A unified approach for handling convection terms in finite volumes and mimetic discretization methods for elliptic problems

BEIRÃO DA VEIGA

Dipartimento di Matematica 'F. Enriques', Università degli Studi di Milano, via Saldini 50, 20133 Milano, Italy
lourenco.beirao@unimi.it

JÉRÔME DRONIOU*

Institut de Mathématiques et de Modélisation de Montpellier, Université Montpellier 2, CC 051, Place Eugène Bataillon, 34095 Montpellier cedex 5, France
*Corresponding author. Email: droniou@math.univ-montp2.fr

GIANMARCO MANZINI

Istituto di Matematica Applicata e Tecnologie Informatiche—CNR, via Ferrata 1, 27100 Pavia, Italy
manzini@imati.cnr.it

We study the numerical approximation to the solution of the steady convection–diffusion equation. The diffusion term is discretized by using the hybrid mimetic method (HMM), which is the unified formulation for the hybrid finite-volume (FV) method, the mixed FV method and the mimetic finite-difference method recently proposed in Droniou *et al.* (2010, *Math. Models Methods Appl. Sci.*, **20**, 265–295). In such a setting we discuss several techniques to discretize the convection term that are mainly adapted from the literature on FV or FV schemes. For this family of schemes we provide a full proof of convergence under very general regularity conditions of the solution field and derive an error estimate when the scalar solution is in $H^2(\Omega)$. Finally, we compare the performance of these schemes on a set of test cases selected from the literature in order to document the accuracy of the numerical approximation in both diffusion- and convection-dominated regimes. Moreover, we numerically investigate the behaviour of these methods in the approximation of solutions with boundary layers or internal regions with strong gradients.

Keywords: finite-volume methods; mimetic finite-difference methods; convection-diffusion equation; convection-dominated flows; convergence analysis; error estimates.

1. Introduction

Many physical models of fluid flows involve partial differential equations (PDEs) with both convection and diffusion terms, such as the Navier–Stokes equations, flows in porous media, etc. Analytical solutions are not normally available for real applications and numerical approximations must be devised in some way. For this purpose efficient numerical schemes based on finite and mixed finite element and two-point finite volumes (FVs) have been developed for the numerical treatment of the diffusive part of the equation. In such a framework a great amount of work has been done to investigate the

connections between the lowest order Raviart–Thomas mixed finite-element ($RT_0 - P_0$) methods and various cell-centred FV and finite-difference numerical formulations on meshes of simplexes and quadrilaterals/hexahedrons. The relationship between the mixed finite-element method and the cell-centred finite-difference method on rectangular meshes was first established in Russell & Wheeler (1983) and further developed in subsequent papers (see, for example, Arbogast *et al.*, 1998). Basically, it can be shown that, by applying appropriate quadrature rules to the numerical formulation in the RT_0 space on rectangles, the vector variable (the velocity) is eliminated, thus reducing the method to a positive-definite cell-centred finite-difference method for the scalar variable (the pressure). Using this approach, classical cell-centred finite-difference methods on rectangular meshes are easily retrieved based on a nine-point stencil for full tensor coefficients and a five-point stencil for a scalar (diagonal) tensor. Similar results are also obtainable on regular hexahedral meshes. These developments led to the formulation of enhanced cell-centred finite differences (cf. Arbogast *et al.*, 1998) that can handle general shape elements (triangles, quadrilaterals and hexahedra) and are suitable for full tensor coefficients. A similar relationship exists between the $RT_0 - P_0$ scheme and the two-point FV formulation on triangular meshes using triangle circumcentres. This relationship was originally established by Baranger *et al.* (1996) for two-dimensional diffusion problems with scalar coefficients. This approach has been further developed in Younes *et al.* (2004), which investigates the case of a full diffusion tensor in two and three dimensions on meshes of simplexes.

Nonetheless, practical situations, such as those encountered in petroleum engineering, require computational grids that are not structured or simple enough to make use of the methods mentioned above. Thus, alternative and more sophisticated techniques have been developed in the last decade to approximate the solution to diffusive equations on general grids. In this framework we mention, for instance, the discontinuous Galerkin method (Arnold *et al.*, 2002; Riviere, 2008, and references therein), the multi-point flux approximation (Aavatsmark *et al.*, 1998a,b; Wheeler & Yotov, 2006), the mimetic finite-difference (MFD) method (Berndt *et al.*, 2001; Hyman *et al.*, 2002; Brezzi *et al.*, 2005a,b, 2007, 2009; Kuznetsov *et al.*, 2005; Beirão da Veiga, 2008; Beirão da Veiga & Manzini, 2008a,b; Cangiani & Manzini, 2008; Beirão da Veiga *et al.*, 2009b; Lipnikov *et al.*, 2009, and references therein), the hybrid FV method (Eymard *et al.*, 2009) and the mixed FV method (Droniou & Eymard, 2006; Chainais-Hillairet & Droniou, 2007; Droniou & Eymard, 2009). Strict correlations also exist among these numerical approximations and with respect to the lowest order mixed finite-element method, and it is not surprising that sometimes the lowest order schemes may belong to more than one of these families of methods. For example, the first-order discontinuous Galerkin scheme can be easily reinterpreted as an FV method. The lowest order Raviart–Thomas scheme on grids of simplexes (triangles in two dimensions and tetrahedrons in three dimensions) is a member of the family of MFD methods (cf. Cangiani & Manzini, 2008). Note, however, that on meshes of quadrilaterals and hexahedrons no connection has yet been established between the MFD method in mixed form and the mixed finite-element method. We also mention the paper by Vohralik (2006) that outlined the relationship existing between the multi-point flux approximation and the mixed finite-element method.

A remarkable fact has been recently discovered in Droniou *et al.* (2010), which showed that after some generalization, a unified formulation exists for three of the methods cited above, namely, the hybrid FV method, the mixed FV method and the MFD method. Consequently, these three methods are members of the same family of discretization techniques. Following Droniou *et al.* (2010), we will refer to such a family of numerical methods as *hybrid mimetic mixed* methods or use the abbreviation HMM.

Since the HMM method can be considered as the meeting of two different frameworks, namely, the mimetic/finite-element and the FV ones, the convective term can be naturally discretized using quite different techniques depending on the adopted point of view on the scheme. There are, indeed, two

possible approaches: either the diffusive flux is approximated, and then some form of centred or upwind approximation of the convection term is considered in the discretization of the divergence equation, or the total flux, which includes both diffusive and convective terms, is approximated, which leads to a centred-type approximation of the convection terms. The first approach is, perhaps, more popular in the finite-difference and FV practitioner community (cf. Chainais-Hillairet & Droniou, 2009; Droniou & Eymard, 2009), while the second approach seems to be more popular in the finite-element practitioner community. Nevertheless, it is worth mentioning that both approaches have been considered in the framework of mixed finite-element methods (see Douglas & Roberts, 1982, 1985; Jaffre & Roberts, 1985).

In the MFD setting a numerical discretization of the full diffusion and convection fluxes has been proposed by Cangiani *et al.* (2009). A proper reformulation of the mimetic scheme as a conforming method, using the finite-dimensional subspace of $H(\mathrm{div}, \Omega)$ given by the lifting of the degrees of freedom of the vector variable, makes it possible to perform the convergence analysis in a very similar way to that presented in Douglas & Roberts (1985).

From this overview we can conclude that several numerical discretizations of the convection–diffusion equations that may fit in the HMM setting have been proposed in the literature. However, no systematic study has been carried out so far on the possible ways, and related advantages and drawbacks, in which a convective term can be treated numerically by using the more general HMM formulation. It is our main goal in this work to perform such an investigation in order to assess the behaviour of such methods both theoretically and numerically.

The plan of the paper is as follows. In Section 2 we recall the principles of the HMM schemes for the pure diffusion equation, and we discuss how to discretize the convection term, using some centred, upwind or exponential fitting-like choice in accordance with a two-point FV flux formula (or, from the point of view of finite elements, see Jaffre (1984) and Dawson & Aizinger (1999)). We also show that the numerical approximation proposed in Cangiani *et al.* (2009), possibly with a stabilization term, is an HMM method to which the theoretical analysis of the present paper applies. In Section 3 we provide full proofs of convergence under very general regularity conditions when the mesh size tends to zero and derive error estimates in suitable mesh-dependent norms when the scalar solution is in $H^2(\Omega)$. Section 4 is devoted to the presentation and discussion of how various instances of the HMM discretizations perform when applied to a set of standard test cases for the convection–diffusion equations, including the approximation of solutions with boundary and internal layers. Finally, conclusions are given in Section 5.

2. The HMM formulation for convection–diffusion problems

2.1 *The mathematical model*

Let us consider the steady convection–diffusion equation

$$-\mathrm{div}(\Lambda \nabla p) + \mathrm{div}(Vp) = f \quad \text{in } \Omega, \tag{2.1}$$

$$p = g^{\mathrm{D}} \quad \text{on } \partial\Omega \tag{2.2}$$

under the following hypotheses:

(H1) Ω is a bounded, open, polygonal subset of $\mathbf{R}^d$ with $d \geqslant 1$;

(H2) $\Lambda : \Omega \to M_d(\mathbf{R})$ is a bounded, measurable, symmetric and uniformly elliptic tensor;

(H3) $f \in L^2(\Omega)$;

(H4) $V \in C^1(\overline{\Omega})^d$ is such that $\operatorname{div}(V) \geqslant 0$.

Moreover, let us introduce the diffusive flux and the total flux as follows:

$$F = -\Lambda \nabla p \quad \text{and} \quad \widetilde{F} = F + Vp. \tag{2.3}$$

For simplicity, we will restrict the presentation of a the methods and the theoretical analysis to the case of a homogeneous Dirichlet boundary condition by setting $g^{\mathrm{D}} = 0$ in (2.2) and we will consider the nonhomogeneous case in the numerical experiments of Section 4.

Under assumptions (H1)–(H4), the existence and uniqueness of a weak solution in $H_0^1(\Omega)$ to (2.1) and (2.2) with $g^{\mathrm{D}} = 0$ is completely standard since the bilinear form associated with this problem is continuous and coercive.

REMARK 2.1 The C^1 regularity assumption on V in (H4) can be weakened for the convergence study (see Section 3.1.3). We assume the smoothness of the convection field in order to simplify a little bit some (already lengthy) technical arguments and also to prove error estimates.

2.2 *Mesh notation and regularity*

Let us begin with the definition of an admissible discretization of Ω and the related notation.

DEFINITION 2.2 (Admissible discretization). An admissible discretization of Ω is given by the triplet $\mathcal{D}_h = (\Omega_h, \mathcal{E}_h, \mathcal{P}_h)$, where the *mesh size h* will be defined in the following and where the following conditions hold.

- Here Ω_h is a finite family of nonempty open polygonal disjoint subsets E of Ω, the *cells* of the mesh, such that $\overline{\Omega} = \bigcup_{E \in \Omega_h} \overline{E}$.

- Here $\mathcal{E}_h$ is a finite family of nonempty open disjoint subsets e of $\overline{\Omega}$, the *faces* of the mesh, such that, for all $e \in \mathcal{E}_h$, there exists an affine hyperplane $\mathcal{A}$ of $\mathbf{R}^d$ and a cell $E \in \Omega_h$ such that $e \subset (\overline{E} \backslash E) \cap \mathcal{A}$. We also assume that the following hold:

  - for all $E \in \Omega_h$ there exists a subset ∂E of $\mathcal{E}_h$ such that $\overline{E} \backslash E = \bigcup_{e \in \partial E} \overline{e}$;

  - for all $e \in \mathcal{E}_h$ either we have that $e \subset \partial \Omega$ or we have that $e \subset \overline{E} \cap \overline{E'}$ for some pair of elements $E, E' \in \Omega_h$ with $E \neq E'$.

- Here $\mathcal{P}_h$ is a family of points of Ω indexed by E, that is, $\mathcal{P}_h = (x_E)_{E \in \Omega_h}$, and such that each mesh cell E is star shaped with respect to x_E.

REMARK 2.3 When all of the mesh cells are convex shaped a convenient choice for the points $(x_E)_{E \in \Omega_h}$ is given, for instance, by the centre of gravity of the cells.

The d-dimensional measure of each cell E is denoted by $|E|$ and the cell size by h_E. As usual, the mesh size is given by $h = \sup_{E \in \Omega_h} h_E$. For consistency of notation, $|e|$ and h_e denote the $(d-1)$-dimensional measure of face e and the face diameter. For each face $e \in \mathcal{E}_h$ we let $\bar{x}_e$ denote the barycentre of e and $\mathbf{n}_E^e$ its normal direction pointing out of E. Moreover, with each face e we associate the unit normal vector $\mathbf{n}^e$, whose orientation is arbitrarily chosen when e is an internal face, and assumed to be pointing out of Ω when e is a boundary face. We denote the set of the internal faces by $\mathcal{E}_{h,\mathrm{int}}$, that is, $\mathcal{E}_{h,\mathrm{int}} = \{e \in \mathcal{E}_h \text{ for } e \not\subset \partial\Omega\}$, and the set of the boundary faces by $\mathcal{E}_{h,\mathrm{ext}}$, that is, $\mathcal{E}_{h,\mathrm{ext}} = \{e \in$

$\mathcal{E}_h$ for $e \subset \partial\Omega$}. We will find it convenient to denote the two cells that share an internal face e by E and E' and, where required, to fix the orientation of e so that $\mathbf{n}_E^e \cdot \mathbf{n}^e = 1$. Finally, we introduce the following geometric quantities that will be useful in the definition of the numerical convection flux in Section 2.4.1:

$$d_{E,e} \text{ is the } = \text{ distance between } x_E \text{ and the hyperplane containing } e$$

and

$$d_e = \begin{cases} d_{E,e} + d_{E',e} & \text{for any internal face } e \in \mathcal{E}_{h,\text{int}}, \\ d_{E,e} & \text{for any boundary face } e \in \mathcal{E}_{h,\text{ext}}. \end{cases}$$

Figure 1 illustrates some of this notation.

The proof of convergence for $h \to 0$ that we present in Section 3 requires the following very mild geometrical assumptions on the meshes of $\mathcal{D}_h$.

- (G1) Every mesh cell E is star shaped with respect to the corresponding point x_E.
- (G2) For any internal face $e \in \mathcal{E}_{h,\text{int}}$ let us introduce $\mathcal{M}_e = \{E, E'\}$, that is, the cells on the opposite side of e. Then the quantity

$$\text{regul}(\mathcal{D}_h) = \max \left(\max_{e \in \mathcal{E}_{h,\text{int}}, (E,E') \in \mathcal{M}_e} \frac{d_{E,e}}{d_{E',e}}, \max_{E \in \Omega_h, e \in \partial E} \frac{h_E}{d_{E,e}}, \max_{E \in \Omega_h} \text{Card}(\partial E) \right),$$

which expresses the mesh regularity, is *uniformly* bounded from above for $h \to 0$.

In the mimetic framework a similar condition is often used, and we state this as follows.

- (ME) [*Star-shaped elements*] There exists a positive number τ^* such that each element E is star shaped with respect to *all* the points of a ball of radius $\tau^* h_E$ centred at x_E.

Stronger conditions on the mesh regularity are required to derive an error estimate for the HMM approximations to the exact solution and flux. We formulate these mesh regularity conditions for $d = 3$. The restriction to other dimensions is straightforward.
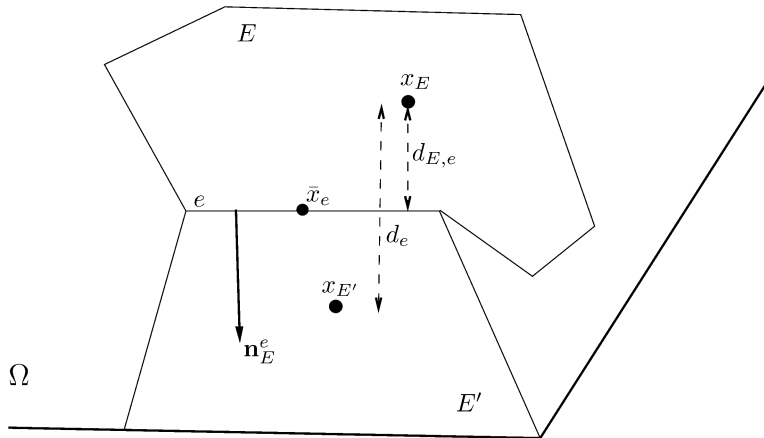

Fig. 1. Mesh notation.

(HG) [*Shape regularity*] There exist two positive real numbers N_s and ρ_s, independent of h, such that every mesh Ω_h of the sequence admits a subpartition into tetrahedrons $\mathcal{S}_h$ such that we have the following.

   (HG1) The decomposition of every polyhedron $E \in \Omega_h$, denoted by $\mathcal{S}_{h|E}$, is formed by at most N_s tetrahedrons, and each vertex of Ω_h is a vertex of $\mathcal{S}_h$.

   (HG2) Every tetrahedron of $\mathcal{S}_h$ is *shape regular* in the sense that the ratio between r_T, the radius of its inscribed sphere, and h_T, its diameter, is bounded from below by ρ_s. Formally, we have that

$$\forall\, T \in \mathcal{S}_h : \frac{r_T}{h_T} \geqslant \rho_s > 0.$$

From the above assumptions, several properties of the mesh that are useful in the error analysis of the mimetic formulation can be derived. For the sake of the reader's convenience, we list them below for future reference in the paper.

(M1) There exist two positive integers N_E and N_e that are independent of h, $E \in \Omega_h$ and $e \in \mathcal{E}_h$ and such that every element E has $\mathrm{Card}(\partial E) \leqslant N_E$ faces, and every face e has $\mathrm{Card}(\partial e) \leqslant N_e$ edges.

(M2) For any mesh element $E \in \Omega_h$ the quantities $|E|$, $|e|$ for $e \in \partial E$ and $|l|$ for each edge $l \in \partial e$ properly scale with respect to h_E. In particular, there exists a positive constant a^* such that

$$a^* h_E^{d-1} \leqslant |e|, \quad a^* h_E \leqslant h_e, \quad a^* h_E^{d-2} \leqslant |l|.$$

(M3) There exists a constant C^{Ag} that is independent of h_E and such that (Brezzi *et al.*, 2005a)

$$\sum_{e \in \partial E} \|\phi\|_{L^2(e)}^2 \leqslant C^{\mathrm{Ag}}(h_E^{-1}\|\phi\|_{L^2(E)}^2 + h_E|\phi|_{H^1(E)}^2) \tag{2.4}$$

for any function $\phi \in H^1(E)$. We will refer to (2.4) as the *Agmon inequality*.

(M4) For any function $q \in H^2(E)$ there exists a *linear polynomial* $\mathcal{L}_1(q)$ interpolating q and a constant C, independent of h_E, such that (Brenner & Scott, 1994)

$$\|q - \mathcal{L}_1(q)\|_{L^2(E)} + h_E|q - \mathcal{L}_1(q)|_{H^1(E)} \leqslant Ch_E^2|q|_{H^2(E)}. \tag{2.5}$$

2.3 *Discretization of the diffusion term*

To approximate (2.1) and (2.2) we introduce the space of the *discrete scalar fields* Q_h and the space of the *discrete flux fields* X_h. The discrete scalars $q \in Q_h$ are defined by taking one degree of freedom per cell, denoted by q_E, that is, $q = (q_E)_{E \in \Omega_h}$. Therefore the space Q_h can be identified with the space of the piecewise constant polynomials defined on Ω_h. Similarly, the *discrete fluxes* are defined by taking one degree of freedom per face per element, denoted by F_E^e, that is, $F = (F_E^e)_{E \in \Omega_h}^{e \in \partial E}$, which represents the normal flux across the face e in the direction $\mathbf{n}_E^e$. We require that every flux $F \in X_h$ satisfies the following *flux conservation property* at any internal face:

$$\forall\, e \in \mathcal{E}_{h,\mathrm{int}}, e \subset \partial E \cup \partial E' : F_E^e + F_{E'}^e = 0, \tag{2.6}$$

so that the elements of X_h only possess one degree of freedom per face, and the sign of F_E^e depends on the orientation of the face e with respect to E. The restriction of the F to the cell $E \in \Omega_h$ is denoted by $F_E = (F_E^e)_{e \in \partial E}$ and represents the collection of the normal fluxes in the directions $\mathbf{n}_E^e$ for $e \in \partial E$. The

set of these vector fields forms the linear space X_E. Throughout the paper we will also make use of the symbol $\widehat{X}_h$ to denote the linear space of the *discontinuous fluxes*, that is, of the vectors having the same form $F = (F_E^e)_{E \in \Omega_h}^{e \in \partial E}$ but that do not satisfy condition (2.6). Note that X_h is a linear subspace of $\widehat{X}_h$.

The next ingredient of the HMM formulation is the *discrete divergence operator* $\text{div}_h : \widehat{X}_h \to Q_h$, which is defined as follows:

$$\forall\, G \in \widehat{X}_h, \quad \forall\, E \in \Omega_h : \quad (\text{div}_h(G))_E = \frac{1}{|E|} \sum_{e \in \partial E} |e| G_E^e. \tag{2.7}$$

With any sufficiently regular vector field G and scalar field q we associate the interpolated fields $G^I \in X_h$ and $q^I \in Q_h$ that are given by

$$\forall\, e \in \mathcal{E}_h : (G^I)^e = \frac{1}{|e|} \int_e G \cdot \mathbf{n}^e \quad \text{and} \quad \forall\, E \in \Omega_h : (q^I)_E = \frac{1}{|E|} \int_E q. \tag{2.8}$$

REMARK 2.4 The definition of the discrete divergence operator in (2.7) is consistent with the Gauss divergence theorem for the interpolations of (2.8), so that the following commutation property holds:

$$(\text{div}(G))^I = \text{div}_h(G^I). \tag{2.9}$$

We endow Q_h with the usual $L^2(\Omega)$ scalar product for piecewise constant functions, that is, $[\cdot, \cdot]_{Q_h} := [\cdot, \cdot]_{L^2}$. On the other hand, X_h and $\widehat{X}_h$ are equipped with the scalar product

$$[F, G]_{\widehat{X}_h} = \sum_{E \in \Omega_h} [F_E, G_E]_E \tag{2.10}$$

that assembles the locally defined scalar products $[\cdot, \cdot]_E$. The local scalar products $([\cdot, \cdot]_E)$ satisfy the following coercivity and consistency assumptions.

(S1) There exist two positive constants σ_* and σ^*, independent of the mesh size h, such that, for every mesh cell E, we have

$$\sigma_* |E| \sum_{e \in \partial E} (G_E^e)^2 \leqslant [G, G]_E \leqslant \sigma^* |E| \sum_{e \in \partial E} (G_E^e)^2 \quad \forall\, G \in X_h \ .$$

(S2) For every element E we have that

$$[(\Lambda_E \nabla q^1)^I, G]_E = -[\text{div}_h(G), q^1]_{L^2(E)} + \sum_{e \in \partial E} G_E^e \int_e q^1$$

for all $G \in X_h$ and all linear polynomials q^1, where Λ_E is the cell average of Λ.

REMARK 2.5 Here Λ_E is actually an approximation of $\Lambda|_E$, the restriction of the diffusion tensor Λ to the cell E. To prove the convergence of the numerical solution in Section 3.1, we only require that the diffusion tensor Λ satisfy the regularity assumption (H2), while we need a stronger regularity condition to derive the error estimates of Section 3.2. In this latter case we will find it convenient to assume (H2) and also that Λ be locally Lipschitz continuous on Ω_h, that is, for all $E \in \Omega_h$, the components of $\Lambda|_E$ are Lipschitz continuous functions on E. Consequently, Λ_E can be any constant approximation of $\Lambda|_E$ such that the estimate

$$\max_{i,j=1,d} \sup_{x \in E} |(\Lambda_E)_{ij} - \Lambda_{ij}(x)| = \mathcal{O}(h)$$

holds.

The construction of a family of scalar products satisfying the above assumptions when x_E is the centre of gravity of E can be found in Brezzi *et al.* (2005b). Moreover, in this case it was proved in Droniou *et al.* (2010) that (S1) and (S2) lead necessarily to the following form:

$$\forall (F_E, G_E) \in X_E : [F_E, G_E]_E = |E| \Lambda_E \mathbf{v}_E(F_E) \cdot \mathbf{v}_E(G_E) + T_E(G_E)^{\mathrm{T}} \mathbb{B}_E T_E(F_E), \qquad (2.11)$$

where

$$\mathbf{v}_E(F_E) = -\frac{1}{|E|} \Lambda_E^{-1} \sum_{e \in \partial E} |e| F_E^e (\bar{x}_e - x_E) \qquad (2.12)$$

is a constant approximation of ∇p on the cell E, $T_E(F_E) = (T_{E,e}(F^e))_{e \in \partial E}$ is given by

$$T_{E,e}(F_E) = F_E^e + \Lambda_E \mathbf{v}_E(F_E) \cdot \mathbf{n}_E^e, \qquad (2.13)$$

and $\mathbb{B}_E$ is a symmetric positive-definite matrix of size $\mathrm{Card}(\partial E)$. More precisely, it turns out that the matrix $\mathbb{B}_E$ satisfies the following coercivity condition that is directly related to (S1).

(C) There exists a positive constant α, that is independent of the mesh size such that, for all $E \in \Omega_h$ and $G_E \in X_E$, we have that

$$\alpha \sum_{e \in \partial E} |e| d_{E,e}(T_{E,e}(G_E))^2 \leqslant T_E(G_E)^{\mathrm{T}} \mathbb{B}_E T_E(G_E) \leqslant \frac{1}{\alpha} \sum_{e \in \partial E} |e| d_{E,e}(T_{E,e}(G_E))^2.$$

If x_E is not the barycentre of E, then the same construction (2.11)–(2.13) still holds provided that (S2) is modified by introducing a suitable integration weight (see Droniou *et al.*, 2010).

The HMM discretization of problem (2.1) and (2.2) with $V = 0$, which provides us with the desired approximation of the diffusion operator, takes the following form:

find $(p_h, F_h) \in Q_h \times X_h$ such that

$$\forall G \in X_h : \quad [F_h, G]_{\widehat{X}_h} = [\mathrm{div}_h(G), p_h]_{Q_h}, \qquad (2.14)$$

$$\forall q \in Q_h : \quad [\mathrm{div}_h(F_h), q]_{Q_h} = [f, q]_{Q_h}. \qquad (2.15)$$

Here $p_h \in Q_h$ and $F_h \in X_h$ are the approximations to p^I and F^I, the interpolations of the exact scalar solution p and its flux $F = -\Lambda \nabla p$.

The HMM method can be easily hybridized through the introduction of $H(\mathcal{E}_h)$, the space of face values $q_{\mathcal{E}_h} = (q_e)_{e \in \mathcal{E}_h} \in \mathbf{R}^{\mathrm{Card}(\mathcal{E}_h)}$ with $q_e = 0$ for $e \in \mathcal{E}_{h,\mathrm{ext}}$, and imposing explicitly the flux conservation property (2.6). The discrete variational form (2.14) and (2.15) with $[\cdot, \cdot]_E$ satisfying (2.11)–(2.13) is equivalent to the following:

find $(p_h, F_h, p_{\mathcal{E}_h}) \in Q_h \times \widehat{X}_h \times H(\mathcal{E}_h)$ such that

$$\forall E \in \Omega_h, \quad \forall G_E \in X_E : \qquad [F_E, G_E]_E = \sum_{e \in \partial E} |e| G_E^e (p_E - p_e), \qquad (2.16)$$

$$\forall E \in \Omega_h : \qquad \sum_{e \in \partial E} |e| F_E^e = \int_E f, \qquad (2.17)$$

$$\forall e \in \mathcal{E}_{h,\mathrm{int}}, : \qquad F_E^e + F_{E'}^e = 0, \qquad (2.18)$$

where E, $E' \in \Omega_h$ are the two elements such that $e \subset \partial E \cap \partial E'$ for every $e \in \mathcal{E}_{h,\text{int}}$. Under suitable assumptions on the regularity of the exact solution p, the additional unknowns $p_{\mathcal{E}_h} = (p_e)_{e \in \mathcal{E}_h}$ approximate the face average of the exact solution over each mesh face. We will formalize this concept through the introduction of $p^J \in H(\mathcal{E}_h)$, the face interpolation of p, in Section 3.2 (see equation (3.43)).

2.4 *Discretization of the convective term*

As discussed in Section 1, two different strategies can be considered for the numerical treatment of the convection term in the HMM discretization of an elliptic problem. In the first strategy, which is reviewed in Section 2.4.1, we introduce some form of centred or upwind approximation of the convection term in the discretization of the divergence equation provided by the HMM method (cf. Chainais-Hillairet & Droniou, 2009; Droniou & Eymard, 2009). In the second strategy, which is reviewed in Section 2.4.2, the total flux, which includes both diffusive and convective terms, is approximated, thus leading to a centred-type approximation of the convection terms (cf. Cangiani *et al.*, 2009). Both approaches were considered for the mixed finite-element method in Douglas & Roberts (1982, 1985) and Jaffre & Roberts (1985). It turns out that, in the new framework of HMM methods, a unified formulation is possible, which is the topic of Section 2.4.3. We end this section with a discussion on an alternative hybridized form of the numerical convection terms (Section 2.4.4).

In the rest of this section we assume that the velocity field V is a continuous function with a continuous derivative, that is, $V \in C^1(\overline{\Omega})^d$. The cell restriction of its interpolation in X_h is given by the set of real numbers $(V_E^e)_{e \in \partial E} \in X_E$ such that

$$\forall e \in \partial E : \quad V_E^e = \frac{1}{|e|} \int_e V \cdot \mathbf{n}_E^e. \tag{2.19}$$

2.4.1 *FV-based discretizations.*

Several discretization schemes for the convection term are available in the FV literature, for example, the second-order centred scheme, the first-order upwind scheme, the θ-scheme, the Scharfetter–Gummel scheme, etc. In these methods, the convection flux of the exact solution field p is approximated through the numerical convection flux of the discrete scalar field $p_h \in Q_h$. This numerical convection flux is given by the collection of real numbers $F_c(p_h) = (F_c(p_h)_E^e)_{E \in \Omega_h, e \in \partial E}$ such that

$$\forall E \in \Omega_h, \quad \forall e \in \partial E : \quad \frac{1}{|e|} \int_e V p \cdot \mathbf{n}_E^e \approx (F_c(p_h))_E^e. \tag{2.20}$$

We list below the schemes that we will explicitly consider in the section on numerical experiments. We let E' be the cell on the other side of e if $e \in \mathcal{E}_{h,\text{int}}$ and assume for notational simplicity that $p_{E'} = 0$ if $e \in \mathcal{E}_{h,\text{ext}}$.

- The *second-order centered scheme* is given by the approximation

$$\frac{1}{|e|} \int_e V p \cdot \mathbf{n}_E^e \approx (F_c(p_h))_E^e = V_E^e \frac{p_E + p_{E'}}{2}.$$

- The *first-order upwind scheme* is given by the approximation

$$\frac{1}{|e|} \int_e V p \cdot \mathbf{n}_E^e \approx (F_c(p_h))_E^e = (V_E^e)^+ p_E - (V_E^e)^- p_{E'},$$

with $s^\pm = \max(\pm s, 0)$.

- The *θ-scheme* is given by the approximation

$$\frac{1}{|e|} \int_e V p \cdot \mathbf{n}_E^e \approx (F_c(p_h))_E^e = (V_E^e)^+ ((1-\theta) p_E + \theta p_{E'}) - (V_E^e)^- ((1-\theta) p_{E'} + \theta p_E)$$

$$= (1-2\theta)((V_E^e)^+ p_E - (V_E^e)^- p_{E'}) + \theta V_E^e (p_E + p_{E'}),$$

with $\theta \in [0, 1/2]$. This choice is clearly intermediate between the centred and the upwind schemes.

- The *Scharfetter–Gummel scheme* (Scharfetter & Gummel, 1969) is given by the approximation

$$\frac{1}{|e|} \int_e V p \cdot \mathbf{n}_E^e \approx (F_c(p_h))_E^e = \frac{1}{d_e} (A_{\mathrm{sg}}(d_e V_E^e) p_E - A_{\mathrm{sg}}(-d_e V_E^e) p_{E'}), \tag{2.21}$$

with

$$A_{\mathrm{sg}}(s) = \frac{-s}{\mathrm{e}^{-s} - 1} - 1. \tag{2.22}$$

Note that the first three approaches above can also be found in the finite-element literature (see, for instance, Jaffre (1984) and Dawson & Aizinger (1999)). As pointed out in Chainais-Hillairet & Droniou (2009), the Scharfetter–Gummel scheme in Scharfetter & Gummel (1969) was written for an isotropic homogeneous material, that is, $\Lambda = I$. In the original formulation, diffusion and convection terms were simultaneously treated to define the numerical flux. Removing the diffusive part in the numerical flux formulation allows us to obtain the formulas (2.21) and (2.22). This definition of a pure convective flux through the simple elimination of the diffusive part is somewhat basic in the general case $\Lambda \neq I$, especially if some eigenvalues of Λ are small. Although the above definition of A_{sg} ensures the L^2-stability of the scheme, it can give quite bad solutions in convection-dominated cases. This fact can be understood if one comes back to the two-point FV scheme for $-\epsilon \Delta p + \mathrm{div}(V p) = f$. Then the choice (2.22) ensures the maximum principle of the scheme only if $\epsilon \geqslant 1$, while the maximum principle is lost numerically if $\epsilon < 1$. When applying the Scharfetter–Gummel method to compute the numerical convective flux, a better choice is provided by locally scaling A_{sg} in accordance with the smallest eigenvalue of Λ. If e is the face between E and E' and λ_e is the smallest eigenvalue of Λ_E and $\Lambda_{E'}$, then we use

$$A_{\mathrm{sg}, \Lambda, e}(s) = \min(1, \lambda_e) A_{\mathrm{sg}} \left(\frac{s}{\min(1, \lambda_e)} \right) \tag{2.23}$$

instead of $A_{\mathrm{sg}}(s)$ in (2.21). In this way the numerical flux automatically and locally adjusts the upwinding of the convection term depending on its strength with respect to the diffusive term without perturbing the consistency property of A_{sg}. Note that $\lambda_e \to 0$ implies that $\lambda_e A_{\mathrm{sg}}(s/\lambda_e) \to s^+$. Therefore, if the local diffusion is very small, then this implementation of the Scharfetter–Gummel method allows the flux to adjust to upwinding automatically, thus bringing enough numerical diffusion to ensure a better stability.

Once an FV-based discretization of the convective term has been chosen, the divergence of the convection term in (2.1), that is, $\mathrm{div}(V p)$, is approximated on E by

$$(\mathrm{div}(V p))_E^I \approx \frac{1}{|E|} \sum_{e \in \partial E} |e| (F_c(p_h))_E^e = \mathrm{div}_h(F_c(p_h))|_E$$

and the HMM approximation to the model problem (2.1) and (2.2) then reads as follows:
find $(p_h, F_h) \in Q_h \times X_h$ such that

$$\forall G \in X_h : \quad [F_h, G]_{\widehat{X}_h} = [\mathrm{div}_h(G), p_h]_{Q_h}, \tag{2.24}$$

$$\forall q \in Q_h : \quad [\mathrm{div}_h(F_h + F_c(p_h)), q]_{Q_h} = [f^I, q]_{Q_h}. \tag{2.25}$$

2.4.2 *MFD-based discretizations.* From a theoretical standpoint, MFDs have only very recently approached problems that are different to the pure diffusion one (see, for instance, Beirão da Veiga *et al.*, 2009a, 2010; Beirão da Veiga, 2010). To our knowledge, the only paper considering the development and error analysis of convection–diffusion equations directly in the framework of MFD is that of Cangiani *et al.* (2009). In this subsection we briefly review the formulation and the major convergence results of the method considered in that paper, and we show how it can be reformulated as an HMM method.

Let $H(\mathrm{div}, \Omega)$ be the space of vector fields all of whose components are square integrable functions and that have square integrable divergence. Formally,

$$H(\mathrm{div}, \Omega) = \{\mathbf{v} \in (L^2(\Omega))^d \text{ such that } \mathrm{div}(\mathbf{v}) \in L^2(\Omega)\}$$

is a Hilbert space when equipped with the scalar product

$$[\mathbf{v}, \mathbf{u}]_{H(\mathrm{div}, \Omega)} = \int_\Omega \mathbf{v} \cdot \mathbf{u} + \int_\Omega \mathrm{div}(\mathbf{v}) \, \mathrm{div}(\mathbf{u})$$

and the corresponding norm

$$\|\mathbf{v}\|^2_{H(\mathrm{div}, \Omega)} = \|\mathbf{v}\|^2_{L^2(\Omega)} + \|\mathrm{div}(\mathbf{v})\|^2_{L^2(\Omega)}.$$

In Cangiani *et al.* (2009) a numerical approximation was considered to the *mixed variational formulation* of problem (2.1) and (2.2), which reads as follows (Brezzi & Fortin, 1991):
find $(\widetilde{F}, p) \in (\mathrm{div}, \Omega) \times L^2(\Omega)$ such that

$$\forall \mathbf{v} \in H(\mathrm{div}, \Omega) : \quad [\Lambda^{-1}\widetilde{F}, \mathbf{v}]_{L^2} - [p, \mathrm{div}(\mathbf{v})]_{L^2} - [\Lambda^{-1}V p, \mathbf{v}]_{L^2} = 0, \tag{2.26}$$

$$\forall q \in L^2(\Omega) : \quad [\mathrm{div}(\widetilde{F}), q]_{L^2} = [f, q]_{L^2}, \tag{2.27}$$

where $\widetilde{F}$ is the total vector flux defined in (2.3).

To discretize the convection term we transform the corresponding variational term as follows:

$$\forall \mathbf{v} \in H(\mathrm{div}, \Omega) : [\Lambda^{-1}V p, \mathbf{v}]_{L^2} \approx \sum_{E \in \Omega_h} \int_E \Lambda_E^{-1} V p \cdot \mathbf{v} \quad \longrightarrow \quad \forall G \in X_h : \sum_{E \in \Omega_h} p_E [V^I, G]_E,$$

where the components of the interpolated velocity field $V^I \in X_h$, that is, $(V^I)^e_E$ for all $E \in \Omega_h$ and $e \in \partial E$, are given by (2.19), and the local scalar products are required to satisfy assumptions (S1) and (S2). The mimetic variational formulation presented in Cangiani *et al.* (2009) reads as follows:
find $(\widetilde{F}_h, p_h) \in X_h \times Q_h$ such that

$$\forall G \in X_h : \quad [\widetilde{F}_h, G]_{\widehat{X}_h} - [p_h, \mathrm{div}_h(G)]_{Q_h} - \sum_{E \in \Omega_h} p_E [V^I, G]_E = 0, \tag{2.28}$$

$$\forall q \in Q_h : \quad [\mathrm{div}_h(\widetilde{F}_h), q]_{Q_h} = [f^I, q]_{Q_h}. \tag{2.29}$$

The convergence analysis of this scheme was carried out in Cangiani *et al.* (2009) under assumptions on the grid regularity that are substantially equivalent to (HG) and (ME). When the scalar solution p is in $H^2(\Omega)$ the analysis provides the following error estimate:

$$\|\widetilde{F}_h - \widetilde{F}^I\|_{\widehat{X}_h} + \|p_h - p^I\|_{Q_h} \leqslant Ch\|p\|_{H^2(\Omega)}, \tag{2.30}$$

where $\|\cdot\|_{\widehat{X}_h}$ and $\|\cdot\|_{Q_h}$ are the norms induced by the inner products of the spaces $\widehat{X}_h$ and Q_h, respectively. It is worth mentioning that the approximation of the scalar variable is superconvergent when the calculation is performed on a wide set of meshes. Superconvergence was also theoretically proved under some stronger assumptions on the regularity of the domain shape, the source term and the velocity field.

Despite convergence being proved for $h \to 0$ this scheme is expected to become unstable when the model problem is dominated by convection. This fact usually manifests through spurious effects like numerical undershoots, overshoots or oscillations that may appear in the approximate solution. To improve stability we modify the divergence equation by introducing a stabilization term that depends on the solution's jumps at mesh faces. We use the symbols E and E' to denote the two distinct cells that share the face e when e is internal and assume the orientation of e to be such that $\mathbf{n}_E^e \cdot \mathbf{n}^e = 1$. Let us now introduce the *jump* of the discrete scalar field $q_h \in Q_h$ as follows:

$$[\![q_h]\!]_e = \begin{cases} q_E - q_{E'} & \text{for } e \in \mathcal{E}_{h,\text{int}}, \\ q_E & \text{for } e \in \mathcal{E}_{h,\text{ext}}. \end{cases} \tag{2.31}$$

Equation (2.29) is substituted by

$$\forall q \in Q_h : [\text{div}_h(\widetilde{F}_h) + J_h(p_h), q]_{Q_h} = [f^I, q]_{Q_h}, \tag{2.32}$$

where the stabilization term $J_h(p_h)$ is given by

$$J_h(p_h)|_E = \frac{\alpha}{2|E|} \sum_{e \in \partial E} |e| |(V^I)_E^e| [\![p_h]\!]_e, \tag{2.33}$$

and α is a non-negative parameter that can be tuned to control the amount of numerical dissipation of the scheme.

This approach formally differs from the method introduced in Section 2.4.1 for FV-based discretizations in that the convection term is numerically treated as part of the mimetic flux equation. However, it is possible to 'extract' an explicit form of the numerical convection flux from the scheme given by equations (2.28) and (2.32) to reformulate it as an HMM method. For this purpose, we define the collection of numbers $F_h = (F_E^e)_{E \in \Omega_h, e \in \partial E}$ by

$$F_E^e = \widetilde{F}_E^e - p_E(V^I)_E^e. \tag{2.34}$$

Equation (2.28) shows that F_h satisfies (2.14) and therefore plays the role of a purely diffusive flux. Moreover, noting that the stabilization term $J_h(p_h)$ is locally written as a balance of fluxes, that is, a discrete divergence, allows us to identify the convective flux as

$$(F_c(p_h))_E^e = p_E(V^I)_E^e + \frac{\alpha}{2}|(V^I)_E^e|(p_E - p_{E'}) \tag{2.35}$$

(we let $p_{E'} = 0$ if e is a boundary edge), so that (2.32) is simply given by $\mathrm{div}_h(F_h + F_c(p_h)) = f^I$. The stabilized MFD scheme (2.28) and (2.32) can therefore be written as follows:

find $p_h \in Q_h$ and $F_h \in \widehat{X}_h$ such that

$$\forall\, G \in X_h : \quad [F_h, G]_{\widehat{X}_h} = [\mathrm{div}_h(G), p_h]_{Q_h}, \tag{2.36}$$

$$\forall\, q \in Q_h : \quad [\mathrm{div}_h(F_h + F_c(p_h)), q]_{Q_h} = [f^I, q]_{Q_h}, \tag{2.37}$$

$$\forall\, e \in \mathcal{E}_{h,\mathrm{int}} : \quad (F_h + (F_c(p_h))^e_E + (F_h + (F_c(p_h))^e_{E'} = 0. \tag{2.38}$$

Note that the diffusive flux F_h and the convective flux $F_c(p_h)$ are not conservative in the sense of (2.6) when considered separately and therefore belong to the linear space $\widehat{X}_h$. However, their sum, that is, $F_h + F_c(p_h)$, is conservative since it belongs to X_h in view of equation (2.38).

2.4.3 *Unified setting.* A unified formulation exists for the numerical discretization of the convection term. This formulation includes the FV-based discretizations, as was noted in Chainais-Hillairet & Droniou (2009), and the MFD-based discretization (2.36)–(2.38). This fact makes it possible to simplify the software implementation and carry out a unified theoretical analysis.

Let us consider two functions $A, B : \mathbf{R} \to \mathbf{R}$ and choose the numerical convection flux as the collection of real numbers

$$F_c(p_h) = (F_c(p_h)^e_E)_{E \in \Omega_h, e \in \partial E} \tag{2.39}$$

such that

$$\forall\, E \in \Omega_h, \quad \forall\, e \in \partial E : (F_c(p_h))^e_E := \frac{1}{d_e}(A(d_e V^e_E) p_E + B(d_e V^e_E) p_{E'}). \tag{2.40}$$

Since in the MFD discretization of the convection term these flux components are not conservative, the diffusive flux components cannot be conservative either and conservation must be imposed on the total flux. The generic HMM approximation to the model problem (2.1) and (2.2) is thus written as follows:

find $p_h \in Q_h$ and $F_h \in \widehat{X}_h$ such that

$$\forall\, G \in X_h : \quad [F_h, G]_{\widehat{X}_h} = [\mathrm{div}_h(G), p_h]_{Q_h}, \tag{2.41}$$

$$\forall\, q \in Q_h : \quad [\mathrm{div}_h(F_h + F_c(p_h)), q]_{Q_h} = [f^I, q]_{Q_h}, \tag{2.42}$$

$$\forall\, e \in \mathcal{E}_{h,\mathrm{int}} : \quad (F_h + (F_c(p_h))^e_E + (F_h + (F_c(p_h))^e_{E'} = 0. \tag{2.43}$$

The schemes presented in Sections 2.4.1 and 2.4.2 can all be included in this general setting, with the following choices of A and B:

- *centred scheme*: $A(s) = A_{\mathrm{ce}}(s) := \frac{s}{2}$ and $B(s) = -A_{\mathrm{ce}}(-s) = \frac{s}{2}$;

- *upwind scheme*: $A(s) = A_{\mathrm{up}}(s) := s^+$ and $B(s) = -A_{\mathrm{up}}(-s) = -s^-$;

- *θ-scheme*: $A(s) = A_\theta(s) := (1 - 2\theta)A_{\mathrm{up}}(s) + 2\theta A_{\mathrm{ce}}(s)$ and $B(s) = -A_\theta(-s)$;

- *Scharfetter–Gummel scheme*: $A(s) = A_{\mathrm{sg}}(s)$ defined by (2.22) and $B(s) = -A_{\mathrm{sg}}(-s)$, and the locally scaled Scharfetter–Gummel scheme is obtained by using $A_{\mathrm{sg}, \Lambda, e}$ defined by (2.23) instead of A_{sg};

- *stabilized MFD scheme*: $A(s) = s + \frac{\alpha}{2}|s|$ and $B(s) = -\frac{\alpha}{2}|s|$.

The first four choices in (2.40) lead to a conservative definition of the numerical convection flux, whereas the last one does not. However, in all of the cases mentioned above, total conservation is ensured by (2.43). We note that all of these choices of A and B satisfy the following properties:

(AB1) $A : \mathbf{R} \to \mathbf{R}$ and $B : \mathbf{R} \to \mathbf{R}$ are Lipschitz-continuous functions and $A(0) = B(0) = 0$;

(AB2) $A(s) + B(s) = s$ for any real number s;

(AB3) One of the following two alternatives holds:

  (AB3-s) $A(s) + B(-s) = 0$ and $A(s) - B(s) \geqslant 0$ for any real number s;

  (AB3-w) the function $s \to A(s) + B(-s)$ is odd and there exists $C > 0$ such that $A(s) - B(s) \geqslant -C|s|$ for any real number s.

We refer to (AB3-s) as the *strong* (AB3) condition and to (AB3-w) as the *weak* (AB3) condition. Assumption (AB3-s) is satisfied by all of the FV-based discretizations listed above, whereas the MFD-based discretization satisfies (AB3-w). In fact, condition $A(s) + B(-s) = 0$ in (AB3-s) is the one ensuring the conservation of the numerical convection flux (2.40). On the other hand, the numerical convection flux extracted from the MFD-based formulation satisfies (AB3-w) and hence is not conservative. We will see in Section 3 that assumptions (AB1)–(AB3) are enough to carry out the theoretical analysis of the scheme in (2.39)–(2.43), with slightly different results depending on which alternative in (AB3) is satisfied.

REMARK 2.6 It is worth noting that equation (2.42) can be rewritten in an FV form as the following cell-based flux balance equation:

$$\forall E \in \Omega_h : \quad \sum_{e \in \partial E} |e|(F_E^e + (F_c(p_h))_E^e) = \int_E f. \qquad (2.44)$$

REMARK 2.7 We could also choose, in (2.40), different functions $A = A^e$ and $B = B^e$ for each edge e, provided that all these functions satisfy (AB1)–(AB3) and that their Lipschitz constants remain uniformly bounded as the mesh size tends to 0. This setting would allow the scheme to make a finer tuning of the numerical diffusion due to upwinding, thus better adapting the scheme behaviour to the location inside the domain or the local geometry of the mesh.

2.4.4 *An alternative hybrid discretization of the convection term.* An alternative discretization of the convection term is possible by using the hybridized value p_e in (2.40) instead of $p_{E'}$, an idea introduced in Arnold & Brezzi (1985). In such a case we define the numerical convection flux of the discrete scalar field described by $(p_h, p_{\mathcal{E}_h}) \in Q_h \times H(\mathcal{E}_h)$ as the collection of real numbers

$$F_{c,\mathcal{E}_h}(p_h, p_{\mathcal{E}_h}) = ((F_{c,\mathcal{E}_h}(p_h, p_{\mathcal{E}_h}))_E^e)_{E \in \Omega_h, e \in \partial E} \qquad (2.45)$$

such that

$$\forall E \in \Omega_h, \forall e \in \partial E : (F_{c,\mathcal{E}_h}(p_h, p_{\mathcal{E}_h}))_E^e = \frac{1}{d_e}(A(d_e V_E^e)p_E + B(d_e V_E^e)p_e). \qquad (2.46)$$

The substantial difference with the preceding choice (2.40) is that no property of A and B ensures that the fluxes $F_{c,\mathcal{E}_h}(p_h, p_{\mathcal{E}_h})$ are conservative (and they are not in general). However, this will not bring any additional difficulty in the theoretical study provided that the following weaker form of (AB3) is considered.

(AB3h) One of the following strong or weak alternatives holds:

(AB3h-s) $A(s) - B(s) \geqslant 0$ for any real number s;

(AB3h-w) there exists $C > 0$ such that $A(s) - B(s) \geqslant -C|s|$ for any real number s.

The hybrid HMM formulation can then be written as follows:

find $(p_h, F_h, p_{\mathcal{E}_h}) \in Q_h \times \widehat{X}_h \times H(\mathcal{E}_h)$ such that

$$\forall E \in \Omega_h, \quad \forall G_E \in X_E : \qquad [F_E, G_E]_E = \sum_{e \in \partial E} |e| G_E^e (p_E - p_e), \tag{2.47}$$

$$\forall E \in \Omega_h : \qquad \sum_{e \in \partial E} |e| (F_E^e + (F_{c,\mathcal{E}_h}(p_h, p_{\mathcal{E}_h}))_E^e) = \int_E f, \tag{2.48}$$

$$\forall e \in \mathcal{E}_{h,\text{int}} : \quad (F_h + (F_{c,\mathcal{E}_h}(p_h, p_{\mathcal{E}_h}))_E^e + (F_h + (F_{c,\mathcal{E}_h}(p_h, p_{\mathcal{E}_h}))_{E'}^e = 0, \tag{2.49}$$

where the local scalar products used in (2.47) satisfy (S1) and (S2) and may thus be given in the form (2.11)–(2.13).

REMARK 2.8 An important advantage of discretizing the convective fluxes by using (2.45) and (2.46) instead of (2.39) and (2.40) is that the unknowns p_h and F_h in the resulting numerical formulation (2.47)–(2.49) can be eliminated by *static condensation*, that is, through a local Gaussian elimination (this classical technique is not directly applicable to (2.39) and (2.40)). This procedure, which is common for hybrid mixed finite elements, provides a reduced linear system in the face unknowns $p_{\mathcal{E}_h}$. Moreover, when the discretization of the convection term increases significantly the numerical diffusion, as, for example, in the case of the upwind scheme, the hybrid version of the HMM method is likely to be less diffusive than that provided by (2.39) and (2.40).

3. Theoretical study

In the present section we develop the theoretical analysis for the class of methods that we wish to investigate in this work. In Section 3.1 we prove the convergence of the numerical approximations to the exact solution and its gradient. The analysis is based on a compactness argument, which is common in the FV literature, under the weaker assumptions of mesh regularity (G1) and (G2) (see also Definition 2.2). In Section 3.2 we prove an $\mathcal{O}(h)$ convergence rate for the numerical approximation of both the scalar solution and the flux. The analysis involves stability and consistency arguments, which are in the MFD (and finite-element method) literature, under the stronger mesh regularity assumptions (HG) and (ME).

Let us introduce the mesh-dependent norms for the spaces X_h and Q_h. Let $\mathcal{D}_h$ be an admissible mesh in accordance with Definition 2.2 that satisfies (G1) and (G2) or, alternatively, (HG) and (ME). The scalar product in $\widehat{X}_h$ induces the norm

$$\|G\|_{\widehat{X}_h}^2 = [G, G]_{\widehat{X}_h} \quad \forall G \in \widehat{X}_h, \tag{3.1}$$

and its local counterpart

$$\|G\|_E^2 = [G_E, G_E]_E \quad \forall G_E \in X_E. \tag{3.2}$$

The elements of Q_h can be identified with the Ω_h-piecewise constant functions and the scalar product in Q_h is, in fact, the L^2 scalar product for such functions. Therefore it is quite natural to consider

the L^2 norm. However, we will also find it useful to carry out the analysis by using the discrete H_0^1-like norm

$$\|q_h\|_{1,\mathcal{D}_h} = \left(\sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| d_{E,e} \left(\frac{|q_E - q_{E'}|}{d_e} \right)^2 \right)^{1/2} \qquad \forall q_h \in Q_h, \tag{3.3}$$

where E' is the cell on the other side of $e \in \partial E \cap \mathcal{E}_{h,\text{int}}$ and, to ease notation, we take $q_{E'} = 0$ if $e \in \partial E \cap \mathcal{E}_{h,\text{ext}}$. We will also need the following discrete H^1 norm on $Q_h \times H(\mathcal{E}_h)$:

$$\|(q_h, q_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h} = \left(\sum_{E \in \Omega_h} \sum_{e \in \partial E} \frac{|e|}{d_{E,e}} |q_E - q_e|^2 \right)^{1/2} \qquad \forall (q_h, q_{\mathcal{E}_h}) \in Q_h \times H(\mathcal{E}_h). \tag{3.4}$$

It is easy to see that this norm is stronger than (3.3). More precisely, if $\theta \geqslant \text{regul}(\mathcal{D}_h)$, then there exists a constant C that is only dependent on θ such that, for all $(q_h, q_{\mathcal{E}_h}) \in Q_h \times H(\mathcal{E}_h)$, we have that

$$\|q_h\|_{1,\mathcal{D}_h} \leqslant C\|(q_h, q_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h}. \tag{3.5}$$

In the following developments we will number all constants whose value may be zero depending on which alternative is considered in (AB3), that is, the strong (AB3-s) or the weak (AB3-w) condition. We will also use the symbol $\lesssim$ to indicate an upper bound that holds up to a positive multiplicative constant that is independent of h. However, we will trace explicitly the constants where required by the proofs or that may be zero depending on the choice of assumption (AB3).

LEMMA 3.1 Let us assume that (H1)–(H3) hold. Let $\mathcal{D}_h$ be an admissible discretization of Ω such that $\theta \geqslant \text{regul}(\mathcal{D}_h)$, and let $F_c(q)$ be the convective flux of $q \in Q_h$ given by (2.39) and (2.40) for the vector field $V \in C^1(\overline{\Omega})^d$ with A and B satisfying assumptions (AB1)–(AB3). Then there exists a non-negative constant $C_1 \geqslant 0$ that only depends on θ, V, A *and* B such that

$$\forall (q, q_{\mathcal{E}_h}) \in Q_h \times H(\mathcal{E}_h) :$$

$$\frac{1}{2} \int_\Omega q^2 \text{div}(V) \leqslant \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| (F_c(q))_E^e (q_E - q_e) + C_1 h \|(q, q_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h}^2, \tag{3.6}$$

and where $C_1 = 0$ if (AB3-s) holds.

Proof. By gathering the sum by faces, we transform the term involving $F_c(q)$ on the right-hand side of (3.6) as follows:

$$\sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| (F_c(q))_E^e (q_E - q_e) = \sum_{e \in \mathcal{E}_h} |e| ((F_c(q))_E^e (q_E - q_e) + (F_c(q))_{E'}^e (q_{E'} - q_e))$$

$$= \sum_{e \in \mathcal{E}_h} |e| (F_c(q))_E^e (q_E - q_{E'}) \tag{3.7}$$

$$+ \sum_{e \in \mathcal{E}_h} |e| ((F_c(q))_E^e + (F_c(q))_{E'}^e)(q_{E'} - q_e).$$

To handle the first term on the right-hand side of (3.7) we note that, by using (2.40) and writing, due to (AB2),

$$
\begin{aligned}
A(d_e V_E^e) &= \tfrac{1}{2}(d_e V_E^e + A(d_e V_E^e) - B(d_e V_E^e)), \\
B(d_e V_E^e) &= \tfrac{1}{2}(d_e V_E^e + B(d_e V_E^e) - A(d_e V_E^e)),
\end{aligned}
\tag{3.8}
$$

we have

$$
(F_c(q))_E^e = \frac{1}{2} V_E^e (q_E + q_{E'}) + \frac{1}{2d_e}(A(d_e V_E^e) - B(d_e V_E^e))(q_E - q_{E'}).
$$

Therefore we infer that

$$
\begin{aligned}
\sum_{e \in \mathcal{E}_h} |e|(F_c(q))_E^e (q_E - q_{E'}) &= \frac{1}{2} \sum_{e \in \mathcal{E}_h} |e| V_E^e (q_E + q_{E'})(q_E - q_{E'}) \\
&\quad + \frac{1}{2} \sum_{e \in \mathcal{E}_h} \frac{|e|}{d_e}(A(d_e V_E^e) - B(d_e V_E^e))(q_E - q_{E'})^2.
\end{aligned}
\tag{3.9}
$$

Then, let us observe that

$$
\sum_{e \in \mathcal{E}_h} |e|(q_E - q_{E'})^2 \lesssim \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e|(q_E - q_e)^2 \lesssim h \|(q, q_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}^2.
\tag{3.10}
$$

By using (AB3), the conservation of $(V_E^e)_{E \in \Omega_h, e \in \partial E}$, the fact that $\sum_{e \in \partial E} |e| V_E^e = \int_E \mathrm{div}(V)$, and inequality (3.10), we obtain the following estimate:

$$
\begin{aligned}
\sum_{e \in \mathcal{E}_h} |e|(F_c(q))_E^e (q_E - q_{E'}) &\geqslant \frac{1}{2} \sum_{e \in \mathcal{E}_h} |e| V_E^e (q_E^2 - q_{E'}^2) - C_2 \sum_{e \in \mathcal{E}_h} |e|(q_E - q_{E'})^2 \\
&\geqslant \frac{1}{2} \sum_{E \in \Omega_h} q_E^2 \sum_{e \in \partial E} |e| V_E^e - C_3 h \|(q, q_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}^2 \\
&\geqslant \frac{1}{2} \int_\Omega q^2 \mathrm{div}(V) - C_3 h \|(q, q_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}^2,
\end{aligned}
\tag{3.11}
$$

where C_2 and C_3 only depend on θ, V, A and B, and $C_2 = C_3 = 0$ if (AB3-s) holds.

From (2.40) and since $V_E^e = -V_{E'}^e$, we have

$$
(F_c(q))_E^e + (F_c(q))_{E'}^e = \frac{1}{d_e}([A(d_e V_E^e) + B(-d_e V_E^e)]q_E + [B(d_e V_E^e) + A(-d_e V_E^e)]q_{E'}).
$$

If (AB3-s) holds, then this quantity is equal to zero (this is the conservation of the convective flux), and if (AB3-w) holds, then we have, due to (AB1), that

$$
|(F_c(q))_E^e + (F_c(q))_{E'}^e| = \frac{1}{d_e}|(A(d_e V_E^e) + B(-d_e V_E^e))(q_E - q_{E'})| \leqslant C_4 \|V\|_\infty |q_E - q_{E'}|
$$

for some C_4 that is only dependent on A and B. Writing $|q_E - q_{E'}| \leqslant |q_E - q_e| + |q_e - q_{E'}|$ and again using inequality (3.10) allows us to estimate the last term of (3.7) as follows:

$$\left| \sum_{e \in \mathcal{E}_h} |e|((F_c(q))^e_E + (F_c(q))^e_{E'})(q_{E'} - q_e) \right| \leqslant C_5 \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e|(q_E - q_e)^2 \leqslant C_5 h \|(q, q_{\mathcal{E}_h})\|^2_{1, \mathcal{D}_h, \mathcal{E}_h}, \tag{3.12}$$

where C_5 only depends on V, A and B, and we have that $C_5 = 0$ if (AB3-s) holds.

The proof terminates by gathering inequalities (3.11) and (3.12) into (3.7). □

3.1 *Convergence of the method*

3.1.1 *Preliminary results.*   Proposition 3.2 below is the key point in the study of the scheme (2.39)–(2.43) since it gives the inequality leading to the basic *a priori* estimates of the solution error. To state this proposition, we first note that, due to (2.41), we can introduce the set of face values $p_{\mathcal{E}_h} \in H(\mathcal{E}_h)$ such that (2.47) holds even if F_h is not conservative. For this purpose, we simply define p_e through $|e|(p_E - p_e) = [F_E, G_E(E, e)]_E$, where $G_E(E, e)_e = 1$ and $G_E(E, e)_{e'} = 0$ for $e \neq e'$. Then, taking the vector $G \in X_h$ that vanishes on all mesh faces except e and is such that $G^e_E = 1$ and $G^e_{E'} = -1$ in (2.41) allows us to show that p_e does not depend on the choice of the cell E such that $e \in \partial E$. This definition also ensures that $p_e = 0$ whenever $e \in \mathcal{E}_{h, \text{ext}}$.

PROPOSITION 3.2   Let us assume that (H1)–(H3) hold. Let $\mathcal{D}_h$ be an admissible discretization of Ω such that $\theta \geqslant \text{regul}(\mathcal{D}_h)$, and let $F_c(q)$ be the convective flux of $q \in Q_h$ given by (2.39) and (2.40) for the vector field $V \in C^1(\overline{\Omega})^d$ with A and B satisfying assumptions (AB1)–(AB3). Then, for all solutions (p_h, F_h) to the HMM scheme (2.41)–(2.43) we have

$$\sum_{E \in \Omega_h} [F_E, F_E]_E + \frac{1}{2} \int_\Omega \text{div}(V) p_h^2 \leqslant \int_\Omega f p_h + C_1 h \|(p_h, p_{\mathcal{E}_h})\|^2_{1, \mathcal{D}_h, \mathcal{E}_h}, \tag{3.13}$$

where C_1, which is the same constant as in Lemma 3.1, is non-negative, only depends on θ, V, A and B, and is zero when (AB3-s) holds.

Proof.   Let us take $q = p_h$ in (2.42), use the flux conservation (2.43) and property (2.47) of face values to obtain the following:

$$\int_\Omega f p_h = \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e|(F^e_E + (F_c(p_h))^e_E) p_E$$

$$= \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e|(F^e_E + (F_c(p_h))^e_E)(p_E - p_e)$$

$$= \sum_{E \in \Omega_h} [F_E, F_E]_E + \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e|(F_c(p_h))^e_E(p_E - p_e). \tag{3.14}$$

The proposition follows by applying Lemma 3.1 with $q = p_h$ and $q_{\mathcal{E}_h} = p_{\mathcal{E}_h}$. □

COROLLARY 3.3   Under the assumptions of Proposition 3.2, if V satisfies (H4) then, for all solutions (p_h, F_h) to the scheme (2.41)–(2.43), we have

$$\|(p_h, p_{\mathcal{E}_h})\|^2_{1, \mathcal{D}_h, \mathcal{E}_h} \lesssim \|f\|_{L^2(\Omega)} \|p_h\|_{L^2(\Omega)} + C_1 h \|(p_h, p_{\mathcal{E}_h})\|^2_{1, \mathcal{D}_h, \mathcal{E}_h}, \tag{3.15}$$

where C_1, which is the same constant as in Lemma 3.1 and Proposition 3.2, is non-negative, only depends on θ, V, A and B, and is zero when (AB3-s) holds.

In particular, for all h small enough (or any h if (AB3-s) holds) the scheme (2.41)–(2.43) has a unique solution.

Proof. We apply Proposition 3.2 and use (H4) and the form (2.11)–(2.13) of the local scalar products $([\cdot, \cdot]_E)_{E \in \Omega_h}$ to write, due to (C),

$$\sum_{E \in \Omega_h} |E| \, |\mathbf{v}_E(F_E)|^2 + \alpha \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| d_{E,e} |T_{E,e}(F_E)|^2 \leqslant \int_\Omega f p_h + C_1 h \|(p_h, p_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}^2. \quad (3.16)$$

From (2.47) and (2.11)–(2.13) we have

$$|e|(p_E - p_e) = |E| \Lambda_E \mathbf{v}_E(F_E) \cdot \mathbf{v}_E(G_E(e)) + T_E(G_E(e))^{\mathrm{T}} \mathbb{B}_E T_E(F_E), \quad (3.17)$$

where $G_E(e) \in X_E$ is equal to 1 on the face e and 0 on the other faces. But $\mathbf{v}_E(G_E(e)) = -\frac{1}{|E|} \Lambda_E^{-1} |e|$ $(\bar{x}_e - x_E)$, and thus, by the bound on regul$(\mathcal{D}_h)$, we have $|\mathbf{v}_E(G_E(e))| \lesssim \frac{|e| d_{E,e}}{|E|}$, and for all $e' \in \partial E$ we have $|T_{E,e'}(G_E(e))| \lesssim |G_E(e)^{e'}| + \frac{|e| d_{E,e}}{|E|}$. In particular, by using the Cauchy–Schwarz inequality and (C) since $\sum_{e' \in \partial E} |e'| d_{E,e'} = d |E|$, we obtain

$$|T_E(G_E(e))^{\mathrm{T}} \mathbb{B}_E T_E(F_E)| \lesssim \left(\sum_{e' \in \partial E} |e'| d_{E,e'} |T_{E,e'}(G_E(e))|^2 \right)^{1/2} \left(\sum_{e' \in \partial E} |e'| d_{E,e'} |T_{E,e'}(F_E)|^2 \right)^{1/2}$$

$$\lesssim \left(|e| d_{E,e} + \frac{|e|^2 d_{E,e}^2}{|E|} \right)^{1/2} \left(\sum_{e' \in \partial E} |e'| d_{E,e'} |T_{E,e'}(F_E)|^2 \right)^{1/2}.$$

When substituted into (3.17), this estimate and $|E| \, |\mathbf{v}_E(G_E(e))| \lesssim |e| d_{E,e}$ lead to

$$|p_E - p_e| \lesssim d_{E,e} |\mathbf{v}_E(F_E)| + \left(\frac{d_{E,e}}{|e|} + \frac{d_{E,e}^2}{|E|} \right)^{1/2} \left(\sum_{e' \in \partial E} |e'| d_{E,e'} |T_{E,e'}(F_E)|^2 \right)^{1/2}.$$

We then obtain, from (3.16), that

$$\sum_{E \in \Omega_h} \sum_{e \in \partial E} \frac{|e|}{d_{E,e}} |p_E - p_e|^2 \lesssim \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| d_{E,e} |\mathbf{v}_E(F_E)|^2$$

$$+ \sum_{E \in \Omega_h} \sum_{e \in \partial E} \left(1 + \frac{|e| d_{E,e}}{|E|} \right) \left(\sum_{e' \in \partial E} |e'| d_{E,e'} |T_{E,e'}(F_E)|^2 \right)$$

$$\lesssim \int_\Omega f p_h + C_1 h \|(p_h, p_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}^2,$$

and the proof of (3.15) is completed.

The existence and uniqueness of the numerical solution readily follow from (3.15). In fact, when the right-hand side f vanishes, this inequality implies that the mesh-dependent norm $\|(p_h, p_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}$ is zero and thus that $(p_h, p_{\mathcal{E}_h})$ are zero at least for a sufficiently small mesh size h. In such a case, the numerical flux F_h is also zero by (2.47). □

REMARK 3.4 (Estimates for the hybrid discretization of the convection). For the hybrid discretization in (2.45)–(2.49) with A and B satisfying (AB1), (AB2) and (AB3h) a result holds that is similar to that given in Proposition 3.2 and Corollary 3.3. However, the proof is simpler. In fact, by using (3.8), we have that

$$\sum_{E \in \Omega_h} \sum_{e \in \partial E} |e|(F_{c,\mathcal{E}_h}(q, q_{\mathcal{E}_h}))_E^e(q_E - q_e) = \frac{1}{2} \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| V_E^e(q_E + q_e)(q_E - q_e)$$

$$+ \frac{1}{2} \sum_{E \in \Omega_h} \sum_{e \in \partial E} \frac{|e|}{d_e}(A(d_e V_E^e) - B(d_e V_E^e))(q_E - q_e)^2.$$

The right-hand side of this equation is similar to the right-hand side of (3.9) with q_e instead of $q_{E'}$ and, reasoning as in the proof of Lemma 3.1, can be bounded from below by the the right-hand side of (3.11). The resulting estimate is then used in (3.14) with $F_{c,\mathcal{E}_h}(p_h, p_{\mathcal{E}_h})$ instead of $F_c(p_h)$ in order to prove Proposition 3.2.

We conclude this preliminary subsection by reporting two technical lemmas that we will use in the analysis of the next subsection. The first lemma is a direct consequence of Eymard *et al.* (2009, Lemmas 5.2 and 5.3) and, for this reason, is given without proof.

LEMMA 3.5 (Discrete Sobolev inequalities). Let $\mathcal{D}_h$ be an admissible discretization of Ω in the sense of Definition 2.2 with $\theta > 0$ and such that $\theta \leqslant \frac{d_{E,e}}{d_{E',e}} \leqslant \theta^{-1}$ for all $e \in \mathcal{E}_{h,\mathrm{int}}$. Let $r = \frac{2d}{d-2}$ if $d > 2$ and $r < +\infty$ if $d = 2$. Then there exists a real positive constant C that only depends on Ω, θ and r such that, for all $q_h \in Q_h$, we have $\|q_h\|_{L^r(\Omega)} \leqslant C \|q_h\|_{1,\mathcal{D}_h}$.

Let $\mathbf{v}_h(F_h)$ denote the piecewise constant function that is equal to $\mathbf{v}_E(F_E)$ on $E \in \Omega_h$ as defined in (2.12).

LEMMA 3.6 (Discrete Rellich theorem). Let $\Lambda: \Omega \to M_d(\mathbf{R})$ be a diffusion tensor satisfying hypothesis (H2). Let $(\mathcal{D}_h)_{h \to 0}$ be a family of admissible discretizations of Ω in the sense of Definition 2.2 with mesh size h tending to 0 and satisfying the regularity assumptions (G1) and (G2). Let $(p_h, p_{\mathcal{E}_h}) \in Q_h \times H(\mathcal{E}_h)$ be a numerical scalar field such that $\|(p_h, p_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h}$ remains bounded as $h \to 0$. Let $F_h = (F_E^e)_{E \in \Omega_h, \, e \in \partial E}$ be a collection of numbers that satisfy equation (2.47) for the assigned $(p_h, p_{\mathcal{E}_h})$, with the local scalar products defined according to (2.11)–(2.13).

Then there exists a scalar field $p \in H_0^1(\Omega)$ such that, up to a subsequence as $h \to 0$, the following hold:

(i) $p_h \to p$ in $L^r(\Omega)$ for all $r < \frac{2d}{d-2}$;

(ii) $\mathbf{v}_h(F_h) \to \nabla p$ weakly in $L^2(\Omega)^d$.

Proof. Using Lemma 5.6 of Eymard *et al.* (2009), Lemma 3.5, Vitali's theorem and the fact that the quantity $\|(p_h, p_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h}$ is uniformly bounded, ensures that $(p_h)_{h \to 0}$ is relatively compact in $L^r(\Omega)$ for all $r < \frac{2d}{d-2}$. After defining the discrete gradient $\widetilde{\nabla}(p_h, p_{\mathcal{E}_h}) : \Omega \to \mathbf{R}^d$ by

$$\forall E \in \Omega_h, \quad \forall x \in E : \widetilde{\nabla}(p_h, p_{\mathcal{E}_h})(x) = \frac{1}{|E|} \sum_{e \in \partial E} |e|(p_e - p_E)\mathbf{n}_E^e,$$

we see from the bound on $\|(p_h, p_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h}$ that $\widetilde{\nabla}(p_h, p_{\mathcal{E}_h})$ remains bounded in $L^2(\Omega)^d$. The technique used to prove Lemma 5.7 of Eymard *et al.* (2009) ensures that, if $p_h \to p$ in $L^2(\Omega)$ (up to a subsequence), then p belongs to $H_0^1(\Omega)$ and $\widetilde{\nabla}(p_h, p_{\mathcal{E}_h})$ is weakly convergent to ∇p in $L^2(\Omega)^d$. The

lemma is therefore true since the argument discussed in Droniou *et al.* (2010, Remark 2.7) implies that $\widetilde{\nabla}(p_h, p_{\mathcal{E}_h}) = \mathbf{v}_h(F_h)$ if $(p_h, p_{\mathcal{E}_h}, F_h)$ are linked through (2.11)–(2.13) and (2.47). $\qquad\square$

3.1.2 *Convergence without regularity assumption.* Let us consider the HMM method on $(\mathcal{D}_h)_{h \to 0}$, a family of meshes that are admissible according to Definition 2.2, with mesh size h tending to 0 and all of which satisfy the regularity conditions (G1) and (G2). We also assume that all of the local scalar products in the scheme formulation are defined by (2.11)–(2.13) through a set of symmetric and positive-definite matrices $(\mathbb{B}_E)_{E \in \Omega_h}$ that satisfy the coercivity condition (C). Moreover, the numerical convection flux $F_c(p_h)$ in (2.42) is constructed by using (2.39) and (2.40) through some instance of the functions A and B that satisfy (AB1)–(AB3). Finally, we recall that $\mathbf{v}_h(F_h) : \Omega \to \mathbf{R}^d$ is the piecewise constant function that is equal to $\mathbf{v}_E(F_E)$ on E for all $E \in \Omega_h$. The convergence result of this subsection is stated in the following theorem.

THEOREM 3.7 Let $p \in H_0^1(\Omega)$ be the weak solution to (2.1) and (2.2) under assumptions (H1)–(H4), and let (p_h, F_h) be the numerical solution to problem (2.41)–(2.43) constructed along the guidelines summarized above. Then, for $h \to 0$, the following hold:

(i) $p_h \to p$ in $L^r(\Omega)$ for all $r < \frac{2d}{d-2}$;

(ii) $\mathbf{v}_h(F_h) \to \nabla p$ in $L^2(\Omega)^d$.

Proof. The proof of Theorem 3.7 is based on compactness tools developed for mixed FV or hybrid FV for the pure diffusion equation (Droniou & Eymard, 2006; Eymard *et al.*, 2009) and on techniques from classical FV schemes (Eymard *et al.*, 2000; Chainais-Hillairet & Droniou, 2009) for handling the numerical convection term. We report the full proof for the sake of completeness since none of these methods has ever been formulated in the new HMM framework.

Step 1: Compactness of the approximate solutions. Using Corollary 3.3, we have $\|(p_h, p_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h}^2$ $\lesssim \|f\|_{L^2(\Omega)} \|p_h\|_{L^2(\Omega)}$ (at least for h small enough if (AB3-s) does not hold). In view of Lemma 3.5 and inequality (3.5), we obtain an upper bound on $\|(p_h, p_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h}$. Then the result of Lemma 3.6 implies the existence of a function $p \in H_0^1(\Omega)$ such that, up to a subsequence, $p_h \to p$ in $L^r(\Omega)$ for all $r < \frac{2d}{d-2}$ and $\mathbf{v}_h(F_h) \to \nabla p$ *weakly* in $L^2(\Omega)^d$.

Step 2: The limit function p is the weak solution to (2.1) and (2.2). Since the exact solution is unique, this step allows us to prove the convergence to p of the whole sequence of discrete solutions p_h for $h \to 0$. We take $\varphi \in C_c^\infty(\Omega)$, define $\varphi_h \in Q_h$ by $\varphi_h = \varphi(x_E)$ on $E \in \Omega_h$ and substitute $q = \varphi_h$ into (2.42). Since $F_h + F_c(p_h)$ is conservative, we obtain

$$\int_\Omega f\varphi_h = \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| F_E^e (\varphi(x_E) - \varphi(\bar{x}_e))$$

$$+ \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| (\varphi(x_E) - \varphi(\bar{x}_e)) \frac{1}{d_e} (A(d_e V_E^e) p_E + B(d_e V_E^e) p_{E'})$$

$$= \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| F_E^e (x_E - \bar{x}_e) \cdot \nabla\varphi(x_E) + \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| F_E^e R_{E,e}^h(\varphi)$$

$$+ \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| (\varphi(x_E) - \varphi(\bar{x}_e)) \frac{1}{d_e} (A(d_e V_E^e) + B(d_e V_E^e)) p_E$$

$$+ \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e|(\varphi(x_E) - \varphi(\bar{x}_e)) \frac{1}{d_e} B(d_e V_E^e)(p_{E'} - p_E)$$

$$= T_1 + T_2 + T_3 + T_4, \tag{3.18}$$

where the residual term $R_{E,e}^h(\varphi)$ in T_2 is such that $|R_{E,e}^h(\varphi)| \lesssim d_{E,e} h \|\nabla^2 \varphi\|_\infty$.

By (2.12), we have that

$$T_1 = \sum_{E \in \Omega_h} |E| \Lambda_E \mathbf{v}_E(F_E) \cdot \nabla\varphi(x_E) = \int_\Omega \Lambda \mathbf{v}_h(F_h) \cdot (\nabla\varphi)_h,$$

where $(\nabla\varphi)_h = \nabla\varphi(x_E)$ on $E \in \Omega_h$. The regularity of φ together with the weak convergence of $\mathbf{v}_h(F_h)$ implies that

$$T_1 \longrightarrow \int_\Omega \Lambda \nabla p \cdot \nabla\varphi \quad \text{as } h \longrightarrow 0. \tag{3.19}$$

From (2.13) we have $|F_E^e| \lesssim |T_{E,e}(F_E)| + |\mathbf{v}_E(F_E)|$ and, since $\|p_h\|_{L^2(\Omega)}$ and $\|(p_h, p_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h}$ are bounded, inequality (3.16) implies that

$$\sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| d_{E,e} |F_E^e| \leqslant \left(\sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| d_{E,e} \right)^{1/2} \left(\sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| d_{E,e} |F_E^e|^2 \right)^{1/2} \lesssim 1$$

$\left(\text{recall that } \sum_{e \in \partial E} |e| d_{E,e} = d|E| \right)$. Therefore we obtain that

$$|T_2| \lesssim h \|\nabla^2 \varphi\|_\infty \longrightarrow 0 \quad \text{as } h \longrightarrow 0. \tag{3.20}$$

Assumption (AB2) makes it possible to show that

$$T_3 = \sum_{E \in \Omega_h} p_E \sum_{e \in \partial E} |e|(\varphi(x_E) - \varphi(\bar{x}_e)) V_E^e$$

$$= \sum_{E \in \Omega_h} p_E \varphi(x_E) \sum_{e \in \partial E} |e| V_E^e - \sum_{E \in \Omega_h} p_E \sum_{e \in \partial E} |e| \varphi(\bar{x}_e) V_E^e$$

$$= \int_\Omega p_h \varphi_h \text{div}(V) - \sum_{E \in \Omega_h} p_E \sum_{e \in \partial E} \int_e \varphi V \cdot \mathbf{n}_E^e + \sum_{E \in \Omega_h} p_E \sum_{e \in \partial E} \int_e (\varphi - \varphi(\bar{x}_e)) V \cdot \mathbf{n}_E^e$$

$$= \int_\Omega p_h \varphi_h \text{div}(V) - \int_\Omega p_h \text{div}(\varphi V) + \sum_{E \in \Omega_h} p_E \sum_{e \in \partial E} \int_e (\varphi - \varphi(\bar{x}_e)) V \cdot \mathbf{n}_E^e.$$

The regularity of φ and the convergence of p_h ensure that, as $h \to 0$, the first two terms on this right-hand side tend to $\int_\Omega p\varphi\text{div}(V)$ and $\int_\Omega p\text{div}(\varphi V)$. As for the last term, using the fact that $\int_e (\varphi - \varphi(\bar{x}_e)) V \cdot \mathbf{n}_E^e$ vanishes for boundary faces (φ has a compact support) and is conservative for interior

faces (i.e., by changing E to E', the cell on the other side of e, only the sign is changed), we find that

$$\left| \sum_{E \in \Omega_h} p_E \sum_{e \in \partial E} \int_e (\varphi - \varphi(\bar{x}_e)) V \cdot \mathbf{n}_E^e \right| = \left| \sum_{E \in \Omega_h} \sum_{e \in \partial E} (p_E - p_e) \int_e (\varphi - \varphi(\bar{x}_e)) V \cdot \mathbf{n}_E^e \right|$$

$$\lesssim h \|\nabla \varphi\|_\infty \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| |p_E - p_e|.$$

But the Cauchy–Schwarz inequality and the bound on $\|(p_h, p_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}$ give that

$$\sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| |p_E - p_e| \leqslant (d|\Omega|)^{1/2} \|(p_h, p_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h} \lesssim 1, \tag{3.21}$$

and thus $\sum_{E \in \Omega_h} p_E \sum_{e \in \partial E} \int_e (\varphi - \varphi(\bar{x}_e)) V \cdot \mathbf{n}_E^e$ tends to 0 with h. We deduce that

$$T_3 \longrightarrow \int_\Omega p \varphi \operatorname{div}(V) - \int_\Omega p \operatorname{div}(\varphi V) = -\int_\Omega V p \cdot \nabla \varphi \quad \text{as } h \longrightarrow 0. \tag{3.22}$$

To handle T_4 we start by noting that assumption (AB1) implies that $\frac{1}{d_e} |B(d_e V_E^e)| \lesssim 1$. Thus, writing $p_{E'} - p_E = p_{E'} - p_e + p_e - p_E$ and using (3.21), we obtain

$$|T_4| \lesssim h \|\nabla \varphi\|_\infty \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| (|p_{E'} - p_e| + |p_e - p_E|)$$

$$\lesssim 2h \|\nabla \varphi\|_\infty \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| |p_E - p_e| \longrightarrow 0 \quad \text{as } h \longrightarrow 0. \tag{3.23}$$

Eventually, the convergence properties (3.19), (3.20), (3.22) and (3.23) allow us to obtain the limit of (3.18) for $h \to 0$ and show that p is the weak solution to (2.1) and (2.2).

Step 3: Strong convergence of the gradient.   Estimate (3.13) and relation (2.11) imply that

$$\int_\Omega \Lambda \mathbf{v}_h(F_h) \cdot \mathbf{v}_h(F_h) + \frac{1}{2} \int_\Omega \operatorname{div}(V) p_h^2 \leqslant \int_\Omega f p_h + Ch \|(p_h, p_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}^2.$$

Taking the upper limit of this inequality, recalling that $\|(p_h, p_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}$ stays bounded and noting that p_h is strongly convergent to p in $L^2(\Omega)$ and that p is the weak solution to (2.1) and (2.2) leads to

$$\limsup_{h \to 0} \int_\Omega \Lambda \mathbf{v}_h(F_h) \cdot \mathbf{v}_h(F_h) + \frac{1}{2} \int_\Omega \operatorname{div}(V) p^2 \leqslant \int_\Omega f p = \int_\Omega \Lambda \nabla p \cdot \nabla p + \frac{1}{2} \int_\Omega \operatorname{div}(V) p^2,$$

from which we deduce that

$$\limsup_{h \to 0} \int_\Omega \Lambda \mathbf{v}_h(F_h) \cdot \mathbf{v}_h(F_h) \leqslant \int_\Omega \Lambda \nabla p \cdot \nabla p. \tag{3.24}$$

Since $\mathbf{w} \to (\int_\Omega \Lambda \mathbf{w} \cdot \mathbf{w})^{1/2}$ is a norm in $L^2(\Omega)^d$ that is equivalent to the usual norm, equation (3.24) proves that the weak convergence $\mathbf{v}_h(F_h) \to \nabla p$ in $L^2(\Omega)^d$ is, in fact, strong. $\square$

REMARK 3.8 Adjusting these same arguments makes it possible to prove a similar convergence result for the hybrid HMM formulation (2.47)–(2.49) that is based on the numerical convection flux (2.45) and (2.46) instead of (2.39) and (2.40).

3.1.3 *About the regularity assumption on V.*   Often, the velocity field V is not given but comes from the resolution of another problem (see, e.g., Chainais-Hillairet & Droniou, 2007). In this case it is not obvious that it satisfies the regularity assumption $V \in C^1(\overline{\Omega})^d$. We can in general ensure that $V \in H(\text{div}, \Omega)$, but nothing more. How does this impact on the preceding convergence study?

We first of all have to be able to define the fluxes V_E^e of the velocity. This is, in general, quite straightforward either using (2.19) and the fact that V belongs to $H(\text{div}, \Omega)$ or even more directly by looking at the discretization of the equation providing V (this discretization also usually provides the fluxes of the velocity, as in Chainais-Hillairet & Droniou, 2007). The minimal requirement on these fluxes is their conservativity, namely

$$\forall e \in \mathcal{E}_{h,\text{int}}, : V_E^e + V_{E'}^e = 0$$

(where $E, E' \in \Omega_h$ are the two elements such that $e \subset \partial E \cap \partial E'$ for every $e \in \mathcal{E}_{h,\text{int}}$) and their compatibility with the coercitivity assumption $\text{div}(V) \geqslant 0$, namely

$$\forall E \in \Omega_h, : \sum_{e \in \partial E} |e| V_E^e \geqslant 0$$

(usually, $\sum_{e \in \partial E} |e| V_E^e$ plays the role of an approximation of $\int_E \text{div}(V)$). Under these two requirements and the strong version of assumption (AB3) (i.e., (AB3s)), it is then easy to see that the *a priori* estimates still hold (see Lemma 3.1, Proposition 3.2 and Corollary 3.3).

As for the convergence (Theorem 3.7), we have to check if T_3 and T_4 behave well. For T_3 we need that

$$\forall E \in \Omega_h, : \sum_{e \in \partial E} |e| V_E^e = \int_E \text{div}(V)$$

(or at least that $\sum_{e \in \partial E} |e| V_E^e$ approximates $\int_E \text{div}(V)$ as the size of the mesh tends to 0), which is usually the case from the definition of V_E^e using (2.19) or an expression of these fluxes coming from the resolution of another elliptic equation. For T_3 we also need that, for any smooth function φ with compact support, denoting by $\Phi_h : \Omega \to \mathbf{R}$ the function defined by

$$\forall E \in \Omega_h, \quad \forall x \in E : \Phi_h(x) = \sum_{e \in \partial E} \varphi(\bar{x}_e) |e| V_E^e,$$

the function Φ_h weakly converges in $L^2(\Omega)$, as $h \to 0$, to $\text{div}(\varphi V)$. Since, for any $W \in H(\text{div}, \Omega)$, defining $W_E^e = \frac{1}{|e|} \int_e W \cdot \mathbf{n}_E^e$ (in the usual weak sense), we have

$$|W_E^e|^2 \leqslant Ch^{-d} \|W\|_{L^2(E)}^2 + Ch^{-d+2} \|\text{div}(W)\|_{L^2(E)}^2 \tag{3.25}$$

(this is the usual Agmon scaling of trace estimates), the estimates we provide in the proof of Theorem 3.7 on the last part of T_3 indicate that Φ_h behaves as needed if V_E^e comes from (2.19) with $V \in H(\text{div}, \Omega)$. If these velocity fluxes come from the approximation of another elliptic equation, then the expected behaviour of Φ_h is usually a straightforward consequence of the properties of the scheme used on this other equation (see, e.g., Chainais-Hillairet & Droniou, 2007).

For T_4 we require that

$$\sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| d_{E,e} |V_E^e|^2 \tag{3.26}$$

remains bounded as the size of the mesh tends to 0. When V_E^e comes from the approximation of an elliptic equation (i.e., V_E^e is the 'F_E^e' of this other equation), then this estimate is usually a basic one (for HMM methods, for example, it is a direct consequence of (S1) or (C)). If V_E^e is constructed from (2.19) with $V \in H(\text{div}, \Omega)$, then (3.25) shows that (3.26) also remains bounded independently of the mesh size.

In other words, although the preceding study has been made, for the sake of simplicity, with regular velocity fields, it is easy to adapt to more realistic fields, and the convergence results still hold for these fields.

3.2 *Error estimates*

In the theoretical developments of this section we assume that (HG) and (ME) hold.

Now, we consider the bilinear form

$$\mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; G'_h, q'_h, q'_{\mathcal{E}_h}) = [G_h, G'_h]_{\widehat{X}_h} - \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| (G'_h)_E^e (q_E - q_e)$$

$$+ [\text{div}_h (G_h + F_c(q_h)), q'_h]_{Q_h} - \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| (G_h + F_c(q_h))_E^e q'_e \tag{3.27}$$

for all couples of triplets $(G_h, q_h, q_{\mathcal{E}_h})$ and $(G'_h, q'_h, q'_{\mathcal{E}_h})$ in $\widehat{X}_h \times Q_h \times H(\mathcal{E}_h)$. Problem (2.41)–(2.43) can be reformulated as follows:

find $(F_h, p_h, p_{\mathcal{E}_h}) \in \widehat{X}_h \times Q_h \times H(\mathcal{E}_h)$ such that

$$\mathcal{B}(F_h, p_h, p_{\mathcal{E}_h}; G_h, q_h, q_{\mathcal{E}_h}) = [f^I, q_h]_{Q_h} \quad \forall (G_h, q_h, q_{\mathcal{E}_h}) \in \widehat{X}_h \times Q_h \times H(\mathcal{E}_h). \tag{3.28}$$

In order to prove the convergence result we need the following stability lemma.

LEMMA 3.9 Assume (AB1)–(AB3) with either h small enough if (AB3-w) holds or any h if (AB3-s) holds. For any triple $(G_h, q_h, q_{\mathcal{E}_h}) \in \widehat{X}_h \times Q_h \times H(\mathcal{E}_h)$ there exists a triple $(G'_h, q'_h, q'_{\mathcal{E}_h}) \in \widehat{X}_h \times Q_h \times H(\mathcal{E}_h)$ with

$$\|G'_h\|_{\widehat{X}_h} + \|q'_h\|_{1, \mathcal{D}_h} + \|(q'_h, q'_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h} \leqslant 1 \tag{3.29}$$

for which the following holds:

$$\mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; G'_h, q'_h, q'_{\mathcal{E}_h}) \gtrsim \|G_h\|_{\widehat{X}_h} + \|q_h\|_{1, \mathcal{D}_h} + \|(q_h, q_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}. \tag{3.30}$$

Proof. A straightforward calculation shows that

$$\mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; G_h, q_h, q_{\mathcal{E}_h}) = \|G_h\|_{\widehat{X}_h}^2 + \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| (F_c(q_h))_E^e (q_E - q_e). \tag{3.31}$$

Since $\text{div}(V) \geqslant 0$, applying Lemma 3.1 yields the inequality

$$\mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; G_h, q_h, q_{\mathcal{E}_h}) \geqslant \|G_h\|_{\widehat{X}_h}^2 - C_1 h \|(q_h, q_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}^2. \tag{3.32}$$

The non-negative real constant C_1, which is provided by Lemma 3.1, is zero if assumption (AB3-s) holds.

Let us consider the *nonconservative* vector field $\widehat{G}_h \in \widehat{X}_h$ given by

$$\forall E \in \Omega_h, \quad \forall e \in \partial E : (\widehat{G}_h)_E^e = \frac{q_e - q_E}{d_{E,e}}.$$

Since (M2) implies that $|E| \lesssim |e| d_{E,e}$, we have that

$$\|\widehat{G}_h\|_{\widehat{X}_h}^2 = \sum_{E \in \Omega_h} \|\widehat{G}_E\|_E^2 \leqslant \sigma^* \sum_{E \in \Omega_h} \sum_{e \in \partial E} \frac{|E|}{d_{E,e}^2} (q_E - q_e)^2 \leqslant \widehat{C} \|(q_h, q_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}^2, \tag{3.33}$$

where $\widehat{C} > 0$ is independent of h and only depends on the constant σ^* of assumption (S1) and on the mesh regularity constants of (M2). We infer, from the Cauchy–Schwarz inequality and Young's inequality, that

$$|[G_h, \widehat{G}_h]_{\widehat{X}_h}| \leqslant \|G_h\|_{\widehat{X}_h} \|\widehat{G}_h\|_{\widehat{X}_h} \leqslant \frac{\widehat{C}}{2} \|G_h\|_{\widehat{X}_h}^2 + \frac{1}{2} \|(q_h, q_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}^2.$$

By using the definitions (3.27) and (3.4) we obtain that

$$\mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; \widehat{G}_h, 0, 0) = [G_h, \widehat{G}_h]_{\widehat{X}_h} - \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| (\widehat{G}_h)_E^e (q_E - q_e)$$

$$= [G_h, \widehat{G}_h]_{\widehat{X}_h} + \sum_{E \in \Omega_h} \sum_{e \in \partial E} \frac{|e|}{d_{E,e}} (q_E - q_e)^2$$

$$\geqslant -\frac{\widehat{C}}{2} \|G_h\|_{\widehat{X}_h}^2 + \frac{1}{2} \|(q_h, q_{\mathcal{E}_h})\|_{1, \mathcal{D}_h, \mathcal{E}_h}^2. \tag{3.34}$$

In the following development it is natural to use the H^1-like norm for the elements of Q_h given by

$$\|q_h\|_{1,h}^2 = \sum_{e \in \mathcal{E}_h} |e| h_e^{-1} ([\![q_h]\!]_e)^2, \tag{3.35}$$

where $[\![q_h]\!]_e$ is the jump of q_h at the edge e that is defined according to (2.31). Assumptions (HG) and (ME) imply that the mesh-dependent norm $\|\cdot\|_{1,h}$ in (3.35) is *uniformly* equivalent to the norm $\|\cdot\|_{1,\mathcal{D}_h}$ in (3.3), that is, there exists two positive constants ν_* and ν^* that are independent of the mesh size h such that we have

$$\nu_* \|\cdot\|_{1,\mathcal{D}_h} \leqslant \|\cdot\|_{1,h} \leqslant \nu^* \|\cdot\|_{1,\mathcal{D}_h} \tag{3.36}$$

for every instance of the admissible mesh family $(\mathcal{D}_h)_h$. As in Beirão da Veiga *et al.* (2009b), let us consider the *conservative* vector field $\widetilde{G}_h \in X_h$ given by

$$\forall E \in \Omega_h, \forall e \in \partial E : (\widetilde{G}_h)_E^e = h_e^{-1} (q_{E'} - q_E).$$

Since (M2) implies that $|E| \lesssim |e| h_e$, we have that

$$\|\widetilde{G}_h\|_{\widehat{X}_h}^2 = \sum_{E \in \Omega_h} \|\widetilde{G}_E\|_E^2 \leqslant \sigma^* \sum_{E \in \Omega_h} \sum_{e \in \partial E} |E| h_e^{-2} (q_{E'} - q_E)^2 \leqslant \widetilde{C} \|q_h\|_{1,h}^2, \tag{3.37}$$

where $\widetilde{C}$ is independent of h and only depends on σ^* and the mesh regularity constants of (M2). We then apply the Cauchy–Schwarz inequality and Young's inequality to obtain

$$|[G_h, \widetilde{G}_h]_{\widehat{X}_h}| \leqslant \|G_h\|_{\widehat{X}_h}\|\widetilde{G}_h\|_{\widehat{X}_h} \leqslant \frac{\widetilde{C}}{2}\|G_h\|^2_{\widehat{X}_h} + \frac{1}{2\widetilde{C}}\|\widetilde{G}_h\|^2_{\widehat{X}_h} \leqslant \frac{\widetilde{C}}{2}\|G_h\|^2_{\widehat{X}_h} + \frac{1}{2}\|q_h\|^2_{1,h}.$$

By using the definition of $\widetilde{G}_h$ and the norm definition (3.4), we obtain that

$$\sum_{E \in \Omega_h} \sum_{e \in \partial E} |e|(\widetilde{G}_h)^e_E(q_E - q_e) = \sum_{e \in \mathcal{E}_h} |e|((\widetilde{G}_h)^e_E q_E + (\widetilde{G}_h)^e_{E'}q_{E'}) + \sum_{e \in \mathcal{E}_h} |e|((\widetilde{G}_h)^e_E + (\widetilde{G}_h)^e_{E'})q_e$$

$$= -\sum_{e \in \mathcal{E}_h} |e|h_e^{-1}(q_{E'} - q_E)^2 = -\|q_h\|^2_{1,h}.$$

Therefore we have that

$$\mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; \widetilde{G}_h, 0, 0) = [G_h, \widetilde{G}_h]_{\widehat{X}_h} + \|q_h\|^2_{1,h} \geqslant -\frac{\widetilde{C}}{2}\|G_h\|^2_{\widehat{X}_h} + \frac{1}{2}\|q_h\|^2_{1,h}. \tag{3.38}$$

Let $G'_h = \theta G_h + \widehat{G}_h + \widetilde{G}_h$ for some value of θ, $q'_h = q_h$ and $q'_{\mathcal{E}_h} = q_{\mathcal{E}_h}$. From (3.32), (3.34) and (3.38) the following holds:

$$\mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; G'_h, q'_h, q'_{\mathcal{E}_h})$$

$$= \theta\mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; G_h, q_h, q_{\mathcal{E}_h}) + \mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; \widehat{G}_h, 0, 0) + \mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; \widetilde{G}_h, 0, 0)$$

$$\geqslant \left(\theta - \frac{\widehat{C}}{2} - \frac{\widetilde{C}}{2}\right)\|G_h\|^2_{\widehat{X}_h} + \frac{1}{2}\|q_h\|^2_{1,h} + \left(\frac{1}{2} - \theta C_1 h\right)\|(q_h, q_{\mathcal{E}_h})\|^2_{1,\mathcal{D}_h,\mathcal{E}_h}. \tag{3.39}$$

Now, we take $\theta = (1 + \widehat{C} + \widetilde{C})/2$ and we obtain the inequality

$$\|G_h\|^2_{\widehat{X}_h} + \|q_h\|^2_{1,h} + \|(q_h, q_{\mathcal{E}_h})\|^2_{1,\mathcal{D}_h,\mathcal{E}_h} \lesssim \mathcal{B}(G_h, q_h, q_{\mathcal{E}_h}; G'_h, q'_h, q'_{\mathcal{E}_h}), \tag{3.40}$$

which holds for h small enough under assumption (AB3-w) and for any h under assumption (AB3-s) because $C_1 = 0$ in this case. Using inequalities (3.33) and (3.37) allows us to obtain

$$\|G'_h\|_{\widehat{X}_h} + \|q'_h\|_{1,h} + \|(q'_h, q'_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h} \leqslant \theta\|G_h\|_{\widehat{X}_h} + (1 + \sqrt{\widehat{C}})\|q_h\|_{1,h}$$

$$+ (1 + \sqrt{\widetilde{C}})\|(q_h, q_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h}. \tag{3.41}$$

The inequalities (3.29) and (3.30) in Lemma 3.9 follow from (3.40) and (3.41) by rescaling the three discrete fields G'_h, q'_h and $q'_{\mathcal{E}_h}$ by the positive factor $\max(\theta, 1 + \sqrt{\widehat{C}}, 1 + \sqrt{\widetilde{C}})(\|G_h\|_{\widehat{X}_h} + \|q_h\|_{1,h} + \|(q_h, q_{\mathcal{E}_h})\|_{1,\mathcal{D}_h,\mathcal{E}_h})$. □

The following technical lemma provides us with an estimate for the interpolation of a vector field that is locally in $(H^1(E))^d$.

LEMMA 3.10 Let $G \in (H^1(E))^d$ and let G^I be the interpolated field (2.8). Then we have that

$$\|G^I\|_E \lesssim \|G\|_{L^2(E)} + h_E|G|_{H^1(E)}. \tag{3.42}$$

Proof. Using the stability condition of assumption (S1), the Agmon inequality from (M3), and the scaling $|E|/|e| \lesssim h_E$, which is a consequence of (M2), it readily follows that

$$\|(G^I)\|_E^2 \lesssim |E| \sum_{e \in \partial E} (G_E^e)^2 \lesssim |E| \sum_{e \in \partial E} |e|^{-1} \|G\|_{L^2(e)}^2 \lesssim h_E (h_E^{-1} \|G\|_{L^2(E)}^2 + h_E |G|_{H^1(E)}^2)$$

$$\lesssim \|G\|_{L^2(E)}^2 + h_E^2 |G|_{H^1(E)}^2,$$

from which the lemma's statement immediately follows.   □

We can now prove the main result of this subsection, which is stated in the following theorem. This theorem provides a bound on the approximation error that is defined by comparing the numerical solution $(p_h, F_h, p_{\mathcal{E}_h}) \in Q_h \times \widehat{X}_h \times H(\mathcal{E}_h)$ with the interpolations p^I and F^I of the exact solution and flux given by (2.8), and with the interpolated field $p^J = \{(p^J)^e\}^{e \in \mathcal{E}_h} \in H(\mathcal{E}_h)$ given by

$$\forall\, e \in \mathcal{E}_h : (p^J)^e = \frac{1}{|e|} \int_e p. \tag{3.43}$$

THEOREM 3.11 Let p be the solution of the continuous problem (2.1) and (2.2) under assumptions (H1)–(H4) with Λ locally Lipschitz continuous on Ω_h (cf. Remark 2.5) and F given by (2.3). Let (F_h, p_h) be the solution of problem (2.41) and (2.42) under assumptions (HG), (ME) and (AB1)–(AB3) with either h small enough if (AB3-w) holds or any h if (AB3-s) holds. Then we have that

$$\|F_h - F^I\|_{\widehat{X}_h} + \|p_h - p^I\|_{1, \mathcal{D}_h} + \|(p_h - p^I, p_{\mathcal{E}_h} - p^J)\|_{1, \mathcal{D}_h, \mathcal{E}_h} \lesssim h \|p\|_{H^2(\Omega)}. \tag{3.44}$$

Proof. Let us consider the triplet of error fields $(F_h - F^I, p_h - p^I, p_{\mathcal{E}_h} - p^J) \in \widehat{X}_h \times Q_h \times H(\mathcal{E}_h)$. Due to Lemma 3.9, there exists a triplet $(G_h, q_h, q_{\mathcal{E}_h}) \in \widehat{X}_h \times Q_h \times H(\mathcal{E}_h)$ with

$$\|G_h\|_{\widehat{X}_h} + \|q_h\|_{1, \mathcal{D}_h} + \|q_{\mathcal{E}_h}\|_{1, \mathcal{D}_h, \mathcal{E}_h} \leqslant 1 \tag{3.45}$$

such that

$$\|F_h - F^I\|_{\widehat{X}_h} + \|p_h - p^I\|_{1, \mathcal{D}_h} + \|(p_h - p^I, p_{\mathcal{E}_h} - p^J)\|_{1, \mathcal{D}_h, \mathcal{E}_h}$$

$$\lesssim \mathcal{B}(F_h - F^I, p_h - p^I, p_{\mathcal{E}_h} - p^J; G_h, q_h, q_{\mathcal{E}_h}). \tag{3.46}$$

By using equations (3.28), a straightforward calculation gives

$$\mathcal{B}(F_h - F^I, p_h - p^I, p_{\mathcal{E}_h} - p^J; G_h, q_h, q_{\mathcal{E}_h}) = T_1 + T_2 + T_3, \tag{3.47}$$

where

$$T_1 = [p^I, \mathrm{div}_h(G_h)]_{Q_h} - [F^I, G_h]_{\widehat{X}_h} - \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e|(p^J)^e G_E^e,$$

$$T_2 = [f^I, q_h]_{Q_h} - [\mathrm{div}_h(F^I) + \mathrm{div}_h(F_c(p^I)), q_h]_{Q_h},$$

$$T_3 = \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e|(F^I + F_c(p^I))_E^e q_e.$$

For convenience, we will separately bound T_1 and $T_2 + T_3$. For this purpose, let us first introduce the discontinuous Ω_h-piecewise linear function p^1, which is such that $p^1|_E$ is the L^2 orthogonal projection of p on the linear polynomials defined on $E \in \Omega_h$. Let us start by noting that $\|p - p^1\|_{L^2(E)} \leqslant \|p - q^1\|_{L^2(E)}$ for any linear polynomial q^1 defined in E. Hence, taking $q^1 = \mathcal{L}_1(p)$, the linear interpolation of p on E provided by (M4), allows us to use the estimate for the interpolation error. Moreover, adding and subtracting $\mathcal{L}_1(p)$, applying the triangular inequality, using (M4) and a standard inverse inequality yields that

$$|p - p^1|_{H^1(E)} \leqslant |p - \mathcal{L}_1(p)|_{H^1(E)} + |\mathcal{L}_1(p) - p^1|_{H^1(E)} \lesssim h_E |p|_{H^2(E)} + h_E^{-1} \|\mathcal{L}_1(p) - p^1\|_{L^2(E)}. \quad (3.48)$$

The second term in the last inequality of (3.48) is obtained by adding and subtracting p, applying the triangular inequality and using the estimate for the interpolation error of (M4) as follows

$$\|\mathcal{L}_1(p) - p^1\|_{L^2(E)} \leqslant \|\mathcal{L}_1(p) - p\|_{L^2(E)} + \|p - p^1\|_{L^2(E)} \lesssim h_E^2 |p|_{H^2(E)}. \quad (3.49)$$

Substituting (3.49) into (3.48) yields the final inequality

$$\|p - p^1\|_{L^2(E)} + h_E |p - p^1|_{H^1(E)} \lesssim h_E^2 \|p\|_{H^2(E)}. \quad (3.50)$$

We recall that, for convenience, we may identify the elements of Q_h with the piecewise constant functions whose restriction to each cell E is the degree of freedom of that cell. Therefore it is possible to reformulate the first term of T_1 as an L^2 scalar product, so that $[p^I, \mathrm{div}_h(G_h)]_{Q_h} = [p, \mathrm{div}_h(G_h)]_{L^2}$. Then we split T_1 into four subterms by recalling that $F = -\Lambda \nabla p$, adding and subtracting $(\Lambda_E \nabla p)^I$ and $(\Lambda_E \nabla p^1)^I$, using the local consistency assumption (S2), and noting that $|e|(p^J)^e = \int_e p$. We therefore obtain the following:

$$T_1 = [p - p^1, \mathrm{div}_h(G_h)]_{L^2} + [p^1, \mathrm{div}_h(G_h)]_{L^2} - [F^I, G_h]_{\widehat{X}_h} - \sum_{E \in \Omega_h} \sum_{e \in \partial E} G_E^e \int_e p$$

$$= T_{1,1} + [p^1, \mathrm{div}_h(G_h)]_{L^2} + [(\Lambda \nabla p)^I, G_h]_{\widehat{X}_h} - \sum_{E \in \Omega_h} \sum_{e \in \partial E} G_E^e \int_e p$$

$$= T_{1,1} + [p^1, \mathrm{div}_h(G_h)]_{L^2} + \sum_{E \in \Omega_h} [(\Lambda_E \nabla p)^I, G_h]_E$$

$$+ \sum_{E \in \Omega_h} [((\Lambda - \Lambda_E) \nabla p)^I, G_h]_E - \sum_{E \in \Omega_h} \sum_{e \in \partial E} G_E^e \int_e p$$

$$= T_{1,1} + \sum_{E \in \Omega_h} ([\mathrm{div}_h(G_h), p^1]_{L^2(E)} + [(\Lambda_E \nabla p^1)^I, G_h]_E) + \sum_{E \in \Omega_h} [(\Lambda_E \nabla (p - p^1))^I, G_h]_E$$

$$+ \sum_{E \in \Omega_h} [((\Lambda - \Lambda_E) \nabla p)^I, G_h]_E - \sum_{E \in \Omega_h} \sum_{e \in \partial E} G_E^e \int_e p$$

$$= T_{1,1} + \sum_{E \in \Omega_h} \sum_{e \in \partial E} G_E^e \int_e (p^1 - p) + \sum_{E \in \Omega_h} [(\Lambda_E \nabla (p - p^1))^I, G_h]_E$$

$$+ \sum_{E \in \Omega_h} [((\Lambda - \Lambda_E) \nabla p)^I, G_h]_E = T_{1,1} + T_{1,2} + T_{1,3} + T_{1,4}.$$

To estimate $T_{1,1}$ let us first note that the definition of div_h, the scalings $h_E^d \leqslant |E| \lesssim h_E^d$ and $|e| \lesssim h_E^{d-1}$ from (M2) and assumption (S1) imply that

$$\|\mathrm{div}_h(G_h)\|_{L^2(E)}^2 = |E||(\mathrm{div}_h(G_h))_E|^2 \lesssim \frac{1}{|E|} \sum_{e \in \partial E} |e|^2 (G_E^e)^2 \lesssim h_E^{-2} \|G_h\|_E^2 . \qquad (3.51)$$

Thus, using the Cauchy–Schwarz inequality for each scalar product in X_E, error estimate (3.50), inequality (3.51), the Cauchy–Schwarz inequality again and finally noting that (3.45) implies that $\|G_h\|_{\widehat{X}_h} \leqslant 1$, yields that

$$T_{1,1} \lesssim \sum_{E \in \Omega_h} \|p - p^1\|_{L^2(E)} \|\mathrm{div}_h(G_h)\|_{L^2(E)} \lesssim \sum_{E \in \Omega_h} (h_E^2 |p|_{H^2(E)})(h_E^{-1} \|G_h\|_E)$$

$$\lesssim h \left(\sum_{E \in \Omega_h} |p|_{H^2(E)}^2 \right)^{1/2} \left(\sum_{E \in \Omega_h} \|G_h\|_E^2 \right)^{1/2} \lesssim h|p|_{H^2(\Omega)}. \qquad (3.52)$$

The second term is bounded using a scaling argument and inequality (3.45). We obtain that

$$T_{1,2} \lesssim h\|p\|_{H^2(\Omega)} \|G_h\|_{\widehat{X}_h} \lesssim h\|p\|_{H^2(\Omega)}. \qquad (3.53)$$

To get an upper bound for $T_{1,3}$ we use the Cauchy–Schwarz inequality for the local scalar product in X_E, the result of Lemma 3.10, an upper bound on Λ_E that easily follows from the upper bound of Λ in (H2), the Cauchy–Schwarz inequality again, the estimate of the interpolation error given by (3.50) that follows from (M4) and the fact that $\|G_h\|_{\widehat{X}_h} \leqslant 1$ due to inequality (3.45). We obtain the following chain of inequalities:

$$T_{1,3} = \sum_{E \in \Omega_h} [(\Lambda_E \nabla(p - p^1))^I, G_h]_E$$

$$\leqslant \sum_{E \in \Omega_h} \|(\Lambda_E \nabla(p - p^1))^I\|_E \|G_h\|_E$$

$$\lesssim \sum_{E \in \Omega_h} (\|\Lambda_E \nabla(p - p^1)\|_{L^2(E)} + h_E |\Lambda_E \nabla(p - p^1)|_{H^1(E)}) \|G_h\|_E$$

$$\lesssim \sum_{E \in \Omega_h} (|p - p^1|_{H^1(E)} + h_E |p|_{H^2(E)}) \|G_h\|_E$$

$$\lesssim \left(\sum_{E \in \Omega_h} |p - p^1|_{H^1(E)}^2 + h_E^2 |p|_{H^2(E)}^2 \right)^{1/2} \left(\sum_{E \in \Omega_h} \|G_h\|_E^2 \right)^{1/2}$$

$$\lesssim \left(\sum_{E \in \Omega_h} h_E^2 |p|_{H^2(E)}^2 \right)^{1/2} \|G_h\|_{\widehat{X}_h}$$

$$\lesssim h|p|_{H^2(\Omega)}. \qquad (3.54)$$

Using the Cauchy–Schwarz inequality and inequality (3.42), we get

$$T_{1,4} \lesssim \sum_{E \in \Omega_h} \|((\Lambda - \Lambda_E)\nabla p)^I\|_E \|G_h\|_E$$

$$\lesssim \left(\sum_{E \in \Omega_h} \|((\Lambda - \Lambda_E)\nabla p)^I\|_E^2 \right)^{1/2}$$

$$\lesssim \left(\sum_{E \in \Omega_h} \|(\Lambda - \Lambda_E)\nabla p\|_{L^2(E)}^2 + h_E^2 |(\Lambda - \Lambda_E)\nabla p|_{H^1(E)}^2 \right)^{1/2}. \tag{3.55}$$

Due to the definition of Λ_E and since the restriction $\Lambda|_E$ belongs to $W^{1,\infty}(E)$ for all $E \in \Omega_h$, we obtain that

$$\|\Lambda - \Lambda_E\|_{L^\infty(E)} + h_E |\Lambda - \Lambda_E|_{W^{1,\infty}(E)} \lesssim h_E \quad \forall\, E \in \Omega_h.$$

Combining the above bound with (3.55) easily yields

$$T_{1,4} \lesssim h \|p\|_{H^2(\Omega)}. \tag{3.56}$$

Combining (3.52)–(3.54) and (3.56) yields the following upper bound for T_1:

$$T_1 \lesssim h \|p\|_{H^2(\Omega)}. \tag{3.57}$$

To get an upper bound for $T_2 + T_3$ we note that the commuting property of the divergence operator (2.9) (cf. also Remark 2.4), the flux definition given in (2.3), and the model's equation (2.1) allow us to write

$$\text{div}_h(F^I) = (\text{div}(F))^I = f^I - (\text{div}(Vp))^I. \tag{3.58}$$

Equation (3.58) makes it possible to reformulate T_2 as follows:

$$T_2 = [(\text{div}(Vp))^I - \text{div}_h(F_c(p^I)), q_h]_{Q_h}. \tag{3.59}$$

As before, we identify the elements of Q_h with the space of Ω_h-piecewise constant functions, and the scalar product in Q_h with the L^2 scalar product. Then we split T_2 into two subterms by applying the divergence theorem to each cell's contribution and adding and subtracting the term $V_E^e p$, to give

$$T_2 = \sum_{E \in \Omega_h} \sum_{e \in \partial E} q_E \int_e (V \cdot \mathbf{n}_E^e - V_E^e)p + \sum_{E \in \Omega_h} \sum_{e \in \partial E} q_E \int_e (V_E^e p - (F_c(p^I))_E^e) = T_{2,1} + T_{2,2}. \tag{3.60}$$

Noting that $V_E^e + V_{E'}^e = 0$ and $\mathbf{n}_E^e + \mathbf{n}_{E'}^e = 0$ for any $e \in \mathcal{E}_{h,\text{int}}$, and using definition (2.31) for the jump of q_h, that is, $[\![q_h]\!]_e$, allows us to reformulate $T_{2,1}$ as follows:

$$T_{2,1} = \sum_{e \in \mathcal{E}_h} [\![q_h]\!]_e \int_e (V \cdot \mathbf{n}_E^e - V_E^e)p,$$

where $E = E(e)$ is the unique cell for which e belongs to the boundary and such that $\mathbf{n}_E^e \cdot \mathbf{n}^e = 1$. Due to the definition of V_E^e, on each edge e the quantity $(V \cdot \mathbf{n}_E^e - V_E^e)$ is orthogonal to constants. Therefore we can write

$$T_{2,1} = \sum_{e \in \mathcal{E}_h} [\![q_h]\!]_e \int_e (V \cdot \mathbf{n}_E^e - V_E^e)(p - \overline{p}_e), \tag{3.61}$$

where $\overline{p}_e$ is the average of p on e. Applying the Hölder inequality to each face's term, the interpolation estimates for the face's velocity, the Cauchy–Schwarz inequality and the equivalence between norms $\|\cdot\|_{1,h}$ and $\|\cdot\|_{1,\mathcal{D}_h}$ gives

$$T_{2,1} \lesssim \sum_{e \in \mathcal{E}_h} |[\![q_h]\!]_e| \, h_e^{\frac{d-1}{2}} \, \|(V \cdot \mathbf{n}_E^e - V_E^e)(p - \overline{p}_e)\|_{L^2(e)}$$

$$\lesssim \sum_{e \in \mathcal{E}_h} h_e^{\frac{d-2}{2}} |[\![q_h]\!]_e| \, h_e^{\frac{3}{2}} |V|_{W^{1,\infty}(\Omega)} \|p - \overline{p}_e\|_{L^2(e)}$$

$$\lesssim |V|_{W^{1,\infty}(\Omega)} \left(\sum_{e \in \mathcal{E}_h} h_e^{d-2} |[\![q_h]\!]_e|^2 \right)^{1/2} \left(\sum_{e \in \mathcal{E}_h} h_e^3 \|p - \overline{p}_e\|_{L^2(e)}^2 \right)^{1/2}$$

$$\lesssim |V|_{W^{1,\infty}(\Omega)} \|q_h\|_{1,\mathcal{D}_h} \left(\sum_{e \in \mathcal{E}_h} h_e^5 \|\nabla p\|_{L^2(e)}^2 \right)^{1/2}, \tag{3.62}$$

where in the last line we have also used a standard approximation result. Now, the Agmon inequality for ∇p (cf. (M3) with $\phi = \nabla p$) and the fact that for $e \subseteq \partial E \cap \partial E'$ we have that $h_e \leqslant \max(h_E, h_{E'})$ imply that

$$\sum_{e \in \mathcal{E}_h} h_e^5 \|\nabla p\|_{L^2(e)}^2 \lesssim \sum_{E \in \Omega_h} \sum_{e \in \partial E} h_E^5 \|\nabla p\|_{L^2(e)}^2 \lesssim \sum_{E \in \Omega_h} h_E^5 (h_E^{-1} |p|_{H^1(E)}^2 + h_E |p|_{H^2(E)}^2).$$

The bound for $T_{2,1}$ readily follows by recalling (3.45), to give

$$T_{2,1} \lesssim h^2 \|p\|_{H^1(\Omega)} + h^3 \|p\|_{H^2(\Omega)} \lesssim h^2 \|p\|_{H^2(\Omega)}, \tag{3.63}$$

where we have included the data factor $|V|_{W^{1,\infty}(\Omega)}$ in the inequality's constant. As a byproduct, we observe here that, from (3.63) and the estimate (3.68), it becomes clear that the error coming from the approximation of the datum V is a higher-order term.

Now, let us search for an upper bound for $T_{2,2} + T_3$. First, note that, since V^I and F^I are conservative fields, we have

$$\sum_{E \in \Omega_h} \sum_{e \in \partial E} q_e \int_e V_E^e p = \sum_{e \in \mathcal{E}_h} q_e (V_E^e + V_{E'}^e) \int_e p = 0 \quad \text{and} \quad \sum_{E \in \Omega_h} \sum_{e \in \partial E} |e| (F^I)_E^e q_e = 0,$$

and thus

$$T_{2,2} + T_3 = \sum_{E \in \Omega_h} \sum_{e \in \partial E} (q_E - q_e) \int_e (V_E^e p - (F_c(p^I))_E^e). \tag{3.64}$$

Moreover, assumption (A2) implies that $V_E^e = \frac{1}{d_e}(A(d_e V_E^e) + B(d_e V_E^e))$ and therefore, by using definition (2.40) and the triangle inequality, a straightforward calculation gives

$$\|V_E^e p - (F_c(p^I))_E^e\|_{L^2(e)}^2 \leqslant \left\|(p - (p^I)_E)\frac{A(d_e V_E^e)}{d_e}\right\|_{L^2(e)}^2 + \left\|(p - (p^I)_{E'})\frac{B(d_e V_E^e)}{d_e}\right\|_{L^2(e)}^2. \quad (3.65)$$

From (A1) and the definition of V_E^e it easily follows that $\max(|A(d_e V_E^e)|, |B(d_e V_E^e)|) \lesssim d_e$. Then, by using the Agmon inequality of (M3) and the standard first-order interpolation estimate for cell averages, that is, $\|p - (p^I)_E\|_{L^2(E)} \lesssim h_E|p|_{H^1(E)}$, we have that

$$\left\|(p - (p^I)_E)\frac{A(d_e V_E^e)}{d_e}\right\|_{L^2(e)}^2 \lesssim \|p - (p^I)_E\|_{L^2(e)}^2$$

$$\lesssim h_E^{-1}\|p - (p^I)_E\|_{L^2(E)}^2 + h_E|p - (p^I)_E|_{H^1(E)}^2$$

$$\lesssim h_E|p|_{H^1(E)}^2. \quad (3.66)$$

A similar inequality can be derived by repeating the same argument for the second term on the right-hand side of (3.65) when $e \in \mathcal{E}_{h,\mathrm{int}}$ and noting that the second term is zero if e is a boundary face. Finally, we obtain

$$\|V_E^e p - (F_c(p^I))_E^e\|_{L^2(e)}^2 \lesssim h|p|_{H^1(E\cup E')}^2.$$

Therefore, by using a Hölder inequality on the faces and an l^2 Cauchy–Schwarz inequality, from (3.64) we obtain

$$T_{2,2} + T_3 \lesssim h^{\frac{1}{2}} \sum_{e \in \mathcal{E}_h} |q_E - q_e||e|^{\frac{1}{2}}|p|_{H^1(E\cup E')} \lesssim h^{\frac{1}{2}} \|q_h\|_{1,\mathcal{D}_h,\mathcal{E}_h} \left(\sum_{e \in \mathcal{E}_h} h_e|p|_{H^1(E\cup E')}^2\right)^{1/2}. \quad (3.67)$$

Recalling (3.45) yields

$$T_{2,2} + T_3 \lesssim h\|p\|_{H^1(\Omega)}. \quad (3.68)$$

Combining (3.63) and (3.68), we have the bound for $T_2 + T_3$, and also considering (3.46), (3.47) and (3.57), we conclude the proof. □

From Theorem 3.11 we immediately get two corollaries that we state without proof.

COROLLARY 3.12 Under the same hypotheses as in Theorem 3.11, the following holds:

$$\|\widetilde{F}_h - \widetilde{F}^I\|_{\widehat{X}_h} \lesssim h\|p\|_{H^2(\Omega)}, \quad (3.69)$$

where the total fluxes are defined through $\widetilde{F^I} = -(\Lambda\nabla p + Vp)^I$ and $\widetilde{F}_h = F_h + F_c(p_h)$.

COROLLARY 3.13 Under the same hypotheses as in Theorem 3.11 (and applying Lemma 3.5) the following holds:

$$\|p^I - p_h\|_{L^r(\Omega)} \lesssim h\|p\|_{H^2(\Omega)},$$

where $r = \frac{2d}{d-2}$ if $d > 2$ and $r < +\infty$ if $d = 2$.

REMARK 3.14 Repeating the same arguments makes it possible to prove a similar error estimate for the hybrid HMM formulation (2.47)–(2.49) that is based on the numerical convection flux (2.45) and (2.46).

REMARK 3.15 It must be noted that the proofs in this paper are not uniform with respect to the Péclet number, that is, the estimates degenerate when the convection becomes dominant. On the other hand, uniform estimates cannot be derived under the general framework considered here since it also includes methods that are not stable in the limit. Nevertheless, the general approach used here can be followed in order to develop uniform error estimates for certain methods. For example, we believe that a uniform error bound can be developed for the upwind scheme starting from a uniform version of the stability results in Proposition 3.2 and Corollary 3.3. A deeper theoretical investigation of the convection-dominated case will be the objective of future communications.

4. Numerical experiments

In this section we present a number of examples of problem (2.1) and (2.2), whose solutions are computed over uniform and nonuniform meshes. The performance of these discretization methods is investigated by evaluating the rate of convergence when the meshes are refined and the shock-capturing capability when strong layers develop in the convection-dominated regime.

For this purpose, we consider the sequence of meshes of the mesh family $\mathcal{M}1$ on $\Omega =]0, 1[\times]0, 1[$. These meshes are built by remapping the position (ξ, η) of the nodes of an $n \times n$ uniform grid of quadrilaterals into final positions (x, y) through

$$x = \xi + \frac{1}{10} \sin(2\pi \xi) \sin(2\pi \eta), \tag{4.1}$$

$$y = \eta + \frac{1}{10} \sin(2\pi \xi) \sin(2\pi \eta). \tag{4.2}$$

Then we split each quadrilateral-shaped cell into two triangles, which gives *the primal mesh*, and then we connect the barycentres of adjacent triangular cells by a straight segment. We complete the mesh construction at the domain boundary $\partial \Omega$ by connecting the barycentres of triangular cells close to $\partial \Omega$ to the midpoints of boundary edges and connecting these latter points to the boundary vertices of the primal mesh. For this mesh family, the base mesh of the refinement process is obtained by setting $n = 10$. Refined meshes are generated by doubling this parameter and repeating the construction procedure. The plots of Fig. 2 illustrate the base mesh and the first refined mesh of $\mathcal{M}1$. Details about the mesh characteristics are reported in Table 1.

The numerical implementation is partially based on P2MESH (Bertolazzi & Manzini, 2002), a $C++$ public domain library designed to manage data structures of unstructured meshes in the implementation of solvers of PDEs. For convenience, we will use the labels listed below to refer to the different instances of the HMM family of schemes considered in our numerical experiments. In each one of these schemes the diffusion term is discretized along the lines described in Section 2.3, while the numerical treatment of the convection term differs as specified in the following descriptions of the schemes:

- HMM-Cnt, two-point centred flux formula;

- HMM-Upw, two-point upwind flux formula;

- HMM-SG, two-point Scharfetter–Gummel formula with local adjustment (2.23);

- HMM-(no stabilization), central mimetic method without any form of stabilization;

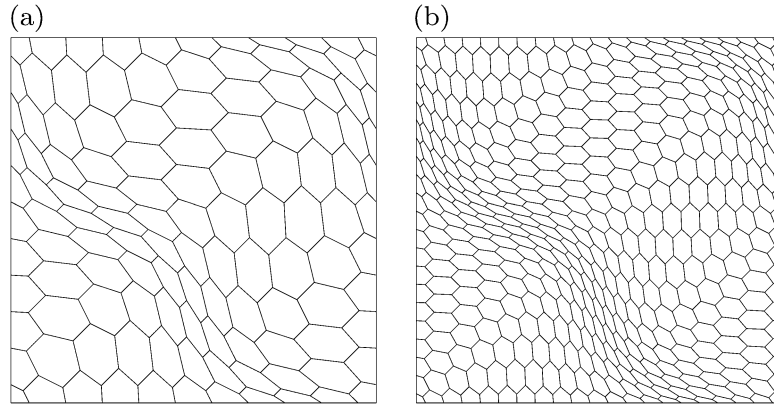- HMM-Jmp, central mimetic method with jump stabilization (2.32).

FIG. 2. (a) The base mesh and (b) its first refinement for the mesh family $\mathcal{M}1$. The mesh construction parameter n is initially taken equal to 10 and doubled at each refinement step.

TABLE 1 *Mesh parameters of the mesh sequence $\mathcal{M}1$*

r	N_E	N_e	N_V	h
0	121	400	280	9.477×10^{-2}
1	441	1400	960	4.843×10^{-2}
2	1681	5200	3520	2.445×10^{-2}
3	6561	20000	13440	1.225×10^{-2}
4	25921	78400	52480	6.130×10^{-3}
5	103041	310400	207360	3.066×10^{-3}

Here r is the refinement level (0 refers to the base mesh), N_E is the number of cells, N_e is the number of mesh edges and N_V is the number of mesh vertices.

4.1 *Accuracy*

In this test case the forcing term f in (2.1) and the boundary condition function g^D in (2.2) are set according to the exact solution

$$p(x, y) = \left(x - e^{\frac{2(x-1)}{\nu}} \right) \left(y^2 - e^{\frac{3(y-1)}{\nu}} \right) \tag{4.3}$$

and $V = (2, 3)^T$. We assume that the diffusion tensor Λ is given by the identity matrix scaled by the positive real factor ν. By taking $\nu = 10^{-4}$, the problem is strongly convection dominated and the solution is characterized by an exponential boundary layer near the top and right sides of Ω.

We are mainly interested in showing that the shock-capturing capability does not significantly deteriorate the convergence behaviour where the solution is smooth enough, that is, away from the boundaries where the layer develops. As pointed out in Bertolazzi & Manzini (2004), Rapin & Lube (2004) (to which we also refer the reader interested in the comparison with performance of the mixed hybrid finite element and different kind of FV schemes on this test case) Manzini & Russo (2008) and Coudière & Manzini (2010), the errors due to the approximation of the solution gradient in the narrow strip around the boundary where the layer develops are so large that including them in the error measurements would prevent any convergence at all. For this reason, we restrict the error

measurement to the subdomain $[0, 0.95] \times [0, 0.95]$. Convergence rates are measured by the relative errors

$$\mathcal{E}_{Q_h} = \frac{\|p^I - p_h\|_{Q_h}}{\|p^I\|_{Q_h}} \quad \text{and} \quad \mathcal{E}_{X_h} = \frac{\|\widetilde{F^I} - \widetilde{F_h}\|_{\widehat{X}_h}}{\|\widetilde{F^I}\|_{\widehat{X}_h}}, \tag{4.4}$$

where, in the second error definition, we use the total fluxes $\widetilde{F^I}$ and $\widetilde{F_h}$ that are defined in Corollary 3.12. Practically speaking, the quantity $\mathcal{E}_{Q_h}$ is a measure of the approximation error of cell averages and is calculated by using a mesh-dependent L^2-like norm. In turn, the error $\mathcal{E}_{X_h}$ compares the edge-based flux $\widetilde{F^I}$ with the numerical flux $\widetilde{F_h}$ through the mesh-dependent norm induced in X_h by the mimetic scalar product.

In Fig. 3 we present the log–log plots of the errors $\mathcal{E}_{Q_h}$ (on the left) and $\mathcal{E}_{X_h}$ (on the right) versus the characteristic mesh size h. Herein, we compare the convergence behaviour of the various implementations of the HMM schemes considered in this paper. The actual order of accuracy shown by these methods is reflected by the slopes of the experimental error curves and can be approximately evaluated by comparison with the 'theoretical' slopes represented in the bottom-left corner of each plot (see also the caption's comment). These plots document the *optimal* convergence behaviour of all the numerical approximations in the diffusive regime (see the upper plots). When the problem becomes convection dominated, that is, for the smallest value of the diffusion coefficient, convergence is still provided for both scalar and flux unknowns by all methods except HMM-(no stabilization).
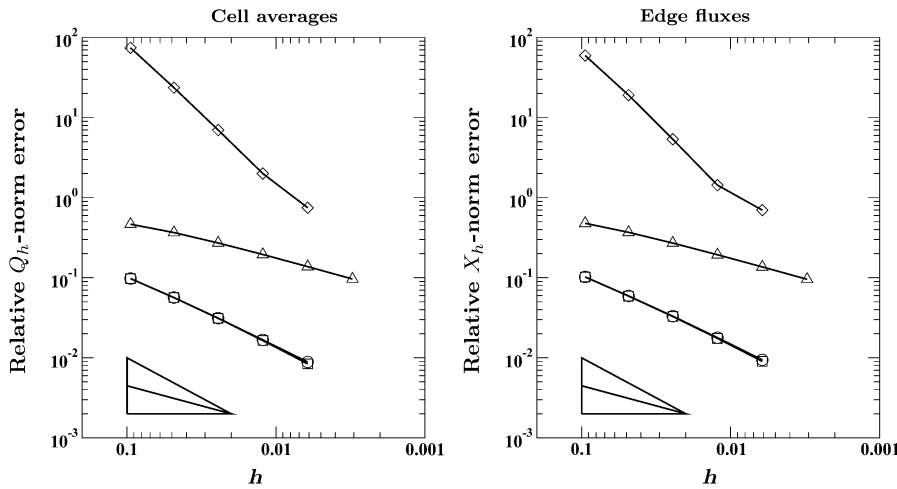
When we use HMM-SG, HMM-Cnt and HMM-(no stabilization) in the diffusive regime a super-convergence effect is visible for the approximation of the scalar and the flux variables. The numerical approximation of the scalar unknown is second-order accurate, while we have $\mathcal{O}(h^{3/2})$ for the flux variable. In contrast, both HMM-Upw and HMM-Jmp provide a first-order accurate approximation for both p and F. It is also worth noting that the error curves of HMM-SG and HMM-Cnt almost coincide. Moreover, the errors from HMM-(no stabilization) are a little bit smaller than those obtained by the centred schemes of FV type. Instead, in this test case the scheme HMM-Upw gives better results than HMM-Jmp.

In the convection-dominated case, that is, for $\nu = 10^{-4}$, the central approximation HMM-(no stabilization) is not at all convergent on the meshes considered by $\mathcal{M}1$. In contrast, HMM-Cnt is convergent, but the numerical solution (not shown in the paper) is affected by large amplitude oscillations that almost completely destroy the solution's profile. This fact is consistent with the error curves displayed in Fig. 3. The numerical approximation of the scalar and flux variables provided by the methods HMM-Upw and HMM-SG is linearly convergent, while the one provided by HMM-Jmp seems to converge at a rate proportional to $\mathcal{O}(h^{1/2})$, even if this last effect might be due to an insufficient mesh resolution.

REMARK 4.1 As noted at the end of Section 2.4.1, in the convection-dominated regime the function A given by (2.23) is numerically nearly indistinguishable from the upwind functions A_{up}. Figures 3 and 4 confirm this. On the other hand, in the diffusion regime the modified Scharfetter–Gummel scheme has better convergence properties than the upwind scheme. This behaviour is a very interesting characteristic of the choice (2.23) when the convection term is discretized by (2.40). The scheme then automatically adjusts to provide either a good order of convergence in the diffusive regime or enough numerical diffusion to stabilize the calculation in the convection-dominated regime. Note that, if one takes A and B to satisfy, (AB1), (AB2) and (AB3-s) and to be such that $A(s) \sim s$ as $s \to +\infty$, $A(s)$ has a finite limit as $s \to -\infty$ and $A(s)$ is regular around $s = 0$, then a scheme using such functions modified in the same way as (2.23) is expected to show the same kind of behaviour (this has been numerically tested on several choices of such functions). Note also that this approach does not hold for the upwind scheme since $A_{\mathrm{up}}(s)$ is *not* regular at $s = 0$.
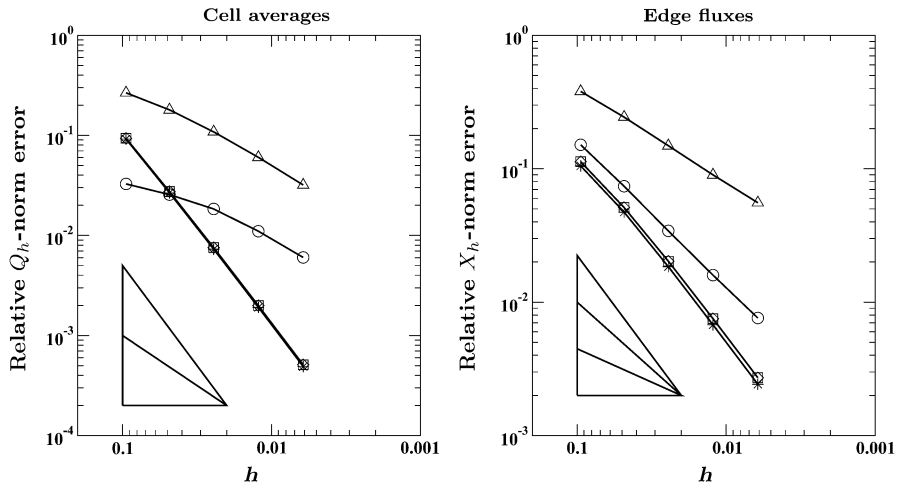
Diffusive regime: $V = (2,3)^T$ and $\nu = 1$; slopes are h^2 and h in the left plot, and $h^{3/2}$, h, $h^{1/2}$ in the right plot.

Convection-dominated regime: $V = (2,3)^T$ and $\nu = 10^{-4}$; slopes are h and $h^{1/2}$ in both plots.

Symbols: ◯ HMM-Upw, □ HMM-SG, ◇ HMM-Cnt, △ HMM-Jmp, ⋆ HMM-(no stabilization).

FIG. 3. Test case 1: error curves for the numerical approximation of an exact solution that is smooth in the diffusive regime (top) and shows an exponential boundary layer on the right and top sides of the computational domain in the convection-dominated regime (bottom). The approximation errors are measured on the reduced domain $[0, 0.95] \times [0, 0.95]$, that is, away from the critical region where the layer may develop. All calculations are performed on the mesh sequence $\mathcal{M}1$.

4.2 *Shock-capturing behaviour*

Shock-capturing behaviour is investigated by solving (2.1) and (2.2) in the convection-dominated regime. The exact solution may be characterized by boundary layers of exponential and parabolic type and is
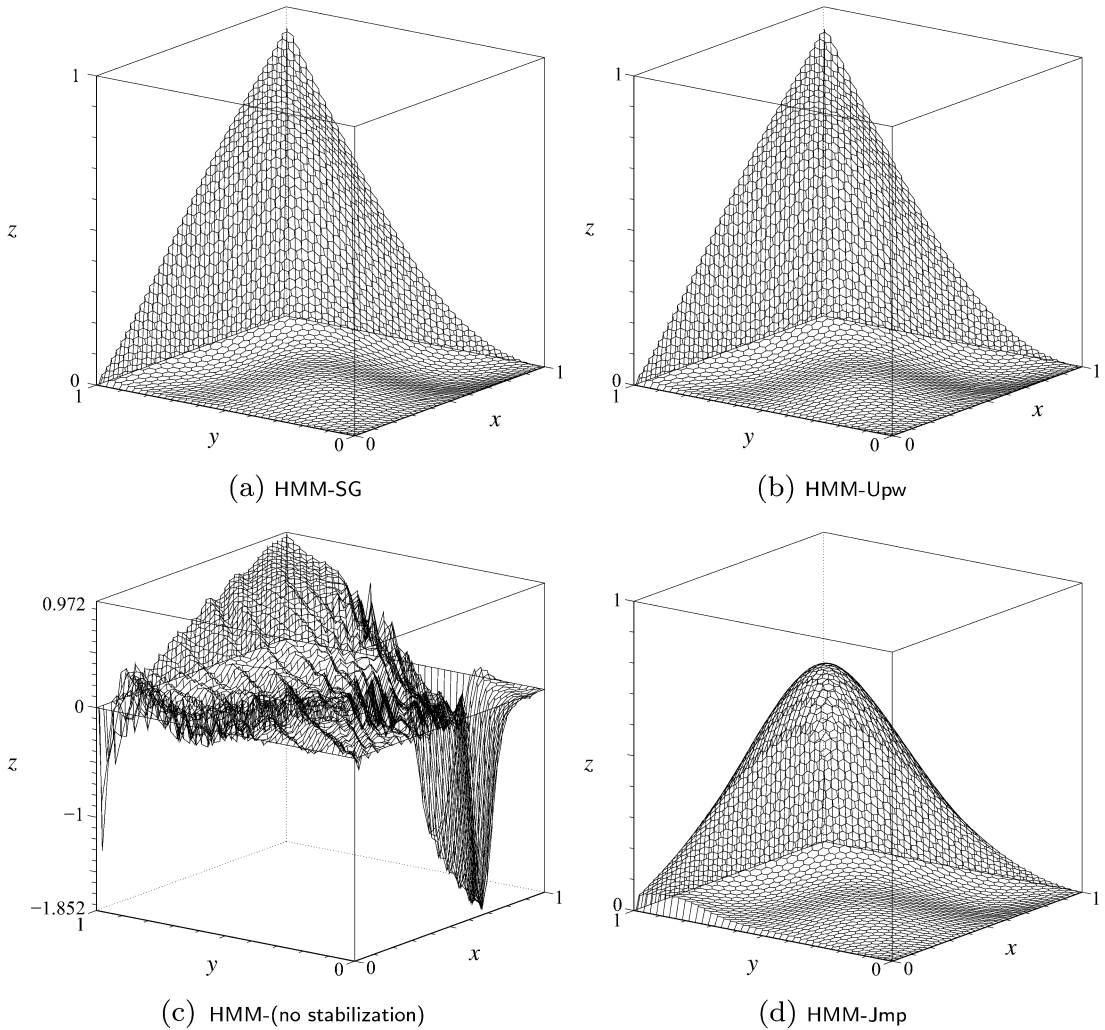
FIG. 4. Shock-capturing test case: the exact solution has an exponential boundary layer on the right and top sides of the computational domain. The calculations are performed on the second mesh of the mesh family $\mathcal{M}1$ by taking the constant velocity field $V = (2, 3)^{\mathrm{T}}$ and $\nu = 10^{-4}$. The numerical solution is displayed at the mesh vertices through linear interpolation. Severe oscillations are visible in plot (c) when we use the scheme HMM-(no stabilization), that is, the central mimetic discretization without any stabilization (note the different scale along the z-axis). This phenomenon disappears in plot (d) when jump stabilization is turned on using the scheme HMM-Jmp.

approximated on the sequence of meshes of $\mathcal{M}1$. The numerical solution is plotted at mesh vertices. Vertex values are obtained by interpolating the approximate cell averages provided by the scheme.

4.2.1 *Exponential boundary layers.*

We experimentally investigate how these methods approximate a solution with an exponential boundary layer that forms on those sides of the domain boundary where V points outward. For this purpose, we solve problem (2.1) and (2.2) with the same data that was used for the accuracy benchmark test in the convection-dominated regime, that is, for $\nu = 10^{-4}$.

In Fig. 4 we compare the numerical solutions produced by the following implementations: HMM-SG, HMM-Upw, HMM-(no stabilization) and HMM-Jmp.

In plots (a) and (b) the nonoscillatory solutions produced by the schemes HMM-SG and HMM-Upw, respectively, are displayed. In contrast, from plot (c) it is evident that, when the calculation is performed using the HMM-(no stabilization) scheme without any stabilization, the numerical solution suffers from severe oscillations. These oscillations disappear when we introduce a stabilizing term in the divergence equation that is based on the solution's jump at the mesh edges. However, a great numerical diffusion is introduced by this form of upwinding and the resolution of the boundary layer is poor and generally worse than that obtained through the other HMM implementations.

4.2.2 *Exponential and parabolic boundary layers.*   On $\Omega =]0, 1[\times]0, 1[$ we numerically solve (2.1) and (2.2) with the Dirichlet boundary conditions

$$p(x, 0) = (1 - x)^3, \quad p(x, 1) = (1 - x)^2, \quad p(0, y) = 1, \quad p(1, y) = 0,$$

and $V = (1, 0)^T$ in the convection-dominated regime for $\nu = 10^{-4}$. The solution has an exponential boundary layer at the side $x = 1$ and two parabolic boundary layers at $y = 0$ and $y = 1$. Figure 5 shows the numerical results obtained from calculations using HMM-SG, HMM-Upw, HMM-(no stabilization) and HMM-Jmp. The behaviour is similar to the behaviour documented in Section 4.2.1 for the case of a single exponential layer.

4.3 *Strongly anisotropic heterogeneous and convection-dominated case*

In this third example we consider the test case proposed in Ern *et al.* (2009) where the problem (2.1) and (2.2) is solved for a strongly anisotropic and heterogeneous diffusion tensor and a rotating convection field. A zeroth-order term proportional to p is also present in the model's equations, and its discretization is straightforward (see, e.g., Cangiani *et al.*, 2009). The domain $\Omega =]0, 1] \times]0, 1[$ is split into four subdomains $\Omega_1 =]0, 2/3[\times]0, 2/3[$, $\Omega_2 =]0, 2/3[\times]2/3, 1[$, $\Omega_2 =]2/3, 1[\times]2/3, 1[$ and $\Omega_4 =]2/3, 1[\times]0, 2/3[$. The diffusion tensor is diagonal in each subdomain and is characterized by a very small value along one principal direction as follows:

$$\Lambda = \begin{pmatrix} 10^{-6} & 0 \\ 0 & 1 \end{pmatrix} \quad \text{in } \Omega_1 \text{ and } \Omega_3$$

and

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 10^{-6} \end{pmatrix} \quad \text{in } \Omega_2 \text{ and } \Omega_4.$$

Note that the directions along which diffusion is small are interchanged for adjacent subdomains. Convection is given by the clockwise rotating solenoidal field $V(x, y) = 40(x(2y - 1)(x - 1), -y(2x - 1)(y - 1))$ and the right-hand side is a gaussian bump positioned at a distance of 0.35 from the domain centre, given by $f(x, y) = 10^{-2} \exp(-(r - 0.35)^2/0.005)$ with $r^2 = (x - 0.5)^2 + (y - 0.5)^2$. This problem is convection dominated, thus requiring some sort of upwinding in the numerical treatment of the convection term. Moreover, the exact solution is continuous, but internal layers form near the interfaces that separate the subdomains due to the small diffusion value in the switching directions. The strong solution gradients cannot be resolved by the attainable grid sizes and the numerical approximations are expected to be discontinuous at the internal interfaces.

B. DA VEIGA *ET AL.*

(a) HMM-SG
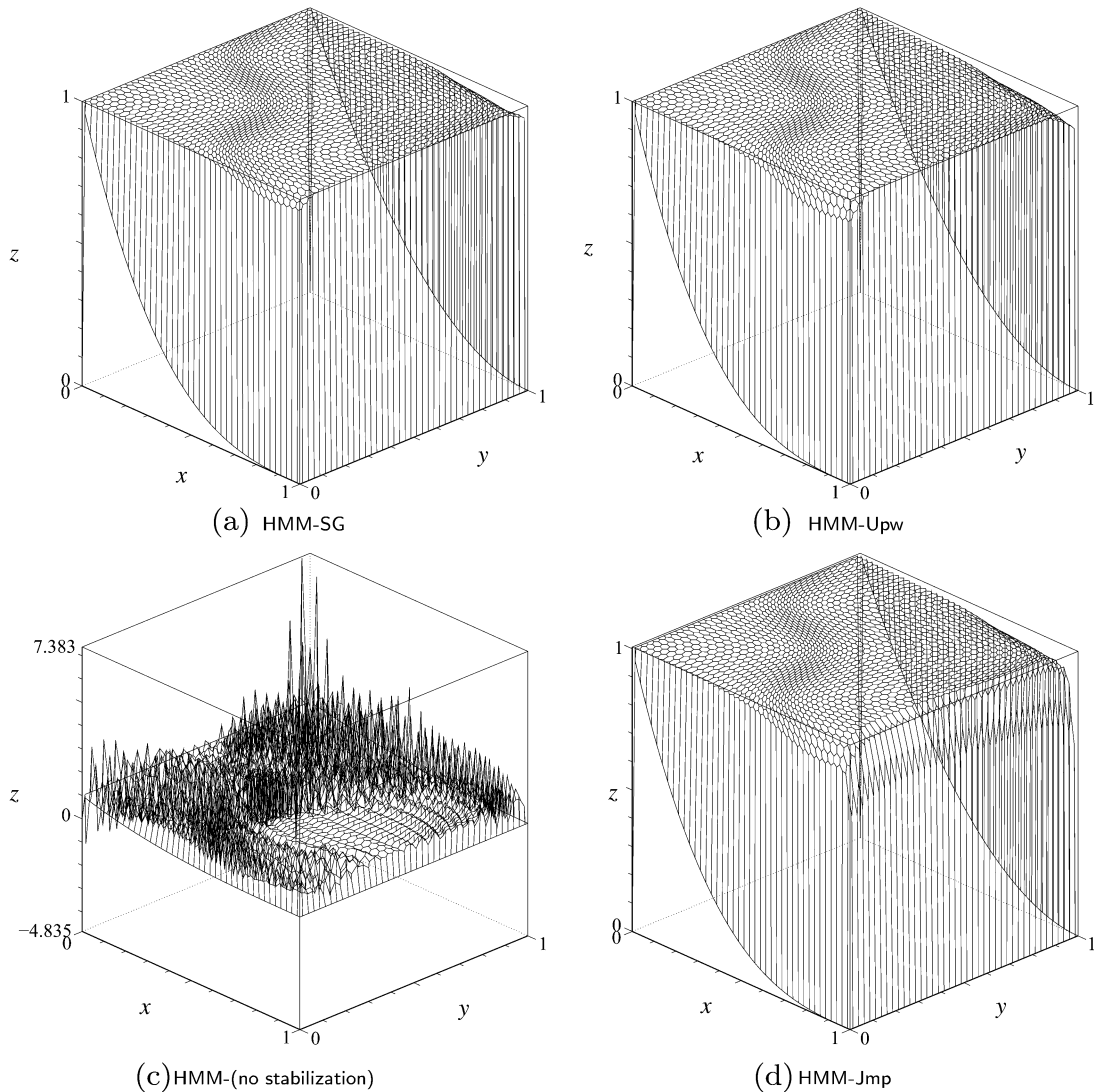
(b) HMM-Upw

(c) HMM-(no stabilization)

(d) HMM-Jmp

FIG. 5. Shock-capturing test case: the exact solution has an exponential boundary layer on the right side and two parabolic layers on the top and bottom sides of the computational domain. The calculations are performed on the second mesh of the mesh family $\mathcal{M}1$ by taking the constant velocity field $V = (1, 0)^{\mathrm{T}}$ and $\nu = 10^{-4}$. The numerical solution is displayed at the mesh vertices through linear interpolation. Severe oscillations are visible in plot (c) when we use the scheme HMM-(no stabilization), that is, the central mimetic discretization without any stabilization (note the different scale along the z-axis). This phenomenon disappears in plot (d) when jump stabilization is turned on using the scheme HMM-Jmp.

In the test cases presented in Sections 4.2.1 and 4.2.2 there was no significant difference between the numerical approximations provided by the cell-based version of the upwind scheme HMM-Upw (see (2.41) and (2.42)), and its edge-based version (see (2.47)–(2.49)). This is no longer the case herein, as illustrated in Fig. 6. The calculations, which use the two alternative versions of the HMM-Upw scheme,

(a) Grid pattern

(b) Exact solution

(c) Cell-based upwind HMM-Upw
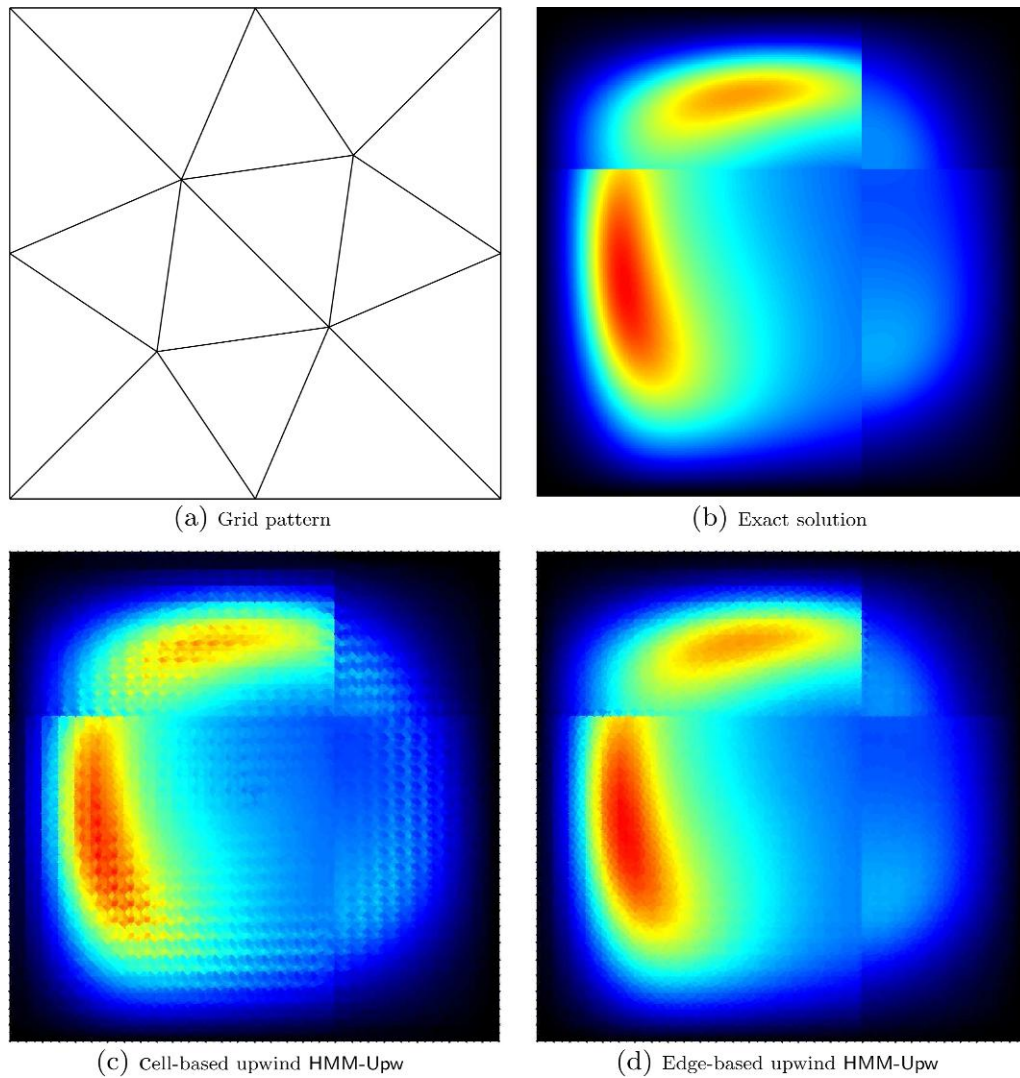
(d) Edge-based upwind HMM-Upw

FIG. 6. Strongly anisotropic heterogeneous and convection-dominated test case: on a coarse mesh the cell-based upwinding of the convection provokes spurious oscillations that are completely absent in the edge-based upwinding discretization.

are performed on a grid obtained by a 30×30 periodic reproduction of the pattern shown in Fig. 6(a). Since the *exact* solution of this problem is unknown, a reference solution is calculated, for comparison's sake, on a very fine cartesian grid. The reference solution is displayed in Fig. 6(b). The numerical solution provided by the cell-based upwind scheme is shown in Fig. 6(c) and is clearly affected by spurious oscillations. In contrast, this undesirable effect is almost completely absent in the numerical solution provided by the edge-based upwind scheme, which is shown in Fig. 6(d).

It is also worth mentioning the behaviour of these two different implementations of the HMM-Upw scheme as far as minimum and maximum principles are concerned. For this purpose, we recall that

the numerical solutions obtained by first-order upwind two-point FV schemes in convection-dominated problems are characterized by numerical properties such as positivity, monotonicity, etc. A thorough inspection of our numerical results reveals that both cell-based and edge-based schemes respect the minimum value, which is zero for the reference solution, and provide 6.6×10^{-4} and 6.9×10^{-4}, respectively, for the maximum value against a reference value of approximately 6.7×10^{-4}. Nonetheless, we noted a minimum value of approximately -1.1×10^{-5}, which corresponds to a numerical undershoot of around 1.6%, when we applied the cell-based scheme on a different mesh given by splitting every other rectangular cell of a 120×60 regular partition of Ω into two subtriangles. On this latter mesh, the edge-based upwind scheme was still seen to respect the zero minimum value. We do not show the other solution plots for these latter calculations because their behaviours are very similar to those of the solutions shown in Fig. 6. From these qualitative comparisons we deduce that the edge-based upwind scheme may be more stable and accurate than the cell-based upwind scheme. We also remark that the edge-based upwind scheme has the advantage of being fully hybridizable, thus leading to a linear system in the edge unknowns through local variable eliminations such as, for example, in the static condensation of mixed finite elements. For these reasons, the edge-based upwind scheme may be preferable when dealing with stiff problems on coarse meshes.

REMARK 4.2 As a final comment, we observe that, in all of the developed tests, the schemes HMM-(no stabilization) and HMM-Jmp, which satisfy only (AB3-w), do not show particular pathologies for coarse meshes. Therefore, at least on the basis of the presented tests, the h-small-enough condition appearing in Theorem 3.11 does not seem to pose a true limitation in practice.

5. Conclusions

We have presented a new family of methods for the numerical approximation to the solution of the steady convection–diffusion equation. These methods, which are referred to as HMM methods, are based on a unified formulation for the hybrid FV method, the mixed FV method and the MFD method, and differ mainly in the approximation of the convection term. In particular, we considered centred, upwind, weighted and locally scaled Scharfetter–Gummel-type discretizations for which we provided a full proof of convergence under very general regularity conditions of the solution field and derived an error estimate when the scalar solution is in $H^2(\Omega)$.

In the last part of the paper we numerically compared the performance of these schemes on a set of test cases selected from the literature in both diffusion- and convection-dominated regimes. As expected, the methods, including a centred-type discretization of the convective term, showed a better behaviour in the test cases dominated by diffusion, exhibiting a superconvergence in the approximation of both scalar and vector variables. On the other hand, such schemes showed a strong loss of convergence rate in the convection-dominated tests, while on those same tests the methods with upwinding or stabilization exhibited a better behaviour. Finally, we showed a test with strong anisotropy and jumps in the coefficients. The results seem to suggest that the hybridized formulation gives more stable results for this kind of problem.

References

AAVATSMARK, I., BARKVE, T., BOE, O. & MANNSETH, T. (1998a) Discretization on unstructured grids for inhomogeneous, anisotropic media. Part I: derivation of the methods. *SIAM J. Sci. Comput.*, **19**, 1700–1716.

AAVATSMARK, I., BARKVE, T., BOE, O. & MANNSETH, T. (1998b) Discretization on unstructured grids for inhomogeneous, anisotropic media. Part II: discussion and numerical results. *SIAM J. Sci. Comput.*, **19**, 1717–1736.

ARBOGAST, T., DAWSON, C. N., KEENAN, P. T., WHEELER, M. F. & YOTOV, I. (1998) Enhanced cell-centered finite differences for elliptic equations on general geometry. *SIAM J. Sci. Comput.*, **19**, 404–425.

ARNOLD, D. N. & BREZZI, F. (1985) Mixed and nonconforming finite element methods. Implementation, post processing and error estimates. *RAIRO Modél Math. Anal. Numér.*, **19**, 7–32.

ARNOLD, D. N., BREZZI, F., COCKBURN, B. & MARINI, L. D. (2002) Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, **39**, 1749–1779.

BARANGER, J., MAITRE, J.-F. & OUDIN, F. (1996) Connection between finite volume and mixed finite element methods. *RAIRO Modél. Math. Anal. Numér.*, **30**, 445–465.

BEIRÃO DA VEIGA, L. (2008) A residual based error estimator for the mimetic finite difference method. *Numer. Math.*, **108**, 387–406.

BEIRÃO DA VEIGA, L. (2010) A mimetic finite difference method for linear elasticity. ESAIM *Math. Model. Numer. Anal.*, **44**, 231–250.

BEIRÃO DA VEIGA, L., GYRYA, V., LIPNIKOV, K. & MANZINI, G. (2009a) Mimetic finite difference method for the Stokes problem on polygonal meshes. *J. Comput. Phys.*, **228**, 7215–7232.

BEIRÃO DA VEIGA, L., LIPNIKOV, K. & MANZINI, G. (2010) The mimetic finite difference method for the steady Stokes problem on polyhedral meshes. *SIAM J. Numer. Anal.*, Available as Technical Report 6PV09/5/0, IMATI-CNR, Pavia, Italy (2009).

BEIRÃO DA VEIGA, L., LIPNIKOV, K. & MANZINI, G. (2009b) Convergence analysis of the high-order mimetic finite difference method. *Numer. Math.*, **113**, 325–356.

BEIRÃO DA VEIGA, L. & MANZINI, G. (2008a) A higher-order formulation of the mimetic finite difference method. *SIAM J. Sci. Comput.*, **31**, 732–760.

BEIRÃO DA VEIGA, L. & MANZINI, G. (2008b) An a posteriori error estimator for the mimetic finite difference approximation of elliptic problems. *Int. J. Numer. Methods Eng.*, **76**, 1696–1723.

BERNDT, M., LIPNIKOV, K., MOULTON, J. D. & SHASHKOV, M. (2001) Convergence of mimetic finite difference discretizations of the diffusion equation. *East-West J. Numer. Math.*, **9**, 253–284.

BERTOLAZZI, E. & MANZINI, G. (2002) Algorithm 817 P2MESH: generic object-oriented interface between 2-D unstructured meshes and FEM/FVM-based PDE solvers. *ACM Trans. Math. Softw.*, **28**, 101–132.

BERTOLAZZI, E. & MANZINI, G. (2004) A cell-centered second-order accurate finite volume method for convection–diffusion problems on unstructured meshes. *Math. Models Methods Appl. Sci.*, **8**, 1235–1260.

BRENNER, S. & SCOTT, L. (1994) *The Mathematical Theory of Finite Element Methods*. Berlin: Springer.

BREZZI, F., BUFFA, A. & LIPNIKOV, K. (2009) Mimetic finite differences for elliptic problems. *Math. Model Numer. Anal.*, **43**, 277–295.

BREZZI, F. & FORTIN, M. (1991) *Mixed and Hybrid Finite Element Methods*. New York: Springer.

BREZZI, F., LIPNIKOV, K. & SHASHKOV, M. (2005a) Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. *SIAM J. Numer. Anal.*, **43**, 1872–1896.

BREZZI, F., LIPNIKOV, K., SHASHKOV, M. & SIMONCINI, V. (2007) A new discretization methodology for diffusion problems on generalized polyhedral meshes. *Comput. Methods Appl. Mech. Eng.*, **196**, 3682–3692.

BREZZI, F., LIPNIKOV, K. & SIMONCINI, V. (2005b) A family of mimetic finite difference methods on polygonal and polyhedral meshes. *Math. Models Methods Appl. Sci.*, **15**, 1533–1551.

CANGIANI, A. & MANZINI, G. (2008) Flux reconstruction and pressure post-processing in mimetic finite difference methods. *Comput. Methods Appl. Mech. Eng.*, **197**, 933–945.

CANGIANI, A., MANZINI, G. & RUSSO, A. (2009) Convergence analysis of a mimetic finite difference method for general second-order elliptic problems. *SIAM J. Numer. Anal.*, **47**, 2612–2637.

CHAINAIS-HILLAIRET, C. & DRONIOU, J. (2007) Convergence analysis of a mixed finite volume scheme for an elliptic-parabolic system modeling miscible fluid flows in porous media. *SIAM J. Numer. Anal*, **45**, 2228–2258.

CHAINAIS-HILLAIRET, C. & DRONIOU, J. (2009) Finite volume schemes for non-coercive elliptic problems with Neumann boundary conditions. *IMA J. Numer. Anal.* (to appear) doi: 10.1093/imanum/drp009.

COUDIÈRE, Y. & MANZINI, G. (2010) The discrete duality finite volume method for convection–diffusion problems. *SIAM J. Numer. Anal.*, **47**, 4163–4192.

DAWSON, C. & AIZINGER, V. (1999) Upwind-mixed methods for transport equations. *Comput. Geosci.*, **3**, 93–110.

DOUGLAS JR, J. & ROBERTS, J. E. (1982) Mixed finite element methods for second order elliptic problems. *Math. Appl. Comput.*, **1**, 91–103.

DOUGLAS JR, J. & ROBERTS, J. E. (1985) Global estimates for mixed methods for second order elliptic equations. *Math. Comput.*, **44**, 39–52.

DRONIOU, J. & EYMARD, R. (2006) A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numer. Math.*, **105**, 35–71.

DRONIOU, J. & EYMARD, R. (2009) Study of the mixed finite volume method for Stokes and Navier–Stokes equations. *Numer. Methods Partial Differ. Equ.*, **25**, 137–171.

DRONIOU, J., EYMARD, R., GALLOUËT, T. & HERBIN, R. (2010) A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *Math. Models Methods Appl. Sci.*, **20**, 265–295.

ERN, A., STEPHANSEN, A. F. & ZUNINO, P. (2009) A discontinuous Galerkin method with weighted averages for advection–diffusion equations with locally small and anisotropic diffusivity. *IMA J. Numer. Anal.*, **29**, 235–256.

EYMARD, R., GALLOUËT, T. & HERBIN, R. (2000) The finite volume method. *Techniques of Scientific Computing, Part III. Handbook for Numerical Analysis* (P. Ciarlet & J. L. Lions eds). The Netherlands: North-Holland, pp. 715–1022.

EYMARD, R., GALLOUËT, T. & HERBIN, R. (2009) Discretization of heterogeneous and anisotropic diffusion problems on general non-conforming meshes, SUSHI: a scheme using stabilisation and hybrid interfaces. *IMA J. Numer. Anal.*, **30**, 1009–1043.

HYMAN, J., MOREL, J., SHASHKOV, M. & STEINBERG, S. (2002) Mimetic finite difference methods for diffusion equations. *Comput. Geosci.*, **6**, 333–352.

JAFFRE, J. (1984) Decentrage et elements finis mixtes pour les equations de diffusion–convection. *Calcolo*, **21**, 171–197.

JAFFRE, J. & ROBERTS, J. E. (1985) Upstream weighting and mixed finite elements in the simulation of miscible displacements. *RAIRO Modél. Math. Anal. Numér.*, **19**, 443–460.

KUZNETSOV, Y., LIPNIKOV, K. & SHASHKOV, M. (2005) The mimetic finite difference method on polygonal meshes for diffusion-type problems. *Comput. Geosci.*, **8**, 301–324.

LIPNIKOV, K., SHASHKOV, M. & YOTOV, I. (2009) Local flux mimetic finite difference methods. *Numer. Math.*, **112**, 115–152.

MANZINI, G. & RUSSO, A. (2008) A finite volume method for advection–diffusion problems in convection-dominated regimes. *Comput. Methods Appl. Mech. Eng.*, **197**, 1242–1261.

RAPIN, G. & LUBE, G. (2004) A stabilized scheme for the Lagrange multiplier method for advection–diffusion equations. *Math. Models Methods Appl. Sci.*, **14**, 1035–1060.

RIVIERE, B. (2008) *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation*. Philadelphia, PA: SIAM.

RUSSELL, T. F. & WHEELER, M. F. (1983) Finite element and finite difference methods for continuous flows in porous media. *The Mathematics of Reservoir Simulation* (R. E. Ewing ed.). Philadelphia, PA: SIAM, pp. 35–106.

SCHARFETTER, D. L. & GUMMEL, H. K. (1969) Large signal analysis of a silicon read diode. *IEEE Trans. Electron Devices*, **16**, 64–77.

VOHRALIK, M. (2006) Equivalence between lowest-order mixed finite element and multi-point finite volume methods on simplicial meshes. *Math. Model. Numer. Anal.*, **40**, 367–391.

WHEELER, M. F. & YOTOV, I. (2006) A multipoint flux mixed finite element method. *SIAM J. Numer. Anal.*, **44**, 2082–2106.

YOUNES, A., ACKERER, P. & CHAVENT, G. (2004) From mixed finite elements to finite volumes for elliptic PDEs in two and three dimensions. *Int. J. Numer. Methods Eng.*, **59**, 365–388.