INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 2, Number 3, Pages 241–282

MORTAR ADAPTIVITY IN MIXED METHODS FOR FLOW IN POROUS MEDIA

MAŁGORZATA PESZYŃSKA

Abstract. We define an error indicator for mixed mortar formulation of flow in porous media. The mixed mortar domain decomposition method for single-phase flow problems was defined by Arbogast et al; it relies on coupling of subdomain problems using mortar Lagrange multipliers defined as continuous piecewise linears on the subdomain interface. The accuracy and efficiency of the resulting interface formulation depends on the number of mortar degrees of freedom which we propose to adapt using error indicators involving jump of the flux across the interface. Rigorous a-posteriori analysis and proof of reliability of the estimator are established for single-phase 2D flow problems with diagonal coefficients for $\mathbf{RT}_{[0]}$ spaces on rectangular grids. Computational experiments demonstrate the application of the estimator. Next, the algorithm and indicator are extended to the two-phase flow case which is illustrated with numerical examples. We focus on adapting the mortar grid while keeping subdomain grids fixed. Full mortar adaptivity is discussed elsewhere.

Key Words. Single-phase flow in porous media, multi-phase flow, mixed finite elements, a-posteriori error estimation, mortars, domain decomposition, adaptivity

1. Introduction

This paper is devoted to grid adaptivity for a family of heterogeneous domain decomposition methods based on the *mixed mortar* finite element method.

The method was introduced in [7] and it provides a rigorous optimally convergent discretization technique for the elliptic equation

$$-\nabla \cdot (K\nabla p) = f, \quad x \in \Omega, \tag{1}$$

with $\Omega \subset \mathbb{R}^d, d = 2, 3$. Here K denotes the diffusion coefficient, f denotes the source/sink terms, and p is the unknown pressure.

In the mortar domain decomposition method, the region Ω is decomposed into individual non-overlapping subdomains $\Omega_i, i = 1, \ldots n$ which are separated by the union of interfaces Γ on which *mortar* grids and unknowns are introduced. The subdomains are gridded independently; subdomain problems which are the local counterparts of (1) can be solved essentially independently from one another but are coupled by mortars; see [14, 12] for mortar formulation when subdomain problems are solved with standard Galerkin (conforming) methods.

In the mixed mortar method the subdomain problems are solved using mixed finite element methods thereby providing a locally conservative approximation to

Received by the editors April 15, 2004 and, in revised form, October 22, 2004.

2000 *Mathematics Subject Classification.* 65N30,65M60,65N50,65N55,76S05,76M10.

This research was partly supported by the DOE grant DE-FG03-99ER25371, and the NSF grants: SBR 9873326, ITR EIA-0121523, NPACI 10181410.

both pressure and velocity unknowns $\mathbf{u} := -K\nabla p$. The method relies on introducing *mortar Lagrange multipliers* on the interface Γ which provide Dirichlet boundary conditions for the subdomain problems. Additionally, the subdomain problems are coupled by the requirement that the global velocities be *weakly* continuous across Γ which relaxes the global continuity (of normal components across any smooth surface) of exact velocities. This weak-continuity condition averages the jumps of velocities and is defined relative to the discrete space of Lagrange multipliers on the interface which are defined on a *mortar grid* characterized by the parameter h_m, or by the number of mortar degrees of freedom $n_m \approx O(\frac{1}{h_m})$.

Let us be given a collection of rectangular partitions of Ω_i with associated grid parameter $h = \max_i h_i$. In principle, h_m can be selected independently of h, as long as certain lower and upper bound conditions hold. These guarantee, respectively, the unique solvability of the mixed mortar formulation, and the optimal approximation properties of weakly continuous velocities which in turn are necessary for the optimal rate of convergence of the method, the same as for discretization *without* mortars. For this optimal convergence rate which, for lowest order Raviart-Thomas spaces $\mathbf{RT}_{[0]}$, is $O(h)$ in both the pressure and velocity unknowns [46, 17], h_m should depend linearly on h; the approximation error increases, in general, with the proportionality constant.

The number of mortar unknowns n_m on Γ determines the complexity of the *interface problem*. Recall that, in a classical domain decomposition setting, the algorithm for approximation of (1) can be written in terms of the interface unknowns and as such solved by an iterative algorithm which requires, in each iteration, solution of subdomain problems which are responding to the current guess of Dirichlet data. In the mixed mortar algorithms the number of iterations on the interface in general grows with n_m, unless optimal preconditioners can be applied.

The mixed mortar method has been the cornerstone of several major reservoir simulation projects. Recall that (1) can be used as a model for single-phase flow in a reservoir Ω. Its natural extension is to the multi-phase flow; the algorithm has been integrated within the IPARS (Integrated Parallel Accurate Reservoir Simulator) framework [51, 44, 49, 41, 35]. The attractiveness of the mortar approach lies in that it makes the subdomain problems independent from one another. It is only the interface Lagrange multipliers (Dirichlet data) and the resulting fluxes (the Neumann data) which provide "communication" between subdomains. As such, the subdomain problems can be considered as "black-boxes" thereby allowing for local adaptivity of time-stepping [43], grids and solvers, and mainly, the physical models [37]. The latter coupling is a form of *multiphysics* and is an instance of *heterogeneous domain decomposition*; see [50].

The difficulties in practical application of the overall procedure lie in finding optimal preconditioners for the general multi-phase solver on the interface; see [57]. However, in spite of the large complexity of the interface solver, the mixed mortar approach has been extremely successful when applied to a large class of real reservoir problems. It is important to note that in all the successful cases we found a relatively small n_m sufficient for a good level of accuracy and at the same time mandatory for an acceptable degree of computational complexity.

In [45] the mixed mortar method gave rise to the *mortar upscaling* method. Here the subdomain grids are kept fixed but n_m varies, thereby providing a variable degree of local conservation of mass or of weak-continuity of the fluxes.

It is in the research reported in [43, 37, 45] that the need to define the "right" mortar grid and to control the error due to only weak continuity of fluxes became

very important. It was clear that the quantitative criteria to determine the mortar grid, that is, the error indicators, must somehow involve a measure of the defect introduced by the weak continuity condition, and that these measures must involve the jump of the computed fluxes across Γ. This intuition was readily confirmed by numerical experiments. At that time, several new results applicable to a-posteriori error indicators for mixed methods in $\mathbb{R}^2$ [56, 18] and for mortar formulations for conforming methods [54, 53, 55] became available. These have provided the basis for our rigorous analysis.

In this paper we combine the ideas coming from [56, 18, 54, 53, 55, 43, 37, 45] to rigorously justify the construction of error indicators for mixed mortar solution of single-phase problems; we also extend these ideas to two-phase flow. We use the algorithmic and implementation basis referenced above to provide the computational results which demonstrate the strength of our approach.

The plan of the paper is as follows. In Section 2 we recall the original statement of the mixed mortar formulation for (1) and introduce a view of mortar grid parametrization which departs from one in [7]. In Section 3 we construct an error indicator for $\mathbf{RT}_{[0]}$ mixed finite element spaces over a quadrangulated subdomain with non-homogeneous boundary conditions and diagonal K, and prove an upper bound for the error; these results are a detailed but natural extension of [56, 18] where triangular grids, Raviart-Thomas $\mathbf{RT}_0$ spaces on triangles, and homogeneous boundary conditions were used. Since the mixed mortar algorithm comes from application of boundary conditions to local subdomains, the construction in Section 3 provides the building blocks for the main theoretical result of this paper developed in Section 4; it is the definition of a reliable *residual a-posteriori error estimator* η for (1) in the mixed mortar form. In Section 5 we illustrate the applicability of this result to single-phase flow computations. Next, Section 6 introduces briefly the mixed mortar algorithm for two-phase flow and appropriate extensions of η, and Section 7 presents the computational results on mortar adaptivity for two-phase flow.

In all computational results we are only concerned with adaptivity of mortars. In general, both subdomain and mortar grids should be adapted as is done in [25, 30] for Galerkin methods, but this will not be pursued here.

2. Mixed mortar formulation for single phase flow

In this section we recall the statement and the main results concerning the mixed mortar formulation from [7]. We start by introducing notation, following [7, 56, 18, 54], and proceed in Section 2.2 to define the problem as in [7]. In Section 2.3 we discuss the parametrization of mortar grid.

2.1. Notation. Consider a set $\omega \subset \mathbb{R}^2$ with $\omega \ni x = (x_1, x_2)$. Denote $\nabla = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2})$ and $\mathbf{Curl} := (\frac{\partial}{\partial x_2}, -\frac{\partial}{\partial x_1})$ so that for $\psi : \omega \mapsto \mathbb{R}$ we have $\nabla\psi := (\frac{\partial \psi}{\partial x_1}, \frac{\partial \psi}{\partial x_2}) = (\psi_1, \psi_2)$ and $\mathbf{Curl}\psi = (\frac{\partial \psi}{\partial x_2}, -\frac{\partial \psi}{\partial x_1}) = (\psi_2, -\psi_1)$. For a vector valued function $\mathbf{q} : \omega \mapsto \mathbb{R}^2$, $\mathbf{q} = (q_1, q_2)$ we write $div\mathbf{q} := \nabla \cdot \mathbf{q} = \frac{\partial q_1}{\partial x_1} + \frac{\partial q_2}{\partial x_2} = q_{1,1} + q_{2,2}$. Also, we use the symbol $curl\mathbf{q} := \frac{\partial q_1}{\partial x_2} - \frac{\partial q_2}{\partial x_1} = q_{1,2} - q_{2,1}$. We also introduce $\mathbf{T} := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. then we notice $\mathbf{Curl}\phi = \mathbf{T}\nabla\phi$, and $curl\mathbf{q} = -\nabla \cdot \mathbf{Tq}$.

On the boundary $\partial\omega$ of ω we denote by $\mathbf{n}_\omega = (n_1, n_2)$ the unit outward normal to the boundary $\partial\omega$ and by $\mathbf{t}_\omega := \mathbf{Tn} = (n_2, -n_1)$ the unit tangent vector whose orientation is consistent with the clockwise orientation of $\partial\omega$. We drop the subscript when it does not lead to a confusion. Clearly $\mathbf{Curl}\phi \cdot \mathbf{n} = -\nabla\phi \cdot \mathbf{t}$.

When integrating over ω, we abbreviate dx_1dx_2 by dx and use the symbol ds in boundary integrals. Also, we use the notation $< \phi, \psi >_{\partial\omega} := \int_{\partial\omega} \phi(s)\psi(s)ds$ for boundary integrals.

Now we recall the Lebesgue and Sobolev function spaces [2]. As usual, $L^2(\omega)$ denotes the space of square integrable functions (distributions) over ω; it is a Hilbert space with the product $(\phi, \psi)_\omega := \int_\omega \phi(x)\psi(x)dx$ and with the norm $\| \phi \|_{0,\omega} := \sqrt{(\phi, \phi)_\omega}$. (We shall drop the subscript $_\omega$ when this does not lead to confusion.) Next, $H^1(\omega)$ denotes the space of functions which, along with their (weak) derivatives, are in $L^2(\omega)$; $H^1_0(\omega)$ is the subset of $H^1(\omega)$ consisting of functions whose boundary values in the sense of traces are zero on $\partial\omega$. We recall that the traces of $H^1(\omega)$ functions are in $H^{1/2}(\partial\omega)$, provided $\partial\omega$ is smooth. For details on standard notation and the properties of Sobolev spaces, see [2]. The space $\mathbf{H}(\mathbf{div}; \omega)$ is the space of vector valued functions whose components, along with their divergence, are in $L^2(\omega)$, and the associated norm is $\| \mathbf{q} \|_{\mathbf{H}(\mathbf{div};\omega)} := \sqrt{\int_\omega \mathbf{q} \cdot \mathbf{q} dx + (\nabla \cdot \mathbf{q})^2 dx}$. We recall that normal traces of elements of $\mathbf{H}(\mathbf{div}; \Omega)$ across any smooth curve are continuous [17]. We shall also use another norm on $\mathbf{H}(\mathbf{div}; \omega)$. Let us be given a positive definite uniformly bounded tensor $K = K(x),\ x \in \omega$, such as the one discussed below. Then one defines

$$\| \mathbf{q} \|_{K,\omega} := \sqrt{\int_\omega K^{-1}\mathbf{q} \cdot \mathbf{q} dx + (\nabla \cdot \mathbf{q})^2 dx}. \tag{2}$$

This norm is equivalent to the standard $\mathbf{H}(\mathbf{div}; \Omega)$ norm; the constants of equivalence are established below.

Next, by $Y(A) \bigoplus Y(B)$ we mean the set of functions on $A \cup B$ whose restrictions to sets A, B belong to the spaces $Y(A)$ and $Y(B)$, respectively. In contrast, $Y + Z$ denotes a set of linear combinations of functions from Y and Z.

Finally, the projections to a subspace $Y \subseteq X$ are defined in the usual way: $x \in X, \pi_Y x \in Y$ and $(x - \pi_Y x, y)_Y = 0, \forall y \in Y$.

2.2. Statement of the problem. Consider a fixed open bounded domain $\Omega \subset \mathbb{R}^2$ and the flow of a single phase incompressible fluid whose pressure is denoted by $p = p(x),\ x \in \Omega$. The flow is governed by Darcy's law which defines the velocity (flux) $\mathbf{u}$ as proportional to the gradient of the pressure p

$$\mathbf{u} \quad = \quad -K\nabla p. \tag{3}$$

The (hydraulic conductivity) coefficient K combines the (absolute) permeability constant and the viscosity of the fluid. Here we ignore the presence of gravity; gravity can be handled by replacing p with a potential variable.

The conductivity tensor $K(x)$ is assumed to be symmetric, essentially bounded and uniformly positive definite. In fact, in this paper we assume that K is diagonal, that is,

$$K(x) := \begin{bmatrix} K_1(x) & 0 \\ 0 & K_2(x) \end{bmatrix}.$$

Denote the upper and lower bounds on the eigenvalues of this coefficient over a set ω by

$$K_\omega := \min_{x \in \omega} \{K_1(x), K_2(x)\}, \tag{4}$$

$$\kappa_\omega := \min_{x \in \omega} \{K_1^{-1}(x).K_2^{-1}(x)\}. \tag{5}$$

Obviously, $\kappa_\omega^{-1} = \max_{x \in \omega} \{K_1(x), K_2(x)\}\}$ and $K_\omega^{-1} = \max_{x \in \omega} \{K_1^{-1}(x), K_2^{-1}(x)\}\}$.

While Darcy's law (3) expresses the conservation of momentum, the conservation of mass of the fluid (continuity equation) is expressed by

$$\nabla \cdot \mathbf{u} = f \tag{6}$$

where f represents the sources and sinks of the fluid.

The system (3),(6) can be rewritten as an elliptic equation for the pressure (1) and solved as such, or it can be solved in the mixed formulation (3), (6). In either case, boundary conditions on $\partial\Omega$ need to be specified.

In theoretical derivation of the error estimator in Sections 3, 4 we assume, for simplicity, that the pressure p satisfies the Dirichlet boundary condition

$$p|_{\partial\Omega} = g. \tag{7}$$

In general, other boundary conditions could be considered. In particular, a homogenous Neumann no-flow boundary condition

$$\mathbf{u} \cdot \mathbf{n}|_{\partial\Omega} = 0, \tag{8}$$

is useful in applications; it is used in our computational examples.

We seek a weak solution to the boundary-value problem (3), (6), (7). Set $W := L^2(\Omega)$ and $\mathbf{V} := \mathbf{H}(\mathbf{div};\Omega)$. The weak solution $(\mathbf{u}, p) \in \mathbf{V} \times W$ satisfies

$$(K^{-1}\mathbf{u}, \mathbf{v}) = (p, \nabla \cdot \mathbf{v}) - < g, \mathbf{v} \cdot \mathbf{n} >, \quad \forall \mathbf{v} \in \mathbf{V} \tag{9}$$

$$(\nabla \cdot \mathbf{u}, w) = (f, w), \quad \forall w \in W \tag{10}$$

This problem has a unique solution if $f \in W, g \in H^{1/2}(\partial\Omega)$, K is uniformly bounded and elliptic and if Ω is bounded, see [17], IV.1.2. From now on shall assume

Assumption 1. *K is diagonal and there exists $K_\Omega > 0$ and $\kappa_\Omega > 0$.*

Assumption 2. *Ω is open, bounded, convex, with polygonal boundary $\partial\Omega$, $f \in W$.*

For convergence analysis, it is necessary to assume that p is smoother than $p \in W$, which in turn, by (9), and if K is smooth, increases the smoothness of $\mathbf{u}$. In [7] it is assumed that $p \in H^{3/2+\varepsilon}, \varepsilon > 0$. In [56, 18] the problem (9)–(10) is assumed 2-regular that is, K, Ω, f, g are smooth enough so that $p \in H^2(\Omega)$. This is guaranteed, for example, in addition to (A1) and (A2), if K is differentiable and $g \in H^{3/2}(\partial\Omega)$, see ([26], Thm 8.13), exhaustive theory in [28], and ([15], Prop.2.2). We note that in general reservoir problems the piecewise constant coefficients and point sources are not smooth enough for the theory considered here.

Nevertheless, from now on we shall assume

Assumption 3. *The data K, g, are smooth enough so that $p \in H^2(\Omega), \mathbf{u} \in (H^1(\Omega))^d$.*

Finally, we note that with (A1) we have

$$\kappa_\Omega \parallel \mathbf{q} \parallel^2_{\mathbf{H}(\mathbf{div};\Omega)} \leq \parallel\!\!\parallel \mathbf{q} \parallel\!\!\parallel^2_{K,\Omega} \leq K_\Omega^{-1} \parallel \mathbf{q} \parallel^2_{\mathbf{H}(\mathbf{div};\Omega)} \tag{11}$$

2.2.1. Multiblock decomposition of Ω. Assume that there is a (sufficiently smooth) multiblock decomposition of Ω into n smooth nonoverlapping *subdomains* Ω_i, $i = 1, \ldots n$, each of which satisfies (A2). Denote by $\Gamma_{ij} := \partial\Omega_i \cap \partial\Omega_j$ the (possibly empty) interface between two subdomains Ω_i and Ω_j and for each i, by $\Gamma_i := \bigcup_{j=1}^n \Gamma_{ij}$. Finally, define the *interface* $\Gamma := \bigcup_i \Gamma_i$ which can also be written as $\Gamma = \bigcup_{i,j} \Gamma_{ij}$.

From the point of view of a subdomain, its boundary $\partial\Omega_i$ is made of the *subdomain interface* $\Gamma_i := \partial\Omega_i \cap \Gamma$ and of parts $\partial\Omega_i \setminus \Gamma = \partial\Omega_i \cap \partial\Omega$ that are part of $\partial\Omega$. Clearly, for some i, $\partial\Omega_i \setminus \Gamma = \emptyset$ and for some i, j, $\Gamma_{ij} = \emptyset$.

For simplicity of the exposition below and to remove unnecessary complications with notation, we shall assume in fact the following

Assumption 4. *Assume that each Ω_i is a rectangle (i.e., a* macroelement*) and that each Γ_{ij}, for every i, j, is a union of straight segments.*

On each subdomain Ω_i we define locally the function spaces and relate them to the globally defined spaces $W, \mathbf{V}$. The space $W = L^2(\Omega)$ can be straightforwardly decomposed as $W = \bigoplus_{i=1}^n W_i$ with the spaces of local test functions $W_i = L^2(\Omega_i)$ and with a naturally defined scalar product $(\cdot,\cdot)_i$ being a shorthand for $(\cdot,\cdot)_{\Omega_i}$ and with the associated norm $\|\cdot\|_{0,i}$.

The velocity space over $\Omega = \bigcup_{i=1}^n \Omega_i$ is constructed as follows

$$\mathbf{V^{-1}} := \bigoplus_{i=1}^n \mathbf{V}_i$$

with the local spaces $\mathbf{V}_i := \mathbf{H(div};\Omega_i)$. Note that $\mathbf{V} \subset \mathbf{V^{-1}}$ but $\mathbf{V} \neq \mathbf{V^{-1}}$ unless $n = 1$, since the normal components of functions from $\mathbf{V^{-1}}$ are not, in general, continuous across Γ. This is reflected by the superscript $^{-1}$ in the notation $\mathbf{V^{-1}}$ as in [54, 53, 55]. In the sequel we shall define an intermediate space $\mathbf{V^0}$ in the sense $\mathbf{V} \subset \mathbf{V^0} \subset \mathbf{V^{-1}}$ whose members satisfy a *weak continuity* condition with respect to the space of discrete Lagrange multipliers on Γ.

The local spaces are related to the global weak problem as follows: if $(\mathbf{u}, p) \in (\mathbf{V}, W) \cap (\mathbf{V}, H^1(\Omega))$ satisfies the global problem (9)-(10), then one can show, see [27], also ([17], IV.1.3), that the following *local* weak problem defined over Ω_i is satisfied:

$$\begin{aligned} (12)\quad (K^{-1}\mathbf{u}, \mathbf{v})_i &= (p, \nabla\cdot\mathbf{v})_i \\ &- < p, \mathbf{v}\cdot\mathbf{n}_i >_{\partial\Omega_i\setminus\partial\Omega} - < g, \mathbf{v}\cdot\mathbf{n} >_{\partial\Omega_i\cap\partial\Omega},\ \ \forall \mathbf{v} \in \mathbf{V}_i \\ (13)\quad (\nabla\cdot\mathbf{u}, w)_i &= (f, w)_i,\ \ w \in W_i \end{aligned}$$

with $p \in H^1(\Omega_i)$.

In turn, if the local problems are satisfied with $\mathbf{u}$ defined locally over each Ω_i that is, $\mathbf{u} \in \mathbf{V^{-1}}, p \in W$, then, when equations (12) are added over $i = 1, \ldots n$, and if we take $\mathbf{v} \in \mathbf{V^{-1}}$, then we see that the right hand side of the summed equations (12) contains the term

$$(14)\ \sum_i < p, \mathbf{v}\cdot\mathbf{n}_i >_{\partial\Omega_i\setminus\partial\Omega} = \sum_i < p, \mathbf{v}\cdot\mathbf{n}_i >_{\Gamma_i} = \sum_{i,j} < p, \mathbf{v}\cdot\mathbf{n}_i + \mathbf{v}\cdot\mathbf{n}_j >_{\Gamma_{ij}}$$

The global problem (9) is recovered if we require that $u \in \mathbf{V}$ and that the term (14) vanishes. This is achieved if the test functions $\mathbf{v}$ have continuous normal components across Γ that is, $\mathbf{v} \in \mathbf{V}$, because then the jump

$$(15)\qquad [\mathbf{v}\cdot\mathbf{n}]_{ij}(s) := (\mathbf{v}|_{\Omega_i}\cdot\mathbf{n}_i)(s) + (\mathbf{v}|_{\Omega_j}\cdot\mathbf{n}_j)(s),\ \ s \in \Gamma_{ij}$$

vanishes pointwise:

$$(16)\qquad [\mathbf{v}\cdot\mathbf{n}]_{ij}(s) = 0,\ \ \forall s \in \Gamma_{i,j},\ \ \forall \mathbf{v} \in \mathbf{V}.$$

In the discrete counterpart of these equations, the requirement that the jump term vanishes pointwise will be replaced by a requirement that the discrete counterpart of the term (14) vanishes. This condition will give rise to the definition of the weakly continuous velocities from a space $\mathbf{V^0}$.

We refer to [27] or ([17], III.1.2) for more details on this classical domain decomposition setup.

2.2.2. Discrete problem. Assume that each Ω_i is covered by a rectangular partition $\mathcal{T}_i$ so that $\bar{\Omega}_i = \bigcup_{E\in\mathcal{T}_i} \bar{E}$ and each E is a rectangle. This partition is a special case of a quadrilateral partition (*quadrangulation* [13]). We can assume that the edges $e \subset \partial E \in \partial\mathcal{T}_i := \bigcup_{E\in\mathcal{T}_i} \partial E$ are aligned with coordinate axes. Also, each E is an image of the reference element $\hat{E} = [0,1]^2$ under an affine reference map $\mathbf{F}_E : \hat{E} \mapsto E$ with $E \ni x = \mathbf{F}\hat{x} = b + B(\hat{x_1}, \hat{x_2})^T$ with a diagonal matrix B.

Denote by $h_E = diam(E)$ which for rectangles is the maximum length of the edges of an element E and define $h_i := \max_{E\in\mathcal{T}_i} h_E$. With this $\mathcal{T}_i$ can be referred to $\mathcal{T}_{h_i}$.

Assume that there is a uniform bound, $\forall E \in \mathcal{T}_i$, $\frac{h_E}{\rho_E} \le C^i_{\mathcal{T}}$ where ρ_E is the diameter of a circle inscribed in E. This means that for each i the family $\mathcal{T}_i$ is *regular* in the sense of [22, 13].

In this multiblock (domain) decomposition of Ω, there is no *global* quadrangulation of Ω as in general the grids on Ω_i do not have to match. However, since n is finite, one can naturally define $h = \max_i h_i$ and $C_{\mathcal{T}} := \max_i C^i_{\mathcal{T}}$. When referring to the collection of quadrangulations $\mathcal{T}_i$, we shall use the notation $\mathcal{T}_h$ or simply $\mathcal{T}$.

Next, we denote by $P_{k,l}(\omega)$ the set of polynomials of degree k in the first and of degree l in the second variable of $x \in \omega \subset \mathbb{R}^2$ and by $P_k(\omega) := P_{k,k}(\omega)$.

In this paper we are concerned with one type only of the mixed finite element spaces associated with each $\mathcal{T}_i$. These are the finite dimensional spaces $W_{h_i} \subset W_i$ and $\mathbf{V}_{h,i} \subset \mathbf{V}_i$. Specifically, the "discrete pressure space" W_{h_i} is the space of functions which are $P_0(E)$, $E \in \mathcal{T}_i$. In turn, the "discrete velocity space" $\mathbf{V}_{h,i}$ is an $\mathbf{RT}_{[0]}$ space, that is, on each element $E \in \mathcal{T}_i$, each $\mathbf{V}_{h,i} \ni \mathbf{q}$, $\mathbf{q}|_E \in P_{1,0}(E) \times P_{0,1}(E)$. We recall that this choice of spaces, for each i, satisfies the Babuska-Brezzi condition ([17], III.3.2). We refer to [17] for more details on these and other mixed finite element spaces.

The *mortar mixed FE solution* is defined as in [7]. First we define $\mathbf{V^{-1}}_h = \bigoplus_{i=1}^n \mathbf{V}_{h,i}$. We note that $\mathbf{V^{-1}}_h \subset \mathbf{V^{-1}}$ but that, in general (if grids are not matching), $\mathbf{V^{-1}}_h \cap \mathbf{V}$ contains not much more than the zero element, as the members of the spaces $\mathbf{V}_{h,i}$ and $\mathbf{V}_{h,j}$ do not have matching normal traces on $\Gamma_{i,j}$.

To handle non-matching grids and to define some form of a coupling across the interface the *mortar spaces* were introduced. The goal is, on one hand, to relax the continuity of the discrete solution across Γ which cannot be satisfied across Γ because of non-matching grids, and on the other hand, to couple the subdomain solutions. In addition, introduction of the mortar unknowns (or Lagrange multipliers) on the interface Γ allows one to formulate the discrete problem for (1) in terms of the interface unknowns and to entirely decouple the subdomain (local) problems from one another; their interaction is only through the mortar unknowns. For the conforming mortar methods see, e.g., [14, 12] and references therein. In addition, see [30] and references therein for discussion of iterative methods on the interface with optimal preconditioners.

For mixed mortar methods the Lagrange multiplier "mortar pressures" supply the Dirichlet boundary conditions on Γ applied to subdomain problems. In turn, the Neumann data (values of fluxes across Γ) need to be matched somehow to achieve the aforementioned global flow. In fact, averages of their projections onto some space are matched: this condition relaxes the requirement $\mathbf{v}_h \in \mathbf{V}$ which cannot be satisfied because of nonmatching grids, and, on the other hand, it restricts the functions from $\mathbf{V^{-1}}_h$ so that some global mass conservation across Γ can be achieved.

Following [7], in the mixed mortar method the *mortar pressures* are defined as members of a finite dimensional subset $\Lambda_{h_m} \subset L^2(\Gamma)$ associated with the *mortar grid* $\mathcal{T}_m$ which is defined locally on each Γ_{ij}, with a discretization parameter h_m. Recall $n_m \approx \frac{1}{h_m}$ is the number of mortar degrees of freedom. We consider the spaces $\Lambda_{h_m,ij} \subset L^2(\Gamma_{ij})$ spanned by continuous piecewise linears over the grid $\mathcal{T}_m$ restricted to Γ_{ij}, and their direct sum $\Lambda_{h_m} := \bigoplus_{i,j} \Lambda_{h_m,ij}$.

The mortar grid parameter h_m may, in principle, be defined *independently* of the subdomain discretization parameter(s) h_i as long as certain approximation properties to be discussed in Section 2.3 are satisfied.

The space Λ_{h_m} provides the means to define a weak continuity condition for normal traces which replaces (16) so that not the jump (15) itself but rather its weighted average with respect to the discrete weights from Λ_{h_m} vanishes

$$\mathbf{V^0} = \{\mathbf{v} \in \mathbf{V^{-1}} \quad : \quad \sum_i < \mathbf{v}\cdot\mathbf{n}_i, \mu >_{\Gamma_i} = \sum_{i,j} < [\mathbf{v}\cdot\mathbf{n}]_{ij}, \mu >_{\Gamma_{ij}} = 0, \quad \forall \mu \in \Lambda_h\} \tag{17}$$

Note that $\mathbf{V} \subset \mathbf{V^0} \subset \mathbf{V^{-1}}$.

Also, we define $\mathbf{V^0}_h := \mathbf{V^{-1}}_h \cap \mathbf{V^0}$. As stated in [7], $\mathbf{V^0}_h$ is difficult to characterize directly, but it is nonempty, and it provides the appropriate 'home' for an optimally convergent approximation of $\mathbf{u}$. More precisely, in the *discrete mixed mortar* problem formulation one seeks $\mathbf{u}_h \in \mathbf{V^0}_h,\ p_h \in W_h,\ \lambda_h \in \Lambda_{h_m}$, which solve the discrete equivalents of local problems (12), (13).

$$\begin{aligned} (K^{-1}\mathbf{u}_h, \mathbf{v}_h)_i &= (p_h, \nabla\cdot\mathbf{v}_h)_i \\ &- < \lambda_{h_m}, \mathbf{v}_h\cdot\nu_i >_{\Gamma_i} - < g, \mathbf{v}_h\cdot\nu >_{\partial\Omega_i\setminus\Gamma}, \quad \mathbf{v}_h \in \mathbf{V}_{h,i}, \end{aligned} \tag{18}$$

$$(\nabla\cdot\mathbf{u}_h, w_h)_i = (f, w_h)_i, w_h \in W_{h,i}, \tag{19}$$

These local equations are coupled by the mortar pressures λ_h in (18) and by the condition of weak continuity of fluxes $\mathbf{u}_h \in \mathbf{V^0}_h$.

Equivalently, the system can be rewritten by adding over all subdomains and by using test functions from $\mathbf{V^0}_h$, whereby the resulting jump term involving mortar "pressure" λ_{h_m} arising on the rhs of (18) is eliminated by virtue of weak continuity of the test functions. The attractive feature of such a rewrite is that it takes appearance of a global system similar to (9), (10), with the coupling conditions imposed on the test functions: $\mathbf{u}_h \in \mathbf{V^0}_h, p_h \in W_h$ such that

$$(K^{-1}\mathbf{u}_h, \mathbf{v}_h) = (p_h, \nabla\cdot\mathbf{v}_h) - < g, \mathbf{v}_h\cdot\mathbf{n} >_{\partial\Omega} \quad \mathbf{v}_h \in \mathbf{V^0}_h, \tag{20}$$

$$(\nabla\cdot\mathbf{u}_h, w_h) = (f, w_h), \quad w_h \in W_h. \tag{21}$$

If we assume (A1)–(A5) and additional assumptions on h_m to be discussed below, then the scheme (20)–(21) has a unique solution as shown in ([7], Lemma 2.1) and is asymptotically convergent at an optimal rate $O(h)$ in p and u as shown by a-priori estimates for the velocity and pressure in ([7], Thereoms 4.2, 5.1), respectively. We discuss the assumptions on h_m next.

2.3. Parametrization of the mortar grid. Here we consider the parametrization of the mortar grid $\mathcal{T}_m$ with respect to the subdomain grids $\mathcal{T}_h$. Recall that these are associated, respectively, with parameters h_m and h. Assume

$$h_m = h_m(h) \tag{22}$$

where h_m is monotone function of h. Obviously, we are interested in h_m depending linearly or less on h in order to keep n_m from growing too much.

In [7] the following two assumptions on $h_m(h)$ are made. Recall the definition of the L^2 projection $Q_{h,i} : L^2(\Gamma_i) \mapsto \mathbf{V}_{h,i} \cdot \nu_i|_{\Gamma_i}$ and its "global"counterpart Q_h coming from the direct sum over all subdomains. We also consider the projection P_{h_m} defined from $L^2(\Gamma)$ onto Λ_{h_m}.

Assumption 5. *([7], 3.18) Assume that for any $\mu \in \Lambda_h$ there is a constant C_I independent of h such that*

$$\| \mu \|_{0,\Gamma_{i,j}} \leq C_I(\| Q_{h,i}\mu \|_{0,\Gamma_{i,j}} + \| Q_{h,j}\mu \|_{0,\Gamma_{i,j}}) \tag{23}$$

on every interface $\Gamma_{i,j}, i \neq j$.
([7], 3.1) Also assume that for $\psi \in L^2(\Gamma)$,

$$\| \psi - P_{h_m}\psi \|_{-s,\Gamma_{ij}} \leq C_{II} \| \psi \|_{r,\Gamma_{ij}} h^{r+s} \tag{24}$$

These two assumptions provide, respectively, a lower bound for $h_m(h)$ for any fixed h, and an upper bound for $h_m(h)$ as $h \to 0$. The first assumption guarantees unique solvability of (18)–(19) and the second guarantees the optimal approximation properties of the mortar space which in turn are necessary for the optimal convergence of the mortar algorithm.

Remark 2.1. *Since the approximation properties of P_{h_m} depend only on the grid parameter h_m, condition* (24) implicitly *requires*

$$h_m = N_M h \tag{25}$$

where $N_M \equiv const$.
Furthermore, the computational examples in [7] *recommend $N_M \approx 2$ by stating that "the mortar grid on each interface is a coarsening by two in each direction of the trace of either one of the subdomain grids".*

Below we show that that the choice of $N_M \equiv const > 2$ implies (23) and by (25), naturally also (24). Next we show that a sublinear nonconstant dependence $h_m(h)$ may yield convergence of the mortar algorithm, albeit at suboptimal rate. In summary, for a fixed h the mortar grid parameter h_m should be chosen adaptively as shown in Section 4.

2.3.1. Lower bound on mortar parametrization. Here we explore the dependence of C_I in (23) on h_m by way of a simple example. For simplicity assume $|\Gamma_{ij}| = 1$ or that all calculations are done in the reference domain $(0, 1)$. Also, assume matching grids $h = h_i = h_j$, and (25) with $N_M \in \mathbb{N}$.

Consider $\mu \in \Lambda_h$. We have $\mu(x) = \sum_k \psi_k(x)\mu_k$, where ψ_k are the usual continuous piecewise linear nodal basis functions associated with nodes of the mortar grid on $\Gamma_{i,j}$ and where μ_k are the weights. We have that the support of a basis function $supp(\psi_k) = [(k-1)h_M, kh_m]$. On the other hand, $Q_{h,i}\mu(x) = \sum_k \mu_k Q_{h,i}\psi_k(x)$ and $Q_{h,i}\psi_k(x)$ is a linear combination of the characteristic functions over the subdomain grid $\chi_m(x) =: \chi_{[(m-1)h,mh]}(x)$ with weights given by midpoint values of ψ_k over the support of each χ_m. To derive C_I we compute

$$\| \mu \|^2_{0,\Gamma_{i,j}} = \sum_k \mu_k^2 \int_{supp(\psi_k)} (\psi_k(x))^2 dx, \tag{26}$$

$$\| Q_{h,i}\mu \|^2_{0,\Gamma_{i,j}} = \sum_{k,m} \mu_k^2 \int_{supp(\psi_k)\cap supp(\chi_m)} ((Q_{h,i}\psi_k(x))^2 dx. \tag{27}$$

It is straighforward to compute

$$\int_{supp(\psi_k)} (\psi_k(x))^2 dx = \frac{h_m}{3}, \tag{28}$$

and

$$\int_{supp(\psi_k)}((Q_{h,i}\psi_k(x))^2dx = \sum_{m=1}^{N_M}\int_{supp(\psi_k)\cap supp(\chi_m)}((Q_{h,i}\psi_k(x))^2dx$$
$$= \frac{h^3}{h_m^2}(\frac{N_M^3}{3} - \frac{N_M}{12} = \frac{h_m}{3}(1-\frac{1}{4N_M^2}),$$

From these calculations, comparing Equations (26) and (27) we get

$$C_I(N_M) = (2(1-\frac{1}{4N_M^2}))^{-1/2}. \tag{29}$$

This example shows that C_I is a decreasing function of N_M. In order for (23) to be satisfied, one should choose $N_M \geq N_{M,min}$; this is consistent with the recommendation in ([7], Remark 2.2) that mortar grid should not be too fine with respect to the subdomain grids.

2.3.2. Upper bound on parametrization of h_m. As we mentioned, the approximation properties of the projection operator P_h onto piecewise linears Λ_{h_m} over grid $\mathcal{T}_m$ depend on h_m through the estimate, for smooth enough ψ,

$$\| \psi - P_h\psi \|_{-s,\Gamma_{i,j}} \leq C \| \psi \|_{r,\Gamma_{ij}} h_m^{r+s}$$

If (25) holds and N_M is fixed as $h \to 0$, then (24) holds and for nonnegative $r+s$, C_{II} in (24) is an non-decreasing function of N_M.

2.3.3. Sublinear parametrization of $h_m(h)$. Assume first $h \to 0$ with $h_m \equiv const.$ Then $N_M = N_M(h)$ blows up and so does the constant C_{II}.

Now assume that h_m is a monotone sublinear function of h

$$h_m = O(h^l), \tag{30}$$

with $0 < l \leq 1$ (when $l = 1$ we have (25)).

Consider the use of (24) with (30) instead of (25) in the proofs of the convergence results in [7]. We find that, for $\mathbf{RT}_{[0]}$ spaces, the assumption (30) instead of (24) leads to the order of convergence for velocities of order

$$O(h_m^{3/2}h^{-1/2}) = O(h^{\frac{1}{2}(3l-1)})$$

hence, the mixed mortar scheme converges as long as $l > \frac{1}{3}$. We skip the details and remark that obviously, the convergence of the scheme is suboptimal for $l < 1$ asymptotically when $h \to 0$.

Remark 2.2. *From now on we shall assume that* (23), (24) *hold. In particular, we shall assume that* (25) *holds locally on each part Γ_{ij} of the union of interfaces Γ. The linear dependence of h_m on h will be reflected in notation $\Lambda_h \equiv \Lambda_{h_m}$ which we adopt from now on. In order to determine optimal value of n_m or h_m per each interface Γ_{ij} we shall establish a-posteriori error indicators.*

3. Residual error estimator for space $\mathbf{RT}_{[0]}$ and for non-homogeneous Dirichlet conditions

In this section we formally define an *a-posteriori* error estimator for the velocity in the mixed formulation for $\mathbf{RT}_{[0]}$ spaces and non-homogeneous boundary condition. Specifically, we consider the following Dirichlet problem in a weak mixed form $\mathbf{u} \in \mathbf{V}, p \in W$ such that

$$(K^{-1}\mathbf{u}, \mathbf{v}) = (p, \nabla\cdot\mathbf{v}) - < g, \mathbf{v}\cdot\nu >_{\partial\Omega}, \quad \forall \mathbf{v} \in \mathbf{V}, \tag{31}$$

$$(\nabla\cdot\mathbf{u}, w) = (f, w), \quad \forall w \in W, \tag{32}$$

and its mixed FE solution, over a conforming partition $\mathcal{T}_h$ of Ω into rectangular elements E, with $\mathbf{V}_h, W_h$ being $\mathbf{RT}_{[0]}$ spaces: $\mathbf{u}_h \in \mathbf{V}_h, p \in W_h$ such that

$$(K^{-1}\mathbf{u}_h, \mathbf{v}_h) = (p_h, \nabla \cdot \mathbf{v}_h) - < g_h, \mathbf{v}_h \cdot \nu >_{\partial\Omega}, \quad \forall \mathbf{v}_h \in \mathbf{V}_h, \tag{33}$$

$$(\nabla \cdot \mathbf{u}_h, w_h) = (f, w_h), \quad \forall w_h \in W_h. \tag{34}$$

This setup and the results of this section can be applied in two different ways. First, (31)-(32) can be seen as the global weak problem (9)–(10) with $n = 1$ posed over Ω, and (33)-(34) as its discrete counterpart (20)-(21), each enhanced by non-homogeneous boundary conditions g and g_h, respectively, with $g \not\equiv g_h$.

Second, this setup can be seen as (31)-(32) representing the local continuous problem (12)-(13) posed over Ω_i, derived from the global problem, for some fixed i, with $1 \leq i \leq n, n > 1$. Similarly, (33)-(34) is its discrete counterpart (18)-(19). In this context we assume that on the Γ_i part of the boundary $\partial\Omega_i$ the value g is given as a trace of the global solution $p|_{\Gamma_i}$ and is smooth enough. Likewise, in the discrete problem, we assume that a fixed $\lambda_h|_{\Gamma_i}$ which plays the role of g_h is given, This interpretation is convenient for our subsequent analysis of the mortar formulation.

The results of this section are essentially the same for both applications therefore they are combined together. We shall now assume that a generic g and g_h are given and proceed to develop an a-posteriori error estimator and show that it provides an upper bound for the error.

Here we recall that a-posteriori error estimators for Galerkin finite elements have been introduced first in [11] and have been a topic of many theoretical and applied studies; see the recent monographs [48, 4] and references therein. A-posteriori estimators for mixed finite elements were first developed for the Laplace problem in [16] where mesh-dependent norms were used, and just recently for (1) in [56, 18]. Here we follow these two latter works.

Both papers [56, 18] develop an estimator for velocity for 2D problems with $g \equiv g_h \equiv 0$. The residual estimator in [56] is appropriate for triangular grids and $\mathbf{RT}_0$ spaces, and in [56] it is shown to be locally (and globally) equivalent to other estimators. The estimator in [18] is also appropriate for triangular grids but is defined for various higher order mixed spaces; it reduces to the one in [56] in the case of $\mathbf{RT}_0$ spaces. Both of the estimator(s) of interest to this paper are of residual type, and their construction and analysis are based on the Helmholtz decomposition of the vectors in $\mathbb{R}^2$ into their solenoidal and (weakly) irrotational parts. Specifically, the space $\mathbf{H}(\mathbf{div}; \Omega)$ is decomposed into

$$\mathbf{H}(\mathbf{div}; \Omega) = (\mathbf{H}(\mathbf{div}; \Omega) \cap \mathbf{X^0}) \cup (\mathbf{H}(\mathbf{div}; \Omega) \cap \mathbf{X^K}), \tag{35}$$

$$\mathbf{X^0} = \{\mathbf{q} \in \mathbf{H}(\mathbf{div}; \Omega) : \nabla \cdot \mathbf{q} = 0\}, \tag{36}$$

$$\mathbf{X^K} = \{\mathbf{r} \in \mathbf{H}(\mathbf{div}; \Omega) : \int_\Omega K^{-1}\mathbf{r} \cdot \mathbf{q} = 0, \forall \mathbf{q} \in \mathbf{X^0}\}. \tag{37}$$

This decomposition is crucial for the derivation of the estimator and for the proof of the upper bounds.

Here our main objective is to treat rectangular grids, $\mathbf{RT}_{[0]}$ spaces, and non-homogeneous boundary terms g and g_h. While extension of the results in [56, 18] to our case is in itself not difficult, we supply all the details for the sake of completeness. This allows, in particular, to see where the generalization of these results to more general case(s) fails.

The structure of the results below follows the one in similar works: first we define the residual $r(\mathbf{v})$ as a linear functional on an element $\mathbf{v} \in \mathbf{V}$ which, when evaluated

for $\mathbf{v} = \mathbf{e}$ with the error in velocity $\mathbf{e} := \mathbf{u} - \mathbf{u}_h$ gives rise to the natural norm $\| \mathbf{e} \|_K$ of the error.

Then we perform integration by parts, which allows to extract, in each term of the estimator, a part which can be calculated directly from the finite element solution and which gives rise to the definition of the estimator η, and a part which can be estimated using some norm of the error. The estimates are achieved with help of a local regularization operator. Recall that in [56, 18] the Clément operator P_C appropriate for triangular meshes is used. In this paper we use the local regularization P_{BG} defined by Bernardi and Girault in [13] which works for triangular and quadrilateral meshes. Since the residual is akin to the square of this natural norm of the error, this procedure leads to an upper bound for the error

$$\| \mathbf{e} \|_K \leq C\eta. \tag{38}$$

3.1. Definition of residual and residual calculations. Here we follow the notation and assumptions introduced in Section 2.2.2 and define the residual $r(\mathbf{v})$, for $\mathbf{v} \in \mathbf{H}(\mathbf{div}; \Omega)$, as the following linear functional

$$r(\mathbf{v}) \quad := \quad -(K^{-1}\mathbf{u}_h, \mathbf{v}) + (p_h, \nabla \cdot \mathbf{v}) - < g_h, \mathbf{v} \cdot \mathbf{n} >_{\partial\Omega} . \tag{39}$$

Using the quadrangulation $\mathcal{T}_h$ of Ω, denoting by e the edges of $E \in \mathcal{T}_h$ and writing $\partial\Omega = \bigcup_{\mathbf{e}\cap\partial\Omega} e$ we can write

$$r(\mathbf{v}) \quad = \quad -\sum_E (K^{-1}\mathbf{u}_h, \mathbf{v})_E + \sum_E (p_h, \nabla \cdot \mathbf{v})_E - \sum_{e\cap\partial\Omega} < g_h, \mathbf{v} \cdot \mathbf{n} >_e$$

By (33) we note that $r(\mathbf{v}_h) = 0, \ \forall \mathbf{v}_h \in \mathbf{V}_h$.

We proceed to define local (elementwise) contributions to the residual. The key is the above mentioned decomposition of the space $\mathbf{H}(\mathbf{div}; \Omega)$. For any $\mathbf{v} \in \mathbf{H}(\mathbf{div}; \Omega)$, we have $\mathbf{v} = \mathbf{v}^0 + \mathbf{v}^K$ where $\mathbf{v}^0 \in \mathbf{H}(\mathbf{div}; \Omega) \cap \mathbf{X^0}$ and where $\mathbf{v}^K \in \mathbf{H}(\mathbf{div}; \Omega) \cap \mathbf{X^K}$. In addition, a solenoidal vector $\mathbf{v}^0 = \mathbf{Curl}\phi$ for some $\phi = \phi(\mathbf{v}^0) \in H^1(\Omega)$, with ϕ determined up to a constant. In fact, we fix ϕ to be such that $(\phi, 1) = 0$ that is, to have zero average over Ω. This decomposition of $\mathbf{H}(\mathbf{div}; \Omega)$ is discussed in [56] and [18].

Now one calculates by integration by parts over an element E, for any vector $\mathbf{q} = (q_1, q_2)$, and a test function $\mathbf{v}^0 \in \mathbf{H}(\mathbf{div}; \Omega) \cap \mathbf{X^0}$, and an associated $\phi = \phi(\mathbf{v}^0)$

$$\begin{aligned} -\int_E \mathbf{q} \cdot \mathbf{v}^0 dx \quad &= \quad -\int_E \mathbf{q} \cdot \mathbf{Curl}\phi dx = \int_E (q_{1,2}\phi - q_{2,1}\phi)dx - \int_{\partial E} (q_1 n_2 \phi - q_2 n_1 \phi) ds \\ &= \quad \int_E curl\mathbf{q}\phi dx - \sum_{e\cap\partial E} \int_e \mathbf{q} \cdot \mathbf{t}\phi ds. \end{aligned}$$

As it is done in [56], for $\mathbf{q} = K^{-1}\mathbf{u}_h \in \mathbf{RT}_0$, the integral $\int_E curl\mathbf{q}\phi dx$ vanishes. Indeed, $\mathbf{RT}_0$ elements have the general form of $(P_0(E))^2 + xP_0(E)$. In other words, we have $\mathbf{q}|_E = (K_{1,1}^{-1}(\alpha_1 + \beta x_1), K_{2,2}^{-1}(\alpha_2 + \beta x_2))$ for some element-wise chosen constants $\alpha_1, \alpha_2, \beta$ and trivially $q_{1,2} = 0, q_{2,1} = 0$. If K is nondiagonal but symmetric, then still $q_{1,2} - q_{2,1} = 0$. However, if K is non-symmetric and/or when spaces of higher order than $\mathbf{RT}_0$ are used, then the terms involving $curl(K^{-1}\mathbf{u}_h)|_E$ may not vanish (see [18]).

For $\mathbf{q} \in \mathbf{RT}_{[0]}$ which is the space of interest in this paper, we have that elementwise $\mathbf{q}|_E = (K_{1,1}^{-1}(\alpha_1 + \beta_1 x_1), K_{2,2}^{-1}(\alpha_2 + \beta_2 x_2))$, and by reasoning as above, if K

is diagonal, $curl(\mathbf{q}) = curl(K^{-1}\mathbf{u}_h) = 0$. In summary, we get

$$-\int_E K^{-1}\mathbf{u}_h \cdot \mathbf{v}^0 dx = \sum_{e\subset\partial E}\int_e (K^{-1}_{1,1}(u_h)_1 n_2\phi - K^{-1}_{2,2}(u_h)_2 n_1\phi)ds$$
$$= \sum_{e\subset\partial E}\int_e K^{-1}\mathbf{u}_h \cdot \mathbf{t}\phi ds.$$

Now we take a sum over all elements E in the calculation of the residual value $r(\mathbf{v}_0)$ by (39), with $\mathbf{v}^0 \in \mathbf{H}(\mathbf{div};\Omega)\cap\mathbf{X^0}$ and note that the second term in (39) vanishes by virtue of $\nabla\cdot\mathbf{v}^0 = 0$, and we get

$$r(\mathbf{v}^0) = \sum_E\sum_{e\subset\partial E}\int_e K^{-1}\mathbf{u}_h\cdot\mathbf{t}\phi ds - \sum_{e\cap\partial\Omega} < g_h, \mathbf{v}^0\cdot\mathbf{n} >_e \tag{40}$$
$$= \sum_{e\setminus\partial\Omega}\int_e [K^{-1}\mathbf{u}_h\cdot\mathbf{t}]\phi ds + \sum_{e\cap\partial\Omega}\int_e (K^{-1}\mathbf{u}_h\cdot\mathbf{t}\phi - g_h\mathbf{v}^0\cdot\mathbf{n})ds$$

where the jump $[\cdot]$ is defined as in (15) on the edges $e\setminus\partial\Omega$ of elements E which are interior to Ω.

3.1.1. Calculation of $\| \mathbf{e} \|_K$. Decompose the error $\mathbf{e} = \mathbf{e}^0 + \mathbf{e}^K$ and note that we have, by the definition of the norms (2) and properties of $\mathbf{e}^0$ and $\mathbf{e}^K$,

$$\| \mathbf{e} \|^2_K = (K^{-1}(\mathbf{e}^0+\mathbf{e}^K), \mathbf{e}^0+\mathbf{e}^K) + (\nabla\cdot\mathbf{e}^K, \nabla\cdot\mathbf{e}^K) \tag{41}$$
$$= (K^{-1}\mathbf{e}^0,\mathbf{e}^0) + (K^{-1}\mathbf{e}^K,\mathbf{e}^K) + (\nabla\cdot\mathbf{e}^K,\nabla\cdot\mathbf{e}^K) = \| \mathbf{e}^0 \|^2_K + \| \mathbf{e}^K \|^2_K .$$

Next, by subtracting (34) from (32) which, from the properties of the $L^2(\Omega)$ projection $\Pi_0 : W \mapsto W_h$ can be rewritten as $(\nabla\cdot\mathbf{u}_h, w) = (\nabla\cdot\mathbf{u}_h, \Pi_0 w) = (f, \Pi_0 w) = (\Pi_0 f, w)$, we get

$$(\nabla\cdot\mathbf{e}, w) = (\nabla\cdot(\mathbf{u}-\mathbf{u}_h), w) = (f - \Pi_0 f, w), \quad \forall w \in W.$$

By applying Cauchy-Schwarz inequality twice we get

$$\| \nabla\cdot\mathbf{e} \|_0 = \| f - \Pi_0 f \|_0$$

which is also true on any element E so that

$$\| \nabla\cdot\mathbf{e} \|_{0,E} = \| f - \Pi_0 f \|_{0,E} .$$

Finally we see that $\| \nabla\cdot\mathbf{e}^K \|_0 = \| \nabla\cdot\mathbf{e} \|_0 = \| f - \Pi_0 f \|_0$. In the sequel we shall use the fact that on $\mathbf{H}(\mathbf{div};\Omega)\cap\mathbf{X^K}$ the natural norm in $\mathbf{H}(\mathbf{div};\Omega)$ is equivalent to the norm $\| \nabla\cdot^K \|_0$, hence, $\| \mathbf{e}^K \|_{\mathbf{H}(\mathbf{div};\Omega)} \le C_{eq} \| \nabla\cdot\mathbf{e}^K \|_0$, and by (11) we have

$$\| \mathbf{e}^K \|_K \le \max(1, K^{-1}_\Omega) C_{eq} \| \nabla\cdot\mathbf{e}^K \|_0 \le C_K \| f - \Pi_0 f \|_0 \tag{42}$$

where we have set $C_K := \max(1, K^{-1}_\Omega) C_{eq}$.

Now we compute the norm of $\| \mathbf{e}^0 \|_K$ using the definition of the residual $r(\mathbf{v})$. By adding $r(\mathbf{v})$ to both sides of (31), we see that the error satisfies

$$(K^{-1}(\mathbf{u}-\mathbf{u}_h), \mathbf{v}) - (p-p_h, \nabla\mathbf{v}) + < g-g_h, \mathbf{v}\cdot\mathbf{n} >_{\partial\Omega} = r(\mathbf{v}), \quad \mathbf{v}\in\mathbf{H}(\mathbf{div};\Omega)$$

hence, restricting the test functions to $\mathbf{X^0}$ we get

$$(K^{-1}(\mathbf{u}-\mathbf{u}_h), \mathbf{v}^0) = - < g-g_h, \mathbf{v}^0\cdot\mathbf{n} >_{\partial\Omega} + r(\mathbf{v}^0), \quad \mathbf{v}^0 \in \mathbf{H}(\mathbf{div};\Omega)\cap\mathbf{X^0}.$$

Now use $\mathbf{v} = \mathbf{e}$ to get, by (2) and the above calculations

$$\| \mathbf{e}^0 \|^2_K = (K^{-1}\mathbf{e}, \mathbf{v}^0) = (K^{-1}(\mathbf{u}-\mathbf{u}_h), \mathbf{v}^0) \tag{43}$$
$$= - < g-g_h, \mathbf{v}^0\cdot\mathbf{n} >_{\partial\Omega} + r(\mathbf{v}^0).$$

Upon inserting the expression (40) into (43) and canceling terms we obtain, with a $\phi = \phi(\mathbf{v}^0)$

$$\begin{aligned} \| \mathbf{e}^0 \|_K^2 &= \sum_{e \setminus \partial\Omega} \int_e [K^{-1}\mathbf{u}_h \cdot \mathbf{t}]\phi ds \\ &+ \sum_{e \cap \partial\Omega} \int_e (K^{-1}\mathbf{u}_h \cdot \mathbf{t}\phi(\mathbf{v}^0) - g\mathbf{v}^0 \cdot \mathbf{n}) ds \end{aligned} \tag{44}$$

Now we need to decompose these expressions as products of a "computable" quantity and of one that can be estimated by $\| \mathbf{e} \|_K$.

3.2. Local regularization operator. To achieve this aim, one considers a suitable finite dimensional projection $\mathbf{v}_h^0$ of $\mathbf{v}^0$. In this context, in [56, 18], the Clément regularization operator (quasi-interpolant) $P_C\phi$ of ϕ is used which, locally $P_C\phi|_E \in P_1(E)$, so that $\mathbf{Curl}\phi_h \in \mathbf{RT}_0$. In other words, P_C is "compatible" with the mixed space $\mathbf{RT}_0$. Also, since in [56, 18] $g \equiv g_h \equiv 0$, one has $r(\mathbf{Curl}\phi_h) = 0$, and the term $r(\mathbf{Curl}\phi_h)$ can be subtracted from $r(\mathbf{v}^0)$ without changing $\| \mathbf{u} - \mathbf{u}_h \|_K^2$.

In our case, for $\mathbf{RT}_{[0]}$ spaces, we use $\phi_h := P_{BG}\phi$ where P_{BG} is the local regularization operator on quadrilateral grids defined in [13]. In case of rectangular elements, $\phi_h := P_{BG}\phi$ is the unique element of $Q_1(E)$ constructed, roughly speaking, by averaging the values of ϕ over the neighborhood $\triangle E$ of the element E. See [13] for details which in fact are carried through in the reference domain. The approximation $\phi_h := P_{BG}\phi$ of ϕ satisfies the local estimates (respectively, as ([13], eq. 4.2, 4.10), for $k = 1$)

$$\| \phi - \phi_h \|_{0,E} \leq c_{BG} h_E |\phi|_{1,\triangle E} \tag{45}$$

$$\| \phi - \phi_h \|_{0,e} \leq C_{BG}\sqrt{h_E}\, \| \phi \|_{1,\triangle E}, \quad \forall e \in \partial E \tag{46}$$

There exists a natural "compatibility" between P_{BG} and $\mathbf{RT}_{[0]}$ spaces, at least on rectangular meshes. Indeed, for any $\phi_h \in Q_1(E)$ if each E is a rectangle, we have $\mathbf{Curl}\phi_h \in \mathbf{RT}_{[0]}$. To see that, consider a basis function $\psi_h \in Q_1(E)$. Recall these are defined as $\psi_h = \hat{\psi}_h \circ F_E^{-1}$ where $\hat{\psi}_h(\hat{x_1}, \hat{x_2})$ is defined on the reference element $\hat{E}$ and is one of bilinear polynomials $(1-\hat{x}_1)(1-\hat{x}_2), \hat{x}_1(1-\hat{x}_2), \hat{x}_1\hat{x}_2, (1-\hat{x}_1)\hat{x}_2$. To show $\mathbf{Curl}\phi_h \in \mathbf{RT}_{[0]}$ it is enough to show that $\mathbf{Curl}\hat{\psi}_h \in \mathbf{RT}_{[0]}$ holds on $\hat{E}$. This can be done by computing $\nabla\hat{\psi}_h$ from which directly we see that $\mathbf{Curl}\hat{\phi}_h = T\nabla\hat{\phi}_h \in \mathbf{RT}_{[0]}$. See also [17], Lemma III.3.3.

With this compatibility established, for a given $\mathbf{v}^0$ and its associated $\phi(\mathbf{v}^0)$, we let $\phi_h = P_{BG}(\phi)$, and denote $\mathbf{v}_h^0 := \mathbf{Curl}\phi_h$. Since $\mathbf{v}_h^0 \in \mathbf{RT}_{[0]}$, by (33), we have $r(\mathbf{v}_h^0) = 0$. Thus, we can subtract $r(\mathbf{v}_h^0)$ from the rhs of (43) to find

$$\begin{aligned} \| \mathbf{e}^0 \|_K^2 &= - < g - g_h, \mathbf{v}^0 \cdot \mathbf{n} >_{\partial\Omega} + r(\mathbf{v}^0) - r(\mathbf{v}_h^0) \\ &= - < g - g_h, \mathbf{v}^0 \cdot \mathbf{n} >_{\partial\Omega} + r(\mathbf{v}^0 - \mathbf{v}_h^0) \\ &= \sum_{e \setminus \partial\Omega} \int_e [K^{-1}\mathbf{u}_h \cdot \mathbf{t}](\phi - \phi_h) ds \\ &+ \sum_{e \cap \partial\Omega} \int_e (K^{-1}\mathbf{u}_h \cdot \mathbf{t}(\phi - \phi_h) \\ &- g_h(\mathbf{v}^0 - \mathbf{v}_h^0) \cdot \mathbf{n} - (g - g_h)\mathbf{v}^0 \cdot \mathbf{n}) ds \\ &:= \sum_{e \setminus \partial\Omega} \int_e (I_{int}) ds + \sum_{e \cap \partial\Omega} \int_e (B_I + B_{II} + B_{III}) ds \end{aligned} \tag{47}$$

If $g \equiv g_h \equiv 0$, then all the boundary terms B_I, B_{II}, B_{III} vanish and the interior terms are estimated as in [56, 18].

3.3. Interior estimates. Use $\mathbf{v}^0 = \mathbf{u}^0 - \mathbf{u}_h^0$, by Cauchy-Schwarz inequality and (46)

$$\begin{aligned} \sum_{e\setminus\partial\Omega}\int_e I_{int}ds &= \sum_{e\setminus\partial\Omega}\int_e [K^{-1}\mathbf{u}_h\cdot\mathbf{t}](\phi-\phi_h)ds \\ &\leq \sum_{e\setminus\partial\Omega} \parallel [K^{-1}\mathbf{u}_h\cdot\mathbf{t}] \parallel_{0,e} \parallel (\phi-\phi_h) \parallel_{0,e} \\ &\leq \sum_{e\setminus\partial\Omega} \parallel [K^{-1}\mathbf{u}_h\cdot\mathbf{t}] \parallel_{0,e} C_{BG}\sqrt{h_E} \parallel \phi \parallel_{1,\triangle E} \end{aligned} \tag{48}$$

Now, by Poincaré inequality (recall that the average of ϕ over Ω is zero) we have $\parallel \phi \parallel_{0,\Omega} \leq C_P|\phi|_{1,\Omega} = C_P \parallel \mathbf{v}^0 \parallel_{0,\Omega} \leq \frac{C_P}{\sqrt{\kappa_\Omega}} \vert\!\vert\!\vert \mathbf{v}^0 \vert\!\vert\!\vert_{K,\Omega}$. With this, and by Cauchy-Scharz inequality for sums, and noting that there can be at most four elements in $\triangle E$ for every E, the terms in (48) can be estimated as follows

$$\sum_{e\setminus\partial\Omega}\int_e I_{int}ds \leq C_{int}\frac{1}{\sqrt{\kappa_\Omega}}\sqrt{\sum_{e\setminus\partial\Omega} h_E \parallel [K^{-1}\mathbf{u}_h\cdot\mathbf{t}] \parallel_{0,e}^2} \vert\!\vert\!\vert \mathbf{v}^0 \vert\!\vert\!\vert_{K,\Omega}, \tag{49}$$

where $C_{int} := C_{BG}\sqrt{4(1+C_P^2)}$ and which is essentially the same as appropriate terms in [56, 18]. We note in passing that, if the norm on the rhs of (46) can be replaced by the 1-seminorm, (as in [23] or in [56]), one can improve the above bound by replacing κ_Ω by a local estimate of $\kappa_{\triangle E}$ and the error estimte, instead of (49) would read as follows

$$\sum_{e\setminus\partial\Omega}\int_e I_{int}ds \leq C_{int}\sqrt{\sum_{e\setminus\partial\Omega} \frac{h_E}{\kappa_E} \parallel [K^{-1}\mathbf{u}_h\cdot\mathbf{t}] \parallel_{0,e}^2} \vert\!\vert\!\vert \mathbf{v}^0 \vert\!\vert\!\vert_{K,\Omega}\,.$$

At this time however we are unable to determine whether it can be done without further loss of generality.

3.4. Boundary terms. Here we consider the general case of $g \not\equiv g_h \not\equiv 0$ and proceed to estimate the boundary terms B_I, B_{II}, B_{III}. We only consider integration over those edges e which lie on $\partial\Omega$ and we drop the notation $\mathbf{e}\cap\partial\Omega$ for brevity.

First we consider the following calculation on an edge e which belongs to $\partial\Omega$. Denote the two endpoints of e by P_e^1, P_e^2. Let ρ and ψ be two functions defined on e, smooth in the interior of e and defined meaningfully pointwise at its endpoints. We have by integration by parts,

$$\int_e \rho\mathbf{Curl}\psi\cdot\mathbf{n}ds = \rho\psi|_{P_e^1}^{P_e^2} - \int_e \mathbf{Curl}\rho\cdot\mathbf{n}\psi ds = \rho\psi|_{P_e^1}^{P_e^2} - \int_e(-\nabla\rho\cdot\mathbf{t})\psi ds. \tag{50}$$

If ρ and ψ are continuous on $\partial\Omega$, we sum over all edges $e\in\partial\Omega$, the pointwise values cancel and we get

$$\sum_e\int_e \rho\mathbf{Curl}\psi\cdot\mathbf{n}ds = \sum_e\int_e \mathbf{Curl}\rho\cdot\mathbf{n}\psi ds = -\sum_e\int_e \nabla\rho\cdot\mathbf{t}\psi ds. \tag{51}$$

This calculation suggests how to handle terms arising in B_I, B_{II}, B_{III} and how to distinguish different cases depending on the definition and smoothness of g_h.

We consider two cases: (i) g_h is the piecewise linear interpolant of g, and (ii) g_h is the piecewise constant $L^2(\partial\Omega)$ projection of g. In each case the discrete values of g_h are defined with respect to the trace of the triangulation $\mathcal{T}_h$ of Ω on $\partial\Omega$.

3.4.1. Case (i). In the first case, we see that g_h and $\phi - \phi_h$ are smooth enough for calculation (51) to work. Hence, we obtain

$$\sum_e \int_e B_{II} ds = \sum_e \int_e g_h \mathbf{Curl}(\phi - \phi_h) \cdot \mathbf{n} ds = \sum_e \int_e \mathbf{Curl} g_h \cdot \mathbf{n} (\phi - \phi_h) ds$$

and, after it is combined with the term B_I, we get

$$\sum_e \int_e (K^{-1}\mathbf{u}_h \cdot \mathbf{t}(\phi - \phi_h) - g_h(\mathbf{v}^0 - \mathbf{v}_h^0) \cdot \mathbf{n}) ds = \sum_e \int_e (K^{-1}\mathbf{u}_h + \nabla g_h) \cdot \mathbf{t}(\phi - \phi_h) ds$$

which can be estimated in a manner similar to the interior terms in (48), and we get

$$\sum_e \int_e (B_I + B_{II}) ds \leq C_I \frac{1}{\sqrt{\kappa_\Omega}} \sqrt{\sum_e h_e \parallel (K^{-1}\mathbf{u}_h + \nabla g_h) \cdot \mathbf{t} \parallel_{0,e}^2} \parallel\!\!\parallel \mathbf{v}^0 \parallel\!\!\parallel_{K,\Omega},$$

where now $C_I := C_{BG}\sqrt{2(1 + C_P^2)}$ and where h_e denotes the maximum of h_E for elements E adjacent to the element containing e.

Next, to deal with the term B_{III}, we use the average $\bar{\phi}$ of ϕ over each element E adjacent to the edge e and extended by 0 to the rest of Ω. Clearly $\mathbf{Curl}\bar{\phi} \cdot \mathbf{n}|_e = 0$ and we subtract this from the term B_{III} and use (50) in the following

$$\begin{aligned} \int_e (g - g_h)\mathbf{v}^0 \cdot \mathbf{n} ds &= \int_e (g - g_h)(\mathbf{v}^0 - \mathbf{Curl}\bar{\phi}) \cdot \mathbf{n} ds \\ = (g - g_h)(\phi - \bar{\phi})|_{P_e^1}^{P_e^2} &- \int_e \mathbf{Curl}(g - g_h) \cdot \mathbf{n}(\phi - \bar{\phi}) ds \end{aligned}$$

where each of the terms $(g - g_h)(\phi - \bar{\phi})|_{P_e^1}$ vanishes by virtue of g_h interpolating g at the endpoints of e. Hence we get, for the term B_{III},

$$\begin{aligned} \sum_e B_{III} ds &= \sum_e \int_e (g - g_h)\mathbf{v}^0 \cdot \mathbf{n} ds = \sum_e \int_e \mathbf{Curl}(g - g_h) \cdot \mathbf{n}(\phi - \bar{\phi}) ds \\ = \sum_e \int_e -\nabla(g - g_h) \cdot \mathbf{t}(\phi - \bar{\phi}) ds &\leq \sum_e \parallel \nabla(g - g_h) \cdot \mathbf{t} \parallel_{0,e} \parallel \phi - \bar{\phi} \parallel_{0,e} \end{aligned}$$

which follows from the Cauchy-Schwarz inequality. Now, by the approximating property of $\bar{\phi}$

$$\parallel \phi - \bar{\phi} \parallel_{0,e} \leq C_{ap}|e|^{1/2} \parallel \nabla\phi \cdot \mathbf{t} \parallel_{0,E} \leq C_{ap}\sqrt{h_E} \parallel \mathbf{Curl}\phi \cdot \mathbf{n} \parallel_{0,E} = C_{ap}\sqrt{h_E} \parallel \mathbf{v}^0 \parallel_{0,E}$$

where we have used the fact that the length $|e| \leq h_E$, we get,

$$\sum_e \int_e (g - g_h)\mathbf{v}^0 \cdot \mathbf{n} ds \leq C_{ap} \sqrt{\sum_e \frac{h_E}{\kappa_E} \parallel \nabla(g - g_h) \cdot \mathbf{t} \parallel_{0,e}^2} \parallel\!\!\parallel \mathbf{v}^0 \parallel\!\!\parallel_K .$$

Combining all these estimates together we get

$$\sum_{e \cap \partial\Omega} \int_e (B_I + B_{II} + B_{III}) ds \leq \max\{C_I, C_{ap}\} \frac{\sqrt{2}}{\sqrt{\kappa_\Omega}} \tag{52}$$

$$\sqrt{\sum_e h_E \left(\parallel (K^{-1}\mathbf{u}_h + \nabla g_h) \cdot \mathbf{t} \parallel_{0,e}^2 + \parallel \nabla(g - g_h) \cdot \mathbf{t} \parallel_{0,e}^2 \right)} \parallel\!\!\parallel \mathbf{v}^0 \parallel\!\!\parallel_{K,\Omega} .$$

3.4.2. Case (ii): Here we assume that g_h is piecewise constant on $\partial\Omega$ in fact, that it is the L^2 projection of g. In addition, we assume that g is defined pointwise and is continuous on $\partial\Omega$.

Since g_h is the average of g, and since $\mathbf{v}_h^0 \cdot \mathbf{n}$ is constant on each e, we have $\int_e (g-g_h)\mathbf{v}_h^0 \cdot \mathbf{n} ds = 0$ and this latter term can be added to the term B_{III}. Together they can be rewritten as

$$\begin{aligned} \int_e B_{III} ds &= \int_e B_{III} ds + \int_e (g-g_h)\mathbf{v}_h^0 \cdot \mathbf{n} ds \\ &= -\int_e (g-g_h)(\mathbf{v}^0 - \mathbf{v}_h^0) \cdot \mathbf{n} ds \\ &= (g-g_h)(\phi-\phi_h)|_{P_e^1}^{P_e^2} + \int_e \mathbf{Curl}(g-g_h) \cdot \mathbf{n}(\phi-\phi_h) ds. \end{aligned} \tag{53}$$

Consider now the term B_{II}. When integrated by parts, B_{II} gives rise to the same integral $\int_e \mathbf{Curl} g_h \cdot \mathbf{n}(\phi-\phi_h) ds$ as in Case (i) and to pointwise terms

$$g_h(\phi-\phi_h)|_{P_1^e}^{P_2^e}$$

which cannot be assumed to vanish when summed over $\partial\Omega$. But, these pointwise values will cancel with those arising from the term B_{III} as appears from (53). Then there remains the sum of pointwise terms $g(\phi-\phi_h)|_{P_1^e}^{P_2^e}$ which vanishes by virtue of smoothness of g and of $\phi-\phi_h$. In summary, we are left with

$$\begin{aligned} \sum_e \int_e & (B_I + B_{II} + B_{III}) ds \\ &= \sum_e \int_e (K^{-1}\mathbf{u}_h \cdot \mathbf{t}(\phi-\phi_h) - g_h(\mathbf{v}^0 - \mathbf{v}_h^0) \cdot \mathbf{n} - (g-g_h)\mathbf{v}^0 \cdot \mathbf{n}) ds \\ &= \sum_e \int_e (K^{-1}\mathbf{u}_h + \nabla g_h) \cdot \mathbf{t}(\phi-\phi_h) ds + \int_e \mathbf{Curl}(g-g_h) \cdot \mathbf{n}(\phi-\phi_h) ds \end{aligned}$$

whose individual pieces can be estimated in the same way as it was done in Case (i) and one gets

$$\begin{aligned} \sum_e \int_e & (B_I + B_{II} + B_{III}) ds \le C_I \frac{1}{\sqrt{\kappa_\Omega}} \\ & \sqrt{\sum_e h_E (\| K^{-1}\mathbf{u}_h + \nabla g_h) \cdot \mathbf{t} \|_{0,e}^2 + \| \nabla(g-g_h) \cdot \mathbf{t} \|_{0,e}^2)} \, \| \mathbf{v}^0 \|_{K,\Omega}, \end{aligned} \tag{54}$$

3.4.3. Summary. Combining (41), (42), (43), (47), (48), and (52) or (54) we obtain the following main result of this section.

Theorem 3.1. *Assume that $\mathcal{T}_h$ is a regular partition of Ω and that Assumptions 1, 2, 3, 4 are satisfied. Then the following upper bound for the mixed finite element solution of* (33) (34) *for $\mathbf{RT}_{[0]}$ spaces holds*

$$\| e \|_{K,\Omega} \le C(K,\Omega)\eta_\Omega, \tag{55}$$

where the residual estimator η_Ω is defined as

$$
\begin{aligned}
\eta_\Omega^2 &:= \eta_{f,\Omega}^2 + \eta_{int,\Omega}^2 + \eta_{II,\partial\Omega}^2 + \eta_{III,\partial\Omega}^2 \\
\eta_{f,\Omega}^2 &:= \sum_{E\in\mathcal{T}} \| f - \Pi_0 f \|_{0,E}^2, \\
\eta_{int,\Omega}^2 &:= \sum_{e\in\partial\mathcal{T}, e\setminus\partial\Omega} h_e \| [K^{-1}\mathbf{u}_h \cdot \mathbf{t}] \|_{0,e}^2 \\
\eta_{II,\partial\Omega}^2 &:= \sum_{e\cap\partial\Omega} h_e \| (K^{-1}\mathbf{u}_h + \nabla g_h)\cdot \mathbf{t} \|_{0,e}^2 \\
\eta_{III,\partial\Omega}^2 &:= \sum_{e\cap\partial\Omega} h_e \| \nabla(g - g_h)\cdot \mathbf{t} \|_{0,e}^2
\end{aligned}
\tag{56}
$$

and where the constant $C(K,\Omega)$ does not depend on h and where h_e denotes the maximum length of the edges of any rectangle E' adjacent to the rectangle E such that the edge $e \subset \partial E$.

Here we do not attempt to establish a lower bound for the error; this has been done for $\mathbf{RT}_0$ spaces and homogeneous boundary conditions in [56]. The proof of the lower bound in [18] is not done and the reader is referred to [48].

Remark 3.1. *One can also derive a bound for the total error which, in addition to error in velocity includes the error $\| p - p_h \|_0$ in the pressure variable. This is done in* [56] *by decomposing*

$$\| p - p_h \|_0 \leq \| p - \Pi_o p \|_0 + \| \Pi_o p - p_h \|_0 .$$

The first part is estimated on each element from approximation properties of Π_o by $h_E \| \nabla p \|_{0,E}$ which leads to a computable part and an error part by way of $\nabla p = K^{-1}\mathbf{u}_h + K^{-1}(\mathbf{u} - \mathbf{u}_h)$. The second part is estimated in terms of $\mathbf{e}$ by duality, see ([17]*, II.2.7) or* [10]*. In summary, additional interior terms involving $\| K^{-1}\mathbf{u}_h \|_{0,E}$ need to be included in η but this direction will not be pursued here.*

Remark 3.2. *It is interesting to compare the structure of the parts of the estimator to the one for conforming methods. In particular, while the η_f is standard and appears similarly in the Galerkin a-posteriori estimates, the counterpart of η_{int} for conforming methods is the measure of the jump of the normal component of the gradient of the solution. On the other hand, the term η_{II} measures how badly the "discrete" gradient of the boundary data approximates the gradient of the discrete pressure. In this sense, this term is similar to the previous one. Finally, the last term in the estimator is rather standard as it measures the consistency error between g and g_h ("data error"in* [3]*).*

Most important for our subsequent analysis of mortar formulation are the η_{II} and η_{III} parts of the estimator.

4. Residual error estimator for mortar formulation

In this section we define an a-posteriori error estimator $\eta(n)$ for the problem (18)-(19) with n subdomains and then we prove an upper bound for the error using $\eta(n)$. We are mainly interested in handling terms arising on the interface Γ and specifically, in the part of the error estimator related to the jump of the numerically computed flux across Γ.

If $n = 1$, the estimator $\eta(n)$ derived below reduces, of course, to the one discussed in Theorem 3.1. For $n > 1$, $\eta(n)$ includes the subdomain interior terms akin to the terms $\eta_{f,i}, \eta_{int,i}$ in (56) and which are calculated on every subdomain Ω_i, the

boundary terms defined on $\partial\Omega_i \cap \partial\Omega$, if nonempty, and the interface terms defined on the interface Γ. The latter combine the terms $\eta_{II,i}, \eta_{III,i}$ in (56). All the terms arise naturally from the definition of the residual.

To prove the upper bound for the velocity error in the (broken) norm on $\mathbf{V^{-1}}$

$$\sum_i \| \mathbf{e} \|^2_{K,i} = \sum_i \| \mathbf{e}^0 \|^2_{K,i} + \| \mathbf{e}^K \|^2_{K,i}, \tag{57}$$

where similarly as in (42) one can show

$$\sum_i \| \mathbf{e}^K \|^2_{K,i} \leq \sum_i C_{K,i} \| f - \Pi_{0,i} f \|_{i,0}, \tag{58}$$

we follow some of the ideas developed in [54] for the mortar formulations for standard Galerkin method. There, as a distinct feature, the various hierarchical and residual estimators considered include terms which estimate the error on subdomains as well as terms which pertain to the interface problems and provide a 'measure of nonconformity' arising from the use of non-matching grids and mortar spaces. To handle these, an important technical assumption called the *saturation assumption* is used; it is based on asymptotic a-priori estimates derived for mortar formulations. The quasi-interpolant used in the proofs of upper bounds is similar to one by Clément, and it is adapted to various choices of basis functions on the interface.

In this paper, the proof of the upper bound in our case is analogous to [54] but also very different since the underlying mixed spaces and mortar methods have an entirely different structure.

4.1. Definition of residual and calculations. First we define, for every $\mathbf{v} \in \mathbf{V^{-1}}$, that is, for $\mathbf{v}_i$ locally in $\mathbf{V}_i$, the residual,

$$r(\mathbf{v}) := \sum_i r_i(\mathbf{v}_i) \tag{59}$$

$$\begin{aligned} r_i(\mathbf{v}_i) &:= -(K^{-1}\mathbf{u}_h, \mathbf{v}_i)_i + (p_h, \nabla \cdot \mathbf{v}_i)_i \\ &- < g_h, \mathbf{v}_i \cdot \nu >_{\partial\Omega_i \setminus \Gamma_i} - < \lambda_h, \mathbf{v}_i \cdot \nu >_{\Gamma_i} . \end{aligned} \tag{60}$$

Notice that the residual $r_i(\mathbf{v}_i)$ has been defined analogously to (39).

On each subdomain we can decompose $\mathbf{v}_i \in \mathbf{V}_i$ in such a way that $\mathbf{v}_i = \mathbf{v}_i^0 + \mathbf{v}_i^K$ with the decomposition being local to Ω_i. Also, the corresponding $\phi_i = \phi_i(\mathbf{v}_i^0) \in H^1(\Omega_i)$ but because of lack of continuity of its normal flux across Γ, $\phi(\mathbf{v}^0) \notin H^1(\Omega)$. Next, $\mathbf{v}^K$ is K-weakly orthogonal to $\mathbf{v}^0$ only locally, that is, $(K^{-1}\mathbf{v}^K, \mathbf{q})_i = 0$, for any $\mathbf{q} \in \mathbf{H(div}; \Omega_i) \cap \mathbf{X^0}$. Finally, its regularization $\phi_{i,h} = P_{BG}\phi_i$ is local to Ω_i.

Now we proceed with residual calculations similar to those that gave (44). We use $\mathbf{v} = \mathbf{u} - \mathbf{u}_h$ and its restriction $\mathbf{v}_i$ on Ω_i, to get

$$\begin{aligned} \| \mathbf{v}^0 \|^2_{K,\Omega_i} &= \sum_{e \setminus \partial\Omega_i} \int_e [K^{-1}\mathbf{u}_h \cdot \mathbf{t}] \phi_i ds \\ &+ \sum_{e \cap (\partial\Omega_i \setminus \Gamma_i)} \int_e (K^{-1}\mathbf{u}_h \cdot \mathbf{t} \phi_i(\mathbf{v}_i^0) - g\mathbf{v}_i^0 \cdot \mathbf{n}) ds \\ &+ \sum_{e \cap \Gamma_i} \int_e (K^{-1}\mathbf{u}_h \cdot \mathbf{t} \phi_i(\mathbf{v}_i^0) - p\mathbf{v}_i^0 \cdot \mathbf{n}) ds \end{aligned} \tag{61}$$

where we have replaced g by p on the interface part Γ_i of $\partial\Omega_i$. Next, as it was done to get (47), we take into account the discrete problem (18), subtract the **Curl** of the

local regularization $\phi_{i,h} = P_{BG}\phi_i$, and calculate individual parts of the residual, for each i, as follows:

$$
\begin{aligned}
(62) \qquad \| \mathbf{e}^0 \|^2_{K,\Omega_i} &= - < p - \lambda_h, \mathbf{v}_i^0 \cdot \mathbf{n} >_{\Gamma_i} - < g - g_h, \mathbf{v}_i^0 \cdot \mathbf{n} >_{\partial\Omega_i \setminus \Gamma_i} \\
&+ r(\mathbf{v}_i^0 - \mathbf{v}_{i,h}^0) \\
&= \sum_{e \setminus \partial\Omega_i} \int_e [K^{-1}\mathbf{u}_h \cdot \mathbf{t}](\phi_i - \phi_{i,h}) ds \\
&+ \sum_{e \cap (\partial\Omega_i \setminus \Gamma_i)} \int_e (K^{-1}\mathbf{u}_h \cdot \mathbf{t}(\phi_i - \phi_{i,h}) \\
&- g_h(\mathbf{v}_i^0 - \mathbf{v}_{i,h}^0) \cdot \mathbf{n} - (g - g_h)\mathbf{v}_i^0 \cdot \mathbf{n}) ds \\
&+ \sum_{e \cap \Gamma_i} \int_e (K^{-1}\mathbf{u}_h \cdot \mathbf{t}(\phi_i - \phi_{i,h}) \\
&- \lambda_h(\mathbf{v}_i^0 - \mathbf{v}_{i,h}^0) \cdot \mathbf{n} - (p - \lambda_h)\mathbf{v}_i^0 \cdot \mathbf{n}) ds \\
&:= \sum_{e \in \partial\mathcal{T}_i, e \subset \Omega_i} \int_e I_{int,i} ds \\
&+ \sum_{e \in \partial\mathcal{T}_i, e \subset \partial\Omega_i \setminus \Gamma_i} \int_e (B_{i,I} + B_{i,II} + B_{i,III}) ds \\
&+ \sum_{e \in \partial\mathcal{T}_i, e \subset \Gamma_i} \int_e (D_{i,I} + D_{i,II} + D_{i,III}) ds
\end{aligned}
$$

where the definitions of the terms $B_{I,i}, \dots D_{III,i}$ follow naturally as in (47) except that B_* terms apply on the external parts of the boundary $\partial\Omega_i \cap \partial\Omega$ and D_* apply on its interface part Γ_i.

When adding (62) over all subdomains $i = 1, \dots n$, we find that $\sum_i \| \mathbf{e}^0 \|^2_{K,\Omega_i}$ is made of parts interior to subdomains $I_{int,i}$, the external boundary parts $B_{I,i} \dots B_{III,i}$ and of the interface parts $D_{I,i} \dots D_{III,i}$. The handling of the former is exactly the same as in Sections 3.3 and 3.4. We focus therefore on the latter which involves terms on the interface Γ.

In fact, we show below that the terms $D_{I,i}$ and $D_{II,i}$ are combined together for each i and treated as "boundary terms" in the sense of calculations done in Section 3.4 and that they act "along" Γ, see Remark 3.2. On the other hand, the terms $D_{III,i}$ when added over all i give rise to the part of the estimator dealing with the jump "across" each Γ_{ij} which is most interesting to us as it provides the natural measure of "inconsistency" or "defect" between discrete velocities defined on on both sides of Γ_{ij} which as we recall are only weakly continuous.

4.2. Interface terms. In what follows it will be convenient to consider an *intersection of grids*. Recall that each Ω_i is covered by quadrangulation $\mathcal{T}_i$ whose trace $\partial\mathcal{T}_i \cap \partial\Omega_i$ provides a grid on each Γ_i. Also, each interface Γ_{ij} has its own associated mortar grid $\mathcal{T}_m$. In the calculations below it will be convenient to use a refinement of these grids, or in other words, a partition $\mathcal{T}_{i\cap m}$ and $\mathcal{T}_{i\cap j}$ of Γ_{ij} so that

$$\mathcal{T}_{i\cap m} := \{\xi : \ \xi = e_i \cap e_m, \ e_i \subset \partial E_i \cap \Gamma_{ij}, \ E_i \in \mathcal{T}_i, \ e_m \in \mathcal{T}_m\}$$

and

$$\mathcal{T}_{i\cap j} := \{\chi : \ \chi = e_i \cap e_j, \ e_i \subset \partial E_i \cap \Gamma_{ij}, \ E_i \in \mathcal{T}_i, \ e_j \subset \partial E_j \cap \Gamma_{ij}, \ E_j \in \mathcal{T}_j, \}$$

Note that these new partitions are nothing but refinements of the original grids. By abuse of notation, we could write $\mathcal{T}_{i\cap j} = \partial\mathcal{T}_i \cap \partial\mathcal{T}_j$, and $\mathcal{T}_{i\cap m} = \partial\mathcal{T}_i \cap \mathcal{T}_m$.

4.2.1. Interface terms $D_{i,I} + D_{i,II}$ "along" Γ. Here we consider calculations for each subdomain i.

Recall that λ_h is piecewise linear on the grid $\mathcal{T}_m$ on Γ_{ij} and it is therefore also smooth on $\mathcal{T}_{i\cap m}$. Using the same calculations as in Section 3.4.2, we rewrite the sum over e as a sum over $\xi \in \mathcal{T}_{i\cap m}$ and consider the contributions from $D_{i,I} + D_{i,II}$ as

$$\sum_{\xi\cap\Gamma_i} \int_e (K^{-1}\mathbf{u}_h + \nabla\lambda_h \cdot \mathbf{t}(\phi_i - \phi_{i,h}) \tag{63}$$

where we understand by $\nabla\lambda_h \cdot \mathbf{t}$ the directional derivative alongside Γ of λ_h. In this calculation which follows from integration by parts, we are left, as in (53), with additional pointwise terms $\lambda_h(\phi - \phi_h)$ at those endpoints $P^i_{\partial\Omega} := \bar{\Gamma}_i \cap (\partial\bar{\Omega})$ of Γ_i which may lie on $\partial\Omega$. We ignore these pointwise values as well as those in the corner points where some Γ_{ij} and Γ_{kl} intersectas their handling is implementation dependent.

More precisely, when summing over the edges ξ we have $\sum_\xi \int_\xi -\lambda_h(\mathbf{v}^0_i - \mathbf{v}^0_{i,h}) \cdot \mathbf{n}ds = \sum_\xi \int_\xi \nabla\lambda_h \cdot \mathbf{t}(\phi - \phi_h)ds + \sum_{P_{\partial\Omega}} \lambda_h(\phi - \phi_h)|_{P_{\partial\Omega}}$. The pointwise values at $P^i_{\partial\Omega}$ must be combined with $B_{i,II}$. In practice, handling of these terms which enter the $\eta_{\partial\Omega}$ part of the estimator depends upon the type of approximation of boundary condition imposed on $\partial\Omega$ and its details will not be discussed.

To handle the terms arising from (63), we proceed as in Section 3.4.2 by noticing that on each ξ the norm $\| \phi - \phi_h \|_{0,\xi}$ can be estimated from above by $\| \phi - \phi_h \|_{0,e}$, for some e containing ξ. Therefore we get the same type of terms with the part of the estimator

$$\eta^2_{\Gamma,|,\xi} := h_E \| (K^{-1}\mathbf{u}_h + \nabla\lambda_h) \cdot \mathbf{t} \|^2_{0,\xi}$$

akin to the terms $\eta_{II,\Omega}$ in (56). Naturally

$$\eta^2_{\Gamma,|} := \sum_i \sum_{\xi\cap\Gamma_i} \eta^2_{\Gamma,|,\xi}$$

and we conclude by stating the result which can be proven using the same technique as the one applied in Section 3.4.2:

$$\sum_i \sum_{\xi\cap\Gamma_i} \int_e (D_{I,i} + D_{II,i})ds \le C\eta_{\Gamma,|} \| \mathbf{e} \|_K, \tag{64}$$

where the constant C depends on Ω and K but not on h.

Remark 4.1. *In implementation, the values of λ_h used as a boundary condition for the local discrete problem* (18),(19) *are replaced by its piecewise constant projection onto the grid of Ω_i expressed by $Q_{h,i}\lambda_h$. Then the calculation as above is no more valid as $\mathbb{Q}_{h,i}\lambda_h$ is piecewise constant. In such a case, we propose to consider, on a union of two edges $e_1 \cup e_2$, instead of the quantity $\nabla\lambda_h$, its discrete "gradient"* $\nabla_h\lambda_h\,|_{e_1\cup e_2} := \frac{Q_{h_i}\lambda_h|_{e_2} - Q_{h_i}\lambda_h|_{e_1}}{h_{e_1,e_2}}$. *It is not hard to see that, a similar estimate as above will still be valid. Also then, the norm over ξ in the definition of $\eta_{\Gamma,|}$ can be replaced by the norm over e_i which is much easier to compute in implementation.*

We note in passing that the use of the discrete gradient is related to the definition of an interpolation operator I_i which is discussed in Section 4.3.

4.2.2. Interface terms $D_{i,III} + D_{j,III}$ "across" Γ. First we note that for $\mathbf{v} = \mathbf{e} = \mathbf{u} - \mathbf{u}_h$ we have $\mathbf{v} \in \mathbf{V} + \mathbf{V^0}_h$ that is, the error is a linear combination of vector functions whose normal components are at least weakly continuous across Γ. Also, $\mathbf{u}^0 \in \mathbf{V} \cap \mathbf{X^0}$, and $\mathbf{u}^0_h \in \mathbf{V^0}_h \cap \mathbf{X^0}$. Hence,

$$\sum_i \sum_{e \cap \Gamma_i} \int_e D_{III,i} ds = \sum_i \sum_{e \cap \Gamma_i} \int_e -(p - \lambda_h)\mathbf{v}_i^0 \cdot \mathbf{n} ds$$
$$= -\sum_i \sum_{e \cap \Gamma_i} \int_e (\lambda_h - p)\mathbf{u}_h^0 \cdot \mathbf{n}_i ds.$$

Using the intersection of grids $\mathcal{T}_{i \cap j}$ we can write

$$\sum_i \sum_{e \cap \Gamma_i} \int_e D_{III,i} ds = -\sum_{i,j} \sum_{\chi \cap \Gamma_{ij}} \int_\chi (\lambda_h - p)[\mathbf{u}_h^0 \cdot \mathbf{n}]_{ij} ds. \tag{65}$$

The crux now is to find an appropriate estimate of $(\lambda_h - p)$ in some norm of the error $\mathbf{e}$ and to find appropriate characterization of the jump quantity $[\mathbf{u}_h^0 \cdot \mathbf{n}]_{ij}$ which is constant on each χ. The former will be done by borrowing from the theory developed in [27, 24], and by applying some a-priori estimates from [7]. The latter will be done by defining a norm of $[\mathbf{u}_h^0 \cdot \mathbf{n}]_{ij}$ weighted by averages of the length of the edges on both "sides"of χ which will give rise to the "across-Γ" part $\eta_{\Gamma,*}$ of the estimator η.

Remark 4.2. *Note that if the source f restricted elemntwise to the support of the discrete solution $\mathbf{u}_h$ is zero, which is frequently the case of wells in a reservoir, then $\mathbf{u}_h^0$ coincides with $\mathbf{u}_h$.*

4.3. Estimates on Γ. In order to estimate the terms in (65), we first note that it can be rewritten as

$$\sum_i \sum_{e \cap \Gamma_i} \int_e D_{III,i} ds = -\sum_i \int_{\Gamma_i} (\lambda_h - p)\mathbf{u}_h^0 \cdot \mathbf{n}_i ds.$$

which resembles the definition of a bilinear form $(\cdot,\cdot)_{d_h}$ in [7]. That form provides some "measure of inconsistency" due to the non-matching grids and mortar spaces: it evaluates the jump of the fluxes arising from a set of solutions in subdomains subject to homogeneous boundary conditions on $\partial\Omega$ and no source/sink terms, and subject to a prescribed boundary condition on Γ given by $\lambda_h - p$. The a-priori results in [7] provide a convergence estimate for $\| \lambda_h - p \|_{d_h}$ in the norm arising from $(\cdot,\cdot)_{d_h}$.

We cannot use this characterization directly, mainly because it does not provide a natural way to define a computable part of the estimator. Rather, we define a norm which can be shown to be equivalent to $\| \cdot \|_{d_h}$. To this aim, we recall an interpolation operator I_i which is a composition of the interpolation operator $I_h^{\Omega_I}$ from [24] and the projection $Q_{h,i}$ from [7], see ([7], Remark 6.1). More precisely, for a function $r \in L^2(\Gamma)$, $I_i r$ is the piecewise linear interpolant of the piecewise constant averages of r defined on the grid $\mathcal{T}_i|_{\Gamma_i}$.

We now define the following discrete seminorm on Γ which in fact is a norm on Λ_h for piecewise constants on the grids $\partial\mathcal{T}_i$ on Γ

$$|r|_M^2 := \sum_i |r|_{M,\Gamma_i}^2 := \sum_i \sum_{e \in \partial\mathcal{T}_i|_{\Gamma_i}} \frac{1}{h_e} \| I_i r \|_{0,e}^2$$

where $h_e = |e|$ is the length of an edge $e \in \mathcal{T}_i|_{\Gamma_i}$.

Remark 4.3. *It is shown in* [24] *that*

$$|I_i r|^2_{1/2,\Gamma_i} \approx \sum_{e\in\partial\mathcal{T}_i|_{\Gamma_i}} \frac{1}{h_e} \parallel I_i r \parallel^2_{0,e}$$

where the $\approx$ *sign means equivalence of norms. In this result the scaling by a power of* $|e| \approx h_e$ *is typical to the dimension of* Γ *and* Ω *in* $\mathbb{R}^d, d = 2$*, and is reminiscent of the fact that the discrete* l_2 *norm of nodal values of* $I_i r$ *scales like* $\frac{1}{\sqrt{|e|}}$ *with respect to the* $L^2(\Gamma)$ *norm of* r*. We can conclude, from the definition of* M*-seminorm that it is equivalent to this discrete* $H^{1/2}(\Gamma_i)$ *seminorm.*

Remark 4.4. *In* [7]*, Remark 6.1, the* $|I_i \cdot|_{1/2,\Gamma_i}$ *(and by Remark 4.3, the* M*-seminorm defined above) is discussed as being equivalent to the* $|\cdot|_{d_h}$ *seminorm. This equivalence is reconfirmed by numerical experiments shown in (*[7]*, Section 8.1).*

In conclusion from these two Remarks we may write $|r|^2_M \approx |r|_{d_h}$. To see how this can be exploited we proceed to estimate (65).

Denote by $\hat{r}_\omega$ the average value of r over ω and use the fact that, upon neglecting higher order terms, $\parallel \hat{r}_e \parallel_{0,e} \approx \parallel I_i r|_e \parallel_{0,e}$, for an edge $e \in \partial\mathcal{T}_i$. We can then derive the following estimate, by grouping the terms $\chi \subset e_i \cap e_j$, for some $e_i \in \partial\mathcal{T}_i$ and $e_j \in \partial\mathcal{T}_j$, and applying Cauchy-Schwarz inequality for sums repeatedly,

$$\begin{aligned} < r, s_h >_{\Gamma_{ij}} &= \sum_\chi \int_\chi r s_h ds \le \sum_\chi \parallel s_h \parallel_{0,\chi} \parallel \hat{r}_\chi \parallel_{0,\chi} \\ &\le \frac{1}{2}\left(\sum_{e_i\in\partial\mathcal{T}_i, e_i\subset\Gamma_{ij}} \frac{1}{\sqrt{h_{e_i}}} \parallel I_i r \parallel_{0,e_i} \sqrt{h_{e_i}} \parallel s_h \parallel_{0,e_i} \right. \\ &+ \left. \sum_{e_j\in\partial\mathcal{T}_j, e_j\subset\Gamma_{ij}} \frac{1}{\sqrt{h_{e_j}}} \parallel I_j r \parallel_{0,e} \sqrt{h_{e_j}} \parallel s_h \parallel_{0,e_j} \right) \end{aligned}$$

Next, upon summing over all i, j and using the definition of M-norm, and splitting the norms (squared) of s_h over e_i, e_j as sums of norms (squared) over χ, we get

$$\begin{aligned} < r, s_h >_\Gamma &\le \frac{1}{2} \parallel r \parallel_M \times \\ &\sqrt{\sum_{i,j} \left(\sum_{e_i\in\partial\mathcal{T}_i, e_i\subset\Gamma_{ij}} h_{e_i} \parallel s_h \parallel^2_{0,e_i} + \sum_{e_j\in\partial\mathcal{T}_j, e_j\subset\Gamma_{ij}} h_{e_j} \parallel s_h \parallel^2_{0,e_j} \right)} \\ &\le \frac{\sqrt{2}}{2} \parallel r \parallel_M \sqrt{\sum_{i,j} \left(\sum_{\chi\in\mathcal{T}_{i\cap j}, \chi\subset\Gamma_{ij}} \frac{h_{e_i} + h_{e_j}}{2} \parallel s_h \parallel^2_{0,\chi} \right)} \end{aligned}$$

These calculations suggest a natural definition of the "across-Γ"part of the error estimator as

$$\eta^2_{\Gamma,*} := \sum_{i,j} \left(\sum_{\chi\in\mathcal{T}_{i\cap j}, \chi\subset\Gamma_{ij}} \frac{h_{e_i} + h_{e_j}}{2} \parallel [\mathbf{u}^0_h \cdot \mathbf{n}]_{ij} \parallel^2_{0,\chi} \right). \tag{66}$$

We note in passing here that the expression used in the definition of $\eta_{\Gamma,*}$ can be seen, by its scaling, as the discrete dual $H^{-1/2}$ norm, see [10], and henceforth, as the norm M^{-1} "dual" to M. This characterization will not be pursued.

Combining the above calculations and applying them to (65), and following the Remarks 4.3,4.4 we arrive at

$$\sum_i \sum_{e\cap\Gamma_i} \int_e D_{III,i} ds \leq C \parallel p - \lambda_h \parallel_M \eta_{\Gamma,*} \tag{67}$$

where the constant C is independent of h and where it lumps the constants arising in the calculations as well as those throgh equivalence of norms.

Finally, we need an estimate for $\parallel p-\lambda_h \parallel_M$ in some norm of $\mathbf{e}$. This is provided by the fundamental a-priori result in ([7], Theorem 4.2) which states that the error $\sum_i \parallel \mathbf{u} - \mathbf{u}_h \parallel_{\mathbf{H(div};\Omega_i)} = O(h)$. At the same time, Theorem 6.2, [7], says, that for diagonal tensor K, and rectangular grids, we have $\parallel p-\lambda_h \parallel_{d_h} = O(h^{3/2})$. Recalling the equivalence of $\|\!| \cdot |\!\|_{K,i}$ and $\| \cdot \|_{\mathbf{H(div};\Omega_i)}$ which hold on every Ω_i, we combine these results in the following Assumption.

Assumption 6. *For sufficiently small h, the following estimate holds*

$$\parallel p - \lambda_h \parallel_M \leq C_{sat} \|\!| \mathbf{u} - \mathbf{u}_h |\!\|_K \, . \tag{68}$$

This assumption is similar in its application to the *saturation assumption* in [54]. With (68) we can complete the estimate of the "across Γ" terms (67) and get

$$\sum_i \sum_{e\cap\Gamma_i} \int_e D_{III,i} ds \leq C \eta_{\Gamma,*} \|\!| \mathbf{e} |\!\|_K \tag{69}$$

where the constant C lumps constants from (68) and from calculations leading to (67) and is independent of h but is dependent on K and Ω.

4.4. Main result. By combining (57), (58), (62), (64), (69), and proceeding similarly as in Section 3, we arrive at the following theorem which shows that the estimator η is *reliable*.

Theorem 4.1. *Let Assumptions 1, 2, 3, 4 and 5, 6 be satisfied. Then the following upper bound holds:*

$$\begin{aligned} \sum_i \|\!| \mathbf{e} |\!\|^2_{K,i} \quad &\leq \quad \eta^2(n) := \sum_i \eta^2_{f,i} + \sum_i \eta^2_{int,i} \\ &+ \quad \sum_i \eta^2_{II,\partial\Omega_i\cap\partial\Omega} + \sum_i \eta^2_{III,\partial\Omega_i\cap\partial\Omega} \\ &+ \quad \sum_i \eta^2_{\Gamma,|,i} + \eta^2_{\Gamma,*}. \end{aligned} \tag{70}$$

We de not pursue the proof that the estimator is *efficient*, or consideration of the total error, see discussion following Theorem 3.1. Instead we devote the rest of the paper to examples and extensions concerning the terms $\eta_{\Gamma,*}$ which are most interesting as they measure the "defect" arising from mortar spaces.

5. Computational results for single phase flow

Results reported in this paper were obtained with the implementation of the mortar algorithm designed and integrated within the general multiblock and multiphysics capabilities of the IPARS framework [51]. The grids in the framework are rectangular cell-centered; identification of finite difference schemes on these grids through numerical integration with mixed methods on $\mathbf{RT}_{[0]}$ spaces is done as in [47, 8] and Peaceman well models [39] are employed. Boundary conditions and

solvers applied are described in [42, 32, 33], respectively. Various physical models are available in the framework; in this paper we focus on a single-phase and a two-phase implicit and sequential models.

For single-phase flow, the IPARS implementation of the mortar algorithm, recalled in Section 5.1, is essentially a straightforward port of the code used in [7]. Asymptoptic convergence of this algorithm was demonstrated in [7] and will not be discussed here. The algorithm was enhanced for the needs of this paper by computation of error indicators on the mortar interface which helps to find an optimal mortar grid $\mathcal{T}_m$ for a given subdomain grid $\mathcal{T}_h$.

The extensions of the algorithm to multi-phase flow are discussed, and related computational results are reported in Section 7. Details of the algorithm and its implementation for multi-phase flow and multinumerics and multiphysics couplings are described in detail in [43, 44, 36, 45, 50, 35].

Remark 5.1. *For the sake of proper interpretation of computational results discussed below, it is important to point out the following. IPARS framework, among other features, offers visualization and post-processing capabilities of cell-centered unknowns designed to be flexible and robust for execution with multiple blocks and on multiple processors. While it provides continuous contours across processor boundaries, the visualization tool used for presentation of results in this paper works locally on any given block and does not exploit any meta-information concerning the global continuity of the solution across the interface* Γ. *While this could be done by some form of interpolation or kriging, also on non-matchning grids, it would result in ad-hoc smearing of* real *values. The "lack-of-continuity" phenomenon gets weaker with refinement of subdomain grids* $\mathcal{T}_h$, *but we advocate some degree of caution when interpreting the visual results; see Figure 1.*

Finally, visualization of velocity/flux type unknowns gets even trickier: the visualization tool requires projecting any values to a node-centered visualization grid. Since velocities in $\mathbf{RT}_{[0]}$ *spaces are naturally defined at mid-points of the edges, the post-processed fluxes are at best first order accurate in the interior of subdomains and appear unnatural at the interface* Γ. *For that reason we do not show the velocity field in our computational examples.*

5.1. Algorithm for single-phase flow. Now we discuss the computational realization of the system (18)-(19) following [27, 7]. It is done in a classical domain decomposition setting where the unknowns are the Lagrange multipliers on the interface sought by an iterative algorithm.

Specifically, consider a given $\lambda_h \in \Lambda_h$. For any such λ_h, the solution $\mathbf{u}_{h,i},\ p_h|_{\Omega_i}$ to the local problems (12)-(13) exists and is unique. One can therefore associate with λ_h the pointwise values of the jump of normal components of $\mathbf{u}_h \in \mathbf{V^{-1}}_h$ across Γ which we denote by $\boldsymbol{B}(\lambda_h)$. We have

$$(\boldsymbol{B}(\lambda_h), \mu)_\Gamma = \sum_i < \mathbf{u}_{h,i} \cdot \mathbf{n}_i, \mu >_{\Gamma_i}, \quad \mu \in \Lambda. \tag{71}$$

Now take $\mu \in \Lambda_h$; with such test functions $\boldsymbol{B}(\lambda_h)$ can be identified with its projection $\boldsymbol{B}_h(\lambda_h)|_{\Gamma_{ij}} := \Pi_h[\mathbf{u}_h]_{ij}$ onto Λ_h.

Using this characterization we see that $\mathbf{u}_h \in \mathbf{V^{-1}}_h$ satisfies (18)-(19) if actually $\mathbf{u}_h \in \mathbf{V^0}_h$, or, equivalently, if $\forall \mu \in \Lambda_h,\ (\boldsymbol{B}(\lambda_h), \mu)_\Gamma = (\boldsymbol{B}_h(\lambda_h), \mu)_\Gamma = 0$. Hence, the following problem is equivalent to solving (18)-(19)

$$\text{find } \lambda_h : \quad \boldsymbol{B}_h(\lambda_h) = \mathbf{0}, \tag{72}$$

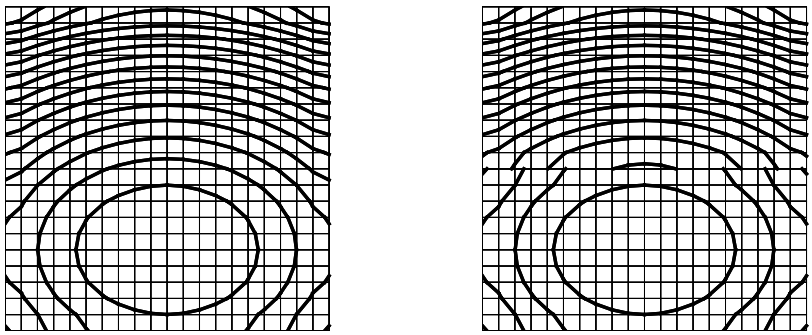


FIGURE 1. Visualization issues for multiblock mortar decomposition. Left: contours of a continuous variable. Right: contours of the same variable displayed on a multiblock grid with *matching* grids lack continuity as they are rendered independently on each block. This "artificial" discontinuity decreases with the grid size and is purely a visualization artifact.

To solve (72) one employs an iterative algorithm, which starts with an initial guess λ_h^0 and which terminates when a given convergence criterium

$$\| \boldsymbol{B}_h(\lambda_h) \|_{0,\Gamma} \leq \varepsilon \tag{73}$$

is satisfied.

In view of our subsequent discussion, it is convenient to represent the solution algorithm as follows

Iterate $I = 0, 1, \ldots I_{conv}$: given current guess λ_h^I
- a) solve in each block $\Omega_i,\ i = 1 \ldots n$:
  - use $\lambda_h^I|_{\Gamma_i}$ as a boundary condtition for the pressure in each block i,
  - find $(p_h, \mathbf{u}_h^{(I)})|_{\Omega_i}$
- b) compute $\boldsymbol{B}_h(\lambda_h^I)$, verify (73), if satisfied, quit;
- c) if (73) not satisfied, determine a new guess λ_h^{I+1}; $I := I + 1$, go to a).

FIGURE 2. Multiblock algorithm for single-phase flow

In implementation, one can exploit linearity of the problem, and solve only a positive-definite sub-problem of (73) using a CG solver; see [27, 7]. In a general setting applicable to multi-phase flow problems, a non-symmetric solver from GMRES family is applied to solve the algorithm in Figure 2. The choice and properties of that solver critically influence the computational complexity of the overall mortar algorithm which in general increases with the number n_m of mortar degrees of freedom. Application of a preconditioner may weaken this dependence; in particular, an optimal preconditioner may render I_{conv} independent of n_m. In this direction, see the discussion on balancing preconditioner in [24, 40], or the multigrid algorithms described in [52].

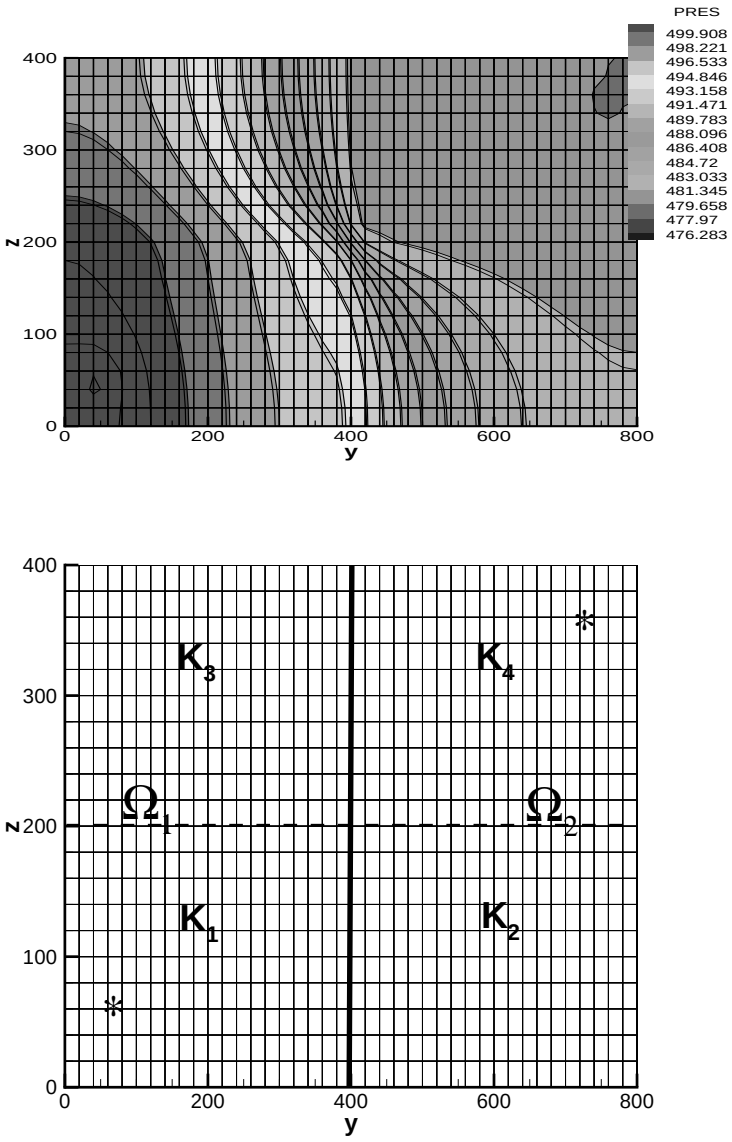


FIGURE 3. Single-phase flow example. Left: pressure solution on one block (without mortars). Right: permeability field characterized by constants K_1, K_2, K_3, K_4, position of wells marked with a "*", and decomposition of the domain into two blocks Ω_1, Ω_2. Note different aspect ratio on the right.

Without optimal preconditioners I_{conv} increases at least linearly with n_m which remains true in the more general setting for multi-phase flow. Hence, it is of substantial interest to find a "good" mortar grid for which the balance of accuracy and complexity is acceptable. This is discussed below.

Remark 5.2. *We note that, for the sake of this paper, we conciously ignore the following elements of the implementation which render the* computed *solution* $\tilde{\mathbf{u}}_h$

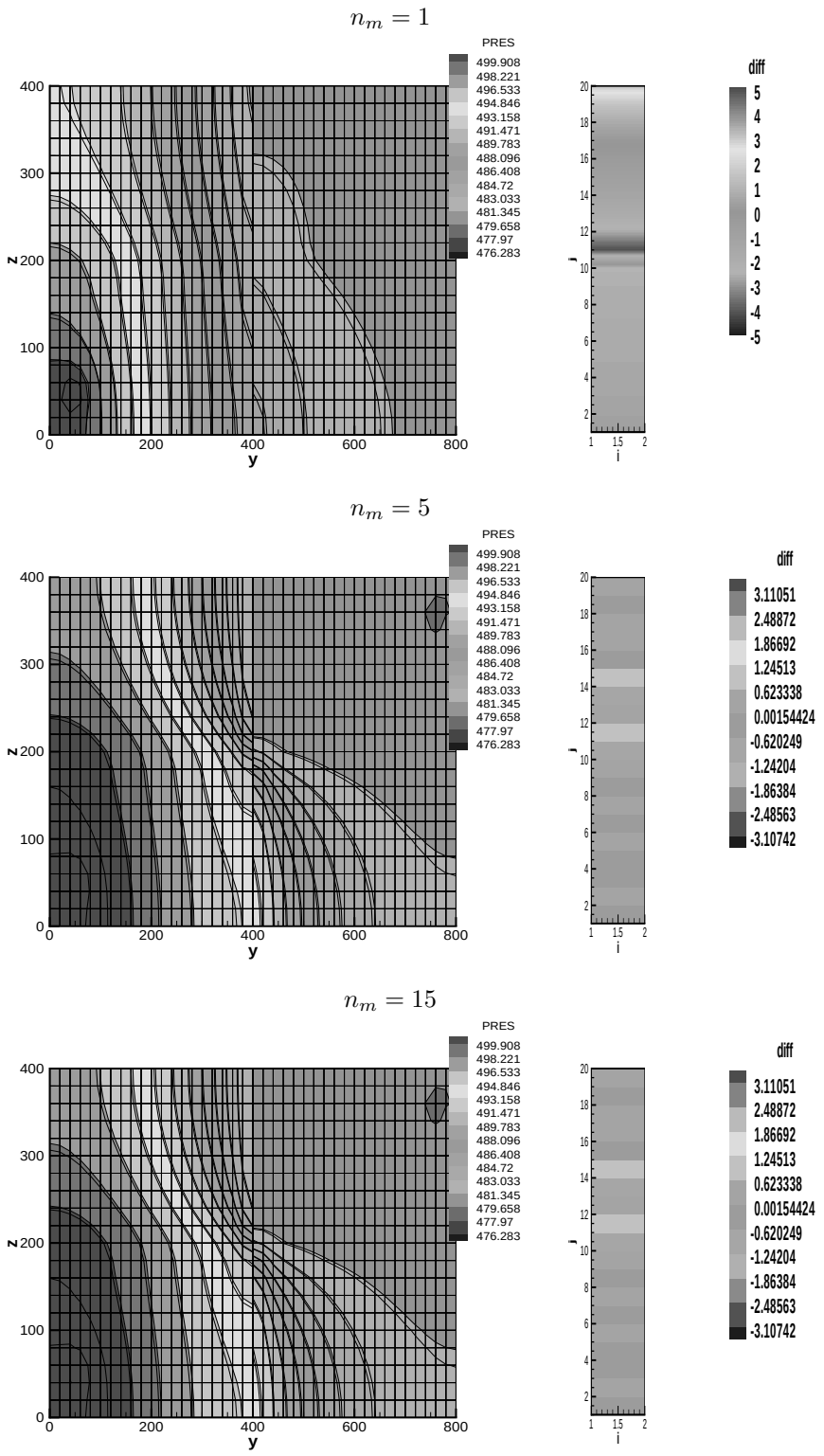


FIGURE 4. Single-phase flow example, mortar solution. Left: pressure solution. Right: contours of the error quantity $\eta_{\Gamma,*}$ computed pointwise

different from $\mathbf{u}_h$ *considered above. These elements include effects of Peaceman well models, numerical integration and in particular those of identification of mixed* $\mathbf{RT}_{[0]}$ *solution with one of cellcentered discretization, iterative (inexact) solution of subdomain problems, and those of inexact solution to the interface problem.*

5.2. Computational experiment. Now we focus on the main interest of this paper which is to determine the "right" mortar grid $\mathcal{T}_m$ for a given subdomain discretization $\mathcal{T}_h$. To this aim, consider a simple computational experiment where the flow in the region $\Omega = (0, 8) \times (0, 800) \times (0, 400)$ is driven by an injection and a production well, located as shown in Figure 3, each with bottom hole pressure (BHP) specified at $510[psi], 475[psi]$ respectively. The permeability is given by the constants $K_1 = 0.5, K_2 = 5, K_3 = 2, K_4 = 200$, with isotropy in both horizontal directions. The case is really two-dimensional but is simulated with a 3D code; the additional parameter is the anisotropy ratio of K equal to $1/50$ in the vertical direction. The grid is $2 \times (2 \times 20) \times 20$.

We note that instead of (7) we use here the no-flow boundary conditions (8). In addition, k and f are not smooth, hence, Assumption 3 is not satisfied. In fact, the pressure p it is at most $H^{3/2+\varepsilon}$, see ([15], Prop. 2.2). Still, the estimator η in Theorem 4.1 is well defined even though formal convergence proof does not hold.

The solution is clearly smooth locally inside each of the subdomains where $K_k, k = 1, \ldots 4$ is locally constant and away from the wells. However, the pressure has very high gradients near the center of Ω where there is a sharp jump of K in all directions. If the grid on Ω were to be adapted, the refinement would focus in this region. Assume however that we are keeping $\mathcal{T}_h$ fixed and that we vary $\mathcal{T}_m$ in the mortar multiblock decomposition shown in Figure 3.

Specifically, Ω is decomposed into two blocks Ω_1, Ω_2, with $\Gamma = (0, 8) \times \{400\} \times (0, 400)$. On Γ the subdomain grids simply match. For mortar grid we use uniform partition with $n_m = 1, 5, 15$ referred to as the coarse, fine, and finest mortar grids, respectively. The problem is solved in each case with very low tolerance $\varepsilon = 1.e-6$ on the interface. The pressure solution for each case is shown in Figure 4. We do not show the velocity field; by Remark 3.1 and 5.1 the images of scalar unknowns whose error is coupled to the velocity error is a better illustration for discussed phenomena.

Clearly, as can be seen from the pressure profiles compared to the one-block solution in Figure 3, the coarse mortar grid gives rise to a solution whose features are not adequately captured close to Γ, even if we factor in the distorting visualization effects discussed in Figure 1. However, the fine and the finest mortar grid appear to deliver an appropriate level of agreement with the one-block solution. Note that the grids across Γ are purposely chosen to match; this allows us to see clearly the effects of a weak-continuity condition of fluxes from $\mathbf{V^{-1}}$ versus those in $\mathbf{V^0}$, as in Figure 3, and its dependence on n_m. While obviously the quality of solution improves with increasing n_m, the weak-continuity never "converges" to the continuity condition: on one hand, there is an upper bound on n_m for a given $\mathcal{T}_h$ following from condition (23), and on the other hand, even for matching grids $\mathcal{T}_i, \mathcal{T}_j$, the composition of the projections $\Pi_h \circ Q_{h,i}$ or its transpose are never idempotent.

In general, a comparison of a numerical solution with an analytical solution is not available and one uses a-posteriori error estimators to guide the selection of an optimal grid. In the context of the mortar grid, a comparison of the mortar multiblock solution with one-block solution is likely also unavailable, due to the complexity of the size of the overall problem. Even if it is available, a rigorous

comparison may be hard to make due to different grids and different models applied to individual subdomains.

5.2.1. The effect of $\eta_{\Gamma,*}$. Consider the choice of an optimal mortar grid, while keeping $\mathcal{T}_h$ fixed. The error bound in Theorem 4.1 includes, in addition to $\eta_{\Gamma,*}$, the subdomain and the external boundary parts of the estimator, and therefore, no rigorous quantitative comparison can be made between the energy error and the error quantity η_{Γ_*}. We focus therefore on qualitative information: the error quantity on the interface delivers valuable information which can be used to determine the "best" mortar grid and, at the very least, the most suitable n_m.

Figure 4, right column, shows the local values of $\eta_{\Gamma,*}$ across Γ which are actually computed on the interesection of the grids $\mathcal{T}_{1\cap 2}$ which in this case is just the trace of the matching grids $\mathcal{T}_i$. One can see that there is a substantial support of the error component for coarse mortar grid which becomes localised in case of fine and finest mortar grids. These individual values can be used to determine the optimal grid. As can be seen in Figure 4, the values of $\eta_{\Gamma,*}$ correlate with the quality of the solution.

In summary, these results show the essential practical features of the mortar mixed algorithm as concerns dependence on n_m. First, the weak continuity of fluxes in (72) is not sufficient for a good level of local mass conservation when very coarse mortar grid is used. However it is not mandatory to use the finest mortar grid; an intermediate discretization may deliver satisfactory results. In fact, one can find the "right" mortar grid $\mathcal{T}_m$ from qualitative analysis of the error quantity $\eta_{\Gamma,*}$ shown per mortar grid element. In addition, for $n > 2$ the l_2 discrete norm over each interface Γ_{ij}, could be used to determine on which interface Γ_{ij} the mortar grid should be adapted. This is done in the context of multi-phase flow in Section 6.

6. Multi-phase flow model and mixed mortar algorithm

In this section we present an extension of the results above to multi-phase flow. First we introduce the model and the modified algorithm. The extension is obviously nontrivial as it includes transient nonlinear effects of viscous, capillary, and compressible flow of multiple components, and requires solution of a system of coupled parabolic/hyperbolic equations rather than a single elliptic problem as (1).

6.1. Model for multi-phase flow. Here we briefly introduce the model for the flow of multiple fluid phases which extendx (3) and (6). First, each of the phases α has its associated value of saturation S_α. One has

$$\sum_\alpha S_\alpha \equiv 1. \tag{74}$$

The pressures of the individual phases are denoted by p_α. Their difference is called the capillary pressure $P_{c,\alpha,\alpha'}$, it is dependent on phase saturations and is given as *rock property* data. Each phase has its associated density ρ_α. Relative permeabilities k_α and viscosity μ_α are, in general, dependent on the saturations and pressure. These are given as rock property data as well; for examples and details see [29].

The multi-phase extension of Darcy's law (3) reads

$$\mathbf{u}_\alpha \quad = \quad -K\frac{k_\alpha}{\mu_\alpha}\nabla p_\alpha. \tag{75}$$

while the conservation of mass (6) for immiscible flow is replaced by the equation, written for each α separately,

$$\frac{\partial(\phi\rho_\alpha S_\alpha)}{\partial t} + \nabla\cdot(\rho_\alpha \mathbf{u}_\alpha) \quad = \quad f_\alpha \tag{76}$$

where ϕ denotes the porosity of the porous medium, generally pressure-dependent and here assumed as dependent on position $x \in \Omega$ only.

This general multi-phase flow model is given in [38, 34] and more recently in [29]; it gives rise to a family of multi-phase and multi-component flow models ranging from models of unsaturated groundwater flow through two-phase oil-water flow models and black-oil models to a full compositional model. In the latter cases, the models are complemented by thermodynamic statements which relate the pressures and compositions of the fluids. In this paper we focus on two-phase flow only applicable to oil-water flow with $\alpha = w, o$ or flow in vadose zone with $\alpha = w, a$ where w, o, a denote phase subscripts for water, oil, and air, respectively. For simplicity we consider incompressible examples for oil-water reservoirs only.

Equations (75), (76) are written for each phase $\alpha = o, w$ and are coupled by (74) as well as by the capillary pressure relationship. In the case of two immiscible fluids $\alpha = w, o$, this relationship is given by

$$p_o - p_w := P_{cow}(S_w) \tag{77}$$

where P_{cow} is a non-increasing function of the wetting phase saturation S_w, which is singular at the left end of its domain and possibly as well at the right end. The model is complemented by initial conditions and boundary conditions. The former are, for example, hydrostatic equilibrium conditions and the latter are no-flow conditions which extend (8).

For this general multi-phase flow model case, the well-posedness results for single-phase flow do not extend. Results that do apply rely on characterization of the system (75)–(76) as a coupled system of two nonlinear generally parabolic equations which may locally change type to elliptic or hyperbolic. When the system is solved for, say, a pressure unknown and a saturation unknown, these variables generally exhibit non-smooth behavior in space and in time. In particular, the low regularity of pressure is mainly due to the jumps in the permeability coefficients and to high gradients around wells. At the same time, the saturation variable behaves like a solution to a nonlinear weakly diffusive conservation law and hence it exhibits fronts of shock and rarefaction type. See [5, 9, 6, 19, 20, 21] for discussion of well-posedness of two-phase flow problems and references to other works.

6.2. Numerical methods for multi-phase flow. As a consequence of low regularity of solutions to multi-phase flow models, the construction and analysis of appropriate numerical methods is difficult. The existing algorithms differ by the spatial and temporal discretization techniques. The former includes various finite element and finite volume discretizations as well as block- or cell-centered finite difference techniques. The latter includes various implicit, semi-implicit and explicit time discretizations as well as their time-adaptive variants. We refer the reader to [38, 29] for an introduction and a review of established methods and to [1] and references therein for recent advances.

The common characteristics of numerical methods for multi-phase flow relevant to the topic of this paper is the tremendous complexity of the overall system, especially in heterogeneous medium with many phases or fluid components present and when globally implicit methods are used. Efficient solvers are mandatory; see [33] for recent progress in this direction. On the other hand, the locally adaptive

gridding, time-stepping, model selection and solver techniques offer an attractive alternative, by allowing one to choose locally the best discretization method for the desired level of accuracy and efficiency.

This paper is devoted to one such technique based on mortar formulation, whose multi-phase and multi-physics algorithmic extension has been shown to be effective in this direction, see [43, 51, 44, 45, 37, 35].

6.3. Mortar algorithm for multi-phase flow. When extending the algorithm shown in Figure 2 to multi-phase flow, first we have to account for the transient behavior of the flow. For simplicity we consider a generic fully implicit form of time-stepping, and associate t with the time step value. In other words, we first consider a time- and space-discrete version of (75) and (76) using $\mathbf{RT}_{[0]}$ spaces, with upwinding applied to nonlinear terms on Ω. In fact, we consider its local variant on each Ω_i, similarly as it was done for single-phase flow. The local solutions are coupled by interface conditions posed in terms of unknowns $\mathbf{L}_h^t$ on Γ. Each local problem is solved for the set of unknowns $\Upsilon_h^t|_{\Omega_i}$ which include pressures, fluxes, and saturations. The problem in every block is nonlinear and is solved by a Newton method, to a certain tolerance. See [45] for details of the discretization and [42] for details of the model and its boundary conditions, and [32, 33] for discussion of subdomain solvers. Finally, see [43] for alternative to this fully implicit time-stepping procedure on the interface and in the subdomains. Some new results on non-implicit time-stepping on the interface are available but will not be discussed here.

Now we discuss $\mathbf{L}_h^t \in (\Lambda_h)^2$ (or more general $\mathbf{L}^t$ pointwise) which has to provide boundary conditions for the flow. If pressure is one of the two unknowns in $\mathbf{L}^t$, it has to be complemented by another phase variable. The choice of variables included in $\mathbf{L}^t$ is delicate and can be made adaptive in time, see a discussion in [41, 42]. Here we assume a fixed choice at every time step t, $\mathbf{L}^t \equiv \mathbf{L} = (P_o, S_o)$ pointwise which is consistent in this case with the subdomain unknowns Υ^t.

As concerns fluxes, we now consider $\boldsymbol{B}^t(\mathbf{L}_h^t) = ([\mathbf{u}_{h,w}^t], [\mathbf{u}_{h,o}^t])$ or any combination of these quantities, as may be convenient in various time-stepping procedures on the interface. In reality, for compressible flow, not the phase velocities $\mathbf{u}_\alpha$ but rather the mass fluxes $\rho_\alpha \mathbf{u}_\alpha$ are used in $\boldsymbol{B}$. We denote by $\boldsymbol{B}_h^t$ the discrete value of $\boldsymbol{B}^t(\mathbf{L}_h^t)$ identified through a projection onto $(\Lambda_h)^2$, as was done in Section 5.1.

In summary, at every time step t we solve the system

$$\boldsymbol{B}_h^t(\mathbf{L}_h^t) = \mathbf{0}$$

which is of size 2 times the number of mortar unknowns. This problem is, in general, nonlinear, and is solved by a Newton-Krylov algorithm to a given tolerance that is, we solve

$$\| \boldsymbol{B}_h^t(\mathbf{L}_h^t) \|_{0,\Gamma} \leq \varepsilon, \tag{78}$$

see [57] and extensions to multiphysics couplings in [35]. The algorithm naturally exploits the transient behavior of the problem and uses solutions from the previous time step on interface $\mathbf{L}_h^{t-1}$ as well as on subdomains Υ_h^{t-1} as initial guess for the nonlinear ieration in step t.

The algorithm is schematically shown in Figurre 5.

Time step t: set $\mathbf{L}_h^{t,1} := \mathbf{L}_h^{t-1}$ and use $\Upsilon_h^{t-1}|_{\Omega_b}$ as initial condition
Iterate $I = 1, \dots I_{conv}$: given current guess $\mathbf{L}_h^{t,I}$
a) solve in all blocks $i = 1, \dots n$:
i)... ii) ... use $\mathbf{L}_h^{t,I}$ as boundary condtitions for block i
iii)... iv) ... find $\Upsilon_h^{t,I}|_{\Omega_b}$
b) compute $\boldsymbol{B}_h^t(\mathbf{L}_h^{t,I})$, verify convergence and determine new guess $\mathbf{L}_h^{t,I}$
Time step t: we have computed $\mathbf{L}_h^t := \mathbf{L}_h^{t,I_{conv}}$ and $\Upsilon_h^t|_{\Omega_i} := \Upsilon_h^{t,I_{conv}}|_{\Omega_i}, i = 1, \dots n$

FIGURE 5. Multiblock algorithm for multi-phase flow

7. Computational results for multi-phase flow

Our focus in this section is as before on the jump quantity $\boldsymbol{B}$ across Γ and how it is related to the error in the computed solution. In the context of multi-phase flow, the very definition of the error quantity may vary, depending on the properties of the exact solution and on the goals of the computation. Whereas for single phase flow the error naturally included the phase pressure and velocities, in the case of multi-phase flow it clearly must include another phase property in addition to pressure and perhaps all other phase or component fluxes. While the behavior of phase pressures in case of constant flow conditions is not very dynamic, the behavior of phase saturations, in contrast, exhibits strong transient character. In waterflooding experiments starting from equilibrium conditions, such as those considered here, the saturation fronts start at injection wells and gradually advance to the production wells. Finally, the injection and production rates, which are dependent on the imposed well bottom hole pressure as well as on the current pressure and saturation values, are the most interesting; they can be considered as the natural *quantity of interest* in the language of error estimators [4].

The case studies whose results are reported below were designed to show the sensitivity of the multiblock mortar solution and its correlation to the error quantity $\eta_{\Gamma,*}$ to the mortar grid in a realistic reservoir simulation setting.

7.1. Setup of experiments. First we discuss the choice of the permeability field. In Section 5.1 we reported results for single phase flow in a reservoir with a piecewise constant permeability field, where the mortar interface Γ was aligned with some of the lines of discontinuity of the coefficient K as is classically done in domain decomposition; see also examples in [45]. Here we choose a more realistic setting in which K is heterogeneous and possibly uncertain; in fact, we generate K using a simple geostatistic package [31]. We consider three cases labeled as H, HI, HX, respectively; each represents a single realization of a permeability field. The cases differ by the correlation length of the assumed probability distribution and range from large to small to "zero" (uncorrelated), respectively. The fields are shown in Figure 6 and are based on the same $1 \times 30 \times 30$ grid. They were scaled to have approximately the same mean.

Now we discuss the flow conditions of all experiments. The flow is driven by an injection well and by a production well; consequently, the waterflooding proceeds from an injection well to the production well. The simulation lasts until 1000 days which is close to the water breakthrough time for all cases.

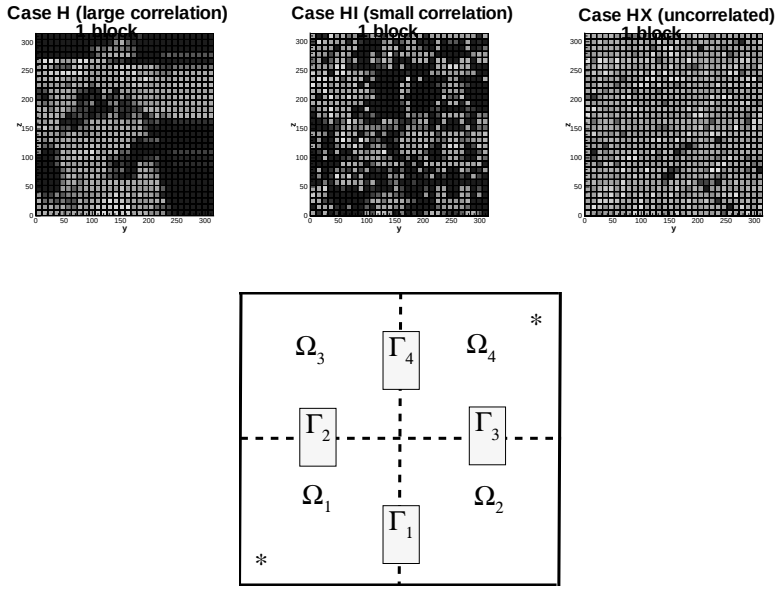


FIGURE 6. Design of two-phase flow experiments. Permeability fields for cases H, HI, HX and the domain decomposition.

The mortar multiblock algorithm is executed on each of the domains as follows. First, we obtain the one-block solution ($n = 1$, no mortar interface) which we use as the reference case. Next, Ω is decomposed into $n = 4$ subdomains with parts of the interfaces Γ numbered as $1, 2, 3, 4$, see Figure 6. We are interested in solving the problem both with various number n_m of mortar degress of freedom assigned per interface, and also with varying degree of tolerance ε in (78). In fact, it turns out that these two directions are related, see Section 7.3.

7.2. Saturations. First, we discuss the pointwise results. Consider the contours of water saturation S_w at the end of simulation shown in Figure 7. These can be considered representative for the problem as they are strongly coupled to the pressure solution and in turn to the *weakly continuous* velocity solution. In a way, saturation profiles give a sense of the "worst" scenario, especially that for variables with convective character such as saturations, the visualization artifacts discussed in Remark 5.1 are worse than those for steady-state variables.

For each case, we now compare the $n = 1$ solution to the $n = 4$ with varying n_m. We see that for case H the results for the coarse mortar grid are much worse than those for the finer grids. On the other hand, for cases HI, HX the difference between results for varying n_m is much "smoother". This can be clearly explained with the following reasoning. Cases with large correlation length such as H (or the case reported in Section 5.1) give rise to large-scale variations of the flow, and in order to be handled correctly in domain decomposition setting, the mortar grid should be aligned with the material properties or should be chosen fine enough *locally* near where these variations occur. See also related results in [45]. On the other hand, cases with smaller correlation length give rise to more "small-scale" variations of the solution which are not harmed as much by averaging introduced by the mortars.

7.3. Well rates. As mentioned before, the well rates reported from reservoir simulations are important for several reasons. First, they capture in some sense the average in space and transient behavior of the field without having to post-process and maintain large volumes of pointwise data. Second, they are practically the only quantities that can be realistically measured in a real field. Last, it is the well rates which determine the economic value of a field and its operating conditions and therefore are most important from the applications point of view.

In the model (76), the well injection and production (mass) rates appear as the source/sink terms f_α, respectively. For incompressible flow, at any time the net amount of volumetric rates must be zero, and well rates are not independent from one another. Taking into account reasonable pressure gradients, the same is approximately true for compressible flow.

One typically considers just a few of the quantities. Here we consider the oil production rate f_o which is most dependent on the global dynamics of pressure, and the water/oil ratio $\frac{f_w}{f_o}$, which clearly shows the water breakthrough time and is most sensitive to the local heterogeneities and to the arrival of saturation fronts. These two quantities for all three choices H, HI, HX of K as well as for various choices of n_m (left) and of ε (right) are shown in Figure 8.

When analyzing these results, we first note the behavior of oil production rate f_o which is large approximately throughout the first half of the simulation and decreases around the breakthrough time. This behavior is typical for fixed BHP-specified wells.

Next, as concerns the sensitivity of the well rates to the mortar discretization parameter n_m, conclusions similar to those arising from analysis of saturations in Section 7.2 can be drawn. That is, the case H appears to be non-smoothly sensitive to n_m whereas cases HI, HX remain less sensitive. It is interesting but not very surprising that essentially the same can be said concerning sensitivity to the interface tolerance criteria in (78) as shown by the sequence of graphs in Figure 8, right.

In summary, in case H it appears important how "closely" the flux continuity is approximated by weak-continuity. On the other hand, this is less significant in cases HI, HX. Note that these sensitivities were determined in experiments where n_m and ε were chosen uniformly on all interfaces $\Gamma_1, \dots \Gamma_4$.

In the next section, guided by the error indicator $\eta_{\Gamma,*}$, we show how the quality of the solution can be improved by refining the mortar grid locally, that is, on a selected Γ_k.

7.4. Error indicator for multi-phase flow. Let us focus now on the case H as it is the one that can benefit most. Consider $n_m = 1$, uniformly on all interfaces and consider the norm of the jump in the oil flux $\boldsymbol{B}_{h,o}$ computed locally per each interface $\Gamma_k, k = 1, \dots 4$, at each time step, plotted in Figure 9, left. Note that the same simplifications as those mentioned in Remark 5.2 apply.

From the plot we see that the largest values of that functional arise on interfaces Γ_1, Γ_4. This is related to the fact that the local variation of K along (and across) these two interfaces is more prominent than at Γ_2, Γ_3; see Figure 6.

The plot suggests therefore the need to adapt n_m on these two interfaces. After that is done, the simulation is redone and new error indicators plotted (in the same Figure) appear all approximately the same, that is, having the same norm.

To see that this refinement procedure was successful, we show the well rates recomputed for the adapted case and plotted along with all the other uniform mortar grid results in Figure 9, right.

Note that the norm of $\boldsymbol{B}_{h,w}$ is not shown; however, it has similar overall behavior as $\boldsymbol{B}_{h,o}$. The delicate differences between the behavior of $\boldsymbol{B}_{h,w}$ and $\boldsymbol{B}_{h,o}$ can be exploited in a fully adaptive procedure discussed next.

7.5. Summary of adaptivity of mortars for multi-phase flow. The following mortar adaptive procedure arises as a logical generalization of the experiments shown above. First, given a certain geological model, one should choose a practical subdomain decomposition that is consistent with the major geological features of the reservoir. In particular, the subdomain interfaces Γ, whenever possible, should fall along any major contrasts of rock types. Second, given practical computational constraints including the size of memory and estimating realistically the time of computations, one should choose subdomain grids $\mathcal{T}_h$ which may be non-matching across Γ. Third, one should choose the coarsest mortar grid and most liberal tolerance that seem practical in computations and compute a corresponding transient solution. Then, guided by possible imbalance of mortar error indicators as shown in Section 7.4, one should increase n_m locally per interface Γ_{ij}. As a last step one could decide on the fine details of the mortar grid per interface Γ_{ij} as guided by the pointwise values of $\eta_{\Gamma,*}$ as shown in the Section 5.1.

Obviously, n_m and convergence criteria could also vary in time which complements the implementation of optimal solvers and preconditioners; this has been succesfully attempted in large reservoir simulation projects referenced above albeit in a static setting - a dynamic adaptive approach will be likely more efficient.

We close by remarking that some preliminary results not included here suggest that the second and third steps above could be done using a grid coarser than the target grid $\mathcal{T}_h$ where some form of upscaling from geological model may be necessary. Finally, a fully adaptive procedure including both subdomain and mortar grids as well as model and time-stepping can be performed but this is outside our scope.

Acknowledgments

All of the computational results and a significant part of the theoretical results presented in this paper were obtained while the author was at The University of Texas at Austin as a member of Center for Subsurface Modeling directed by Prof. Mary F. Wheeler whose support is gratefully acknowledged.

In addition, as mentioned in Section 5, we used the multiblock implementation within the IPARS framework. The author would like to thank all the colleagues who contributed to the IPARS implementation and in particular, John A. Wheeler, Mary F. Wheeler, Qin Lu, Manish Parashar, Yuri Vassilevski, and Ivan Yotov.

References

[1] In Z. Chen and R. Ewing, editors, *Fluid flow and transport in porous media: mathematical and numerical treatment (South Hadley, MA, 2001)*, volume 295 of *Contemp. Math.* Amer. Math. Soc., Providence, RI, 2002.

[2] Robert A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[3] Mark Ainsworth and Donald W. Kelly. A posteriori error estimators and adaptivity for finite element approximation of the non-homogeneous Dirichlet problem. *Adv. Comput. Math.*, 15(1-4):3–23 (2002), 2001. A posteriori error estimation and adaptive computational methods.

[4] Mark Ainsworth and J. Tinsley Oden. A posteriori error estimation in finite element analysis. *Comput. Methods Appl. Mech. Engrg.*, 142(1-2):1–88, 1997.

[5] H. W. Alt and E. di Benedetto. Nonsteady flow of water and oil through inhomogeneous porous media. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 12(3):335–392, 1985.

[6] T. Arbogast. The existence of weak solutions to single porosity and simple dual-porosity models of two-phase incompressible flow. *Nonlinear Analysis, Theory, Methods and Applications*, 19:1009–1031, 1992.

[7] T. Arbogast, L. C. Cowsar, M. F. Wheeler, and I. Yotov. Mixed finite element methods on non-matching multiblock grids. *SIAM J. Numer. Anal.*, 37:1295–1315, 2000.

[8] T. Arbogast, M. F. Wheeler, and I. Yotov. Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite differences. *SIAM J. Numer. Anal.*, 34(2):828–852, 1997.

[9] Todd Arbogast, Mary F. Wheeler, and Nai-Ying Zhang. A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous media. *SIAM J. Numer. Anal.*, 33(4):1669–1687, 1996.

[10] D. N. Arnold and F. Brezzi. Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. *RAIRO Modél. Math. Anal. Numér.*, 19(1):7–32, 1985.

[11] Ivo Babuška and Werner C. Rheinboldt. A posteriori error analysis of finite element solutions for one-dimensional problems. *SIAM J. Numer. Anal.*, 18(3):565–589, 1981.

[12] F. Ben Belgacem. The mortar finite element method with Lagrange multiliers. *Numer. Math.*, 84(2):173–197, 1999.

[13] C. Bernardi and V. Girault. A local regularization operator for triangular and quadrilateral finite elements. *SIAM J. Numer. Anal.*, 35(5):1893–1916 (electronic), 1998.

[14] C. Bernardi, Y. Maday, and A. T. Patera. A new nonconforming approach to domain decomposition: the mortar element method. In H. Brezis and J. L. Lions, editors, *Nonlinear partial differential equations and their applications*. Longman Scientific & Technical, UK, 1994.

[15] C. Bernardi and R. Verfürth. Adaptive finite element methods for elliptic equations with non-smooth coefficients. *Numer. Math.*, 85(4):579–608, 2000.

[16] D. Braess and R. Verfürth. A posteriori error estimators for the Raviart-Thomas element. *SIAM J. Numer. Anal.*, 33(6):2431–2444, 1996.

[17] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Springer-Verlag, New York, 1991.

[18] Carsten Carstensen. A posteriori error estimate for the mixed finite element method. *Math. Comp.*, 66(218):465–476, 1997.

[19] Zhangxin Chen. Degenerate two-phase incompressible flow. I. Existence, uniqueness and regularity of a weak solution. *J. Differential Equations*, 171(2):203–232, 2001.

[20] Zhangxin Chen. Degenerate two-phase incompressible flow. II. Regularity, stability and stabilization. *J. Differential Equations*, 186(2):345–376, 2002.

[21] Zhangxin Chen and Natalia L. Khlopina. Degenerate two-phase incompressible flow problems. i. regularization and numerical results. *Commun. Appl. Anal.*, 5(3):319–334, 2001.

[22] Philippe G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4.

[23] Ph. Clément. Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér.*, 9(R-2):77–84, 1975.

[24] Lawrence C. Cowsar, Jan Mandel, and Mary F. Wheeler. Balancing domain decomposition for mixed finite elements. *Math. Comp.*, 64(211):989–1015, 1995.

[25] B. Engelmann, R. H. W. Hoppe, Y. Iliash, Y. A. Kuznetsov, Y. Vassilevski, and B. Wohlmuth. Adaptive finite element methods for domain decomposition on nonmatching grids. In *Parallel solution of partial differential equations (Minneapolis, MN, 1997)*, pages 57–83. Springer, New York, 2000.

[26] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1983.

[27] R. Glowinski and M. F. Wheeler. Domain decomposition and mixed finite element methods for elliptic problems. In R. Glowinski, G. H. Golub, G. A. Meurant, and J. Periaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 144–172. SIAM, Philadelphia, 1988.

[28] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.

[29] R. Helmig. *Multiphase flow and transport processes in the Subsurface*. Springer, 1997.

[30] R. Hoppe, Yu. Iliash, Yu. Kuznetsov, Yu. Vassilevski, and B. Wohlmuth. Analysis and parallel implementation of adaptive mortar element methods. *East-West J. Numer. Math.*, 6(3):223–248, 1998.

[31] J. W. Jennings, Jr., S. C. Ruppel, and W. B. Ward. Geostatistical analysis of permeability data and modeling of fluid-flow effects in carbonate outcrops. *Society of Petroleum Engineers Reservoir Evaluation and Engineering*, 3(4):292–303, 2000.

[32] S. Lacroix, Y. Vassilevski, and M. F. Wheeler. Iterative solvers of the Implicit Parallel Accurate Reservoir Simulator (IPARS). *Numerical Linear Algebra with Applications*, 4:537–549, 2001.

[33] S. Lacroix, Yu. Vassilevski, M. Wheeler, and J. Wheeler. Iterative solution methods for modeling multiphase flow in porous media fully implicitly. *SIAM J. Sci. Comp.*, 25(3):905–926, 2003.

[34] L. W. Lake. *Enhanced oil recovery.* Prentice Hall, 1989.

[35] Q. Lu, M. Parashar, M. Peszyńska, and M. F. Wheeler. Parallel implementation of multiphysics multi-block for multi-phase flow in subsurface. In preparation.

[36] Q. Lu, M. Peszyńska, and M. F. Wheeler. A parallel multi-block black-oil model in multimodel implementation. In *2001 SPE Reservoir Simulation Symposium*, Houston, Texas, 2001. SPE 66359.

[37] Q. Lu, M. Peszyńska, and M. F. Wheeler. A parallel multi-block black-oil model in multimodel implementation. *SPE Journal*, 7(3):278–287, September 2002. SPE 79535.

[38] D. W. Peaceman. *Fundamentals of numerical reservoir simulation.* Elsevier Scientfic Publishing Company, Amsterdam-Oxford-New York, first edition, 1977.

[39] D. W. Peaceman. Interpretation of well-block pressure in numerical reservior simulation with non-square grid blocks and anisotropic permeability. *Tran. AIME*, 275:10–22, 1983.

[40] Gergina Pencheva and Ivan Yotov. Balancing domain decomposition for mortar mixed finite element methods. *Numer. Linear Algebra Appl.*, 10(1-2):159–180, 2003. Dedicated to the 60th birthday of Raytcho Lazarov.

[41] M. Peszyńska. Advanced techniques and algorithms for reservoir simulation III. Multiphysics coupling for two phase flow in degenerate conditions. In J. Chadam, A. Cunningham, R. E. Ewing, P. Ortoleva, and M. F. Wheeler, editors, *IMA Volume 131: Resource Recovery, Confinement, and Remediation of Environmental Hazards*, pages 21–40. Springer, 2002.

[42] M. Peszyńska, E. Jenkins, and M. F. Wheeler. Boundary conditions for fully implicit two-phase flow model. In Xiaobing Feng and Tim P. Schulze, editors, *Recent Advances in Numerical Methods for Partial Differential Equations and Applications*, volume 306 of *Contemporary Mathematics Series*, pages 85–106. American Mathematical Society, 2002.

[43] M. Peszyńska, Q. Lu, and M. F. Wheeler. Coupling different numerical algorithms for two phase fluid flow. In J. R. Whiteman, editor, *MAFELAP Proceedings of Mathematics of Finite Elements and Applications*, pages 205–214, Uxbridge, U.K., 1999. Brunel University.

[44] M. Peszyńska, Q. Lu, and M. F. Wheeler. Multiphysics coupling of codes. In L. R. Bentley, J. F. Sykes, C. A. Brebbia, W. G. Gray, and G. F. Pinder, editors, *Computational Methods in Water Resources*, pages 175–182. A. A. Balkema, 2000.

[45] M. Peszyńska, M. F. Wheeler, and I. Yotov. Mortar upscaling for multiphase flow in porous media. *Computational Geosciences*, 6:73–100, 2002.

[46] R. A. Raviart and J. M. Thomas. A mixed finite element method for 2nd order elliptic problems. In *Mathematical Aspects of the Finite Element Method, Lecture Notes in Mathematics*, volume 606, pages 292–315. Springer-Verlag, New York, 1977.

[47] T. F. Russell and M. F. Wheeler. Finite element and finite difference methods for continuous flows in porous media. In R. E. Ewing, editor, *The Mathematics of Reservoir Simulation*, pages 35–106. SIAM, Philadelphia, 1983.

[48] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques.* Wiley-Teubner, Chichester, 1996.

[49] M. F. Wheeler. Advanced techniques and algorithms for reservoir simulation, II: The multiblock approach in the integrated parallel accurate reservoir simulator (IPARS). In John Chadam, Al Cunningham, Richard E. Ewing, Peter Ortoleva, , and Mary Fanett Wheeler, editors, *IMA Volumes in Mathematics and its Applications, Volume 131: Resource Recovery, Confinement, and Remediation of Environmental Hazards.* Springer, 2002.

[50] M. F. Wheeler and M. Peszyńska. Computational engineering and science methodologies for modeling and simulation of subsurface applications. *Advances in Water Resources*, 25(8-12):1147–1173, 2002.

[51] M. F. Wheeler, J. A. Wheeler, and M. Peszyńska. A distributed computing portal for coupling multi-physics and multiple domains in porous media. In L. R. Bentley, J. F. Sykes, C. A. Brebbia, W. G. Gray, and G. F. Pinder, editors, *Computational Methods in Water Resources*, pages 167–174. A. A. Balkema, 2000.

[52] Mary F. Wheeler and Ivan Yotov. Multigrid on the interface for mortar mixed finite element methods for elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 184(2-4):287–302, 2000. Vistas in domain decomposition and parallel processing in computational mechanics.

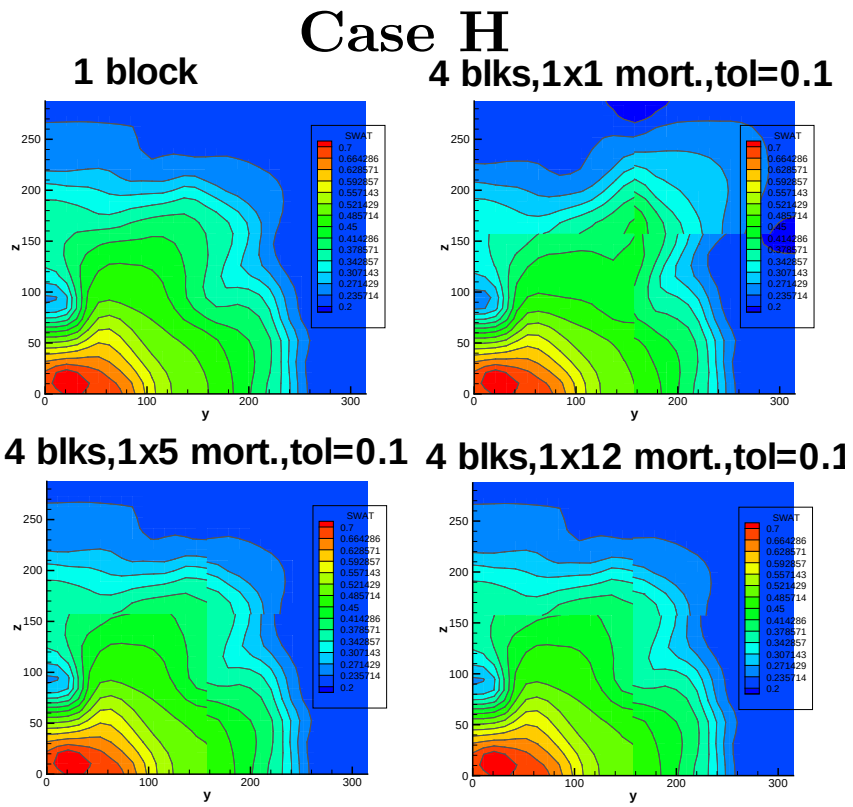
[53] Barbara I. Wohlmuth. Hierarchical a posteriori error estimators for mortar finite element methods with Lagrange multipliers. *SIAM J. Numer. Anal.*, 36(5):1636–1658 (electronic), 1999.

[54] Barbara I. Wohlmuth. A residual based error estimator for mortar finite element discretizations. *Numer. Math.*, 84(1):143–171, 1999.

[55] Barbara I. Wohlmuth. A mortar finite element method using dual spaces for the Lagrange multiplier. *SIAM J. Numer. Anal.*, 38(3):989–1012 (electronic), 2000.

[56] Barbara I. Wohlmuth and Ronald H. W. Hoppe. A comparison of a posteriori error estimators for mixed finite element discretizations by Raviart-Thomas elements. *Math. Comp.*, 68(228):1347–1378, 1999.

[57] Ivan Yotov. Advanced techniques and algorithms for reservoir simulation. IV. Multiblock solvers and preconditioners. In *Resource recovery, confinement, and remediation of environmental hazards (Minneapolis, MN, 2000)*, volume 131 of *IMA Vol. Math. Appl.*, pages 41–55. Springer, New York, 2002.



Department of Mathematics, Oregon State University Corvallis, OR 97330
E-mail: mpesz@math.oregonstate.edu
URL: http://www.math.oregonstate.edu/∼mpesz/

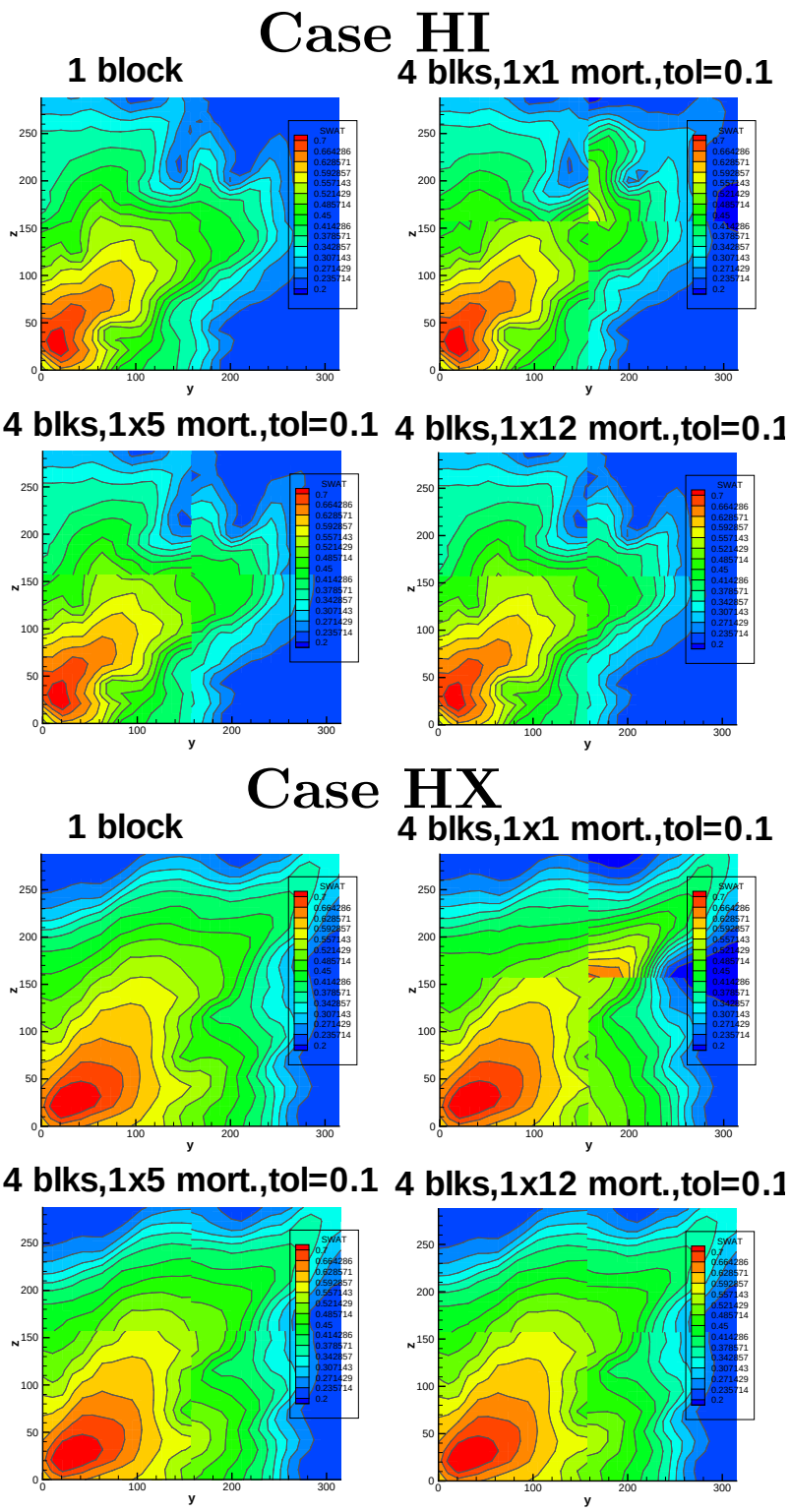


FIGURE 7. Two-phase flow. Saturations after 1000 days, from top to bottom: cases H, HI, HX

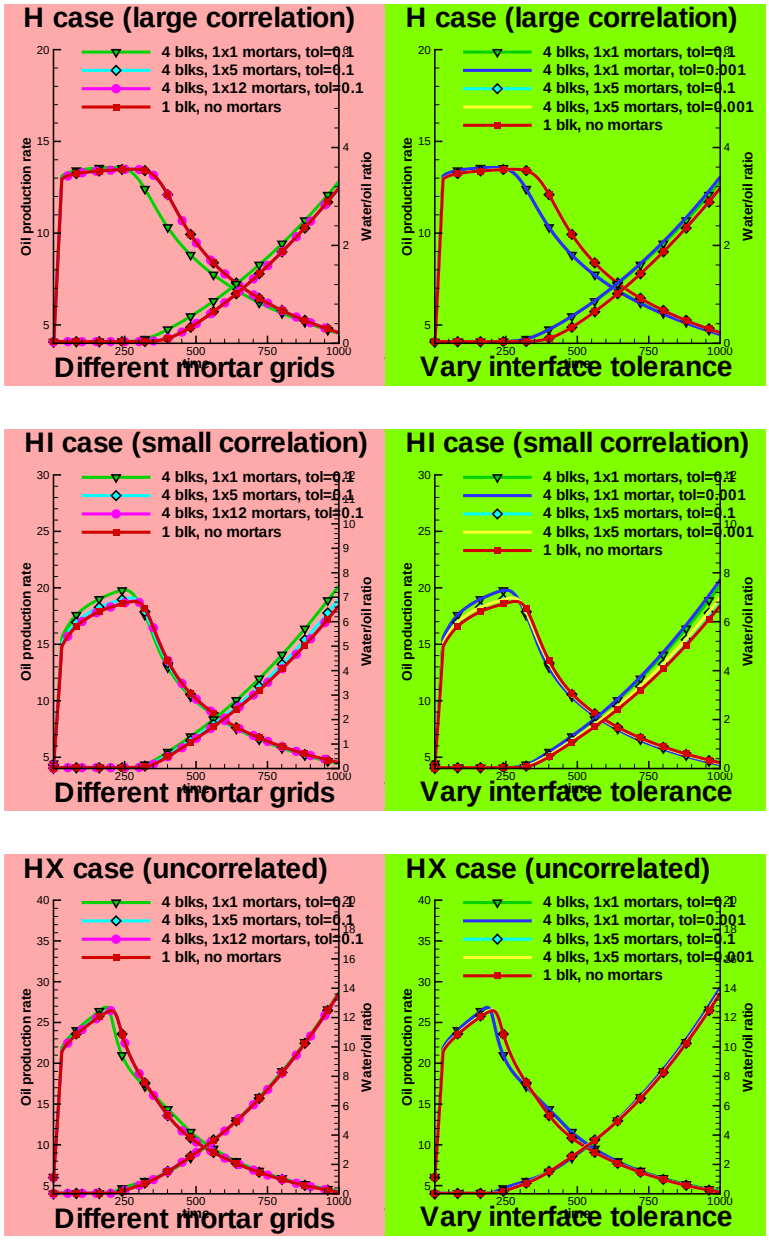


FIGURE 8. Well rates for the two-phase flow example. Sensitivity to mortar grid (left column) and to convergence tolerance ε (right column). Top to bottom: cases H, HI, HX

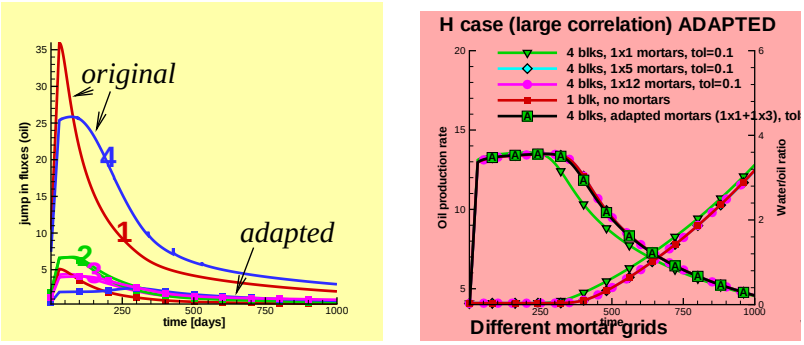


FIGURE 9. Application of the error indicator $\eta_{\Gamma,*}$ for two phase flow. Left: error indicator. Right: well rates. Results are shown for the case before and after the mortar grid was adapted