Hermite analogs of the lowest order Raviart–Thomas mixed method for convection–diffusion equations

V. Ruas, F. A. Radu

Hermite analogs of the lowest order Raviart-Thomas mixed method for convection-diffusion equations

V. Ruas[1,2]

[1] Sorbonne Universités, UPMC Univ Paris 06 & CNRS, UMR 7190, IJRDA, 75005 Paris, France
e-mail: *vitoriano.ruas@upmc.fr*

[2] CNPq research grant holder, Graduate school of Metrology for Quality and Innovation, PUC-Rio, Rio de Janeiro, Brazil
e-mail: *vitoriano.ruas@pq.cnpq.fr*

F. A. Radu[3]

[3] Department of Mathematics, University of Bergen, Bergen, Norway
e-mail: *florin.radu@math.uib.no*

Abstract

The Raviart-Thomas mixed finite element method of the lowest order [19] commonly known as the RT_0 method, is a well-established and popular numerical tool to solve diffusion-like problems providing flux continuity across inter-element boundaries. Douglas & Roberts extended the method to the case of more general second order boundary value problems including the convection-diffusion equations (cf. this journal [10]). The main drawback of these methods however is the poor representation of the primal variable by piecewise constant functions. The Hermite analog of the RT_0 method for treating pure diffusion phenomena proposed in [21] proved to be a valid alternative to attain higher order approximation of the primal variable, while keeping intact the matrix structure and the quality of the discrete flux variable of the original RT_0 method. Non trivial extensions of this method are studied here, that can be viewed as Hermite analogs of the two Douglas & Roberts' versions of the RT_0 method, to solve convection-diffusion equations. A detailed convergence study is carried out for one of the Hermite methods, and numerical results illustrate the performance of both of them, as compared to each other and to the corresponding mixed methods.

September 14, 2017

1 Introduction

A rather great amount of numerical solution techniques for the convection-diffusion equations are available today. Nevertheless the fact that these equations lie on the basis of the mathematical modeling of countless physical phenomena, keeps encouraging specialists in the search for efficient methodology to solve this class of problems. This is particularly true of convection dominated processes, which often reveal demerits of widespread computational techniques, even when the problem to solve is linear.

This work is primarily aimed at carrying out a complete mathematical study of the Hermite finite element method proposed in [23], to solve the convection-diffusion equations. More specifically such a method is an extension to the convection-diffusion equations of the Hermite finite element method introduced in [21] for pure diffusion problems. We recall that the latter method can be regarded as a variant of the well-established Raviart-Thomas mixed method of the lowest order, also known as the RT_0 method, to solve the diffusion equations in two- and three-dimensional space. The method of [23] in turn is a Hermite analog applied to the case of the convection-diffusion equations, of one of the RT_0 method's extensions proposed by Douglas & Roberts in this journal [10] to solve more general second order boundary-value problems. A Hermite analog of the other version of the RT_0 method due to the latter authors, also studied in [10], is considered in this work, though from a purely numerical point of

view.

Historically Hermite finite elements have mostly been used to solve fourth order partial differential equations, because minimum continuity of solution derivatives across inter-element boundaries is required in this case. However the construction of such elements can be rather laborious, as shown in [8]. It is noticeable in this respect that the recent technique of the virtual element led to feasible constructions of C^k functions for $k \geq 1$ on meshes consisting of polygonal elements of arbitrary shape [3], though by means of polynomials of rather high degree.

On the other hand Hermite interpolation has been showing to be a good alternative to solve several kinds of field problems modeled by second order boundary value problems in many respects. An outstanding demonstration of such an assertion is provided by the isogeometric analysis (IGA) introduced in the last decade (see e.g. [12]). In this case advantage is taken from data satisfying high continuity requirements supplied by CAD, for a subsequent finite element analysis. However IGA in connection with triangular or tetrahedral meshes is incipient, in spite of the undeniable geometric flexibility of this type of partitions. This is a good reason to study Hermite finite elements methods defined upon triangles or tetrahedra to solve second order partial differential equations, which are low order and easy to implement at a time. That is what we do in this work, by focusing more particularly the representation of fluxes for the simulation of phenomena or processes of the convection-diffusion type.

In practice quantities directly depending on partial derivatives of the variable in terms of which an equation is expressed, i.e., the primal variable, are often more important than this unknown itself. Among them one might quote the flux in a porous medium flow or in heat flow. As far as methods allowing to enforce the continuity of normal derivatives or normal fluxes across the boundaries of triangular or tetrahedral cells are concerned, both mixed finite elements and finite volumes have been playing a prominent role since long. In particular the RT_0 method is a popular numerical tool to solve diffusion-like problems providing flux continuity across inter-element boundaries. As recalled above, Douglas & Roberts [10] extended the method to linear second order boundary-value problems including the convection-diffusion equations. The main drawback of these methods however is the poor representation of the primal variable by piecewise constant functions. That is why many authors attempted to enhance the quality of approximation of the primal variable through post-processing or hybridization, among other techniques.

In contrast to such approaches a Hermite analog of the RT_0 method can be used in the direct solution of the diffusion equations. It proved to be a well adapted alternative to attain higher convergence rates without increasing the computational effort, as shown in [22]. This rather well-off experience encouraged the authors to further study two finite element methods of the Hermite type based on a quadratic interpolation, to solve the convection-diffusion equations. Both can be viewed as Hermite analogs of the Douglas & Roberts' extensions of the RT_0 method. However the former have either identical or better convergence properties than the latter, according to the norm under consideration. Once again this is achieved at practically the same implementation cost.

An outline of this paper is as follows. In Section 2 we introduce some notations and specify the model problem, which all the studies conducted here apply to. The method proposed in [23] is described in more detail in Section 3, where we recall that it is uniformly stable with respect to a suitable working norm. In Section 4 we apply these results, which immediately lead to first order error estimates in the same norm. We also prove, through duality techniques, that the method's convergence order in the L^2-norm is two, in contrast to the first order ones that hold for the mixed extension of the RT_0 method to convection-diffusion-reaction equations in non divergence form proposed in [10]. In Section 5 we consider a variant of the method studied in Sections 3 and 4, which can be viewed as the Hermite analog of the Douglas & Roberts mixed method [10], applying to the C-D equations in divergence form. Both methods are compared, either to each other or to the corresponding Douglas & Roberts methods, by checking their convergence properties and accuracy in different senses, in the light of numerical experiments reported in Section 6. In Section 7 we conclude with some comments on the whole work.

2 Notations and model problem

Let Ω be a bounded Lipschitz domain of $\Re^N$, $N = 2, 3$, with boundary Γ. Referring to [1], in the sequel we employ the following notations: S being a proper subset of Ω, we denote the standard norm of Sobolev spaces $H^m(S)$ (resp. $W^{m,p}(S)$ for $1 \leq p \leq \infty$, $p \neq 2$), for any non negative integer m by $\| \cdot \|_{m,S}$ (resp. $\| \cdot \|_{m,p,S}$) including $L^2(S) \equiv H^0(S)$ (resp. $L^p(S) \equiv W^{0,p}(S)$). For $m > 0$ the standard semi-norm of $H^m(\Omega)$ (resp. $W^{m,p}(\Omega)$ for $1 \leq p \leq \infty$, $p \neq 2$), i.e. the standard norm of $H_0^m(\Omega)$ (resp. $W_0^{m,p}(\Omega)$ for $1 \leq p \leq \infty$, $p \neq 2$), is denoted by $| \cdot |_m$ (resp. $| \cdot |_{m,p}$). Further for any non negative integer m and $1 \leq p \leq \infty$, $p \neq 2$, the standard norm of $W^{m,p}(\Omega)$ will be denoted by $\| \cdot \|_{m,p}$; moreover $\forall f, g \in L^2(S)$, $(f, g)_S := \int_S fg \, dS$ and $\| f \|_S := [(f, f)_S]^{1/2}$ and we set $(f, g) := \int_\Omega fg \, dx$ $\forall f, g \in L^2(\Omega)$ and $\| f \| := [(f, f)]^{1/2}$.

Let f be a given function in $L^2(\Omega)$, $\mathcal{K}$ be a tensor assumed to be constant, symmetric and positive-definite and $\mathbf{w} \in [C^0(\bar{\Omega})]^N$ denote a velocity field. In this work we study as a model the following equation:

$$\begin{cases} \text{Find } u \in H_0^1(\Omega) \text{ such that} \\ -\nabla \cdot \mathcal{K} \nabla u + \mathbf{w} \cdot \nabla u = f \text{ in } \Omega. \end{cases} \tag{1}$$

Equation (1) is assumed to have a unique solution, which is guaranteed in some important particular cases. For instance, the case where the divergence of $\mathbf{w}$ is bounded in $\bar{\Omega}$ and $C_D := \| \nabla \cdot \mathbf{w} \|_{0,\infty}$ is sufficiently small. Indeed by the Divergence Theorem we easily obtain:

$$-(\nabla \cdot \mathcal{K} \nabla u, u) + (\mathbf{w} \cdot \nabla u, u) = (\mathcal{K} \nabla u, \nabla u) - (\nabla \cdot \mathbf{w} u, u)/2 \geq \lambda_{min} \| \nabla u \|^2 - C_D \| u \|^2 /2,$$

where λ_{min} is the smallest eigenvalue of $\mathcal{K}$. It follows that whenever $\lambda_{min} > C_D C_P^2/2$, where C_P is the constant of the Friedrichs-Poincaré inequality $\| v \| \leq C_P \| \nabla v \| \, \forall v \in H_0^1(\Omega)$, (1) has a unique solution by the Lax-Milgram Theorem. Notice that in many applications $\mathbf{w}$ is solenoidal, and in this case existence and uniqueness of a solution to (1) always hold true. Another situation where this problem has a unique solution arises when a convection-diffusion process is taking place at a low Péclet number Pé (see e.g. [15]). Just to make ideas clear we can take Pé $= \| \mathbf{w} \|_{0,\infty} C_P/(2\lambda_{min})$. Then in case Pé $< 1/2$, again by the Lax-Milgram Theorem, (1) will have a unique solution.

3 A Hermite solution method

Henceforth we assume that Ω is a polygon if $N = 2$ or a polyhedron if $N = 3$, and that we are given a finite element partition $\mathcal{T}_h$ of Ω, consisting of triangles or tetrahedra according to the value of N, and belonging to a regular family of partitions (cf. [8]). h denotes the maximum diameter of the elements of $\mathcal{T}_h$.

In the following paragraphs we define two finite element spaces U_h and V_h associated with $\mathcal{T}_h$. Let $\mathbf{w}_h$ be the constant field in each element of $T \in \mathcal{T}_h$ whose value in T is $\mathbf{w}(\mathbf{x}_T)$, where $\mathbf{x}_T$ is the position vector of the centroid of T, and $\mathbf{w}_h^1$ be the standard continuous piecewise linear interpolate of $\mathbf{w}$ at the vertices of $\mathcal{T}_h$. We further introduce the operators $\Pi_T : L^2(T) \longrightarrow L^2(T)$ given by $\Pi_T[v] := \int_T v dx / meas(T)$ for $T \in \mathcal{T}_h$, and $\Pi_h : L^2(\Omega) \longrightarrow L^2(\Omega)$ by $\Pi_h[v]_{|T} = \Pi_T[v_{|T}] \; \forall T \in \mathcal{T}_h$. Now throughout this work we will work with the following,

Local algebraic structure of the Hermite finite element spaces:
Every function $v \in V_h$ (resp. $\in U_h$) is such that in each element $T \in \mathcal{T}_h$ it is expressed by

$$v_{|T} = \mathbf{x}^{\mathbf{t}} \{ \mathcal{K}^{-1} [a\mathbf{x}/2 + \mathbf{b}] \} + d, \tag{2}$$

where $\mathbf{x}$ represents the space variable, $\mathbf{b}$ is a constant vector of $\Re^N$ and a and d are two real coefficients. In every N-simplex T we associate with a quadratic function v of the form (2):

Sets $\mathcal{D}_T$ and $\mathcal{E}_T$ of local degrees of freedom for the Hermite finite element spaces:
F being an edge if $N = 2$ or a face if $N = 3$ belonging to the boundary ∂T of an N-simplex T, and $\mathbf{n}_F$ being the unit normal vector on F oriented in a given manner for each $F \subset \partial T$, we set:

$$\left\{ \begin{array}{l} \mathcal{D}_T := \{\cup_F \mathcal{U}_F\} \cup \mathcal{U}_T \text{ where} \\ \mathcal{U}_F(v) = \int_F \mathcal{K}\nabla v \cdot \mathbf{n}_F ds/meas(F); \\ \mathcal{U}_T(v) = \int_T v\, dx/meas(T). \end{array} \right. \tag{3}$$

$$\left\{ \begin{array}{l} \mathcal{E}_T := \{\cup_F \mathcal{V}_F\} \cup \mathcal{V}_T \text{ where} \\ \mathcal{V}_F(v) = \int_F (\mathcal{K}\nabla v + \mathbf{w}_h \Pi_T[v]) \cdot \mathbf{n}_F ds/meas(F); \\ \mathcal{V}_T(v) = \int_T v\, dx/meas(T). \end{array} \right. \tag{4}$$

The canonical basis functions associated with these sets of degrees of freedom are as follows. First we note that $\forall v \in V_h$ or $\in U_h$, $\nabla v_{|T}$ for $T \in \mathcal{T}_h$ is expressed by $\mathcal{K}^{-1}[a_T\mathbf{x} + \mathbf{b}_T]$ for certain $a_T \in \Re$ and $\mathbf{b}_T \in \Re^N$. Then the flux variable $\mathcal{K}\nabla v_{|T}$ is of the form $a_T\mathbf{x} + \mathbf{b}_T$, and from a well-known property of the lowest order Raviart-Thomas mixed element a_T and $\mathbf{b}_T$ can be uniquely determined for prescribed $\mathcal{V}_F(v)$ (resp $\mathcal{U}_F(v)$), $\forall F \subset \partial T$. Indeed by construction the flux variable for the Hermite element is locally defined by functions of the same form as for the lowest order Raviart-Thomas element. Once a_T and $\mathbf{b}_T$ are known, we determine the value of the additive constant d_T to complete the expression of $v_{|T}$, by enforcing the condition $\Pi_T[v] = 0$. As for the basis function corresponding to the degree of freedom $\mathcal{V}_T$ and $\mathcal{U}_T$, the values of a_T and $\mathbf{b}_T$ are obtained as follows: $a_T = 0$ for both $\mathcal{V}_T$ and $\mathcal{U}_T$ while $\mathbf{b}_T = -\mathcal{K}\mathbf{w}_h$ for $\mathcal{V}_T$ and $\mathbf{b}_T = \mathbf{0}$ for $\mathcal{U}_T$. Then the value of d_T is adjusted in such a way that the mean value of the corresponding quadratic function equals one. More precisely, for $\mathcal{V}_T$ we have $d_T = \mathcal{K}\mathbf{w}_h \cdot \mathbf{x}_T + 1$ and for $\mathcal{U}_T$ we have $d_T = 1$.
This should be enough to determine the $N + 2$ basis functions associated with a given N-simplex T, corresponding to the sets $\mathcal{U}_T$ and $\mathcal{V}_T$ of degrees of freedom, for spaces U_h and V_h respectively, since the RT_0 method is well-known (cf. [25]). However for the sake of clarity we exhibit them below.
T being an element of $\mathcal{T}_h$ let $\mathbf{x}_i^T$ be the position vector of the i-th vertex S_i^T of T, F_i^T be the face of T opposite to S_i^T and h_i^T be the length of the corresponding height of T, for $i = 1, \ldots, N + 1$. We have:

Local basis functions φ_i^T for space U_h:
The local basis function φ_i^T associated with the degree of freedom $\mathcal{U}_{F_i^T}$ and the basis function φ_{N+2}^T associated with the degree of freedom $\mathcal{U}_T$ are given by:

$$\left\{ \begin{array}{l} \varphi_i^T = \mathbf{x}^{\mathbf{t}}\{\mathcal{K}^{-1}[a_i^T\mathbf{x}/2 + \mathbf{b}_i^T]\} + d_i^T \text{ for } i = 1, \ldots, N+2, \text{ where} \\ \left. \begin{array}{l} a_i^T = [h_i^T]^{-1}, \\ \mathbf{b}_i^T = -\mathbf{x}_i^T a_i^T, \\ d_i^T = -\int_T \mathbf{x}^{\mathbf{t}}\{\mathcal{K}^{-1}[a_i^T\mathbf{x}/2 + \mathbf{b}_i^T]\}\, dx/meas(T), \end{array} \right\} \text{for } i = 1, \ldots, N+1; \\ a_{N+2}^T = 0, \\ \mathbf{b}_{N+2}^T = (0, \ldots, 0)^{\mathbf{t}}, \\ d_{N+2}^T = 1. \end{array} \right. \tag{5}$$

Local basis functions ψ_i^T for space V_h:
Akin to the case of U_h, the ψ_i^T's are functions of the form (2) for $i = 1, \ldots, N + 2$. Since by definition the mean values of all the first $N + 1$ φ_i^T's vanish, the local basis functions ψ_i^T of V_h associated with the degree of freedom $\mathcal{V}_{F_i^T}$ for $i = 1, \ldots, N + 1$, together with its local basis function ψ_{N+2}^T associated with the degree of freedom $\mathcal{V}_T$ are given by:

$$\left\{ \begin{array}{l} \psi_i^T = \varphi_i^T \text{ for } i = 1, \ldots, N+1 \\ \psi_{N+2}^T = [\mathbf{x}_T - \mathbf{x}]^{\mathbf{t}}\mathcal{K}^{-1}\mathbf{w}_h + 1. \end{array} \right. \tag{6}$$

Next we define,

Hermite finite element spaces U_h and V_h:
As already stated in our previous definitions for element's local algebraic structure, for every interface F (an inner edge for $N = 2$ and an inner face for $N = 3$) of two elements in $\mathcal{T}_h$, $\mathbf{n}_F$ is oriented in a given manner for both of them. Then every function in $v \in V_h$ (resp. $\in U_h$) is such that its restriction to every $T \in \mathcal{T}_h$ is a $N + 2$ coefficient quadratic function of the form (2), whose degrees of freedom of the type $\mathcal{V}_F$ (resp. $\mathcal{U}_F$) coincide on both sides of every interface F of a pair of elements in $\mathcal{T}_h$.

We proceed by setting the discrete variational problem (7) below, aimed at approximating (1), whose bi-linear form a_h and linear form L_h are given by (8):

Find $u_h \in U_h$ such that for all $v \in V_h$

$$a_h(u_h, v) = L_h(v), \tag{7}$$

holds, where $\forall u \in U_h$ and $\forall v \in V_h$,

$$\left\{ \begin{array}{ll} a_h(u,v) := & \displaystyle\sum_{T\in\mathcal{T}_h} [(\nabla\cdot\mathcal{K}\nabla u - \mathbf{w}_h^1\cdot\nabla u, \Pi_T[v])_T \\ & +(\nabla u, \mathcal{K}\nabla v + \mathbf{w}_h\Pi_T[v])_T + (u, \nabla\cdot\mathcal{K}\nabla v)_T]; \\ L_h(v) := & -(f, \Pi_h[v]). \end{array} \right. \tag{8}$$

It is noteworthy that whenever $\mathbf{w}_h^1 = \mathbf{w}_h$ the two terms involving $\mathbf{w}$ in the expression of a_h cancel out. Hence apparently problem's dependence on the convective velocity is not taken into account by formulation (7) if $\mathbf{w}$ happens to be constant. However one should bear in mind that even in this case such a dependence remains implicit therein, through the construction of space V_h.

Now let us consider the space

$$V := \{v | v \in H^1(\Omega); \nabla\cdot\mathcal{K}\nabla v \in L^2(\Omega)\}.$$

Clearly a_h can be extended to $(U_h + V) \times (V_h + V)$. Then we further introduce the functional $\| \cdot \|_h$: $U_h + V_h + V \longrightarrow \Re$ given by:

$$\| v \|_h := \left[(\Pi_h v, \Pi_h v) + \sum_{T\in\mathcal{T}_h} \{(\nabla v, \nabla v)_T + (\nabla\cdot\mathcal{K}\nabla v, \nabla\cdot\mathcal{K}\nabla v)_T\} \right]^{1/2}. \tag{9}$$

The expression $\| \cdot \|_h$ obviously defines a norm over V, U_h and V_h. In this manner, it is not difficult to establish the continuity of a_h over $(U_h + V) \times (V_h + V)$ with a mesh independent constant M (cf. the proof of **Proposition 3.1** hereafter):

$$a_h(u, v) \leq M \| u \|_h \| v \|_h . \tag{10}$$

On the other hand there is no way for a_h to be coercive. Hence we resort to an inf-sup condition for a_h over $U_h \times V_h$ [2], which directly implies that (7) has a unique solution. More specifically the following stability result was proved in [23].

Proposition 3.1 *([23]) If h is sufficiently small and $\mathbf{w} \in [W^{1,\infty}(\Omega)]^N$, there exists a constant $\alpha > 0$ independent of h such that*

$$\forall u \in U_h \quad \sup_{v\in V_h\setminus\{0\}} \frac{a_h(u, v)}{\| v \|_h} \geq \alpha \| u \|_h . \tag{11}$$

In the next section we derive estimates for $\| u - u_h \|_h$ using a modified Strang Lemma for non coercive problems given in [13]. In this aim we have to consider the following auxiliary problem:

Find $u_h^* \in U_h$ such that for all $v \in V_h$

$$a_h^*(u_h^*, v) = L_h(v), \tag{12}$$

holds, where $\forall u \in U_h + V$ and $\forall v \in V_h + V$,

$$\begin{aligned} a_h^*(u,v) := \quad & \sum_{T \in \mathcal{T}_h} [(\nabla \cdot \mathcal{K}\nabla u - \mathbf{w} \cdot \nabla u, \Pi_T[v])_T \\ & +(\nabla u, \mathcal{K}\nabla v + \mathbf{w}_h \Pi_T[v])_T + (u, \nabla \cdot \mathcal{K}\nabla v)_T]. \end{aligned} \tag{13}$$

Similarly to the case of problem (7) (cf. [23]) we can prove,

Theorem 3.2 *Problem (12) has a unique solution and moreover there exists a constant C^* independent of h such that*

$$\| u_h^* \|_h \leq C^* \| f \| . \tag{14}$$

Before proving **Theorem 3.2** we establish a stability result for problem (12).

Proposition 3.3 *If h is sufficiently small and $\mathbf{w} \in [W^{1,\infty}(\Omega)]^N$, there exists a constant $\alpha^* > 0$ independent of h such that*

$$\forall u \in U_h \quad \sup_{v \in V_h \setminus \{0\}} \frac{a_h^*(u,v)}{\| v \|_h} \geq \alpha^* \| u \|_h . \tag{15}$$

Proof: Given $u \in U_h$ define $v := v_1 + v_2 + v_3$, where $v_i \in V_h$ for $i = 1, 2, 3$ are defined as follows:
$v_1 = \theta_1 w_1$, θ_1 being a non negative constant to be specified, and w_1 being defined by $\Pi_h[w_1] = \Pi_h[u]$, together with $(\mathcal{K}\nabla w_1 + \mathbf{w}_h \Pi_h[w_1]) \cdot \mathbf{n}_T = (\mathcal{K}\nabla u) \cdot \mathbf{n}_T$ for every $T \in \mathcal{T}_h$, where $\mathbf{n}_T$ is the outer normal on ∂T.
v_2 equals $\theta_2 \nabla \cdot \mathcal{K}\nabla u$ in every $T \in \mathcal{T}_h$, where θ_2 is a non negative constant to be specified.
v_3 is constructed by applying Theorem 4 of [19]. According to it there exists a field $\mathbf{p} \in \mathbf{Q}_h := \{\mathbf{q} \mid \exists u \in U_h$ such that $\mathbf{q}_{|T} = \mathcal{K}\nabla u_{|T} \; \forall T \in \mathcal{T}_h\}$, satisfying for a constant $\tilde{C}$ independent of h:

$$\nabla \cdot \mathbf{p} = \Pi_h[u] \text{ in } \Omega; \qquad \| \mathbf{p} \| \leq \tilde{C} \| \Pi_h[u] \| . \tag{16}$$

Then recalling that the normal traces over the faces of the elements in $\mathcal{T}_h$ of fields belonging to $\mathbf{Q}_h$ are constant [19] v_3 is defined in such a way that $\forall T \in \mathcal{T}_h$, $(\mathcal{K}\nabla v_3 + \mathbf{w}_h \Pi_T[v_3]) \cdot \mathbf{n}_T = \mathbf{p} \cdot \mathbf{n}_T \; \forall T \in \mathcal{T}_h$ and $\Pi_h[v_3] = -\theta_1 \Pi_h[u]$.

It is clear that $\nabla \cdot \mathcal{K}\nabla w_1 = \nabla \cdot \mathcal{K}\nabla u$. Moreover by construction we have,

$$\oint_{\partial T} \mathcal{K}\nabla w_1 \cdot \mathbf{n}_T w_1 \, dS - (\nabla \cdot \mathcal{K} w_1, w_1)_T = (\mathcal{K}\nabla w_1, \nabla w_1)_T = (\mathcal{K}\nabla u - \mathbf{w}_h \Pi_T[u], \nabla w_1)_T \; \forall T \in \mathcal{T}_h. \tag{17}$$

Then λ and Λ being the smallest and the largest eigenvalue of $\mathcal{K}$, after straightforward manipulations it follows that

$$\| \nabla w_1 \|_T \leq \lambda^{-1}(\Lambda \| \nabla u \|_T + \| \mathbf{w} \|_{0,\infty} \| \Pi_T[u] \|_T). \tag{18}$$

This implies that for $\tilde{C}_1 = \lambda^{-1}(\Lambda^2 + \| \mathbf{w} \|_{0,\infty}^2)^{1/2}$, $\sum_{T \in \mathcal{T}_h} \| \nabla w_1 \|_T^2 \leq \tilde{C}_1^2 \| u \|_h^2$, which immediately yields,

$$\| v_1 \|_h \leq C_1 \| u \|_h, \text{ with } C_1 = \theta_1 (1 + \tilde{C}_1^2)^{1/2}. \tag{19}$$

As for v_2 we trivially have,

$$\| v_2 \|_h \leq C_2 \| u \|_h, \text{ with } C_2 = \theta_2. \tag{20}$$

On the other hand by construction v_3 fulfills $\nabla \cdot [\mathcal{K}\nabla v_3]_{|T} = \nabla \cdot \mathbf{p}_{|T} = \Pi_T[u], \forall T \in \mathcal{T}_h$, and hence,

$$\lambda \| \nabla v_3 \|_T^2 \leq (\mathcal{K}\nabla v_3, \nabla v_3)_T = \oint_{\partial T} (\mathcal{K}\nabla v_3) \cdot \mathbf{n}_T v_3 dS - (v_3, \nabla \cdot \mathbf{p})_T = (\theta_1 \mathbf{w}_h \Pi_T[u] + \mathbf{p}, \nabla v_3)_T \tag{21}$$

It easily follows that

$$\| v_3 \|_h \leq C_3 \| \Pi_h[u] \| \leq C_3 \| u \|_h, \text{ where } C_3 = [(\tilde{C} + \theta_1 \| \mathbf{w} \|_{0,\infty})^2 \lambda^{-2} + \theta_1^2 + 1]^{1/2}. \quad (22)$$

Now taking into account (16) and (17), after straightforward calculations we obtain,

$$a_h^*(u, v_1) = \theta_1 \sum_{T \in \mathcal{T}_h} \{2(\nabla \cdot \mathcal{K}\nabla u, \Pi_T[u])_T + (\mathcal{K}\nabla u, \nabla u)_T - (\mathbf{w} \cdot \nabla u, \Pi_T[u])_T\}; \quad (23)$$

$$a_h^*(u, v_2) = \theta_2 \sum_{T \in \mathcal{T}_h} \{\| \nabla \cdot \mathcal{K}\nabla u \|_T^2 + ([\mathbf{w}_h - \mathbf{w}] \cdot \nabla u, \nabla \cdot \mathcal{K}\nabla u)_T\}; \quad (24)$$

$$a_h^*(u, v_3) = \sum_{T \in \mathcal{T}_h} \{\| \Pi_T[u] \|_T^2 + \theta_1[(\mathbf{w} \cdot \nabla u, \Pi_T[u])_T - (\nabla \cdot \mathcal{K}\nabla u, \Pi_T[u])_T] + (\mathbf{p}, \nabla u)_T\}. \quad (25)$$

Then, recalling the definition of $\mathbf{w}_h$, there exists a mesh independent constant C_W (cf. [8]) such that $\| \mathbf{w} - \mathbf{w}_h \|_{0,\infty,T} \leq C_W h \| \nabla \mathbf{w} \|_{0,\infty,T} \; \forall T \in \mathcal{T}_h$. Using this fact, together with (23)-(24)-(25), simple manipulations lead to:

$$\left\{ \begin{array}{l} a_h^*(u, v) \geq \| \Pi_h[u] \|^2 + \displaystyle\sum_{T \in \mathcal{T}_h} \left[\theta_1 \lambda \| \nabla u \|_T^2 + \frac{\theta_2}{2} \| \nabla \cdot \mathcal{K}\nabla u \|_T^2 - \| \mathbf{p} \|_T \| \nabla u \|_T \right] \\ - \displaystyle\sum_{T \in \mathcal{T}_h} \left[\theta_1 \| \nabla \cdot \mathcal{K}\nabla u \|_T \| \Pi_T[u] \|_T) + \theta_2 C_W^2 h^2 |\mathbf{w}|_{1,\infty}^2 \| \nabla u \|_T^2 / 2 \right] \\ \geq \| \Pi_h[u] \|^2 / 4 + \displaystyle\sum_{T \in \mathcal{T}_h} \{(\theta_1 \lambda - \tilde{C}^2 - \theta_2 C_W^2 h^2 |\mathbf{w}|_{1,\infty}^2 / 2) \| \nabla u \|_T^2 \\ + (\theta_2/2 - \theta_1^2/2) \| \nabla \cdot \mathcal{K}\nabla u \|_T^2 \} \end{array} \right. \quad (26)$$

Now if we assume that $h^2 \leq \beta (C_W |\mathbf{w}|_{1,\infty})^{-2}$ with $\beta \leq 4\lambda^2 / [D + (D^2 + 8\lambda^2)^{1/2}]$ for $D = 1 + 4\tilde{C}^2$, we may choose $\theta_1 > 0$ satisfying $\theta_1 \lambda - \tilde{C}^2 - \beta\theta_2/2 \geq 1/4$ with $\theta_2 = 1/2 + \theta_1^2$. It follows from (26), (19), (20) and (22), that,

$$a_h^*(u, v) \geq \| u \|_h^2 / 4; \qquad \| v \|_h \leq C \| u \|_h, \text{ with } C = [3(C_1^2 + C_2^2 + C_3^2)]^{1/2}. \quad (27)$$

This immediately yields (15) with $\alpha^* = 1/(4C)$. ■

Proof of **Theorem 3.2**: Since V_h is a finite dimensional space, according to [2] the existence and uniqueness of a solution to (12) follows from (15). Moreover, combining (12) and (15) we easily obtain,

$$\alpha^* \| u_h^* \|_h \leq \sup_{v \in V_h \setminus \{0\}} \frac{L_h(v)}{\| \Pi_h[v] \|}. \quad (28)$$

Since $L_h(v) = \int_\Omega f \Pi_h[v] dx$ from (28) we finally derive (14) with $C^* = [\alpha^*]^{-1}$. ■

4 Convergence results

Henceforth we denote by ∇_h the operator from $V + U_h + V_h$ onto $L^2(\Omega)$ defined by

$$[\nabla_h w]_{|T} = \nabla [w_{|T}] \; \forall T \in \mathcal{T}_h, \; \forall w \in V + U_h + V_h.$$

Notice that for any function $u \in V + U_h$, $\nabla \cdot \mathcal{K}\nabla u$ is well-defined in $L^2(\Omega)$ (cf. [25]) and hence there is no need to use the operator ∇_h in this case.

In order to study the convergence of u_h to u in appropriate norms we first note that from the properties of V_h and equation (1) we easily infer that u satisfies

$$a_h^*(u,v) = L_h(v)\ \forall v \in V_h. \tag{29}$$

From the continuity of a_h^* and the uniform stability result proved in **Proposition 3.3**, we may apply the generalized First and Second Strang's inequality for the weakly coercive case, namely, inequality (32) of [13]. In the case under study this writes,

$$\| u - u_h \|_h \leq \frac{M^*}{\alpha^*} \inf_{w \in U_h} \| u - w \|_h + \frac{1}{\alpha} \sup_{v \in V_h \setminus \{0\}} \frac{[a_h^* - a_h](u_h^*, v)}{\| v \|_h} \tag{30}$$

where M^* is a constant such that

$$a_h^*(u,v) \leq M^* \| u \|_h \| v \|_h \ \ \forall u \in V + U_h,\ \forall v \in V + V_h. \tag{31}$$

Proposition 4.1 *There exists a constant M^* independent of h such that (31) holds.*

Proof: Since $(u, \nabla \cdot \mathcal{K}\nabla v)_T = (\Pi_T[u], \nabla \cdot \mathcal{K}\nabla v)_T\ \forall T \in \mathcal{T}_h$ and $\forall v \in V_h$, we trivially have,

$$\left\{ \begin{array}{l} a_h^*(u,v) \leq (\| \nabla \cdot \mathcal{K}\nabla_h u \| + \| \mathbf{w} \|_{0,\infty} \| \nabla_h u \|) \| \Pi_h[v] \| + \\ \| \nabla_h u \| (\Lambda \| \nabla_h v \| + \| \mathbf{w} \|_{0,\infty} \| \Pi_h[v] \|) + \sum_{T \in \mathcal{T}_h} \| \Pi_T[u] \|_T \| \nabla \cdot \mathcal{K}\nabla v \|_T . \end{array} \right. \tag{32}$$

(32) immediately yields (31) with $M^* = 2\max[1, 2 \| \mathbf{w} \|_{0,\infty}, \Lambda]$. ■

Next we prove the validity of the following a priori error estimate for the method under study:

Theorem 4.2 *Assume that $\mathbf{w} \in [W^{1,\infty}(\Omega)]^N$ and h is sufficiently small. Then if $u \in H^2(\Omega)$ and $f \in H^1(\Omega)$ there exists a mesh independent constant $C^{'}$ such that,*

$$\| u - u_h \|_h \leq C^{'} h[\| u \|_2 + \| f \|_1]. \tag{33}$$

Proof: By standard results applying to the RT_0 method, and since $\| \Pi_h[u - u_h] \|$ is obviously bounded above by a mesh independent constant times $h|u|_1$, for a suitable constant C_I independent of h it holds,

$$\inf_{w \in V_h} \| u - w \|_h \leq C_I h[\| u \|_2 + |f|_1]. \tag{34}$$

On the other hand we have $|[a_h^* - a_h](u_h^*, v)| = |([\mathbf{w} - \mathbf{w}_h^1] \cdot \nabla_h u_h^*, \Pi_h[v])|$. Hence for a mesh independent constant C_W^* such that $\| \mathbf{w} - \mathbf{w}_h^1 \|_{0,\infty} \leq C_W^* h |\mathbf{w}|_{1,\infty}$ we derive,

$$|[a_h^* - a_h](u_h^*, v)| \leq C_W^* h |\mathbf{w}|_{1,\infty} \| \nabla_h u_h^* \| \| v \|_h . \tag{35}$$

Taking into account (14), (30)-(34)-(35) readily yield (33), $C^{'}$ being a mesh independent constant. ■

Next we give a fundamental result of this work:

Theorem 4.3 *If Ω is convex, $\mathbf{w} \in [W^{2,4}(\Omega)]^N$, $u \in H^2(\Omega)$, $f \in H^1(\Omega)$ and h is sufficiently small, there exists a constant $C^{''}$ independent of h such that,*

$$\| u - u_h \| \leq C^{''} h^2 (\| u \|_2 + \| f \|_1). \tag{36}$$

Proof: The proof is a non-trivial extension of the proof of Theorem 2.3 in [21], where the quadratic convergence was shown for the pure diffusion case. We first observe that $\mathbf{w} \in [W^{1,\infty}(\Omega)]^N$ according to the Sobolev Embedding Theorem [1]. Moreover using the definitions of a_h and Π_h, together with the continuity of the normal components of $\mathcal{K}\nabla v_h + \mathbf{w}_h \Pi_h[v_h]$ on ∂T for $v_h \in V_h$, we easily obtain,

$$a_h(u - u_h, v_h) + ((\mathbf{w}_h^1 - \mathbf{w}) \cdot \nabla u, \Pi_h[v_h]) = 0, \tag{37}$$

for all $v_h \in V_h$. Similarly, owing to the continuity of the normal components of $\mathcal{K}\nabla u_h$ on ∂T (cf. [21]):

$$(u-u_h, \nabla\cdot\mathcal{K}\nabla v) = a_h(u-u_h, v)+(\nabla\cdot\mathcal{K}\nabla(u-u_h), v-\Pi_h[v])+((\mathbf{w}_h^1-\mathbf{w}_h)\cdot\nabla_h(u-u_h), \Pi_h[v]), \quad (38)$$

for all $v \in \{v | v \in H_0^1(\Omega), \nabla\cdot\mathcal{K}\nabla v \in L^2(\Omega)\}$. By using the Aubin-Nitsche trick and (38) we can write

$$\| u-u_h \| = \sup_{v\in D(\Omega)\setminus\{0\}} \frac{(u-u_h, \nabla\cdot\mathcal{K}\nabla v)}{\| \nabla\cdot\mathcal{K}\nabla v \|} = \sup_{v\in D(\Omega)\setminus\{0\}} \frac{a_h(u-u_h, v) + (\nabla\cdot\mathcal{K}\nabla(u-u_h), v-\Pi_h[v]) + ((\mathbf{w}_h^1-\mathbf{w}_h)\cdot\nabla_h(u-u_h), \Pi_h[v])}{\| \nabla\cdot\mathcal{K}\nabla v \|} \quad (39)$$

where $D(\Omega) := V \cap H_0^1(\Omega)$. Let now $u \neq u_h$ (if $u = u_h$ then (36) trivially holds) and $B_D(\mathbf{0}, 1) := \{v | v \in D(\Omega),\ \| \nabla\cdot\mathcal{K}\nabla v \| = 1\}$. We know that there exists $v_0 \in B_D(\mathbf{0}, 1)$ such that $\nabla\cdot\mathcal{K}\nabla v_0 = \frac{u-u_h}{\| u-u_h \|}$ (cf. [11]). Thus it is easy to see that due to (39) it holds,

$$\| u-u_h \| = a_h(u-u_h, v_0)+(\nabla\cdot\mathcal{K}\nabla(u-u_h), v_0-\Pi_h[v_0])+((\mathbf{w}_h^1-\mathbf{w}_h)\cdot\nabla_h(u-u_h), \Pi_h[v_0]). \quad (40)$$

By combining (37) and (40) we further get for any $v_h \in V_h$

$$\begin{aligned}\| u-u_h \| = a_h(u-u_h, v_0-v_h) + (\nabla\cdot\mathcal{K}\nabla(u-u_h), v_0-\Pi_h[v_0]) \\ +((\mathbf{w}_h^1-\mathbf{w}_h)\cdot\nabla_h(u-u_h), \Pi_h[v_0]) - ((\mathbf{w}_h^1-\mathbf{w})\cdot\nabla u, \Pi_h[v_h]).\end{aligned} \quad (41)$$

Since $D(\Omega) \subset H^2(\Omega)$ in case Ω is convex (cf. [11]), we can define the standard interpolate $I_h v \in V_h$ of every $v \in D(\Omega)$, based on the degrees of freedom of V_h. Taking $v_h = I_h v_0$ in (41), we get,

$$\begin{aligned}\| u-u_h \| = a_h(u-u_h, v_0-I_h v_0) + (\nabla\cdot\mathcal{K}\nabla(u-u_h), v_0-\Pi_h[v_0]) \\ +((\mathbf{w}_h^1-\mathbf{w}_h)\cdot\nabla_h(u-u_h), \Pi_h[v_0]) - ((\mathbf{w}_h^1-\mathbf{w})\cdot\nabla u, \Pi_h[I_h v_0]).\end{aligned} \quad (42)$$

By using the continuity (10) of a_h and the definition of v_0, together with the approximation properties of I_h for h sufficiently small (see [21], p. 239 for details), we have for a suitable h-independent constant $\underline{C}$,

$$a_h(u-u_h, v_0-I_h v_0) \leq \frac{3}{4}\| u-u_h \| + \underline{C}\| u-u_h \|_h [\| v_0-I_h v_0 \| + \| \nabla_h(v_0-I_h v_0) \|]. \quad (43)$$

Applying to (42) and (43) standard results to estimate $\| v_0-\Pi_h[v_0] \|$ together with $\| v_0-I_h v_0 \| + \| \nabla_h(v_0-I_h v_0) \|$, recalling (33) we easily conclude that there exists another constant $\bar{C}$ independent of h such that,

$$\| u-u_h \| \leq \bar{C}h^2(|u|_2+|f|_1)+4[|((\mathbf{w}_h^1-\mathbf{w}_h)\cdot\nabla_h(u-u_h), \Pi_h[v_0])| + |((\mathbf{w}_h^1-\mathbf{w})\cdot\nabla u, \Pi_h[I_h v_0])|]. \quad (44)$$

We proceed by estimating the two terms in brackets on the right hand side of (44) denoted by T_1 and T_2. First we note that from the Sobolev Embedding Theorem and the convexity of Ω (cf. [14]), there exist constants C_∞ and C'_∞ depending only on Ω such that

$$\| \Pi_h[v_0] \|_{0,\infty} \leq \| v_0 \|_{0,\infty} \leq C_\infty \| v_0 \|_2 \leq C'_\infty.$$

Thus for a mesh independent constant C_2 we have:

$$T_1 \leq \| \mathbf{w}_h^1-\mathbf{w}_h \| \| \nabla_h(u-u_h) \| \| \Pi_h[v_0] \|_{0,\infty} \leq C_2 h^2 |u|_2. \quad (45)$$

For deriving the estimate (45) we used the fact that both $\mathbf{w}_h^1$ and $\mathbf{w}_h$ are interpolates of $\mathbf{w}$, the result (33) and the boundedness of $\| \Pi_h[v_0] \|_{0,\infty}$. C_2 is the product of C'_∞ with another constant not depending on h and the semi-norm of $\mathbf{w}$ in $[H^1(\Omega)]^N$.

The term T_2 can be estimated in a similar way. Using now the fact that $\mathbf{w}_h^1$ is a piecewise linear interpolate of $\mathbf{w}$, the regularity of the solution u and standard properties of Π_h and I_h, there holds

$$T_2 \leq \| \mathbf{w}_h^1 - \mathbf{w} \|_{0,4} \| \nabla u \|_{0,4} \| \Pi_h[I_h v_0] \| \leq C_3 h^2 \| u \|_2, \tag{46}$$

where C_3 equals an h-independent constant times $|\mathbf{w}|_{2,4}$. Putting together (44)-(45)-(46), we obtain the quadratic convergence (36). ■

5 A variant for the equations in divergence form

Like in [10] it is possible to consider a variant of the method described in Section 3 applying to the case where the normal component of the total flux $-\mathcal{K}\nabla u + \mathbf{w}u$ is continuous across the element interfaces. In the case of the mixed formulation this corresponds to introducing the auxiliary variable $\mathbf{p}$ given by the above expression, and write the C-D equation equation (1) in divergence form, namely

Find u satisfying $u = 0$ on Γ and $\mathbf{p}$ such that

$$\left\{ \begin{array}{rcll} \nabla \cdot \mathbf{p} - \nabla \cdot \mathbf{w}\, u & = & f & \text{in } \Omega, \\ \mathbf{p} + \mathcal{K}\nabla u - \mathbf{w}u & = & 0 & \text{in } \Omega. \end{array} \right. \tag{47}$$

Recalling the space $\mathbf{H}(div;\Omega) := \{\mathbf{q} \mid \mathbf{q} \in [L^2(\Omega)]^N,\ \nabla \cdot \mathbf{q} \in L^2(\Omega)\}$, a natural weak (variational) formulation equivalent to system (47) is given in [10], that is,

Find $u \in L^2(\Omega)$ and $\mathbf{p} \in \mathbf{H}(div;\Omega)$ such that for all $v \in L^2(\Omega)$, and for all $\mathbf{q} \in \mathbf{H}(div;\Omega)$,

$$\left\{ \begin{array}{rcl} (\nabla \cdot \mathbf{p}, v) - (\nabla \cdot \mathbf{w}\, u, v) & = & (f, v) \\ (\mathcal{K}^{-1}\mathbf{p}, \mathbf{q}) - (u, \nabla \cdot \mathbf{q}) - (\mathcal{K}^{-1}\mathbf{w}\, u, \mathbf{q}) & = & 0. \end{array} \right. \tag{48}$$

The extension of RT_0 to the C-D equation considered in [10] consists of using the Raviart-Thomas interpolation of the lowest order to represent $\mathbf{p}$ and $\mathbf{q}$ - i.e. to approximate $\mathbf{H}(div;\Omega)$ -, and the space of constant functions in each element of the partition $\mathcal{T}_h$ to represent u and v. In contrast, here we shall mimic (48) by resorting to the space U_h, after adding up both relations in (48). More specifically we take in each element $T \in \mathcal{T}_h$, $\mathbf{q}_{|T} = \mathcal{K}\nabla v_{|T}$ for $v \in U_h$. Now u_h will be searched for in a space W_h defined hereafter. First we have to construct field a $\tilde{\mathbf{w}}_h$ to replace $\mathbf{w}_h$ (cf. Section 3), in order to preserve optimality of the approximation of u. In this aim it suffices that $\tilde{\mathbf{w}}_h$ be of the form $c\mathbf{x} + \mathbf{d}$ in each $T \in \mathcal{T}_h$ for suitable real number c and real vector $\mathbf{d} = [d_1, \ldots, d_N]^{\mathbf{t}}$. This representation is compatible with the requirement that the normal component of the flux variable $\mathbf{p} = -\mathcal{K}\nabla + \mathbf{w}u$ be continuous across the mesh edges at discrete level. The natural choice of $\tilde{\mathbf{w}}_h$ is certainly the interpolate of $\mathbf{w}$ in the Raviart-Thomas (RT_0) space. Now we define the

Hermite finite element space W_h:

W_h is the space of functions v of the form $\mathbf{x}^{\mathbf{t}}\{\mathcal{K}^{-1}[a\mathbf{x}/2 + \mathbf{b}]\} + d$ in every $T \in \mathcal{T}_h$, such that the mean normal flux $\int_F (-\mathcal{K}\nabla v + \tilde{\mathbf{w}}_h \Pi_h[v]) \cdot \mathbf{n}_F ds/meas(F)$ is continuous across all the inner edges or faces F of the partition. This is about all that is needed to complete the definition of W_h. Indeed using a procedure very similar to the one in Section 3 (cf. (5)) it is possible to uniquely determine the $N+2$ local basis functions η_i^T for each N-simplex $T \in \mathcal{T}_h$ related to the above set of degrees of freedom completed with the function mean value in T, defining W_h locally. More precisely, setting $[\tilde{\mathbf{w}}_h]_{|T} = a_w^T\mathbf{x} + \mathbf{b}_w^T$, since by definition $\Pi_T \eta_i^T = 0$ for $i = 1, \ldots, N+1$ and $\Pi_T \eta_{N+2}^T = 1$, recalling (5) we have:

$$\left\{ \begin{array}{l} \eta_i^T = \varphi_i^T \text{ for } i = 1, \ldots, N+1, \\ \eta_{N+2}^T = \mathbf{x}^{\mathbf{t}}\{\mathcal{K}^{-1}[a_w^T\mathbf{x}/2 + \mathbf{b}_w^T]\} + 1 - \int_T \mathbf{x}^{\mathbf{t}}\{\mathcal{K}^{-1}[a_w^T\mathbf{x}/2 + \mathbf{b}_w^T]\} dx/meas(T). \end{array} \right. \tag{49}$$

Now we replace in (48) :

- u with $\Pi_h[u_h]$;

- $\mathbf{w}$ with $\tilde{\mathbf{w}}_h$;
- $\mathbf{p}$ with $-\mathcal{K}\nabla_h u_h + \tilde{\mathbf{w}}_h \Pi_h[u_h]$ (taking $u_h \in W_h$);
- $\mathbf{q}$ with $-\mathcal{K}\nabla_h v$ (taking $v \in U_h$);
- f with $\Pi_h[f]$.

This leads to the following equation:

$$\begin{array}{l} \sum_{T\in\mathcal{T}_h} [(\nabla\cdot\{\mathcal{K}\nabla u_h - \tilde{\mathbf{w}}_h\Pi_T[u_h]\}, v)_T + (\nabla\cdot\tilde{\mathbf{w}}_h\Pi_T[u_h], v)_T + \\ (\mathcal{K}\nabla u_h - \tilde{\mathbf{w}}_h\Pi_T[u_h], \nabla v)_T + (\Pi_T[u_h], \nabla\cdot\mathcal{K}\nabla v)_T + \\ (\tilde{\mathbf{w}}_h\Pi_T[u_h], \nabla v)_T] = -(\Pi_h[f], v)\ \forall v \in U_h. \end{array} \tag{50}$$

After straightforward simplifications, and taking into account that $(\Pi_h[f], v) = (f, \Pi_h[v])$, we come up with the following Hermite finite element counterpart of (1):

Find $u_h \in W_h$ such that for all $v \in U_h$

$$\tilde{a}_h(u_h, v) = L_h(v), \tag{51}$$

where $\forall u \in V + W_h$ and $\forall v \in V + U_h$,

$$\left\{ \begin{array}{l} \tilde{a}_h(u,v) := \sum_{T\in\mathcal{T}_h} [(\nabla\cdot\mathcal{K}\nabla u, v)_T + (\nabla u, \mathcal{K}\nabla v)_T + (u, \nabla\cdot\mathcal{K}\nabla v)_T] \\ L_h(v) := -(f, \Pi_h[v]). \end{array} \right. \tag{52}$$

At a first glance (52) seems to indicate that the velocity $\mathbf{w}$ does not appear in formulation (51). Nonetheless $\mathbf{w}$ remains implicit therein through the definition of space W_h.

The fact that problem (51) has a unique solution can be established quite similarly to problem (7). The convergence results that hold for this method can be proved very much like in the case of the method defined in Section 3. The main difference is that it is necessary to require a little more regularity of $\nabla\cdot\mathbf{w}$, namely, that this function lies in $W^{1,\infty}(\Omega)$. Apart from this assumption, the results are qualitatively equivalent, in the sense that a priori error estimates completely analogous to those of **Theorem 4.2** and **Theorem 4.3** apply to problem (51) as well. As far as this work is concerned, resulting properties among others we have not formally established here, are illustrated by means of numerical examples given in the following section.

6 Numerical experiments

In this section we present some numerical results obtained with the methods described in Sections 3 and 5 for two test problems, which particularly highlight their behavior. The following nomenclature is used for the different numerical methods being experimented:

- Method A - Douglas & Roberts version in non divergence form of mixed method RT_0;
- Method HA - Hermite analog of Method A (cf. Section 3);
- Method B - Douglas & Roberts version in divergence form of mixed method RT_0;
- Method HB - Hermite analog of Method B (cf. Section 5).

In this section we will denote by e_h the error function $u - u_h$, and by $\Delta_h g$ the function defined by $[\Delta_h g]_{|T} = \Delta[g_{|T}]\ \forall T \in \mathcal{T}_h$, for any function g whose laplacian is well defined in the interior of every mesh element. Moreover the expression *pseudo maximum norm* will stand for the maximum absolute

value of a function at the centroids of the mesh elements.

Test-problem 1: In these experiments Ω is the unit square and a manufactured solution u is given by $u(x_1, x_2) = (x_1 - x_1^2)(x_2 - x_2^2)/4$. This together with the choice $\mathcal{K} = \mathcal{I}$ and $\mathbf{w} =$Pé$[x_1^2, x_2^2]^{\mathrm{t}}/\sqrt{2}$ where Pé is the Péclet number, produces a right hand side datum f. A sequence of uniform meshes was employed with $2L^2$ triangles, for $L = 8, 16, 32, 64$, constructed by first subdividing Ω into L^2 equal squares and then each one of these squares into two triangles by means of their diagonals parallel to the line $x_1 = x_2$. Quite abusively we denote by h the spacing $1/L$.

In Tables 1, 2, 3 and 4 we display the absolute errors in four different respects for increasing values of L, of the approximate solutions obtained with methods A, HA, B and HB for Pé$= 1$. More precisely the absolute errors of u, ∇u and $\Delta u = \nabla \cdot \mathcal{K} \nabla u$ measured in the norm of $L^2(\Omega)$, and in the pseudo maximum norm, are shown in Tables 1, 2, 3, 4, respectively. In Tables 5, 6, 7, 8 the same kind of results are displayed for Pé$= 100$.

h	Method A	Method B	Method HA	Method HB
1/8	0.10909896 x 10^{-2}	0.10917040 x 10^{-2}	0.17683545 x 10^{-3}	0.17991224 x 10^{-3}
1/16	0.54815217 x 10^{-3}	0.54824708 x 10^{-3}	0.44953727 x 10^{-4}	0.45890562 x 10^{-4}
1/32	0.27439727 x 10^{-3}	0.27440932 x 10^{-3}	0.11286733 x 10^{-4}	0.11531497 x 10^{-4}
1/64	0.13723841 x 10^{-3}	0.13723992 x 10^{-3}	0.28250216 x 10^{-5}	0.28866263 x 10^{-5}

Table 1: Test-problem 1 with Pé=1 - L^2 errors of u for methods A, B, HA, HB

h	Method A	Method B	Method HA	Method HB
1/8	0.45950033 x 10^{-2}	0.47137052 x 10^{-2}	0.46006556 x 10^{-2}	0.47021526 x 10^{-2}
1/16	0.23211704 x 10^{-2}	0.23834158 x 10^{-2}	0.23219012 x 10^{-2}	0.23772504 x 10^{-2}
1/32	0.11636060 x 10^{-2}	0.11951094 x 10^{-2}	0.11636981 x 10^{-2}	0.11919738 x 10^{-2}
1/64	0.58218263 x 10^{-3}	0.59798255 x 10^{-3}	0.58219418 x 10^{-3}	0.59640799 x 10^{-3}

Table 2: Test-problem 1 with Pé=1 - L^2 errors of ∇u for methods A, B, HA, HB

h	Method A	Method B	Method HA	Method HB
1/8	0.12129518 x 10^{-1}	0.12201600 x 10^{-1}	0.12131302 x 10^{-1}	0.12201600 x 10^{-1}
1/16	0.60914089 x 10^{-2}	0.61274572 x 10^{-2}	0.60916246 x 10^{-2}	0.61274572 x 10^{-2}
1/32	0.30490236 x 10^{-2}	0.30670547 x 10^{-2}	0.30490504 x 10^{-2}	0.30670547 x 10^{-2}
1/64	0.15249263 x 10^{-2}	0.15339429 x 10^{-2}	0.15249297 x 10^{-2}	0.15339429 x 10^{-2}

Table 3: Test-problem 1 with Pé=1 - L^2 errors of Δu for methods A, B, HA, HB

From Tables 1 through 8 one can infer that:

- Methods A and HA are fairly equivalent to Methods B and HB in all respects for a low Péclet number.
- Methods A and HA are superior to Methods B and HB in all respects when the Péclet number is not low.

h	Method A	Method B	Method HA	Method HB
1/8	0.10886098 x 10^{-3}	0.15656954 x 10^{-3}	0.16950256 x 10^{-3}	0.24286177 x 10^{-3}
1/16	0.29448711 x 10^{-4}	0.47404901 x 10^{-4}	0.44100102 x 10^{-4}	0.66030394 x 10^{-4}
1/32	0.78792148 x 10^{-5}	0.12849987 x 10^{-4}	0.11178619 x 10^{-4}	0.17190429 x 10^{-4}
1/64	0.20428256 x 10^{-5}	0.33363926 x 10^{-5}	0.28130033 x 10^{-5}	0.43809703 x 10^{-5}

Table 4: Test-problem 1 with Pé=1 - Maximum errors of u at centroids for methods A, B, HA, HB

h	Method A	Method B	Method HA	Method HB
1/8	0.10964223 x 10^{-2}	0.37372813 x 10^{-2}	0.27608031 x 10^{-3}	0.45450302 x 10^{-2}
1/16	0.54833181 x 10^{-3}	0.12935252 x 10^{-2}	0.46012246 x 10^{-4}	0.14346673 x 10^{-2}
1/32	0.27441391 x 10^{-3}	0.41669269 x 10^{-3}	0.10350476 x 10^{-4}	0.37772184 x 10^{-3}
1/64	0.13724039 x 10^{-3}	0.15861798 x 10^{-3}	0.25386722 x 10^{-5}	0.95399515 x 10^{-4}

Table 5: Test-problem 1 with Pé=100 - L^2 errors of u for methods A, B, HA, HB

h	Method A	Method B	Method HA	Method HB
1/8	0.59939985 x 10^{-2}	0.12889435 x 10^{+0}	0.77493931 x 10^{-2}	0.11984783 x 10^{+0}
1/16	0.25481792 x 10^{-2}	0.59439183 x 10^{-1}	0.28082657 x 10^{-2}	0.56152848 x 10^{-1}
1/32	0.11934791 x 10^{-2}	0.27928892 x 10^{-1}	0.12388051 x 10^{-2}	0.26448385 x 10^{-1}
1/64	0.58595099 x 10^{-3}	0.13745079 x 10^{-1}	0.59256341 x 10^{-3}	0.13029605 x 10^{-1}

Table 6: Test-problem 1 with Pé=100 - L^2 errors of ∇u for methods A, B, HA, HB

h	Method A	Method B	Method HA	Method HB
1/8	0.17192369 x 10^{+0}	0.24082822 x 10^{+0}	0.67758963 x 10^{-1}	0.24082822 x 10^{+0}
1/16	0.10035771 x 10^{+0}	0.12162465 x 10^{+0}	0.59883623 x 10^{-1}	0.12162465 x 10^{+0}
1/32	0.51102598 x 10^{-1}	0.60964240 x 10^{-1}	0.42990111 x 10^{-1}	0.60964240 x 10^{-1}
1/64	0.15249263 x 10^{-2}	0.15339429 x 10^{-2}	0.15249297 x 10^{-2}	0.15339429 x 10^{-2}

Table 7: Test-problem 1 with Pé=100 - L^2 errors of Δu for methods A, B, HA, HB

h	Method A	Method B	Method HA	Method HB
1/8	0.54600580 x 10^{-3}	0.70291040 x 10^{-2}	0.55050954 x 10^{-3}	0.81134902 x 10^{-2}
1/16	0.88757842 x 10^{-4}	0.20869472 x 10^{-2}	0.11114041 x 10^{-3}	0.24367046 x 10^{-2}
1/32	0.14427021 x 10^{-4}	0.54988868 x 10^{-3}	0.21157560 x 10^{-4}	0.65466283 x 10^{-3}
1/64	0.26841141 x 10^{-5}	0.13998920 x 10^{-3}	0.37752993 x 10^{-5}	0.16575725 x 10^{-3}

Table 8: Test-problem 1 with Pé=100 - Maximum errors of u at centroids for methods A, B, HA, HB

- The theoretical results of Section 4 for Method HA were confirmed in the case of both a low and a moderate Péclet number.
- The numerical convergence rate in the pseudo maximum norm is approximately two for all the four methods for Pé=1;
- The numerical convergence rate for Pé=100 in the pseudo maximum norm is significantly greater than two for methods A and HA, but remains close to two for methods B and HB.

- For Pé$=100$ the mixed methods are a little more accurate than their Hermite counterparts, as far as errors at the triangle centroids are concerned.

- Both ∇u and Δu tend to be equally approximated by a mixed method and its Hermite counterpart.

The numerical results for Methods B and HB deteriorated substantially as we switched to higher Péclet numbers, which was partially the case of Method HA, while most of the results obtained with Method A remained quite reasonable. Taking $L = 64$ we illustrate the behaviour of Methods A and HA in Tables 9 and 10, respectively, for increasing Péclet numbers. More precisely we took Pé$= 10^{2k}$ for $k = 0, 1, 2, 3$, for which we display the absolute errors of u, ∇u, Δu measured in the norm of $L^2(\Omega)$ and the absolute error of u measured in the pseudo maximum norm.

Pé	$\|e_h\|$	$\|\nabla_h e_h\|$	$\|\Delta_h e_h\|$	$\max_T \|e_h(\mathbf{x}_T)\|$
1	0.13723841 x 10^{-3}	0.58218263 x 10^{-3}	0.15249263 x 10^{-2}	0.20428256 x 10^{-5}
100	0.13724039 x 10^{-3}	0.58595099 x 10^{-3}	0.25587661 x 10^{-1}	0.26841141 x 10^{-5}
10000	0.18370239 x 10^{-2}	0.51526514 x 10^{+0}	0.15093095 x 10^{+3}	0.52745011 x 10^{-1}
1000000	0.20738981 x 10^{-3}	0.57979017 x 10^{-1}	0.22150639 x 10^{+2}	0.12137294 x 10^{-2}

Table 9: Absolute errors for Test-problem 1 solved by Method A with $h = 1/64$

Pé	$\|e_h\|$	$\|\nabla_h e_h\|$	$\|\Delta_h e_h\|$	$\max_T \|e_h(\mathbf{x}_T)\|$
1	0.28250216 x 10^{-5}	0.58219418 x 10^{-3}	0.15249297 x 10^{-2}	0.28130033 x 10^{-5}
100	0.25386722 x 10^{-5}	0.59256341 x 10^{-3}	0.24402972 x 10^{-1}	0.37752993 x 10^{-5}
10000	0.10089341 x 10^{+4}	0.26546637 x 10^{+6}	0.75214576 x 10^{+7}	0.57326794 x 10^{+4}
1000000	0.13681695 x 10^{+0}	0.26569231 x 10^{+2}	0.66544379 x 10^{+2}	0.27070200 x 10^{+1}

Table 10: Absolute errors for Test-problem 1 solved by Method HA with $h = 1/64$

Tables 9 and 10 indicate that Method A is more stable than its Hermite counterpart, as Pé increases.

Test-problem 2: In order to observe the behaviour of the four methods being checked, in the presence of a curved boundary, in this test-problem the domain is a disk with unit radius. The manufactured solution u is given by $u(x_1, x_2) = (1 - x_1^2 - x_2^2)/4$. Taking again $\mathcal{K} = \mathcal{I}$, the right hand side function $f = 1 - (x_1^2 + x_2^2)/2$ corresponds to a convective velocity $\mathbf{w} =$Pé$[x_1, x_2]^{\mathrm{t}}$. For symmetry reasons the computational domain Ω is only the quarter of disk given by $x_1 > 0$ and $x_2 > 0$. A sequence of quasi-uniform meshes with $2L^2$ triangles was employed for $L = 8, 16, 32, 64$, constructed by mapping the meshes of Test-problem 1 into the actual meshes of Ω using the transformation of cartesian into polar coordinates in the way described in [20]. For this procedure we have $h = 1/L$. We denote by Ω_h the approximation of Ω consisting of the union of the triangles in $\mathcal{T}_h$.
In Tables 11, 12, 13, 14 we display the absolute errors in four different respects for increasing values of L, of the approximate solutions obtained with methods A, HA, B and HB for Pé$= 1$. The errors and corresponding notations are the same as in Tables 1 through 8. More precisely the absolute errors of u, ∇u and $\Delta u = \nabla \cdot \mathcal{K} \nabla u$ measured in the norm of $L^2(\Omega_h)$ together with the pseudo maximum norm are shown in Tables 11, 12, 13, 14, respectively.

From these tables we infer that:

- Methods A and HA are superior to Methods B and HB in all respects, except in the approximation of (constant) Δu, which is almost exactly approximated by all the four methods.

h	Method A	Method B	Method HA	Method HB
1/8	0.92106378 x 10^{-2}	0.12088608 x 10^{-1}	0.35534315 x 10^{-3}	0.82905447 x 10^{-2}
1/16	0.46131717 x 10^{-2}	0.91517998 x 10^{-2}	0.88950447 x 10^{-4}	0.80213268 x 10^{-2}
1/32	0.23075591 x 10^{-2}	0.82633266 x 10^{-2}	0.22251853 x 10^{-4}	0.79642825 x 10^{-2}
1/64	0.11539009 x 10^{-2}	0.80263174 x 10^{-2}	0.55709298 x 10^{-5}	0.79504439 x 10^{-2}

Table 11: Test-problem 2 with Pé=1 - L^2 errors of u for methods A, B, HA, HB

h	Method A	Method B	Method HA	Method HB
1/8	0.93229273 x 10^{-8}	0.49035715 x 10^{-1}	0.93226950 x 10^{-8}	0.49803379 x 10^{-1}
1/16	0.93341685 x 10^{-8}	0.48942597 x 10^{-1}	0.93341123 x 10^{-8}	0.49096187 x 10^{-1}
1/32	0.93369805 x 10^{-8}	0.48911400 x 10^{-1}	0.93369665 x 10^{-8}	0.48942095 x 10^{-1}
1/64	0.93376835 x 10^{-8}	0.48902954 x 10^{-1}	0.93376800 x 10^{-8}	0.48909457 x 10^{-1}

Table 12: Test-problem 2 with Pé=1 - L^2 errors of ∇u for methods A, B, HA, HB

h	Method A	Method B	Method HA	Method HB
1/8	0.26390408 x 10^{-7}	0.26390408 x 10^{-7}	0.26390250 x 10^{-7}	0.26390408 x 10^{-7}
1/16	0.26406317 x 10^{-7}	0.26406317 x 10^{-7}	0.26406293 x 10^{-7}	0.26406317 x 10^{-7}
1/32	0.26410294 x 10^{-7}	0.26410294 x 10^{-7}	0.26410290 x 10^{-7}	0.26410294 x 10^{-7}
1/64	0.26411289 x 10^{-7}	0.26411289 x 10^{-7}	0.26411287 x 10^{-7}	0.26411289 x 10^{-7}

Table 13: Test-problem 2 with Pé=1 - L^2 errors of Δu for methods A, B, HA, HB

h	Method A	Method B	Method HA	Method HB
1/8	0.12068400 x 10^{-3}	0.22261595 x 10^{-1}	0.40130052 x 10^{-3}	0.25015968 x 10^{-1}
1/16	0.30221612 x 10^{-4}	0.24492318 x 10^{-1}	0.10040660 x 10^{-3}	0.25368398 x 10^{-1}
1/32	0.75493397 x 10^{-5}	0.25794826 x 10^{-1}	0.25126450 x 10^{-4}	0.26020645 x 10^{-1}
1/64	0.18777283 x 10^{-5}	0.26328682 x 10^{-1}	0.63028673 x 10^{-5}	0.26385490 x 10^{-1}

Table 14: Test-problem 2 with Pé=1 - Maximum errors of u at centroids for methods A, B, HA, HB

- Methods A and HA do not seem to be affected by the curved boundary approximation by polygons, while this seems to be case of Methods B and HB.
- Method A and HA approximate both ∇u and Δu to machine precision; this is an expected behavior since both functions in this test-problem can be exactly represented by the same underlying incomplete linear and constant interpolation for both methods.
- The approximations of u by Method HA converge as an $O(h^2)$ in $L^2(\Omega_h)$, while those computed by Method A converge as an $O(h)$, the best we can hope for.
- The numerical convergence rate in the pseudo maximum norm is approximately two for both Method A and Method HA with an advantage of the former over the latter in terms of accuracy.

Akin to Test-problem 1, we checked the behaviour of Methods A and HA as the Péclet number increases. Here again we took $L = 64$ and Pé$= 10^{2k}$ for $k = 0, 1, 2, 3$. The resulting errors measured in the same manner as in Tables 9 and 10 are displayed in Table 15 for Method A and in Table 16 for

Method HA.

Pé	$\|e_h\|_{\Omega_h}$	$\|\nabla_h e_h\|_{\Omega_h}$	$\|\Delta_h e_h\|_{\Omega_h}$	$\max_T \|e_h(\mathbf{x}_T)\|$
1	0.11539009 x 10^{-2}	0.93376835 x 10^{-8}	0.26411289 x 10^{-7}	0.18777283 x 10^{-5}
100	0.11539009 x 10^{-2}	0.93376819 x 10^{-8}	0.26411285 x 10^{-7}	0.18777283 x 10^{-5}
10000	0.11539009 x 10^{-2}	0.93377017 x 10^{-8}	0.26411346 x 10^{-7}	0.18777283 x 10^{-5}
1000000	0.11539009 x 10^{-2}	0.96367921 x 10^{-8}	0.34526072 x 10^{-7}	0.18774905 x 10^{-5}

Table 15: Absolute errors for Test-problem 2 solved by Method A with $h = 1/64$

Pé	$\|e_h\|_{\Omega_h}$	$\|\nabla_h e_h\|_{\Omega_h}$	$\|\Delta_h e_h\|_{\Omega_h}$	$\max_T \|e_h(\mathbf{x}_T)\|$
1	0.55709298 x 10^{-5}	0.93376800 x 10^{-8}	0.26411287 x 10^{-7}	0.63028673 x 10^{-5}
100	0.55708924 x 10^{-5}	0.92319279 x 10^{-8}	0.26157933 x 10^{-7}	0.63028012 x 10^{-5}
10000	0.55694727 x 10^{-5}	0.52231986 x 10^{-8}	0.15104352 x 10^{-7}	0.62996183 x 10^{-5}
1000000	0.55709889 x 10^{-5}	0.93376916 x 10^{-8}	0.26406856 x 10^{-7}	0.63029895 x 10^{-5}

Table 16: Absolute errors for Test-problem 2 solved by Method HA with $h = 1/64$

From Tables 15 and 16 we observe that both Method A and Method HA are accurate to machine precision, irrespective of the Péclet number, as far as the approximations of ∇u and Δu are concerned. The approximations of u in $L^2(\Omega_h)$ and in the pseudo maximum norm do not seem to be affected by the Péclet number either in this test-problem for both methods. In the former sense Method HA is much more accurate than Method A as expected, while in the latter sense Method A is slightly more precise than Method HA. Notice that this test-problem is a little peculiar, since the exact solution is a quadratic function, whose gradient can be exactly represented by the gradient of the underlying interpolating functions. Actually this also happens to the approximation of the function itself by Method HA, but in this case other sources of errors came into play, such as numerical integration (see also Remark 3 hereafter).

Test-problem 3: Finally, we consider a test problem with a non-polynomial analytical solution. The computational domain Ω is the unit square with boundary layers along $x = 1$ and $y = 1$. We are solving equation (1) with homogeneous Dirichlet boundary conditions, $\mathbf{w} = (1, 1)^T$ and $\mathcal{K} = v\mathcal{I}$ for $v = \dfrac{1}{100}$. The right hand side is given by $f(x, y) := s_v(x)g(y) + s_v(y)g(x) + sin(\pi x) + sin(\pi y)$, where $s_v : [0, 1] \to \mathbb{R}$ is the solution of the equation $-v[s_v]'' + [s_v]' = 1$ on $[0, 1]$ and satisfying $s_v(0) = s_v(1) = 0$. The function $g : [0, 1] \to \mathbb{R}$ is defined through $g(z) := \pi[cos(\pi z) + \pi v sin(\pi z)]$. In this setting, equation (1) admits the analytical solution $u(x, y) = s_v(x)sin(\pi y) + s_v(y)sin(\pi x)$. The computations were done with uniform $2 \times n \times n$ meshes, for $n = 25, n = 50$ and $n = 100$, corresponding to a mesh diameter $h = \sqrt{2}/25, \sqrt{2}/50$ and $\sqrt{2}/100$. The results are presented in Tables 17-18 below.

h	$\|e_h\|$	$\|\nabla_h(\mathcal{K}e_h)\|$	$\|\nabla_h \cdot \mathcal{K}\nabla_h e_h\|$	$\max_T \|e_h(\mathbf{x}_T)\|$
$\sqrt{2}/25$	0.88378563 x 10^{-1}	0.77159915 x 10^{-1}	0.78473709 x 10^{1}	0.38808847 x 10^{0}
$\sqrt{2}/50$	0.37831058 x 10^{-1}	0.39786417 x 10^{-1}	0.50624144 x 10^{1}	0.13186931 x 10^{0}
$\sqrt{2}/100$	0.18041177 x 10^{-1}	0.20250530 x 10^{-1}	0.27814490 x 10^{1}	0.41014836 x 10^{-1}

Table 17: Absolute errors for Test-problem 3 solved by Method A

h	$\|e_h\|$	$\|\nabla_h(\mathcal{K}e_h)\|$	$\|\nabla_h \cdot \mathcal{K}\nabla_h e_h\|$	$\max_T \|e_h(\mathbf{x}_T)\|$
$\sqrt{2}/25$	0.27252781x 10^{7}	0.24839776 x 10^{7}	0.21605994 x 10^{9}	0.25252501 x 10^{8}
$\sqrt{2}/50$	0.54136232 x 10^{-1}	0.73013269 x 10^{-1}	0.91471443 x 10^{1}	0.45183509 x 10^{0}
$\sqrt{2}/100$	0.10370033 x 10^{-1}	0.23350942 x 10^{-1}	0.31967471 x 10^{1}	0.10196511 x 10^{0}

Table 18: Absolute errors for Test-problem 3 solved by Method HA

We remark that for $n = 25$ and $n = 50$ method A is more accurate than its Hermite variant HA in the L^2 sense. However, for $n = 100$ the opposite occurs, and the expected second order convergence starts playing a role. The method A converges only linearly (in the L^∞ sense this is less clear up to this level of refinement). The H^1 errors are very big owing to the sharp boundary layers. The conclusion is that, for a moderate Péclet number the Hermite method HA is the best option, provided an affordable fine mesh is used. Of course if the Péclet number increases some stabilization is required, like for any other method, but addressing this issue is beyond the aim of this paper. Notice that the results for $n = 25$ indicate that the assumption on the magnitude of h for stability of method HA is not superfluous.

7 Concluding remarks

We summarize this work with a few conclusions and remarks.

Remark 1 The Hermite method described in Section 3 works as well as the corresponding Douglas and Roberts extension of the RT_0 element, as far as the fluxes are concerned. On the other hand the former behaves much better in terms of the error of the primal variable in $L^2(\Omega)$. ■

Remark 2 According to the theoretical results derived in this work, numerical convergence in case the Péclet number is high could only be observed for meshes much finer than those used in Test-problem 1 and in [23]. However running tests with such meshes may become unrealistic. Therefore the authors intend to study modifications of the variational formulations employed in this work, in order to obtain stable solutions within acceptable accuracy, even in the case of high Péclet numbers, without resorting to excessive mesh refinement. The work of Park and Kim [16] for RT_0 discretizations could be a inspiring one in this connection. ■

Remark 3 The bilinear forms and the linear form L_h considered in this work do not really reduce to those in [21] in the case where $\mathbf{w} \equiv \mathbf{0}$. This is because somehow we wanted to incorporate numerical quadrature to the variational formulations in use, which is mandatory if f is not easy to integrate. However, in this case we should rather take $L_h(v) := (f_h, v)$ for a suitable f_h defined through point values of f only. Assuming for instance that $f \in H^2(\Omega)$, f_h can be chosen to be a piecewise linear interpolate of f in every $T \in \mathcal{T}_h$. Second order convergence results in $L^2(\Omega)$ can still be proven to hold for such a choice, using the well-known analysis of variational crimes [24] and [13]. The same qualitative results can also be obtained by using a suitable quadrature formula to compute the integral of the function $g := fv$ in every element of the mesh. For more details we also refer to [24], or to many other text books on the finite element method. ■

Remark 4 The Hermite methods studied in this work can be viewed as a technique to improve the accuracy of the primal variable computations with mixed element RT_0 without resorting to post-processing (see e.g. [7], [17]) or hybridization (see e.g. [4] for the diffusion equation and [18] for convection-diffusion problems). Incidentally the method proposed in [18] can be applied also to BDM elements to obtain optimal estimates [5] improving in this way the classical BDM method for convection-diffusion equations [9] or for stabilization purposes [18, 6] like in several previous work on the subject. Notice that our method allows to achieve better accuracy directly from the numerical solution procedure, at

negligible additional cost. Thus it seems worthwhile searching for Hermite analogs of BDM methods as well in the future. ■

Remark 5 A comparison between the methods described in Section 3 (Method HA) and in Section 5 (Method HB), whose study was the main object of this work, advocates in favor of the former in all respects. ■

Remark 6 As a by-product of the numerical experimentation presented in Section 6 the two versions of the Douglas & Roberts extension of the RT_0 method were also compared to each other, in the framework of the solution of convection-diffusion equations at different Péclet numbers. We observed that the method in non divergence form (Method A) is substantially more accurate and reliable than the method in divergence form (Method B). To the best of authors' knowledge this kind of numerical study had not been carried out before. ■

Acknowledgment: The first author is thankful for the financial support provided by CNPq through grant 307996/2008-5 and the second author gratefully acknowledges the support of Statoil through the Akademia agreement.

References

[1] R.A. Adams, Sobolev Spaces, Academic Press, New York, 1975.

[2] I. Babuška, The finite element method with Lagrange multipliers, Numerische Mathematik, 20 (1973), 179-192.

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

[4] F. Brezzi and M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag, N.Y., 1991.

[5] F. Brunner, F. A. Radu, M. Bause and P. Knabner, Optimal order convergence of a modified BDM1 mixed finite element scheme for reactive transport in porous media., Adv. Water Resour. 35 (2012), 163–171.

[6] F. Brunner, F. A. Radu and P. Knabner, Analysis of an upwind-mixed hybrid finite element method for transport problems, SIAM J. Numer. Anal. 52 (2014), 83-102.

[7] E. Burman and B. Stamm, Bubble stabilized discontinuous Galerkin method for parabolic and elliptic problems, Numerische Mathematik 116 (2010), 213-241.

[8] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North Holland, Amsterdam, 1978.

[9] A. Demlow, Suboptimal and Optimal Convergence in Mixed Finite Element Methods, SIAM J. Numer. Anal. 39 (2002), 1938–1953.

[10] J. Douglas Jr and J.E. Roberts, Mixed finite element methods for second order elliptic problems, Computational and Applied Mathematics 1 (1982), 91-103.

[11] L. C. Evans, **Partial Differential Equations**, Graduate Studies in Mathematics 19 (2nd ed.), Providence, RI, American Mathematical Society, 2010.

[12] J.A. Cottrell, J. Austin, T.J.R. Hughes and Y. Bazilevs, Isogeometric Analysis: Toward Integration of CAD and FEA, John Wiley & Sons, 2009.

[13] J.A. Cuminato and V. Ruas, Unification of distance inequalities for linear variational problems, Computational and Applied Mathematics, 34-3 (2015), 1009-1033.

[14] P. Grisvard, Elliptic Problems in Non Smooth Domains, Pitman, 1985.

[15] T. Ikeda, **Maximum Principle in Finite Element Models for Convection-diffusion Phenomena**, Lecture Notes in Numerical ans Applied Analysis Vol. 4, H. Fujita and M. Yamaguti eds, North-Holland Mathematical Studies 76, North Holland/Kinokuniya, 1983.

[16] D. Kim and E.-J. Park, A posteriori error estimators for the upstream weighting mixed methods for convection-diffusion problems, Computer Methods in Applied Mechanics and Engineering 197 (2008), 806-820.

[17] C. Lovadina and R. Stenberg, A posteriori error analysis of mixed finite element methods for second order elliptic equations, Math. Comp. 75 (2006), 1659-1674.

[18] F. A. Radu, N. Suciu, J. Hoffmann, A. Vogel, O. Kolditz, C-H. Park and S. Attinger, Accuracy of numerical simulations of contaminant transport in heterogeneous aquifers: a comparative study, Adv. Water Resour. 34 (2011), 47–61.

[19] P.A. Raviart and J.M. Thomas, Mixed Finite Element Methods for Second Order Elliptic Problems, Lecture Notes in Mathematics, Springer Verlag 1977, 292-315.

[20] V. Ruas, Automatic generation of triangular finite element meshes, Computer and Mathematics with Applications, 5-2 (1979), 125-140.

[21] V. Ruas, Hermite finite elements for second order boundary value problems with sharp gradient discontinuities, Journal of Computational and Applied Mathematics 246 (2013), 234-242.

[22] V. Ruas, D. Brandão and M. Kischinhevsky, Hermite finite elements for diffusion phenomena, Journal of Computational Physics 235 (2013), 542-564.

[23] V. Ruas and P. Trales, A Hermite finite element method for convection-diffusion equations, AIP Proceedings of the 11th International Conference Numerical Analysis and Applied Mathematics, T. Simos et al. ed., Rhodes, Greece, 2013.

[24] G.J. Strang and G. Fix, An Analysis of the Finite Element Method, Prentice Hall, Englewood Cliffs, 1973.

[25] J.-M. Thomas, Sur l'analyse numérique des méthodes d'éléments finis hybrides et mixtes, Thèse de Doctorat d'Etat, Université Pierre et Marie Curie, Paris 6, 1977.