Journal of Scientific Computing (2019) 80:420–443
https://doi.org/10.1007/s10915-019-00945-y

An HDG Method for the Time-dependent Drift–Diffusion Model of Semiconductor Devices

Gang Chen[1] · **Peter Monk**[2] · **Yangwen Zhang**[2]

Received: 4 November 2018 / Revised: 9 February 2019 / Accepted: 17 March 2019 /
Published online: 3 April 2019

Abstract
We propose a hybridizable discontinuous Galerkin (HDG) finite element method to approximate the solution of the time dependent drift–diffusion problem. This system involves a nonlinear convection diffusion equation for the electron concentration u coupled to a linear Poisson problem for the electric potential ϕ. The non-linearity in this system is the product of the $\nabla\phi$ with u. An improper choice of a numerical scheme can reduce the convergence rate. To obtain optimal HDG error estimates for ϕ, u and their gradients, we utilize two different HDG schemes to discretize the nonlinear convection diffusion equation and the Poisson equation. We prove optimal order error estimates for the semidiscrete problem. We also present numerical experiments to support our theoretical results.

Keywords Hybridizable discontinuous Galerkin method · Drift–diffusion · Error analysis · Optimal convergence rate

1 Introduction

Drift–diffusion equations play an important role in modeling the movement of charged particles particularly in semiconductor physics [1,2,10,28,45–47,53]. Besides the applications to semiconductors, these kinds of PDEs have many applications in the simulation of batteries [54,64], charged particles in biology [52,65] and physical chemistry [7,30,43,44].

We consider the following model time dependent drift–diffusion equation posed on a Lipschitz polyhedral domain $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$: we seek to determine the unknown electron density u and the electric potential ϕ that satisfy

✉ Peter Monk
monk@udel.edu

Gang Chen
cglwdm@uestc.edu.cn

Yangwen Zhang
ywzhangf@udel.edu

1 College of Mathematics, Sichuan University, Chengdu 610064, China

2 Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA

$$
\begin{aligned}
u_t - \Delta u + \nabla\cdot(u\nabla\phi) &= 0 && \text{in } \Omega\times(0,T], && \text{(1a)}\\
-\varepsilon\Delta\phi + u &= 0 && \text{in } \Omega\times(0,T], && \text{(1b)}\\
u &= g_u && \text{on } \partial\Omega\times(0,T], && \text{(1c)}\\
\phi &= g_\phi && \text{on } \partial\Omega\times(0,T], && \text{(1d)}\\
u(\cdot,0) &= u_0 && \text{in } \Omega, && \text{(1e)}
\end{aligned}
$$

where ε is a constant and typically small in real applications. In our analysis, we have not analyzed the ε dependence of the coefficients. This will be considered in future work. To simplify the presentation, we set $\varepsilon = 1$ in the rest of the paper. We shall discuss the smoothness assumptions on g_u, g_ϕ and u_0 needed for our analysis later in the paper. Applications of the drift–diffusion model often involve more complicated versions of the above model, for example including additional particle transport equations (for example, for holes) and recombination terms. However the above system contains the principle difficulty from the point of view of proving convergence: the term $\nabla\cdot(u\nabla\phi)$.

Theoretical and numerical studies for this type of partial differential equation (PDE) have a long history. For the theoretical analysis of the drift–diffusion system we refer to [5,6,34,35,46,56] and the references therein. Computational studies started in the 1960s [29,39] and many discretization methods have been used for the drift–diffusion system in the past decades. For an extensive body of literature devoted to this subject we refer to, e.g., the finite difference method [31,40,50,55], the finite volume method [3,4,12–14], the standard finite element method (FEM) [36,52,62], and mixed FEM [37,41]. Furthermore, there are many new models in which the drift–diffusion equation coupled with other PDEs; such as Stokes [42], Navier–Stokes [61] and Darcy flow [32]. However these extensions are outside the scope of this paper.

The product of the gradient of the electric potential, $\nabla\phi$ with electron concentration u in (1a) can cause a reduction in the convergence rate of the solution if the numerical schemes for the two equations are not properly devised. In [62], the authors obtained an optimal convergence rate in H^1 norm but a suboptimal in L^2 norm by using the standard FEM. To overcome the convergence order reduction, a new method was proposed to discretize the system (1): mixed FEM for Poisson equation (1b) and standard FEM for (1a). This scheme provides optimal error estimates for u and ϕ in both the H^1 and $H(\text{div})$ norms. Very recently, the authors in [37] obtained an optimal convergence rate by using mixed FEM for both (1a) and (1b).

In the drift–diffusion model, typically, the magnitude of $\nabla\phi$ is huge (see [9]). Therefore, it is natural to consider the discontinuous Galerkin (DG) method to discretize the system (1). In [51], a local DG (LDG) method was used to study a 1D drift-diffsuion equation, they obtained an optimal convergence rate by using an important relationship between the gradient and interface jump of the numerical solution with the independent numerical solution of the gradient in the LDG methods; see [63, Lemma 2.4] and [51, Lemma 4.3]. However, to the best of our knowledge, the inequality in [63, Lemma 2.4] is not straightforward to extend to high dimensions.

Moreover, the number of degrees of freedom for the DG or LDG methods is much larger compared to standard FEM; this is the main drawback of DG methods. Hybridizable discontinuous Galerkin (HDG) methods were originally proposed in [25] to remedy this issue. The global system of HDG methods only involve the degrees of freedom on the interfaces between elements. Therefore, HDG methods have a significantly smaller number of degrees of freedom in the global system compared to DG methods, LDG methods or mixed FEM. Moreover, HDG methods keep the advantages of DG methods, which are suitable for the drift term if $\nabla\phi$ is large. For more information about HDG methods for convection diffusion problems; see, e.g., [17–19,33,59].

There are many different HDG schemes, see for example [20–25,48]. Among all of these methods, two are most popular, following standard terminology we call them are HDG_k and HDG(A) in the rest of the paper. The HDG_k method uses polynomials of degree k to approximate the solution, the flux, and the trace on the interfaces between elements together with a positive stabilization parameter chosen to be $\mathcal{O}(1)$. The HDG(A) method uses polynomial degree $k+1$ to approximate the solution, polynomial degree k to approximate the flux and uses the so called Lehrenfeld-Schöberl stabilization function, see [48, Remark 1.2.4]. These two methods were used to study the Poisson equation in [27,49,57], the linear elasticity [22,58], the convection diffusion equation in [18,19,59], the Stokes equation in [26,38] and the Navier–Stokes equation in [11,60].

The goal of this paper is to design an HDG scheme by the appropriate choice of HDG spaces such that the overall scheme is optimally convergent and to prove semi-discrete optimal convergence rates in d spatial dimensions ($d = 2, 3$). The result is a new HDG scheme for the drift–diffusion system with attractive convergence properties. We shall assume that a suitably regular solution of the drift–diffusion system exists. For existence theory, see for example the book of Markowich [53].

To develop our HDG method, we write the drift–diffusion system as a first order system by introducing new variable $\boldsymbol{q}$ and $\boldsymbol{p}$ such that $\boldsymbol{q} + \nabla u = 0$, $\boldsymbol{p} + \nabla\phi = 0$. Then (1), becomes the problem of finding $(u, \boldsymbol{q}, \phi, \boldsymbol{p})$ such that

$$\boldsymbol{q} + \nabla u = 0 \quad \text{in } \Omega \times (0, T], \tag{2a}$$

$$\boldsymbol{p} + \nabla\phi = 0 \quad \text{in } \Omega \times (0, T], \tag{2b}$$

$$u_t + \nabla\cdot\boldsymbol{q} - \nabla\cdot(\boldsymbol{p}u) = 0 \quad \text{in } \Omega \times (0, T], \tag{2c}$$

$$\varepsilon\nabla\cdot\boldsymbol{p} + u = 0 \quad \text{in } \Omega \times (0, T], \tag{2d}$$

$$u = g_u \quad \text{on } \partial\Omega \times (0, T], \tag{2e}$$

$$\phi = g_\phi \quad \text{on } \partial\Omega \times (0, T], \tag{2f}$$

$$u(\cdot, 0) = u_0 \quad \text{in } \Omega. \tag{2g}$$

In this work, we only

We can now introduce our HDG formulation by first defining the mesh. Let $\mathcal{T}_h$ denote a collection of disjoint simplexes K that partition Ω and let $\partial\mathcal{T}_h$ be the set $\{\partial K : K \in \mathcal{T}_h\}$. Here h denotes the maximum diameter of the simplices in $\mathcal{T}_h$. Since we will need to use an inverse inequality in our analysis, we assume that the mesh is shape regular and quasi-uniform.

We denote by $\mathcal{E}_h$ the set of all faces (or edges in when $d = 2$) in the mesh. Then we define the set of interior and boundary faces or edges denoted $\mathcal{E}_h^o$ and $\mathcal{E}_h^\partial$ respectively. From now on, to simplify terminology, we shall refer to elements of $\mathcal{E}_h$ as faces, even if $d = 2$. For each face e we say $e \in \mathcal{E}_h^o$ is an interior face if the Lebesgue measure of $e = \partial K^+ \cap \partial K^-$ for some pair of elements $K^+, K^- \in \mathcal{T}_h$ is non-zero, similarly, $e \in \mathcal{E}_h^\partial$ is a boundary face if the Lebesgue measure of $e = \partial K \cap \partial\Omega$ is non-zero. We 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 $(\cdot, \cdot)_K$ denotes the $L^2(K)$ inner product and $\langle\cdot, \cdot\rangle_{\partial K}$ denotes the L^2 inner product on ∂K.

The HDG method uses discontinuous finite element spaces $\boldsymbol{Q}_h$, V_h, $\widehat{V}_h$, $\boldsymbol{S}_h$, Ψ_h, $\widehat{\Psi}_h$ that we shall discuss shortly. Assuming these are given, the HDG method seeks $(\boldsymbol{q}_h, u_h, \widehat{u}_h) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(g_u)$ and $(\boldsymbol{p}_h, \phi_h, \widehat{\phi}_h) \in \boldsymbol{S}_h \times \Psi_h \times \widehat{\Psi}_h(g_\phi)$ satisfying

$$(\boldsymbol{q}_h, \boldsymbol{r}_1)_{\mathcal{T}_h} - (u_h, \nabla \cdot \boldsymbol{r}_1)_{\mathcal{T}_h} + \langle \widehat{u}_h, \boldsymbol{r}_1 \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0, \tag{3a}$$

$$(\boldsymbol{p}_h, \boldsymbol{r}_2)_{\mathcal{T}_h} - (\phi_h, \nabla \cdot \boldsymbol{r}_2)_{\mathcal{T}_h} + \langle \widehat{\phi}_h, \boldsymbol{r}_1 \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} = 0, \tag{3b}$$

for all $(\boldsymbol{r}_1, \boldsymbol{r}_2) \in \boldsymbol{Q}_h \times \boldsymbol{S}_h$, together with

$$\begin{aligned} &(u_{h,t}, w_1)_{\mathcal{T}_h} - (\boldsymbol{q}_h, \nabla w_1)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, w_1 \rangle_{\partial \mathcal{T}_h} + (\boldsymbol{p}_h u_h, \nabla w_1)_{\mathcal{T}_h} \\ &\quad - \langle \widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n} \widehat{u}_h, w_1 \rangle_{\partial \mathcal{T}_h} = 0, \end{aligned} \tag{3c}$$

$$- (\boldsymbol{p}_h, \nabla w_2)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n}, w_2 \rangle_{\partial \mathcal{T}_h} + (u_h, w_2)_{\mathcal{T}_h} = 0 \tag{3d}$$

for all $(w_1, w_2) \in V_h \times \Psi_h$. The boundary fluxes must satisfy

$$\langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, \mu_1 \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} = 0, \tag{3e}$$

$$\langle \widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n}, \mu_2 \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} = 0 \tag{3f}$$

for all $(\mu_1, \mu_2) \in \widehat{V}_h(0) \times \widehat{\Psi}_h(0)$. The numerical fluxes $\widehat{\boldsymbol{q}}_h$ and $\widehat{\boldsymbol{p}}_h$ will be specified later.

As in [11,51], we shall need the following energy estimate

$$\begin{aligned} &\|\nabla u_h\|_{\mathcal{T}_h} + \|h_K^{-1/2}(u_h - \widehat{u}_h)\|^2_{\partial \mathcal{T}_h} \\ &\quad \le C \left(\|\boldsymbol{q}_h\|^2_{\mathcal{T}_h} + \|h_K^{-1/2}(\Pi_k^\partial u_h - \widehat{u}_h)\|^2_{\partial \mathcal{T}_h} \right). \end{aligned} \tag{4}$$

where Π_k^∂ is the L^2 projection defined in (12). Inequality (4) cannot hold for the HDG$_k$ method unless we take the stabilization function to be h_K^{-1}. However, in this case we only have a suboptimal convergence rate for the flux $\boldsymbol{q}$. Hence we need to use the HDG(A) method to approximate the Eq. (1a), i.e., we choose

$$\begin{aligned} \boldsymbol{Q}_h &:= \{\boldsymbol{v}_h \in [L^2(\Omega)]^d : \boldsymbol{v}_h|_K \in [\mathcal{P}^k(K)]^d, \quad \forall K \in \mathcal{T}_h\}, \\ V_h &:= \{v_h \in L^2(\Omega) : v_h|_K \in \mathcal{P}^{k+1}(K), \quad \forall K \in \mathcal{T}_h\}, \\ \widehat{V}_h(g) &:= \{\widehat{v}_h \in L^2(\mathcal{E}_h) : \widehat{v}_h|_e \in \mathcal{P}^k(e), \quad \forall e \in \mathcal{E}_h, \widehat{v}_h|_{\mathcal{E}_h^\partial} = \Pi_k^\partial g\}, \end{aligned}$$

where $\mathcal{P}^k(K)$ denotes the set of polynomials of degree at most k on the element K (similarly $\mathcal{P}^k(\mathcal{E}_h)$ denotes the set of polynomials of degree at most k on the faces in the mesh). Moreover, the numerical trace of the flux on $\partial \mathcal{T}_h$ is defined as

$$\widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n} = \boldsymbol{q}_h \cdot \boldsymbol{n} + h_K^{-1}(\Pi_k^\partial u_h - \widehat{u}_h), \tag{5}$$

where Π_k^∂ denotes L^2 projection onto $\mathcal{P}^k(\mathcal{E}_h)$ which can be done face by face.

To avoid a reduction in the convergence rate for the solution u_h, the polynomial degree of the space V_h for u_h and the space $\boldsymbol{S}_h$ for $\boldsymbol{p}_h$ need to be the same, i.e.,

$$\boldsymbol{S}_h := \{\boldsymbol{v}_h \in [L^2(\Omega)]^d : \boldsymbol{v}_h|_K \in [\mathcal{P}^{k+1}(K)]^d, \quad \forall K \in \mathcal{T}_h\}.$$

If we choose the HDG(A) method to discretize (1b) we would need to use polynomials of degree $k+2$ to approximate ϕ, but in this case, we get a suboptimal convergence rate for ϕ. Therefore, we use HDG$_{k+1}$ to discretize (1b) and so choose

$$\Psi_h := \{v_h \in L^2(\Omega) : v_h|_K \in \mathcal{P}^{k+1}(K), \quad \forall K \in \mathcal{T}_h\},$$
$$\widehat{\Psi}_h(g) := \{\widehat{v}_h \in L^2(\mathcal{E}_h) : \widehat{v}_h|_e \in \mathcal{P}^{k+1}(e), \quad \forall e \in \mathcal{E}_h, \widehat{v}_h|_{\mathcal{E}_h^\partial} = \Pi_{k+1}^\partial g\}.$$

and the numerical trace of the flux on $\partial\mathcal{T}_h$ is defined as

$$\widehat{\boldsymbol{p}}_h \cdot \boldsymbol{n} = \boldsymbol{p}_h \cdot \boldsymbol{n} + \tau(\phi_h - \widehat{\phi}_h), \tag{6}$$

where τ is a positive $\mathcal{O}(1)$ function and the initial condition $u_h(0)$ will be specified in Sect. 3.1. If needed, τ can be chosen to provide upwind stabilization as in [59].

The organization of the paper is as follows. In Sect. 2, we present our main results and some useful projections. Then the proof of the main results is given in Sect. 3. In Sect. 4, we provide some numerical experiments to support our theoretical results.

2 Main Result and Preliminary Material

In this section, we first present the main result in Sect. 2.1 for the semidiscrete HDG formulation (3). Next, we provide preliminary material in Sect. 2.2, which are required for the analysis.

We use the standard notation $W^{m,p}(D)$ for Sobolev spaces on D with norm $\|\cdot\|_{m,p,D}$ and seminorm $|\cdot|m, p, D$. We also write $H^m(D)$ instead of $W^{m,2}(D)$, and we omit the index p in the corresponding norms and seminorms. Moreover, we omit the index m when $m = 0$.

Throughout, we assume the data and the solution of (1) are smooth enough for our analysis.

2.1 Main Result

The proof of our main error estimate relies on the use of duality arguments and requires sufficient regularity for the solution of the corresponding "adjoint"problem. In particular:

Assumption 2.1 Assume that the component $\boldsymbol{p}$ of the solution of (2) is such that $\boldsymbol{p} \in H^1((0,T), \boldsymbol{W}_1^\infty(\Omega))$. Let $M > 0$ be such that for all time $t \in (0,T)$

$$M \geq \|\nabla \cdot \boldsymbol{p}(t)\|_{0,\infty} + 2\|\partial_t \boldsymbol{p}(t)\|_{0,\infty}. \tag{7}$$

If $\boldsymbol{p} = 0$, set $M = 0$. Then, for $\Theta \in L^2(\Omega \times (0,T))$, let $(\boldsymbol{\Phi}, \Psi)$ be the solution of

$$\begin{aligned} \boldsymbol{\Phi} + \nabla\Psi &= 0 \quad \text{in } \Omega, \\ M\Phi + \nabla \cdot \boldsymbol{\Phi} + \boldsymbol{p} \cdot \nabla\Psi &= \Theta \quad \text{in } \Omega, \\ \Psi &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{8}$$

We assume the so lution $(\boldsymbol{\Psi}, \Phi)$ has the following regularity

$$\|\boldsymbol{\Phi}\|_{H^1(\Omega)} + \|\Psi\|_{H^2(\Omega)} \leq C_{\text{reg}}\|\Theta\|_{\mathcal{T}_h}. \tag{9}$$

Remark 2.2 It is well known that the above regularity holds if the domain is convex, which is usually the case in solar cell applications.

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

Theorem 2.3 *Suppose that Assumption 2.1 holds and that the mesh is quasi-uniform. Assume in addition that*

$$(\boldsymbol{q}, u) \in H^2((0,T), \boldsymbol{H}^{k+1}(\Omega)) \times H^2((0,T), H^{k+2}(\Omega)),$$
$$(\boldsymbol{p}, \phi) \in H^2((0,T), \boldsymbol{H}^{k+2}(\Omega)) \times H^2((0,T), H^{k+3}(\Omega))$$

for $k \geq 0$ *solve* (2). *Let* $(\boldsymbol{q}_h, u_h, \boldsymbol{p}_h, \phi_h) \in \boldsymbol{Q}_h \times V_h \times \boldsymbol{S}_h \times \Psi_h$ *be the solution of the semi-discrete HDG equations* (3). *Then we have*

$$\|u - u_h\|_{\mathcal{T}_h} + \|\phi - \phi_h\|_{\mathcal{T}_h} + \|\boldsymbol{p} - \boldsymbol{p}_h\|_{\mathcal{T}_h} \leq Ch^{k+2}$$

for all $t \in [0, T]$, *and*

$$\sqrt{\int_0^T \|\boldsymbol{q} - \boldsymbol{q}_h\|^2_{\mathcal{T}_h} dt} \leq Ch^{k+1}.$$

Remark 2.4 The error estimates in Theorem 2.3 are optimal for the variables $\boldsymbol{q}$, u, $\boldsymbol{p}$ and ϕ. Since the global degrees of freedom are the numerical traces, then from the point of view of global degrees of freedom, the error estimates for the variable u is superconvergent, which, to our knowledge, is the first time this has been proved in the literature.

2.2 Preliminary Material

We first introduce the HDG$_k$ projection operator $\Pi_h(\boldsymbol{p}, \phi) := (\boldsymbol{\Pi}_V \boldsymbol{p}, \Pi_W \phi)$ defined in [27], where $\boldsymbol{\Pi}_V \boldsymbol{p}$ and $\Pi_W \phi$ denote components of the projection of $\boldsymbol{p}$ and ϕ into $\boldsymbol{S}_h$ and Ψ_h, respectively. For each element $K \in \mathcal{T}_h$, the projection is determined by the equations

$$(\boldsymbol{\Pi}_V \boldsymbol{p}, \boldsymbol{r})_K = (\boldsymbol{p}, \boldsymbol{r})_K, \quad \forall \boldsymbol{r} \in [\mathcal{P}_k(K)]^d, \tag{10a}$$
$$(\Pi_W \phi, w)_K = (\phi, w)_K, \quad \forall w \in \mathcal{P}_k(K), \tag{10b}$$
$$\langle \boldsymbol{\Pi}_V \boldsymbol{p} \cdot \boldsymbol{n} + \tau \Pi_W \phi, \mu \rangle_e = \langle \boldsymbol{p} \cdot \boldsymbol{n} + \tau \phi, \mu \rangle_e, \quad \forall \mu \in \mathcal{P}_{k+1}(e) \tag{10c}$$

for all faces e of the simplex K. The approximation properties of the HDG$_k$ projection (10) are given in the following result from [27]:

Lemma 2.5 *Suppose* $k \geq 0$, $\tau|_{\partial K}$ *is nonnegative and* $\tau_K^{\max} := \max \tau|_{\partial K} > 0$. *Then the system* (10) *is uniquely solvable for* $\boldsymbol{\Pi}_V \boldsymbol{p}$ *and* $\Pi_W \phi$. *Furthermore, there is a constant* C *independent of* K *and* τ *such that*

$$\|\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}\|_K \leq Ch_K^{\ell_{\boldsymbol{p}}+1} |\boldsymbol{p}|_{\ell_{\boldsymbol{p}}+1,K} + Ch_K^{\ell_\phi+1} \tau_K^* |\phi|_{\ell_\phi+1,K}, \tag{11a}$$
$$\|\Pi_W \phi - \phi\|_K \leq Ch_K^{\ell_\phi+1} |\phi|_{\ell_\phi+1,K} + C\frac{h_K^{\ell_{\boldsymbol{p}}+1}}{\tau_K^{\max}} |\nabla \cdot \boldsymbol{p}|_{\ell_{\boldsymbol{p}},K} \tag{11b}$$

for $\ell_{\boldsymbol{p}}, \ell_\phi$ *in* $[0, k+1]$. *Here* $\tau_K^* := \max \tau|_{\partial K \setminus e^*}$, *where* e^* *is a face of* K *at which* $\tau|_{\partial K}$ *is maximum.*

We next define the standard L^2 projections $\boldsymbol{\Pi}_k^o : [L^2(\Omega)]^d \to \boldsymbol{Q}_h$, $\Pi_{k+1}^o : L^2(\Omega) \to V_h$, and $\Pi_k^\partial : L^2(\mathcal{E}_h) \to \widehat{V}_h$, which satisfy

$$\begin{aligned} (\boldsymbol{\Pi}_k^o \boldsymbol{q}, \boldsymbol{r}_1)_K &= (\boldsymbol{q}, \boldsymbol{r}_1)_K, \quad \forall \boldsymbol{r}_1 \in [\mathcal{P}_k(K)]^d, \\ (\Pi_{k+1}^o u, w_1)_K &= (u, w_1)_K, \quad \forall w_1 \in \mathcal{P}_{k+1}(K), \\ \langle \Pi_k^\partial u, \mu_1 \rangle_e &= \langle u, \mu_1 \rangle_e, \quad \forall \mu_1 \in \mathcal{P}_k(e). \end{aligned} \tag{12}$$

In the analysis, we use the following classical results [16, Lemma 3.3]:

$$\|\boldsymbol{q} - \boldsymbol{\Pi}_k^o \boldsymbol{q}\|_{\mathcal{T}_h} \le C h^{k+1} \|\boldsymbol{q}\|_{k+1,\Omega}, \quad \|u - \Pi_{k+1}^o u\|_{\mathcal{T}_h} \le C h^{k+2} \|u\|_{k+2,\Omega}, \tag{13a}$$

$$\|u - \Pi_{k+1}^o u\|_{\partial\mathcal{T}_h} \le C h^{k+\frac{3}{2}} \|u\|_{k+2,\Omega}, \quad \|w\|_{\partial\mathcal{T}_h} \le C h^{-\frac{1}{2}} \|w\|_{\mathcal{T}_h}, \quad \forall w \in V_h. \tag{13b}$$

To shorten lengthy equations, we rewrite the HDG formulation (3) in the following compact form: find $(\boldsymbol{q}_h, u_h, \widehat{u}_h) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(g_u)$ and $(\boldsymbol{p}_h, \phi_h, \widehat{\phi}_h) \in \boldsymbol{S}_h \times \Psi_h \times \widehat{\Psi}_h(g_\phi)$ such that

$$(\partial_t u_h, w_1)_{\mathcal{T}_h} + \mathcal{A}(\boldsymbol{q}_h, u_h, \widehat{u}_h; \boldsymbol{r}_1, w_1, \mu_1) + \mathcal{C}(\boldsymbol{p}_h, \widehat{\boldsymbol{p}}_h; u_h, \widehat{u}_h; w_1) = 0, \tag{14a}$$

$$\mathcal{B}(\boldsymbol{p}_h, \phi_h, \widehat{\phi}_h; \boldsymbol{r}_2, w_2, \mu_2) + (u_h, w_2)_{\mathcal{T}_h} = 0, \tag{14b}$$

for all $(\boldsymbol{r}_1, \boldsymbol{r}_2, w_1, w_2, \mu_1, \mu_2) \in \boldsymbol{Q}_h \times \boldsymbol{S}_h \times V_h \times \Psi_h \times \widehat{V}_h(0) \times \widehat{\Psi}_h(0)$, where the HDG bilinear forms $\mathcal{A}$, $\mathcal{B}$ and the trilinear form $\mathcal{C}$ are defined by

$$\begin{aligned}
&\mathcal{A}(\boldsymbol{q}_h, u_h, \widehat{u}_h; \boldsymbol{r}_1, w_1, \mu_1) \\
&\quad = (\boldsymbol{q}_h, \boldsymbol{r}_1)_{\mathcal{T}_h} - (u_h, \nabla \cdot \boldsymbol{r}_1)_{\mathcal{T}_h} + \langle \widehat{u}_h, \boldsymbol{r}_1 \cdot \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} + (\nabla \cdot \boldsymbol{q}_h, w_1)_{\mathcal{T}_h} \\
&\qquad - \langle \boldsymbol{q}_h \cdot \boldsymbol{n}, \mu_1 \rangle_{\partial\mathcal{T}_h} + \langle h_K^{-1} (\Pi_k^\partial u_h - \widehat{u}_h), \Pi_k^\partial w_1 - \mu_1 \rangle_{\partial\mathcal{T}_h}
\end{aligned} \tag{14c}$$

for all $(\boldsymbol{q}_h, u_h, \widehat{u}_h, \boldsymbol{r}_1, w_1, \mu_1) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(g_u) \times \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(0)$,

$$\begin{aligned}
&\mathcal{B}(\boldsymbol{p}_h, \phi_h, \widehat{\phi}_h; \boldsymbol{r}_2, w_2, \mu_2) \\
&\quad = (\boldsymbol{p}_h, \boldsymbol{r}_2)_{\mathcal{T}_h} - (\phi_h, \nabla \cdot \boldsymbol{r}_2)_{\mathcal{T}_h} + \langle \widehat{\phi}_h, \boldsymbol{r}_2 \cdot \boldsymbol{n} \rangle_{\partial\mathcal{T}_h} + (\nabla \cdot \boldsymbol{p}_h, w_2)_{\mathcal{T}_h} \\
&\qquad - \langle \boldsymbol{p}_h \cdot \boldsymbol{n}, \mu_2 \rangle_{\partial\mathcal{T}_h} + \langle \tau (\phi_h - \widehat{\phi}_h), w_2 - \mu_2 \rangle_{\partial\mathcal{T}_h}
\end{aligned} \tag{14d}$$

for all $(\boldsymbol{p}_h, \phi_h, \widehat{\phi}_h, \boldsymbol{r}_2, w_2, \mu_2) \in \boldsymbol{S}_h \times \Psi_h \times \widehat{\Psi}_h(g_\phi) \times \boldsymbol{S}_h \times \Psi_h \times \widehat{\Psi}_h(0)$,

$$\mathcal{C}(\boldsymbol{p}, \widehat{\boldsymbol{p}}; u_h, \widehat{u}_h; w_1) = (\boldsymbol{p} u_h, \nabla w_1)_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{p}} \cdot \boldsymbol{n} \widehat{u}_h, w_1 \rangle_{\partial\mathcal{T}_h} \tag{14e}$$

for all $(u_h, \widehat{u}_h, w_1, \mu_1) \in V_h \times \widehat{V}_h(g_u) \times V_h \times \widehat{V}_h(0)$.

Next, we present basic properties of the operators $\mathcal{A}$ and $\mathcal{B}$.

Lemma 2.6 *For any* $(\boldsymbol{q}_h, u_h, \widehat{u}_h, \boldsymbol{r}_1, w_1, \mu_1) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(0) \times \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(0)$ *and* $(\boldsymbol{p}_h, \phi_h, \widehat{\phi}_h, \boldsymbol{r}_2, w_2, \mu_2) \in \boldsymbol{S}_h \times \Psi_h \times \widehat{\Psi}_h(0) \times \boldsymbol{S}_h \times \Psi_h \times \widehat{\Psi}_h(0)$, *we have*

$$\begin{aligned}
\mathcal{A}(\boldsymbol{q}_h, u_h, \widehat{u}_h; -\boldsymbol{r}_1, w_1, \mu_1) &= \mathcal{A}(\boldsymbol{r}_1, w_1, \mu_1; -\boldsymbol{q}_h, u_h, \widehat{u}_h), \\
\mathcal{B}(\boldsymbol{p}_h, \phi_h, \widehat{\phi}_h; -\boldsymbol{r}_2, w_2, \mu_2) &= \mathcal{B}(\boldsymbol{r}_2, w_2, \mu_2; -\boldsymbol{p}_h, \phi_h, \widehat{\phi}_h),
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{A}(\boldsymbol{q}_h, u_h, \widehat{u}_h; \boldsymbol{q}_h, u_h, \widehat{u}_h) &= \|\boldsymbol{q}_h\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2} (\Pi_k^\partial u_h - \widehat{u}_h)\|_{\partial\mathcal{T}_h}^2, \\
\mathcal{B}(\boldsymbol{p}_h, \phi_h, \widehat{\phi}_h; \boldsymbol{p}_h, \phi_h, \widehat{\phi}_h) &= \|\boldsymbol{p}_h\|_{\mathcal{T}_h}^2 + \|\sqrt{\tau} (\phi_h - \widehat{\phi}_h)\|_{\partial\mathcal{T}_h}^2.
\end{aligned}$$

The proof of Lemma 2.6 is straightforward, hence we omit it here.

The proof of the following two lemmas are found in [59, Lemma 3.2] and [8, Equation (1.3)], respectively.

Lemma 2.7 *If* $(\boldsymbol{q}_h, u_h, \widehat{u}_h)$ *satisfies the Eq.* (3a)*, then we have*

$$\|\nabla u_h\|_{\mathcal{T}_h} + \|h_K^{-1/2}(u_h - \widehat{u}_h)\|_{\partial\mathcal{T}_h} \le C\left(\|\boldsymbol{q}_h\|_{\mathcal{T}_h} + \|h_K^{-1/2}(\Pi_k^\partial u_h - \widehat{u}_h)\|_{\partial\mathcal{T}_h}\right).$$

Lemma 2.8 (Piecewise Poincáre-Friedrichs' inequality) *Let* $v_h \in H^1(\mathcal{T}_h)$*, then we have*

$$\|v_h\|_{\mathcal{T}_h}^2 \le C\left(\|\nabla v_h\|_{\mathcal{T}_h}^2 + |\langle v_h, 1\rangle_{\partial\Omega}|^2 + \sum_{e\in\mathcal{E}_h^o} |e|^{d/(1-d)}\left(\int_e [\![v_h]\!]\, ds\right)^2\right),$$

where $|e|$ *denotes the measure of* e.

By Lemma 2.8, we immediately have the following lemma.

Lemma 2.9 (HDG Poinćare inequality) *If* $(v_h, \widehat{v}_h) \in V_h \times \widehat{V}_h(0)$*, then we have*

$$\|v_h\|_{\mathcal{T}_h}^2 \le C\left(\|\nabla v_h\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2}(\Pi_k^\partial v_h - \widehat{v}_h)\|_{\partial\mathcal{T}_h}^2\right).$$

Proof By Lemma 2.9, $\widehat{v}_h$ is zero on $\partial\Omega$ and is single valued on interior faces. We have

$$\begin{aligned}
\|v_h\|_{\mathcal{T}_h}^2 &\le C\left(\|\nabla v_h\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2}[\![v_h]\!]\|_{\mathcal{E}_h}^2\right)\\
&= C\left(\|\nabla v_h\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2}[\![v_h - \Pi_k^\partial v_h + \Pi_k^\partial v_h - \widehat{v}_h]\!]\|_{\mathcal{E}_h}^2\right)\\
&\le C\left(\|\nabla v_h\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2}(v_h - \Pi_k^\partial v_h)\|_{\partial\mathcal{T}_h}^2 + \|h_K^{-1/2}(\Pi_k^\partial v_h - \widehat{v}_h)\|_{\partial\mathcal{T}_h}^2\right)\\
&\le C\left(\|\nabla v_h\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2}(\Pi_k^\partial v_h - \widehat{v}_h)\|_{\partial\mathcal{T}_h}^2\right).
\end{aligned}$$

□

3 Proof of Theorem 2.3

To prove Theorem 2.3, we follow a similar strategy to that in [15]. We first bound the error between the solution of an HDG elliptic projection defined in (15) and the solution of the system (1a). Then we bound the error between the solution of the HDG elliptic projection (15) and the HDG formulation (14a) and the error between the solution of the system (1b) and the solution of the HDG formulation (14b). A simple application of the triangle inequality then gives a bound on the error between the solution of the HDG formulation (14) and the system (1). First, we present the HDG elliptic projection.

3.1 HDG Elliptic Projection and Basic Estimates

For $t \in [0, T]$, let $(\boldsymbol{q}_{Ih}, u_{Ih}, \widehat{u}_{Ih}) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(g_u)$ be the solution of

$$\begin{aligned}
&M(u_{Ih}, w_1)_{\mathcal{T}_h} + \mathcal{A}(\boldsymbol{q}_{Ih}, u_{Ih}, \widehat{u}_{Ih}; r_1, w_1, \mu_1) + \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; u_{Ih}, \widehat{u}_{Ih}; w_1)\\
&\quad = (Mu - u_t, w_1)_{\mathcal{T}_h}
\end{aligned} \tag{15}$$

for all $(r_1, w_1, \mu_1) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(0)$ where M is a given constant such that (7) holds.

Take the partial derivative of (15) with respect to t, hence, $(\partial_t \boldsymbol{q}_{Ih}, \partial_t u_{Ih}, \partial_t \widehat{u}_{Ih}) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(\partial_t g_u)$ is the solution of

$$
\begin{aligned}
&M(\partial_t u_{Ih}, w_1)_{\mathcal{T}_h} + \mathcal{A}(\boldsymbol{\partial_t q}_{Ih}, \partial_t u_{Ih}, \partial_t \widehat{u}_{Ih}; r_1, w_1, \mu_1) \\
&\quad + \mathcal{C}(\partial_t \boldsymbol{p}, \partial_t \boldsymbol{p}; u_{Ih}, \widehat{u}_{Ih}; w_1) + \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \partial_t u_{Ih}, \partial_t \widehat{u}_{Ih}; w_1) \\
&= (M u_t - u_{tt}, w_1)_{\mathcal{T}_h}
\end{aligned} \tag{16}
$$

for all $(r_1, w_1, \mu_1) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(0)$.

We choose the initial condition $u_h(0) = u_{Ih}(0)$ for the purposes of analysis. In fact, the initial condition $u_h(0)$ can be chosen to be the L^2 projection of u_0, i.e., $\Pi_k^o u_0$.

The following result, Theorem 3.1, gives the error between the solution of an HDG elliptic projection (15) and the solution of the system (1a) and the proofs are given in Sect. 6.

Theorem 3.1 *For any $t \in [0, T]$, if the elliptic regularity inequality (9) holds and h is small enough, then we have the following error estimates*

$$
\|u - u_{Ih}\|_{\mathcal{T}_h} \le C h^{k+2} \|u\|_{k+2}, \tag{17a}
$$

$$
\|\boldsymbol{q} - \boldsymbol{q}_{Ih}\|_{\mathcal{T}_h} + \|h_K^{-1/2}(\Pi_k^\partial u_{Ih} - \widehat{u}_{Ih})\|_{\partial \mathcal{T}_h} \le C h^{k+1} \|u\|_{k+2}. \tag{17b}
$$

In addition, we have

$$
\|\partial_t u - \partial_t u_{Ih}\|_{\mathcal{T}_h} \le C h^{k+2} \|\partial_t u\|_{k+2}. \tag{17c}
$$

3.2 Error Equation Between the HDG Formulation (14) and the HDG Elliptic Projection (15)

To bound the error between the solution of the HDG elliptic projection (15) and the system (14a), and the error between the solution of the HDG formulation (14b) and the system (1b). We first derive the error equation summarized in the next lemma. To simplify notation, we define

$$
\begin{aligned}
&\xi_h^{\boldsymbol{q}} = \boldsymbol{q}_{Ih} - \boldsymbol{q}_h, \quad \xi_h^u = u_{Ih} - u_h, \quad \xi_h^{\widehat{u}} = \widehat{u}_{Ih} - \widehat{u}_h, \\
&\xi_h^{\boldsymbol{p}} = \boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}_h, \quad \xi_h^\phi = \Pi_W \phi - \phi_h, \quad \xi_h^{\widehat{\phi}} = \Pi_{k+1}^\partial \phi - \widehat{\phi}_h.
\end{aligned}
$$

Lemma 3.2 *For any $(\boldsymbol{r}_1, w_1, \mu_1, \boldsymbol{r}_2, w_2, \mu_2) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(0) \times \boldsymbol{S}_h \times \Psi_h \times \widehat{\Psi}_h(0)$, we have the following error equation*

$$
\begin{aligned}
&(\partial_t \xi_h^u, w_1)_{\mathcal{T}_h} + \mathcal{A}(\xi_h^{\boldsymbol{q}}, \xi_h^u, \xi_h^{\widehat{u}}; \boldsymbol{r}_1, w_1, \mu_1) \\
&\quad = (\partial_t (u_{Ih} - u), w_1)_{\mathcal{T}_h} + M(u - u_{Ih}, w_1)_{\mathcal{T}_h} \\
&\qquad - \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \xi_h^u, \xi_h^{\widehat{u}}; w_1) - \mathcal{C}(\boldsymbol{p} - \boldsymbol{p}_h, \boldsymbol{p} - \widehat{\boldsymbol{p}}_h; u_h, \widehat{u}_h; w_1),
\end{aligned} \tag{18a}
$$

$$
\mathcal{B}(\xi_h^{\boldsymbol{p}}, \xi_h^\phi, \xi_h^{\widehat{\phi}}; \boldsymbol{r}_2, w_2, \mu_2) = (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{r}_2)_{\mathcal{T}_h} - (u - u_h, w_2)_{\mathcal{T}_h}. \tag{18b}
$$

Proof We first prove (18a). Subtracting Eq. (14a) from (15) and using the definition of $\mathcal{A}$ and $\mathcal{C}$ we get

$$
\begin{aligned}
&M(u_{Ih}, w_1)_{\mathcal{T}_h} + \mathcal{A}(\xi_h^{\boldsymbol{q}}, \xi_h^u, \xi_h^{\widehat{u}}; \boldsymbol{r}_1, w_1, \mu_1) + \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; u_{Ih}, \widehat{u}_{Ih}; w_1) \\
&\quad - (\partial_t u_h, w_1)_{\mathcal{T}_h} - \mathcal{C}(\boldsymbol{p}_h, \widehat{\boldsymbol{p}}_h; u_h, \widehat{u}_h; w_1) \\
&\quad = (M u - u_t, w_1)_{\mathcal{T}_h}.
\end{aligned}
$$

This gives

$$
\begin{aligned}
&(\partial_t \xi_h^u, w_1)_{\mathcal{T}_h} + \mathcal{A}(\xi_h^{\boldsymbol{q}}, \xi_h^u, \xi_h^{\widehat{u}}; \boldsymbol{r}_1, w_1, \mu_1) \\
&\quad + \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; u_h, \widehat{u}_h; w_1) - \mathcal{C}(\boldsymbol{p}_h, \widehat{\boldsymbol{p}}_h; u_{Ih}, \widehat{u}_{Ih}; w_1) \\
&\quad = (\partial_t u_{Ih}, w_1)_{\mathcal{T}_h} - (u_t, w_1)_{\mathcal{T}_h} + M(u - u_{Ih}, w_1)_{\mathcal{T}_h}.
\end{aligned}
$$

We note that the nonlinear operator $\mathcal{C}$ is linear for each variables, hence we have

$$
\begin{aligned}
&\mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; u_{Ih}, \widehat{u}_{Ih}; w_1) - \mathcal{C}(\boldsymbol{p}_h, \widehat{\boldsymbol{p}}_h; u_h, \widehat{u}_h; w_1) \\
&\quad = \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; u_{Ih}, \widehat{u}_{Ih}; w_1) - \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; u_h, \widehat{u}_h; w_1) \\
&\qquad + \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; u_h, \widehat{u}_h; w_1) - \mathcal{C}(\boldsymbol{p}_h, \widehat{\boldsymbol{p}}_h; u_h, \widehat{u}_h; w_1) \\
&\quad = \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \xi_h^u, \xi_h^{\widehat{u}}; w_1) + \mathcal{C}(\boldsymbol{p} - \boldsymbol{p}_h, \boldsymbol{p} - \widehat{\boldsymbol{p}}_h; u_h, \widehat{u}_h; w_1).
\end{aligned}
$$

This implies

$$
\begin{aligned}
&(\partial_t \xi_h^u, w_1)_{\mathcal{T}_h} + \mathcal{A}(\xi_h^{\boldsymbol{q}}, \xi_h^u, \xi_h^{\widehat{u}}; \boldsymbol{r}_1, w_1, \mu_1) \\
&\quad = (\partial_t (u_{Ih} - u), w_1)_{\mathcal{T}_h} + M(u - u_{Ih}, w_1)_{\mathcal{T}_h} \\
&\qquad - \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \xi_h^u, \xi_h^{\widehat{u}}; w_1) - \mathcal{C}(\boldsymbol{p} - \boldsymbol{p}_h, \boldsymbol{p} - \widehat{\boldsymbol{p}}_h; u_h, \widehat{u}_h; w_1).
\end{aligned}
$$

Next, we prove (18b). By the definition of $\mathcal{B}$ in (14d), we have

$$
\begin{aligned}
&\mathcal{B}(\boldsymbol{\Pi}_V \boldsymbol{p}, \Pi_W \phi, \Pi_{k+1}^{\partial} \phi; \boldsymbol{r}_2, w_2, \mu_2) \\
&\quad = (\boldsymbol{\Pi}_V \boldsymbol{p}, \boldsymbol{r}_2)_{\mathcal{T}_h} - (\Pi_W \phi, \nabla \cdot \boldsymbol{r}_2)_{\mathcal{T}_h} + \langle \Pi_{k+1}^{\partial} \phi, \boldsymbol{r}_2 \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
&\qquad + (\nabla \cdot \boldsymbol{\Pi}_V \boldsymbol{p}, w_2)_{\mathcal{T}_h} - \langle \boldsymbol{\Pi}_V \boldsymbol{p} \cdot \boldsymbol{n}, \mu_2 \rangle_{\partial \mathcal{T}_h} \\
&\qquad + \langle \tau(\Pi_W \phi - \Pi_{k+1}^{\partial} \phi), w_2 - \mu_2 \rangle_{\partial \mathcal{T}_h} - \langle \boldsymbol{p} \cdot \boldsymbol{n}, \mu_2 \rangle_{\partial \mathcal{T}_h}.
\end{aligned}
$$

By the definition of $\boldsymbol{\Pi}_V$ and Π_W in (10) we get

$$
\begin{aligned}
&\mathcal{B}(\boldsymbol{\Pi}_V \boldsymbol{p}, \Pi_W \phi, \Pi_{k+1}^{\partial} \phi; \boldsymbol{r}_2, w_2, \mu_2) \\
&\quad = (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{r}_2)_{\mathcal{T}_h} + (\boldsymbol{p}, \boldsymbol{r}_2)_{\mathcal{T}_h} - (\phi, \nabla \cdot \boldsymbol{r}_2)_{\mathcal{T}_h} + \langle \phi, \boldsymbol{r}_2 \cdot \boldsymbol{n} \rangle_{\partial \mathcal{T}_h} \\
&\qquad + (\nabla \cdot (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}), w_2)_{\mathcal{T}_h} + (\nabla \cdot \boldsymbol{p}, w_2)_{\mathcal{T}_h} + \langle (\boldsymbol{p} - \boldsymbol{\Pi}_V \boldsymbol{p}) \cdot \boldsymbol{n}, \mu_2 \rangle_{\partial \mathcal{T}_h} \\
&\qquad + \langle \tau(\Pi_W \phi - \Pi_{k+1}^{\partial} \phi), w_2 - \mu_2 \rangle_{\partial \mathcal{T}_h} \\
&\quad = \mathcal{B}(\boldsymbol{p}, \phi, \phi; \boldsymbol{r}_2, w_2, \mu_2) + (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{r}_2)_{\mathcal{T}_h} + (\nabla \cdot (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}), w_2)_{\mathcal{T}_h} \\
&\qquad + \langle (\boldsymbol{p} - \boldsymbol{\Pi}_V \boldsymbol{p}) \cdot \boldsymbol{n}, \mu_2 \rangle_{\partial \mathcal{T}_h} + \langle \tau(\Pi_W \phi - \Pi_{k+1}^{\partial} \phi), w_2 - \mu_2 \rangle_{\partial \mathcal{T}_h}.
\end{aligned}
$$

Since

$$
\begin{aligned}
(\nabla \cdot (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}), w_2)_{\mathcal{T}_h} &= \langle (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}) \cdot \boldsymbol{n}, w_2 \rangle_{\partial \mathcal{T}_h} - (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \nabla w_2)_{\mathcal{T}_h} \\
&= \langle (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}) \cdot \boldsymbol{n}, w_2 \rangle_{\partial \mathcal{T}_h}.
\end{aligned}
$$

We have

$$
\begin{aligned}
&\mathcal{B}(\boldsymbol{\Pi}_V \boldsymbol{p}, \Pi_W \phi, \Pi_{k+1}^{\partial} \phi; \boldsymbol{r}_2, w_2, \mu_2) \\
&\quad = \mathcal{B}(\boldsymbol{p}, \phi, \phi; \boldsymbol{r}_2, w_2, \mu_2) + (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{r}_2)_{\mathcal{T}_h} \\
&\qquad + \langle (\boldsymbol{p} - \boldsymbol{\Pi}_V \boldsymbol{p}) \cdot \boldsymbol{n}, \mu_2 - w_2 \rangle_{\partial \mathcal{T}_h} + \langle \tau(\Pi_W \phi - \Pi_{k+1}^{\partial} \phi), w_2 - \mu_2 \rangle_{\partial \mathcal{T}_h}.
\end{aligned}
$$

Using the analogue of Eq. (14b) for the exact solution, and (10) we get

$$
\mathcal{B}(\boldsymbol{\Pi}_V \boldsymbol{p}, \Pi_W \phi, \Pi_{k+1}^{\partial} \phi; \boldsymbol{r}_2, w_2, \mu_2) = (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{r}_2)_{\mathcal{T}_h} - (u, w_2)_{\mathcal{T}_h}.
$$

Therefore, subtracting Eq. (14b) we have the following error equation

$$\mathcal{B}(\xi_h^{\boldsymbol{p}}, \xi_h^{\phi}, \xi_h^{\widehat{\phi}}; \boldsymbol{r}_2, w_2, \mu_2) = (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{r}_2)_{\mathcal{T}_h} - (u - u_h, w_2)_{\mathcal{T}_h}.$$

□

3.2.1 L^2 Error Estimates for p and ϕ

Lemma 3.3 *We have the following estimate*

$$\|\xi_h^{\boldsymbol{p}}\|_{\mathcal{T}_h}^2 + \|\sqrt{\tau}(\Pi_{k+1}^{\partial}\xi_h^{\phi} - \xi_h^{\widehat{\phi}})\|_{\partial\mathcal{T}_h}^2 \le \|u - u_h\|_{\mathcal{T}_h}\|\xi_h^{\phi}\|_{\mathcal{T}_h}.$$

Proof We take $(\boldsymbol{r}_2, w_2, \mu_2) = (\xi_h^{\boldsymbol{p}}, \xi_h^{\phi}, \xi_h^{\widehat{\phi}})$ in (18b) to get

$$\mathcal{B}(\xi_h^{\boldsymbol{p}}, \xi_h^{\phi}, \xi_h^{\widehat{\phi}}; \xi_h^{\boldsymbol{p}}, \xi_h^{\phi}, \xi_h^{\widehat{\phi}}) = -(u - u_h, \xi_h^{\phi})_{\mathcal{T}_h} \le \|u - u_h\|_{\mathcal{T}_h}\|\xi_h^{\phi}\|_{\mathcal{T}_h}.$$

On the other hand, by Lemma 2.6 we have

$$\|\xi_h^{\boldsymbol{p}}\|_{\mathcal{T}_h}^2 + \|\sqrt{\tau}(\xi_h^{\phi} - \xi_h^{\widehat{\phi}})\|_{\partial\mathcal{T}_h}^2 \le \|u - u_h\|_{\mathcal{T}_h}\|\xi_h^{\phi}\|_{\mathcal{T}_h}.$$

□

If we directly apply Lemma 2.9 to get the estimate of $\|\xi_h^{\phi}\|_{\mathcal{T}_h}$, we will obtain only suboptimal convergence rates. To obtain optimal rates we use the dual problem introduced in Eq. (8) with $\boldsymbol{p} = 0$ and $M = 0$ and assume the regularity estimate (9).

We follow the proof of Lemma 3.2 to get the following lemma.

Lemma 3.4 *Let $(\boldsymbol{\Phi}, \Psi)$ solve (8) with $\boldsymbol{p} = 0$ and $M = 0$ having data Θ. Then for any $(\boldsymbol{r}_2, w_2, \mu_2) \in \boldsymbol{S}_h \times \Psi_h \times \widehat{\Psi}_h(0)$, we have the following equation*

$$\mathcal{B}(\boldsymbol{\Pi}_V \boldsymbol{\Phi}, \Pi_W \Psi, \Pi_{k+1}^{\partial}\Psi; \boldsymbol{r}_2, w_2, \mu_2) = (\boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi}, \boldsymbol{r}_2)_{\mathcal{T}_h} + (\Theta, w_2)_{\mathcal{T}_h}.$$

Using this lemma we can now estimate ξ_h^{ϕ} in terms of $u - u_h$ and other consistency terms.

Lemma 3.5 *For any $t \in [0, T]$, if the elliptic regularity inequality (9) holds, then we have the following error estimates*

$$\|\xi_h^{\phi}\|_{\mathcal{T}_h}^2 \le Ch^2\|\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}\|_{\mathcal{T}_h}^2 + C\|u - u_h\|_{\mathcal{T}_h}^2.$$

Proof Consider the dual problem (8) with $\boldsymbol{p} = 0$ and $M = 0$ and $\Theta = \xi_h^{\phi}$. We take $(\boldsymbol{r}_2, w_2, \mu_2) = (-\boldsymbol{\Pi}_V \boldsymbol{\Phi}, \Pi_W \Psi, \Pi_{k+1}^{\partial}\Psi)$ in Eq. (18b) of Lemma 3.2 to get

$$\begin{aligned}
&\mathcal{B}(\xi_h^{\boldsymbol{p}}, \xi_h^{\phi}, \xi_h^{\widehat{\phi}}; -\boldsymbol{\Pi}_V \boldsymbol{\Phi}, \Pi_W \Psi, \Pi_{k+1}^{\partial}\Psi) \\
&\quad = -(\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{\Pi}_V \boldsymbol{\Phi})_{\mathcal{T}_h} - (u - u_h, \Pi_W \Psi)_{\mathcal{T}_h}.
\end{aligned} \tag{19}$$

On the other hand, by Lemmas 2.6 and 3.4, we have

$$\begin{aligned}
&\mathcal{B}(\xi_h^{\boldsymbol{p}}, \xi_h^{\phi}, \xi_h^{\widehat{\phi}}; -\boldsymbol{\Pi}_V \boldsymbol{\Phi}, \Pi_W \Psi, \Pi_{k+1}^{\partial}\Psi) \\
&\quad = \mathcal{B}(\boldsymbol{\Pi}_V \boldsymbol{\Phi}, \Pi_W \Psi, \Pi_{k+1}^{\partial}\Psi; -\xi_h^{\boldsymbol{p}}, \xi_h^{\phi}, \xi_h^{\widehat{\phi}}) \\
&\quad = -(\boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi}, \xi_h^{\boldsymbol{p}})_{\mathcal{T}_h} + \|\xi_h^{\phi}\|_{\mathcal{T}_h}^2.
\end{aligned} \tag{20}$$

Comparing the above two equalities (19) and (20) gives

$$\begin{aligned}
\|\xi_h^\phi\|_{\mathcal{T}_h}^2 &= (\boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi}, \xi_h^{\boldsymbol{p}})_{\mathcal{T}_h} - (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{\Pi}_V \boldsymbol{\Phi})_{\mathcal{T}_h} - (u - u_h, \Pi_W \Psi)_{\mathcal{T}_h} \\
&= (\boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi}, \xi_h^{\boldsymbol{p}})_{\mathcal{T}_h} - (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi})_{\mathcal{T}_h} \\
&\quad - (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{\Phi})_{\mathcal{T}_h} - (u - u_h, \Pi_W \Psi)_{\mathcal{T}_h} \\
&= (\boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi}, \xi_h^{\boldsymbol{p}})_{\mathcal{T}_h} - (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi})_{\mathcal{T}_h} \\
&\quad + (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \nabla \Psi)_{\mathcal{T}_h} - (u - u_h, \Pi_W \Psi)_{\mathcal{T}_h} \\
&= (\boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi}, \xi_h^{\boldsymbol{p}})_{\mathcal{T}_h} - (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{\Pi}_V \boldsymbol{\Phi} - \boldsymbol{\Phi})_{\mathcal{T}_h} \\
&\quad + (\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}, \nabla (\Psi - \Pi_W \Psi))_{\mathcal{T}_h} - (u - u_h, \Pi_W \Psi)_{\mathcal{T}_h} \\
&\le Ch^2 \|\xi_h^{\boldsymbol{p}}\|_{\mathcal{T}_h}^2 + Ch^2 \|\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}\|_{\mathcal{T}_h}^2 + \|u - u_h\|_{\mathcal{T}_h} \|\Pi_W \Psi\|_{\mathcal{T}_h}.
\end{aligned}$$

By the continuous dependence result (9) and the projection property of Π_W in (11b) we get

$$\|\Pi_W \Psi\|_{\mathcal{T}_h} \le \|\Pi_W \Psi - \Psi\|_{\mathcal{T}_h} + \|\Psi\|_{\mathcal{T}_h} \le C\|\Psi\|_{H^2(\Omega)} \le C\|\Theta\|_{\mathcal{T}_h} = C\|\xi_h^\phi\|_{\mathcal{T}_h}.$$

By Lemma 3.3 and the Cauchy-Schwarz inequality we obtain the result of the lemma:

$$\|\xi_h^\phi\|_{\mathcal{T}_h}^2 \le Ch^2 \|\boldsymbol{\Pi}_V \boldsymbol{p} - \boldsymbol{p}\|_{\mathcal{T}_h}^2 + C\|u - u_h\|_{\mathcal{T}_h}^2.$$

□

As a consequence of the above result, a simple application of the triangle inequality and Lemmas 3.3 and 3.5 gives the following bounds of $\|\phi - \phi_h\|_{\mathcal{T}_h}$ and $\|\boldsymbol{p} - \boldsymbol{p}_h\|_{\mathcal{T}_h}$:

Lemma 3.6 *Let $(\boldsymbol{p}, \phi)$ and $(\boldsymbol{p}_h, \phi_h)$ be the solutions of (2) and (3), respectively. For any $t \in [0, T]$, if the elliptic regularity inequality (9) holds, then we have the following error estimates*

$$\|\phi - \phi_h\|_{\mathcal{T}_h} + \|\boldsymbol{p} - \boldsymbol{p}_h\|_{\mathcal{T}_h} \le C_1 h^{k+2} + C\|u - u_h\|_{\mathcal{T}_h},$$

where C_1 depends on the $H^{k+1}(\Omega)$ norm of $\boldsymbol{p}$ at each time.

3.3 L^2 Error Estimates for u

Having the result of Lemma 3.6 it remains to estimate $u - u_h$. The fundamental estimate is contained in the next lemma.

Lemma 3.7 *If h small enough, then there exists $t_h^\star \in [0, T]$ such that for all $t \in [0, t_h^\star]$ we have*

$$\|\xi_h^u\|_{\mathcal{T}_h}^2 + \int_0^t \left(\|\xi_h^{\boldsymbol{q}}\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2}(\Pi_k^\partial \xi_h^u - \xi_h^{\widehat{u}})\|_{\partial \mathcal{T}_h}^2 \right) dt \le Ch^{2k+4}.$$

Proof We take $(\boldsymbol{r}_1, w_1, \mu_1) = (\xi_h^{\boldsymbol{q}}, \xi_h^u, \xi_h^{\widehat{u}})$ in (18a) to get

$$\begin{aligned}
&(\partial_t \xi_h^u, \xi_h^u)_{\mathcal{T}_h} + \|\xi_h^{\boldsymbol{q}}\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2}(\Pi_k^\partial \xi_h^u - \xi_h^{\widehat{u}})\|_{\partial \mathcal{T}_h}^2 \\
&\quad = (\partial_t (u_{Ih} - u), \xi_h^u)_{\mathcal{T}_h} + M(u - u_{Ih}, \xi_h^u)_{\mathcal{T}_h} \\
&\qquad - ((\boldsymbol{p} - \boldsymbol{p}_h) u_h, \nabla \xi_h^u)_{\mathcal{T}_h} + \langle (\boldsymbol{p} - \widehat{\boldsymbol{p}}_h) \cdot \boldsymbol{n} \widehat{u}_h, \xi_h^u - \xi_h^{\widehat{u}} \rangle_{\partial \mathcal{T}_h} \\
&\qquad - (\boldsymbol{p} \xi_h^u, \nabla \xi_h^u)_{\mathcal{T}_h} + \langle \boldsymbol{p} \cdot \boldsymbol{n} \xi_h^{\widehat{u}}, \xi_h^u \rangle_{\partial \mathcal{T}_h} \\
&\quad =: R_1 + R_2 + R_3 + R_4 + R_5 + R_6.
\end{aligned} \tag{21}$$

We note that $\xi_h^u(0) = u_h(0) - u_{Ih}(0) = 0$. Let $t = 0$ in (21) to get

$$\|\xi_h^{\boldsymbol{q}}(0)\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2}(\Pi_k^\partial \xi_h^u(0) - \xi_h^{\widehat{u}}(0))\|_{\partial\mathcal{T}_h}^2 = 0.$$

This implies $\xi_h^{\widehat{u}}(0) = \xi_h^u(0) = 0$. Hence we have $\widehat{u}_h(0) = \widehat{u}_{Ih}(0)$. By Theorem 3.1 we have

$$\begin{aligned}
\|\Pi_{k+1}^o u(0) - u_h(0)\|_{\mathcal{T}_h} &= \|\Pi_{k+1}^o u(0) - u_{Ih}(0)\|_{\mathcal{T}_h} \le Ch^{k+2},\\
\|\Pi_k^\partial u(0) - \widehat{u}_h(0)\|_{\partial\mathcal{T}_h} &= \|\Pi_k^\partial u(0) - \widehat{u}_{Ih}(0)\|_{\partial\mathcal{T}_h} \le Ch^{k+3/2}.
\end{aligned}$$

For h small enough these estimates imply that

$$\|u(t) - \Pi_{k+1}^o u(t)\|_{L^\infty(\Omega)} \le 1/2 \text{ and } \|u(t) - \Pi_k^\partial u(t)\|_{L^\infty(\mathcal{E}_h)} \le 1/2 \text{ for all } t \in [0, T]. \tag{22}$$

Let $\mathcal{M} = \max_{(t,x)\in[0,T]\times\Omega} |u(t,x)|$, then the inverse inequality gives

$$\begin{aligned}
\|u_h(0)\|_{L^\infty(\Omega)} &\le Ch^{-d/2}\|\Pi_{k+1}^o u(0) - u_h(0)\|_{\mathcal{T}_h}\\
&\quad + \|\Pi_{k+1}^o u(0) - u(0)\|_{L^\infty(\Omega)} + \|u(0)\|_{L^\infty(\Omega)}\\
&\le Ch^{k+2-d/2} + \mathcal{M} + 1/2,\\
\|\widehat{u}_h(0)\|_{L^\infty(\mathcal{E}_h)} &\le Ch^{1/2-d/2}\|\Pi_k^\partial u(0) - \widehat{u}_h(0)\|_{\mathcal{T}_h}\\
&\quad + \|\Pi_k^\partial u(0) - u(0)\|_{L^\infty(\mathcal{E}_h)} + \|u(0)\|_{L^\infty(\mathcal{E}_h)}\\
&\le Ch^{k+2-d/2} + \mathcal{M} + 1/2.
\end{aligned}$$

Also, since the error Eq. (18a) is continuous with respect to the time t, then there exists $t_h^\star \in [0, T]$ such that for h small enough,

$$\|u_h\|_{L^\infty(\Omega)} + \|\widehat{u}_h\|_{L^\infty(\mathcal{E}_h)} \le 2\mathcal{M} + 2. \tag{23}$$

By the Cauchy-Schwarz inequality, Theorem 3.1 and Lemma 2.7 we get

$$\begin{aligned}
R_1 + R_2 &\le Ch^{k+2}\|\xi_h^u\|_{\mathcal{T}_h}\\
&\le Ch^{2k+4} + \frac{1}{8}\left(\|\xi_h^{\boldsymbol{q}}\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2}(\Pi_k^\partial \xi_h^u - \xi_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}^2\right).
\end{aligned}$$

For the term R_3, by the Cauchy-Schwarz, Lemmas 3.6, 2.9 and 2.7 we get

$$\begin{aligned}
R_3 &\le C\|\boldsymbol{p} - \boldsymbol{p}_h\|_{\mathcal{T}_h}\|\nabla\xi_h^u\|_{\mathcal{T}_h}\\
&\le C\|\boldsymbol{p} - \boldsymbol{p}_h\|_{\mathcal{T}_h}^2 + \frac{1}{C}\|\nabla\xi_h^u\|_{\mathcal{T}_h}^2\\
&\le Ch^{2k+4} + C\|u - u_h\|_{\mathcal{T}_h}^2 + \frac{1}{C}\|\nabla\xi_h^u\|_{\mathcal{T}_h}^2\\
&\le Ch^{2k+4} + C\|\xi_h^u\|_{\mathcal{T}_h}^2 + \frac{1}{8}\left(\|\xi_h^{\boldsymbol{q}}\|_{\mathcal{T}_h}^2 + \|h_K^{-1/2}(\Pi_k^\partial \xi_h^u - \xi_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}^2\right).
\end{aligned}$$

Also, applying Lemma 2.7 again to obtain

$$\begin{aligned} R_4 &= \langle (\boldsymbol{p} - \widehat{\boldsymbol{p}}_h) \cdot \boldsymbol{n} \widehat{u}_h, \xi_h^u - \xi_h^{\widehat{u}} \rangle_{\partial \mathcal{T}_h} \\ &\le C \| h_K^{1/2} (\boldsymbol{p} - \widehat{\boldsymbol{p}}_h) \|_{\partial \mathcal{T}_h} \| h_K^{-1/2} (\xi_h^u - \xi_h^{\widehat{u}}) \|_{\partial \mathcal{T}_h} \\ &\le C h^{2k+4} + C \| \xi_h^u \|_{\mathcal{T}_h}^2 + \frac{1}{8} \left(\| \xi_h^{\boldsymbol{q}} \|_{\mathcal{T}_h}^2 + \| h_K^{-1/2} (\Pi_k^\partial \xi_h^u - \xi_h^{\widehat{u}}) \|_{\partial \mathcal{T}_h}^2 \right). \end{aligned}$$

For the last two terms $R_5 + R_6$, by the Assumption 2.1, we know that $\boldsymbol{p}$ is bounded. Next, integration by parts to get

$$\begin{aligned} R_5 + R_6 &= -(\boldsymbol{p} \xi_h^u, \nabla \xi_h^u)_{\mathcal{T}_h} + \langle \boldsymbol{p} \cdot \boldsymbol{n} \xi_h^{\widehat{u}}, \xi_h^u \rangle_{\partial \mathcal{T}_h} \\ &= -\frac{1}{2} \langle \boldsymbol{p} \cdot \boldsymbol{n} (\xi_h^u - \xi_h^{\widehat{u}}), \xi_h^u - \xi_h^{\widehat{u}} \rangle_{\mathcal{T}_h} - (\nabla \cdot \boldsymbol{p} \xi_h^u, \xi_h^u)_{\mathcal{T}_h} \\ &\le \frac{1}{8} \| h_K^{-1/2} (\Pi_k^\partial \xi_h^u - \xi_h^{\widehat{u}}) \|_{\partial \mathcal{T}_h}^2 + \| \nabla \cdot \boldsymbol{p} \|_{L^\infty(\Omega)} \| \xi_h^u \|_{\mathcal{T}_h}^2. \end{aligned}$$

Sum the above estimates of $\{R_i\}_{i=1}^6$ to get

$$(\partial_t \xi_h^u, \xi_h^u)_{\mathcal{T}_h} + \| \xi_h^{\boldsymbol{q}} \|_{\mathcal{T}_h}^2 + \| h_K^{-1/2} (\Pi_k^\partial \xi_h^u - \xi_h^{\widehat{u}}) \|_{\partial \mathcal{T}_h}^2 \le C h^{2k+4} + C \| \xi_h^u \|_{\mathcal{T}_h}^2. \tag{24}$$

Integrating both sides of (24) on $[0, t_h^*]$ we finally obtain

$$\begin{aligned} &\| \xi_h^u (t_h^*) \|_{\mathcal{T}_h}^2 + \int_0^{t_h^*} \left(\| \xi_h^{\boldsymbol{q}} \|_{\mathcal{T}_h}^2 + \| h_K^{-1/2} (\Pi_k^\partial \xi_h^u - \xi_h^{\widehat{u}}) \|_{\partial \mathcal{T}_h}^2 \right) dt \\ &\quad \le C h^{2k+4} + C \int_0^{t_h^*} \| \xi_h^u \|_{\mathcal{T}_h}^2 dt. \end{aligned}$$

The use of Gronwall's inequality gives the desired result. □

Lemma 3.8 *For h small enough, the result in Lemma 3.7 holds on the whole time interval* $[0, T]$.

Proof Fix $h^* > 0$ so that Lemma 3.7 is true for all $h \le h^*$, and assume t_h^* is the largest value for which (23) is true for all $h \le h^*$. Define the set $\mathbb{A} = \{h \in [0, h^*] : t_h^* \ne T\}$. If the result is not true, then $\mathbb{A}$ is nonempty, $\inf\{h : h \in \mathbb{A}\} = 0$, and also

$$\| u_h \|_{L^\infty(\Omega)} + \| \widehat{u}_h \|_{L^\infty(\mathcal{E}_h)} = 2\mathcal{M} + 2 \quad \text{for all } h \in \mathcal{A}. \tag{25}$$

However, by the inverse inequality and since Lemma 3.7 holds, we have

$$\| u_h \|_{L^\infty(\Omega)} + \| \widehat{u}_h \|_{L^\infty(\mathcal{E}_h)} \le C h^{2-d/2} + 2\mathcal{M} + 1 \quad \text{for all } h \in \mathcal{A}.$$

Since C does not depend on h, there exists $h_1^* \le h^*$ such that $\| u_h \|_{L^\infty(\Omega)} + \| \widehat{u}_h \|_{L^\infty(\mathcal{E}_h)} < 2\mathcal{M} + 2$ for all $h \in \mathbb{A}$ such that $h \le h_1^*$. This contradicts (25), and therefore $t_h^* = T$ for all h small enough. □

The above lemma, the triangle inequality, and Lemma 3.3 complete the proof of Theorem 2.3.

4 Numerical Results

In this section we present some numerical results in two spatial dimensions.

Example 4.1 We begin with an example with an exact solution in order to illustrate the convergence theory. The domain is the unit square $\Omega = [0, 1] \times [0, 1] \subset \mathbb{R}^2$ and homogeneous Dirichlet boundary conditions are applied on the boundary. The source terms f_1, f_2 and the initial condition are chosen so that $\varepsilon = 0.1$ and the exact solution $u = \cos(t)\sin(x)\cos(y)$ and $\phi = \sin(t)\cos(x)\sin(y)$. The second order backward differentiation formula (BDF2) is applied for the time discretization and for the space discretization we choose polynomial degrees $k = 0$ or $k = 1$ (used in the definition of the discrete spaces in Sect. 1). The time step is chosen to be $\Delta t = h$ when $k = 0$ and $\Delta t = h^{3/2}$ when $k = 1$. We report the errors at the final time $T = 1$. The observed convergence rates match our theory (Tables 1, 2).

Next, we test an example without a convergence rate but that show the performance of the HDG method. We take $k = 0$, the domain is also the unit square $\Omega = [0, 1] \times [0, 1] \subset \mathbb{R}^2$

Table 1 History of convergence for $\boldsymbol{q}_h$ and $\boldsymbol{p}_h$ for Example 4.1 under uniform mesh refinement

Degree	$\frac{h}{\sqrt{2}}$	$\\|\boldsymbol{q}-\boldsymbol{q}_h\\|_{0,\Omega}$		$\\|\boldsymbol{p}-\boldsymbol{p}_h\\|_{0,\Omega}$	
		Error	Rate	Error	Rate
$k=0$	2^{-1}	4.2730E−02		6.4843E−03	
	2^{-2}	2.2386E−02	0.93	1.9113E−03	1.76
	2^{-3}	1.1265E−02	0.99	5.1822E−04	1.88
	2^{-4}	5.6455E−03	1.00	1.3592E−04	1.93
	2^{-5}	2.8248E−03	1.00	3.4881E−05	1.96
$k=1$	2^{-1}	2.9547E−03		3.8888E−04	
	2^{-2}	7.5335E−04	1.97	5.4882E−05	2.82
	2^{-3}	1.9796E−04	1.93	7.5341E−06	2.86
	2^{-4}	5.0451E−05	1.97	9.8858E−07	2.93
	2^{-5}	1.2748E−05	1.98	1.2705E−07	2.96

Table 2 History of convergence for u_h and ϕ_h for Example 4.1 under uniform mesh refinement

Degree	$\frac{h}{\sqrt{2}}$	$\\|u-u_h\\|_{0,\Omega}$		$\\|\phi-\phi_h\\|_{0,\Omega}$	
		Error	Rate	Error	Rate
$k=0$	2^{-1}	1.8339E−02		1.0205E−02	
	2^{-2}	4.9503E−03	1.89	2.3408E−03	2.12
	2^{-3}	1.2423E−03	2.00	4.9774E−04	2.23
	2^{-4}	3.1156E−04	2.00	1.1131E−04	2.16
	2^{-5}	7.7965E−05	2.00	2.6001E−05	2.09
$k=1$	2^{-1}	1.8339E−02		4.0894E−04	
	2^{-2}	2.3140E−04	2.98	3.7700E−05	3.43
	2^{-3}	2.9565E−05	2.97	4.6167E−06	3.03
	2^{-4}	3.7026E−06	3.00	5.3872E−07	3.01
	2^{-5}	4.6363E−07	3.00	6.6418E−08	3.02

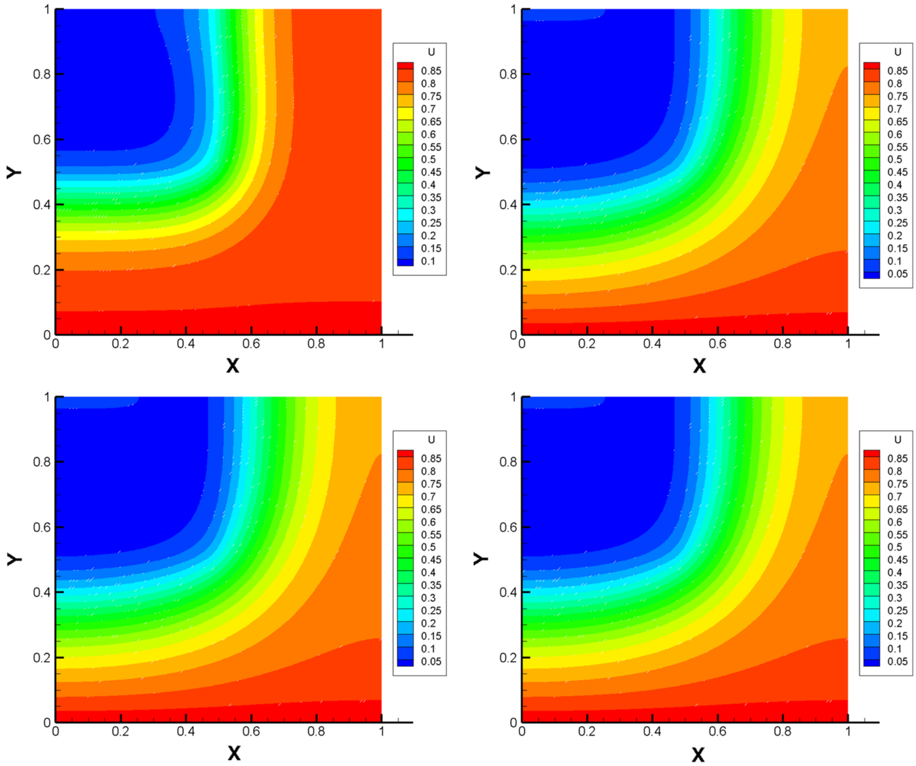

Fig. 1 From left to right, from top to bottom are the contour plots of u_h at time: $T = 0.01, 0.4, 0.7, 1$ for Example 4.2

and partition into 20,000 triangles, i.e., $h = \sqrt{2}/100$. BDF2 is applied for time discretization and the time step $\Delta t = 1/1000$, at each time step, we utilized the Newton's method to solve the nonlinear system.

Example 4.2 This example has non-homogeneous Dirichlet data and demonstrates that our HDG scheme can handle this case. We take $\varepsilon = 10^{-2}$ and the source terms $f_1 = 0$ and

$$f_2 = \begin{cases} -0.8 & (0, 0.5) \times (1/2, 1), \\ 0.8 & \text{else}. \end{cases}$$

The Dirichlet boundary condition $g_u = 0.9$, $g_\phi = 1.1$ on $\{y = 0\}$, and $g_u = 0.1$, $g_\phi = -1.1$ on $\{y = 1, 0 \le x \le 0.25\}$. Elsewhere we impose homogeneous Neumann boundary conditions. Initial condition $u_0 = (1 + f_2)/2$. A similar example was studied in [3] by a finite volume method. We plot the solutions u_h and ϕ_h at different final time T; see Figs. 1 and 2.

5 Conclusion

In this work, we proposed an HDG method for the drift–diffusion equation. We proved optimal semi-discrete error estimates for all variables; moreover, from the point view of

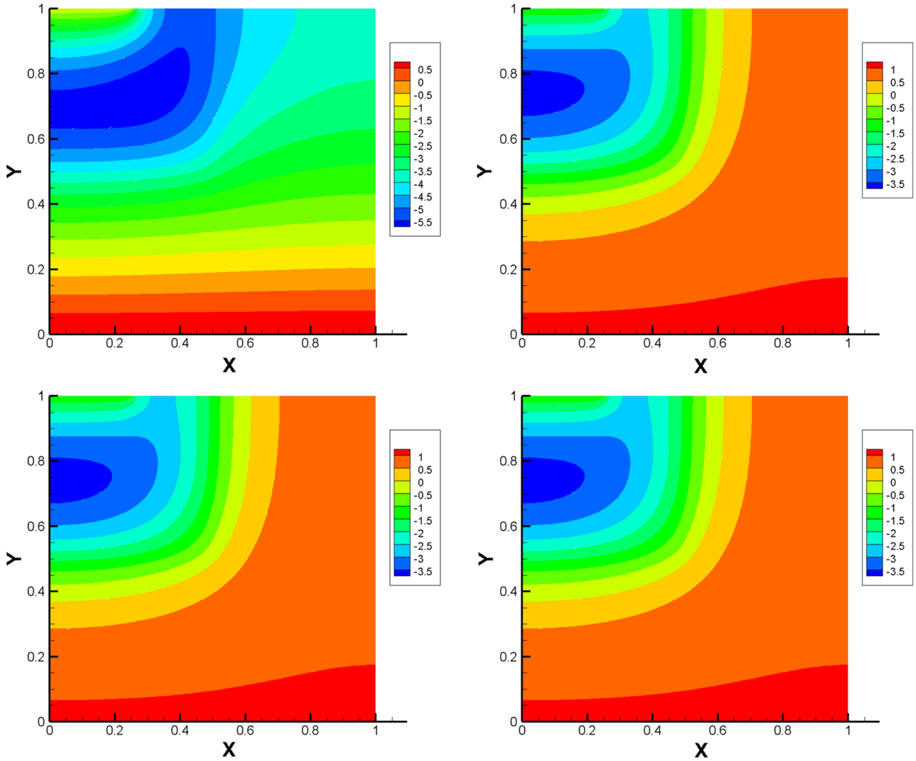

Fig. 2 From left to right, from top to bottom are the contour plots of ϕ_h at time: $T = 0.01, 0.4, 0.7, 1$ for Example 4.2

degrees of freedom, we obtained a superconvergent convergence rate for the variable u. As far as we are aware, this is the first such result in the literature.

Clearly it would be desirable to prove convergence without the need to assume an inverse assumption. Equally, it would be useful to prove fully discrete estimates using, for example BDF2 in time.

This is the first of a series of papers in which we develop efficient HDG methods for drift–diffusion equation, including devising HDG methods when ε approaches to zero. We have a great interest in the numerical solution of steady state drift–diffusion equation, and we will explore this problem in our future papers.

Acknowledgements G. Chen is supported by National natural science Foundation of China (NSFC) under Grant Number 11801063 and China Postdoctoral Science Foundation under Grant Number 2018M633339. The research of P. B. Monk and Y. Zhang is partially supported by the US National Science Foundation (NSF) under Grant Number DMS-1619904.

6 Appendix

In this section we give a proof for (17a) and (17b). The proof of (17c) is similar and we do not provide details.

6.1 Error Equations

We start be deriving equations satisfied by standard projections [see (12)] of the exact solution.

Lemma 6.1 *Let* $(\boldsymbol{q}, u)$ *be components of the solution of* (2)*, then we have*

$$
\begin{aligned}
&M(\Pi_{k+1}^{o} u, w_1)_{\mathcal{T}_h} + \mathcal{A}(\boldsymbol{\Pi}_k^{o}\boldsymbol{q}, \Pi_{k+1}^{o} u, \Pi_k^{\partial} u; \boldsymbol{r}_1, w_1, \mu_1) + \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \Pi_{k+1}^{o} u, \Pi_k^{\partial} u; w_1)\\
&\quad = (Mu - u_t, w_1)_{\mathcal{T}_h} + \langle(\boldsymbol{\Pi}_k^{o}\boldsymbol{q} - \boldsymbol{q})\cdot\boldsymbol{n}, w_1 - \mu_1\rangle_{\partial\mathcal{T}_h} + (\boldsymbol{p}(\Pi_{k+1}^{o} u - u), \nabla w_1)_{\mathcal{T}_h}\\
&\qquad - \langle \boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^{\partial} u - u), w_1 - \mu_1\rangle_{\partial\mathcal{T}_h} + \langle h_K^{-1}(\Pi_{k+1}^{o} u - u), \Pi_k^{\partial} w_1 - \mu_1\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

holds for all $(\boldsymbol{r}_1, w_1, \mu_1) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(0)$.

Proof By the definition of $\mathcal{A}$ and $\mathcal{C}$ in (14c) and (14e) respectively, the projections and integrating by parts, we get

$$
\begin{aligned}
&\mathcal{A}(\boldsymbol{\Pi}_k^{o}\boldsymbol{q}, \Pi_{k+1}^{o} u, \Pi_k^{\partial} u; \boldsymbol{r}_1, w_1, \mu_1)\\
&\quad = \langle(\boldsymbol{\Pi}_k^{o}\boldsymbol{q} - \boldsymbol{q})\cdot\boldsymbol{n}, w_1 - \mu_1\rangle_{\partial\mathcal{T}_h} + (\nabla\cdot\boldsymbol{q}, w_1)_{\mathcal{T}_h}\\
&\qquad + \langle h_K^{-1}(\Pi_{k+1}^{o} u - u), \Pi_k^{\partial} w_1 - \mu_1\rangle_{\partial\mathcal{T}_h},
\end{aligned}
$$

where we have also used (2a). In addition,

$$
\mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \Pi_{k+1}^{o} u, \Pi_k^{\partial} u; w_1) = (\boldsymbol{p}\Pi_{k+1}^{o} u, \nabla w_1)_{\mathcal{T}_h} - \langle \boldsymbol{p}\cdot\boldsymbol{n}\Pi_k^{\partial} u, w_1\rangle_{\partial\mathcal{T}_h}.
$$

Hence, again using the projections, we have

$$
\begin{aligned}
&M(\Pi_{k+1}^{o} u, w_1)_{\mathcal{T}_h} + \mathcal{A}(\boldsymbol{\Pi}_k^{o}\boldsymbol{q}, \Pi_{k+1}^{o} u, \Pi_k^{\partial} u; \boldsymbol{r}_1, w_1, \mu_1) + \mathcal{C}(\boldsymbol{p}; \Pi_{k+1}^{o} u, \Pi_k^{\partial} u; w_1)\\
&\quad = (Mu, w_1)_{\mathcal{T}_h} + \langle(\boldsymbol{\Pi}_k^{o}\boldsymbol{q} - \boldsymbol{q})\cdot\boldsymbol{n}, w_1 - \mu_1\rangle_{\partial\mathcal{T}_h} + (\nabla\cdot\boldsymbol{q}, w_1)_{\mathcal{T}_h}\\
&\qquad + \langle h_K^{-1}(\Pi_{k+1}^{o} u - u), \Pi_k^{\partial} w_1 - \mu_1\rangle_{\partial\mathcal{T}_h} + (\boldsymbol{p}\Pi_{k+1}^{o} u, \nabla w_1)_{\mathcal{T}_h} - \langle \boldsymbol{p}\cdot\boldsymbol{n}\Pi_k^{\partial} u, w_1\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

Since, using (2c), $\nabla\cdot\boldsymbol{q} = \nabla\cdot(\boldsymbol{p}u) - u_t$, then we have

$$
(\nabla\cdot\boldsymbol{q}, w_1)_{\mathcal{T}_h} = -(u_t, w_1)_{\mathcal{T}_h} + \langle \boldsymbol{p}\cdot\boldsymbol{n}u, w_1\rangle_{\partial\mathcal{T}_h} - (\boldsymbol{p}u, \nabla u)_{\mathcal{T}_h}.
$$

This implies that

$$
\begin{aligned}
&M(\Pi_{k+1}^{o} u, w_1)_{\mathcal{T}_h} + \mathcal{A}(\boldsymbol{\Pi}_k^{o}\boldsymbol{q}, \Pi_{k+1}^{o} u, \Pi_k^{\partial} u; \boldsymbol{r}_1, w_1, \mu_1) + \mathcal{C}(\boldsymbol{p}; \Pi_{k+1}^{o} u, \Pi_k^{\partial} u; w_1)\\
&\quad = (Mu - u_t, w_1)_{\mathcal{T}_h} + \langle(\boldsymbol{\Pi}_k^{o}\boldsymbol{q} - \boldsymbol{q})\cdot\boldsymbol{n}, w_1 - \mu_1\rangle_{\partial\mathcal{T}_h} + (\boldsymbol{p}(\Pi_{k+1}^{o} u - u), \nabla w_1)_{\mathcal{T}_h}\\
&\qquad - \langle \boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^{\partial} u - u), w_1 - \mu_1\rangle_{\partial\mathcal{T}_h} + \langle h_K^{-1}(\Pi_{k+1}^{o} u - u), \Pi_k^{\partial} w_1 - \mu_1\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

and completes the proof of the lemma. □

To simplify notation, we define

$$
\eta_h^{\boldsymbol{q}} := \boldsymbol{\Pi}_k^{o}\boldsymbol{q} - \boldsymbol{q}_{Ih}, \quad \eta_h^{u} := \Pi_{k+1}^{o} u - u_{Ih}, \quad \eta_h^{\widehat{u}} := \Pi_k^{\partial} u - \widehat{u}_{Ih}.
$$

We then subtract the equation in Lemma 6.1 from (15) to get the following lemma.

Lemma 6.2 *Under the conditions of Lemma 6.1, we have the error equation*

$$
\begin{aligned}
&M(\eta_h^{u}, w_1)_{\mathcal{T}_h} + \mathcal{A}(\eta_h^{\boldsymbol{q}}, \eta_h^{u}, \eta_h^{\widehat{u}}; \boldsymbol{r}_1, w_1, \mu_1) + \mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \eta_h^{u}, \eta_h^{\widehat{u}}; w_1)\\
&\quad = \langle(\boldsymbol{\Pi}_k^{o}\boldsymbol{q} - \boldsymbol{q})\cdot\boldsymbol{n}, w_1 - \mu_1\rangle_{\partial\mathcal{T}_h} + (\boldsymbol{p}(\Pi_{k+1}^{o} u - u), \nabla w_1)_{\mathcal{T}_h}\\
&\qquad - \langle \boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^{\partial} u - u), w_1\rangle_{\partial\mathcal{T}_h} + \langle h_K^{-1}(\Pi_{k+1}^{o} u - u), \Pi_k^{\partial} w_1 - \mu_1\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
\tag{26}
$$

holds for all $(\boldsymbol{r}_1, w_1, \mu_1) \in \boldsymbol{Q}_h \times V_h \times \widehat{V}_h(0)$.

6.2 Main Error Estimate

We can now prove (17b).

Lemma 6.3 *For h small enough, we have the error estimates*

$$\|\boldsymbol{q}-\boldsymbol{q}_{Ih}\|_{\mathcal{T}_h}+\|h_K^{-1/2}(\Pi_k^{\partial}u_{Ih}-\widehat{u}_{Ih})\|_{\partial\mathcal{T}_h}\le Ch^{k+1}|u|_{k+2}.$$

Proof We take $(\boldsymbol{r}_1, w_1, \mu_1)=(\eta_h^{\boldsymbol{q}}, \eta_h^u, \eta_h^{\widehat{u}})$ in (26). First

$$\begin{aligned}\mathcal{A}(\eta_h^{\boldsymbol{q}}, \eta_h^u, \eta_h^{\widehat{u}}; \eta_h^{\boldsymbol{q}}, \eta_h^u, \eta_h^{\widehat{u}}) &= \|\eta_h^{\boldsymbol{q}}\|_{\mathcal{T}_h}^2\\ &\quad+\|h_K^{-1/2}(\Pi_k^{\partial}\eta_h^u-\eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}^2.\end{aligned}$$

Next,

$$\begin{aligned}&M(\eta_h^u, \eta_h^u)_{\mathcal{T}_h}+\mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \eta_h^u, \eta_h^{\widehat{u}}; \eta_h^u)\\ &\quad= M(\eta_h^u, \eta_h^u)_{\mathcal{T}_h}+(\boldsymbol{p}\eta_h^u, \nabla\eta_h^u)_{\mathcal{T}_h}-\langle \boldsymbol{p}\cdot\boldsymbol{n}\eta_h^{\widehat{u}}, \eta_h^u\rangle_{\partial\mathcal{T}_h}\\ &\quad=\left(M-\frac{1}{2}\nabla\cdot\boldsymbol{p}, \eta_h^u\eta_h^u\right)_{\mathcal{T}_h}+\frac{1}{2}\langle \boldsymbol{p}\cdot\boldsymbol{n}\eta_h^u, \eta_h^u\rangle_{\partial\mathcal{T}_h}-\langle \boldsymbol{p}\cdot\boldsymbol{n}\eta_h^{\widehat{u}}, \eta_h^u\rangle_{\partial\mathcal{T}_h}\\ &\quad=\left(M-\frac{1}{2}\nabla\cdot\boldsymbol{p}, \eta_h^u\eta_h^u\right)_{\mathcal{T}_h}+\frac{1}{2}\langle \boldsymbol{p}\cdot\boldsymbol{n}(\eta_h^u-\eta_h^{\widehat{u}}), \eta_h^u-\eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\\ &\quad\ge\frac{M}{2}\|\eta_h^u\|_{\mathcal{T}_h}-\frac{1}{2}\||\boldsymbol{p}\cdot\boldsymbol{n}|(\Pi_k^{\partial}\eta_h^u-\eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}^2-Ch\|\boldsymbol{p}\|_{0,\infty}\|\nabla\xi_h^u\|_{\mathcal{T}_h}^2.\end{aligned}$$

For h small enough, we obtain

$$\begin{aligned}&M(\eta_h^u, \eta_h^u)_{\mathcal{T}_h}+\mathcal{A}(\eta_h^{\boldsymbol{q}}, \eta_h^u, \eta_h^{\widehat{u}}; \eta_h^{\boldsymbol{q}}, \eta_h^u, \eta_h^{\widehat{u}})+\mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \eta_h^u, \eta_h^{\widehat{u}}; \eta_h^u)\\ &\quad\ge\frac{1}{2}\left(M\|\eta_h^u\|_{\mathcal{T}_h}^2+\|\eta_h^{\boldsymbol{q}}\|_{\mathcal{T}_h}^2+\|h_K^{-1/2}(\Pi_k^{\partial}\eta_h^u-\eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}^2\right).\end{aligned}$$

On the other hand,

$$\begin{aligned}&M(\eta_h^u, \eta_h^u)_{\mathcal{T}_h}+\mathcal{A}(\eta_h^{\boldsymbol{q}}, \eta_h^u, \eta_h^{\widehat{u}}; \eta_h^{\boldsymbol{q}}, \eta_h^u, \eta_h^{\widehat{u}})+\mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \eta_h^u, \eta_h^{\widehat{u}}; \eta_h^u)\\ &\quad=\langle(\boldsymbol{\Pi}_k^o\boldsymbol{q}-\boldsymbol{q})\cdot\boldsymbol{n}, \eta_h^u-\eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h}+(\boldsymbol{p}(\Pi_{k+1}^o u-u), \nabla\eta_h^u)_{\mathcal{T}_h}\\ &\qquad-\langle\boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^{\partial}u-u), \eta_h^u-\eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h}+\langle h_K^{-1}(\Pi_{k+1}^o u-u), \Pi_k^{\partial}\eta_h^u-\eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\\ &\quad=: R_1+R_2+R_3+R_4.\end{aligned}$$

Next, we estimate $\{R_i\}_{i=1}^4$ term by term. For the first term R_1, Lemma 2.7 gives

$$\begin{aligned}R_1&\le Ch^{k+1}|\boldsymbol{q}|_{k+1}\|h_K^{-1/2}(\eta_h^u-\eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h},\\ &\le Ch^{k+1}|\boldsymbol{q}|_{k+1}\left(\|\eta_h^{\boldsymbol{q}}\|_{\mathcal{T}_h}+\|h_K^{-1/2}(\Pi_k^{\partial}\eta_h^u-\eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}\right).\end{aligned}$$

For the term R_2, by Lemmas 2.9 and 2.7 to get

$$\begin{aligned}R_2&\le Ch^{k+2}|u|_{k+2}\|\nabla\eta_h^u\|_{\mathcal{T}_h}\\ &\le Ch^{k+2}|u|_{k+2}\left(\|\eta_h^{\boldsymbol{p}}\|_{\mathcal{T}_h}+\|h_K^{-1/2}(\Pi_k^{\partial}\eta_h^u-\eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}\right).\end{aligned}$$

For the term R_3, we use Lemma 2.7 to get

$$
\begin{aligned}
R_3 &= \langle \boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^{\partial}u - u), \eta_h^u - \eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h} \\
&\le Ch^{k+1}|u|_{k+1}\|h_K^{-1/2}(\eta_h^u - \eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h} \\
&\le Ch^{k+1}|u|_{k+1}\left(\|\eta_h^{\boldsymbol{p}}\|_{\mathcal{T}_h} + \|h_K^{-1/2}(\Pi_k^{\partial}\eta_h^u - \eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}\right).
\end{aligned}
$$

Moreover, for the last term we have

$$
R_4 \le Ch^{k+1}|u|_{k+1}\|h_K^{-1/2}(\Pi_k^{\partial}\eta_h^u - \eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}.
$$

Use the Cauchy-Schwarz inequality for the above estimates of $\{R_i\}_{i=1}^4$, we get

$$
\|\eta_h^{\boldsymbol{q}}\|_{\mathcal{T}_h} + \|h_K^{-1/2}(\Pi_k^{\partial}\eta_h^u - \eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h} \le Ch^{k+1}|u|_{k+2}.
$$

Use of the triangle inequality and estimates (13a) and (13b) completes the estimate. $\square$

6.3 Duality Arguments

To obtain a L^2 norm estimate of $\|\eta_h^u\|_{\mathcal{T}_h}$, we use the dual problem (8) with corresponding a priori estimate (9). To perform the error analysis, the main difficulty is to deal with the nonlinearity. We define a new form $\mathcal{C}^\star$ which is related to the trilinear form $\mathcal{C}$:

$$
\mathcal{C}^\star(\boldsymbol{p},\boldsymbol{p};u_h,\widehat{u}_h;w_1) = -(\boldsymbol{p}u_h,\nabla w_1)_{\mathcal{T}_h} + \langle \boldsymbol{p}\cdot\boldsymbol{n}\widehat{u}_h, w_1\rangle_{\partial\mathcal{T}_h} - (\nabla\cdot\boldsymbol{p}u_h, w_1)_{\mathcal{T}_h}. \quad (27)
$$

Next, we give a property of the operators $\mathcal{C}$ and $\mathcal{C}^\star$. We omit the proof since it is very straightforward.

Lemma 6.4 *For all $(u_h,\widehat{u}_h,w_1,\mu_1)\in V_h\times\widehat{V}_h(0)\times V_h\times\widehat{V}_h(0)$, we have*

$$
\mathcal{C}(\boldsymbol{p},\boldsymbol{p};u_h,\widehat{u}_h;w_1) + \mathcal{C}^\star(\boldsymbol{p},\boldsymbol{p};w_1,\mu_1;-u_h) = \langle \boldsymbol{p}\cdot\boldsymbol{n}(u_h-\widehat{u}_h), w_1-\mu_1\rangle_{\partial\mathcal{T}_h}.
$$

Similarly to Lemma 6.1, we have the following lemma.

Lemma 6.5 *Assuming M is chosen sufficiently large, let $(\Phi,\boldsymbol{\Psi})$ solve (8) then we have the equation*

$$
\begin{aligned}
&M(\Pi_{k+1}^o\Phi, w_1)_{\mathcal{T}_h} + \mathcal{A}(\boldsymbol{\Pi}_k^o\boldsymbol{\Psi}, \Pi_{k+1}^o\Phi, \Pi_k^{\partial}\Phi; \boldsymbol{r}_1, w_1, \mu_1) + \mathcal{C}^\star(\boldsymbol{p},\boldsymbol{p};\Pi_{k+1}^o\Phi,\Pi_k^{\partial}\Phi;w_1) \\
&\quad= (\Theta, w_1) + \langle(\boldsymbol{\Pi}_k^o\boldsymbol{\Psi} - \boldsymbol{\Psi}\cdot)\boldsymbol{n}, w_1-\mu_1\rangle_{\partial\mathcal{T}_h} + \langle h_K^{-1}(\Pi_{k+1}^o\Phi - \Phi), \Pi_k^{\partial}w_1 - \mu_1\rangle_{\partial\mathcal{T}_h} \\
&\qquad- (\boldsymbol{p}(\Pi_{k+1}^o\Phi - \Phi), \nabla w_1)_{\mathcal{T}_h} + \langle \boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^{\partial}\Phi - \Phi), w_1\rangle_{\partial\mathcal{T}_h} \\
&\qquad- (\nabla\cdot\boldsymbol{p}(\Pi_{k+1}^o\Phi - \Phi), w_1)_{\mathcal{T}_h}.
\end{aligned}
$$

holds for all $(\boldsymbol{r}_1,w_1,\mu_1)\in\boldsymbol{Q}_h\times V_h\times\widehat{V}_h(0)$.

With the above preparation we can now derive estimate (17a).

Theorem 6.6 *Let u and u_{Ih} be the solutions of (2) and (15), respectively. If h is small enough, then we have the error estimate*

$$
\|u - u_{Ih}\|_{\mathcal{T}_h} \le Ch^{k+2}\|u\|_{k+2}.
$$

Proof We take $(\boldsymbol{r}_1, w_1, \mu_1) = (\eta_h^{\boldsymbol{q}}, -\eta_h^u, -\eta_h^{\widehat{u}})$ and $\Theta = -\eta_h^u$ in Lemma 6.5 to get

$$
\begin{aligned}
&- M(\Pi_{k+1}^o\Phi, \eta_h^u)_{\mathcal{T}_h} + \mathcal{A}(\boldsymbol{\Pi}_k^o\boldsymbol{\Psi}, \Pi_{k+1}^o\Phi, \Pi_k^\partial\Phi; \eta_h^{\boldsymbol{q}}, -\eta_h^u, -\eta_h^{\widehat{u}}) + \mathcal{C}^\star(\boldsymbol{p}; \Pi_{k+1}^o\Phi, \Pi_k^\partial\Phi; -\eta_h^u)\\
&\quad = -M(\eta_h^u, \Pi_{k+1}^o\Phi)_{\mathcal{T}_h} - \mathcal{A}(\eta_h^{\boldsymbol{q}}, \eta_h^u, \eta_h^{\widehat{u}}; -\boldsymbol{\Pi}_k^o\boldsymbol{\Psi}, \Pi_{k+1}^o\Phi, \Pi_k^\partial\Phi) + \mathcal{C}^\star(\boldsymbol{p}; \Pi_{k+1}^o\Phi, \Pi_k^\partial\Phi; -\eta_h^u)\\
&\quad = -M(\eta_h^u, \Pi_{k+1}^o\Phi)_{\mathcal{T}_h} - \langle(\boldsymbol{\Pi}_k^o\boldsymbol{q} - \boldsymbol{q})\cdot\boldsymbol{n}, \Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h} - (\boldsymbol{p}(\Pi_{k+1}^o u - u), \nabla\Pi_{k+1}^o\Phi)_{\mathcal{T}_h}\\
&\qquad + \langle\boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^\partial u - u), \Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h} - \langle h_K^{-1}(\Pi_{k+1}^o u - u), \Pi_k^\partial\Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h}\\
&\qquad + M(\eta_h^u, \Pi_{k+1}^o\Phi)_{\mathcal{T}_h} + \mathcal{C}(\boldsymbol{p}; \eta_h^u, \eta_h^{\widehat{u}}; \Pi_{k+1}^o\Phi) + \mathcal{C}^\star(\boldsymbol{p}; \Pi_{k+1}^o\Phi, \Pi_k^\partial\Phi; -\eta_h^u).
\end{aligned}
$$

By Lemma 6.4 we have

$$
\begin{aligned}
&\mathcal{C}(\boldsymbol{p}, \boldsymbol{p}; \eta_h^u, \eta_h^{\widehat{u}}; \Pi_{k+1}^o\Phi) + \mathcal{C}^\star(\boldsymbol{p}, \boldsymbol{p}; \Pi_{k+1}^o\Phi, \Pi_k^\partial\Phi; -\eta_h^u)\\
&\quad = \langle\boldsymbol{p}\cdot\boldsymbol{n}(\eta_h^u - \eta_h^{\widehat{u}}), \Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

This implies

$$
\begin{aligned}
&- M(\Pi_{k+1}^o\Phi, \eta_h^u)_{\mathcal{T}_h} + \mathcal{A}(\boldsymbol{\Pi}_k^o\boldsymbol{\Psi}, \Pi_{k+1}^o\Phi, \Pi_k^\partial\Phi; \eta_h^{\boldsymbol{q}}, -\eta_h^u, -\eta_h^{\widehat{u}})\\
&\quad + \mathcal{C}^\star(\boldsymbol{p}, \boldsymbol{p}; \Pi_{k+1}^o\Phi, \Pi_k^\partial\Phi; -\eta_h^u)\\
&\quad = -\langle(\boldsymbol{\Pi}_k^o\boldsymbol{q} - \boldsymbol{q})\cdot\boldsymbol{n}, \Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h} - (\boldsymbol{p}(\Pi_{k+1}^o u - u), \nabla\Pi_{k+1}^o\Phi)_{\mathcal{T}_h}\\
&\qquad + \langle\boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^\partial u - u), \Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h}\\
&\qquad - \langle h_K^{-1}(\Pi_{k+1}^o u - u), \Pi_k^\partial\Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h}\\
&\qquad + \langle\boldsymbol{p}\cdot\boldsymbol{n}(\eta_h^u - \eta_h^{\widehat{u}}), \Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h}.
\end{aligned}
$$

On the other hand, we have

$$
\begin{aligned}
&- M(\Pi_{k+1}^o\Phi, \eta_h^u)_{\mathcal{T}_h} + \mathcal{A}(\boldsymbol{\Pi}_k^o\boldsymbol{\Psi}, \Pi_{k+1}^o\Phi, \Pi_k^\partial\Phi; \eta_h^{\boldsymbol{q}}, -\eta_h^u, -\eta_h^{\widehat{u}})\\
&\quad + \mathcal{C}^\star(\boldsymbol{p}, \boldsymbol{p}; \Pi_{k+1}^o\Phi, \Pi_k^\partial\Phi; -\eta_h^u)\\
&= -\|\eta_h^u\|_{\mathcal{T}_h}^2 - \langle(\boldsymbol{\Pi}_k^o\boldsymbol{\Psi} - \boldsymbol{\Psi}\cdot)\boldsymbol{n}, \eta_h^u, -\eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h} - \langle h_K^{-1}(\Pi_{k+1}^o\Phi - \Phi), \Pi_k^\partial\eta_h^u - \eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\\
&\quad + (\boldsymbol{p}(\Pi_{k+1}^o\Phi - \Phi), \nabla\eta_h^u)_{\mathcal{T}_h} - \langle\boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^\partial\Phi - \Phi), \eta_h^u - \eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\\
&\quad + (\nabla\cdot\boldsymbol{p}(\Pi_{k+1}^o\Phi - \Phi), \eta_h^u)_{\mathcal{T}_h}.
\end{aligned}
$$

Comparing the above two equations, we get

$$
\begin{aligned}
\|\eta_h^u\|_{\mathcal{T}_h}^2 &= -\langle\boldsymbol{\Pi}_k^o\boldsymbol{q}\cdot\boldsymbol{n} - \boldsymbol{q}\cdot\boldsymbol{n}, \Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h} - \langle h_K^{-1}(\Pi_{k+1}^o u - u), \Pi_k^\partial\Pi_{k+1}^o\Phi\\
&\quad - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h} - (\boldsymbol{p}(\Pi_{k+1}^o u - u), \nabla\Pi_{k+1}^o\Phi)_{\mathcal{T}_h}\\
&\quad + \langle\boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^\partial u - u), \Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h}\\
&\quad + \langle\boldsymbol{p}\cdot\boldsymbol{n}(\eta_h^u - \eta_h^{\widehat{u}}), \Pi_{k+1}^o\Phi - \Pi_k^\partial\Phi\rangle_{\partial\mathcal{T}_h}\\
&\quad - \langle\boldsymbol{\Pi}_k^o\boldsymbol{\Psi}\cdot\boldsymbol{n} - \boldsymbol{\Psi}\cdot\boldsymbol{n}, \eta_h^{\widehat{u}} - \eta_h^u\rangle_{\partial\mathcal{T}_h} - \langle h_K^{-1}(\Pi_{k+1}^o\Phi - \Phi), \Pi_k^\partial\eta_h^u - \eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\\
&\quad + (\boldsymbol{p}(\Pi_{k+1}^o\Phi - \Phi), \nabla\eta_h^u)_{\mathcal{T}_h} - \langle\boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^\partial\Phi - \Phi), \eta_h^u - \eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h}\\
&\quad + (\nabla\cdot\boldsymbol{p}(\Pi_{k+1}^o\Phi - \Phi), \eta_h^u)_{\mathcal{T}_h}\\
&= \sum_{i=1}^{10} S_i.
\end{aligned}
$$

We estimate $\{S_i\}_{i=1}^{10}$ as follows (we omit some of the details):

$$
\begin{aligned}
S_1 &= -\langle \boldsymbol{\Pi}_k^o \boldsymbol{q}\cdot\boldsymbol{n} - \boldsymbol{q}\cdot\boldsymbol{n}, \Phi - \Pi_{k+1}^o\Phi\rangle_{\partial\mathcal{T}_h} \le Ch^{k+2}|\boldsymbol{q}|_{k+1}\|\Phi\|_2,\\
S_2 &= -\langle h_K^{-1}(\Pi_{k+1}^o u - u), \Pi_{k+1}^o\Phi - \Phi\rangle_{\partial\mathcal{T}_h} \le Ch^{k+2}|u|_{k+2}\|\Phi\|_2,\\
S_3 &= -(\boldsymbol{p}(\Pi_{k+1}^o u - u), \nabla\Pi_{k+1}^o\Phi)_{\mathcal{T}_h} \le Ch^{k+2}|u|_{k+2}|\Phi|_1,\\
S_4 &= \langle \boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^\partial u - u), \Pi_{k+1}^o\Phi - \Phi\rangle_{\partial\mathcal{T}_h} \le Ch^{k+2}|u|_{k+1}\|\Phi\|_2,\\
S_5 &\le C\|h_K^{-1/2}(\eta_h^u - \eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h} h|\Phi|_1 \le Ch^{k+2}|u|_{k+2}|\Phi|_1,\\
S_6 &\le Ch\|h_K^{-1/2}(\eta_h^u - \eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}\|\boldsymbol{\Psi}\|_1 \le Ch^{k+2}\|\boldsymbol{\Psi}\|_1,\\
S_7 &\le Ch\|h_K^{-1/2}(\eta_h^u - \eta_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}\|\boldsymbol{\Phi}\|_2 \le Ch^{k+2}|u|_{k+2}\|\boldsymbol{\Phi}\|_2,\\
S_8 &\le Ch^2\|\Phi\|_2\|\nabla\eta_h^u\|_{\mathcal{T}_h} \le Ch^{k+2}\|\Phi\|_2|u|_{k+2},\\
S_9 &= -\langle \boldsymbol{p}\cdot\boldsymbol{n}(\Pi_k^\partial\Phi - \Phi), \eta_h^u - \eta_h^{\widehat{u}}\rangle_{\partial\mathcal{T}_h} \le Ch^{k+2}|\Phi|_1|u|_{k+2},\\
S_{10} &\le Ch^2\|\Phi\|_2\|\eta_h^u\|_{\mathcal{T}_h}.
\end{aligned}
$$

Summing the above estimates, we get

$$\|\eta_h^u\|_{\mathcal{T}_h}^2 \le Ch^{k+2}\|u\|_{k+2}\|\eta_h^u\|_{\mathcal{T}_h} + Ch^2\|\eta_h^u\|_{\mathcal{T}_h}^2.$$

Let h be small enough, we have

$$\|\eta_h^u\|_{\mathcal{T}_h} \le Ch^{k+2}\|u\|_{k+2}.$$

A simple application of the triangle inequality finishes the proof. □

References

1. Baltes, H.P., Popovic, R.S.: Integrated semiconductor magnetic field sensors. Proc. IEEE **74**(8), 1107–1132 (1986)
2. Bank, R.E., Rose, D.J., Fichtner, W.: Numerical methods for semiconductor device simulation. SIAM J. Sci. Stat. Comput. **4**(3), 416–435 (1983)
3. Bessemoulin-Chatard, M., Chainais-Hillairet, C., Vignal, M.-H.: Study of a finite volume scheme for the drift-diffusion system. Asymptotic behavior in the quasi-neutral limit. SIAM J. Numer. Anal. **52**(4), 1666–1691 (2014)
4. Bessemoulin-Chatard, M.: A finite volume scheme for convection–diffusion equations with nonlinear diffusion derived from the Scharfetter–Gummel scheme. Numer. Math. **121**(4), 637–670 (2012)
5. Biler, P., Dolbeault, J.: Long time behavior of solutions of Nernst–Planck and Debye–Hückel drift-diffusion systems. Ann. Henri Poincaré **1**(3), 461–472 (2000)
6. Biler, P., Hebisch, W., Nadzieja, T.: The Debye system: existence and large time behavior of solutions. Nonlinear Anal. **23**(9), 1189–1209 (1994)
7. Bin, T., Chen, M., Xie, Y., Zhang, L., Eisenberg, B., Benzhuo, L.: A parallel finite element simulator for ion transport through three-dimensional ion channel systems. J. Comput. Chem. **34**(24), 2065–2078 (2013)
8. Brenner, S.C.: Poincaré–Friedrichs inequalities for piecewise H^1 functions. SIAM J. Numer. Anal. **41**(1), 306–324 (2003)
9. Brezzi, F., Marini, L.D., Pietra, P.: Two-dimensional exponential fitting and applications to drift–diffusion models. SIAM J. Numer. Anal. **26**(6), 1342–1355 (1989)
10. Burgler, J.F., Bank, R.E., Fichtner, W., Smith, R.K.: A new discretization scheme for the semiconductor current continuity equations. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. **8**(5), 479–489 (1989)
11. Cesmelioglu, A., Cockburn, B., Qiu, W.: Analysis of a hybridizable discontinuous Galerkin method for the steady-state incompressible Navier–Stokes equations. Math. Comp. **86**(306), 1643–1670 (2017)
12. Chainais-Hillairet, C., Liu, J.-G., Peng, Y.-J.: Finite volume scheme for multi-dimensional drift–diffusion equations and convergence analysis. M2AN Math. Model. Numer. Anal. **37**(2), 319–338 (2003)

13. Chainais-Hillairet, C., Peng, Y.-J.: Convergence of a finite-volume scheme for the drift-diffusion equations in 1D. IMA J. Numer. Anal. **23**(1), 81–108 (2003)
14. Chainais-Hillairet, C., Peng, Y.-J.: Finite volume approximation for degenerate drift–diffusion system in several space dimensions. Math. Models Methods Appl. Sci. **14**(3), 461–481 (2004)
15. Chen, G., Cockburn, B., Singler, J.R., Zhang, Y.: Superconvergent interpolatory HDG methods for reaction diffusion equations. Part I: HDG-k methods (In preparation)
16. Chen, G., Singler, J., Zhang, Y.: An HDG method For dirichlet boundary control of convection dominated diffusion PDEs. SIAM J. Numer. Anal
17. Chen, H., Li, J., Qiu, W.: Robust a posteriori error estimates for HDG method for convection–diffusion equations. IMA J. Numer. Anal. **36**(1), 437–462 (2016)
18. Chen, Y., Cockburn, B.: Analysis of variable-degree HDG methods for convection–diffusion equations. Part I: general nonconforming meshes. IMA J. Numer. Anal. **32**(4), 1267–1293 (2012)
19. Chen, Y., Cockburn, B.: Analysis of variable-degree HDG methods for convection–diffusion equations. Part II: semimatching nonconforming meshes. Math. Comp. **83**(285), 87–111 (2014)
20. Cockburn, B., Fu, G.: Superconvergence by M-decompositions. Part II: construction of two-dimensional finite elements. ESAIM Math. Model. Numer. Anal. **51**(1), 165–186 (2017)
21. Cockburn, B., Fu, G.: Superconvergence by M-decompositions. Part III: construction of three-dimensional finite elements. ESAIM Math. Model. Numer. Anal. **51**(1), 365–398 (2017)
22. Cockburn, B., Fu, G.: Devising superconvergent HDG methods with symmetric approximate stresses for linear elasticity by M-decompositions. IMA J. Numer. Anal. **38**(2), 566–604 (2018)
23. Cockburn, B., Fu, G., Qiu, W.: A note on the devising of superconvergent HDG methods for Stokes flow by M-decompositions. IMA J. Numer. Anal. **37**(2), 730–749 (2017)
24. Cockburn, B., Fu, G., Sayas, F.J.: Superconvergence by M-decompositions. Part I: general theory for HDG methods for diffusion. Math. Comp. **86**(306), 1609–1641 (2017)
25. Cockburn, B., Gopalakrishnan, J., Lazarov, R.: Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. SIAM J. Numer. Anal. **47**(2), 1319–1365 (2009)
26. Cockburn, B., Gopalakrishnan, J., Nguyen, N.C., Peraire, J., Sayas, F.-J.: Analysis of HDG methods for Stokes flow. Math. Comp. **80**(274), 723–760 (2011)
27. Cockburn, B., Gopalakrishnan, J., Sayas, F.-J.: A projection-based error analysis of HDG methods. Math. Comp. **79**(271), 1351–1367 (2010)
28. Datta, S.: Electronic Transport in Mesoscopic Systems. Cambridge University Press, Cambridge (1997)
29. De Mari, A.: An accurate numerical steady-state one-dimensional solution of the pn junction. Solid-State Electr. **11**(1), 33–58 (1968)
30. Eisenberg, B., Hyon, Y., Liu, C.: Energy variational analysis of ions in water and channels: field theory for primitive models of complex ionic fluids. J. Chem. Phys. **133**(10), 104104 (2010)
31. Flavell, A., Machen, M., Eisenberg, B., Kabre, J., Liu, C., Li, X.: A conservative finite difference scheme for Poisson–Nernst–Planck equations. J. Comput. Electr. **13**(1), 235–249 (2014)
32. Frank, F., Knabner, P.: Convergence analysis of a BDF2/mixed finite element discretization of a Darcy–Nernst–Planck–Poisson system. ESAIM Math. Model. Numer. Anal. **51**(5), 1883–1902 (2017)
33. Guosheng, F., Qiu, W., Zhang, W.: An analysis of HDG methods for convection–dominated diffusion problems. ESAIM Math. Model. Numer. Anal. **49**(1), 225–256 (2015)
34. Gajewski, H.: On existence, uniqueness and asymptotic behavior of solutions of the basic equations for carrier transport in semiconductors. Z. Angew. Math. Mech. **65**(2), 101–108 (1985)
35. Gajewski, H., Gröger, K.: On the basic equations for carrier transport in semiconductors. J. Math. Anal. Appl. **113**(1), 12–35 (1986)
36. Gao, H., He, D.: Linearized conservative finite element methods for the Nernst–Planck–Poisson equations. J. Sci. Comput. **72**(3), 1269–1289 (2017)
37. Gao, H., Sun, P.: A linearized local conservative mixed finite element method for Poisson-Nernst-Planck equations. J. Sci. Comput. **77**, 793 (2018)
38. Gong, W., Hu, W., Mateos, M., Singler, J.R., Zhang, Y.: An HDG method for tangential Dirichlet boundary control of stokes equations I: high regularity, Numerische Mathematik
39. Gummel, H.K.: A self-consistent iterative scheme for one-dimensional steady state transistor calculations. IEEE Trans. Electr. Dev. **11**(10), 455–465 (1964)
40. He, D., Pan, K.: An energy preserving finite difference scheme for the Poisson–Nernst–Planck system. Appl. Math. Comput. **287**(288), 214–223 (2016)
41. He, M., Sun, P.: Error analysis of mixed finite element method for Poisson–Nernst–Planck system. Numer. Methods Partial Differ. Equ. **33**(6), 1924–1948 (2017)
42. He, M., Sun, P.: Mixed finite element analysis for the Poisson–Nernst–Planck/Stokes coupling. J. Comput. Appl. Math. **341**, 61–79 (2018)

43. Horng, T.-L., Lin, T.-C., Liu, C., Eisenberg, B.: Pnp equations with steric effects: a model of ion flow through channels. J. Phys. Chem. B **116**(37), 11422–11441 (2012)
44. Hsieh, C.-Y., Hyon, Y.K., Lee, H., Lin, T.-C., Liu, C.: Transport of charged particles: entropy production and maximum dissipation principle. J. Math. Anal. Appl. **422**(1), 309–336 (2015)
45. Jerome, J.W.: Consistency of semiconductor modeling: an existence/stability analysis for the stationary Van Roosbroeck system. SIAM J. Appl. Math. **45**(4), 565–590 (1985)
46. Jerome, J.W.: Analysis of Charge Transport. A Mathematical Study of Semiconductor Devices. Springer, Berlin (1996)
47. Jerome, J.W.: Analysis of Charge Transport: A Mathematical Study of Semiconductor Devices. Springer, Berlin (2012)
48. Lehrenfeld, C.: Hybrid Discontinuous Galerkin methods for solving incompressible flow problems, PhD Thesis (2010)
49. Li, B., Xie, X.: Analysis of a family of HDG methods for second order elliptic problems. J. Comput. Appl. Math. **307**, 37–51 (2016)
50. Liu, H., Wang, Z.: A free energy satisfying finite difference method for Poisson–Nernst–Planck equations. J. Comput. Phys. **268**, 363–376 (2014)
51. Liu, Y.X., Shu, C.-W.: Analysis of the local discontinuous Galerkin method for the drift–diffusion model of semiconductor devices. Sci. China Math. **59**(1), 115–140 (2016)
52. Lu, B., Holst, M.J., McCammon, J.A., Zhou, Y.C.: Poisson–Nernst–Planck equations for simulating biomolecular diffusion–reaction processes I: finite element solutions. J. Comput. Phys. **229**(19), 6979–6994 (2010)
53. Markowich, P.A.: The Stationary Semiconductor Device Equations. Computational Microelectronics. Springer, Vienna (1986)
54. Meng, D., Zheng, B., Lin, G., Sushko, M.L.: Numerical solution of 3D Poisson–Nernst–Planck equations coupled with classical density functional theory for modeling ion and electron transport in a confined environment. Commun. Comput. Phys. **16**(5), 1298–1322 (2014)
55. Mirzadeh, M., Gibou, F.: A conservative discretization of the Poisson–Nernst–Planck equations on adaptive Cartesian grids. J. Comput. Phys. **274**, 633–653 (2014)
56. Mock, M.S.: An initial value problem from semiconductor device theory. SIAM J. Math. Anal. **5**, 597–612 (1974)
57. Oikawa, I.: A hybridized discontinuous Galerkin method with reduced stabilization. J. Sci. Comput. **65**(1), 327–340 (2015)
58. Qiu, W., Shen, J., Shi, K.: An HDG method for linear elasticity with strong symmetric stresses. Math. Comp. **87**(309), 69–93 (2018)
59. Qiu, W., Shi, K.: An HDG method for convection diffusion equation. J. Sci. Comput. **66**(1), 346–357 (2016)
60. Qiu, W., Shi, K.: A superconvergent HDG method for the incompressible Navier–Stokes equations on general polyhedral meshes. IMA J. Numer. Anal. **36**(4), 1943–1967 (2016)
61. Schmuck, M.: Analysis of the Navier–Stokes–Nernst–Planck–Poisson system. Math. Models Methods Appl. Sci. **19**(6), 993–1015 (2009)
62. Sun, Y., Sun, P., Zheng, B., Lin, G.: Error analysis of finite element method for Poisson–Nernst–Planck equations. J. Comput. Appl. Math. **301**, 28–43 (2016)
63. Wang, H., Shu, C.-W., Zhang, Q.: Stability and error estimates of local discontinuous Galerkin methods with implicit-explicit time-marching for advection–diffusion problems. SIAM J. Numer. Anal. **53**(1), 206–227 (2015)
64. Wu, J., Srinivasan, V., Xu, J., Wang, C.Y.: Newton–Krylov-multigrid algorithms for battery simulation. J. Electrochem. Soc. **149**(10), A1342–A1348 (2002)
65. Xu, S., Chen, M., Majd, S., Yue, X., Liu, C.: Modeling and simulating asymmetrical conductance changes in gramicidin pores. Mol. Based Math. Biol. **2**(1), 34–55 (2014)