}^u, \varepsilon_{jh}^{\widehat u}\right) \\
&\ge (c_j\varepsilon_{jh}^{\boldsymbol q}, \varepsilon_{jh}^{\boldsymbol q})_{\mathcal T_h} + \left\|h_K^{-\frac12}(P_M\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u})\right\|_{\partial\mathcal T_h}^2 \\
&\quad - \frac12\left\langle \boldsymbol\beta_j\cdot\boldsymbol n(\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u}), \varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u}\right\rangle_{\partial\mathcal T_h}.
\end{aligned}
$$

Then, taking $(\boldsymbol r_j, w_j, \mu_j) = (\varepsilon_{jh}^{\boldsymbol q}, \varepsilon_{jh}^u, \varepsilon_{jh}^{\widehat u})$ in (A.3) and the stability (3.7) with $(\gamma, w, \mu) = (\boldsymbol\beta_j, \varepsilon_{jh}^u, \varepsilon_{jh}^{\widehat u})$, we have

$$
\begin{aligned}
&(c_j\varepsilon_{jh}^{\boldsymbol q}, \varepsilon_{jh}^{\boldsymbol q})_{\mathcal T_h} + \frac12\left\|h_K^{-\frac12}(P_M\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u})\right\|_{\partial\mathcal T_h}^2 \\
\le & Ch\|\nabla\varepsilon_{jh}^u\|_{\mathcal T_h}^2 + \left(c_j(\boldsymbol\Pi_k\boldsymbol q_j - \boldsymbol q_j), \varepsilon_{jh}^{\boldsymbol q}\right)_{\mathcal T_h} + \left\langle(\boldsymbol\Pi_k\boldsymbol q_j - \boldsymbol q_j)\cdot\boldsymbol n, \varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u}\right\rangle_{\partial\mathcal T_h} \\
&- \left\langle h_K^{-1}(\Pi_{k+1}u_j - u_j), P_M\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u}\right\rangle_{\partial\mathcal T_h} - \left(\boldsymbol\beta_j(\Pi_{k+1}u_j - u_j), \nabla\varepsilon_{jh}^u\right)_{\mathcal T_h} \\
&- \left(\nabla\cdot\boldsymbol\beta_j(\Pi_{k+1}u_j - u_j), \varepsilon_{jh}^u\right)_{\mathcal T_h} + \left\langle\boldsymbol\beta_j\cdot\boldsymbol n(P_M u_j - u_j), \varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u}\right\rangle_{\partial\mathcal T_h} \\
=:& \sum_{i=1}^7 R_i.
\end{aligned}
$$

Next, we estimate $\{R_i\}_{i=1}^7$ term by term. First, by A.4 and Young's inequality, we have

$$
\begin{aligned}
R_1 &\le Ch\|\varepsilon_{jh}^{\boldsymbol q}\|_{\mathcal T_h}^2 + Ch\left\|h_K^{-\frac12}(P_M\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u})\right\|_{\partial\mathcal T_h}^2 + C\|\boldsymbol\Pi_k\boldsymbol q_j - \boldsymbol q_j\|_{\mathcal T_h}^2, \\
R_3 &\le C\left\|h_K^{\frac12}(\boldsymbol\Pi_k\boldsymbol q_j - \boldsymbol q_j)\right\|_{\partial\mathcal T_h}^2 + \frac1{16}\|\varepsilon_{jh}^{\boldsymbol q}\|_{\mathcal T_h}^2 \\
&\quad + \frac1{16}\left\|h_K^{-\frac12}(P_M\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u})\right\|_{\partial\mathcal T_h}^2 + C\|\boldsymbol\Pi_k\boldsymbol q_j - \boldsymbol q_j\|_{\mathcal T_h}^2, \\
R_5 &\le C\|\Pi_{k+1}u_j - u_j\|_{\mathcal T_h}^2 + \frac1{16}\|\varepsilon_{jh}^{\boldsymbol q}\|_{\mathcal T_h}^2 \\
&\quad + \frac1{16}\left\|h_K^{-\frac12}(P_M\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat u})\right\|_{\partial\mathcal T_h}^2 + C\|\boldsymbol\Pi_k\boldsymbol q_j - \boldsymbol q_j\|_{\mathcal T_h}^2,
\end{aligned}
$$

$$
\begin{aligned}
R_7 \le\, & C\Big\|h_K^{\frac{1}{2}}\big(P_M u_j - u_j\big)\Big\|_{\partial\mathcal{T}_h}\Big\|h_K^{-\frac{1}{2}}\big(\varepsilon_{jh}^{u} - \varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h} \\
\le\, & C\Big\|h_K^{\frac{1}{2}}(P_M u_j - u_j)\Big\|_{\partial\mathcal{T}_h}^2 + \frac{1}{16}\big\|\varepsilon_{jh}^{\boldsymbol{q}}\big\|_{\mathcal{T}_h}^2 \\
& + \frac{1}{16}\Big\|h_K^{-\frac{1}{2}}\big(P_M\varepsilon_{jh}^{u} - \varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}^2 + C\big\|\boldsymbol{\Pi}_k\boldsymbol{q}_j - \boldsymbol{q}_j\big\|_{\mathcal{T}_h}^2.
\end{aligned}
$$

Young's inequality for the terms $R_2$ and $R_4$,

$$
\begin{aligned}
R_2 &\le C\big\|\boldsymbol{\Pi}_k\boldsymbol{q}_j - \boldsymbol{q}_j\big\|_{\mathcal{T}_h}^2 + \frac{1}{16}\big\|\varepsilon_{jh}^{\boldsymbol{q}}\big\|_{\mathcal{T}_h}^2, \\
R_4 &\le C\Big\|h_K^{-\frac{1}{2}}\big(\Pi_{k+1}u_j - u_j\big)\Big\|_{\partial\mathcal{T}_h}^2 + \frac{1}{16}\Big\|h_K^{-\frac{1}{2}}\big(P_M\varepsilon_{jh}^{u} - \varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}^2.
\end{aligned}
$$

For the term $R_6$, using the Poincaré inequality Lemmas 3.7 and A.4, we have

$$
\begin{aligned}
R_6 \le\, & C\big\|\Pi_{k+1}u_j - u_j\big\|_{\mathcal{T}_h}\left(\big\|\nabla\varepsilon_{jh}^{u}\big\|_{\mathcal{T}_h} + \Big\|h_K^{-\frac{1}{2}}\big(\varepsilon_{jh}^{u} - \varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}\right) \\
\le\, & C\big\|\Pi_{k+1}u_j - u_j\big\|_{\mathcal{T}_h}^2 + \frac{1}{16}\big\|\varepsilon_{jh}^{\boldsymbol{q}}\big\|_{\mathcal{T}_h}^2 + \frac{1}{16}\Big\|h_K^{-\frac{1}{2}}\big(P_M\varepsilon_{jh}^{u} - \varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}^2 \\
& + C\big\|\boldsymbol{\Pi}_k\boldsymbol{q}_j - \boldsymbol{q}_j\big\|_{\mathcal{T}_h}^2.
\end{aligned}
$$

Sum all the estimates above, and let $h$ small enough, we get

$$
\begin{aligned}
& \big\|\varepsilon_{jh}^{\boldsymbol{q}}\big\|_{\mathcal{T}_h}^2 + \Big\|h_K^{-\frac{1}{2}}\big(P_M\varepsilon_{jh}^{u} - \varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}^2 \\
\le & C\big\|\boldsymbol{\Pi}_k\boldsymbol{q}_j - \boldsymbol{q}_j\big\|_{\mathcal{T}_h}^2 + C\Big\|h_K^{\frac{1}{2}}\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j - \boldsymbol{q}_j\big)\Big\|_{\partial\mathcal{T}_h}^2 + C\Big\|h_K^{-\frac{1}{2}}\big(\Pi_{k+1}u_j - u_j\big)\Big\|_{\partial\mathcal{T}_h}^2 \\
& + C\Big\|h_K^{\frac{1}{2}}\big(P_M u_j - u_j\big)\Big\|_{\partial\mathcal{T}_h}^2 + C\big\|\Pi_{k+1}u_j - u_j\big\|_{\mathcal{T}_h}^2. \qquad \square
\end{aligned}
$$

As a consequence, a simple application of the triangle inequality gives optimal convergence rates for $\|\boldsymbol{q}_j - \overline{\boldsymbol{q}}_{jh}\|_{\mathcal{T}_h}$:

**Lemma A.6.** *We have*

$$
\big\|\boldsymbol{q}_j - \overline{\boldsymbol{q}}_{jh}\big\|_{\mathcal{T}_h} \le \big\|\boldsymbol{q}_j - \boldsymbol{\Pi}_k\boldsymbol{q}_j\big\|_{\mathcal{T}_h} + \big\|\boldsymbol{\Pi}_k\boldsymbol{q}_j - \overline{\boldsymbol{q}}_{jh}\big\|_{\mathcal{T}_h} \le Ch^{k+1}.
$$

### A.1.3 Step 3: Estimate for $\varepsilon_{jh}^{u}$ by a duality argument

The next step is the consideration of the dual problems:

$$
\begin{aligned}
& c_j\boldsymbol{\Phi}_j + \nabla\Psi_j = 0 && \text{in } \Omega, \\
& \nabla\cdot\boldsymbol{\Phi}_j - \boldsymbol{\beta}_j\cdot\nabla\Psi_j = \Theta_j && \text{in } \Omega, \\
& \Psi_j = 0 && \text{on } \partial\Omega.
\end{aligned} \tag{A.5}
$$

**Elliptic regularity.** Since the domain $\Omega$ is convex, we have the following regularity estimate

$$\|\boldsymbol{\Phi}_{j[H^1(\Omega)]^d}\| + \|\Psi_{jH^2(\Omega)}\| \le C_{\rm reg}\|\Theta_{jL^2(\Omega)}\|. \tag{A.6}$$

With the above dual problems (A.5) and regularity (A.6), we can derive the following error estimates.

**Lemma A.7.** *For $h$ small enough, we have*

$$\begin{aligned}\|\varepsilon_{jh}^u\|_{\mathcal{T}_h} \le & Ch^{\frac{3}{2}}\|\boldsymbol{\Pi}_k\boldsymbol{q}_j - \boldsymbol{q}_j\|_{\partial\mathcal{T}_h} + Ch\|\boldsymbol{\Pi}_k\boldsymbol{q}_j - \boldsymbol{q}_j\|_{\mathcal{T}_h} + Ch\|\varepsilon_{jh}^{\boldsymbol{q}}\|_{\mathcal{T}_h}\\ &+ Ch\Big\|h_K^{-1}\big(\Pi_{k+1}u_j - u_j\big)\Big\|_{\partial\mathcal{T}_h} + Ch\Big\|h_K^{-\frac{1}{2}}\big(\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}\\ &+ Ch\Big\|h_K^{-\frac{1}{2}}\big(P_M\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h} + Ch^{\frac{3}{2}}\|P_Mu_j - u_j\|_{\partial\mathcal{T}_h} + C\|\Pi_{k+1}u_j - u_j\|_{\mathcal{T}_h}.\end{aligned}$$

*Proof.* Consider the dual problem (A.5) and let $\Theta_j = \varepsilon_{jh}^u$, we take $(\boldsymbol{r}_j, w_j, \mu_j) = (-\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j, \Pi_{k+1}\Psi_j, P_M\Psi_j)$ in Eq. (A.3) in Lemma A.3, we have

$$\begin{aligned}&\mathscr{B}_j\left(\varepsilon_{jh}^{\boldsymbol{q}}, \varepsilon_{jh}^u, \varepsilon_{jh}^{\widehat{u}}; -\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j, \Pi_{k+1}\Psi_j, P_M\Psi_j\right)\\ =&\mathscr{C}_j\left(\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j, \Pi_{k+1}\Psi_j, P_M\Psi_j; -\varepsilon_{jh}^{\boldsymbol{q}}, \varepsilon_{jh}^u, \varepsilon_{jh}^{\widehat{u}}\right)\\ &+ \left\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}}\big), \Pi_{k+1}\Psi_j - P_M\Psi_j\right\rangle_{\partial\mathcal{T}_h}\\ =&-\left(c_j\big(\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j - \boldsymbol{\Phi}_j\big), \varepsilon_{jh}^{\boldsymbol{q}}\right)_{\mathcal{T}_h} + \left\langle\big(\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j - \boldsymbol{\Phi}_j\big)\cdot\boldsymbol{n}, \varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}}\right\rangle_{\partial\mathcal{T}_h}\\ &+ \left\langle h_K^{-1}(\Pi_{k+1}\Psi_j - \Psi_j), P_M\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}}\right\rangle_{\partial\mathcal{T}_h} + \|\varepsilon_{jh}^u\|_{\mathcal{T}_h}^2\\ &+ \left(\boldsymbol{\beta}_j\big(\Pi_{k+1}\Psi_j - \Psi_j\big), \nabla\varepsilon_{jh}^u\right)_{\mathcal{T}_h} - \left\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(P_M\Psi_j - \Psi_j\big), \varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}}\right\rangle_{\partial\mathcal{T}_h}\\ &+ \left\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(\varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}}\big), \Pi_{k+1}\Psi_j - P_M\Psi_j\right\rangle_{\partial\mathcal{T}_h}.\end{aligned} \tag{A.7}$$

On the other hand, by (A.3), we have

$$\begin{aligned}&\mathscr{B}_j\left(\varepsilon_{jh}^{\boldsymbol{q}}, \varepsilon_{jh}^u, \varepsilon_{jh}^{\widehat{u}}; -\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j, \Pi_{k+1}\Psi_j, P_M\Psi_j\right)\\ =&-\left(c_j\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j - \boldsymbol{q}_j\big), \boldsymbol{\Pi}_k\boldsymbol{\Phi}_j\right)_{\mathcal{T}_h} + \left\langle\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j - \boldsymbol{q}_j\big)\cdot\boldsymbol{n}, \Pi_{k+1}\Psi_j - P_M\Psi_j\right\rangle_{\partial\mathcal{T}_h}\\ &+ \left\langle h_K^{-1}\big(\Pi_{k+1}u_j - u_j\big), P_M\Pi_{k+1}\Psi_j - P_M\Psi_j\right\rangle_{\partial\mathcal{T}_h} - \left(\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j - u_j\big), \nabla\Pi_{k+1}\Psi_j\right)_{\mathcal{T}_h}\\ &- \left(\nabla\cdot\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j - u_j\big), \Pi_{k+1}\Psi_j\right)_{\mathcal{T}_h} + \left\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(P_Mu_j - u_j\big), \Pi_{k+1}\Psi_j - P_M\Psi_j\right\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

Since there holds

$$\begin{aligned}&\left\langle\big(\boldsymbol{\Pi}_k^o\boldsymbol{q}_j - \boldsymbol{q}_j\big)\cdot\boldsymbol{n}, P_M\Psi\right\rangle_{\partial\mathcal{T}_h}\\ =&\langle\boldsymbol{\Pi}_k^o\boldsymbol{q}_j\cdot\boldsymbol{n}, P_M\Psi\rangle_{\partial\mathcal{T}_h} - \langle\boldsymbol{q}_j\cdot\boldsymbol{n}, P_M\Psi\rangle_{\partial\mathcal{T}_h} = \langle\boldsymbol{\Pi}_k\boldsymbol{q}_j\cdot\boldsymbol{n}, P_M\Psi\rangle_{\partial\mathcal{T}_h},\\ &\left\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(P_Mu_j - u_j\big), P_M\Psi_j\right\rangle_{\partial\mathcal{T}_h} = 0 = \left\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(P_Mu_j - u_j\big), \Psi_j\right\rangle_{\partial\mathcal{T}_h}.\end{aligned}$$

This gives

$$\begin{aligned}
&\mathscr{B}_j\left(\varepsilon_{jh}^{\boldsymbol{q}},\varepsilon_{jh}^{u},\varepsilon_{jh}^{\widehat{u}};-\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j,\Pi_{k+1}\Psi_j,P_M\Psi_j\right)\\
=&-\Big(c_j\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big),\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j\Big)_{\mathcal{T}_h}+\Big\langle\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big)\cdot\boldsymbol{n},\Pi_{k+1}\Psi_j-\Psi_j\Big\rangle_{\partial\mathcal{T}_h}\\
&+\Big\langle h_K^{-1}\big(\Pi_{k+1}u_j-u_j\big),P_M\Pi_{k+1}\Psi_j-P_M\Psi_j\Big\rangle_{\partial\mathcal{T}_h}-\Big(\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j-u_j\big),\nabla\Pi_{k+1}\Psi_j\Big)_{\mathcal{T}_h}\\
&-\Big(\nabla\cdot\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j-u_j\big),\Pi_{k+1}\Psi_j\Big)_{\mathcal{T}_h}+\Big\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(P_Mu_j-u_j\big),\Pi_{k+1}\Psi_j-\Psi_j\Big\rangle_{\partial\mathcal{T}_h}. \qquad (A.8)
\end{aligned}$$

Comparing the above two equalities (A.7) and (A.8), we have

$$\begin{aligned}
\big\|\varepsilon_{jh}^{u}\big\|_{\mathcal{T}_h}^2=&-\Big(c_j\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big),\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j\Big)_{\mathcal{T}_h}+\Big(c_j\big(\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j-\boldsymbol{\Phi}_j\big),\varepsilon_{jh}^{\boldsymbol{q}}\Big)_{\mathcal{T}_h}\\
&+\Big\langle\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big)\cdot\boldsymbol{n},\Pi_{k+1}\Psi_j-\Psi_j\Big\rangle_{\partial\mathcal{T}_h}\\
&+\Big\langle h_K^{-1}\big(\Pi_{k+1}u_j-u_j\big),P_M\Pi_{k+1}\Psi_j-P_M\Psi_j\Big\rangle_{\partial\mathcal{T}_h}\\
&-\Big\langle\big(\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j-\boldsymbol{\Phi}_j\big)\cdot\boldsymbol{n},\varepsilon_{jh}^{u}-\varepsilon_{jh}^{\widehat{u}}\Big\rangle_{\partial\mathcal{T}_h}-\Big\langle h_K^{-1}\big(\Pi_{k+1}\Psi_j-\Psi_j\big),P_M\varepsilon_{jh}^{u}-\varepsilon_{jh}^{\widehat{u}}\Big\rangle_{\partial\mathcal{T}_h}\\
&-\Big(\boldsymbol{\beta}_j\big(\Pi_{k+1}\Psi_j-\Psi_j\big),\nabla\varepsilon_{jh}^{u}\Big)_{\mathcal{T}_h}+\Big\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(P_M\Psi_j-\Psi_j\big),\varepsilon_{jh}^{u}-\varepsilon_{jh}^{\widehat{u}}\Big\rangle_{\partial\mathcal{T}_h}\\
&-\Big\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(\varepsilon_{jh}^{u}-\varepsilon_{jh}^{\widehat{u}}\big),\Pi_{k+1}\Psi_j-P_M\Psi_j\Big\rangle_{\partial\mathcal{T}_h}-\Big(\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j-u_j\big),\nabla\Pi_{k+1}\Psi_j\Big)_{\mathcal{T}_h}\\
&-\Big(\nabla\cdot\boldsymbol{\beta}_j\big(\Pi_{k+1}u_j-u_j\big),\Pi_{k+1}\Psi_j\Big)_{\mathcal{T}_h}+\Big\langle\boldsymbol{\beta}_j\cdot\boldsymbol{n}\big(P_Mu_j-u_j\big),\Pi_{k+1}\Psi_j-\Psi_j\Big\rangle_{\partial\mathcal{T}_h}\\
=:&\sum_{i=1}^{12}R_i.
\end{aligned}$$

Next, we estimate $\{R_i\}_{i=1}^{12}$ term by term. First,

$$\begin{aligned}
&R_1+R_2\\
=&-\Big(\big(c_j-\Pi_0c_j\big)\big(\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big),\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j\Big)_{\mathcal{T}_h}+\Big(c_j\big(\boldsymbol{\Pi}_k\boldsymbol{\Phi}_j-\boldsymbol{\Phi}_j\big),\varepsilon_{jh}^{\boldsymbol{q}}\Big)_{\mathcal{T}_h}\\
\le& Ch|c_j|_{1,\infty}\big\|\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big\|_{\mathcal{T}_h}\big\|\varepsilon_{jh}^{u}\big\|_{\mathcal{T}_h}+Ch\big\|\varepsilon_{jh}^{\boldsymbol{q}}\big\|_{\mathcal{T}_h}\big\|\varepsilon_{jh}^{u}\big\|_{\mathcal{T}_h}.
\end{aligned}$$

Then, we have

$$\begin{aligned}
&R_3+R_4+R_5+R_6+R_9\\
\le& Ch^{\frac{3}{2}}\Big(\big\|\boldsymbol{\Pi}_k\boldsymbol{q}_j-\boldsymbol{q}_j\big\|_{\partial\mathcal{T}_h}+\Big\|h_K^{-1}\big(\Pi_{k+1}u_j-u_j\big)\Big\|_{\partial\mathcal{T}_h}+\Big\|h_K^{-\frac{1}{2}}\big(\varepsilon_{jh}^{u}-\varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}\\
&+C\Big\|h_K^{-\frac{1}{2}}\big(P_M\varepsilon_{jh}^{u}-\varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}+C\Big\|h_K^{-\frac{1}{2}}\big(\varepsilon_{jh}^{u}-\varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}\Big)\big\|\varepsilon_{jh}^{u}\big\|_{\mathcal{T}_h}.
\end{aligned}$$

For the term $R_7$, by Lemma 3.7, we get

$$R_7\le Ch^2\left(\big\|\varepsilon_{jh}^{\boldsymbol{q}}\big\|_{\mathcal{T}_h}+\Big\|h_K^{-\frac{1}{2}}\big(P_M\varepsilon_{jh}^{u}-\varepsilon_{jh}^{\widehat{u}}\big)\Big\|_{\partial\mathcal{T}_h}\right)\big\|\varepsilon_{jh}^{u}\big\|_{\mathcal{T}_h}.$$

For the terms $R_8$ and $R_{12}$, we have

$$
\begin{aligned}
&R_8 + R_{12} \\
&= \Big\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n}\big(P_M \Psi_j - \Psi_j\big), \varepsilon_{jh}^u \Big\rangle_{\partial\mathcal{T}_h} + \Big\langle \boldsymbol{\beta} \cdot \boldsymbol{n}\big(P_M u_j - u_j\big), \Pi_{k+1}\Psi_j \Big\rangle_{\partial\mathcal{T}_h} \\
&= \Big\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n}\big(P_M \Psi_j - \Psi_j\big), \varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}} \Big\rangle_{\partial\mathcal{T}_h} + \Big\langle \boldsymbol{\beta}_j \cdot \boldsymbol{n}\big(P_M u_j - u_j\big), \Pi_{k+1}\Psi_j - \Psi_j \Big\rangle_{\partial\mathcal{T}_h} \\
&\le C \left( h \Big\| h_K^{-\frac{1}{2}} (P_M \varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}}) \Big\|_{\partial\mathcal{T}_h} + h^{\frac{3}{2}} \big\| P_M u_j - u_j \big\|_{\partial\mathcal{T}_h} \right) \big\| \varepsilon_{jh}^u \big\|_{\mathcal{T}_h}.
\end{aligned}
$$

For the terms $R_{10}$ and $R_{11}$, we use the boundness of $\Pi_{k+1}$ to get

$$
\begin{aligned}
&R_{10} + R_{11} \\
&\le C \big\| \Pi_{k+1} u_j - u_j \big\|_{\mathcal{T}_h} \Big( \big\| \nabla \Pi_{k+1} \Psi_j \big\|_{\mathcal{T}_h} + \big\| \Pi_{k+1}\Psi_j \big\|_{\mathcal{T}_h} \Big) \\
&\le C \big\| \Pi_{k+1} u_j - u_j \big\|_{\mathcal{T}_h} \Big( \big\| \nabla (\Pi_{k+1} \Psi_j - \Psi_j) \big\|_{\mathcal{T}_h} + \big\| \nabla \Psi_j \big\|_{\mathcal{T}_h} + \big\| \Pi_{k+1}\Psi_j \big\|_{\mathcal{T}_h} \Big) \\
&\le C \big\| \Pi_{k+1} u_j - u_j \big\|_{\mathcal{T}_h} \big\| \varepsilon_{jh}^u \big\|_{\mathcal{T}_h}.
\end{aligned}
$$

Thus, combining all the estimates above give

$$
\begin{aligned}
\big\| \varepsilon_{jh}^u \big\|_{\mathcal{T}_h} \le{}& C h^{\frac{3}{2}} \big\| \boldsymbol{\Pi}_k \boldsymbol{q}_j - \boldsymbol{q}_j \big\|_{\partial\mathcal{T}_h} + C h \big\| \boldsymbol{\Pi}_k \boldsymbol{q}_j - \boldsymbol{q}_j \big\|_{\mathcal{T}_h} + C h \big\| \varepsilon_{jh}^{\boldsymbol{q}} \big\|_{\mathcal{T}_h} \\
&+ C h \Big\| h_K^{-1} \big( \Pi_{k+1} u_j - u_j \big) \Big\|_{\partial\mathcal{T}_h} + C h \Big\| h_K^{-\frac{1}{2}} \big( \varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}} \big) \Big\|_{\partial\mathcal{T}_h} \\
&+ C h \Big\| h_K^{-\frac{1}{2}} \big( P_M \varepsilon_{jh}^u - \varepsilon_{jh}^{\widehat{u}} \big) \Big\|_{\partial\mathcal{T}_h} + C h^{\frac{3}{2}} \big\| P_M u_j - u_j \big\|_{\partial\mathcal{T}_h} \\
&+ C \big\| \Pi_{k+1} u_j - u_j \big\|_{\mathcal{T}_h}.
\end{aligned}
$$

As a consequence, a simple application of the triangle inequality gives optimal convergence rates for $\|u_j - \overline{u}_{jh}\|_{\mathcal{T}_h}$ and $\|h_K^{\frac{1}{2}}(\overline{u}_{jh} - \widehat{\overline{u}}_{jh})\|_{\partial\mathcal{T}_h}$.

**Lemma A.8.** *For $h$ small enough, we have*

$$
\begin{aligned}
\big\| u_j - \overline{u}_{jh} \big\|_{\mathcal{T}_h} &\le \big\| \Pi_{k+1} u_j - u_j \big\|_{\mathcal{T}_h} + \big\| \Pi_{k+1} u_j - \overline{u}_{jh} \big\|_{\mathcal{T}_h} \le C h^{k+2}, \\
\Big\| h_K^{\frac{1}{2}} (\overline{u}_{jh} - \widehat{\overline{u}}_{jh}) \Big\|_{\partial\mathcal{T}_h} &\le \Big\| h_K^{\frac{1}{2}} \varepsilon_{jh}^n \Big\|_{\partial\mathcal{T}_h} + \Big\| h_K^{\frac{1}{2}} \big( P_M u_j - \widehat{\overline{u}}_{jh} \big) \Big\|_{\partial\mathcal{T}_h} \\
&\quad + \Big\| h_K^{\frac{1}{2}} \big( \Pi_{k+1} u_j - P_M u_j \big) \Big\|_{\partial\mathcal{T}_h} \le C h^{k+1}.
\end{aligned}
$$

## References

[1] G. CHEN, B. COCKBURN, J. SINGLER, AND Y. ZHANG, *Superconvergent interpolatory HDG methods for reaction diffusion equations I: An* $\mathrm{HDG}_k$ *method*, J. Sci. Comput., 81 (2019), pp. 2188-2212.

[2] G. CHEN, P. MONK, AND Y. ZHANG, *An HDG method for the time-dependent drift-diffusion model of semiconductor devices*, J. Sci. Comput., 80 (2019), pp. 420-443.
[3] G. CHEN, L. PI, L. XU, AND Y. ZHANG, *A superconvergent ensemble HDG method for parameterized convection diffusion equations*, SIAM J. Numer. Anal., 57 (2019), pp. 2551-2578.
[4] Y. CHEN AND B. COCKBURN, *Analysis of variable-degree HDG methods for convection-diffusion equations. Part II: Semimatching nonconforming meshes*, Math. Comp., 83 (2014), pp. 87-111.
[5] B. COCKBURN, Z. FU, A. HUNGRIA, L. JI, M. A. SÁNCHEZ, AND F.-J. SAYAS, *Stormer-Numerov HDG methods for acoustic waves*, J. Sci. Comput., 75 (2018), pp. 597-624.
[6] B. COCKBURN, J. GOPALAKRISHNAN, AND R. LAZAROV, *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1319-1365.
[7] B. COCKBURN AND K. SHI, *Conditions for superconvergence of HDG methods for Stokes flow*, Math. Comp., 82 (2013), pp. 651-671.
[8] J. A. FIORDILINO, *A second order ensemble timestepping algorithm for natural convection*, SIAM J. Numer. Anal., 56 (2018), pp. 816-837.
[9] M. GUNZBURGER, N. JIANG, AND M. SCHNEIER, *An ensemble-proper orthogonal decomposition method for the nonstationary Navier-Stokes equations*, SIAM J. Numer. Anal., 55 (2017), pp. 286-304.
[10] M. GUNZBURGER, N. JIANG, AND M. SCHNEIER, *A higher-order ensemble/proper orthogonal decomposition method for the nonstationary Navier-Stokes equations*, Int. J. Numer. Anal. Model., 15 (2018), pp. 608-627.
[11] M. GUNZBURGER, N. JIANG, AND Z. WANG, *A second-order time-stepping scheme for simulating ensembles of parameterized flow problems*, Comput. Methods Appl. Math., (2017).
[12] M. GUNZBURGER, N. JIANG, AND Z. WANG, *An efficient algorithm for simulating ensembles of parameterized flow problems*, IMA J. Numer. Anal., 39 (2019), pp. 1180-1205.
[13] W. HU, J. SHEN, J. R. SINGLER, Y. ZHANG, AND X. ZHENG, *A superconvergent HDG method for distributed control of convection diffusion PDEs*, J. Sci. Comput., 76 (2018), pp. 1436-1457.
[14] N. JIANG, *A higher order ensemble simulation algorithm for fluid flows*, J. Sci. Comput., 64 (2015), pp. 264-288.
[15] N. JIANG, *A second-order ensemble method based on a blended backward differentiation formula timestepping scheme for time-dependent Navier-Stokes equations*, Numer. Methods Partial Differential Equations, 33 (2017), pp. 34-61.
[16] N. JIANG AND W. LAYTON, *An algorithm for fast calculation of flow ensembles*, Int. J. Uncertain. Quantif., 4 (2014), pp. 273-301.
[17] N. JIANG AND W. LAYTON, *Numerical analysis of two ensemble eddy viscosity numerical regularizations of fluid motion*, Numer. Methods Partial Differential Equations, 31 (2015), pp. 630-651.
[18] N. JIANG AND H. TRAN, *Analysis of a stabilized CNLF method with fast slow wave splittings for flow problems*, Comput. Methods Appl. Math., 15 (2015), pp. 307-330.
[19] C. LEHRENFELD, *Hybrid discontinuous Galerkin methods for solving incompressible flow problems*, (2010). PhD Thesis.
[20] Y. LUO AND Z. WANG, *An ensemble algorithm for numerical solutions to deterministic and random parabolic PDEs*, SIAM J. Numer. Anal., 56 (2018), pp. 859-876.
[21] I. OIKAWA, *A hybridized discontinuous Galerkin method with reduced stabilization*, J. Sci. Comput., 65 (2015), pp. 327-340.
[22] W. QIU, J. SHEN, AND K. SHI, *An HDG method for linear elasticity with strong symmetric*

*stresses*, Math. Comp., 87 (2018), pp. 69-93.

[23] W. QIU AND K. SHI, *An HDG method for convection diffusion equation*, J. Sci. Comput., 66 (2016), pp. 346-357.

[24] W. QIU AND K. SHI, *A superconvergent HDG method for the incompressible Navier-Stokes equations on general polyhedral meshes*, IMA J. Numer. Anal., 36 (2016), pp. 1943-1967.

[25] M. A. SÁNCHEZ, C. CIUCA, N. C. NGUYEN, J. PERAIRE, AND B. COCKBURN, *Symplectic Hamiltonian HDG methods for wave propagation phenomena*, J. Comput. Phys., 350 (2017), pp. 951-973.

[26] F. VIDAL-CODINA, N. C. NGUYEN, S.-H. OH, AND J. PERAIRE, *A hybridizable discontinuous Galerkin method for computing nonlocal electromagnetic effects in three-dimensional metallic nanostructures*, J. Comput. Phys., 355 (2018), pp. 548-565.