

## UNIFORM-IN-TIME SUPERCONVERGENCE OF HDG METHODS FOR THE HEAT EQUATION

BRANDON CHABAUD AND BERNARDO COCKBURN

ABSTRACT. We prove that the superconvergence properties of the hybridizable discontinuous Galerkin method for second-order elliptic problems do hold uniformly in time for the semidiscretization by the same method of the heat equation provided the solution is smooth enough. Thus, if the approximations are piecewise polynomials of degree  $k$ , the approximation to the gradient converges with the rate  $h^{k+1}$  for  $k \geq 0$  and the  $L^2$ -projection of the error into a space of lower polynomial degree *superconverges* with the rate  $\sqrt{\log(T/h^2)} h^{k+2}$  for  $k \geq 1$  uniformly in time. As a consequence, an element-by-element postprocessing converges with the rate  $\sqrt{\log(T/h^2)} h^{k+2}$  for  $k \geq 1$  also uniformly in time. Similar results are proven for the Raviart-Thomas and the Brezzi-Douglas-Marini mixed methods.

### 1. INTRODUCTION

In this paper, we obtain uniform-in-time superconvergence error estimates for the semidiscretization by hybridizable discontinuous Galerkin (HDG) methods of the model parabolic equation

$$(1.1a) \quad u_t - \Delta u = f \quad \text{in } \Omega \times \mathbb{R}^+,$$

$$(1.1b) \quad u = g \quad \text{on } \partial\Omega \times \mathbb{R}^+,$$

$$(1.1c) \quad u(t=0) = u_0 \quad \text{on } \Omega,$$

where  $\Omega$  is a polyhedral domain of  $\mathbb{R}^n$ .

Let us describe our results. The particular HDG method we consider is the so-called local discontinuous Galerkin-hybridizable (LDG-H) method; see [6]. When applied to the Poisson equation, this method uses polynomials of degree  $k \geq 0$  to approximate  $u$  and each component of the flux  $\mathbf{q} = -\nabla u$ . The first analysis of a LDG-H method was obtained in [5]. Soon after, the analysis of a wide class of discontinuous Galerkin methods including the LDG-H methods was obtained in [8]. Recently, a new analysis of the LDG-H methods was proposed [7] which is based

---

Received by the editor January 27, 2010 and, in revised form, July 26, 2010.

2010 *Mathematics Subject Classification*. Primary 65M60, 35K05.

*Key words and phrases*. Discontinuous Galerkin methods, hybridization, superconvergence, parabolic problems.

Part of this work was done when the first author was at the School of Mathematics, University of Minnesota.

The second author was partially supported by the National Science Foundation (Grant DMS-0712955) and by the Minnesota Supercomputing Institute. Part of this work was done when this author was visiting the Research Institute for Mathematical Sciences, Kyoto University, Japan, during the Fall of 2009.

on the use of suitably chosen projections. In all of the above-mentioned papers, it was shown that the LDG-H methods have convergence properties shared with the Raviart-Thomas (RT) [17] and the Brezzi-Douglas-Marini (BDM) [2] methods. Indeed, it was shown that the approximation to  $\mathbf{q}$  converges with order  $k + 1$  for  $k \geq 0$  and that the  $L^2$ -projection of the error in  $u$  into a space of lower polynomial degree *superconverges* with order  $k + 2$  for  $k \geq 1$ . As a consequence, an element-by-element postprocessing could be computed which converged to  $u$  with order  $k + 2$  for  $k \geq 1$ . In this paper, we show that these results also hold in our setting *uniformly in time* provided the exact solution is smooth enough; though the estimate of the error of the postprocessed approximation has the additional factor  $\sqrt{\log(T/h^2)}$ .

Let us place our results in historical context. In 1981, [16], the first error estimates for the semidiscrete RT method of arbitrary degree on polygonal domains and for the semidiscrete RT method of degree  $k = 1$  for smooth domains in  $\mathbb{R}^2$  were obtained. In 1987, [18], the semidiscrete RT methods of arbitrary degree were studied in  $\mathbb{R}^2$  and optimal error estimates in negative-order norms were obtained by using the so-called quasiprojection method introduced in [9]. Superconvergent approximations were then obtained (for translation invariant meshes and periodic boundary conditions) by convolution as suggested in [1]. These were the first superconvergence estimates for semidiscrete mixed methods. In 1989, [22], superconvergence was obtained for the gradient of the approximation provided by the continuous Galerkin method with piecewise-linear approximations; nonsmooth initial conditions  $u_0$  were considered. In 1993, [13], superconvergence of the flux  $\mathbf{q}$  along the Gaussian lines for rectangular RT methods was proven for smooth solutions. In 1998, [4], the results of the pioneering work done in [16] for the RT method of degree  $k = 1$  for the homogeneous equation with nonsmooth initial data  $u_0$  and smooth domains in  $\mathbb{R}^2$  were extended to the case of RT methods of arbitrary degree  $k$ . The extension of the results in [13] to this setting was achieved at the same time. In 2009, [23], superconvergence of the so-called semidiscrete  $H^1$ -Galerkin mixed method was obtained for rectangular elements.

There are some similarities between our aim and approach and those in [13, 4]. Indeed, in [13, 4], the superconvergence results for mixed methods for second-order elliptic problems obtained in [10, 11, 14] were extended to the semidiscrete mixed methods. In a similar manner, here we extend to our parabolic setting the superconvergence results for the discretization by HDG methods of second-order elliptic problems [5, 8, 7].

In [13, 4], estimates of a suitably defined projection of the errors are obtained by using *energy* techniques and, in [4], also by *parabolic duality* arguments. Similarly, we consider the projection introduced in [7] and proceed to estimate the projection of the errors. Moreover, we also use energy techniques and parabolic duality arguments. The energy techniques are very well known and can be found in the monograph [21]. The parabolic duality technique was also used in 1991, [12], in the framework of adaptive methods. Here, we use a variation of the duality arguments used in [12] and [4]. It is based on an estimate of the  $L^1(0, T; L^2(\Omega))$ -norm of the solution of the dual problem and incorporates the fact that the projection of the error we are trying to estimate lies in a finite dimensional space.

Finally, note that, since the results we obtain here for the HDG methods easily carry over to mixed methods, our approach provides a new way to obtain superconvergence results for those methods in the case of smooth solutions. To the

knowledge of the authors, the closest results in the available literature are those obtained in [4]. However, note that here we are considering polyhedral domains in any space dimension and inhomogeneous equations, whereas in [4] the domains are planar and smooth, and the equations are homogeneous.

The paper is organized as follows. In section 2, we state and discuss our main results for the HDG method under consideration. Their detailed proof is displayed in sections 3, 4 and also in the Appendix, where some auxiliary results are obtained. We end in section 5 by discussing extensions of our results to the RT and BDM mixed methods.

## 2. MAIN RESULTS

**2.1. The HDG method.** To describe the HDG method under consideration, we begin by discretizing the domain  $\Omega$  by a triangulation  $\mathcal{T}_h$  made of simplexes. For simplicity, we take the triangulation to be conforming. We denote by  $\mathcal{E}_h$  the union of faces  $F$  of the simplexes  $K$  of the triangulation  $\mathcal{T}_h$ .

For each time  $t$  on the interval  $[0, T]$ , the method yields a scalar approximation  $u_h(t)$  to  $u(t)$ , a vector approximation  $\mathbf{q}_h(t)$  to  $\mathbf{q}(t)$ , and a scalar approximation  $\widehat{u}_h(t)$  to the trace of  $u(t)$  on element boundaries, in spaces of the form

$$(2.1a) \quad W_h = \{w \in L^2(\Omega) : w|_K \in W(K) \quad \forall K \in \mathcal{T}_h\},$$

$$(2.1b) \quad \mathbf{V}_h = \{\mathbf{v} \in [L^2(\Omega)]^n : \mathbf{v}|_K \in \mathbf{V}(K) \quad \forall K \in \mathcal{T}_h\},$$

$$(2.1c) \quad M_h = \{\mu \in L^2(\mathcal{E}_h) : \mu|_F \in M(F) \quad \forall F \in \mathcal{E}_h\},$$

respectively, where

$$W(K) = \mathcal{P}_k(K), \quad \mathbf{V}(K) = \mathcal{P}_k(K), \quad \text{and} \quad M(F) = \mathcal{P}_k(F).$$

Here  $\mathcal{P}_k(K) := [\mathcal{P}_k(K)]^n$  and  $\mathcal{P}_k(K)$  is the space of polynomials of total degree at most  $k$ .

The HDG method provides approximations  $u_h$  in  $W_h$ ,  $\mathbf{q}_h$  in  $\mathbf{V}_h$ , and the numerical trace  $\widehat{u}_h$  in  $M_h$ , which are determined by requiring that

$$(2.2a) \quad (\mathbf{q}_h, \mathbf{r})_{\mathcal{T}_h} - (u_h, \nabla \cdot \mathbf{r})_{\mathcal{T}_h} + \langle \widehat{u}_h, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} = 0,$$

$$(2.2b) \quad (\partial_t u_h, w)_{\mathcal{T}_h} - (\mathbf{q}_h, \nabla w)_{\mathcal{T}_h} + \langle \widehat{\mathbf{q}}_h \cdot \mathbf{n}, w \rangle_{\partial \mathcal{T}_h} = (f, w)_{\mathcal{T}_h},$$

$$(2.2c) \quad \langle \widehat{u}_h, \mu \rangle_{\partial \Omega} = \langle g, \mu \rangle_{\partial \Omega},$$

$$(2.2d) \quad \langle \widehat{\mathbf{q}}_h \cdot \mathbf{n}, \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} = 0,$$

$$(2.2e) \quad u_h|_{t=0} = \Pi_W u_0,$$

hold for all  $\mathbf{r} \in \mathbf{V}_h$ ,  $w \in W_h$ , and  $\mu \in M_h$ , with a numerical trace for the flux defined by

$$(2.2f) \quad \widehat{\mathbf{q}}_h = \mathbf{q}_h + \tau (u_h - \widehat{u}_h) \mathbf{n}, \quad \text{on } \partial \mathcal{T}_h,$$

for some nonnegative stabilization function  $\tau$  defined on  $\partial \mathcal{T}_h$ , which we assume to be piecewise constant on  $\partial \mathcal{T}_h$ . As explained in [6], these methods are called the LDG-hybridizable (LDG-H) methods because the above numerical trace is that of the LDG method applied separately on each mesh element. The projection operator  $\Pi_W$  is the one introduced in [7]; it will be defined later.

Above and throughout, we use the notation

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

where we write  $(u, v)_D = \int_D uv \, dx$  whenever  $D$  is a domain of  $\mathbb{R}^n$ , and  $\langle u, v \rangle_D = \int_D uv \, dx$  whenever  $D$  is a domain of  $\mathbb{R}^{n-1}$ . For vector functions  $\mathbf{v}$  and  $\mathbf{w}$ , the notations are similarly defined with the integrand being the dot product  $\mathbf{v} \cdot \mathbf{w}$ . This completes the definition of the method.

**The postprocessing.** Finally, as done in [5, 8, 7], we describe the postprocessing to compute the new approximation  $u_h^*$ ; we follow [15, 19, 20]. Let  $\mathcal{W}_{k+1}(K)$  denote the  $L^2(K)$ -orthogonal complement of  $\mathcal{P}_{k-1}(K)$  in  $\mathcal{P}_{k+1}(K)$ . Then on the simplex  $K \in \mathcal{T}_h$ , we take  $u_h^*$  as the element of  $\mathcal{P}_{k+1}(K)$  defined as the only solution of

$$(2.3a) \quad (u_h^*, w)_K = (u_h, w)_K \quad \text{for all } w \in \mathcal{P}_{k-1}(K),$$

$$(2.3b) \quad (\nabla u_h^*, \nabla w)_K = -(\mathbf{q}_h, \nabla w)_K \quad \text{for all } w \in \mathcal{W}_{k+1}(K).$$

**2.2. Estimates of the projection of the errors.** Here we state and discuss our main result, namely, upper bounds for each component of the projection  $\Pi_h$  of the errors:

$$(\varepsilon_h^q, \varepsilon_h^u) := \Pi_h(\mathbf{q} - \mathbf{q}_h, u - u_h).$$

The projection  $\Pi_h$  was introduced in [7] for the study of HDG methods for elliptic problems. Let us describe it.

**The projection.** The projection  $\Pi_h$  into  $\mathbf{V}_h \times W_h$  was introduced in [7]; it is defined as follows. Given  $(\mathbf{q}, u) \in \mathbf{H}^1(\mathcal{T}_h) \times H^1(\mathcal{T}_h)$ , where

$$H^1(\mathcal{T}_h) = \prod_{K \in \mathcal{T}_h} H^1(K), \quad \mathbf{H}^1(\mathcal{T}_h) = H^1(\mathcal{T}_h)^n,$$

the function  $\Pi_h(\mathbf{q}, u) = (\Pi_V \mathbf{q}, \Pi_W u)$  on an arbitrary simplex  $K \in \mathcal{T}_h$  is the element of  $\mathbf{V}_h \times W_h$  which solves

$$(2.4a) \quad (\Pi_V \mathbf{q}, \mathbf{v})_K = (\mathbf{q}, \mathbf{v})_K \quad \text{for all } \mathbf{v} \in \mathcal{P}_{k-1}(K),$$

$$(2.4b) \quad (\Pi_W u, w)_K = (u, w)_K \quad \text{for all } w \in \mathcal{P}_{k-1}(K),$$

$$(2.4c) \quad \langle \Pi_V \mathbf{q} \cdot \mathbf{n} + \tau \Pi_W u, \mu \rangle_F = \langle \mathbf{q} \cdot \mathbf{n} + \tau u, \mu \rangle_F \quad \text{for all } \mu \in \mathcal{P}_k(F),$$

for all faces  $F$  of the simplex  $K$ .

Note that  $\Pi_h(\mathbf{q}_h, u_h) = (\mathbf{q}_h, u_h)$  and so  $(\varepsilon_h^q, \varepsilon_h^u) := (\Pi_V \mathbf{q} - \mathbf{q}_h, \Pi_W u - u_h)$ .

**Norms.** We use  $\|\cdot\|_D$  to denote the  $L^2(D)$ -norm for any  $D$ . If  $D = \Omega$ , we simply write  $\|\cdot\|$ . We denote the norm and seminorm on any Sobolev space  $X$  by  $\|\cdot\|_X$  and  $|\cdot|_X$ , respectively. We also denote  $\|\cdot\|_{X(0,T;Y(\Omega))}$  by  $\|\cdot\|_{X(Y)}$ , and set

$$\begin{aligned} \|(\mathbf{q}, u)\|_{1,T,\Omega} &:= \|\mathbf{q}\|_{L^2(L^2)} + \|u_t\|_{L^1(L^2)}, \\ \|(\mathbf{q}, u)\|_{2,T,\Omega} &:= \|\mathbf{q}(0)\| + \|\mathbf{q}_t\|_{L^1(L^2)} + \|u_t\|_{L^2(L^2)}. \end{aligned}$$

**Elliptic regularity.** For some of the estimates, we are going to assume that the domain  $\Omega$  is such that for any function  $\Psi \in H_0^1(\Omega)$  we have the elliptic regularity inequality

$$(2.5) \quad \|\Psi\|_{H^2(\Omega)} \leq C_{\text{reg}} \|\Delta\Psi\|.$$

It is well known that this holds whenever  $\Omega$  is a convex polyhedral domain.

When we use this assumption, we are also going to use some quantities we define next. For each simplex  $K \in \mathcal{T}_h$ ,  $\rho_K$  denotes the radius of the largest ball included in the simplex  $K$ . We set  $\rho := \min_{K \in \mathcal{T}_h} \rho_K$ . We also set  $\kappa > 1$  to be the solution of

$$(2.6) \quad \kappa \log \kappa = T/\rho^2.$$

In some of our estimates, the factor  $\sqrt{\log \kappa}$  appears. Note, however, that this factor is not big. For example, if  $T/\rho^2 \approx 10^{100}$ , we have that  $\sqrt{\log \kappa} \approx 15$ . For quasiuniform triangulations,  $\rho \approx h$  and so, when  $T \approx 1$ , this means that  $h \approx 10^{-50}$ , which would correspond to an extremely refined triangulation of  $\Omega$ .

Finally, we introduce an auxiliary projection. We denote by  $P_\ell$  the  $L^2$ -projection into the space  $W_{h,\ell}$  defined as

$$W_{h,\ell} := \{w \in L^2(\Omega) : w|_K \in \mathcal{P}_\ell(K) \quad \forall K \in \mathcal{T}_h\}.$$

We are now ready to state our main result.

**Theorem 2.1.** *For any  $T > 0$  and any  $k \geq 0$ , we have*

$$\begin{aligned} \|\varepsilon_h^u\|_{L^\infty(L^2)} &\leq C \|\mathbf{q}, u - \Pi_h(\mathbf{q}, u)\|_{1,T,\Omega}, \\ \|\varepsilon_h^q\|_{L^\infty(L^2)} &\leq C \|\mathbf{q}, u - \Pi_h(\mathbf{q}, u)\|_{2,T,\Omega}. \end{aligned}$$

Moreover, if the elliptic regularity inequality (2.5) holds and  $k \geq 1$ , we have

$$\|P_{k-1}\varepsilon_h^u\|_{L^\infty(L^2)} \leq C C_{\text{reg}} \sqrt{\log \kappa} h \|\mathbf{q}, u - \Pi_h(\mathbf{q}, u)\|_{2,T,\Omega}.$$

Here  $C$  is a constant independent of the exact solution and of the discretization parameters.

This result, which is proven in section 3, extends to our parabolic setting, similar results for HDG methods for second-order elliptic problems [7]. Note the appearance of the factor  $\sqrt{\log \kappa}$  as a consequence of the parabolic duality argument we use in our approach. Note also that the convergence properties of the postprocessing  $u_h^*$  are determined by those of  $P_{k-1}u_h$  and  $\mathbf{q}_h$ , as can be seen from the equations (2.3) defining it. This fact is going to be fully exploited in the next corollary. For the sake of completeness, we provide another  $L^\infty(L^2)$  estimate of  $\varepsilon_h^u$  in section 4. It is of the same form of that of  $P_{k-1}\varepsilon_h^u$  in our main result, but its constant  $C$  is more sensitive to variations of the stabilization function  $\tau$ ; see Theorem 4.1.

An immediate consequence of this result is the following.

**Corollary 2.1.** *For any  $T > 0$  and any  $k \geq 0$ , we have*

$$\begin{aligned} \|u - u_h\|_{L^\infty(L^2)} &\leq \|\Pi_w u - u\|_{L^\infty(L^2)} + C \|\mathbf{q}, u - \Pi_h(\mathbf{q}, u)\|_{1,T,\Omega}, \\ \|\mathbf{q} - \mathbf{q}_h\|_{L^\infty(L^2)} &\leq \|\mathbf{q} - \Pi_v \mathbf{q}\|_{L^\infty(L^2)} + C \|\mathbf{q}, u - \Pi_h(\mathbf{q}, u)\|_{2,T,\Omega}. \end{aligned}$$

Moreover, if the elliptic regularity inequality (2.5) holds and  $k \geq 1$ , we have

$$\begin{aligned} \|u - u_h^*\|_{L^\infty(L^2)} &\leq C h^{k+2} |u|_{L^\infty(0,T;H^{k+2}(\Omega))} \\ &\quad + C C_{\text{reg}} \sqrt{\log \kappa} h \|(\mathbf{q}, u) - \Pi_h(\mathbf{q}, u)\|_{2,T,\Omega}, \end{aligned}$$

where  $C$  is a constant independent of the exact solution and of the discretization parameters.

To obtain actual rates of convergence, we only have to combine this result with the following approximation result.

**Theorem 2.2** ([7]). *Suppose  $k \geq 0$ ,  $\tau|_{\partial K}$  is nonnegative and  $\tau_K^{\max} := \max \tau|_{\partial K} > 0$ . Then the system (2.4) is uniquely solvable for  $\Pi_V \mathbf{q}$  and  $\Pi_W u$ . Furthermore, there is a constant  $C$  independent of  $K$  and  $\tau$  such that*

$$\begin{aligned} \|\Pi_V \mathbf{q} - \mathbf{q}\|_K &\leq C h_K^{\ell_q+1} |\mathbf{q}|_{H^{\ell_q+1}(K)} + C h_K^{\ell_u+1} \tau_K^* |u|_{H^{\ell_u+1}(K)}, \\ \|\Pi_W u - u\|_K &\leq C h_K^{\ell_u+1} |u|_{H^{\ell_u+1}(K)} + C \frac{h_K^{\ell_q+1}}{\tau_K^{\max}} |\nabla \cdot \mathbf{q}|_{H^{\ell_q}(K)}, \end{aligned}$$

for  $\ell_u, \ell_q$  in  $[0, k]$ . Here  $\tau_K^* := \max \tau|_{\partial K \setminus F^*}$ , where  $F^*$  is a face of  $K$  at which  $\tau|_{\partial K}$  is maximum.

In particular, we can see that if the stabilization function  $\tau$  is uniformly bounded, the approximation error of the projection is of order  $k+1$  in both variables. Hence, Corollary 2.1 implies that, if the exact solution is smooth enough, the HDG approximation  $(\mathbf{q}_h, u_h)$  converges with order  $k+1$  for any  $k \geq 0$  uniformly in time. Moreover, if  $k \geq 1$  and the domain  $\Omega$  satisfies the elliptic regularity condition (2.5), then  $u_h^*$  converges with the rate of  $\sqrt{\log \kappa} h^{k+2}$  uniformly in time.

### 3. PROOFS

In this section, we provide a detailed proof of the uniform-in-time estimates of the projection of the errors of Theorem 2.1. We proceed in several steps. In the first step, we obtain the equations satisfied by the projection of the errors; the orthogonality properties satisfied by the projection  $\Pi_h$  play a crucial role in rendering this set of equations suitable for the analysis. The following two steps are devoted to getting the first two estimates by using standard parabolic energy arguments. In steps 4 and 5, we prove the remaining estimate by using a parabolic duality argument.

**Step 1: The equations of the projection of the errors.** The projection of the errors satisfy the following equations. They are stated in terms of the following projection of the error at the boundaries of the elements

$$\varepsilon_h^{\hat{u}} := P_M(u - \hat{u}_h) = P_M u - \hat{u}_h,$$

where  $P_M$  denotes the  $L^2$ -orthogonal projection onto  $M_h$ .

**Lemma 3.1.** *We have*

$$(3.1a) \quad (\boldsymbol{\varepsilon}_h^q, \mathbf{r})_{\mathcal{T}_h} - (\varepsilon_h^u, \nabla \cdot \mathbf{r})_{\mathcal{T}_h} + \langle \varepsilon_h^{\widehat{u}}, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} = (\mathbf{II}_V \mathbf{q} - \mathbf{q}, \mathbf{r})_{\mathcal{T}_h},$$

$$(3.1b) \quad (\partial_t \varepsilon_h^u, w)_{\mathcal{T}_h} - (\boldsymbol{\varepsilon}_h^q, \nabla w)_{\mathcal{T}_h} + \langle \boldsymbol{\varepsilon}_h^{\widehat{q}}, \mathbf{n}, w \rangle_{\partial \mathcal{T}_h} = (\mathbf{II}_W u_t - u_t, w)_{\mathcal{T}_h},$$

$$(3.1c) \quad \langle \varepsilon_h^{\widehat{u}}, \mu \rangle_{\partial \Omega} = 0,$$

$$(3.1d) \quad \langle \boldsymbol{\varepsilon}_h^{\widehat{q}} \cdot \mathbf{n}, \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} = 0,$$

$$(3.1e) \quad \varepsilon_h^u|_{t=0} = 0,$$

for all  $\mathbf{r} \in \mathbf{V}_h$ ,  $w \in W_h$ , and  $\mu \in M_h$ , where

$$(3.1f) \quad \boldsymbol{\varepsilon}_h^{\widehat{q}} \cdot \mathbf{n} := \boldsymbol{\varepsilon}_h^q \cdot \mathbf{n} + \tau(\varepsilon_h^u - \varepsilon_h^{\widehat{u}}) \quad \text{on } \partial \mathcal{T}_h.$$

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

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

for all  $\mathbf{r} \in \mathbf{V}_h$  and  $w \in W_h$ . By the orthogonality properties (2.4a) and (2.4b) of  $\mathbf{II}_V$  and  $\mathbf{II}_W$ , respectively, and since  $P_M$  is the  $L^2$ -projection into  $M_h$  and satisfies the orthogonality property

$$(3.2) \quad \langle \tau(P_M u - u), \mu \rangle_{\partial \mathcal{T}_h} = 0 \quad \text{for all } \mu \in M_h$$

because  $\tau$  is piecewise constant on  $\partial \mathcal{T}_h$ , this implies that

$$\begin{aligned} (\mathbf{II}_V \mathbf{q}, \mathbf{r})_{\mathcal{T}_h} - (\mathbf{II}_W u, \nabla \cdot \mathbf{r})_{\mathcal{T}_h} + \langle P_M u, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} &= (\mathbf{II}_V \mathbf{q} - \mathbf{q}, \mathbf{r})_{\mathcal{T}_h}, \\ (\mathbf{II}_W u_t, w)_{\mathcal{T}_h} - (\mathbf{II}_V \mathbf{q}, \nabla w)_{\mathcal{T}_h} &+ \langle \mathbf{II}_V \mathbf{q} \cdot \mathbf{n} + \tau(\mathbf{II}_W u - P_M u), w \rangle_{\partial \mathcal{T}_h} = (f + \mathbf{II}_W u_t - u_t, w)_{\mathcal{T}_h}, \end{aligned}$$

for all  $\mathbf{r} \in \mathbf{V}_h$  and  $w \in W_h$ . Subtracting the first two equations defining the HDG method, (2.2a) and (2.2b), from the above two equations, respectively, we readily obtain (3.1a) and (3.1b).

The equation (3.1c) follows directly from the equation (2.2c) defining the HDG method and the boundary condition on the exact solution.

To prove (3.1d) we proceed as follows. By the definition of  $\boldsymbol{\varepsilon}_h^{\widehat{q}}$ , (3.1f),

$$\langle \mu, \boldsymbol{\varepsilon}_h^{\widehat{q}} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} = \langle (\mathbf{II}_V \mathbf{q} - \mathbf{q}_h) \cdot \mathbf{n} + \tau(\mathbf{II}_W u - u_h - P_M u + \widehat{u}_h), \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega}.$$

Hence, by the orthogonality property of the projection  $(\mathbf{II}_V, \mathbf{II}_W)$ , (2.4c), and property (3.2) of the projection  $P_M$ , we obtain that

$$\begin{aligned} \langle \mu, \boldsymbol{\varepsilon}_h^{\widehat{q}} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} &= \langle (\mathbf{q} - \mathbf{q}_h) \cdot \mathbf{n} + \tau(u - u_h - u + \widehat{u}_h), \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} \\ &= \langle \mathbf{q} \cdot \mathbf{n}, \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} - \langle \widehat{\mathbf{q}}_h \cdot \mathbf{n}, \mu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega}, \end{aligned}$$

and the equation (3.1d) follows since both of the above terms are zero. Indeed, the first is equal to zero because  $\mathbf{q}$  is in  $H(\text{div}, \Omega)$  and the second because the normal component of  $\widehat{\mathbf{q}}_h$  is single valued by the equation (2.2d) defining the HDG method.

It remains to prove the equation (3.1e). But by the equation (2.2e) defining the HDG method,  $u_h|_{t=0} = \mathbf{II}_W u_0$ , and so

$$\varepsilon_h^u|_{t=0} = \mathbf{II}_W u_0 - u_h|_{t=0} = \mathbf{II}_W u_0 - \mathbf{II}_W u_0 = 0.$$

This completes the proof.  $\square$

**Step 2: Estimate of  $\varepsilon_h^u$  in  $L^\infty(L^2)$  by an energy argument.**

**Lemma 3.2.** *For any  $T > 0$ , we have*

$$\begin{aligned} & \frac{1}{2} \|\varepsilon_h^u(T)\|^2 + \int_0^T [\|\varepsilon_h^q\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)\|_{\partial\mathcal{T}_h}^2] \\ &= \int_0^T [(\mathbf{II}_V \mathbf{q} - \mathbf{q}, \varepsilon_h^q)_{\mathcal{T}_h} + (\mathbf{II}_W u_t - u_t, \varepsilon_h^u)_{\mathcal{T}_h}]. \end{aligned}$$

*Proof.* Taking  $\mathbf{r} := \varepsilon_h^q$  in (3.1a),  $w = \varepsilon_h^u$  in (3.1b),  $\mu = -\widehat{\varepsilon}_h^q \cdot \mathbf{n}$  in (3.1c) and  $\mu = -\widehat{\varepsilon}_h^u$  in (3.1d), and adding the resulting four equations, we get

$$\|\varepsilon_h^q\|^2 + \frac{1}{2} \frac{d}{dt} \|\varepsilon_h^u\|^2 + \Theta_h = (\mathbf{II}_V \mathbf{q} - \mathbf{q}, \varepsilon_h^q)_{\mathcal{T}_h} + (\mathbf{II}_W u_t - u_t, \varepsilon_h^u)_{\mathcal{T}_h},$$

where

$$\begin{aligned} \Theta_h &= -(\varepsilon_h^u, \nabla \cdot \varepsilon_h^q)_{\mathcal{T}_h} + \langle \widehat{\varepsilon}_h^u, \varepsilon_h^q \cdot \mathbf{n} \rangle_{\partial\mathcal{T}_h} - (\varepsilon_h^q, \nabla \varepsilon_h^u)_{\mathcal{T}_h} \\ &\quad + \langle \widehat{\varepsilon}_h^q \cdot \mathbf{n}, \varepsilon_h^u \rangle_{\partial\mathcal{T}_h} - \langle \widehat{\varepsilon}_h^u, \widehat{\varepsilon}_h^q \cdot \mathbf{n} \rangle_{\partial\mathcal{T}_h}. \end{aligned}$$

But, by integration by parts,

$$\begin{aligned} \Theta_h &= -\langle \varepsilon_h^u, \varepsilon_h^q \cdot \mathbf{n} \rangle_{\partial\mathcal{T}_h} + \langle \widehat{\varepsilon}_h^u, \varepsilon_h^q \cdot \mathbf{n} \rangle_{\partial\mathcal{T}_h} + \langle \widehat{\varepsilon}_h^q \cdot \mathbf{n}, \varepsilon_h^u \rangle_{\partial\mathcal{T}_h} - \langle \varepsilon_h^u, \widehat{\varepsilon}_h^q \cdot \mathbf{n} \rangle_{\partial\mathcal{T}_h} \\ &= \langle \widehat{\varepsilon}_h^q \cdot \mathbf{n} - \varepsilon_h^q \cdot \mathbf{n}, \varepsilon_h^u - \widehat{\varepsilon}_h^u \rangle_{\mathcal{T}_h} = \|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)\|_{\partial\mathcal{T}_h}^2, \end{aligned}$$

by the definition of  $\widehat{\varepsilon}_h^q$ , (3.1f). The identity we want to prove follows after integrating in time over the interval  $(0, T)$  and using the fact that  $\varepsilon_h^u(0) = 0$  by (3.1e). This completes the proof.  $\square$

**Corollary 3.1.** *For any  $T > 0$ , we have*

$$\begin{aligned} & \|\varepsilon_h^u(T)\|^2 + \int_0^T [\|\varepsilon_h^q\|^2 + 2\|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)\|_{\partial\mathcal{T}_h}^2] \\ & \leq [\|\mathbf{II}_V \mathbf{q} - \mathbf{q}\|_{L^2(L^2)} + \|\mathbf{II}_W u_t - u_t\|_{L^1(L^2)}]^2. \end{aligned}$$

In this manner, the first estimate of Theorem 2.1 is obtained. To prove this corollary, we are going to use the following auxiliary result. Its proof is detailed in the Appendix; see subsection A.1.

**Proposition 3.1.** *Assume that, for all  $t > 0$ , we have*

$$\zeta^2(t) + \int_0^t Z(s) ds \leq A(t) + \int_0^t B(s) \zeta(s) ds,$$

for some nonnegative functions  $A, B$  in  $L^\infty(\mathbb{R}^+)$ . Then, for any  $T > 0$ ,

$$\zeta^2(T) + \int_0^T Z(s) ds \leq \left( \left[ \max_{0 \leq t \leq T} A(t) \right]^{1/2} + \frac{1}{2} \int_0^T B(s) ds \right)^2.$$

*Proof of Corollary 3.1.* Applying Cauchy-Schwarz inequality to each term of the right-hand side of the identity of Lemma 3.2, we get

$$\begin{aligned} & \frac{1}{2} \|\varepsilon_h^u(T)\|^2 + \int_0^T [\|\varepsilon_h^q\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)\|_{\partial\mathcal{T}_h}^2] \\ & \leq \int_0^T [\|\mathbf{II}_V \mathbf{q} - \mathbf{q}\| \|\varepsilon_h^q\| + \|\mathbf{II}_W u_t - u_t\| \|\varepsilon_h^u\|], \end{aligned}$$



and so

$$\begin{aligned} & \frac{1}{2} \|\varepsilon_h^u(T)\|^2 + \int_0^T \left[ \frac{1}{2} \|\varepsilon_h^q\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \varepsilon_h^{\widehat{u}})\|_{\partial\mathcal{T}_h}^2 \right] \\ & \leq \frac{1}{2} \|\mathbf{II}_V \mathbf{q} - \mathbf{q}\|_{L^2(L^2)}^2 + \int_0^T \|\mathbf{II}_W u_t - u_t\| \|\varepsilon_h^u\|. \end{aligned}$$

The estimate now readily follows by using Proposition 3.1. This completes the proof.  $\square$

**Step 3: Estimate of  $\varepsilon_h^q$  in  $L^\infty(L^2)$  by an energy argument.**

**Lemma 3.3.** *For any  $T > 0$ , we have*

$$\begin{aligned} & \frac{1}{2} [\|\varepsilon_h^q(T)\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \varepsilon_h^{\widehat{u}})(T)\|_{\partial\mathcal{T}_h}^2] + \int_0^T \|\partial_t \varepsilon_h^u\|^2 \\ & = \frac{1}{2} [\|\varepsilon_h^q(0)\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \varepsilon_h^{\widehat{u}})(0)\|_{\partial\mathcal{T}_h}^2] \\ & \quad + \int_0^T [(\mathbf{II}_V \mathbf{q}_t - \mathbf{q}_t, \varepsilon_h^q)_{\mathcal{T}_h} + (\mathbf{II}_W u_t - u_t, \partial_t \varepsilon_h^u)_{\mathcal{T}_h}]. \end{aligned}$$

*Proof.* To prove this result, we use a slightly different set of equations for the projection of the errors than the ones displayed in Lemma 3.1. We keep all of those equations except for (3.1a) and (3.1c), which are replaced by the equations obtained by differentiating them with respect to time:

$$\begin{aligned} (\partial_t \varepsilon_h^q, \mathbf{r})_{\mathcal{T}_h} - (\partial_t \varepsilon_h^u, \nabla \cdot \mathbf{r})_{\mathcal{T}_h} + \langle \partial_t \varepsilon_h^{\widehat{u}}, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial\mathcal{T}_h} &= (\mathbf{II}_V \mathbf{q}_t - \mathbf{q}_t, \mathbf{r})_{\mathcal{T}_h}, \\ (\partial_t \varepsilon_h^u, w)_{\mathcal{T}_h} - (\varepsilon_h^q, \nabla w)_{\mathcal{T}_h} + \langle \widehat{\varepsilon}_h^q \cdot \mathbf{n}, w \rangle_{\partial\mathcal{T}_h} &= (\mathbf{II}_W u_t - u_t, w)_{\mathcal{T}_h}, \\ \langle \partial_t \varepsilon_h^{\widehat{u}}, \mu \rangle_{\partial\Omega} &= 0, \\ \langle \widehat{\varepsilon}_h^q \cdot \mathbf{n}, \mu \rangle_{\partial\mathcal{T}_h \setminus \partial\Omega} &= 0, \\ \varepsilon_h^u|_{t=0} &= 0, \end{aligned}$$

for all  $\mathbf{r} \in \mathbf{V}_h$ ,  $w \in W_h$ , and  $\mu \in M_h$ , where  $\widehat{\varepsilon}_h^q \cdot \mathbf{n} := \varepsilon_h^q \cdot \mathbf{n} + \tau(\varepsilon_h^u - \varepsilon_h^{\widehat{u}})$  on  $\partial\mathcal{T}_h$ .

To prove the identity, just take  $\mathbf{r} := \varepsilon_h^q$ ,  $w = \partial_t \varepsilon_h^u$ ,  $\mu = -\widehat{\varepsilon}_h^q \cdot \mathbf{n}$ , and  $\mu = -\partial_t \varepsilon_h^{\widehat{u}}$  in the above equations, add the resulting four equations, and proceed exactly as in the proof of Lemma 3.2. This completes the proof.  $\square$

**Corollary 3.2.** *For any  $T > 0$ , we have*

$$\begin{aligned} & \|\varepsilon_h^q(T)\|_{\mathcal{T}_h}^2 + \|\sqrt{\tau}(\varepsilon_h^u - \varepsilon_h^{\widehat{u}})(T)\|_{\partial\mathcal{T}_h}^2 + \int_0^T \|\partial_t \varepsilon_h^u\|^2 \\ & \leq [ \|(\mathbf{II}_V \mathbf{q} - \mathbf{q})(0)\| + \|\mathbf{II}_V \mathbf{q}_t - \mathbf{q}_t\|_{L^1(L^2)} + \|\mathbf{II}_W u_t - u_t\|_{L^2(L^2)} ]^2. \end{aligned}$$

Note that this result implies the second estimate of Theorem 2.1.

*Proof.* After applying the Cauchy-Schwarz inequality to terms of the right-hand side of the identity of Lemma 3.3, we obtain

$$\begin{aligned} & \frac{1}{2} [\|\boldsymbol{\varepsilon}_h^q(T)\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)(T)\|_{\partial\mathcal{T}_h}^2] + \int_0^T \|\partial_t \varepsilon_h^u\|^2 \\ & \leq \frac{1}{2} [\|\boldsymbol{\varepsilon}_h^q(0)\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)(0)\|_{\partial\mathcal{T}_h}^2] \\ & \quad + \int_0^T [\|\mathbf{I}_V \mathbf{q}_t - \mathbf{q}_t\| \|\boldsymbol{\varepsilon}_h^q\| + \|\mathbf{I}_W u_t - u_t\| \|\partial_t \varepsilon_h^u\|], \end{aligned}$$

and so,

$$\begin{aligned} & \frac{1}{2} [\|\boldsymbol{\varepsilon}_h^q(T)\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)(T)\|_{\partial\mathcal{T}_h}^2] + \frac{1}{2} \int_0^T \|\partial_t \varepsilon_h^u\|^2 \\ & \leq \frac{1}{2} [\|\boldsymbol{\varepsilon}_h^q(0)\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)(0)\|_{\partial\mathcal{T}_h}^2] \\ & \quad + \int_0^T \|\mathbf{I}_V \mathbf{q}_t - \mathbf{q}_t\| \|\boldsymbol{\varepsilon}_h^q\| + \frac{1}{2} \|\mathbf{I}_W u_t - u_t\|_{L^2(L^2)}^2. \end{aligned}$$

Next, we note that if we differentiate the equation in Lemma 3.2 and evaluate the result at  $t = 0$ , we obtain that

$$\|\boldsymbol{\varepsilon}_h^q(0)\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)(0)\|_{\partial\mathcal{T}_h}^2 = ((\mathbf{I}_V \mathbf{q} - \mathbf{q})(0), \boldsymbol{\varepsilon}_h^q(0))_{\mathcal{T}_h},$$

since  $\varepsilon_h^u(0) = 0$ . This implies that

$$\|\boldsymbol{\varepsilon}_h^q(0)\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)(0)\|_{\partial\mathcal{T}_h}^2 \leq \|(\mathbf{I}_V \mathbf{q} - \mathbf{q})(0)\|^2,$$

and so,

$$\begin{aligned} & \frac{1}{2} [\|\boldsymbol{\varepsilon}_h^q(T)\|^2 + \|\sqrt{\tau}(\varepsilon_h^u - \widehat{\varepsilon}_h^u)(T)\|_{\partial\mathcal{T}_h}^2] + \frac{1}{2} \int_0^T \|\partial_t \varepsilon_h^u\|^2 \\ & \leq \frac{1}{2} \|(\mathbf{I}_V \mathbf{q} - \mathbf{q})(0)\|^2 + \int_0^T \|\mathbf{I}_V \mathbf{q}_t - \mathbf{q}_t\| \|\boldsymbol{\varepsilon}_h^q\| + \frac{1}{2} \|\mathbf{I}_W u_t - u_t\|_{L^2(L^2)}^2. \end{aligned}$$

The inequality now follows by using Proposition 3.1. This completes the proof.  $\square$

**Step 4: Preliminaries to  $L^\infty(L^2)$  estimates by parabolic duality.** To estimate  $P_{k-1} \varepsilon_h^u$  in the  $L^\infty(L^2)$ -norm by duality, the natural approach is to consider the identity

$$\|\varepsilon(T)\| \leq \sup_{\Theta \in \widehat{W}_h} \frac{(\varepsilon(T), \Theta)}{\|\Theta\|},$$

and then estimate the expression  $(\varepsilon(T), \Theta)$  by using the error equations of Step 1 and the solution of the dual problem

$$(3.3a) \quad \boldsymbol{\Phi} + \nabla \Psi = 0 \quad \text{on } \Omega \times (0, T),$$

$$(3.3b) \quad \Psi_t - \nabla \cdot \boldsymbol{\Phi} = 0 \quad \text{on } \Omega \times (0, T),$$

$$(3.3c) \quad \Psi = 0 \quad \text{on } \partial\Omega \times (0, T),$$

$$(3.3d) \quad \Psi(T) = \Theta \quad \text{on } \Omega.$$

Together with the estimates obtained in the previous sections, we will need the parabolic regularity estimates gathered in the following result. Their proof is given in the Appendix; see subsection A.2.

**Proposition 3.2.** *We have that*

$$\begin{aligned} \|\nabla\Psi\|_{L^2(L^2)}^2 &\leq \frac{1}{2} \|\Theta\|^2 && \forall \Theta \in L^2(\Omega), \\ \|\Delta\Psi\|_{L^1(L^2)} &\leq (\delta/2)^{1/2} \|\nabla\Theta\| + \frac{1}{2} (\log(T/\delta))^{1/2} \|\Theta\| && \forall \Theta \in H^1(\Omega), \end{aligned}$$

for all  $\delta \in (0, T)$ .

Unfortunately, we cannot use the second inequality since it assumes that  $\Theta$  lies in  $H^1(\Omega)$ . To overcome this difficulty, we use the fact that  $\varepsilon_h^u$  belongs to a finite dimensional space to obtain the following characterization of its  $L^2$ -norm.

To state it, we need to introduce some notation. Let  $\mathcal{T}_{h'}$  be a triangulation of  $\Omega$  obtained by refining each of the simplexes of the triangulation  $\mathcal{T}_h$ , and let  $W_{h'}^0$  be the space of *continuous* functions which are polynomials of degree  $k$  on each element of  $\mathcal{T}_{h'}$ . Finally let  $P_{h'}$  be the  $L^2$ -projection from  $W_h$  to  $W_{h'}^0$ .

We are now ready to state the result. Its proof is given in the Appendix; see subsection A.3.

**Lemma 3.4.** *For any triangulation  $\mathcal{T}_h$  of  $\Omega$ , we can always find a refinement  $\mathcal{T}_{h'}$  for which we have*

$$\begin{aligned} \|\nabla P_{h'}\theta\| &\leq \frac{C}{\rho} \|\theta\| && \forall \theta \in W_h, \\ \|\varepsilon\| &\leq 2 \sup_{\theta \in W_h} \frac{(\varepsilon, P_{h'}\theta)}{\|\theta\|} && \forall \varepsilon \in W_h. \end{aligned}$$

Here the constant  $C$  depends solely on  $k$  and  $n$ .

As a consequence, we have the following new set of parabolic regularity estimates. Recall that, by equation (2.6),  $\kappa \log \kappa = T/\rho^2$ .

**Corollary 3.3.** *Let  $(\Phi, \Psi)$  be the solution of the dual problem with  $\Theta := P_{h'}\theta$  where  $\theta \in W_h$  and  $P_{h'}$  satisfies Lemma 3.4. Then*

$$\begin{aligned} \|\nabla\Psi\|_{L^2(L^2)}^2 &\leq \frac{1}{2} \|\theta\|^2, \\ \|\Delta\Psi\|_{L^1(L^2)} &\leq C \sqrt{\log \kappa} \|\theta\|. \end{aligned}$$

*Proof.* The first inequality is obtained by setting  $\Theta := P_{h'}\theta$  in the first inequality of Proposition 3.2 and using the fact that  $P_{h'}$  is the  $L^2$ -projection.

The second inequality is obtained as follows. From the second inequality of Proposition 3.2 with  $\Theta := P_{h'}\theta$  and the first inequality of Lemma 3.4, we have

$$\|\Delta\Psi\|_{L^1(L^2)} \leq C (\delta/2\rho^2)^{1/2} \|\theta\| + \frac{1}{2} (\log(T/\delta))^{1/2} \|\theta\| = \left( \frac{C}{\sqrt{2}} + \frac{1}{2} \right) \sqrt{\log \kappa} \|\theta\|,$$

if we take  $\delta := T/\kappa$ . This completes the proof. □

We are now ready to obtain our  $L^\infty(L^2)$  estimates.

**Step 5: Estimate of  $P_{k-1}\varepsilon_h^u$  in  $L^\infty(L^2)$  by parabolic duality.** We begin by obtaining an expression for  $(P_{k-1}\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h}$  in terms of the errors  $\varepsilon_h^q$ ,  $\varepsilon_h^u$  and the solution of the dual problem. In it,  $I_h$  is any interpolation operator from  $L^2(\Omega)$  into  $W_h \cap H_0^1(\Omega)$  and  $P_w$  is the  $L^2$ -projection into  $W_h$ .

**Lemma 3.5.** *For any  $T > 0$ , the quantity  $(P_{k-1}\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h}$  is equal to*

$$\int_0^T [(\varepsilon_h^q, -\Pi^{\text{BDM}} \nabla \Psi + \nabla I_h \Psi)_{\mathcal{T}_h} + (\Pi_V \mathbf{q} - \mathbf{q}, \Pi^{\text{BDM}} \nabla \Psi - \nabla P_w \Psi)_{\mathcal{T}_h} \\ + (\partial_t \varepsilon_h^u, P_{k-1} \Psi - I_h \Psi)_{\mathcal{T}_h} + (\Pi_W u_t - u_t, I_h \Psi - P_{k-1} \Psi)_{\mathcal{T}_h}].$$

*Proof.* Since  $\Psi(T) = \Theta$  by (3.3d) and  $\varepsilon_h^u(0) = 0$  by (3.1e), we have

$$(P_{k-1}\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} = \int_0^T [(\partial_t P_{k-1}\varepsilon_h^u, \Psi)_{\mathcal{T}_h} + (P_{k-1}\varepsilon_h^u, \Psi_t)_{\mathcal{T}_h}] \\ = \int_0^T [(\partial_t \varepsilon_h^u, P_{k-1}\Psi)_{\mathcal{T}_h} + (\varepsilon_h^u, P_{k-1}\nabla \cdot \Phi)_{\mathcal{T}_h}],$$

by the definition of the  $L^2$ -projection  $P_{k-1}$  and by the second equation of the dual problem (3.3b). Now, using a well-known property of the projection  $\Pi^{\text{BDM}}$ , we get that

$$(P_{k-1}\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} = \int_0^T [(\partial_t \varepsilon_h^u, P_{k-1}\Psi)_{\mathcal{T}_h} + (\varepsilon_h^u, \nabla \cdot \Pi^{\text{BDM}} \Phi)_{\mathcal{T}_h}],$$

and by the first error equation (3.1a) with  $\mathbf{r} := \Pi^{\text{BDM}} \Phi$ ,

$$(P_{k-1}\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} = \int_0^T [(\partial_t \varepsilon_h^u, P_{k-1}\Psi)_{\mathcal{T}_h} + (\varepsilon_h^q, \Pi^{\text{BDM}} \Phi)_{\mathcal{T}_h} \\ - (\Pi_V \mathbf{q} - \mathbf{q}, \Pi^{\text{BDM}} \Phi)_{\mathcal{T}_h} + \langle \varepsilon_h^{\hat{u}}, \Pi^{\text{BDM}} \Phi \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h}] \\ = \int_0^T [(\partial_t \varepsilon_h^u, P_{k-1}\Psi)_{\mathcal{T}_h} + (\varepsilon_h^q, \Pi^{\text{BDM}} \Phi)_{\mathcal{T}_h} \\ - (\Pi_V \mathbf{q} - \mathbf{q}, \Pi^{\text{BDM}} \Phi)_{\mathcal{T}_h}],$$

since  $\langle \varepsilon_h^{\hat{u}}, \Pi^{\text{BDM}} \Phi \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} = \langle \varepsilon_h^{\hat{u}}, \Pi^{\text{BDM}} \Phi \cdot \mathbf{n} \rangle_{\partial \Omega} = 0$  because  $\Pi^{\text{BDM}} \Phi \in \mathbf{H}(\text{div})$  and because  $\varepsilon_h^{\hat{u}} = 0$  on  $\partial \Omega$  by the error equation (3.1c).

Let us now work on the second integrand of the right-hand side. By the first equation of the dual problem (3.3a),

$$(\varepsilon_h^q, \Pi^{\text{BDM}} \Phi)_{\mathcal{T}_h} = (\varepsilon_h^q, -\Pi^{\text{BDM}} \nabla \Psi + \nabla I_h \Psi)_{\mathcal{T}_h} - (\varepsilon_h^q, \nabla I_h \Psi)_{\mathcal{T}_h} \\ = (\varepsilon_h^q, -\Pi^{\text{BDM}} \nabla \Psi + \nabla I_h \Psi)_{\mathcal{T}_h} - (\partial_t \varepsilon_h^u, I_h \Psi)_{\mathcal{T}_h} \\ + (\Pi_W u_t - u_t, I_h \Psi)_{\mathcal{T}_h} - \langle \varepsilon_h^{\hat{q}} \cdot \mathbf{n}, I_h \Psi \rangle_{\partial \mathcal{T}_h}$$

by the second error equation (3.1b) with  $w := I_h \Psi$ . Finally, we get that

$$(\varepsilon_h^q, \Pi^{\text{BDM}} \Phi)_{\mathcal{T}_h} = (\varepsilon_h^q, -\Pi^{\text{BDM}} \nabla \Psi + \nabla I_h \Psi)_{\mathcal{T}_h} - (\partial_t \varepsilon_h^u, I_h \Psi)_{\mathcal{T}_h} + (\Pi_W u_t - u_t, I_h \Psi)_{\mathcal{T}_h}$$

since  $\langle \varepsilon_h^{\hat{q}} \cdot \mathbf{n}, I_h \Psi \rangle_{\partial \mathcal{T}_h} = \langle \varepsilon_h^{\hat{q}} \cdot \mathbf{n}, I_h \Psi \rangle_{\partial \Omega} = 0$ . Indeed, the normal component of  $\varepsilon_h^{\hat{q}}$  is single valued by the error equation (3.1d) and  $I_h \Psi = 0$  on  $\partial \Omega$  by the boundary condition of the dual problem (3.3c).

Inserting this expression into the last identity for  $(P_{k-1}\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h}$  and rearranging terms, we obtain

$$(P_{k-1}\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} = \int_0^T [(\varepsilon_h^q, -\Pi^{\text{BDM}} \nabla \Psi + \nabla I_h \Psi)_{\mathcal{T}_h} - (\Pi_V \mathbf{q} - \mathbf{q}, \Pi^{\text{BDM}} \Phi)_{\mathcal{T}_h} \\ + (\partial_t \varepsilon_h^u, P_{k-1}\Psi - I_h \Psi)_{\mathcal{T}_h} + (\Pi_W u_t - u_t, I_h \Psi)_{\mathcal{T}_h}].$$

Following the same argument used in the proof of Lemma 4.1, we get that

$$\begin{aligned} -(\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}, \mathbf{\Pi}^{\text{BDM}} \Phi)_{\mathcal{T}_h} &= (\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}, \mathbf{\Pi}^{\text{BDM}} \nabla \Psi - \nabla P_W \Psi)_{\mathcal{T}_h} \\ (\mathbf{\Pi}_W u_t - u_t, \mathbf{I}_h \Psi)_{\mathcal{T}_h} &= (\mathbf{\Pi}_W u_t - u_t, \mathbf{I}_h \Psi - P_{k-1} \Psi)_{\mathcal{T}_h}, \end{aligned}$$

and the result follows. This completes the proof.  $\square$

We now use the second inequality of Lemma 3.4 with  $\varepsilon := P_{k-1} \varepsilon_h^u(T)$  and  $\Theta := P_h \theta$  for any  $\theta \in W_h$  to obtain the following estimate.

**Corollary 3.4.** *For any  $T > 0$ , we have*

$$\begin{aligned} \|P_{k-1} \varepsilon_h^u(T)\| \leq & H_1(\Theta) \|\varepsilon_h^q\|_{L^\infty(L^2)} + H_2(\Theta) \|\mathbf{q} - \mathbf{\Pi}_V \mathbf{q}\|_{L^\infty(L^2)} \\ & + H_3(\Theta) \|\partial_t \varepsilon_h^u\|_{L^2(L^2)} + H_4(\Theta) \|\mathbf{\Pi}_W u_t - u_t\|_{L^2(L^2)}, \end{aligned}$$

where

$$\begin{aligned} H_1(\Theta) &:= 2 \sup_{\theta \in W_h} \frac{\|\mathbf{\Pi}^{\text{BDM}} \nabla \Psi - \nabla \mathbf{I}_h \Psi\|_{L^1(L^2)}}{\|\theta\|}, \\ H_2(\Theta) &:= 2 \sup_{\theta \in W_h} \frac{\|\mathbf{\Pi}^{\text{BDM}} \nabla \Psi - \nabla P_W \Psi\|_{L^1(L^2)}}{\|\theta\|}, \\ H_3(\Theta) &:= 2 \sup_{\theta \in W_h} \frac{\|P_{k-1} \Psi - \mathbf{I}_h \Psi\|_{L^2(L^2)}}{\|\theta\|}, \\ H_4(\Theta) &:= 2 \sup_{\theta \in W_h} \frac{\|\mathbf{I}_h \Psi - P_{k-1} \Psi\|_{L^2(L^2)}}{\|\theta\|}. \end{aligned}$$

**Lemma 3.6.** *For  $k \geq 1$ , we have*

$$\begin{aligned} H_1(\Theta) &\leq C C_{\text{reg}} \sqrt{\log \kappa} h, \\ H_2(\Theta) &\leq C C_{\text{reg}} \sqrt{\log \kappa} h, \\ H_3(\Theta) &\leq C h, \\ H_4(\Theta) &\leq C h, \end{aligned}$$

where  $C$  is a constant independent of the exact solution and of the discretization parameters.

*Proof.* Using standard approximation properties (see, in particular, the approximation properties of  $\mathbf{\Pi}^{\text{BDM}}$  of [3, Proposition 3.6]), we have that

$$\begin{aligned} \|\mathbf{\Pi}^{\text{BDM}} \nabla \Psi - \nabla \mathbf{I}_h \Psi\| &\leq C h |\Psi|_{H^2(\Omega)}, \\ \|\mathbf{\Pi}^{\text{BDM}} \nabla \Psi - \nabla P_W \Psi\| &\leq C h |\Psi|_{H^2(\Omega)}, \\ \|P_{k-1} \Psi - \mathbf{I}_h \Psi\| &\leq C h \|\nabla \Psi\|, \\ \|\mathbf{I}_h \Psi - P_{k-1} \Psi\| &\leq C h \|\nabla \Psi\|. \end{aligned}$$

The result follows by using the elliptic regularity assumption (2.5) (on the first two inequalities) and the parabolic regularity estimates of Corollary 3.3. This completes the proof.  $\square$

Inserting the bounds of Lemma 3.6 into the inequality of Corollary 3.4, we get the last estimate of Theorem 2.1. This completes the proof of Theorem 2.1.

**Step 6: Estimate of  $u - u_h^*$  in  $L^\infty(L^2)$ .** To end this section, and for the sake of completeness, let us obtain the last estimate of Corollary 2.1.

By the equations defining the postprocessed approximation  $u_h^*$ , (2.3), we have that

$$\begin{aligned} (u - u_h^*, w)_K &= (P_{k-1}\varepsilon_h^u, w)_K && \text{for all } w \in \mathcal{P}_{k-1}(K), \\ (\nabla(u - u_h^*), \nabla w)_K &= -(\varepsilon_h^q, \nabla w)_K && \text{for all } w \in \mathcal{W}_{k+1}(K). \end{aligned}$$

Hence, a simple argument gives that

$$\|u - u_h^*\|_K \leq C h_K^{k+2} |u|_{H^{k+2}(K)} + C \|P_{k-1}\varepsilon_h^u\|_K + C h_K \|\varepsilon_h^q\|_K,$$

and the last estimate of Corollary 2.1 follows.

#### 4. AN ESTIMATE OF $\varepsilon_h^u$ IN $L^\infty(L^2)$

In this section, for the sake of completeness, we provide an estimate of  $\varepsilon_h^u$ .

**4.1. The estimate.** Recall that  $\tau_K^* := \max \tau|_{\partial K \setminus F^*}$ , where  $F^*$  is a face of  $K$  at which  $\tau|_{\partial K}$  is maximum, and  $\tau_K^{\max} := \max \tau|_{\partial K}$ ; see the approximation result of Theorem 2.2 at the end of section 2.

**Theorem 4.1.** *Assume that the elliptic regularity inequality (2.5) holds and that  $k \geq 1$ . Then, for any  $T > 0$ , we have*

$$\|\varepsilon_h^u\|_{L^\infty(L^2)} \leq C h \|\mathbf{q}, u) - \Pi_h(\mathbf{q}, u)\|_{2,T,\Omega}.$$

Here

$$C := C(1 + C_{\text{reg}}(C_{\tau^*} \sqrt{\log \kappa} + C_{\tau^{\max}} h/\rho)),$$

where

$$C_{\tau^*} := \max_{K \in \mathcal{T}_h} \{1, h_K \tau_K^*\} \quad \text{and} \quad C_{\tau^{\max}} := \max_{K \in \mathcal{T}_h} \{1, 1/(h_K \tau_K^{\max})\}.$$

Note that, as long as  $h/\rho$ ,  $C_{\tau^*}$ , and  $C_{\tau^{\max}}$  remain bounded, the above estimate of  $\varepsilon_h^u$  is essentially the same as that of  $P_{k-1}\varepsilon_h^u$  given by Theorem 2.1. For  $h/\rho$  to remain bounded, we need regular and quasiuniform triangulations  $\mathcal{T}_h$ . For  $C_{\tau^*}$  and  $C_{\tau^{\max}}$  to remain bounded, we need to properly choose  $\tau_K^*$  and  $\tau_K^{\max}$ . Taking  $\tau_K^*$  to be order one and  $\tau_K^{\max}$  of order  $1/h_K$  ensures that  $C_{\tau^*}$  and  $C_{\tau^{\max}}$  are uniformly bounded, and that the order of convergence of the projection  $\Pi_h$  is optimal; see Theorem 2.2. Note, however, that if we take  $\tau_K^{\max}$  to be of order one,  $C_{\tau^{\max}}$  is not bounded anymore. In fact, for quasiuniform triangulations, it is of order  $1/h$  and we lose the superconvergence of the projection of the error  $\varepsilon_h^u$ . Note that this is *not* the case for  $P_{k-1}\varepsilon_h^u$ , which does superconverge with order  $k+2$  in this case.

**4.2. Proof of the estimate.** To prove this result, we proceed as in Step 5 in the previous section. Thus, we begin with the following identity.

**Lemma 4.1.** *For any  $T > 0$ , we have*

$$\begin{aligned} (\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} &= - \int_0^T [-(\varepsilon_h^q, \Pi_V \Phi - \Phi)_{\mathcal{T}_h} + (\mathbf{q} - \Pi_V \mathbf{q}, \Pi_V \nabla \Psi - \nabla P_W \Psi)_{\mathcal{T}_h} \\ &\quad + (\partial_t \varepsilon_h^u, \Pi_W \Psi - \Psi)_{\mathcal{T}_h} - (\Pi_W u_t - u_t, \Pi_W \Psi - P_{k-1} \Psi)_{\mathcal{T}_h}]. \end{aligned}$$

*Proof.* To obtain this identity, we begin by noting that

$$(\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} = \int_0^T \frac{d}{dt} (\varepsilon_h^u, \Psi)_{\mathcal{T}_h},$$

since  $\Psi(T) = \Theta$  by (3.3d) and  $\varepsilon_h^u(0) = 0$  by (3.1e). Then

$$(\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} = \int_0^T [(\partial_t \varepsilon_h^u, \Psi)_{\mathcal{T}_h} + (\varepsilon_h^u, \Psi_t)_{\mathcal{T}_h}] = \int_0^T [(\partial_t \varepsilon_h^u, \Psi)_{\mathcal{T}_h} + (\varepsilon_h^u, \nabla \cdot \Phi)_{\mathcal{T}_h}]$$

by the second equation of the dual problem (3.3b). If we now integrate by parts the last term in the right-hand side and use the orthogonality property (2.4a) for  $\Pi_V$ , we get

$$\begin{aligned} (\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} &= - \int_0^T [-(\partial_t \varepsilon_h^u, \Psi)_{\mathcal{T}_h} + (\nabla \varepsilon_h^u, \Pi_V \Phi)_{\mathcal{T}_h} - \langle \varepsilon_h^u, \Phi \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h}] \\ &= - \int_0^T [-(\partial_t \varepsilon_h^u, \Psi)_{\mathcal{T}_h} - (\varepsilon_h^u, \nabla \cdot \Pi_V \Phi)_{\mathcal{T}_h} + \langle \varepsilon_h^u, (\Pi_V \Phi - \Phi) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h}]. \end{aligned}$$

By the first error equation of Lemma 3.1 with  $\mathbf{r} := \Pi_V \Phi$ ,

$$\begin{aligned} (\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} &= - \int_0^T [-(\partial_t \varepsilon_h^u, \Psi)_{\mathcal{T}_h} - (\varepsilon_h^q, \Pi_V \Phi)_{\mathcal{T}_h} - \langle \varepsilon_h^{\hat{u}}, \Pi_V \Phi \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \\ &\quad + (\Pi_V \mathbf{q} - \mathbf{q}, \Pi_V \Phi)_{\mathcal{T}_h} + \langle \varepsilon_h^u, (\Pi_V \Phi - \Phi) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h}] \\ &= - \int_0^T [-(\partial_t \varepsilon_h^u, \Psi)_{\mathcal{T}_h} - (\varepsilon_h^q, \Pi_V \Phi)_{\mathcal{T}_h} + (\Pi_V \mathbf{q} - \mathbf{q}, \Pi_V \Phi)_{\mathcal{T}_h} \\ &\quad + \langle \varepsilon_h^u - \varepsilon_h^{\hat{u}}, (\Pi_V \Phi - \Phi) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} - \langle \varepsilon_h^{\hat{u}}, \Phi \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h}] \\ &= - \int_0^T [-(\partial_t \varepsilon_h^u, \Psi)_{\mathcal{T}_h} - (\varepsilon_h^q, \Pi_V \Phi)_{\mathcal{T}_h} + (\Pi_V \mathbf{q} - \mathbf{q}, \Pi_V \Phi)_{\mathcal{T}_h} \\ &\quad + \langle \varepsilon_h^u - \varepsilon_h^{\hat{u}}, (\Pi_V \Phi - \Phi) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h}], \end{aligned}$$

since  $\langle \varepsilon_h^{\hat{u}}, \Phi \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} = \langle \varepsilon_h^{\hat{u}}, \Phi \cdot \mathbf{n} \rangle_{\partial \Omega} = 0$  because  $\Phi \in \mathbf{H}(\text{div})$  and because  $\varepsilon_h^{\hat{u}} = 0$  on  $\partial \Omega$ , by the error equation (3.1c).

Next, let us work only on the second term of the integrand of the right-hand side. We have, by the first equation of the dual problem (3.3a),

$$\begin{aligned} -(\varepsilon_h^q, \Pi_V \Phi)_{\mathcal{T}_h} &= -(\varepsilon_h^q, \Pi_V \Phi - \Phi)_{\mathcal{T}_h} + (\varepsilon_h^q, \nabla \Psi)_{\mathcal{T}_h} \\ &= -(\varepsilon_h^q, \Pi_V \Phi - \Phi)_{\mathcal{T}_h} - (\nabla \cdot \varepsilon_h^q, \Pi_W \Psi)_{\mathcal{T}_h} + \langle \varepsilon_h^q \cdot \mathbf{n}, \Psi \rangle_{\partial \mathcal{T}_h}, \end{aligned}$$

by the orthogonality property (2.4b) of  $\Pi_W$ . Then

$$\begin{aligned} -(\varepsilon_h^q, \Pi_V \Phi)_{\mathcal{T}_h} &= -(\varepsilon_h^q, \Pi_V \Phi - \Phi)_{\mathcal{T}_h} + (\varepsilon_h^q, \nabla \Pi_W \Psi)_{\mathcal{T}_h} + \langle \varepsilon_h^q \cdot \mathbf{n}, \Psi - \Pi_W \Psi \rangle_{\partial \mathcal{T}_h}, \\ &= -(\varepsilon_h^q, \Pi_V \Phi - \Phi)_{\mathcal{T}_h} + (\partial_t \varepsilon_h^u, \Pi_W \Psi)_{\mathcal{T}_h} - (\Pi_W u_t - u_t, \Pi_W \Psi)_{\mathcal{T}_h} \\ &\quad + \langle \varepsilon_h^{\hat{q}} \cdot \mathbf{n}, \Pi_W \Psi \rangle_{\partial \mathcal{T}_h} + \langle \varepsilon_h^q \cdot \mathbf{n}, \Psi - \Pi_W \Psi \rangle_{\partial \mathcal{T}_h}, \end{aligned}$$

by the second error equation with  $w := \Pi_W \varepsilon_h^u$ . Hence

$$\begin{aligned} -(\varepsilon_h^q, \mathbf{\Pi}_V \mathbf{\Phi})_{\mathcal{T}_h} &= -(\varepsilon_h^q, \mathbf{\Pi}_V \mathbf{\Phi} - \mathbf{\Phi})_{\mathcal{T}_h} + (\partial_t \varepsilon_h^u, \Pi_W \Psi)_{\mathcal{T}_h} - (\Pi_W u_t - u_t, \Pi_W \Psi)_{\mathcal{T}_h} \\ &\quad + \langle (\varepsilon_h^{\hat{q}} - \varepsilon_h^q) \cdot \mathbf{n}, \Pi_W \Psi - \Psi \rangle_{\partial \mathcal{T}_h} + \langle \varepsilon_h^{\hat{q}} \cdot \mathbf{n}, \Psi \rangle_{\partial \mathcal{T}_h} \\ &= -(\varepsilon_h^q, \mathbf{\Pi}_V \mathbf{\Phi} - \mathbf{\Phi})_{\mathcal{T}_h} + (\partial_t \varepsilon_h^u, \Pi_W \Psi)_{\mathcal{T}_h} - (\Pi_W u_t - u_t, \Pi_W \Psi)_{\mathcal{T}_h} \\ &\quad + \langle \tau(\varepsilon_h^u - \varepsilon_h^{\hat{u}}), \Pi_W \Psi - \Psi \rangle_{\partial \mathcal{T}_h}, \end{aligned}$$

by the definition of  $\varepsilon_h^{\hat{q}}$ , (3.1f), and because  $\langle \varepsilon_h^{\hat{q}} \cdot \mathbf{n}, \Psi \rangle_{\partial \mathcal{T}_h} = 0$ . Indeed, by the error equation (3.1d), the normal component of  $\varepsilon_h^{\hat{q}}$  is single valued and we have that  $\langle \varepsilon_h^{\hat{q}} \cdot \mathbf{n}, \Psi \rangle_{\partial \mathcal{T}_h} = \langle \varepsilon_h^{\hat{q}} \cdot \mathbf{n}, \Psi \rangle_{\partial \Omega}$ . The fact that this expression is equal to zero follows from the boundary condition for the dual problem (3.3c).

Inserting the above expression in the last identity of  $(\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h}$  and rearranging terms, we get

$$\begin{aligned} (\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} &= - \int_0^T [ -(\varepsilon_h^q, \mathbf{\Pi}_V \mathbf{\Phi} - \mathbf{\Phi})_{\mathcal{T}_h} + (\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}, \mathbf{\Pi}_V \mathbf{\Phi})_{\mathcal{T}_h} \\ &\quad + (\partial_t \varepsilon_h^u, \Pi_W \Psi - \Psi)_{\mathcal{T}_h} - (\Pi_W u_t - u_t, \Pi_W \Psi)_{\mathcal{T}_h} \\ &\quad + \langle \varepsilon_h^u - \varepsilon_h^{\hat{u}}, (\mathbf{\Pi}_V \mathbf{\Phi} - \mathbf{\Phi}) \cdot \mathbf{n} + \tau(\Pi_W \Psi - \Psi) \rangle_{\partial \mathcal{T}_h} ] \\ &= - \int_0^T [ -(\varepsilon_h^q, \mathbf{\Pi}_V \mathbf{\Phi} - \mathbf{\Phi})_{\mathcal{T}_h} + (\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}, \mathbf{\Pi}_V \mathbf{\Phi})_{\mathcal{T}_h} \\ &\quad + (\partial_t \varepsilon_h^u, \Pi_W \Psi - \Psi)_{\mathcal{T}_h} - (\Pi_W u_t - u_t, \Pi_W \Psi)_{\mathcal{T}_h} ], \end{aligned}$$

by the orthogonality property (2.4c). The result now follows since

$$(\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}, \mathbf{\Pi}_V \mathbf{\Phi})_{\mathcal{T}_h} = (\mathbf{q} - \mathbf{\Pi}_V \mathbf{q}, \mathbf{\Pi}_V \nabla \Psi)_{\mathcal{T}_h} = (\mathbf{q} - \mathbf{\Pi}_V \mathbf{q}, \mathbf{\Pi}_V \nabla \Psi - \nabla P_W \Psi)_{\mathcal{T}_h},$$

by the definition of  $P_W \Psi$  and the orthogonality property of  $\mathbf{\Pi}_V$ , (2.4a), and since

$$(\Pi_W u_t - u_t, \Pi_W \Psi)_{\mathcal{T}_h} = (\Pi_W u_t - u_t, \Pi_W \Psi - P_{k-1} \Psi)_{\mathcal{T}_h},$$

by the definition of  $P_{k-1} \Psi$  and the orthogonality property of  $\Pi_W$ , (2.4b). This completes the proof.  $\square$

We immediately have the following consequence of the above result.

**Corollary 4.1.** *For any  $T > 0$ , we have*

$$\begin{aligned} \|\varepsilon_h^u(T)\| &\leq H_1(\Theta) \|\varepsilon_h^q\|_{L^\infty(L^2)} + H_2(\Theta) \|\mathbf{q} - \mathbf{\Pi}_V \mathbf{q}\|_{L^\infty(L^2)} \\ &\quad + H_3(\Theta) \|\partial_t \varepsilon_h^u\|_{L^2(L^2)} + H_4(\Theta) \|\Pi_W u_t - u_t\|_{L^2(L^2)}, \end{aligned}$$

where

$$\begin{aligned} H_1(\Theta) &:= 2 \sup_{\theta \in W_h} \frac{\|\mathbf{\Pi}_V \mathbf{\Phi} - \mathbf{\Phi}\|_{L^1(L^2)}}{\|\theta\|}, \\ H_2(\Theta) &:= 2 \sup_{\theta \in W_h} \frac{\|\mathbf{\Pi}_V \nabla \Psi - \nabla P_W \Psi\|_{L^1(L^2)}}{\|\theta\|}, \\ H_3(\Theta) &:= 2 \sup_{\theta \in W_h} \frac{\|\Pi_W \Psi - \Psi\|_{L^2(L^2)}}{\|\theta\|}, \\ H_4(\Theta) &:= 2 \sup_{\theta \in W_h} \frac{\|\Pi_W \Psi - P_{k-1} \Psi\|_{L^2(L^2)}}{\|\theta\|}, \end{aligned}$$

and  $(\mathbf{\Phi}, \Psi)$  is the solution of the dual problem (3.3) with  $\Theta := P_h \theta$ .



**Lemma 4.2.** *Assume that the elliptic regularity inequality (2.5) holds. Then, for  $k \geq 1$ , we have*

$$\begin{aligned} \|\mathbf{I}_V \Phi - \Phi\| &\leq C C_{\text{reg}} C_{\tau^*} h \|\Delta \Psi\|, \\ \|\nabla \Psi - \nabla P_W \Psi\| &\leq C C_{\text{reg}} h \|\Delta \Psi\|, \\ \|\mathbf{I}_W \Psi - \Psi\| &\leq C C_{\text{reg}} C_{\tau^{\max}} h^2 \|\Delta \Psi\|, \\ \|\Psi - P_{k-1} \Psi\| &\leq C h \|\nabla \Psi\|, \end{aligned}$$

where  $C_{\tau^*} := \max_{K \in \mathcal{T}_h} \{1, h_K \tau_K^*\}$ ,  $C_{\tau^{\max}} := \max_{K \in \mathcal{T}_h} \{1, 1/(h_K \tau_K^{\max})\}$ , and  $C$  is a constant independent of the exact solution and of the discretization parameters.

*Proof.* A direct application of Theorem 2.2 gives

$$\begin{aligned} \|\mathbf{I}_V \Phi - \Phi\| &\leq C C_{\tau^*} h |\Psi|_{H^2(\Omega)}, \\ \|\mathbf{I}_W \Psi - \Psi\| &\leq C C_{\tau^{\max}} h^2 |\Psi|_{H^2(\Omega)}. \end{aligned}$$

Also, since  $P_W$  is the  $L^2$ -projection into  $W_h$ , we have

$$\|\nabla \Psi - \nabla P_W \Psi\| \leq C h |\Psi|_{H^2(\Omega)}.$$

The first three estimates now follow after using the elliptic regularity assumption (2.5). The last inequality follows from the fact that  $P_{k-1}$  is the  $L^2$ -projection into  $W_{h,k-1}$  and that  $k \geq 1$ . This completes the proof.  $\square$

Combining the approximation estimates of the lemma just proved with the parabolic regularity estimates

$$\begin{aligned} \|\Delta \Psi\|_{L^2(L^2)} &\leq \frac{1}{\sqrt{2}} \|\nabla P_h \theta\| \leq \frac{C}{\rho} \|\theta\|, \quad \text{by Lemma 3.4,} \\ \|\Delta \Psi\|_{L^1(L^2)} &\leq C \sqrt{\log \kappa} \|\theta\|, \end{aligned}$$

by Corollary 3.3, we immediately get the following result.

**Lemma 4.3.** *For  $k \geq 1$ , we have*

$$\begin{aligned} H_1(\Theta) &\leq C C_{\text{reg}} C_{\tau^*} \sqrt{\log \kappa} h, \\ H_2(\Theta) &\leq C C_{\text{reg}} C_{\tau^*} \sqrt{\log \kappa} h, \\ H_3(\Theta) &\leq C C_{\text{reg}} C_{\tau^{\max}} h^2 / \rho, \\ H_4(\Theta) &\leq C h + C C_{\text{reg}} C_{\tau^{\max}} h^2 / \rho, \end{aligned}$$

where  $C$  is a constant independent of the exact solution and of the discretization parameters.

The estimate of Theorem 4.1 can now be obtained by inserting the estimates of the above lemma in the estimates of Corollary 4.1. This completes the proof of Theorem 4.1.

## 5. EXTENSIONS

In this section, we extend to the classical RT and BDM methods, the results obtained for the HDG method. In this way, we extend to the parabolic setting considered in this paper similar superconvergence results of the RT and BDM methods for the elliptic case.

**The hybridized RT mixed method.** As pointed out in [6], the hybridized version of the RT method on simplexes is obtained from the one used here for the HDG method by setting  $\tau = 0$  and changing the local space  $\mathbf{V}(K)$  to  $\mathcal{P}_k(K) + \mathbf{x} \mathcal{P}_k(K)$ .

It is well known [3] that we can define the function  $\Pi_h(\mathbf{q}, u) = (\mathbf{\Pi}_V \mathbf{q}, \Pi_W u)$  on an arbitrary simplex  $K \in \mathcal{T}_h$  as the element of  $\mathbf{V}_h \times W_h$  which solves

$$\begin{aligned} (\mathbf{\Pi}_V \mathbf{q}, \mathbf{v})_K &= (\mathbf{q}, \mathbf{v})_K && \text{for all } \mathbf{v} \in \mathcal{P}_{k-1}(K), \\ (\Pi_W u, w)_K &= (u, w)_K && \text{for all } w \in \mathcal{P}_k(K), \\ \langle \mathbf{\Pi}_V \mathbf{q} \cdot \mathbf{n}, \mu \rangle_F &= \langle \mathbf{q} \cdot \mathbf{n}, \mu \rangle_F && \text{for all } \mu \in \mathcal{P}_k(F), \end{aligned}$$

for all faces  $F$  of the simplex  $K$ . Note the remarkable similarity with the definition of the projection  $\Pi_h$  for the HDG method (2.4). As a consequence, the equations of the projection of the errors are exactly those contained in Lemma 3.1. The only difference is that the right-hand side of the second equation, (3.1b), is identically equal to zero because  $\Pi_W$  is nothing but the  $L^2$ -projection into the space  $W_h$ . Thus, all the results in section 3 and those in section 4 hold *without* the terms involving  $\Pi_W u_t - u_t$  in the right-hand sides. Finally, note that the terms involving  $\partial_t \varepsilon_h^u$  also disappear because  $\Pi_W$  is the  $L^2$ -projection into the space  $W_h$ .

In particular, Corollary 3.1 gives

$$\|\varepsilon_h^u(T)\|_{L^\infty(L^2)} \leq \|\mathbf{\Pi}_V \mathbf{q} - \mathbf{q}\|_{L^2(L^2)},$$

Corollary 3.2 gives

$$\|\varepsilon_h^q(T)\|_{L^\infty(L^2)} \leq \|(\mathbf{\Pi}_V \mathbf{q} - \mathbf{q})(0)\| + \|\mathbf{\Pi}_V \mathbf{q}_t - \mathbf{q}_t\|_{L^1(L^2)},$$

and Corollary 4.1 gives

$$\|\varepsilon_h^u(T)\| \leq H_1(\Theta) \|\varepsilon_h^q\|_{L^\infty(L^2)} + H_2(\Theta) \|\mathbf{q} - \mathbf{\Pi}_V \mathbf{q}\|_{L^\infty(L^2)},$$

where

$$\begin{aligned} H_1(\Theta) &:= 2 \sup_{\theta \in W_h} \frac{\|\mathbf{\Pi}_V \Phi - \Phi\|_{L^1(L^2)}}{\|\theta\|}, \\ H_2(\Theta) &:= 2 \sup_{\theta \in W_h} \frac{\|\mathbf{\Pi}_V \nabla \Psi - \nabla P_W \Psi\|_{L^1(L^2)}}{\|\theta\|}. \end{aligned}$$

Thus, we easily obtain the following results.

**Theorem 5.1.** *For any  $T > 0$  and any  $k \geq 0$ , we have*

$$\begin{aligned} \|\varepsilon_h^u\|_{L^\infty(L^2)} &\leq C \|\mathbf{q}, 0) - \Pi_h(\mathbf{q}, 0)\|_{1,T,\Omega}, \\ \|\varepsilon_h^q\|_{L^\infty(L^2)} &\leq C \|\mathbf{q}, 0) - \Pi_h(\mathbf{q}, 0)\|_{2,T,\Omega}. \end{aligned}$$

Moreover, if the elliptic regularity inequality (2.5) holds and  $k \geq 1$ , we have

$$\|\varepsilon_h^u\|_{L^\infty(L^2)} \leq C C_{\text{reg}} h \|\mathbf{q}, 0) - \Pi_h(\mathbf{q}, 0)\|_{2,T,\Omega}.$$

Here  $C$  is a constant independent of the exact solution and of the discretization parameters.

**Corollary 5.1.** *For any  $T > 0$  and any  $k \geq 0$ , we have*

$$\begin{aligned} \|u - u_h\|_{L^\infty(L^2)} &\leq \|\Pi_W u - u\|_{L^\infty(L^2)} + \|\mathbf{q}, 0) - \Pi_h(\mathbf{q}, 0)\|_{1,T,\Omega}, \\ \|\mathbf{q} - \mathbf{q}_h\|_{L^\infty(L^2)} &\leq \|\mathbf{q} - \mathbf{\Pi}_V \mathbf{q}\|_{L^\infty(L^2)} + \|\mathbf{q}, 0) - \Pi_h(\mathbf{q}, 0)\|_{2,T,\Omega}. \end{aligned}$$

Moreover, if the elliptic regularity inequality (2.5) holds and  $k \geq 1$ , we have

$$\begin{aligned} \|u - u_h^*\|_{L^\infty(L^2)} &\leq C h^{k+2} |u|_{L^\infty(0,T;H^{k+2}(\Omega))} \\ &\quad + C C_{\text{reg}} \sqrt{\log \kappa} h \|(\mathbf{q}, 0) - \Pi_h(\mathbf{q}, 0)\|_{2,T,\Omega}, \end{aligned}$$

where  $C$  is a constant independent of the exact solution and of the discretization parameters.

We can still get a superconvergence result for  $k = 0$ . To see this, note that in this case, the identity of Lemma 4.1 becomes

$$\begin{aligned} (\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} &= - \int_0^T [-(\varepsilon_h^q, \Pi_V \Phi - \Phi)_{\mathcal{T}_h} + (\mathbf{q} - \Pi_V \mathbf{q}, \Pi_V \nabla \Psi)_{\mathcal{T}_h}] \\ &= - \int_0^T [-(\varepsilon_h^q, \Pi_V \Phi - \Phi)_{\mathcal{T}_h} + (\mathbf{q} - \Pi_V \mathbf{q}, \Pi_V \nabla \Psi - \nabla \Psi)_{\mathcal{T}_h}] \\ &\quad - \int_0^T (\mathbf{q} - \Pi_V \mathbf{q}, \nabla \Psi)_{\mathcal{T}_h}. \end{aligned}$$

Let us work on the integrand of the last term. Since  $\mathbf{q} - \Pi_V \mathbf{q} \in \mathbf{H}(\text{div})$  and  $\Psi = 0$  on  $\partial\Omega$ , we get that

$$(\mathbf{q} - \Pi_V \mathbf{q}, \nabla \Psi)_{\mathcal{T}_h} = -(\nabla \cdot (\mathbf{q} - \Pi_V \mathbf{q}), \Psi)_{\mathcal{T}_h},$$

and since  $P_W \nabla \cdot \mathbf{q} = \nabla \cdot \Pi_V \mathbf{q}$ ,

$$\begin{aligned} (\mathbf{q} - \Pi_V \mathbf{q}, \nabla \Psi)_{\mathcal{T}_h} &= -((\text{Id} - P_W) \nabla \cdot \mathbf{q}, \Psi)_{\mathcal{T}_h} \\ &= -((\text{Id} - P_W) \nabla \cdot \mathbf{q}, (\text{Id} - P_W) \Psi)_{\mathcal{T}_h} \\ &= ((\text{Id} - P_W)(u_t - f), (\text{Id} - P_W) \Psi)_{\mathcal{T}_h}. \end{aligned}$$

This implies that

$$\begin{aligned} (\varepsilon_h^u(T), \Theta)_{\mathcal{T}_h} &= - \int_0^T [-(\varepsilon_h^q, \Pi_V \Phi - \Phi)_{\mathcal{T}_h} + (\mathbf{q} - \Pi_V \mathbf{q}, \Pi_V \nabla \Psi - \nabla \Psi)_{\mathcal{T}_h}] \\ &\quad - \int_0^T ((\text{Id} - P_W)(u_t - f), (\text{Id} - P_W) \Psi)_{\mathcal{T}_h}, \end{aligned}$$

and we readily get the following result.

**Theorem 5.2.** *If the elliptic regularity inequality (2.5) holds and  $k = 0$ , we have*

$$\begin{aligned} \|\varepsilon_h^u\|_{L^\infty(L^2)} &\leq C_{\text{reg}} \sqrt{\log \kappa} h \|(\mathbf{q}, 0) - \Pi_h(\mathbf{q}, 0)\|_{2,T,\Omega} \\ &\quad + C_{\text{reg}} h \|(\text{Id} - P_W)(u_t - f)\|_{L^2(L^2)} \end{aligned}$$

and

$$\begin{aligned} \|u - u_h^*\|_{L^\infty(L^2)} &\leq C h^{k+2} |u|_{L^\infty(0,T;H^{k+2}(\Omega))} \\ &\quad + C_{\text{reg}} \sqrt{\log \kappa} h \|(\mathbf{q}, 0) - \Pi_h(\mathbf{q}, 0)\|_{2,T,\Omega} \\ &\quad + C_{\text{reg}} h \|(\text{Id} - P_W)(u_t - f)\|_{L^2(L^2)}, \end{aligned}$$

where  $C$  is a constant independent of the exact solution and of the discretization parameters.

Let us end this subsection by pointing out that the same results hold for the RT methods on rectangles and cubes.

**The hybridized BDM mixed method.** In a similar manner, the hybridized version of the BDM method on simplexes is obtained from the one used here for the HDG method by setting  $\tau = 0$  and changing the local space  $W(K)$  to  $\mathcal{P}_{k-1}(K)$ .

It is well known [3] that we can define the function  $\Pi_h(\mathbf{q}, u) = (\mathbf{\Pi}_V \mathbf{q}, \mathbf{\Pi}_W u)$  on an arbitrary simplex  $K \in \mathcal{T}_h$  as the element of  $\mathbf{V}_h \times W_h$  which solves

$$\begin{aligned} (\mathbf{\Pi}_V \mathbf{q}, \nabla w + \mathbf{v})_K &= (\mathbf{q}, \nabla w + \mathbf{v})_K && \text{for all } (w, \mathbf{v}) \in \mathcal{P}_{k-1}(K) \times \mathfrak{X}_k(K), \\ (\mathbf{\Pi}_W u, w)_K &= (u, w)_K && \text{for all } w \in \mathcal{P}_k(K), \\ \langle \mathbf{\Pi}_V \mathbf{q} \cdot \mathbf{n}, \mu \rangle_F &= \langle \mathbf{q} \cdot \mathbf{n}, \mu \rangle_F && \text{for all } \mu \in \mathcal{P}_k(F), \end{aligned}$$

for all faces  $F$  of the simplex  $K$ . Here  $\mathfrak{X}_k(K)$  stands for the space of functions in  $\mathcal{P}_k(K)$  whose divergence is zero and whose normal component on  $\partial K$  is also zero.

Once again, the equations of the projection of the errors are exactly those contained in Lemma 3.1 with the right-hand side of the second equation, (3.1b), equal to zero because  $\mathbf{\Pi}_W$  is nothing but the  $L^2$ -projection into the space  $W_h$ . Thus, all the results in section 3 and those in section 4 hold *without* the terms  $\partial_t \varepsilon_h^u$  and  $\mathbf{\Pi}_W u_t - u_t$  in the right-hand sides.

Proceeding exactly as in the case of the hybridized RT method, we obtain that Theorem 5.1 and Corollary 5.1 are *also* valid for the BDM method *provided* we replace the conditions  $k \geq 0$  and  $k \geq 1$  by  $k \geq 1$  and  $k \geq 2$ , respectively. We must do this because the BDM method is only defined for  $k \geq 1$  and because when  $k = 1$ ,  $\nabla P_W \Psi = 0$ , and  $H_2$  is of order one, not of order  $h$ .

The same results hold for the BDM methods on rectangles and cubes.

**Other extensions.** Let us end by pointing out that the results in this paper can be trivially extended to more general second-order elliptic operators. Extensions to nonsmooth initial data and fully discrete methods constitute the subject of ongoing work.

## APPENDIX A. PROOFS OF SOME AUXILIARY RESULTS

**A.1. Proof of the integral inequality of Proposition (3.1).** Pick any  $T > 0$  and set, for  $t \in [0, T]$ ,  $\chi(t) := \max_{t \in [0, T]} A(t) + \int_0^t B(s) \zeta(s) ds$ . Then  $\frac{d}{dt} \chi(t) = B(t) \zeta(t) \leq B(t) \sqrt{\chi(t)}$ , and so  $\sqrt{\chi(T)} \leq \sqrt{\chi(0)} + \frac{1}{2} \int_0^T B(s) ds$ . The result now easily follows.

**A.2. Proof of Proposition 3.2 on parabolic regularity.** We only prove the last inequality, which is not that well known. For any  $\delta \in (0, T)$ , we have

$$\begin{aligned} \|\Delta \Psi\|_{L^1(0, T; L^2(\Omega))} &= \int_0^{T-\delta} \|\Delta \Psi(s)\|_{L^2(\Omega)} ds + \int_{T-\delta}^T \|\Delta \Psi(s)\|_{L^2(\Omega)} ds \\ &\leq (\log(T/\delta))^{1/2} \left( \int_0^{T-\delta} (T-s) \|\Delta \Psi(s)\|_{L^2(\Omega)}^2 ds \right)^{1/2} \\ &\quad + \delta^{1/2} \left( \int_{T-\delta}^T \|\Delta \Psi(s)\|_{L^2(\Omega)}^2 ds \right)^{1/2}, \end{aligned}$$

and the inequality follows from the standard estimates

$$\int_0^T \|\Delta\Psi(s)\|_{L^2(\Omega)}^2 ds \leq \frac{1}{2} \|\nabla\Theta\|^2 \quad \text{and} \quad \int_0^T (T-s)\|\Delta\Psi(s)\|_{L^2(\Omega)}^2 ds \leq \frac{1}{4} \|\Theta\|^2.$$

This completes the proof of Proposition 3.2.

**A.3. Proof of Lemma 3.4 on the operators  $P_{h'}$ .** To prove this result, we begin by describing how to obtain the refinements of  $\mathcal{T}_h, \mathcal{T}_{h'}$ . Given any integer  $N > 0$ , the elements of the triangulation  $\mathcal{T}_{h'}$  are obtained by refining each simplex  $K \in \mathcal{T}_h$  as follows. Let  $\{\lambda_j\}_{j=1}^{n+1}$  be the set of barycentric coordinates of the simplex  $K$ . Then the refinement of  $K = \{\mathbf{x} : \lambda_i(\mathbf{x}) \in [0, 1], i = 1, \dots, n + 1\}$ , consists of all the nonempty sets of the form  $\bigcap_{j=1}^{n+1} K_{j,\ell_j,N}$ , where  $\ell_j \in \{0, 1, \dots, N - 1\}$  for  $j = 1, \dots, n + 1$ . Here, by  $K_{j,\ell,N}$  we mean the slab  $\{\mathbf{x} \in K : \lambda_j(\mathbf{x}) \in \frac{1}{N}[\ell, \ell + 1]\}$ .

The first inequality can now be obtained as a direct consequence of the construction of the space  $W_{h'}^0$ ; the constant  $C$  only depends on  $N$  and  $n$ . The second inequality follows easily if we prove that

$$(A.1) \quad \|(\text{Id} - P_{h'})\theta\| \leq \frac{1}{2} \|\theta\|,$$

for all  $\theta \in W_h$ . Indeed, since

$$(\varepsilon, \theta) = (\varepsilon, (\text{Id} - P_{h'})\theta) + (\varepsilon, P_{h'}\theta) \leq \frac{1}{2} \|\varepsilon\| \|\theta\| + (\varepsilon, P_{h'}\theta),$$

the second identity follows from the fact that  $\|\varepsilon\| = \sup_{\theta \in W_h} \frac{(\varepsilon, \theta)}{\|\theta\|}$  given that  $\varepsilon \in W_h$ .

It remains to prove the inequality (A.1). To do that, we note that, since  $P_{h'}$  is the  $L^2$ -projection into  $W_{h'}^0$ ,

$$\|(\text{Id} - P_{h'})\theta\| \leq \|\theta - w'\|,$$

for any  $w' \in W_{h'}^0$ . We are thus going to construct a special  $w'$  such that

$$\|\theta - w'\| \leq \frac{1}{2} \|\theta\|.$$

We define  $w'$  as follows. By construction of  $\mathcal{T}_{h'}$ , we can write that

$$K = K_N^i \cup K_N^b,$$

where  $K_N^b := \bigcup_{j=1}^{n+1} K_{j,0,N}$  and  $K_N^i := K \setminus K_N^b$ . On each  $K \in \mathcal{T}_h$ , we set  $w' := \theta$  on  $K_N^i$ . We then define  $w'$  on  $\Omega_{h'}^b := \bigcup_{K \in \mathcal{T}_h} K_N^b$  as the  $L^2$ -projection of  $\theta$  into the affine set

$$W_{h'}^b := \{w' \in L^2(\Omega_{h'}^b) \cap W_{h'}^0 : w' = \theta \text{ on } \partial K_N^i \text{ for all } K \in \mathcal{T}_h\}.$$

Then we have that

$$\|\theta - w'\|^2 = \|\theta - w'\|_{\Omega_{h'}^b}^2 \leq 2 \|\theta\|_{\Omega_{h'}^b}^2 = 2 \sum_{K \in \mathcal{T}_h} \|\theta\|_{K_N^b}^2.$$

By construction, we have that, for  $N > n + 1$ , the simplex  $(1 - (n + 1)/N)K$  is congruent to  $K_N^i$  and so, since the volume of  $K_N^b$  equals the volume of  $K$  minus the volume of  $K_N^i$ , we have that the volume of  $K_N^b$  is equal to  $(1 - (1 - (n + 1)/N)^n)$  times the volume of  $K$ . Hence, a simple scaling argument shows that, for  $N > n + 1$ ,

$$\|\theta - w'\|^2 \leq 2 (1 - (1 - (n + 1)/N)^n) C \|\theta\|^2,$$

for some constant  $C$  depending only on  $k$ . So, for big enough  $N$ , we get that

$$\|\theta - w'\|^2 \leq \frac{1}{4} \|\theta\|^2.$$

This completes the proof of Lemma 3.4.

#### REFERENCES

- [1] J. H. Bramble and A. H. Schatz, *Higher order local accuracy by averaging in the finite element method*, Math. Comp. **31** (1977), 94–111. MR0431744 (55:4739)
- [2] F. Brezzi, J. Douglas, Jr., and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math. **47** (1985), 217–235. MR799685 (87g:65133)
- [3] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag, 1991. MR1115205 (92d:65187)
- [4] H. Chen, R. Ewing, and R. Lazarov, *Superconvergence of mixed finite element methods for parabolic problems with nonsmooth initial data*, Numer. Math. **78** (1998), 495–521. MR1606312 (99c:65180)
- [5] B. Cockburn, B. Dong, and J. Guzmán, *A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems*, Math. Comp. **77** (2008), 1887–1916. MR2429868 (2009d:65166)
- [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), 1319–1365. MR2485455 (2010b:65251)
- [7] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas, *A projection-based error analysis of HDG methods*, Math. Comp. **79** (2010), 1351–1367. MR2629996
- [8] B. Cockburn, J. Guzmán, and H. Wang, *Superconvergent discontinuous Galerkin methods for second-order elliptic problems*, Math. Comp. **78** (2009), 1–24. MR2448694 (2009i:65213)
- [9] J. Douglas, Jr., T. Dupont, and M.F. Wheeler, *A quasi-projection analysis of Galerkin methods for parabolic and hyperbolic equations*, Math. Comp. **32** (1978), 345–362. MR0495012 (58:13780)
- [10] J. Douglas, Jr., *Superconvergence in the pressure in the simulation of miscible displacement*, SIAM J. Numer. Anal. **22** (1985), 962–969. MR799123 (86j:65129)
- [11] R. Durán, *Superconvergence for rectangular mixed finite elements*, Numer. Math. **58** (1990), 287–298. MR1075159 (92a:65289)
- [12] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems I: A linear model problem*, SIAM J. Numer. Anal. **28** (1991), 43–77. MR1083324 (91m:65274)
- [13] R. Ewing and R. Lazarov, *Superconvergence of the mixed finite element approximations of parabolic problems using rectangular finite elements*, East-West J. Numer. Math. **1** (1993), 199–212. MR1253635 (94m:65158)
- [14] R. Ewing, R. Lazarov, and J. Wang, *Superconvergence of the velocity along the Gauss lines in mixed finite element methods*, SIAM J. Numer. Anal. **28** (1991), 1015–1029. MR1111451 (92e:65149)
- [15] L. Gastaldi and R.H. Nochetto, *Sharp maximum norm error estimates for general mixed finite element approximations to second order elliptic equations*, RAIRO Modél. Math. Anal. Numér. **23** (1989), 103–128. MR1015921 (91b:65125)
- [16] C. Johnson and V. Thomée, *Error estimates for some mixed finite element methods for parabolic type problems*, RAIRO Modél. Math. Anal. Numér. **15** (1981), 41–78. MR610597 (83c:65239)
- [17] P. A. Raviart and J. M. Thomas, *A mixed finite element method for second order elliptic problems*, Mathematical Aspects of Finite Element Method, Lecture Notes in Math. 606 (I. Galligani and E. Magenes, eds.), Springer-Verlag, New York, 1977, pp. 292–315. MR0483555 (58:3547)
- [18] C. Squeff, *Superconvergence of mixed finite element methods for parabolic equations*, RAIRO Modél. Math. Anal. Numér. **21** (1987), 327–352. MR896246 (88j:65207)
- [19] R. Stenberg, *A family of mixed finite elements for the elasticity problem*, Numer. Math. **53** (1988), 513–538. MR954768 (89h:65192)
- [20] R. Stenberg, *Postprocessing schemes for some mixed finite elements*, RAIRO Modél. Math. Anal. Numér. **25** (1991), 151–167. MR1086845 (92a:65303)

- [21] V. Thomée, *Galerkin Finite Element Methods for parabolic equations*, Springer-Verlag, 1997. MR1479170 (98m:65007)
- [22] V. Thomée, J.-C. Xu, and N.-Y. Zhang, *Superconvergence of the gradient in piecewise linear finite-element approximation to a parabolic problem*, *SIAM J. Numer. Anal.* **26** (1989), 553–573. MR997656 (90e:65165)
- [23] M. Tripathy and R.K. Sinha, *Superconvergence of  $H^1$ -Galerkin mixed finite element methods for parabolic problems*, *Applicable Analysis* **88** (2009), 1213–1231. MR2568434 (2011a:65329)

DEPARTMENT OF MATHEMATICS, PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, STATE COLLEGE, PENNSYLVANIA 16802

*E-mail address:* `chabaud@math.psu.edu`

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, 206 CHURCH STREET S.E., MINNEAPOLIS, MINNESOTA 55455

*E-mail address:* `cockburn@math.umn.edu`