MIXED FINITE ELEMENT METHODS:
IMPLEMENTATION WITH ONE UNKNOWN PER ELEMENT,
LOCAL FLUX EXPRESSIONS, POSITIVITY, POLYGONAL
MESHES, AND RELATIONS TO OTHER METHODS

MARTIN VOHIRALÍK*
UPMC Univ. Paris 06, UMR 7598,
Laboratoire Jacques-Louis Lions,
75005 Paris, France
CNRS, UMR 7598, Laboratoire Jacques-Louis Lions,
75005 Paris, France
vohralik@ann.jussieu.fr

BARBARA I. WOHLMUTH
Fakultät für Mathematik,
Lehrstuhl für Numerische Mathematik,
Boltzmannstrasse 3, 85748 Garching bei München, Germany
wohlmuth@ma.tum.de

Received 21 July 2010
Revised 16 April 2012
Accepted 3 July 2012
Published 26 October 2012
Communicated by F. Brezzi

In this paper, we study the mixed finite element method for linear diffusion problems. We focus on the lowest-order Raviart–Thomas case. For simplicial meshes, we propose several new approaches to reduce the original indefinite saddle point systems for the flux and potential unknowns to (positive definite) systems for one potential unknown per element. Our construction principle is closely related to that of the so-called multi-point flux-approximation method and leads to local flux expressions. We present a set of numerical examples illustrating the influence of the elimination process on the structure and on the condition number of the reduced matrix. We also discuss different versions of the discrete maximum principle in the lowest-order Raviart–Thomas method. Finally, we recall mixed finite element methods on general polygonal meshes and show that they are a special type of the mimetic finite difference, mixed finite volume, and hybrid finite volume family.

Keywords: Mixed finite element method; local static condensation; local flux expression; discrete maximum principle; polygonal mesh; locally conservative methods.

AMS Subject Classification: 65N30, 76M10, 76M12, 76S05

*Current address: INRIA Paris-Rocquencourt, B.P. 105, 78153 Le Chesnay, France, E-mail: martin.vohralik@inria.fr.
1. Introduction

Let \( \Omega \subset \mathbb{R}^d, d \geq 2, \) be a polygonal (we use this term also in \( \mathbb{R}^d, d \geq 3 \)) domain. We consider the pure diffusion model problem: find the potential \( p : \Omega \rightarrow \mathbb{R} \) such that:

\[
- \nabla \cdot (S \nabla p) = g \quad \text{in } \Omega, \quad \tag{1.1a}
\]

\[
p = 0 \quad \text{on } \partial \Omega, \quad \tag{1.1b}
\]

where \( S : \Omega \rightarrow \mathbb{R}^{d \times d} \) is a symmetric, bounded, and uniformly positive definite diffusion tensor and \( g : \Omega \rightarrow \mathbb{R} \) is a source term. Let \( T_h \) be a matching simplicial mesh of \( \Omega \). For simplicity, we restrict ourselves to the homogeneous Dirichlet boundary condition (1.1b) but our approach can be easily generalized to other boundary conditions. We also assume that \( S \) and \( g \) are piecewise constant on the mesh \( T_h \). We denote by \( u := -S \nabla p \) the flux.

We are interested in the discretization of (1.1a)-(1.1b) by mixed finite elements (MFEs). This consists in finding \( p_h \in \Phi_h \), an approximation to the potential \( p \), and \( u_h \in V_h \), an approximation to the flux \( u \), such that:

\[
(S^{-1} u_h, v_h) - (p_h, \nabla \cdot v_h) = 0 \quad \forall \, v_h \in V_h, \quad \tag{1.2a}
\]

\[
-(\nabla \cdot u_h, \phi_h) = -(g, \phi_h) \quad \forall \, \phi_h \in \Phi_h. \quad \tag{1.2b}
\]

Here, we focus on the lowest-order Raviart–Thomas case, where the degrees of freedom of the potential space \( \Phi_h \) are elementwise constants on \( T_h \), and the degrees of freedom of the flux space \( V_h \) represent the sidewise constant values of the normal component of the flux. The system (1.2a)-(1.2b) can be written in matrix form as

\[
\begin{pmatrix}
A & B^T \\
B & 0
\end{pmatrix}
\begin{pmatrix}
U \\
P
\end{pmatrix} =
\begin{pmatrix}
F \\
G
\end{pmatrix} \quad \tag{1.3}
\]

and is of indefinite, saddle point type. In our case, thanks to the homogeneous Dirichlet boundary condition (1.1b), \( F = 0 \), but we prefer to write the general form (1.3) with \( F \) not necessarily zero. The system (1.3) is well-posed; \( A \) is symmetric and positive definite and \( B \) has a full row rank. Thus, there exists a unique solution, on any simplicial mesh \( T_h \), without any restriction on the shape of the elements; the usual shape-regularity of the mesh is only necessary in convergence proofs. Recall also that the system (1.3) is well-posed in any space dimension. All these classical results can be found, e.g. in Refs. 16, 47, and 54.

There has been a long-standing interest to reduce (1.3) to a system for the potentials \( P \) only. The main motivations are to reduce the number of unknowns, to replace the saddle point system (1.3) by, if possible, a symmetric and positive definite one, and to relate the mixed finite element method to the finite difference and finite volume ones. A possible solution consists in using the first block equation of (1.3) to eliminate the unknowns \( U \) through

\[
U = A^{-1}(F - B^T P). \quad \tag{1.4}
\]
Note that (1.4) represents a global flux expression, since $A^{-1}$ as the inverse of a mass matrix is not sparse. Plugging (1.4) into the second block equation of (1.3), one can solve for $P$ the system
\[ BA^{-1}B^tP = BA^{-1}F - G. \tag{1.5} \]
The matrix $BA^{-1}B^t$ is symmetric and positive-definite; however, it is full and not locally computable. In mixed Schur complement methods, the explicit construction of $BA^{-1}B^t$ is avoided by solving (1.4) in each step of an iterative process.

In the past, various approximate numerical quadratures have been used, see, e.g., Refs. 48, 4, 11, and 8 (we sort the references by time) to reduce (1.3) into a system of the form
\[ \tilde{S}\tilde{P} = \tilde{H}, \tag{1.6} \]
with sparse, locally computable, and possibly symmetric and positive definite matrix $\tilde{S}$. In these approaches, however, because of the numerical quadratures, the new potentials $\tilde{P}$ are, in general, different from the potentials $P$ in (1.3) and one cannot recover the exact potentials $P$.

One-unknown-per-element rewriting of (1.3) without any numerical quadrature in the form
\[ S\bar{P} = H, \tag{1.7} \]
where $\bar{P}$ is a new unknown from which $P$ can be locally recovered, has been achieved in two space dimensions in Refs. 59, 20, and 57 by exploiting an equivalence between lowest-order MFEs and finite volumes. A similar approach, with various extensions, has been taken in Ref. 43. In Ref. 53, a system of the form (1.7) has been obtained with $\bar{P} = P$, i.e. directly for the original unknowns $P$, in two or three space dimensions. In these approaches, in contrast to (1.5), the matrix $S$ is sparse and locally computable and, in contrast to (1.6), one obtains exactly the potentials $P$ of (1.3), by solving (1.7) and possibly performing a local postprocessing step. Intermediately, local flux expressions (enabling to recover the fluxes $U$ of (1.3) on sides of local patches from the potentials $P$ on elements of these patches) have been established in Refs. 59, 20, 57, and 53. For alternative approaches to reduce the number of unknowns in (1.3), we refer to Refs. 6, 51 and the references therein.

The first goal of the present paper is to unify the approaches of Refs. 59, 20, 57 and of Ref. 53, to identify their common principles, and to show that they can be included in the general approach developed herein and in Ref. 55. We tackle this issue in Sec. 4, after summarizing the notation in Sec. 2 and recalling some basics in Sec. 3. In Sec. 3, we also recall the relation of the lowest-order Raviart–Thomas (RT0) mixed finite element method to the Crouzeix–Raviart (CR) nonconforming finite element (NCFE) method and local flux expressions from the Lagrange multipliers.

The second goal of the present paper is to carry out a comparative numerical study of the different one-unknown-per-element RT0 MFE reformulations. In particular, we focus on symmetry, positive definiteness, sparsity, condition number, and
performance of standard direct and iterative solvers in the different approaches in the presence of inhomogeneous and anisotropic diffusion tensors $S$. This numerical study is carried out in Sec. 5.

The third goal of the present paper is to recall and state precisely the different variants of the discrete maximum principle valid in the RT0 MFE method. We do this in Sec. 6.

The fourth goal of the present paper is to recall that, in contrast to a widespread misleading belief, MFEs can be easily defined on general polygonal meshes. We present this result in Sec. 7 on quite general meshes and without losing any accuracy with respect to (1.2a)–(1.2b). The presentation of this section is done for all order schemes. We simply introduce a matching simplicial submesh and use a local static condensation corresponding to the solution of local Dirichlet or Neumann problems.

The last goal of the present paper is to recall known and show new relations between MFE methods and other discretization schemes, namely the two-point finite volume (FV),\textsuperscript{29} mimetic finite difference (MFD),\textsuperscript{17,13} hybrid finite volume,\textsuperscript{26} multi-point flux-approximation (MPFA),\textsuperscript{1,3,19} and related methods.\textsuperscript{28,41,56,14,42} We, in particular, prove in Sec. 7 that the RT0 MFE method on arbitrary polygonal meshes is a particular example of the MFD method. It seems that the only conceptual difference of the present MFE approach with these methods is that in MFEs, one has to construct a simplicial submesh and solve a local problem on each polygonal cell. For some other comparisons between these methods, we refer to Ref. 38 and the references therein. Finally, some conclusions are drawn in Sec. 8.

2. Notation

This section is devoted to collecting in one place the different notation used throughout the paper.

For a given domain $\omega \subset \mathbb{R}^d$, let $L^2(\omega)$ be the space of square-integrable functions over $\omega$ and $(\cdot, \cdot)_{\omega}$ be the $L^2(\omega)$ inner product; we omit the index $\omega$ when $\omega = \Omega$. By $|\omega|$, we denote the Lebesgue measure of $\omega$ and by $|\sigma|$ the $(d-1)$-dimensional Lebesgue measure of a $(d-1)$-dimensional surface $\sigma$ in $\mathbb{R}^d$. The symbol $|S|$ also stands for the cardinality (the number of elements) of a set $S$. Finally, we denote by $(\cdot, \cdot)_{\sigma}$ the $(d-1)$-dimensional $L^2(\sigma)$-inner product on $\sigma \subset \mathbb{R}^{d-1}$.

The elements of the mesh $T_h$ are triangles in two space dimensions, tetrahedra in three space dimensions, and, in general, $d$-simplices. For simplicity, we exclude the pathological cases with just a couple of mesh elements. We suppose that $T_h$ is matching, i.e. such that if $K, L \in T_h$, $K \neq L$, then $K \cap L$ is either an empty set or a common vertex of $K$ and $L$ or a common $d'$-face of the mesh $T_h$, $1 \leq d' \leq d - 1$. We denote by $E_h$ the set of all sides of $T_h$, i.e. the set of $(d-1)$-faces of the mesh $T_h$. We divide $E_h$ into the set $E_h^{\text{int}}$ of interior sides and the set $E_h^{\text{ext}}$ of boundary sides. Let $K \in T_h$. By $E_K$, we denote the set of all sides of $K$ and by $E_K^{\text{int}} := E_K \cap E_h^{\text{int}}$. For $K \in T_h$, let $n_K$ be the outward unit normal vector defined on the sides of $K$, and,
for \( \sigma \in \mathcal{E}_h \), let \( n_\sigma \) stand for the unit normal vector of \( \sigma \) whose orientation is chosen arbitrarily but fixed for interior sides and coinciding with the exterior normal of \( \Omega \) for boundary sides. We denote by \( V_h \) the set of vertices of \( T_h \). For a given vertex \( V \in V_h \), we shall denote by \( T_V \) the patch of the elements of \( T_h \) which share \( V \), by \( \mathcal{E}_V \) those sides of the elements in \( T_V \) contained in \( \mathcal{E}_{h}^{\text{int}} \), and by \( \mathcal{E}_V^{\text{ext}} \) the sides in the interior of \( T_V \). We set \( \mathcal{E}_V^{\text{ext}} := \mathcal{E}_V \backslash \mathcal{E}_V^{\text{int}} \). Figure 1 gives various examples. We will also employ the notation \( T_V^{\text{ext}} \) for all elements from \( T_V \) having a side from the set \( \mathcal{E}_V^{\text{ext}} \); \( T_V^{\text{ext}} = T_V \) apart from the second situation of Fig. 1. We denote by \( \mathcal{E}_{V,K} \) the sides of \( K \) which have \( V \) as vertex. For \( K \in T_h \), let \( x_K \) stand for the barycenter of \( K \) and for \( \sigma \in \mathcal{E}_h \), let \( x_\sigma \) stand for the barycenter of \( \sigma \). We will also employ the notation \( T_\sigma \) for the patch of the elements of \( T_h \) which share the side \( \sigma \in \mathcal{E}_h \). Finally, \( I \) denotes the identity matrix.

Fig. 1. An example of a patch \( T_V \) around a vertex \( V \) in the interior of \( \Omega \) (top), close to the boundary of \( \Omega \) (middle), and on the boundary of \( \Omega \) (bottom).
3. The Lowest-Order Raviart–Thomas Mixed Finite Element Method

We recall here some well-known properties of the RT0 MFE method. First, in Sec. 3.1, we define properly the spaces $\Phi_h$ and $V_h$. Section 3.2 gives the hybridization and Sec. 3.3 recalls its relation to the Crouzeix–Raviart nonconforming finite element, as well as some other methods. In Sec. 3.4, we then present the local flux expressions from the Lagrange multipliers.

3.1. Spaces $\Phi_h$ and $V_h$

The RT0 MFE method is given by (1.2a)–(1.2b) where the spaces $\Phi_h$ and $V_h$ are specified following Ref. 46 for $d = 2$ and Ref. 45 for $d = 3$. The space $\Phi_h$ consists of piecewise constants on $T_h$. For the space $V_h$, there is one basis function $v_\sigma$ associated with each side $\sigma \in E_h$. Let $\sigma$ be an interior side of the mesh $T_h$ shared by the elements $K$ and $L$. Then the associated basis function is given by

$$v_\sigma(x) = \frac{1}{|K|}(x - V_K), \quad x \in K,$$

$$v_\sigma(x) = \frac{1}{|L|}(V_L - x), \quad x \in L,$$

$$v_\sigma(x) = 0 \text{ otherwise},$$

where $V_K$ is the vertex of $K$ opposite to $\sigma$ and $V_L$ the vertex of $L$ opposite to $\sigma$. The orientation of $v_\sigma$ (the order of $K$ and $L$) coincides with the orientation of the normal vector $n_\sigma$. For a boundary side $\sigma$, the support of $v_\sigma$ only consists of the element $K \in T_h$ such that $\sigma \in \partial K$ and $v_\sigma(x) = \frac{1}{|K|}(x - V_K), \quad x \in K$. We refer to Refs. 16 and 47 for more details.

3.2. Hybridization

Let $\sigma$ be an interior side of the mesh $T_h$ shared by the elements $K$ and $L$. The approximate fluxes $u_h$ of (1.2a)–(1.2b) satisfy $u_h|_K \cdot n_\sigma = u_h|_L \cdot n_\sigma$. This constraint can be relaxed using the hybridization technique. The unconstrained flux space is given by $V^H_h := \Pi_{K \in T_h} V_h(K)$, where $V_h(K)$ are the local spaces on each mesh element, and the Lagrange multipliers space $\Psi_h$ is the space of piecewise constants on the interior sides. The hybridized version of (1.2a)–(1.2b) consists in finding $u_h \in V^H_h$, $p_h \in \Phi_h$, and $\lambda_h \in \Psi_h$ such that:

$$\left(\nabla \cdot u_h, v_h\right) - \int_{\partial \Omega} (u_h \cdot n_K, \lambda_h) = 0 \quad \forall v_h \in V^H_h; \quad (3.1a)$$

$$-\left(\nabla \cdot u_h, \phi_h\right) = -\left(g, \phi_h\right) \quad \forall \phi_h \in \Phi_h; \quad (3.1b)$$

$$\sum_{K \in T_h} (u_h \cdot n_K, \psi_h) = 0 \quad \forall \psi_h \in \Psi_h. \quad (3.1c)$$

Here, $\lambda_h$ provides an additional approximation to the potential $p$ on the sides of the mesh. We point out that (1.2a)–(1.2b) and (3.1a)–(3.1c) lead to the same solutions $p_h$ and $u_h$. Due to the different flux ansatz spaces, the algebraic vector associated
with $u_h$ in $V^H_h$ is different from $U$ and denoted by $U^H$. Equations (3.1a)–(3.1c) give in matrix form
\[
\begin{pmatrix}
A^H & (B^H)^t & C^t \\
B^H & 0 & 0 \\
C & 0 & 0
\end{pmatrix}
\begin{pmatrix}
U^H \\
P \\
\Lambda
\end{pmatrix} =
\begin{pmatrix}
F^H \\
G \\
0
\end{pmatrix}.
\] (3.2)

The importance of the hybridization (3.2) lies in the observation that the matrix $A^H$ has now an elementwise block-diagonal structure. Consequently, (3.2) can be reduced by static condensation to the problem
\[
Z \Lambda = E,
\] (3.3)
see Ref. 16 (Sec. V.1). The matrix $Z$ is symmetric and positive definite with a narrow stencil.

3.3. Relation to the Crouzeix–Raviart nonconforming finite element, finite volume, multi-point flux-approximation, and mimetic methods

Let $\Psi_{NC}^h$ be the Crouzeix–Raviart space, see Ref. 25. This is the space of piecewise affine functions on $T_h$ which are continuous in the barycenters of the interior sides and zero in the barycenters of the boundary sides. There is one basis function associated with each interior side $\sigma \in E_h^{int}$, denoted by $\psi_{\sigma}$. It is such that $\psi_{\sigma}(x_{\sigma}) = 1$ and $\psi_{\sigma}(x_{\gamma}) = 0$ for all sides $\gamma \in E_h$ different from $\sigma$.

The CR NCFE method for the problem (1.1a)–(1.1b) reads: find $\lambda_{NC}^h \in \Psi_{NC}^h$ such that:
\[
(\boldsymbol{S} \nabla \lambda_{NC}^h , \nabla \psi_h) = (g, \psi_h) \quad \forall \psi_h \in \Psi_{NC}^h.
\] (3.4)

In a matrix setting, it can be written as
\[
Z \Lambda = E.
\] (3.5)

It follows from Refs. 10, 44, 7, and 22 that the above matrix $Z$ and the right-hand side vector $E$ coincide with the $Z$ and $E$ of (3.3). Thus, the CR NCFE method (3.5) is equivalent to the hybridization (3.3) of the RT0 MFE one. Let $\sigma \in E_h^{int}$ and denote the value of $\lambda_h$ of (3.1a)–(3.1c) on $\sigma$ by $\Lambda_{\sigma}$. Then we can obtain $\lambda_{NC}^h$ of (3.4) by
\[
\lambda_{NC}^h = \sum_{\sigma \in E_h^{int}} \Lambda_{\sigma} \psi_{\sigma}.
\] (3.6)

Other relations have also been established.

Remark 3.1. (Relation to two-point FV) It has been proven in Refs. 59, 20, and 57, see also the references therein, that the RT0 MFE (3.1a)–(3.1c) and the two-point FV discretization coincide in two space dimensions for zero source term $g$. We will in fact recover this result in Sec. 4.4.1 below.
Remark 3.2. (Relation to MPFA) For \( d = 2 \) and \( g = 0 \), it was shown in Refs. 37, 58, based on the results of Ref. 53, that the nonsymmetric MPFA O-method\(^1\) is equivalent to the RT0 MFE written with one unknown per element in the form of Sec. 4.1.2 below.

Remark 3.3. (Relation to MFD) In Refs. 17 (Sec. 5.1) and 18 (Example 1), it has been pointed out that the MFD method on simplicial meshes contains as one of its variants the RT0 MFE method (1.2a)–(1.2b), in the sense that it leads to the same linear system (1.3).

3.4. Local flux expressions from the Lagrange multipliers

Following Refs. 10, 44, and 22, there holds:

\[
\begin{align*}
\mathbf{u}_h|_K &= -S_K \nabla \lambda_h^{\text{NC}}|_K + \frac{g|_K}{d} (\mathbf{x} - \mathbf{x}_K)|_K, \quad (3.7a) \\
p_h|_K &= \lambda_h^{\text{NC}}(\mathbf{x}_K) + \frac{g|_K}{d^2|K|^2} (S^{-1}(\mathbf{x} - \mathbf{x}_K), \mathbf{x} - \mathbf{x}_K)|_K. \quad (3.7b)
\end{align*}
\]

Relation (3.7a) states that there exist local flux expressions from the Lagrange multipliers. Relation (3.7b) means that the original potential approximation \( p_h \) is linked to the Lagrange multipliers \( \lambda_h \) and, additionally, to the source term \( g \). Thus we can locally express both \( \mathbf{u}_h \) and \( p_h \) in terms of the Lagrange multipliers \( \lambda_h \). It is, however, not obvious how to express locally \( \mathbf{u}_h \) from \( p_h \).

4. Reductions to One Unknown Per Element
   and Local Flux Expressions

We present here a unified framework allowing to reduce (1.3), (3.2), or (3.3) in the lowest-order MFE method equivalently to (1.7). Firstly, in Secs. 4.1 and 4.2, we define local problems on patches of elements which have to be solved to give local flux expressions. Section 4.3 then shows how to express the vectors of the Lagrange multipliers \( \Lambda \) or of the fluxes \( \mathbf{U}^H \) from the new potential unknowns \( \bar{P} \), and Sec. 4.4 gives two different ways of obtaining the final one-unknown-per-element system (1.7).

4.1. Definition of the local problems
   by a geometrical interpretation

It follows from (3.6) that \( \lambda_h^{\text{NC}}|_K \) for any \( K \in T_h \) is given by the \( d + 1 \) values \( \Lambda_{\sigma} \); for simplicity, following from (1.1b), we set \( \Lambda_{\sigma} = 0 \) for all \( \sigma \in \mathcal{E}_h^{\text{ext}} \). Let \( \mathbf{z}_K \) be a point arbitrary in \( \mathbb{R}^d \) but such that any \( d \) of the \( d + 1 \) side barycenters \( \mathbf{x}_{\sigma} \) of \( K \) and the point \( \mathbf{z}_K \) do not lie in the same hyperplane. In two space dimensions, this means...
Fig. 2. Triangle $K \in T_h$ and subtriangle $L$ given by two edge midpoints $x_\sigma$ and $x_\gamma$ and the point $z_K$. That $z_K$ does not lie on the boundary of the dashed triangle in Fig. 2. Our main idea is to use the value of $\lambda^NC_h$ in the point $z_K$ as a new unknown,

$$\bar{P}_K := \lambda^NC_h|_K(z_K). \quad (4.1)$$

Let $V$ be any of the vertices of $K$ and let a new simplex $L$ be given by the side barycenters $x_\sigma, \sigma \in E_{V,K}$, and by the point $z_K$, see Fig. 2 (here the points are denoted by $x_\sigma, x_\gamma, \text{and } z_K$). Denote by $\varphi_\sigma, \sigma \in E_{V,K}$, and by $\varphi_K$ the Lagrange basis functions associated with the vertices of $L$; we consider $\varphi_\sigma$ and $\varphi_K$ supported on $K$. Then, using (3.7a), we get

$$u_h|_K = -S_K \nabla \left( \sum_{\sigma \in E_{V,K}} \Lambda_\sigma \varphi_\sigma + \bar{P}_K \varphi_K \right) + \frac{g|_K}{d}(x - x_K)|_K. \quad (4.2)$$

Let now a vertex $V \in \mathcal{V}_h$ be fixed and consider the patch $T_V$ of elements of $T_h$ which share $V$. We refer to Fig. 1 for examples. Consider (3.1c) for $\psi_h$ associated with the internal sides $E_{V}^{int}$ of the patch $T_V$. This gives rise to the following local problem: given $\bar{P}_K, K \in T_V$, find $\Lambda_\sigma, \gamma \in E_{V}^{int}$, such that (4.2) holds together with

$$\sum_{K \in T_\sigma} (u_h \cdot n_K, 1)_\sigma = 0 \quad \forall \sigma \in E_{V}^{int}. \quad (4.3)$$

Suppose that this square linear system is well-posed. Then we can express the Lagrange multipliers $\Lambda_\gamma$ inside the patch $T_V$ as a function of the new potentials $\bar{P}_K$ in the patch $T_V$ and of the sources $g$ in the patch $T_V$. Considering (4.2), we can also get the fluxes $u_h$ in the whole patch $T_V$ as a function of $\bar{P}_K$, i.e. we obtain local flux expressions from the new potentials $\bar{P}_K$.

**Remark 4.1.** (Construction principle and MPFA) Considering a patch of (sub)elements, supposing a piecewise affine, possibly nonconforming, potential approximation, imposing the normal flux continuity, and solving a local linear system on the patch is also the principle of the MPFA method, cf. also Refs. 28, 41, 56, 14, and 42. In the MPFA method, the fluxes on patches of (sub)elements are locally recovered from the potentials $p_h$ strongly by $u_h := -\nabla p_h$. The difference of
our approach is that we recover the fluxes from the potentials by a weak, variationally consistent formulation, following both the potential–flux relation $u = -\nabla \rho$ and the divergence constraint $\nabla \cdot u = g$, see also Ref. 53 (Remarks 2.2 and 2.3).

Let us now elaborate on (4.2)–(4.3). First recall that (4.3) can be equivalently rewritten as

$$
\sum_{K \in T} \langle u_h \cdot n_K, \psi_{\sigma} \rangle_{\partial K} = 0 \quad \forall \sigma \in \mathcal{E}^{\text{int}}.
$$

Here $\psi_{\sigma}$ are the CR nonconforming basis functions, see Sec. 3.3. This follows by the facts that $u_h \cdot n_K$ is constant on any $\sigma \in \mathcal{E}_K$, $\langle 1, \psi_{\sigma} \rangle_{\sigma} = |\sigma|$, and $\langle 1, \psi_{\sigma} \rangle_{\gamma} = 0$ for any side $\gamma$ different from $\sigma$, so that $\langle u_h \cdot n_K, 1 \rangle_{\sigma} = \langle u_h \cdot n_K, \psi_{\sigma} \rangle_{\sigma} = \langle u_h \cdot n_K, \psi_{\sigma} \rangle_{\partial K}$. From this equality, using the Green theorem, we have

$$
\sum_{K \in T} \left\{ (\nabla \cdot u_h, \psi_{\sigma})_K + (u_h, \nabla \psi_{\sigma})_K \right\} = 0 \quad \forall \sigma \in \mathcal{E}^{\text{int}}.
$$

(4.4)

It follows from (1.2b) and the definition of $V_h(K)$ in Sec. 3 that $(\nabla \cdot u_h)|_K = g|_K$. For the other term in the above relation, we employ (4.2). Notice that

$$
(\mathbf{x} - \mathbf{x}_K, \nabla \psi_{\sigma})_K = 0,
$$
as $(\nabla \psi_{\sigma})|_K$ is a constant vector and $\mathbf{x}_K$ is the barycenter of $K$. Extend the notation $\varphi_{\sigma}$ from (4.2) to a Lagrange basis function supported on all $K \in T_h$ such that $\sigma \in \mathcal{E}_K$. The gradients of these basis function are illustrated in the top part of Fig. 3. We then see that (4.2)–(4.3) is equivalent to the following problem: given

![Fig. 3. Gradients of the basis functions (top) and different evaluation points $z_K$ (bottom) in an element $K \in T_h$.](image-url)
\[ P_K, K \in T_V, \text{ find } \Lambda_\gamma, \gamma \in \mathcal{E}_V^{\text{int}}, \text{ such that:} \]
\[ \sum_{K \in T_{\sigma}} \sum_{\gamma \in \mathcal{E}_V^{\text{int}}} \Lambda_\gamma (S \nabla \varphi_\gamma, \nabla \psi_\sigma)_K = \sum_{K \in T_{\sigma}} (g, \psi_\sigma)_K - (\bar{P}_K (S \nabla \varphi_K, \nabla \psi_\sigma)_K) \]
\[ \forall \sigma \in \mathcal{E}_V^{\text{int}}. \quad (4.5) \]

We point out that (4.5) is a Petrov–Galerkin problem, as the basis functions \( \psi_\sigma \) of the test space are different from the basis functions \( \varphi_\gamma \) of the trial space.

The matrix form of (4.5) writes: given \( P_V := \{ P_K \}_{K \in T_V} \), find \( \Lambda_V^{\text{int}} := \{ \Lambda_\gamma \}_{\gamma \in \mathcal{E}_V^{\text{int}}} \) such that:
\[ M_V \Lambda_V^{\text{int}} = \tilde{G}_V - J_V \hat{P}_V, \quad (4.6) \]
where
\[ (M_V)_{\sigma, \gamma} := \sum_{K \in T_{\sigma}} (S \nabla \varphi_\gamma, \nabla \psi_\sigma)_K, \quad (4.7a) \]
\[ (\tilde{G}_V)_\sigma := \sum_{K \in T_{\sigma}} (g, \psi_\sigma)_K, \quad (4.7b) \]
\[ (J_V)_{\sigma, K} := -(S \nabla \varphi_K, \nabla \psi_\sigma)_K. \quad (4.7c) \]

So far, we have not specified the choice of the points \( z_K \). Three different choices of \( z_K \) have been already studied in the literature; however in none of the approaches an abstract framework was provided.

### 4.1.1. \( S \)-circumcenter as the evaluation point

Problem (4.5) was first studied in Refs. 59, 20 for two space dimensions and in Ref. 57 for particular meshes in three space dimensions. One supposes here the existence of a point \( z_K \) such that \( S_{K} \nabla \varphi_\gamma \cdot \nabla \psi_\sigma = 0 \) and \( S_{K} \nabla \varphi_\gamma \cdot \nabla \psi_\gamma = 0 \) for all \( K \in T_h \), with the notation of the top part of Fig. 3, which leads to diagonal matrices \( M_V \). Such a point always exists in two space dimensions and is called the \( S \)-circumcenter of \( K \); the whole mesh \( T_h \) then becomes “\( S \)-orthogonal grid” in the terminology of Ref. 1.

When \( S_{\mid K} = \| s_K \), \( z_K \) gets the circumcenter of \( K \). The bottom part of Fig. 3 gives an illustration for a diffusion tensor given by \( S = \begin{pmatrix} 0.7236 & 0.3804 \\ 0.3804 & 0.4764 \end{pmatrix} \). In the approach of this section, no local linear system needs to be solved and one always obtains a two-point flux expression. The choice of the evaluation point \( z_K \) depends on the diffusion tensor \( S \) but not on the local mesh form. Let \( d = 2 \) and consider a patch \( T_V \) as in Sec. 4.1 and a side \( \sigma \in \mathcal{E}_V^{\text{int}} \). Then, from (4.5), we get
\[ \Lambda_\sigma = \frac{(g, \psi_\sigma)_K + (g, \psi_\sigma)_L - \bar{P}_K (S \nabla \varphi_K, \nabla \psi_\sigma)_K - \bar{P}_L (S \nabla \varphi_L, \nabla \psi_\sigma)_L}{(S \nabla \varphi_\sigma, \nabla \psi_\sigma)_K + (S \nabla \varphi_\sigma, \nabla \psi_\sigma)_L}, \quad (4.8) \]
By the same reasoning as in Sec. 4.1, we come to
\[
(u_h \cdot n_K, 1)_\sigma = -\Lambda_\sigma (S \nabla \varphi, \nabla \psi)_K + (g, \psi)_K - \bar{P}_K (S \nabla \varphi, \nabla \psi)_K. \tag{4.9}
\]
We now use the fact that
\[
S|K \nabla \varphi_K \cdot \nabla \psi = -S|K \varphi_K \cdot \nabla \psi, \tag{4.10}
\]
which follows from the basis functions orthogonalities and from the fact that \(\varphi, \varphi, \gamma\), and \(\varphi_K\) form a partition of unity on \(K\). Denote
\[
\alpha_{K, \sigma} := (S \nabla \varphi, \nabla \psi)_K. \tag{4.11}
\]
It can be checked that this expression does not depend, for a fixed \(K \in T_h\), and \(\sigma \in \mathcal{E}_K\), on the vertex \(V\) of \(\sigma\); (4.11) coincides with Ref. 57 (Eq. (31)). If \(S|K = \|s_K\|\), it follows that:
\[
\alpha_{K, \sigma} = \text{sgn}(z_K) \frac{s_K|\sigma|}{|z_K - x_\sigma|},
\]
where \(\text{sgn}(z_K) = 1\) if \(z_K \in K\) and \(-1\) otherwise. Using the analogous notation \(\alpha_{L, \sigma}\) for the element \(L\), (4.8) can be rewritten as:
\[
\Lambda_\sigma = \frac{(g, \psi)_K + (g, \psi)_L + \bar{P}_K \alpha_{K, \sigma} + \bar{P}_L \alpha_{L, \sigma}}{\alpha_{K, \sigma} + \alpha_{L, \sigma}}. \tag{4.12}
\]
For \(g = 0\), this is the standard two-point FV formula, cf. the first equation on p. 13 of Ref. 31. Note however that, as discussed in Ref. 59 (Sec. 4), this is no more the case when \(g \neq 0\), and the coefficients \(\alpha_{K, \sigma}\) can be negative. Inserting the expression for \(\Lambda_\sigma\) from (4.12) to (4.9), we obtain
\[
(u_h \cdot n_K, 1)_\sigma = \frac{\alpha_{K, \sigma} \alpha_{L, \sigma}}{\alpha_{K, \sigma} + \alpha_{L, \sigma}} (\bar{P}_K - \bar{P}_L) + (g, \psi)_K
\]
\[
- \frac{\alpha_{K, \sigma}}{\alpha_{K, \sigma} + \alpha_{L, \sigma}} ((g, \psi)_K + (g, \psi)_L), \tag{4.13}
\]
i.e. the formula (34) from Ref. 57. When \(g = 0\) and \(S\) is scalar, (4.13) is nothing but the standard two-point FV flux with harmonic averaging of the diffusion coefficient, cf. the second equation on p. 13 of Ref. 31. The influence of the source term \(g\) only disappears in very particular situations; a sufficient condition is, for example, \(S = 1\) and \(g|\mathcal{E}_h| \|z_K - x_\sigma\| \text{sgn}(z_K) = g|L| \|z_L - x_\sigma\| \text{sgn}(z_L)\). For \(\sigma \in \mathcal{E}_h^{\text{ext}}\), one similarly obtains
\[
(u_h \cdot n_K, 1)_\sigma = \alpha_{K, \sigma} \bar{P}_K + (g, \psi)_K. \tag{4.14}
\]

The present approach can degenerate. This happens when \(z_K\) coincides with one of the side barycenters \(x_\sigma\), as illustrated in Fig. 4, left part. Then, with the notation of Fig. 2, the measure of the subtriangle \(L\) becomes 0, \(|L| = 0\). Consequently, one obtains \(|\nabla \varphi| = \infty\). This situation requires a specific treatment, see Sec. 4.4.1 below.
4.1.2. Barycenter as the evaluation point

The choice of the barycenter $z_K = x_K$, cf. the bottom part of Fig. 3, is related to the approach studied in Ref. 53. It turns out that this choice allows for a wider variety of meshes for which the local problems (4.6) are well-posed and that it works in all space dimensions. The matrix $M_V$ is, in general, not diagonal. This leads to the necessity to solve local linear systems and to a multi-point flux expression where the fluxes depend on the potentials $\bar{P}_K$ of the whole patch $T_V$. The choice of the evaluation point $z_K$ depends neither on the diffusion tensor $S$, nor on the local mesh form. This approach can, however, lead to a singular local condensation matrix $M_V$. This happens, e.g. for $S = I$ and the patch $T_V$ illustrated in Fig. 4, right part.

In Ref. 53, one expresses $u_h|_{T_V}$ directly from the original unknowns $P_K := p_h|_K$, $K \in T_V$. When $g = 0$, we have from (3.7b) $P_K = \bar{P}_K = \lambda^{NC}(z_K)$ and thus the local problems (4.5) coincide with those given in Ref. 53. When $g \neq 0$, it follows from (3.7b) that $P_K$ and $\bar{P}_K = \lambda^{NC}(z_K)$ only differ by $\frac{\partial P_K}{\partial \lambda^{NC}}(S^{-1}(x - x_K), x - x_K)_K$. Consequently, the local problems (4.5) and those of Ref. 53 can still be written in the form (4.6) with the same local matrices $M_V$ given by (4.7a) and only differ by the vectors $\bar{P}_V$ and the right-hand side.

Remark 4.2. (Singular matrices in the MPFA method) Recall from Refs. 37, 58 that the R10 MFE with one unknown per element and the nonsymmetric MPFA O-method are equivalent for $d = 2$ and $g = 0$, cf. Remark 3.2. Thus, this variant of the MPFA also gives rise to a singular matrix $M_V$ for the mesh of Fig. 4, right part.

4.1.3. Mesh- and diffusion tensor-dependent evaluation point

In the framework of the CR NCFE method, a new idea has been proposed in Ref. 55 for the solution of the local problems (4.5). It consists in choosing the evaluation point according to the mesh $T_h$ and the diffusion tensor $S$. The choice of the points $z_K$ is done locally in order to: (a) ensure the well-posedness of the local
problems (4.5); (b) influence the properties of the local matrices $M_V$; (c) influence the properties of the final global system matrices (see Secs. 4.4.1 and 4.4.2 below). We refer to Refs. 31, 2 and 27 for some other examples of locally influencing the final system properties.

4.2. Definition of the local problems by an algebraic interpretation

We now generalize the approach of Sec. 4.1 in a purely algebraic way, following Ref. 55.

4.2.1. Potentials viewpoint

Let $K \in T_h$, let $\Lambda_K = \{\Lambda_\gamma\}_{\gamma \in \mathcal{E}_K^\text{int}}$ be the vector of the Lagrange multipliers associated with the interior sides of the element $K$, and let $N_K$ be a $1 \times |\mathcal{E}_K^\text{int}|$ matrix. We define the vector $\bar{P}_K$ in a purely algebraic way by

$$N_K \Lambda_K = \bar{P}_K. \quad (4.15)$$

Whenever additionally the constraint

$$\sum_{\sigma \in \mathcal{E}_K} (N_K)_{\sigma} = 1 \quad (4.16)$$

holds for all interior mesh elements, (4.15) allows for a geometrical interpretation in the sense of Sec. 4.1: then the entries $(N_K)_{\sigma}$ take the values of the CR basis functions $\psi_\sigma$ at the evaluation points $z_K$. As an example, in Sec. 4.1.2, all entries of $N_K$ are equal to $1/(d+1)$, whereas the matrix $N_K$ for Sec. 4.1.1 is specified in Ref. 57 (Eq. (30)).

Now consider a vertex $V \in V_h$ and (4.15) on all elements $K \in T_V^\text{ext}$ (recall that Fig. 1 gives an illustration of the different sets). This gives $|T_V^\text{ext}| = |\mathcal{E}_V^\text{ext}|$ equations involving the unknowns $\Lambda_V := \{\Lambda_\gamma\}_{\gamma \in \mathcal{E}_V}$, with the matrix $N_V$ employing the element matrices $N_K$:

$$N_V \Lambda_V = \bar{P}_V^\text{ext}. \quad (4.17)$$

Consider next the lines in (3.3) associated with $\gamma \in \mathcal{E}_V^\text{int}$. Denoting $Z_V$ the corresponding submatrix of $Z$ and setting $E_V^\text{int} := E_{\gamma \in \mathcal{E}_V^\text{int}}$, this gives a second rectangular linear system, with the number of rows given by $|\mathcal{E}_V^\text{int}|$ and the number of columns is given by $|\mathcal{E}_V^\text{int}|$,

$$Z_V \Lambda_V = E_V^\text{int}. \quad (4.18)$$

Combining (4.17) and (4.18), we obtain a square linear system:

$$\begin{pmatrix} Z_V & \lambda_V \end{pmatrix} \begin{pmatrix} \Lambda_V \\ \bar{P}_V^\text{ext} \end{pmatrix} = \begin{pmatrix} \lambda_V \\ \bar{P}_V^\text{ext} \end{pmatrix}. \quad (4.19)$$
As the equations of (4.18) are taken from the well-posed system (3.3), they are linearly independent. Thus the well-posedness of (4.19) only depends on the matrix $N$. Clearly, in a still more general fashion than in Sec. 4.1.3, the well-posedness of (4.19) can be controlled in function of the local constellation (mesh $T_h$, tensor $S$).

### 4.2.2. Fluxes viewpoint

Starting from the hybridized formulation (3.2) instead of (3.3) and following the same general idea as in Sec. 4.2.1, we can also proceed as follows: consider a vertex $V \in V_h$ and the lines in (3.2) associated with all the elements of the patch $T_V$ and their sides for the first block row, all the elements of the patch $T_V$ for the second block row, and all the interior sides of the patch $T_V$ for the last block row. Set $U^H_V$ for all the elements of the patch $T_V$ and their sides, $P_V$ for all the elements of the patch $T_V$, and $\Lambda_V$ for the sides from $E_V$. This gives the following local rectangular linear system:

$$
\begin{pmatrix}
A^H_V (B^H_V)^t & C^t_V \\
B^H_V & 0 & 0 \\
C_V & 0 & 0
\end{pmatrix}
\begin{pmatrix}
U^H_V \\
P_V \\
\Lambda_V
\end{pmatrix}
= 
\begin{pmatrix}
F^H_V \\
G_V \\
0_V
\end{pmatrix}.
$$

(4.20)

Combining (4.20) with (4.17) gives a local square linear system with $(d + 1)|T_V| + |T_V| + |E^\text{int}_V| + |T^\text{ext}_V|$ equations and unknowns

$$
\begin{pmatrix}
A^H_V (B^H_V)^t & C^t_V \\
B^H_V & 0 & 0 \\
0 & 0 & N_V
\end{pmatrix}
\begin{pmatrix}
U^H_V \\
P_V \\
\Lambda_V
\end{pmatrix}
= 
\begin{pmatrix}
F^H_V \\
G_V \\
0_V
\end{pmatrix}.
$$

(4.21)

Thus we can also obtain the fluxes $U^H_V$ in the whole patch $T_V$ as a function of the new potentials $\bar{P}^\text{ext}_V$.

### 4.3. Recovery of $\Lambda$ and $U^H$ in terms of $\bar{P}$

Let $V \in V_h$ be a vertex and $T_V$ be the corresponding patch. By any of the approaches of Secs. 4.1–4.2, if the corresponding local problems (4.6), (4.19), or (4.21) are well-posed, we obtain local expressions of the Lagrange multipliers $\Lambda^\text{int}_V$ ($\Lambda_V$) and/or local expressions of the fluxes $U^H_V$ from the potentials $P_V$ (and sources in $T_V$). In particular, we infer from (4.6)

$$
\Lambda^\text{int}_V = (M_V)^{-1}(\bar{G}_V - J_V \bar{P}_V)
$$

(4.22)

and from (4.19)

$$
\Lambda_V = \left(\begin{array}{c}
Z_V \\
N_V
\end{array}\right)^{-1} \left(\begin{array}{c}
E^\text{int}_V \\
\bar{P}^\text{ext}_V
\end{array}\right).
$$

(4.23)
Recall that the system matrices in (4.22) and (4.23) are square. A similar expression can be obtained for $U_V^H$ from all (4.6), (4.19), or (4.21).

We now give details on the construction (4.6). Consider (4.22) and run through all vertices of the mesh $T_h$. For every vertex $V$, we have one expression for $\Lambda_{\text{int}}^V$. All these different expressions have to lead to the same values of the Lagrange multipliers $\Lambda$, since the vector $\Lambda$ is the unique solution of (3.3). Thus, we are free to associate a weight $w_{V,\sigma}$ to the expression of $\Lambda^V_{\sigma}$ from every patch $T_V$ where the side $\sigma$ is such that $\Lambda^V_{\sigma}$ is the unknown in the local problem, and combine these expressions with these weights. The only condition is that, for every side $\sigma$, the sum of all its weights is equal to one. Let $V \in V_h$. We define a mapping $\Upsilon_V : \mathbb{R}^{|E|_{\text{int}}^V} \rightarrow \mathbb{R}^{|E|_{\text{int}}^h}$, extending a vector $\Lambda_{\text{int}}^V = \{\Lambda_\sigma\}_{\sigma \in E_{\text{int}}^V}$ of values associated with the sides from $E_{\text{int}}^V$ to a vector of values associated with all the interior sides $E_{\text{int}}^h$ by

$$[\Upsilon_V(\Lambda_{\text{int}}^V)]_{\sigma} = \begin{cases} \Lambda_\sigma & \text{if } \sigma \in E_{\text{int}}^V, \\ 0 & \text{if } \sigma \not\in E_{\text{int}}^V. \end{cases}$$

Next let $W_V$ be a diagonal matrix of size $|E|_{\text{int}}^V \times |E|_{\text{int}}^V$, with the entries given by the weights $w_{V,\sigma}$. With these notations, we have

$$\sum_{V \in V_h} \Upsilon_V(W_V \Lambda_{\text{int}}^V) = \Lambda. \quad (4.24)$$

We now introduce a mapping $\Upsilon_V : \mathbb{R}^{|E|_{\text{int}}^V \times |E|_{\text{int}}^V} \rightarrow \mathbb{R}^{|E|_{\text{int}}^h \times |E|_{\text{int}}^h}$ (with the same name as the previous one, since there is no possibility of confusion), extending a local matrix $M_V$ to a full-size one by zeros by

$$[\Upsilon_V(M_V)]_{\sigma,\gamma} := \begin{cases} (M_V)_{\sigma,\gamma} & \text{if } \sigma \in E_{\text{int}}^V \text{ and } \gamma \in E_{\text{int}}^V, \\ 0 & \text{if } \sigma \not\in E_{\text{int}}^V \text{ or } \gamma \not\in E_{\text{int}}^V. \end{cases}$$

We finally define a mapping $\Theta_V : \mathbb{R}^{|E|_{\text{int}}^V \times |T_V|} \rightarrow \mathbb{R}^{|E|_{\text{int}}^h \times |T_h|}$, filling a full-size representation of a matrix $J_V$ by zeros on the rows associated with the sides that are not from $E_{\text{int}}^V$ and on the columns associated with the elements that are not from $T_V$,

$$[\Theta_V(J_V)]_{\sigma,K} := \begin{cases} (J_V)_{\sigma,K} & \text{if } \sigma \in E_{\text{int}}^V \text{ and } K \in T_V, \\ 0 & \text{if } \sigma \not\in E_{\text{int}}^V \text{ or } K \not\in T_V. \end{cases}$$

With these notations, we obtain from (4.22),

$$\Upsilon_V(W_V \Lambda_{\text{int}}^V) = \Upsilon_V(W_V (M_V)^{-1}(\tilde{G}_V - J_V \tilde{P}_V))$$

$$= \Upsilon_V(W_V (M_V)^{-1})\tilde{G} - \Theta_V(W_V (M_V)^{-1}J_V)\tilde{P}.$$

Now, employing (4.24), we finally come to

$$\Lambda = \tilde{M}^\text{inv} \tilde{G} - \tilde{M}^\text{inv} \tilde{P}, \quad (4.25)$$
with
\[ \tilde{M}_{\text{inv}} := \sum_{V \in V_h} \Upsilon_V (\mathcal{W}_V (M_V)^{-1}), \quad M_{\text{inv}} := \sum_{V \in V_h} \Theta_V (\mathcal{W}_V (M_V)^{-1} \mathcal{J}_V). \]

A similar procedure as above can be applied for the local problems (4.19) or (4.21).

In the same way, in (4.6), (4.19), or (4.21), the fluxes can be expressed. We then arrive at the equivalent of (4.25) in the form (using herein \( F^H = 0 \) for simplicity)
\[ \tilde{U}^H = \tilde{O}_{\text{inv}} G - O_{\text{inv}} \bar{P}. \]  

Let us stress that \( M_{\text{inv}}, \tilde{M}_{\text{inv}}, O_{\text{inv}}, \) and \( \tilde{O}_{\text{inv}} \) are fully computable and sparse matrices, obtained by a weighted combination of the inverses of the local matrices. We would finally like to mention here that in the numerical experiments of Sec. 5, we only use the approach of Sec. 4.1, with all the weights \( w_{V,\sigma} \) of Sec. 4.3 equal to \( 1/d \).

4.4. Prescribing the final system for the potentials \( \bar{P} \) only

We now want to use the above developments in order to write a global system of the form (1.7). It turns out that two different approaches can be used.

4.4.1. Using the equilibrium of the fluxes

The first possibility is to insert (4.26) into the second block equation of (3.2) which yields
\[ -B^H O_{\text{inv}} \bar{P} = G - B^H \tilde{O}_{\text{inv}} G, \]  
i.e. (1.7) with the square matrix \( S := -B^H O_{\text{inv}} \) and the right-hand side vector \( H := G - B^H \tilde{O}_{\text{inv}} G \). Remark that here, the second block equation of (3.2) is used repeatedly.

The approach of Sec. 4.1.1 allows one to state the matrix \( S \) explicitly. Fix \( K \in T_h \) for \( d = 2 \) and recall that (4.11), for \( \sigma \in \mathcal{E}_K \cap \mathcal{E}_h^{\text{int}} \), takes the same values when expressed from the two patches such that \( \sigma \in \mathcal{E}_V^{\text{int}} \). Thus, the weights \( w_{V,\sigma} \) of Sec. 4.3 have no influence in this case and can be chosen arbitrarily. Plugging (4.13)–(4.14) into \( (\nabla \cdot u_h, 1)_K = \sum_{\sigma \in \mathcal{E}_K} (u_h \cdot n_K, 1) = (g, 1)_K \) then gives (1.7) with:
\[ S_{K,L} := -\frac{\alpha_{K,\sigma} \alpha_{L,\sigma}}{\alpha_{K,\sigma} + \alpha_{L,\sigma}}, \]  
\[ S_{K,K} := \sum_{\sigma \in \mathcal{E}_K^\sigma} \frac{\alpha_{K,\sigma} \alpha_{L,\sigma}}{\alpha_{K,\sigma} + \alpha_{L,\sigma}} + \sum_{\sigma \in \mathcal{E}_K^{\text{int}} \cap \mathcal{E}_K} \alpha_{K,\sigma}, \]  
\[ H_K := \sum_{\sigma \in \mathcal{E}_K^\sigma} \frac{\alpha_{K,\sigma}}{\alpha_{K,\sigma} + \alpha_{L,\sigma}} ((g, \psi)_K + (g, \psi)_L). \]

As observed in Refs. 59, 20, and 57, for \( g = 0 \), Eqs. (4.28a)–(4.28c) coincide with the standard two-point FV scheme. 29, 31 The final matrix \( S \) is symmetric and has a
(\(d + 2\))-point stencil, that is, there are at most \(d + 2\) nonzero entries per each row of \(S\). Let \(d = 2\) and \(S|_K = \|s_K\). Then \(S\) is positive definite on Delaunay meshes (a Delaunay mesh is such that the closure of the circumcircle of each simplex does not contain any other simplex vertex) but indefinite otherwise. A critical situation arises, as outlined in Sec. 4.1.1, when two triangles cut a square, see Fig. 4, left part. The local matrices \(M_V\) degenerate and the final problem (4.27) is not well-posed. In order to proceed, an “aggregation” of the two triangles into a square has to be done, leading to one final unknown for each aggregated pair of triangles. We refer for details to Refs. 59, 20, and 57. This approach is called in Sec. 5 below the FV method.

In the approach of Sec. 4.1.2, the final matrix \(S\) is in general nonsymmetric and has a wider stencil (for each \(K \in T_h\), all simplices sharing a node with \(K\) are involved). The family of meshes where \(S\) is positive definite is, however, larger in comparison with the previous case. Recall that, similarly as in the previous case, singular local matrices \(M_V\) can appear, see Sec. 4.1.2. We refer for details to Ref. 53. This approach is called in Sec. 5 below the CMFE method.

### 4.4.2. Using the potential relation

Alternatively, we can insert (4.25) into

\[
N\Lambda = \bar{P},
\]  

(4.29)

where \(N\) is the \(|T_h| \times |\mathcal{E}_{\text{int}}^h|\) matrix with rows formed by the element matrices \(N_K\) of (4.15). This gives

\[
NM^{\text{inv}}\bar{P} + \bar{P} = NM^{\text{inv}}\tilde{G},
\]  

(4.30)

i.e. (1.7) with the square matrix \(S := NM^{\text{inv}} + I\) and right-hand side vector \(H := NM^{\text{inv}}\tilde{G}\). Remark that here, (4.29) is used repeatedly. This approach is studied in detail in Ref. 55.

Let \(d = 2\), fix \(K \in T_h\), and consider the approach of Sec. 4.1.1. Recall (4.10), the notation (4.11), and plug (4.12) into (4.1). This gives (1.7) with:

\[
S_{K,L} := -\frac{\psi_\sigma(z_K)\alpha_{L,\sigma}}{\alpha_{K,\sigma} + \alpha_{L,\sigma}},
\]

(4.31a)

\[
S_{K,K} := 1 - \sum_{\sigma \in \mathcal{E}_{\text{int}}^h} \frac{\psi_\sigma(z_K)\alpha_{K,\sigma}}{\alpha_{K,\sigma} + \alpha_{L,\sigma}},
\]

(4.31b)

\[
H_K := \sum_{\sigma \in \mathcal{E}_{\text{int}}^h} \frac{\psi_\sigma(z_K)}{\alpha_{K,\sigma} + \alpha_{L,\sigma}}((g,\psi_\sigma)_K + (g,\psi_\sigma)_L).
\]

(4.31c)

The matrix \(S\) still has a \((d + 2)\)-point stencil and its positive definiteness and well-posedness depend on the mesh \(T_h\) and the diffusion tensor \(\mathbf{S}\) similarly as in Sec. 4.4.1. Likewise, the weights \(w_{V,\sigma}\) of Sec. 4.3 have no influence in this case and can be chosen arbitrarily. In contrast to Sec. 4.4.1, however, \(S\) is in general nonsymmetric.
S becomes symmetric if the mesh is “symmetric”, consisting of the elements with the same shape, and for $S = I$, as the numerical experiments in Ref. 55 indicate. Then it coincides with that of Sec. 4.4.1. This approach is called in Sec. 5 below the MFEC method.

The other approaches still lead to a nonsymmetric matrix as in Sec. 4.4.1. The stencil involves all simplices sharing a node with a given $K \in T_h$. Positive definiteness and well-posedness depend again on the mesh $T_h$ and the diffusion tensor $S$. The approaches of Secs. 4.1.2 and 4.1.3, respectively, are called the MFEB and MFEO methods in Sec. 5 below.

5. Numerical Experiments

The goal of this section is to carry out a comparative numerical study of the different one-unknown-per-element reformulations of the RT0 MFE method, as presented in Sec. 4. We study the behavior of the different approaches for the homogeneous and isotropic diffusion tensor in Sec. 5.1, for an anisotropic diffusion tensor in Sec. 5.2, and for an inhomogeneous diffusion tensor in Sec. 5.3.

We consider the problem (1.1a) on $\Omega = (0, 1) \times (0, 1)$, with inhomogeneous Dirichlet boundary condition given by the function $p(x, y) = 0.1 y + 0.9$ instead of (1.1b). We perform the calculations on uniform refinements of the mesh viewed in Fig. 5. This mesh is Delaunay, with the minimal and maximal angles equal to 35.4 and 88.7 degrees, respectively. A sink term $g = -0.001$ is prescribed on two elements of the initial mesh. We consider the tensor $S$ in the form

$$
S|_K = \begin{pmatrix}
\cos(\theta_K) & -\sin(\theta_K) \\
\sin(\theta_K) & \cos(\theta_K)
\end{pmatrix}
\begin{pmatrix}
s_K & 0 \\
0 & \nu s_K
\end{pmatrix}
\begin{pmatrix}
\cos(\theta_K) & \sin(\theta_K) \\
-\sin(\theta_K) & \cos(\theta_K)
\end{pmatrix}
$$

for $K \in T_h$, where $\theta_K$ is the angle between the normal to the mesh and the $x$-axis.

Fig. 5. Initial mesh; the coloring indicates the permeability tensor $S$ in cases (5.2) and (5.3).
where we distinguish the following three different forms:

\[ s_K = 1, \quad \forall K \in T_h, \quad \nu = 1, \]  

(5.1)

i.e. the homogeneous isotropic case \((\mathbf{S} = \mathbb{I})\), or

\[ s_K = 1, \quad \forall K \in T_h, \quad \theta_K \in \left\{ \frac{\pi}{5}, \frac{3\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{5}, \frac{\pi}{3} \right\}, \quad \nu = 0, \]  

(5.2)

i.e. the homogeneous, with respect to \(s_K\), but anisotropic case \((\mathbf{S} \text{ is a full-matrix tensor})\), or

\[ s_K \in \{10, 1, 0.1, 0.01, 0.001\}, \quad \nu = 1, \]  

(5.3)

i.e. the inhomogeneous isotropic case \((\mathbf{S} \text{ is a varying multiple of the identity matrix})\). The different gray shades in Fig. 5 correspond to the different choices \(\theta_K\) and \(s_K\) in (5.2) and (5.3), respectively. All the computations were performed in double precision on a notebook with Intel Core2 Duo 2.6 GHz processor and MS Windows Vista operating system. Machine precision was in the power of 1e-16. All the linear system solutions were done with the help of MATLAB 7.0.4.

We test the following methods, which are all equivalent implementations of the RT0 MFE method (1.2a)–(1.2b):

1. **MFEB**: the final system is of the form (1.7), imposed through (4.30), with the barycenter as the evaluation point, cf. Sec. 4.1.2;
2. **MFEC**: the final system is of the form (1.7), imposed through (4.30), with the \(\mathbf{S}\)-circumcenter as the evaluation point, cf. Sec. 4.1.1 (the problem is given by (4.31a)–(4.31c));
3. **MFEO**: the final system is of the form (1.7), imposed through (4.30), while choosing the evaluation point as a function of \(T_h\) and \(\mathbf{S}\), cf. Sec. 4.1.3;
4. **CMFE**: the final system is of the form (1.7), given by the approach of Ref. 53; considering the discrete unknowns \(\mathbf{P}_K\) of (4.1) with the barycenter as the evaluation point, cf. Sec. 4.1.2, the system matrix \(\mathbf{S}\) coincides with that of (1.7) imposed through (4.27);
5. **FV**: the final system is of the form (1.7), imposed through (4.27), with the \(\mathbf{S}\)-circumcenter as the evaluation point, cf. Sec. 4.1.1 (the problem is given by (4.28a)–(4.28c)); this is the approach of Refs. 59, 20, 57 and corresponds to the two-point FV method;
6. **NCFE**: the final system is of the form (3.3), imposed through (3.5), i.e. through the CR NCFE method, see Sec. 3.3.

We show in Fig. 6 the sparsity patterns of the original MFE method (1.3) and of its equivalent reformulations of the above list, for the case of the mesh of Fig. 5. Recall that the matrix sizes are, respectively, the number of mesh elements plus mesh faces, the number of mesh interior faces, and the number of mesh elements. In Tables 2–5, we present various properties of the final matrix systems arising from the different equivalent reformulations. We summarize in Table 1 the abbreviations
used in these tables. Recall that a real matrix $S \in \mathbb{R}^{M \times M}$ is positive definite if $P^tSP > 0$ for all $P \in \mathbb{R}^M$, $P \neq 0$, and negative stable when all its eigenvalues have positive real parts (this is in particular the case for positive definite matrices). The 2-norm condition number of a matrix $S$ is defined by $\|S\|_2\|S^{-1}\|_2$. We also consider the 2-norm condition number after diagonal scaling, by which we mean the minimal of the two 2-norm condition numbers of the two matrices

$$(\text{diag}(S))^{-1/2}S, \quad |\text{diag}(S)|^{-1/2}S|\text{diag}(S)|^{-1/2}.$$  

We also study the computational cost. To do so, we restrict ourselves to standard MATLAB routines and test direct and iterative solvers. We first test the \ direct
Table 1. Abbreviations used in Tables 2–5.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meth.</td>
<td>Method, one of the equivalent MFE formulations</td>
</tr>
<tr>
<td>Un.</td>
<td>Number of unknowns (matrix size)</td>
</tr>
<tr>
<td>Mat.</td>
<td>Matrix</td>
</tr>
<tr>
<td>St.</td>
<td>Stencil (the maximum number of nonzero entries on each matrix row)</td>
</tr>
<tr>
<td>Nonz.</td>
<td>Total number of matrix nonzero entries</td>
</tr>
<tr>
<td>CN</td>
<td>2-norm condition number</td>
</tr>
<tr>
<td>CNS</td>
<td>2-norm condition number after diagonal scaling</td>
</tr>
<tr>
<td>DS</td>
<td>Direct linear solver</td>
</tr>
<tr>
<td>CG</td>
<td>Conjugate gradients iterative linear solver</td>
</tr>
<tr>
<td>PCG</td>
<td>Preconditioned conjugate gradients iterative linear solver</td>
</tr>
<tr>
<td>Bi-CGStab</td>
<td>Bi-conjugate gradients stabilized iterative linear solver</td>
</tr>
<tr>
<td>PBI-CGStab</td>
<td>Preconditioned bi-conjugate gradients stabilized iterative linear solver</td>
</tr>
<tr>
<td>CPU</td>
<td>CPU time of a direct/iterative linear solver</td>
</tr>
<tr>
<td>Iter.</td>
<td>Number of iterations of an iterative linear solver</td>
</tr>
<tr>
<td>IC</td>
<td>CPU time of incomplete Cholesky factorization with a specified drop tolerance</td>
</tr>
<tr>
<td>ILU</td>
<td>CPU time of incomplete LU factorization with a specified drop tolerance</td>
</tr>
<tr>
<td>SPD</td>
<td>Symmetric positive definite</td>
</tr>
<tr>
<td>SID</td>
<td>Symmetric indefinite</td>
</tr>
<tr>
<td>NPD</td>
<td>Nonsymmetric positive definite</td>
</tr>
<tr>
<td>NNS</td>
<td>Nonsymmetric negative stable</td>
</tr>
<tr>
<td>NID</td>
<td>Nonsymmetric indefinite</td>
</tr>
</tbody>
</table>

solver. Such a solver may not be usable for very large systems or may not be suitable for parabolic or nonlinear problems. Thus the behavior of iterative solvers is also very important. We test two iterative methods. If the matrix is symmetric and positive definite, we use the conjugate gradients method. For nonsymmetric matrices, we employ the bi-conjugate gradients stabilized method. Unpreconditioned iterative linear solvers may be rather slow but usually illustrate well the matrix properties and especially the matrix condition number. To accelerate their convergence, we use incomplete Cholesky and incomplete LU factorizations with a specified drop tolerance, cf. Ref. 49. The drop tolerance is always chosen in such a way that the sum of CPU times of the preconditioning and of the solution of the preconditioned system was minimal. We always use a zero start vector and stop the iterative process as soon as the relative residual $\|H - \tilde{P}\|_2/\|H\|_2$, where $\tilde{P}$ is the approximate solution to the system $SP = H$, decreases below $1e^{-8}$.

Some general conclusions may be drawn from the theoretical investigations and from Tables 2–5. NCFE always produces a symmetric positive definite matrix. The FV matrix is always symmetric, but positive definiteness depends on the tensor $S$ and on the mesh $T_h$. All the other methods lead to nonsymmetric matrices. These matrices can be positive definite, negative stable, or indefinite, depending on $S$ and $T_h$. MFEC and FV lead
Table 2. Matrix properties and computational cost of the different equivalent formulations of the MFE method, identity matrix diffusion tensor (5.1), fourth-level mesh refinement.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MFEB</td>
<td>13824</td>
<td>NPD</td>
<td>14</td>
<td>177652</td>
<td>7564</td>
<td>7580</td>
<td>0.27</td>
<td>4.86</td>
<td>324.5</td>
<td>0.81</td>
<td>0.36</td>
<td>9.0</td>
<td></td>
</tr>
<tr>
<td>MFEC</td>
<td>13824</td>
<td>NNS</td>
<td>4</td>
<td>55040</td>
<td>11256</td>
<td>11056</td>
<td>0.09</td>
<td>2.23</td>
<td>372.0</td>
<td>0.42</td>
<td>0.19</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>MFEO</td>
<td>13824</td>
<td>NPD</td>
<td>14</td>
<td>177652</td>
<td>7531</td>
<td>7558</td>
<td>0.26</td>
<td>4.98</td>
<td>270.0</td>
<td>0.80</td>
<td>0.41</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>CMFE</td>
<td>13824</td>
<td>NPD</td>
<td>14</td>
<td>177652</td>
<td>7397</td>
<td>7380</td>
<td>0.27</td>
<td>4.70</td>
<td>312.0</td>
<td>0.83</td>
<td>0.39</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>FV</td>
<td>13824</td>
<td>SPD</td>
<td>4</td>
<td>55040</td>
<td>65722</td>
<td>8898</td>
<td>0.07</td>
<td>3.90</td>
<td>1098.0</td>
<td>0.42</td>
<td>0.17</td>
<td>17.0</td>
<td></td>
</tr>
<tr>
<td>NCFE</td>
<td>20608</td>
<td>SPD</td>
<td>5</td>
<td>102528</td>
<td>14064</td>
<td>9944</td>
<td>0.14</td>
<td>2.92</td>
<td>620.0</td>
<td>1.11</td>
<td>0.56</td>
<td>19.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Matrix properties and computational cost of the different equivalent formulations of the MFE method, identity matrix diffusion tensor (5.1), fifth-level mesh refinement.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MFEB</td>
<td>55296</td>
<td>NPD</td>
<td>14</td>
<td>714740</td>
<td>30289</td>
<td>30373</td>
<td>1.51</td>
<td>33.31</td>
<td>530.0</td>
<td>8.14</td>
<td>3.37</td>
<td>19.5</td>
<td></td>
</tr>
<tr>
<td>MFEC</td>
<td>55296</td>
<td>NNS</td>
<td>4</td>
<td>220672</td>
<td>45016</td>
<td>44643</td>
<td>0.49</td>
<td>17.47</td>
<td>642.5</td>
<td>3.34</td>
<td>1.25</td>
<td>13.5</td>
<td></td>
</tr>
<tr>
<td>MFEO</td>
<td>55296</td>
<td>NPD</td>
<td>14</td>
<td>714740</td>
<td>30143</td>
<td>30285</td>
<td>1.55</td>
<td>37.76</td>
<td>594.5</td>
<td>7.06</td>
<td>2.75</td>
<td>19.5</td>
<td></td>
</tr>
<tr>
<td>CMFE</td>
<td>55296</td>
<td>NPD</td>
<td>14</td>
<td>714740</td>
<td>29630</td>
<td>29551</td>
<td>1.49</td>
<td>34.32</td>
<td>544.5</td>
<td>6.91</td>
<td>2.73</td>
<td>18.5</td>
<td></td>
</tr>
<tr>
<td>FV</td>
<td>55296</td>
<td>SPD</td>
<td>4</td>
<td>220672</td>
<td>263036</td>
<td>35813</td>
<td>0.48</td>
<td>29.52</td>
<td>2170.0</td>
<td>3.08</td>
<td>1.47</td>
<td>20.0</td>
<td></td>
</tr>
<tr>
<td>NCFE</td>
<td>82688</td>
<td>SPD</td>
<td>5</td>
<td>412416</td>
<td>56416</td>
<td>39901</td>
<td>0.90</td>
<td>35.80</td>
<td>1219.0</td>
<td>9.41</td>
<td>3.78</td>
<td>42.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Matrix properties and computational cost of the different equivalent formulations of the MFE method, anisotropic diffusion tensor (5.2), fourth-level mesh refinement.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MFEB</td>
<td>13824</td>
<td>NPD</td>
<td>14</td>
<td>177652</td>
<td>11203</td>
<td>11203</td>
<td>0.28</td>
<td>6.61</td>
<td>448.0</td>
<td>0.98</td>
<td>0.59</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>MFEC</td>
<td>13824</td>
<td>NID</td>
<td>4</td>
<td>55040</td>
<td>2401279</td>
<td>416769</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.45</td>
<td>0.20</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>MFEO</td>
<td>13824</td>
<td>NPD</td>
<td>14</td>
<td>177652</td>
<td>13401</td>
<td>13401</td>
<td>0.27</td>
<td>6.51</td>
<td>440.5</td>
<td>0.95</td>
<td>0.41</td>
<td>10.0</td>
<td></td>
</tr>
<tr>
<td>CMFE</td>
<td>13824</td>
<td>NPD</td>
<td>14</td>
<td>177652</td>
<td>9276</td>
<td>7758</td>
<td>0.28</td>
<td>5.27</td>
<td>350.5</td>
<td>0.84</td>
<td>0.38</td>
<td>9.0</td>
<td></td>
</tr>
<tr>
<td>FV</td>
<td>13824</td>
<td>SID</td>
<td>4</td>
<td>55040</td>
<td>247055</td>
<td>239934</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.45</td>
<td>0.20</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>NCFE</td>
<td>20608</td>
<td>SPD</td>
<td>5</td>
<td>102528</td>
<td>25393</td>
<td>16969</td>
<td>0.18</td>
<td>4.03</td>
<td>850.0</td>
<td>1.12</td>
<td>0.41</td>
<td>30.0</td>
<td></td>
</tr>
</tbody>
</table>

to the fewest total nonzero matrix entries; NCFE has roughly twice and the other methods have roughly three times as many nonzeros.

5.1. Identity matrix diffusion tensor

Tables 2 and 3 present the results for the coefficients (5.1) and respectively fourth- and fifth-level uniform refinements of the mesh of Fig. 5. The condition number of all
Table 5. Matrix properties and computational cost of the different equivalent formulations of the MFE method, inhomogeneous diffusion tensor (5.3), fourth-level mesh refinement.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MFEB</td>
<td>13824</td>
<td>NPD</td>
<td>14</td>
<td>177652</td>
<td>819248</td>
<td>740706</td>
<td>0.28</td>
<td>13.33</td>
<td>897.5</td>
<td>1.05</td>
</tr>
<tr>
<td>MFEC</td>
<td>13824</td>
<td>NNS</td>
<td>4</td>
<td>55040</td>
<td>903789</td>
<td>763849</td>
<td>0.09</td>
<td>5.34</td>
<td>947.5</td>
<td>0.47</td>
</tr>
<tr>
<td>MFEO</td>
<td>13824</td>
<td>NPD</td>
<td>14</td>
<td>177652</td>
<td>820367</td>
<td>799577</td>
<td>0.28</td>
<td>12.45</td>
<td>790.5</td>
<td>1.05</td>
</tr>
<tr>
<td>CMFE</td>
<td>13824</td>
<td>SPD</td>
<td>4</td>
<td>55040</td>
<td>2900730</td>
<td>478974</td>
<td>0.28</td>
<td>102.27</td>
<td>6842.5</td>
<td>1.01</td>
</tr>
<tr>
<td>FV</td>
<td>13824</td>
<td>SPD</td>
<td>4</td>
<td>55040</td>
<td>16387758</td>
<td>497974</td>
<td>0.07</td>
<td>39.41</td>
<td>14101.0</td>
<td>0.44</td>
</tr>
<tr>
<td>NCFE</td>
<td>20608</td>
<td>SPD</td>
<td>5</td>
<td>102528</td>
<td>4797335</td>
<td>670623</td>
<td>0.18</td>
<td>52.42</td>
<td>11226.0</td>
<td>1.22</td>
</tr>
</tbody>
</table>

methods is here roughly comparable; the only (negative) exception is FV. The condition numbers get mutually much closer after the diagonal scaling. The matrix size and sparsity pattern/number of nonzero entries imply that FV/MFEC give smallest CPU times while using the direct solver, followed by NCFE and then all other methods. MFEC behaves best for an unpreconditioned linear solver, actually much better than FV which do has the advantage of a symmetric matrix. MFEB/MFEC/MFEO proposed in the present paper and the related CMFE seem to outperform NCFE for increasing mesh size (a systematically better behavior of CMFE over NCFE is observed in Ref. 53). Concerning preconditioned iterative solvers, FV/MFEC seem to perform roughly twice as fast as MFEB/MFEO/CMFE and roughly three times as fast as NCFE.

5.2. Anisotropic diffusion tensor

Table 4 presents the results for the coefficients (5.2) and the fourth-level uniform refinement of the mesh of Fig. 5. Because of the anisotropy of the diffusion tensor, FV leads to a symmetric indefinite matrix, whereas MFEC to a nonsymmetric indefinite matrix. These matrices are also very badly conditioned, whereby the diagonal scaling does not help too much. Consequently, direct application of the iterative solvers leads to no convergence in 50,000 iterations. All the other methods behave rather similarly to Table 2. Application of the direct solver or of iterative solvers with preconditioning leads to results similar to that of Table 2.

5.3. Inhomogeneous diffusion tensor

Table 5 presents the results for the coefficients (5.3) and the fourth-level uniform refinement of the mesh of Fig. 5. Here FV gives a symmetric positive definite matrix and MFEC a nonsymmetric negative stable matrix. The inhomogeneity of the diffusion tensor however causes an increase of the matrices condition numbers. This increase is severe in CMFE and NCFE, and in particular in FV. Consequently, direct application of the iterative solvers leads to important increase of the CPU time in CMFE, NCFE, and FV. Whereas FV and MFEC behaved similarly in
Table 2, MFEC becomes here almost eight times faster than the FV one. When diagonal scaling is applied, however, the condition numbers of all methods become comparable, whence the preconditioned iterative solvers behave similarly as for the case of Table 2. Also the application of the direct solver leads to results similar to those of Table 2.

5.4. Discussion

Amongst the above-tested equivalent formulations, the MFEC leads to very compact 4-point stencils in two space dimensions and excellent computational performance for problems with a possibly highly inhomogeneous diffusion tensor. The FV has similar properties and leads, in addition, to symmetric matrices, but gives increased matrix condition numbers for highly inhomogeneous diffusion tensors. Both reformulations, however, seem to behave less well for anisotropic diffusion tensors, and, moreover, are only applicable in two space dimensions. Then the MFEB, MFEO, or the previously proposed CMFE reformulations seem to be appealing alternatives to the classical NCFE implementation.

6. The Discrete Maximum Principle

We recall in this section the different variants of the discrete maximum principle valid for the RT0 MFE method. For this purpose, we replace the homogeneous Dirichlet boundary condition (1.1b) by a nonhomogeneous one,

\[ p = f \text{ on } \partial \Omega, \]

with \( f \in H^{1/2}(\partial \Omega) \).

Recall from Sec. 3.3 the notation \( \psi_\sigma \) for the CR basis function of a side \( \sigma \in E_h \) and the equivalence of the lowest-order mixed and nonconforming elements. Then we have the following classical result; in a more general setting, it has been shown in, e.g. Ref. 34 (Theorem 4.5).

**Theorem 6.1.** (Discrete maximum principle for the Lagrange multipliers \( \lambda_h \)) Let

\[ (S \nabla \psi_\gamma, \nabla \psi_\sigma) \leq 0 \]

(6.1)

for all sides \( \sigma \in E^\text{int}_h \) and all \( \gamma \in E_h \) such that \( \sigma \) and \( \gamma \) lie in the same simplex \( K \in T_h \). Let \( g \geq 0 \) and \( f \geq 0 \). Then \( \lambda_h |_\sigma = \Lambda_\sigma \geq 0 \) for all \( \sigma \in E_h \).

Let \( S \) be such that it can be written as \( I \) times a piecewise constant scalar function and let \( T_h \) be weakly acute, i.e. such that the magnitude of the angles between \( n_{K,\sigma}, \sigma \in E_K \), for all \( K \in T_h \) is greater than or equal to \( \pi/2 \) (all angles smaller than or equal to \( \pi/2 \) in two space dimensions). Then (6.1) holds. Thus, the discrete maximum principle holds for the Lagrange multipliers \( \lambda_h \) of (3.1a)–(3.1c) under this condition. Note that one does not necessarily have the discrete maximum principle for the function \( \lambda_{NC}^h \) defined in (3.6), as this function may take values larger than \( \max_{\sigma \in E_h} \Lambda_\sigma \) and smaller than \( \min_{\sigma \in E_h} \Lambda_\sigma \).
Note that by properties of a simplex and by the fact that $\lambda_{NC}^h$ is an affine function,

$$\lambda_{NC}^h(x_K) = \frac{1}{d+1} \sum_{\sigma \in E_K} \Lambda_\sigma$$

for all $K \in T_h$. Recalling the formula (3.7b) relating $\lambda_{NC}^h$ and $p_h$, the discrete maximum principle also holds for the original piecewise constant approximation $p_h$:

**Theorem 6.2.** (Discrete maximum principle for $p_h$) Let the assumptions of Theorem 6.1 be satisfied. Then $p_h \geq 0$.

Finally, we will employ the link between the RT0 MFE and the two-point FV from Sec. 4.4.1. From the form of the final system (1.7) given in this case by (4.28a)–(4.28c), we have the following.

**Theorem 6.3.** (Discrete maximum principle for the circumcenter values of $\lambda_{NC}^h$) Let $d = 2$ and

$$\frac{\alpha_{K,\sigma} \alpha_{L,\sigma}}{\alpha_{K,\sigma} + \alpha_{L,\sigma}} \geq 0 \quad \forall \sigma \in E_h^{int},$$

$$\alpha_{K,\sigma} \geq 0 \quad \forall \sigma \in E_h^{ext}. \tag{6.2a}$$

Next let $g \geq 0$, $f \geq 0$, and

$$H_K \geq 0 \quad \forall K \in T_h, \tag{6.3}$$

where $H$ is the right-hand side vector from (4.28c). Then $\lambda_{NC}^h|_K(z_K) = \bar{P}_K \geq 0$ for all $K \in T_h$.

Let $d = 2$, let $S = I$ for simplicity, let $T_h$ be Delaunay, and let finally each circumcenter of a boundary triangle be inside $\Omega$. Then Eqs. (6.2a)–(6.2b) hold, cf. Ref. 20. (Extensions to triangulations obtained from square meshes are also possible following Ref. 20.) The requirement (6.3) of course holds when $g = 0$ (when the RT0 MFE and the two-point FV coincide), but otherwise, (6.3) adds an additional restriction on the mesh and the diffusion-dispersion tensor $S$. In any case, under the hypotheses of Theorem 6.3, one also has the discrete maximum principle for the values of $\lambda_{NC}^h$ in the circumcenters of the elements of $T_h$.

7. General Polygonal Meshes and Relations to Other Methods

The aim of this section is to recall that MFEs (of arbitrary order) can be used on general polygonal meshes and to study their relations to other locally conservative methods.

Let $\hat{T}_H$ be a mesh consisting of general polygonal elements. We give an example in Fig. 7 (dashed lines). The mesh $\hat{T}_H$ does not need to be matching (can contain hanging nodes), the maximal number of sides of each element $K \in \hat{T}_H$ is not necessarily limited, $\hat{T}_H$ is not necessarily shape-regular, and its elements can be
nonconvex and nonstar-shaped. We only need the existence of a simplicial submesh $T_h$ of $\tilde{T}_H$ (that is, every element of $\tilde{T}_H$ is triangulated by elements of $T_h$) which is matching. An example is given in Fig. 7 (solid lines). We will show that MFEs of arbitrary order can be defined on such polygonal meshes and written with potential unknowns only related to the elements of $\tilde{T}_H$ and fluxes/Lagrange multipliers only related to the sides of $\tilde{T}_H$. For this purpose, we start either from (1.3) or from (3.3).

Recall from Refs. 16 and 47 that any order MFE method from any of the different families on a simplicial mesh $T_h$ can be written under this form.

Theorem 7.1. (Mixed finite elements on polygonal meshes for Lagrange multipliers) Let $\tilde{T}_H$ be an arbitrary polygonal mesh with a matching simplicial submesh $T_h$. Consider any MFE scheme on $T_h$, written for the Lagrange multipliers $\Lambda$ associated with the sides of $T_h$ as

$$Z \Lambda = E. \quad (7.1)$$

Then, $(7.1)$ can be condensed into

$$\tilde{Z} \tilde{\Lambda} = \tilde{E} \quad (7.2)$$

with a sparse, symmetric, and positive definite matrix $\tilde{Z}$. $\tilde{\Lambda}$ is an algebraic vector of Lagrange multipliers related to the sides of $\tilde{T}_H$ only. In the RT0 case (3.3) and when the sides of $T_h$ do not subdivide the sides of $\tilde{T}_H$, there is one unknown in $\tilde{\Lambda}$ per side of $\tilde{T}_H$.

Proof. Note first that (7.2) has the same structure as (7.1). The proof follows by a simple static condensation. Indeed, let $K \in \tilde{T}_H$ be given and denote the unknowns $\Lambda$ of (7.1) corresponding to the sides of $T_h$ which are in the interior of $K$ by $\Lambda_{\text{int}}^K$. Use a similar notation $E_{\text{int}}^K$ for the right-hand side entries of (7.1). Finally, denote the unknowns corresponding to sides of $T_h$ which are on the boundary of $K$ but not on the boundary of $\Omega$ by $\Lambda_{\text{ext}}^K$. Consider the lines of (7.1) associated with such sides.
of \(T_h\) which are in the interior of \(K\). This gives rise to the following local problem: given \(\Lambda^\text{ext}_K\), find \(\Lambda^\text{int}_K\) such that
\[
\int_K \Lambda^\text{int}_K = E^\text{int}_K - \int_K \Lambda^\text{ext}_K - \int_K \Lambda^\text{ext}_K.
\] (7.3)

Note that the system matrix \(\int_K\) is square, symmetric, and positive definite, which implies the well-posedness of the local problem (\(\int_K\) is a submatrix of \(\int\) corresponding to the lines and columns corresponding to the sides of \(T_h\) which are in the interior of \(K\)). Note also that (7.3) is a local Dirichlet problem, which allows to compute \(\Lambda^\text{int}_K\) from \(\Lambda^\text{ext}_K\) and \(E^\text{int}_K\). More precisely, we have
\[
\Lambda^\text{int}_K = (\int_K)^{-1}(E^\text{int}_K - \int_K \Lambda^\text{ext}_K).
\] (7.4)

Next, we repeat this procedure for all \(K \in \hat{T}_H\). We finally use the lines of (7.1) associated with the sides of \(T_h\) which are on the boundary of some \(K \in \hat{T}_H\) but not on the boundary of \(\Omega\), where we insert the expressions (7.4). This gives (7.2). Note that, in contrast to Sec. 4, all equations of (7.1) are used exactly once; this process is called static condensation and clearly leads to the well-posedness of (7.2) with \(\hat{\int}\) being sparse, symmetric, and positive definite.

**Theorem 7.2.** (Mixed finite elements on polygonal meshes for flux and potential unknowns) Let \(\hat{T}_H\) be an arbitrary polygonal mesh with a matching simplicial submesh \(T_h\). Consider any MFE scheme on \(T_h\) leading to
\[
(A \ B^t) (U \ P) = (F \ G).
\] (7.5)

Then, (7.5) can be condensed into a well-posed system of the form
\[
(\hat{A} \ \hat{B}^t) (\hat{U} \ \hat{P}) = (\hat{F} \ \hat{G}),
\] (7.6)

where the system matrix is sparse and of indefinite, saddle point type, with \(\hat{A}\) symmetric and positive definite and \(\hat{B}\) of full row rank. \(\hat{U}\) is an algebraic vector of flux unknowns related to the sides of \(\hat{T}_H\) only, and \(\hat{P}\) is an algebraic vector of potential unknowns with one entry per element of \(\hat{T}_H\). In the \(RT0\) case (1.3) and when the sides of \(T_h\) do not subdivide the sides of \(\hat{T}_H\), there is one unknown in \(\hat{U}\) per side of \(\hat{T}_H\) and one unknown in \(\hat{P}\) per element of \(\hat{T}_H\).

**Proof.** The proof follows by a simple static condensation. We present here briefly its algebraic form; more details are given below in Theorem 7.3. Note first that (7.6) has the same structure as (7.5). Consider such basis of the space \(\Phi_h\) of (1.2a)–(1.2b) which contains the indicator functions \(\phi_K\) of all \(K \in \hat{T}_H\), i.e. the functions equal to one on \(K\) and zero elsewhere (then, all other basis functions of \(\Phi_h\) have zero mean value on each \(K \in \hat{T}_H\)). Let \(K \in \hat{T}_H\) be given and consider the lines of the first block of (7.5), i.e. of \((A \ B^t)\), associated with such sides of \(T_h\) which are in the interior of \(K\). Consider moreover the lines of the second block of (7.5), i.e. of
Let the source term \( g \) be piecewise constant on \( \hat{T}_H \), let the sides of \( T_h \) do not subdivide the sides of \( \hat{T}_H \), and consider the RT0 case (1.3). Then the MFE method (7.6) belongs to the MFD family as characterized in Ref. 17. More generally, both (7.2) and (7.6) belong to the family characterized in Ref. 27.

**Proof.** We first show that (7.6) belongs to the MFD family as characterized in Ref. 17. Let \( K \in \hat{T}_H \) be fixed. Inverting the system matrix in (7.7) gives:

\[
\begin{pmatrix}
U_{\text{int}}^K \\
P_0^K
\end{pmatrix} = \begin{pmatrix}
\mathbb{B}_K^{\text{int,int}} & \mathbb{B}_K^{\text{int,ext}} \\
\mathbb{B}_K^{\text{ext,int}} & 0
\end{pmatrix}^{-1} \begin{pmatrix}
F_{\text{int}}^K - \mathbb{A}_K^{\text{int,int}} U_{\text{int}}^K \\
G_0^K - \mathbb{B}_K^{\text{ext,int}} U_{\text{ext}}^K
\end{pmatrix}.
\]

(7.8)

Let us denote by \( P_{0K}^1 \) the coefficient associated with the \( \phi_K \) basis function (this is the value of the vector \( \hat{P} \) associated with \( K \)). Equations of (7.5) which were not used in (7.7) related to the element \( K \) read

\[
\begin{pmatrix}
\mathbb{B}_K^{\text{ext,int}} \\
\mathbb{B}_K^{\text{1,ext}}
\end{pmatrix} \begin{pmatrix}
U_{\text{ext}}^K \\
U_{\text{int}}^K
\end{pmatrix} + \begin{pmatrix}
\mathbb{B}_K^{\text{1,ext}} \\
0
\end{pmatrix} \begin{pmatrix}
P_0^K \\
P_{0K}^1
\end{pmatrix} = \begin{pmatrix}
P_{\text{ext}}^1 \\
G_1^K
\end{pmatrix}.
\]

(7.9)
Plugging (7.8) into (7.9) gives the following block system for the unknowns $(U^\text{ext}_K, P^1_K)$:

\[
\begin{pmatrix}
\hat{\mathbf{K}}_K \\
\mathbf{B}^\text{int,ext}_K
\end{pmatrix}
\begin{pmatrix}
(U^\text{ext}_K)^	au \\
P^1_K
\end{pmatrix}
= 
\begin{pmatrix}
\hat{\mathbf{G}}_K \\
\mathbf{0}
\end{pmatrix},
\]

(7.10)

where

\[
\hat{\mathbf{K}}_K := \hat{\mathbf{K}}^\text{int,ext}_K - \left(\begin{array}{c|c}
\mathbf{B}^\text{int,ext}_K & \mathbf{0} \\
\hline
\mathbf{0} & \mathbf{0}
\end{array}\right)^\tau 
\begin{pmatrix}
\hat{\mathbf{K}}^\text{int,ext}_K \\
\mathbf{B}^\text{0,ext}_K
\end{pmatrix}^{-1} 
\begin{pmatrix}
\mathbf{B}^\text{0,ext}_K \\
\mathbf{0}
\end{pmatrix},
\]

(7.11)

and

\[
\hat{\mathbf{G}}_K := \hat{\mathbf{G}}^\text{int,ext}_K - \left(\begin{array}{c|c}
\mathbf{B}^\text{int,ext}_K & \mathbf{0} \\
\hline
\mathbf{0} & \mathbf{0}
\end{array}\right)^\tau 
\begin{pmatrix}
\mathbf{B}^\text{0,ext}_K \\
\mathbf{0}
\end{pmatrix}^{-1} 
\begin{pmatrix}
\mathbf{B}^\text{0,ext}_K \\
\mathbf{0}
\end{pmatrix}. 
\]

(7.12)

Let us first remark that the element contribution (7.10) has the same unknowns and matrix structure as in the MFD method. Moreover, the matrices $\mathbf{B}^\text{int,ext}_K$ are identical as in MFDs. Because of the homogeneous Dirichlet boundary condition (1.1b), both $P^\text{int}_K$ and $F^\text{ext}_K$ are zero; moreover, since we have supposed $g$ constant on $K$, $G^0_K = 0$ follows. Thus $\hat{\mathbf{F}}^\text{ext}_K = 0$, and we find that also the right-hand sides are the same. It thus remains to check that the matrix $\hat{\mathbf{K}}_K$, or more precisely the scalar product that it generates, satisfies the requirements of stability (3.10) and consistency (5.1) from Ref. 17 ($\hat{\mathbf{K}}_K$ is obviously symmetric). For this purpose, we will rely on the notion of the lifting operator of Ref. 17 (Theorem 5.1).

Let the notation $\mathbf{V}_h(K)$ stand for the restriction of the space $\mathbf{V}_h$, defined on the simplicial mesh $\mathcal{T}_h$, to the element $K \in \mathcal{T}_H$. Let $U^\text{ext}_K$ be an arbitrary vector associated with the boundary sides $\mathcal{E}^\text{ext}_K$ of the element $K \in \mathcal{T}_H$, representing the exterior normal fluxes through these sides. We will construct a lifting operator which to a given vector $U^\text{ext}_K$ associates a function $\mathbf{u}_{h,K}$ from $\mathbf{V}_h(K)$ such that:

\[
(\mathbf{u}_{h,K} \cdot \mathbf{n}_K)|_{\sigma} = \frac{U^\text{ext}_K}{|\sigma|}, \quad \forall \sigma \in \mathcal{E}^\text{ext}_K, \tag{7.13}
\]

and

\[
\nabla \cdot \mathbf{u}_{h,K} = \frac{1}{|K|} \sum_{\sigma \in \mathcal{E}^\text{ext}_K} U^\text{ext}_\sigma. \tag{7.14}
\]

This corresponds to the problem (5.4) in Ref. 17, with a different scaling (our vector $U^\text{ext}_K$ represents the normal fluxes). Let us decompose $\mathbf{V}_h(K)$ into $\mathbf{V}_h(K)^\text{ext}$, generated by the basis functions of $\mathbf{V}_h(K)$ associated with $\sigma \in \mathcal{E}^\text{ext}_K$ (oriented along the exterior normal $\mathbf{n}_K$), and $\mathbf{V}_h(K)^\text{int}$, generated by the other basis functions of $\mathbf{V}_h(K)$. Prescribe $\mathbf{u}^\text{ext}_{h,K} \in \mathbf{V}_h(K)^\text{ext}$ by the fluxes $U^\text{ext}_\sigma$. Let $\Phi_h(K)$ stand for the restriction of the space $\Phi_h$, defined on the simplicial mesh $\mathcal{T}_h$, to the element $K \in \mathcal{T}_H$. Let $\Phi_h(K)^0$ be its subspace of functions with mean value zero. Then (7.13)–(7.14) can be achieved via the following local Neumann problem (note that the fact
that the right-hand side in (7.14) is constant is important to insert it to the right-hand side of (7.15b): find \( \mathbf{u}_{h,K} \in \mathbf{V}_h(K), \mathbf{u}_{h,K} = \mathbf{u}_{h,K}^{\text{int}} + \mathbf{u}_{h,K}^{\text{ext}} \), and \( p_h^0 \in \Phi_h(K)^0 \) such that:

\[
(\mathbf{S}^{-1} \mathbf{u}_{h,K}^{\text{int}}, \mathbf{w}_{h,K}^{\text{int}})_K - (p_h^0, \nabla \cdot \mathbf{w}_{h,K}^{\text{int}})_K = - (\mathbf{S}^{-1} \mathbf{u}_{h,K}^{\text{ext}}, \mathbf{w}_{h,K}^{\text{int}})_K \\
\forall \mathbf{w}_{h,K}^{\text{int}} \in \mathbf{V}_h(K)^{\text{int}}, \quad (7.15a)
\]

\[
-(\nabla \cdot \mathbf{u}_{h,K}^{\text{int}}, \phi_h^0)_K = (\nabla \cdot \mathbf{u}_{h,K}^{\text{ext}}, \phi_h^0)_K \\
\forall \phi_h^0 \in \Phi_h(K)^0. \quad (7.15b)
\]

Whenever \( U_{\sigma}^{\text{ext}} \) represents the normal fluxes across the sides of \( S_K^{\text{ext}} \) of a constant vector field \( \mathbf{u}_h \), the solution \( \mathbf{u}_{h,K} \) of (7.15a)–(7.15b) obviously coincides with \( \mathbf{u}_h \), so that the consistency condition (5.5) of Ref. 17 is satisfied. Moreover, the stability condition (5.8) of Ref. 17 is also classical for the above local Neumann problem.

It is now crucial to understand that (7.15a)–(7.15b) is nothing but a functional form of (7.7). Indeed, as explained above, \( \mathbf{F}_K^{\text{int}} \) and \( G_K^0 \) are both equal to zero in the present setting. Now let the two vectors \( U_K^{\text{ext}} \) and \( V_K^{\text{ext}} \) be arbitrary. Let \( \mathbf{u}_{h,K} \) be the solution of (7.15a)–(7.15b) for \( U_K^{\text{ext}} \). Similarly, prescribe \( v_{h,K}^{\text{ext}} \in \mathbf{V}_h(K)^{\text{ext}} \) by the fluxes \( V_K^{\text{ext}} \) and define \( \mathbf{v}_h \in \mathbf{V}_h(K), \mathbf{v}_h = v_{h,K}^{\text{ext}} + v_{h,K}^{\text{int}}, \) and \( q_h^0 \in \Phi_h(K)^0 \) by

\[
(\mathbf{S}^{-1} \mathbf{v}_{h,K}^{\text{int}}, \mathbf{w}_{h,K}^{\text{int}})_K - (q_h^0, \nabla \cdot \mathbf{w}_{h,K}^{\text{int}})_K = - (\mathbf{S}^{-1} \mathbf{v}_{h,K}^{\text{ext}}, \mathbf{w}_{h,K}^{\text{int}})_K \\
\forall \mathbf{w}_{h,K}^{\text{int}} \in \mathbf{V}_h(K)^{\text{int}}, \quad (7.16a)
\]

\[
-(\nabla \cdot \mathbf{v}_{h,K}^{\text{int}}, \phi_h^0)_K = (\nabla \cdot \mathbf{v}_{h,K}^{\text{ext}}, \phi_h^0)_K \\
\forall \phi_h^0 \in \Phi_h(K)^0. \quad (7.16b)
\]

Using the equivalence of (7.15a)–(7.15b) and (7.7), we will now show that there holds

\[
(V_K^{\text{ext}})^T \mathbf{A}_K U_K^{\text{ext}} = (\mathbf{S}^{-1} \mathbf{u}_h, \mathbf{v}_h)_K, \quad (7.17)
\]

which is nothing but the definition (5.7) of Ref. 17 (for \( \mathbf{S} \) constant on \( K \)). Thus, all conditions of Ref. 17 (Theorem 5.1) will be satisfied, proving that the MFE method (7.6) belongs to the MFD family.

In order to show (7.17), remark that the matrix \( \mathbf{A}_K \) appears in (7.10) obtained upon plugging (7.7) into (7.9). Let \( p_{h,K} \in \Phi_h(K), p_h = p_{h,K}^1 + p_{h,K}^0, \) with \( p_{h,K}^1 \in \Phi_h(K)^1 \) corresponding to the constant \( P_K^1 \) from (7.9). The functional form of (7.9) states that \( \mathbf{u}_{h,K} \) and \( p_{h,K} \) need to satisfy:

\[
(\mathbf{S}^{-1} \mathbf{u}_{h,K}^{\text{ext}}, \mathbf{w}_{h,K}^{\text{ext}})_K + (\mathbf{S}^{-1} \mathbf{u}_{h,K}^{\text{int}}, \mathbf{w}_{h,K}^{\text{int}})_K \\
-(\nabla \cdot \mathbf{u}_{h,K}^{\text{ext}}, \phi_{h,K}^0)_K = -(\nabla \cdot \mathbf{u}_{h,K}^{\text{int}}, \phi_{h,K}^0)_K \\
-(\nabla \cdot \mathbf{u}_{h,K}^{\text{ext}}, \phi_{h,K}^0)_K = (\mathbf{g}, \phi_{h,K}^0)_K \\
\forall \phi_{h,K}^0 \in \Phi_h(K)^1, \quad (7.18a)
\]

\[
(\mathbf{S}^{-1} \mathbf{u}_{h,K}^{\text{int}}, \mathbf{w}_{h,K}^{\text{int}})_K - (p_{h,K}^1, \nabla \cdot \mathbf{w}_{h,K}^{\text{int}})_K - (p_{h,K}^0, \nabla \cdot \mathbf{w}_{h,K}^{\text{int}})_K = 0 \\
\forall \mathbf{w}_{h,K}^{\text{int}} \in \mathbf{V}_h(K)^{\text{int}}, \quad (7.18a)
\]

\[
-(\nabla \cdot \mathbf{u}_{h,K}^{\text{int}}, \phi_{h,K}^0)_K = -(\nabla \cdot \mathbf{u}_{h,K}^{\text{ext}}, \phi_{h,K}^0)_K \\
\forall \phi_{h,K}^0 \in \Phi_h(K)^0. \quad (7.18b)
\]
noting that \((\nabla \cdot \mathbf{u}_{int}^h,K, \phi^h_{int})_K = 0\) by the Green theorem. Taking \(w_{ext}^h,K = v_{ext}^h,K\) in (7.18a), we see that
\[
(V_{ext}^k)^T K U_{ext}^k = (S^{-1} u_{int}^h,K, v_{ext}^h,K)_K - (p_0^h,K, \nabla \cdot v_{ext}^h,K)_K,
\]
where \(u_{int}^h,K\) and \(p_0^h,K\) depend on \(u_{ext}^h,K\) through (7.15a)–(7.15b). Now taking \(p_0^h,K\) as test function in (7.16b) and \(v_{int}^h,K\) as test function in (7.15a) yields
\[
-(p_0^h,K, \nabla \cdot v_{ext}^h,K)_K = (\nabla \cdot v_{int}^h,K, p_0^h,K)_K = (S^{-1} u_{int}^h,K, v_{int}^h,K)_K.
\]
Combining the two above relations proves (7.17).

We finally show the second assertion of the theorem. In Ref. 27 (Theorem 3.1), it was shown that the MFD,\(^{17}\) mixed finite volume,\(^{26}\) and hybrid finite volume methods\(^{31}\) (or more precisely their generalized forms) on quite general polygonal meshes are all equivalent in the sense that there always exist choices of the parameters of these methods so that they coincide. As we have just demonstrated that MFEs belong to the MFD family, they also belong to this larger family of methods.

We end this section by few remarks.

**Remark 7.2.** (Comparison with mimetic finite difference, mixed finite volume, and hybrid finite volume methods) The MFE methods on polygonal meshes (7.2) and (7.6) have less restrictive assumptions on the mesh \(\mathring{T}_H\) than those needed in Refs. 17, 26, 31, and 27. The only requirement is the existence of the matching simplicial submesh \(T_h\) which needs to be shape-regular for the convergence. Moreover, they admit \(\mathring{S}\) and \(g\) nonconstant inside the polygonal elements \(K \in \mathring{T}_H\) (the alignment is only used with respect to the submesh \(T_h\)). In our opinion, they can be seen as any of the above methods with optimal choice of the free parameter(s) and improved resolution inside the polygonal cells.

**Remark 7.3.** (A priori and a posteriori error estimates) Using Theorems 7.1 and 7.2, (7.2) is only an algebraic rewriting of (7.1), and (7.6) is only an algebraic rewriting of (7.5). Thus, all a priori and a posteriori error estimates known for simplicial meshes are valid for MFE methods also on arbitrary polygonal grids.

**Remark 7.4.** (Discrete maximum principle) In the same sense as in the previous remark, under the conditions discussed in Sec. 6, the discrete maximum principle is valid for MFE methods also on arbitrary polygonal grids.

**Remark 7.5.** (Mixed finite elements with one unknown per polygonal cell) Consider the RT0 case. It turns out that, by the techniques described in Sec. 4, both (7.2) and (7.6) can, at least in certain situations, be reduced to
\[
\mathring{S} \mathring{P} = \mathring{H},
\]
i.e. to a system with one unknown per each polygon in \(\mathring{T}_H\), cf. Ref. 55 (Appendix A).
8. Conclusion

We have introduced in this paper a systematic way of reducing the number of unknowns in the RT0 MFE method from the flux and potential unknowns to (new) potential unknowns only. This gives rise to a whole family of equivalent one-unknown-per-element reformulations of the MFE method, some of which reduce to well-known discretization schemes. Practically, this can lead to important computational savings. At the same time, various local flux expressions are obtained. We have also recalled the validity of the discrete maximum principle and the possibility to formulate MFE methods on general polygonal meshes.

To us, the major conclusion of the present paper is that in all underlying principles, derivation, properties, applicability, and computational cost, the RT0 MFE method is closely related to many other locally conservative discretization methods. Our main belief is that through these different links, the results and tools available in MFE methods and those available in the context of the previously cited methods can become mutually available: for MFE methods which possess a well-explored and solid theoretical background, let us cite, e.g. the well-posedness of (1.3)/(3.2)/(3.3)/(7.1)/(7.5)/(1.7) on simplicial meshes, well-posedness of both (7.2) and (7.6) on arbitrary polygonal meshes, the discrete maximum principle in the lowest-order case, optimal convergence and superconvergence \( a \text{ priori} \) error estimates (see, e.g. Refs. 16, 47, 23, 24, and 54), optimal \( a \text{ posteriori} \) error estimates,\(^5\) multigrid methods (see, e.g. Ref. 15 and 22), parallel implementations (see, e.g. Ref. 35), multiscale and mortar versions (see, e.g. Ref. 9), and convergence and optimality of adaptive methods\(^12,21\); for the previously cited FV-type methods, let us cite in particular convergence analysis for nonlinear (degenerate) parabolic equations and their systems, see, e.g. Refs. 32, 33, and 30.

Acknowledgment

The first author was supported by the GNR MoMaS project “Numerical Simulations and Mathematical Modeling of Underground Nuclear Waste Disposal”, PACEN/CNRS, ANDRA, BRGM, CEA, EdF, IRSN, France.

References

4. A. Agouzal, J. Baranger, J.-F. Maitre and F. Oudin, Connection between finite volume and mixed finite element methods for a diffusion problem with nonconstant


